

ISBN 978-623-6081-03-7 (PDF)



**UMSIDA Press**  
Universitas Muhammadiyah Sidoarjo  
Jl. Mojopahid No. 666B  
Sidoarjo, Jawa Timur



Miftahul Mushlih, M.Sc  
Rafhani Rosyidah., S.Keb., Bd., M.Sc.

# STATISTIKA

## "APLIKASI DI DUNIA KESEHATAN"



**BUKU AJAR MATA KULIAH  
STATISTIKA  
“Aplikasi di Dunia Kesehatan”**

Oleh  
Miftahul Mushlih  
Rafhani Rosyidah



Diterbitkan oleh  
**UMSIDA PRESS**

**UNIVERSITAS MUHAMMADIYAH SIDOARJO**

**2020**

**BUKU AJAR**  
**STATISTIKA “APLIKASI DI DUNIA KESEHATAN”**

**Penulis:**

Miftahul Muslih, M.Sc.

Rafhani Rosyidah, M.Keb.

**ISBN :**

978-623-6081-03-7

**Editor:**

Galuh Ratmana Hanum, M.Si.

**Design Sampul dan Tata Letak:**

Mochammad Nashrullah, S.Pd.

Amy Yoga Prajati, S.Kom.

**Penerbit:**

UMSIDA Press

Anggota IKAPI No. 218/Anggota Luar Biasa/JTI/2019

Anggota APPTI No. 002 018 1 09 2017

**Redaksi**

Universitas Muhammadiyah Sidoarjo

Jl. Mojopahit No 666B

Sidoarjo, Jawa Timur

Cetakan Pertama, September 2020

©Hak Cipta dilindungi undang undang

Dilarang memperbanyak karya tulis ini dengan sengaja, tanpa ijin tertulis dari penerbit.

# Kata Pengantar

Segala puji bagi Allah yang telah memberikan kemudahan kemudahan kepada kami sehingga dapat menyelesaikan buku “Statistika: aplikasi di dunia kesehatan” ini. Tanpa pertolongan dan arahannya-Nya penyusun tidak akan sanggup menyelesaikannya dengan baik.

“Statistika: aplikasi di dunia kesehatan”, Merupakan Buku yang disusun berdasarkan pengalaman dari rujukan maupun penulis sendiri terkait dengan kendala kendala yang ditemui didalam pembelajaran statistika. Semoga Buku ini dapat memberikan pencerahan yang dapat mempermudah dalam proses pembelajaran khususnya bagi yang berkecimpung di dunia kesehatan. Penyusun juga sadar bahwa buku ini masih jauh dari sempurna. kritik dan saran dari pembaca yang membangun sungguh sangat diharapkan.

Pada akhir kata, penulis ingin menyampaikan banyak terimakasih atas terselesaikannya penulisan ini kepada pihak pihak terkait, khususnya kolega Universitas Muhammadiyah Sidoarjo dan teman teman yang memberi saran guna menyempurnakan buku ini.

Tim penulis

# Capaian Pembelajaran

materi	Yang harus dipahami oleh mahasiswa
<b>Konsep Dasar Statistika</b>	<ul style="list-style-type: none"><li>a. Statistika dan Statistik</li><li>b. Ruang Lingkup Statistika</li><li>c. Pengertian data</li><li>d. Sumber data</li><li>e. Jenis data berdasarkan skala pengukuran</li></ul>
<b>Penyajian Data</b>	<ul style="list-style-type: none"><li>a. Jenis-jenis penyajian data<ul style="list-style-type: none"><li>i. Tulisan</li><li>ii. Diagram Batan</li><li>iii. Pictogram Atau Diagram Lambang</li><li>iv. Diagram garis</li><li>v. Diagram Lingkaran</li><li>vi. Tabel Distribusi Frekwensi</li><li>vii. Tabel distribusi frekwensi dengan kelas terbuka</li><li>viii. Histogram Frekwensi</li></ul></li></ul>
<b>Ukuran Pemusatan Data</b>	<ul style="list-style-type: none"><li>a) Interpretasi Pemusatan data<ul style="list-style-type: none"><li>1. Rata-rata atau mean</li><li>2. Modus</li><li>3. Median</li><li>4. Hubungan antara nilai rata-rata, median dan modus</li></ul></li></ul>
<b>Ukuran penyebaran data dan letak ukuran penyebaran data</b>	<ul style="list-style-type: none"><li>a. Range Atau jangkauan<ul style="list-style-type: none"><li>I. Range Atau jangkauan data tunggal</li><li>II. Range Atau jangkauan data distribusi frekwensi kelompok</li></ul></li></ul>

	<ul style="list-style-type: none"> <li>b. Standart Deviasi</li> <li>c. Ragam</li> <li>d. Kuartil, Desil, dan presentil</li> </ul>
<b>Populasi, sampel, sampling, besar sampel</b>	<ul style="list-style-type: none"> <li>a. Populasi dan Sampel</li> <li>b. Besar sampel <ul style="list-style-type: none"> <li>I. Penelitian Laboratorik</li> <li>II. Penelitian Laboratorium secara umum</li> <li>III. Penelitian Survei</li> <li>IV. Untuk populasi besar</li> </ul> </li> <li>c. Metode sampling <ul style="list-style-type: none"> <li>a. Probability Sampling</li> <li>b. Non Probability Sampling</li> </ul> </li> </ul>
<b>Hipotesis &amp; Statistika Inferensial</b>	<ul style="list-style-type: none"> <li>a. Tingkat Kepercayaan</li> <li>b. Hipotesis</li> <li>c. Derajat kebebasan / df (degree of freedom)</li> <li>d. Langkah-langkah uji hipotesis</li> <li>e. Rumusan dan arah hipotesis.</li> <li>f. Statistika Inferensial</li> <li>g. Tahapan Uji Statistika</li> <li>h. Pemilihan uji statistik univariat atau bivariat</li> </ul>
<b>Uji perbedaan mean 2 sampel dengan statistik parametrik (Uji T)</b>	<ul style="list-style-type: none"> <li>a. Uji T Satu Sample (One/ two tail)</li> <li>b. Uji T Dua Sample</li> </ul>
<b>Uji perbedaan mean 2 sampel dengan statistic non parametrik</b>	<ul style="list-style-type: none"> <li>a. Uji wilcoxon</li> <li>b. uji mann-withney</li> </ul>

<b>Uji perbedaan mean lebih dari 2 sampel dengan statistic parametrik</b>	<ul style="list-style-type: none"> <li>a. Anova satu arah dan dua arah</li> <li>b. Penelitian RAK dan RAL</li> </ul>
<b>Uji perbedaan mean lebih dari 2 sampel dengan statistic non parametrik</b>	<ul style="list-style-type: none"> <li>a. Clustal Wallis</li> </ul>
<b>Uji Korelasi dan Regresi</b>	<ul style="list-style-type: none"> <li>a. Koefisien Korelasi</li> <li>b. Pengetesan Signifikansi</li> <li>c. Syarat-Syarat untuk Pengetesan Nilai r</li> <li>d. Apa Gunanya r ?</li> <li>e. Pembagian uji Korelasi</li> </ul>
<b>Analisis regresi</b>	Analisis regresi
<b>Uji Chi-Square</b>	<ul style="list-style-type: none"> <li>a. Uji Homogeneity (Homogenitas) chi square test</li> <li>b. Uji Goodness of fit (Kesesuaian) chi square test</li> <li>c. Independence chi square test</li> </ul>

# Daftar isi

<b>BAB I. Konsep Dasar Statistika .....</b>	<b>1</b>
a. Statistika dan Statistik .....	1
b. Ruang Lingkup Statistika .....	3
c. Pengertian data.....	3
d. Sumber data .....	4
e. Jenis data berdasarkan skala pengukuran .....	4
f. Ringkasan materi .....	5
g. Latihan soal .....	6
h. Daftar Pustaka .....	6
<b>BAB II. Penyajian Data .....</b>	<b>8</b>
a. Jenis jenis penyajian data .....	9
i. Tulisan.....	9
ii. Diagram Batang .....	9
iii. Pictogram Atau Diagram Lambang.....	10
iv. Diagram garis.....	10
v. Diagram Lingkaran.....	11
vi. Tabel Distribusi Frekwensi.....	12
vii. Tabel distribusi frekwensi dengan kelas terbuka.....	17
viii. Histogram Frekwensi .....	19
b) Ringkasan materi .....	21
c) Latihan Soal .....	21
d) Daftar Pustaka .....	22

<b>BAB III. Ukuran Pemusatan Data .....</b>	<b>23</b>
a) Interpretasi Pemusatan data .....	24
1. Rata-rata atau mean .....	25
2. Modus .....	26
3. Median .....	28
4. Hubungan antara nilai rata-rata, median dan modus .....	32
b) Ringkasan Materi .....	33
c) Latihan soal .....	34
d) Daftar Pustaka .....	34

<b>Bab IV. Ukuran penyebaran data dan letak ukuran penyebaran data .....</b>	<b>35</b>
a. Range Atau jangkauan.....	35
I. Range Atau jangkauan data tunggal.....	36
II. Range Atau jangkauan data distribusi frekwensi kelompok 36	
b. Standart Deviasi.....	37
c. Ragam .....	39
d. Kuartil, Desil, dan presentil .....	39
e. Ringkasan Materi.....	42
f. Soal Latihan .....	43
e. Daftar Pustaka .....	44

<b>Bab V. Populasi, sampel, sampling, besar sampel.....</b>	<b>45</b>
a. Populasi dan Sampel .....	47
b. Besar sampel .....	47
I. Penelitian Laboratorik .....	47
II. Penelitian Laboratorium secara umum .....	48
III. Penelitian Survei.....	49
IV. Untuk populasi besar.....	51
c. Metode sampling.....	52
a. Probability Sampling.....	54
b. Non Probability Sampling .....	58

d.	Ringkasan Materi.....	60
e.	Latihan soal.....	61
f.	Daftar Pustaka .....	61
<b>Bab V. Hipotesis &amp; Statistika Inferensial.....</b>		<b>63</b>
a.	Tingkat Kepercayaan .....	64
b.	Hipotesis.....	64
c.	Derajad kebebasan / df (degree of freedom).....	65
d.	Langkah-langkah uji hipotesis .....	65
e.	Rumusan dan arah hipotesis. ....	66
f.	Statistika Inferensial .....	69
g.	Tahapan Uji Statistika .....	69
h.	Pemilihan uji statistik univariat atau bivariat.....	71
i.	Ringkasan Materi.....	73
j.	Latihan soal.....	74
k.	Daftar Pustaka .....	74
<b>BAB VI. Uji perbedaan mean 2 sampel dengan statistik parametrik (Uji T) .....</b>		<b>76</b>
a.	Uji T Satu Sample (One/ two tail) .....	78
b.	Uji T Dua Sample.....	85
c.	Ringakasan materi .....	101
d.	Latihan soal.....	102
e.	Daftar Pustaka .....	103
<b>BAB VII. Uji perbedaan mean 2 sampel dengan statistic non parametrik .....</b>		<b>105</b>
a.	Uji wilcoxon .....	106
b.	uji mann-withney.....	113
c.	Ringakasan materi .....	119
d.	Latihan soal.....	119
f.	Daftar Pustaka .....	120

<b>Bab VIII. Uji perbedaan mean lebih dari 2 sampel dengan statistic parametrik .....</b>	<b>121</b>
a. Anova satu arah dan dua arah.....	124
b. Penelitian RAK dan RAL .....	140
c. Ringkasan materi .....	141
d. Latihan soal.....	141
e. Daftar Pustaka .....	142
<b>Bab IX. Uji perbedaan mean lebih dari 2 sampel dengan statistic non parametrik.....</b>	<b>144</b>
a. <b>Clustal Wallis</b> .....	144
b. Ringkasan materi .....	151
c. Latihan Soal .....	152
d. Daftar pustaka .....	152
<b>Bab X. Uji Korelasi dan Regresi.....</b>	<b>154</b>
a. Koefisien Korelasi .....	156
b. Pengetesan Signifikansi .....	158
c. Syarat-Syarat untuk Pengetesan Nilai $r$ .....	160
d. Apa Gunanya $r$ ? .....	161
e. Pembagian uji Korelasi .....	162
f. Korelasi Pearson .....	163
g. Korelasi Spearman.....	168
h. Korelasi Kendal tau b.....	173
i. Ringkasan materi .....	177
j. Latihan Soal .....	177
e. Daftar pustaka .....	179
<b>Bab XII. Uji Regresi .....</b>	<b>181</b>
a. Analisis Regresi .....	181
b. Ringkasan materi .....	188
c. Latihan soal.....	188

d.	Daftar pustaka .....	190
<b>Bab XIII. Uji Chi-Square .....</b>		<b>191</b>
a.	Uji Homogeneity (Homogenitas) chi square test .....	193
b.	Uji Goodness of fit (Kesesuaian) chi square test .....	196
c.	Independence chi square test .....	199
d.	Ringkasan Materi.....	204
e.	Latihan soal.....	205
a.	Daftar pustaka .....	205



# BAB I. Konsep Dasar Statistika

---

## Tujuan

Mahasiswa memahami konsep dasar, ruang lingkup statistika dan dapat menjelaskan jenis-jenis data

---

## Materi

1. Perbedaan statistika dan statistik
2. Ruang lingkup statistika
3. Pengertian data
4. Jenis data

---

### a. Statistika dan Statistik

Seringkali kita mendengar istilah statistik dan statistika. Lalu Apakah ada perbedaan diantara keduanya?. Perbedaan mendasar di antara 2 kata tersebut di dalam bahasa Inggris statistik selalu diterjemahkan menjadi *statistic* tanpa disertai dengan huruf s. Statistika dalam bahasa Inggris diterjemahkan menjadi *statistics* yang disertai dengan huruf s. Statistik secara sederhana dapat diartikan sebagai sekumpulan angka-angka yang menggambarkan suatu kondisi tertentu sedangkan statistika merupakan ilmu yang mempelajari angka-angka tersebut. Apabila diterjemahkan lebih lanjut maka bio sendiri makhluk hidup. Secara harfiah maka biostatistika dapat didefinisikan sebagai ilmu yang mempelajari tentang angka-angka hasilkan oleh apabila kita lihat kesehatan makan manusia.

Hampir semua profesi membutuhkan statistik. Termasuk diantaranya presiden, pimpinan perusahaan, peneliti, atau bahkan konsumen online. Angka-angka untuk memprediksi suatu kejadian yang akan datang maupun sebagai koreksi suatu kegagalan. Di dunia penelitian dan pendidikan statistik digunakan untuk

menggambarkan dan membuktikan suatu perjanjian dari hasil keadaan lapangan.

Kata-kata statistika berasal dari bahasa latin status yang artinya sama dengan bahasa Italia yaitu *region de stato*. Istilah ini mulai digunakan sekitar tahun 1250 sebelum masehi pada masa pemerintahan Raja Ramses II. Statistik digunakan untuk mencatat kelahiran dan kematian secara sederhana pada suatu negara. Statistika mulai berkembang pada pertengahan abad ke-15 Eropa sekitar tahun 1503 pada zaman Raja Henry ke-8 yang digunakan untuk mencatat peristiwa wa ke sakitan maupun kematian secara lebih teratur. Pada tahun 1603 ground dan Patty Jurnal secara berkala mengenai laporan penduduk yang mengalami sakit maupun kematian. Menganalisis mengenai peristiwa prevalensi, epidemiologi penyakit yang bersifat kronis. Pada tahun 1662 John Graund juga merasakan suatu laporan di jurnal *natural and political observations Medan following index and match you open the bill of mortality*. Laporan tersebut tidak ada hubungannya sama sekali dengan laporan sebelumnya. Peristiwa penerbitan jurnal dengan menggunakan metode statistik ini membuat mendapat julukan sebagai Bapak statistika Kesehatan. Diantara isi dari jurnal jurnal yang diterbitkan oleh ground adalah regulasi fenomena vital pada manusia meliputi kematian kelahiran dan kehidupan, laki-laki selain itu kematian rata-rata yang terjadi pada usia dini lebih tinggi dibandingkan kematian pada masa usia dan yang terakhir adalah kematian di perkotaan jauh lebih tinggi dibandingkan kematian di pedesaan. Semenjak penemuan tersebut muncullah beberapa nama lain seperti Quetelet (1796-1874) yang menyajikan statistika secara lebih modern. Teori probabilitas dikembangkan oleh Laplace Pada tahun 1749. Teori dikembangkan pada saat maraknya cuci yang berada di Inggris sekitar abad ke-18. Dasar ilmu korelasi dan regresi dikembangkan oleh Galton pada tahun 1822-1911. Sedangkan dasar chi Square dipopulerkan oleh Pearson (1857-1936). Fisher (1890-1962) menemukan teori uji hujan atau yang lebih dikenal dengan uji t.

## b. Ruang Lingkup Statistika

Ruang lingkup statistika meliputi statistik deskriptif (statistik deduktif), dan statistika analisis atau inferensial (statistik induktif). Adalah statistik yang meliputi pengumpulan data, pengolahan data, penyajian data dan analisis. Data secara sederhana tidak diikuti dengan uji-uji yang rumit dan membutuhkan taraf kepercayaan ataupun hipotesis. Statistik deskriptif bisa dalam bentuk tabel, grafik, nilai rata-rata, standar deviasi (SD), median, nilai maksimal, nilai minimal atau variasi.

Statistika inferensial merupakan kajian statistik yang mencakup semua secara utuh try out. Mulai dari pengumpulan data, pengolahan data, penyajian data, analisis data dan penarikan kesimpulan berdasarkan hipotesis yang telah disusun. Tujuan dari statistika inferensial yaitu membuat generalisasi dari sekumpulan angka-angka yang yang rumit dan karakter yang spesifik. Secara sederhana yang membedakan antara statistika deskriptif dan statistika inferensial adalah pada statistika deskriptif tidak membutuhkan analisa statistik yang mendalam sedangkan statistika inferensial membutuhkan uji uji statistik untuk menarik kesimpulan dari hipotesis.

## c. Pengertian data

Seringkali kita mendengar sebuah istilah data. Lalu apa yang dimaksud sebagai sebuah data? Data istilah yang berasal dari bahasa latin yaitu *datum*. Istilah baik jamak maupun tunggal. Di dalam pengungkapan sebuah kalimat tidak tepat kita menggunakan data data-data melainkan cukup dengan data saja. Yang perlu digarisbawahi adalah dalam statistik data selalu berupa angka. kurang tepat ketika kita menggambarkan sebuah data tidak dengan angka seperti tinggi, rendah, baik, buruk, sering, jarang atau sedang. Semua harus diganti dengan angka-angka seperti tinggi ditunjukkan dengan angka 1 sedangkan benda ditunjukkan dengan angka 0.

Contoh lain apabila kita menggunakan sering, sedang, dan jarang kita dapat mengganti kategori tersebut skor. Seperti sering diberi skor dua, sedang diberi skor satu, dan jarang diberi skor nol.

#### **d. Sumber data**

Menurut sumbernya data dapat dikelompokkan menjadi data primer, data sekunder, data tersier, dan data kuarter. Data primer merupakan data yang diperoleh secara langsung dengan melakukan pengumpulan secara mandiri si dengan cara mendaftar ataupun interview terhadap seseorang. Data primer seringkali bersifat kolektif yang belum dilakukan analisis. Data sekunder merupakan data hasil olahan dari data primer data sekunder seringkali didapatkan dari instansi pengolah data seperti pusat statistik atau instansi lainnya. Sebagai contoh daftar sekunder adalah banyaknya korban kecelakaan lalu lintas yang ada di satu daerah atau frekuensi penderita penyakit tertentu pada suatu kota atau kabupaten. Data tersier merupakan data yang diperoleh dari olahan data sekunder. Data tersier biasanya didapatkan dengan cara pengertian pada suatu media massa atau jurnal ilmiah. Sedangkan data kuarter merupakan data yang diperoleh dari olahan data tersier. Contoh dari data kuarter adalah jurnal review dari beberapa penelitian. Hal tersebut diambil kan dari beberapa peta khususnya cari jurnal kemudian dijadikan sebuah eh besar sehingga mendapatkan kesimpulan yang lebih akurat.

#### **e. Jenis data berdasarkan skala pengukuran**

Secara sederhana jenis data dapat dikelompokkan menjadi dua yaitu data para matrix dan data non para matrix. Data non para matrix dapat digolongkan menjadi nominal dan ordinal sedangkan data para matrix dapat digolongkan menjadi interval dan rasio.

Perbedaan mendasar antara data nominal dan ordinal adalah data nominal tidak menggambarkan suatu tingkatan tertentu. Tak juga tidak memiliki nilai, tidak memiliki satuan ukuran, titik nol, atau nol mutlak, yang kan pada data yang bersifat koordinat meskipun

perbedaan antar peringkat bersifat relatif namun dari segi sekolah data ordinal dapat di peringkat kan. Contoh dari data nominal adalah nama, warna, tempat lahir, nomor ktp, bahkan nomor hp. Pada intinya meskipun contoh-contoh tersebut memiliki sebuah angka namun demikian angka tersebut tidak memiliki peringkat atau tidak bermakna. Perlu diingat ketika kita berbicara pernah atau nama maka di dalam statistik harus dirubah menjadi angka atau lebih mudah disebut sebagai id. Contoh dari data yang bersifat nominal adalah frekuensi seseorang untuk mencuci tangan sebelum makan data bisa dikategorikan menjadi sering, sedang, atau tidak pernah. Kata sore katagori tersebut meskipun sama-sama harus dirubah menjadi angka sebagai contoh sering diwakili dengan angka 2, sedang diwakili dengan angka satu tidak pernah diwakili dengan angka nol. Namun demikian angka-angka tersebut tidak boleh digambarkan secara acak karena adanya pemberangkatan yang tidak boleh dirubah menjadi 0 atau sering tidak boleh dirubah menjadi angka dua.

**Tabel Karakteristik Skala Pengukuran**

No	Karakteristik	Skala Pengukuran			
		Nominal	Ordinal	Interval	Rasio
1	Sifat dasar skala	Nama/predikat	Peringkat	Peringkat urutan	Peringkat urutan
2	Perbedaan antar peringkat	Tidak bermakna	Relativitas	Bermakna	Bermakna
3	Satuan ukuran	Tidak ada	Tidak ada	Ada	Ada
4	Titik nol	Tidak ada	Tidak ada	Ada	Ada
5	Nol mutlak	Tidak ada	Tidak ada	Tidak ada	Ada
6	Rasio dua ukuran	Tidak bisa	Tidak bisa	Tidak bisa	Bisa

## f. Ringkasan materi

- Statistik merupakan kumpulan dari angka-angka yang berasal dari sebuah pengukuran maupun analisa data.
- Statistika merupakan ilmu yang mempelajari angka-angka untuk dijadikan sebuah intisari atau kesimpulan.
- Data selalu berbentuk angka sehingga data-data sifatnya deskriptif harus berubah pula menjadi angka.
- Jenis data dapat dikategorikan menjadi Nominal, Ordinal, Interval dan Rasio

## g. Latihan soal

- a. tentukan data berikut apakah data diskrit atau kontinu !
  1. Jumlah sapi yang di korban di pondok Al-ishlah
  2. Jumlah pengunjung yang datang di Ramayana perharinya
  3. Berat daging kurban yang di bagikan ke penduduk desa sendangagung
  4. Kecepatan sepeda motor yang di kendarai sueb
  5. Jumlah makanan (tusuk) yang di makan Diana di Warung makan Sate Madura
  6. Nilai ulangan matakuliah Statistika
  
- b. Tentukan data berikut apakah dalam Bentuk skala nominal, ordinal, interval atau ratio!
  1. Skor IQ mahasiswa
  2. Jarak perjalanan mahasiswa dari kos ke Kampus 500m
  3. Nilai mahasiswa dalam ujian statistic
  4. Klasifikasi mahasiswa berdasarkan tempat kelahiran
  5. Jumlah jam belajar mahasiswa per minggu
  6. Luas areal kawasan industry
  7. No HP mahasiswa

## h. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Kustituantio, B.& Badrudin, R. 1994, Statistika 1 (Deskriptif), Gunadarma
4. Murti, Bhisma, Penerapan Metode Statistik Non Parametrik Dalam Ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito



## BAB II. PENYAJIAN DATA

---

### Tujuan

**Mahasiswa memahami metode penyajian data dan dapat membuat penyajian data secara tepat sesuai dengan kebutuhan didalam intrepretasi.**

---

### Materi

---

#### 1. Jenis jenis Penyajian data

---

**P**enyajian data diperlukan untuk menginterpretasikan sebuah hasil hingga mudah untuk dibaca secara lebih ringkas dan informatif. Data yang telah diolah sesuai dengan yang diinginkan, kemudian harus disajikan dalam bentuk penyajian data yang mudah dimengerti maknanya dan mudah diinterpretasi kan hasilnya. Penyajian data harus memperhatikan beberapa hal yaitu:

1. Data harus dapat berbicara sendiri
2. Data harus dapat menyelesaikan fenomena yang disajikan
3. Penyajian data harus mempertimbangkan kelaziman angka dan satuan ukur yang disajikan salsa
4. Satuan harus memperhatikan tujuan penyajian.

Berdasarkan dari beberapa hal diatas maka judul serta penjelasan penjelasan yang menyertai peran penting di dalam sistem penyajian. Judul harus bersifat ada singkat dan dengan sasaran. Begitu juga dengan keterangan. Keterangan yang menyertai judul sama sangat dimungkinkan untuk menjelaskan

fenomena yang tidak dapat dijelaskan secara mendetail di dalam sistem penyajian data. Penggunaan satuan juga harus diperhatikan sehingga seorang mudah untuk membaca. Misalnya satuan untuk tinggi tidak perlu menggunakan satuan meter akan tetapi cukup dengan satuan cm. Untuk data yang disajikan dapat berupa beberapa jenis yaitu sistem scoring (contoh: 87 range 0-100), prosentase (contoh: 87 %), maupun indeks (contoh: 0,87).

## a. Jenis jenis penyajian data

Penyajian data sendiri dapat berupa beberapa metode diantaranya yaitu tulisan, tabel, diagram batang, diagram lambang, diagram garis atau ogive.

### i. Tulisan

Meskipun penyajian data menggunakan tulisan tidak lazim digunakan penyajian data, namun demikian pemaparan data menggunakan tulisan dinilai sangat efektif untuk menunjukkan data yang tidak memiliki kompleksitas yang tinggi. Seperti pemaparan cara suatu daerah yang datanya hanya terdiri atas satu atau dua saja.

- Jumlah mahasiswa fikes sebanyak 521 orang
- Responden dalam penelitian ini berjumlah 89 orang
- Jumlah penduduk desa yang mengalami penyakit diabetes mellitus type II berjumlah 2122 orang

### ii. Diagram Batang

Dengan adanya teknologi atau perkembangan software pembuatan diagram batang memiliki kemudahan tersendiri untuk membuatnya. Hal ini tentu berbeda pada beberapa decade sebelumnya di mana pembuatan data harus menggunakan secara manual. Diagram batang sendiri cocok digunakan untuk mendeskripsikan suatu kenaikan tertentu.



Contoh Diagram Batang

### iii. Pictogram Atau Diagram Lambang

Diagram lambang digunakan untuk mendeskripsikan sesuatu agar memiliki tampilan yang lebih sederhana mempermudah untuk membaca sebuah data. Diagram lambang sendiri seringkali digunakan ada pamflet - pamflet, stiker i, dan media informatif lainnya. Sebagai contoh untuk menggambarkan jumlah kematian pada bayi lahir di tiap lainnya pada suatu daerah

Mortalitas larva udang windu yang diinfeksi bakteri bercahaya  
(*Vibrio harveyi*) selama 24 jam

Stadia Larva	Simbol	Kematian Setelah 24 jam (%)
Zoea		40
Mysis		29
Pasca Larva		16

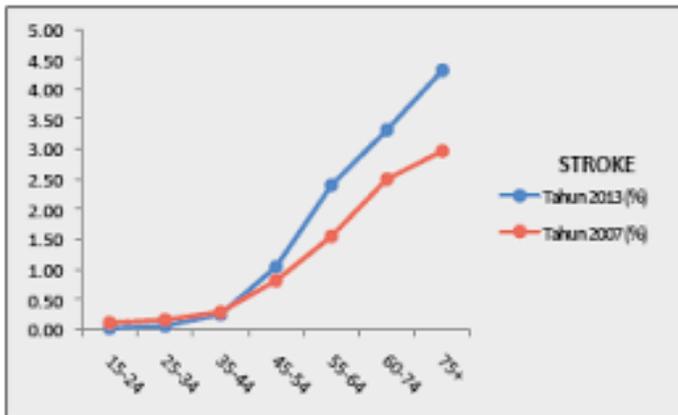
Keterangan =  = 8%

Contoh gambar pictogram

### iv. Diagram garis

Diagram garis merupakan yang seringkali digunakan untuk menyatakan suatu kenaikan atau penurunan sebuah kurva dari waktu ke waktu. Melalui diagram garis pola seseorang akan mudah mengetahui trend dari kenaikan atau penurunan suatu keadaan

Grafik 6. Prevalensi Penyakit Stroke Tahun 2007 dan 2013



Contoh gambar diagram garis

#### v. Diagram Lingkaran

Diagram Lingkaran merupakan diagram yang seringkali digunakan untuk menggambarkan jumlah / proporsi / komposisi dari keseluruhan data. Diagram lingkaran lebih cenderung digunakan untuk analisa data yang lebih sederhana. Penggunaan dengan penampilan yang lebih simple sehingga mudah dipahami oleh pembaca.

JENIS PERKERJAAN ORANG TUA WALI  
MAHASISWA D IV TLM



Contoh gambar diagram Lingkaran

### vi. Tabel Distribusi Frekwensi

Tabel distribusi frekuensi adalah penyajian data yang digolongkan dalam kelas-kelas urutan pada masing-masing. pada penerapannya tabel distribusi frekuensi dibagi menjadi dua kelompok yaitu data tunggal dan data kelompok. Tabel distribusi frekuensi data tunggal menggambarkan frekuensi setiap nilai utuh. Sedangkan tabel distribusi frekuensi kelompok menggambarkan nilai berdasarkan tentang-tentang tertentu. Ada kelemahan penggunaan tabel distribusi frekuensi kelompok yaitu hilangnya beberapa informasi akibat ditemukannya bintang pada tabel. Sebagai contoh kita memiliki nilai statistik sebanyak 50 anak. Data sebagai berikut:

82	66	75	88	90
75	87	91	51	65
75	87	92	55	65
56	76	67	88	92
54	77	87	95	67
53	63	74	89	96

84	96	55	64	77
65	67	68	69	66
64	62	73	64	79
55	62	73	84	92

**a) Table distribusi frekwensi data tunggal**

Tabel distribusi frekuensi data tunggal lebih condong kepada penghitungan data sesuai dengan frekuensinya

Nilai	Frekwensi
51	1
53	1
54	1
55	3
56	1
62	2
63	1
64	3
65	3
66	2
67	3
68	1
69	1
73	2
74	1

Nilai	Frekwensi
75	3
76	1
77	2
79	1
82	1
84	2
87	3
88	2
89	1
90	1
91	1
92	3
95	1
96	2

## b) Table distribusi frekwensi data Kelompok

Untuk membuat tabel distribusi frekuensi data kelompok pada data tersebut maka diperlukan beberapa istilah yang sering digunakan. Yaitu

### 1. Kelas

Kelas merupakan kelompok nilai atau data atau variabel yang membagi sejumlah data. Kelas biasanya disimbolkan dengan huruf K. Untuk menentukan jumlah kelas maka dilakukan perhitungan berdasarkan rumus Sturges

$$K = 1 + 3,3 \log N$$

K merupakan jumlah kelas dan n adalah banyaknya data. Yang perlu diingat di dalam perhitungan hasil perhitungan merupakan data diskrit artinya keseluruhan data yang dihasilkan harus dibulatkan ke atas. Sisanya hasil penghitungan menunjukkan nilai 6,7 bulatkan menjadi 7 dan apabila hasil penghitungan menghasilkan akar 6,2 tetap harus di bulatkan menjadi 7. Pada contoh di atas maka jumlah kelas yang bisa dihitung adalah

$$K = 1 + 3,3 \log N$$

$$K = 1 + 3,3 \log 50$$

$$K = 1 + 5,6$$

$$K = 6,6 \text{ maka di bulatkan menjadi } 7$$

### 2. Rentang dan Interval kelas

**Rentang** adalah selang jarak nilai tertinggi dengan jarak nilai terendah. **Interval Kelas** yaitu yang memisahkan kelas satu dengan kelas yang lain. Rentang disimbolkan dengan huruf R sedangkan interval disimbolkan dengan huruf K Untuk menghitung rentang kelas maka dapat dirumuskan

$$\text{Rentang} = \text{max} - \text{min}$$

$$\text{Interval Kelas} = R / K$$

Dengan demikian cara untuk mengidentifikasi rentang pada data diatas adalah dengan cara mendeteksi nilai maksimal dari data yaitu 96 dan nilai minimal dari data yaitu 51 sehingga penghitungan adalah sebagai berikut

$$\text{Rentang} = \text{max} - \text{min} = 96 - 51 = 45$$

$$\text{Interval Kelas} = R / K = 45 / 6.4 = 7$$

Dari analisa dengan mendeteksi kelas dan rentang maka dapat diperoleh gambaran berikut

no	nilai	frekwensi
1	51 - 57	
2	58 - 64	
3	65 - 71	
4	72 - 78	
5	79 - 85	
6	86 - 92	
7	93 - 99	

Gambar 2. 1 penerapan jumlah kelas dan nilai interval

Dalam penggunaan interval harus dihitung dari angka itu sendiri. Sangat tidak disarankan untuk menerapkan metode pengurangan. Misalnya interval pada kelas 1 adalah 57-51 maka hasilnya akan menjadi 6. Akan tetapi apabila kita menggunakan metode penghitungan 51, 52, 53, 54, 55, 56, 57 maka akan didapatkan jumlah 7.

### 3. Batas Kelas

Nilai-nilai kelas yang membatasi kelas satu dengan yang lain. Ada dua batas kelas, yaitu **batas kelas bawah** yang terdapat pada sisi kiri setiap kelas, serta **batas kelas atas** yang terdapat pada sisi kanan setiap kelas. Berdasarkan contoh diatas maka dapat diketahui beberapa nilai dari batas kelas. Contoh batas kelas atas pada kelas satu adalah 57 sedangkan pada batas kelas bawah pada kelas 1 adalah 51. Demikian pula batas kelas atas pada kelas ke kedua adalah 64 sedangkan pada batas bawah kelas bawah pada kelas ke 2 adalah 58.

### 4. Tepi kelas atau batas riil kelas

Batas kelas yang tidak memiliki lubang/ celah untuk angka tertentu antara kelas yang satu dengan yang lain, yang terdiri atas tepi bawah kelas serta tepi atas kelas. Penentuan tepi kelas adalah dengan mencari titik tengah antara batas atas kelas dengan batas bawah kelas di atasnya. Sebagai contoh nilai tepi kelas atas kelas 1 adalah

$$\text{Tepi kelas atas kelas 1} = \frac{\text{batas atas kelas 1} + \text{batas bawah kelas 2}}{2}$$

$$\text{Tepi kelas atas kelas 1} = \frac{57 + 58}{2} = 57.5$$

$$\text{Tepi kelas atas kelas 2} = \frac{\text{batas atas kelas 2} + \text{batas bawah kelas 3}}{2}$$

$$\text{Tepi kelas atas kelas 2} = \frac{64 + 65}{2} = 64.5$$

## 5. Titik Tengah Kelas

Titik Tengah Kelas angka atau nilai data yang terletak tepat di tengah suatu kelas. Titik tengah adalah representasi kelas yang bersangkutan.

$$TTK = \frac{1}{2} (\text{Batas atas} + \text{Batas bawah})$$

$$TTK = \frac{1}{2} (\text{tepi bawah kelas} + \text{tepi atas kelas})$$

Sehingga dengan memperhatikan rumus tersebut maka titik tengah kelas adalah

$$\text{Titik Tengah Kelas 1} = \frac{1}{2} (\text{Batas atas kelas 1} + \text{Batas bawah kelas 2})$$

$$\text{Titik Tengah Kelas 1} = \frac{1}{2} (51 + 57) = 108/2 = 54$$

### vii. Tabel distribusi frekwensi dengan kelas terbuka

Menggambaran nilai yang tidak terbatas pada titik-titik tertentu. Dengan menggunakan tabel distribusi dengan kelas terbuka maka dapat dimungkinkan memasukkan nilai yang jauh dari kelas biasanya nilai tersebut berada pada di bawah kelas yang pertama atau dari kelas yang terakhir

Nilai	Frekuensi
50 – 59	16
60 – 69	32
70 – 79	20
80 – 89	17

90 Atau lebih	15
Jumlah	100

### a) Tabel Distribusi Kurang dari dan Lebih dari

Tabel distribusi frekuensi kurang dari dan lebih dari memberikan gambaran tentang suatu nilai akumulatif yang dilihat dari kelas paling bawah maupun kelasnya tertinggi. Di Dalam penggunaannya batas yang digunakan adalah batas bawah kelas

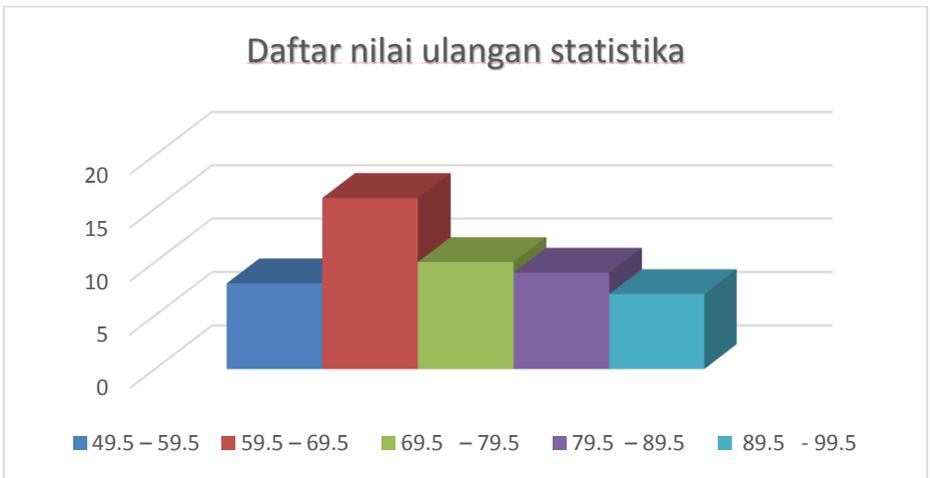
no	nilai			f	<	f	>	f
1	51	-	57	7	<51	0	>51	50
2	58	-	64	6	<58	7	>58	43
3	65	-	71	10	<65	13	>65	37
4	72	-	78	9	<72	23	>72	27
5	79	-	85	4	<79	32	>79	18
6	86	-	92	11	<86	36	>86	14
7	93	-	99	3	<93	47	>93	3
				50	<100	50	<100	0

### viii. Histogram Frekwensi

Mempermudah dalam pembacaan, kelemahan utama adalah tidak dapat diketahui data mentahnya, adapun penjelasan dari tiap bagian diagram adalah:

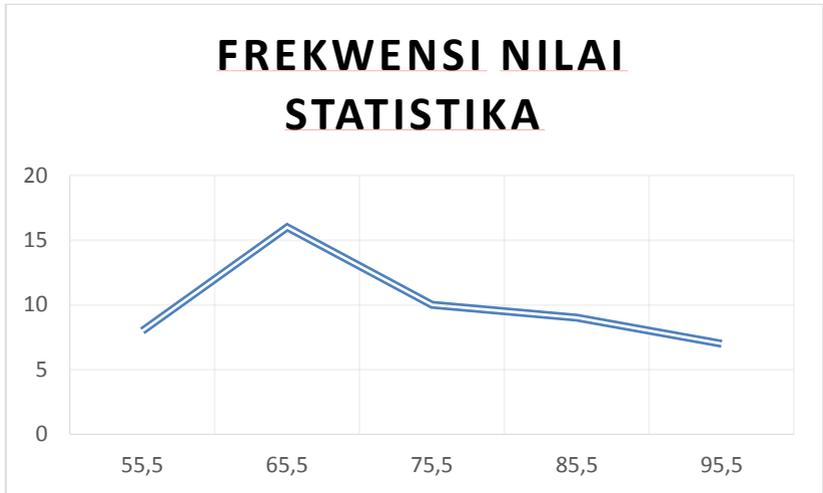
- Sisi tegak menggambarkan frekwensi kelas
- Sisi lebar bidang menggambarkan interval kelas atau lebar kelas

Apabila gambar berdempetan maka yang digunakan adalah tepi kelas, tetapi apabila gambar tidak mepet yang digunakan adalah batas kelas.



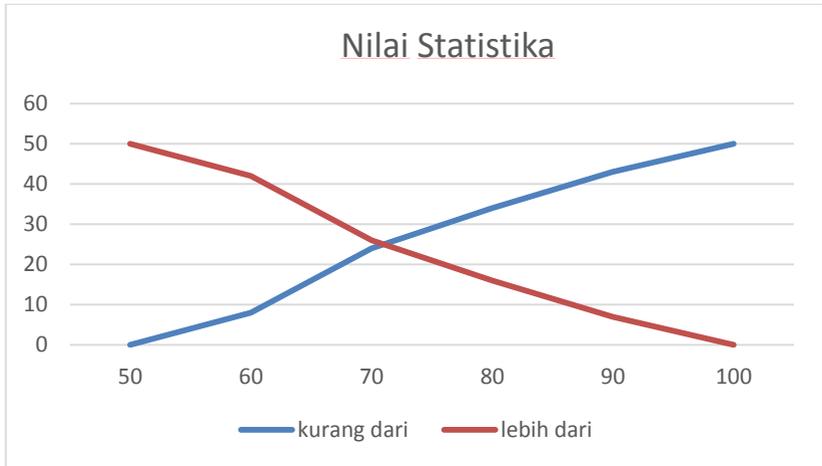
#### a) Poligon Frekwensi

Poligon frekwensi di bangun dengan cara menggabungkan titik-titik tengah KELAS yang ada



#### b) Ogive

Ogive adalah termasuk grafik garis yang menggunakan data dasar tabel distribusi frekuensi kumulatif kurang dari atau lebih dari sama dengan



### b) Ringkasan materi

Penyajian data dapat dilakukan mulai dari metode sederhana sampai kepada metode yang kompleks. Penyajian data dapat dilakukan caranya menggunakan tulisan, gambar batang, pictogram atau diagram lambang, diagram garis diagram lingkaran, atau tabel distribusi frekuensi. Namun demikian penggunaan secara tepat metode penyajian data menjadi kunci utama teresepresentasikan data hasil penelitian secara lebih baik.

### c) Latihan Soal

- Berikut ini data mengai jumlah trombosit pasien DBD (dalam ribu)

128	122	144	149	124	103
86	132	85	83	120	128
96	97	143	149	99	140
79	112	141	114	79	130
81	100	80	141	127	111

120	128	93	120	91	124
82	107	82	146	123	90
142	114	110	142	137	113
129	129	102	127	98	106
81	123	84	144	105	95
148	108	98	106	80	136
147	108	98	123	98	112

Buatlah table distribusi frekuensi kelompok, histogram frekwensi, dan ogive!

2. Jelaskan perbedaan fungsi antara diagram batang, diagram garis dan diagram lingkaran!

#### d) Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Kustituantio, B.& Badrudin, R. 1994, Statistika 1 (Deskriptif), Gunadarma
4. Murti, Bhisma, Penerapan Metode Statistik Non Parametrik Dalam Ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito



## BAB III. Ukuran Pemusatan Data

---

### Tujuan

Mahasiswa dapat memahami dan melakukan penghitungan ukuran pemusatan data baik itu menghitung rata-rata, modus, dan median selain itu Mahasiswa juga memahami hubungan antara nilai-nilai tersebut

---

### Materi

1. Interpretasi Pemusatan data
  2. Mean
  3. Modus
  4. Median
  5. Hubungan antara Mean, Modus, dan Median
- 

Ukuran pemusatan atau ukuran lokasi Pusat data merupakan sembarang ukuran yang menunjukkan nilai pusat gugus data atau nilai tunggal yang mewakili dari Sekumpulan data dan menunjukkan karakteristik secara komplit dari data tersebut. Sebelumnya kita telah mengetahui makna dari sebuah populasi dan sampel. Dalam ilmu statistik populasi memiliki beragam jenis nama lain seperti seluruh dan parameter. Sedangkan sampel memiliki nama lain antara lain sterol, cuplikan, contoh, bagian ngomong atau statistik itu sendiri.

Penggunaan lambang yang digunakan sebagai relatif berbeda misalkan rata-rata disimbolkan sebagai  $\bar{X}$  setengah rata-rata semoga dengan  $x$ . Standar deviasi pada parameter disimbolkan dengan Sigma sedangkan standar deviasi pada statistik disimbolkan

dengan huruf s. Ragam lovarian pada Barometer disimbolkan dengan Sigma kuadrat sedangkan pada statistik disimbolkan dengan x kuadrat. Di dalam bahasa Inggris statistik diterjemahkan menjadi statistics tanpa disertai dengan huruf s. Sedangkan statistika diterjemahkan menjadi statis tik yang disertai dengan huruf s. Secara komplit perbedaan tersebut dapat dilihat pada tabel berikut

### a) Interpretasi Pemusatan data

Anda tentu mengetahui Gambar apa yang ada pada gambar pada berikut ini.



Gambar Penampakan Pisang sebagai estimasi pemusatan data

Semua orang tahu bahwa gambar tersebut merupakan gambar pisang. Lalu bagaimana apabila kita disuruh untuk

mengintegrasikan ukuran dari pisang tersebut. Secara logika maka rata-rata ukuran pisang tersebut berada pada tengah-tengah tandon pisang. Atau mungkin kita berpikir yang lebih kompleks dengan cara mengambil beberapa yaitu satu setengah satu di atas satu dari bawah yang kemudian diambil rata-ratanya. Pemikiran yang kedua tentu memiliki metode yang efektif karena memenuhi persyaratan dari pengambilan data. Sedangkan metode yang pertama tentu tidak bisa digunakan pada beberapa kasus. Kalau misal kasus ini terjadi pada pisang mungkin pada metode pertama masih bisa digunakan. Tapi bagaimana apabila kita disuruh untuk menentukan rata-rata besar (kg) buah semangka yang ada pada satu kebun. Metode yang pertama tentu tidak bisa digunakan sehingga harus digunakan sampling. Terlepas dari itu semua yang kita bahas bukan pada metode ini akan tetapi angka yang dihasilkan dari proses tersebut. Angka yang dihasilkan harus dapat mewakili keseluruhan populasi yang diamati tentu untuk menghasilkan angka tersebut harus sesuai dengan metode ilmiah.

## 1. Rata-rata atau mean

Rata-rata lupakan suatu nilai data observasi Dari sebuah sampel yang dibagi dengan keseluruhan data. Rata-rata  $\bar{x}$  untuk sampel. Untuk menghitung rata-rata dengan cara menggunakan rumus sebagai berikut:

$\mu = \frac{\sum_{i=1}^N x_i}{N}$	<p><math>\mu</math> = rata-rata data observasi</p> <p><math>\Sigma</math> = jumlah</p> <p><math>x_i</math> = nilai data observasi</p> <p><math>N</math> = banyaknya data observasi</p>
------------------------------------	--

Namun demikian di dalam penggunaan data statistik diperlukan kehati-hatian. Contoh Apabila diketahui dari 50 bahan peledak dinamit yang berdasarkan penelitian dinamit tersebut akan meledak rata-rata 3 detik sejak pengaitnya dilepas. Namun Apakah kita saran untuk menunggu 3 detik secara tepat. Tentu kita akan segera mungkin melempar bahkan mungkin 1 detik setelah pengait

lepas. Hal ini dikarenakan kemungkinan dinamit tersebut meledak bisa jadi pada satu setengah detik atau 5 detik atau 4 detik atau 3 detik. Rata-rata merupakan hasil dari sekumpulan dinamik yang kemudian di ambil sebuah kesimpulan. Gambaran tersebut memberikan kita sebuah informasi akan krusialnya sebuah data. Singkat dalam penerapan sebuah penelitian rata-rata(x) selalu bareng dengan standar deviasi. Standar deviasi sendiri merupakan sebuah rentang yang mampu menggambarkan simpangan dari rata-rata.

Mean pada data distribusi frekwensi data kelompok dapat di hitung menggunakan rumus berikut

$$(\sum ( f. Xi))/n, \text{Xi adalah nilai tengah.}$$

Sehingga apabila diterapkan kepada contoh

	<b>nilai</b>	<b>f</b>	<b>xi</b>	<b>xi*f</b>	
	19	30	4	24.5	98
	31	42	4	36.5	146
	43	54	9	48.5	436.5
	55	66	14	60.5	847
	67	78	9	72.5	652.5
	79	90	7	84.5	591.5
	91	102	3	96.5	289.5
	total		50		3061

rata rata                      3061/50 = 61.22

## 2. Modus

Modus merupakan suatu himpunan bilangan yang bilangan tersebut paling sering muncul atau memiliki frekuensi maksimum. Data biasanya hanya terdiri atas 1 modus. Dalam kasus tertentu modus mungkin tidak ada atau lebih dari 1. Modus yang lebih dari

satu bisa disebut bimodal, trimodal, multimodal, dan seterusnya sebagai contoh data berikut

Data	Nilai Modus
12, 13, 14, 13, 12, 13, 21, 23, 12, 32, 43, 23, 12, 32, 12	12
12, 13, 14, 13, 12, 13, 21, 13, 12, 32, 43, 13, 12, 32, 12	12 & 13
12, 13, 14, 12, 13, 14, 12, 13, 14, 12, 13, 14, 12, 13, 14	Tidak bermodus

Modus pada table distribusi frekwensi dapat menggunakan rumus sebagai berikut

$$Modus = L_1 + \left[ \frac{b_1}{b_1 + b_2} \right] \times I$$

$$Modus = L_2 - \left[ \frac{b_2}{b_1 + b_2} \right] \times I$$

Dengan

L1 = batas tepi bawah kelas dari kelas modus.

L2 = batas tepi atas kelas dari kelas modus.

b<sub>1</sub> = selisih frekuensi kelas modus dengan frekuensi kelas sebelumnya

b<sub>2</sub> = selisih frekuensi kelas modus dengan frekuensi kelas sesudahnya

I = panjang kelas/ interval

nilai		f
19	30	4
31	42	4
43	54	9
55	66	14
67	78	9
79	90	7
91	102	3
total		50

Pada contoh kali ini kita coba menganalisis berapakah nilai modusnya. Pertama harus mencari di mana posisi frekuensi yang paling banyak pada contoh tabel distribusi frekuensi kelompok diatas frekuensi yang paling banyak terdapat pada kelas nomor 4 yaitu 14 sehingga ini menjadi patokan penentuan titik modus. sekaligus menjadi patokan bahwa penghitungan modus harus berfokus pada kelas tersebut. Apabila terdapat frekuensi yang sama pada beberapa kelas maka harus dihitung satu persatu dan artinya pada data tersebut terdapat modus yang lebih dari satu.

$$Modus = L_1 + \left[ \frac{b_1}{b_1 + b_2} \right] \times I$$

$$Modus = 54.5 + \left| \frac{5}{5+5} \right| \times 12 = 60.5$$

### 3. Median

Median merupakan data observasi yang berada di tengah-tengah data. Data-data bagian yang sama banyak, nilai data median diberi simbol Md. Di dalam menentukan median terdapat tahapan-tahapan harus ikuti yaitu

1. Urutkan data observasi dari kecil ke besar

2. Tentukan posisi median dengan cara menambahkan nilai populasi dengan angka 1 yang kemudian dibagi menjadi dua

$$\text{letak median} = \frac{N+1}{2}, N = \text{banyaknya data}$$

3. Tentukan nilai median  
Sebagai contoh data tunggal

90	92
84	95
87	89
88	96
76	77
91	79
75	82

Maka langkah yang harus dilakukan adalah dengan mengurutkan data tersebut dari data paling rendah ke data paling tinggi

no urut	data
1	75
2	76
3	77
4	79
5	82
6	84
7	87
8	88
9	89
10	90

11	91
12	92
13	95
14	96

Setelah itu gunakan rumus untuk menentukan nilai mediannya

$$\text{letak median} = \frac{N+1}{2}, N = \text{banyaknya data}$$

Letak median =  $(14 + 1) / 2 = 7,5$  maka nilai menunjukkan pada urutan ke 7,5 sehingga nilai median  $(87 + 88) / 2 = 87.5$

Contoh

$$\text{Median} = L_1 + \left[ \frac{\frac{n}{2} - (\sum f)_a}{f_{med}} \right] \times I$$

$$\text{Median} = L_2 - \left[ \frac{\frac{n}{2} - (\sum f)_b}{f_{med}} \right] \times I$$

Dimana

L1= batas tepi bawah kelas dari kelas median.

L2 = batas tepi atas kelas dari kelas median.

N = banyak data

$(\sum f)_a$  = jumlah frekuensi semua kelas yang lebih rendah dari kelas median

$(\sum f)_b$  = jumlah frekuensi semua kelas yang lebih tinggi dari kelas median

$f_{med}$  = frekuensi kelas median

$l$  = panjang kelas

nilai		f
19	30	4
31	42	4
43	54	9
55	66	14
67	78	9
79	90	7
91	102	3
total		50

Maka untuk menentukan median kita tentukan dahulu posisinya yaitu pada  $50/2 = 25$ . Setelah tahu bahwa nilainya pada 25 maka dilakukan identifikasi dengan cara menentukan table komulatif

nilai		f	Komulatif
19	30	4	4
31	42	4	8
43	54	9	17
55	66	14	31
67	78	9	40
79	90	7	47
91	102	3	50
total		50	

Maka dari identifikasi diperoleh median terdapat pada kelas 4 (18-31) sehingga dapat dilakukan penghitungan pada kelas tersebut .

$$Median = L_1 + \left[ \frac{\frac{n}{2} - (\sum f)_a}{f_{med}} \right] \times I$$

$$Median = 55,5 + \left( \frac{\frac{50}{2} - (17)}{14} \right) \times 12$$

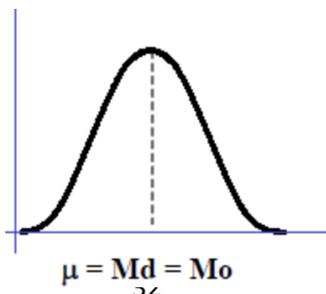
$$Median = 55,5 + 6,85$$

$$Median = 62,35$$

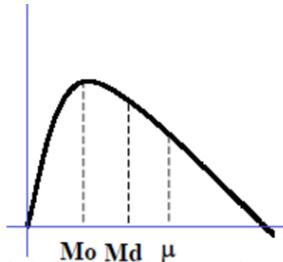
#### 4. Hubungan antara nilai rata-rata, median dan modus

Nilai rata-rata latihan dan modus adalah nilai digunakan untuk menghapus data ketika nilai tersebut sering disebut dengan ukuran kecenderungan terpusat atau measure of central tendency. Hubungan erat ketiganya seringkali digunakan untuk menentukan tingkat kemiringan sebuah kurva apa dengan lain ingin mengetahui kenormalan sebuah kurva. Ketentuan kurva tersebut dapat diuraikan dengan:

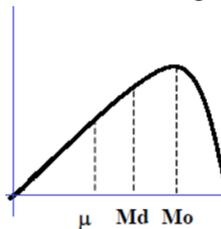
1. Jika rata-rata, median dan modus memiliki nilai yang sama, maka nilai rata-rata, median dan modus akan terletak pada satu titik dalam kurva distribusi frekuensi. Kurva distribusi frekuensi tersebut akan terbentuk simetris.



2. Jika rata-rata lebih besar dari median, dan median lebih besar dari modus, maka pada kurva distribusi frekuensi, nilai rata-rata akan terletak di sebelah kanan, sedangkan median terletak di tengahnya dan modus di sebelah kiri. Kurva distribusi frekuensi akan terbentuk menceng ke kiri.



3. Jika rata-rata lebih kecil dari median, dan median lebih kecil dari modus, maka pada kurva distribusi frekuensi, nilai rata-rata akan terletak di sebelah kiri, sedangkan median terletak di tengahnya dan modus di sebelah kanan. Kurva distribusi frekuensi akan terbentuk menceng ke kanan.



## b) Ringkasan Materi

Ukuran pemusatan data dapat diinterpretasikan dengan menggunakan mean, modus, dan Median. Hubungan dari ketiga parameter tersebut dapat dianalisis menggunakan kurva normal.

### c) Latihan soal

1. Berikut adalah umur lansia pada sebuah penelitian pengaruh pemberian diet terkontrol untuk menurunkan kadar kolesterol.

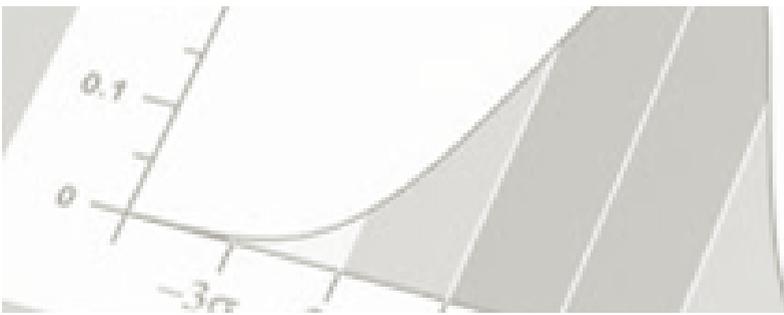
82	54	69	83	70	83
78	79	62	68	54	79
78	65	88	74	81	87
75	57	46	56	72	80
84	72	48	70	74	55
69	74	75	51	55	71

Tentukan Rata-rata Modus, Median!

2. Tentukan Rata-rata Modus, Median dari table distribusi frekwensi kelompok yang telah di buat pada soal nomor 1 bab 2 (penyajian data)!
3. Cobalah buat kurva dari soal nomor 1 untuk menentukan kemiringan dari data tersebut!

### d) Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Kustitunto, B.& Badrudin, R. 1994, Statistika 1 (Deskriptif), Gunadarma
4. Murti, Bhisma, Penerapan Metode Statistik Non Parametrik Dalam Ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito



## Bab IV. Ukuran penyebaran data dan letak ukuran penyebaran data

---

### Tujuan

Mahasiswa dapat memahami dan melakukan penghitungan ukuran pemusatan data baik itu menghitung rata-rata, modus, dan median selain itu Mahasiswa juga memahami hubungan antara nilai-nilai tersebut

---

### Materi

1. Range atau jangkauan
  2. Standart Deviasi
  3. Ragam
  4. Quartil, Desil dan Presentil
- 

**U**ukuran penyebaran data dan letak ukuran penyebaran data merupakan penggambaran persebaran kita yang ada pada suatu kumpulan data.

### a. Range Atau jangkauan

Range/ Jangkauan adalah selisih antara nilai maksimum dan nilai minimum yang terdapat dalam data. Dapat dihitung dengan cara nilai dari Sekumpulan data dikurangi dengan nilai terendah dari data tersebut. Range akan selalu memiliki nilai positif. Jelaskan melalui beberapa cara.  $R = 0$ , menunjukkan bahwa data terbesar sama dengan data terkecil, akibatnya semua data memiliki harga yang sama,  $R$  kecil, memberikan informasi bahwa data akan mengumpul di sekitar pusat data,  $R$  besar, menyatakan bahwa

paling sedikit ada satu data yang harganya berbeda jauh dengan data lainnya

Jangkauan dapat dihitung dengan rumus:

$$R = X \text{ maks} - X \text{ min}$$

## I. Range Atau jangkauan data tunggal

Untuk menghitung nilai range pada data tunggal kita gunakan ilustrasi sebagai berikut:

82	54	69	83	70	83
78	79	62	68	54	79
78	65	88	74	81	87
75	57	46	56	72	80
84	72	48	70	74	55
69	74	75	51	55	71

Mada dari data tersebut dapat terlebih dahulu table distribusi frekwensi kelompoknya

Nilai max: 88

Nilai min: 46

range: Nilai Max- Nilai min = 42

## II. Range Atau jangkauan data distribusi frekwensi kelompok

Data table distribusi frekwensi kelompok=

**R= batas bawah kelas terakhir - batas bawah kelas pertama,**  
atau

**R= nilai tengah kelas tertinggi - nilai tengah kelas terendah**

nilai		f	Kumulatif
19	30	4	4
31	42	4	8
43	54	9	17
55	66	14	31
67	78	9	40
79	90	7	47
91	102	3	50
total		50	

Maka range dari data tersebut adalah

**R= batas bawah kelas terakhir - batas bawah kelas pertama,**

R= 91-19 adalah 72

## b. Standart Deviasi

Standar deviasi atau simpangan standar merupakan akar dari sejumlah deviasi kuadrat dari bilangan-bilangan dibawah bilangan atau angka dari rata-rata deviasi kuadrat. Cara kalimat mungkin susah untuk dipahami Namun demikian untuk menggambarkan standar deviasi bisa dilakukan sebagai berikut data dengan angka 6, 7, dan 8 maka akan memiliki rata rata 7 dan standart deviasi 1. Data 5,7, dan 9 maka akan memiliki rata-rata 7 akan tetapi besar standar deviasi nya berubah menjadi dua. Akan tetapi angka 7, 7, & 7 akan memiliki rata rata 7 standar deviasi. Untuk mempermudah nyaah pemahaman Anda ada angka 1, 7, & 9 maka akan memiliki nilai rata-rata dan standar deviasi 6. Jadi untuk mengungkapkan standar deviasi merupakan tahap yang penting karena menggambarkan kan variasi dari sebuah data. Gambaran di atas menunjukkan rata-rata menunjukkan nilai 7 akan tetapi dengan standart deviasi yang berbeda-beda maka interpretasi data juga akan berbeda-beda. Oleh karena itu hendaknya setiap pemusatan data digambarkan pula nilai standart deviasinya (biasanya diberi tanda  $\pm$ ).

Standart deviasi dapat dihitung berdasarkan proporsinya apakah data berasal dari populasi atau berasal dari sampel. Perhitungannya berbeda apabila menggunakan standar deviasi yang dari populasi kenapa standar deviasi dari sampel. Menghitung standar deviasi dapat menggunakan rumus sebagai berikut:

Untuk standar deviasi pada data tunggal:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Untuk standar deviasi pada data Distribusi frekwensi kelompok:

$$\sigma^2 = \frac{\sum_{i=1}^N f(x_i - \mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^n f(x_i - \bar{x})^2}{n-1}$$

sebagai contoh Berikut ini adalah data tunggal sederhana

Y	y- $\bar{y}$	(Y- $\bar{y}$ ) <sup>2</sup>
5	-5	25
7	-3	9
8	-2	4
9	-1	1
10	0	0
21	11	121
$\Sigma = 60$		160

Maka untuk penghitungan Standart Deviasi pada data tersebut adalah

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

$$s^2 = \frac{160}{5} = 32$$

$$s = \sqrt{32} = 5.657$$

### c. Ragam

Ragam atau varian ragam merupakan jumlah kuadrat dari selisih nilai observasi yang nilai rata-rata observasi. standar deviasi yang kita pasti atas merupakan akar dari ragam. Sehingga rumus untuk mencari standar deviasi maupun ragam adalah mirip dengan standart deviasi. Sebagai ilustrasi diatas ragam dari data adalah

Ragam:

$$s^2 = \frac{160}{5} = 32$$

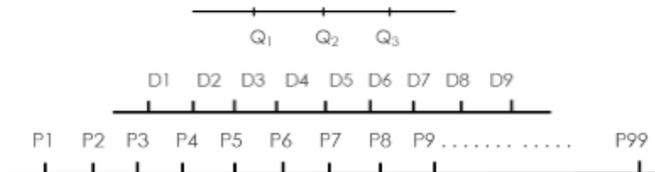
Standart Deviasi

$$s = \sqrt{32} = 5.657$$

### d. Kuartil, Desil, dan presentil

#### 1. Kuartil, Desil dan Presensil Dari data tunggal

Kuartil merupakan gugus data yang apabila dibagi menjadi 4 bagian dengan proporsi sama tanya dan disusun menurut urutan nilainya. Desil merupakan gugus data yang apabila dibagi menjadi 10 bagian dengan proporsi sama tanya dan disusun menurut urutan nilainya. Presentil merupakan gugus data yang apabila dibagi menjadi 100 bagian dengan proporsi sama tanya dan disusun menurut urutan nilainya.



Gambar 1. Ilustrasi penggambaran pembagian Kuartil desil dan presentil

$$Q_k = \frac{k(N+1)}{4}$$

$$D_k = \frac{k(N+1)}{10}$$

$$P_k = \frac{k(N+1)}{100}$$

- Q<sub>k</sub>** = Quartil ke k
- D<sub>k</sub>** = Desil ke k
- P<sub>k</sub>** = Presentil ke k
- K** = kuartil ke1,2,3
- N** = Banyak data

Untuk mempermudah pemahaman marilah kita mencoba menghitung Q= 1 dari nilai ulangan mata kuliah hematologi 13 anak berikut

40	47	40	90	60	70	55	50	35	80	46	15	25
----	----	----	----	----	----	----	----	----	----	----	----	----

Maka langkah awal yang harus dilakukan adalah dengan mengurutkan nilai tersebut dari nilai rendah ke tinggi

Nomor Urut	Nilai
1	15
2	25
3	35
4	40

5	40
6	46
7	47
8	50
9	55
10	60
11	70
12	80
13	90

Sehingga dapat dicari menggunakan rumus berikut

$$Q_1 = \frac{1(N+1)}{4}$$

$$Q_1 = \frac{1(14+1)}{4}$$

$$Q_1 = 3.75$$

Maka posisi Kuartil 1 berada pada data ke 3.75. posisi berada diantara data ke 3 dan ke 4. Maka nilainya diantara 35 dan 40. Maka  $35 + 0.75(40-35)$ . Sehingga nilai dari Kuartil 1 adalah  $35 + 3.75 = 58.75$ .

Untuk menentukan data distribusi kelompok kita coba dari data sebelumnya

	<b>nilai</b>	<b>f</b>	<b>xi</b>	<b>xi*f</b>	
	19	30	4	24.5	98
	31	42	4	36.5	146
	43	54	9	48.5	436.5
	55	66	14	60.5	847
	67	78	9	72.5	652.5
	79	90	7	84.5	591.5
	91	102	3	96.5	289.5
	total		50		3061

Kita tentukan dari Quartil 1 dari data tersebut.

Gambaran sederhana data Kuartil tidak berkelompok

$$Q_N = B_1 + \left[ \frac{n \cdot \frac{k}{4} - (\sum f)_{b1}}{f_Q} \right] \times I$$

Di mana

$B_1$  = batas tepi bawah kelas dari kelas kuartil ke-N

$n$  = banyak data

$(\sum f)_{b1}$  = jumlah frekuensi semua kelas sebelum kelas kuartil ke N

$f_Q$  = frekuensi kelas kuartil ke-N

$I$  = Interval kelas

$$D_N = B_1 + \left[ \frac{n \cdot \frac{k}{10} - (\sum f)_{N1}}{f_D} \right] \times I$$

$$P_N = B_1 + \left[ \frac{n \cdot \frac{k}{100} - (\sum f)_{N1}}{f_D} \right] \times I$$

## e. Ringkasan Materi

Range atau jangkauan mendeskripsikan rentang data dari data terendah sampai data tertinggi. Standart deviasi mendeskripsikan sebaran data, selain itu dapat pula disampaikan dengan nilai Quartil, Desil dan Presentil.

## f. Soal Latihan

1. Berikut adalah nilai SGOT pada pasien yang cuci darah

Inisial Pasien	Nilai SGOT
AS	18
DD	20
KK	10
FAS	14
SA	26
XZ	29
XA	13
LK	24
RE	10
P	33
DS	38
DSA	21
EW	16
SDW	29
LL	29

Tentukan nilai

1. Range
2. SD
3. Ragam
4. Q3
5. D12
6. P21

2. Berikut adalah sebaran umur yang berkunjung pada rumah sakit pada Minggu ke 2 Bulan Juli 2020.

nilai	f	Kumulatif
-------	---	-----------

19	30	4	4
31	42	4	8
43	54	9	17
55	66	14	31
67	78	9	40
79	90	7	47
91	102	3	50
total		50	

Tentukan nilai

1. Range
2. SD
3. Ragam
4. Q3
5. D12
6. P21

#### e) Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Kustituantio, B.& Badrudin, R. 1994, Statistika 1 (Deskriptif), Gunadarma
4. Murti, Bhisma, Penerapan Metode Statistik Non Parametrik Dalam Ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito



## Bab V. Populasi, sampel, sampling, besar sampel

---

### Tujuan

**Mahasiswa dapat memahami dan dapat menentukan karakter populasi, sampel, sampling, besar sampel**

---

### Materi

---

1. **Populasi**
  2. **Sampel**
  3. **Besar sample**
  4. **Jenis jenis sampling**
- 

**A**da dua istilah yang memiliki kata yang mungkin kita sering dengar akan tetapi susah membedakannya. Apa itu? Kata tersebut adalah Jumlah sampel dan besar sampel. Apakah anda tahu? Dalam hal ini akan dibuatkan sebuah simulasi. Si A melakukan sebuah penelitian dengan judul perbedaan kadar SGOT dan SGPT pada perokok dan non perokok pasien komplikasi asma. kemudian dating si B, lalu si B bertanya kepada si A. Berapa jumlah sampel kamu? Si A menjawab 2. Sedangkan si B bertanya lagi berapa besar total sampel kamu?. Jawab Si A “ 80 terdiri dari 60 sampel perokok dan 20 sampel non perokok.

Dari jawaban tersebut dapat diketahui bagaimana pengaruh pemahaman yang kurang dapat menjadi rancu. Missal selama ini istilah jumlah sample seringkali ambigu dengan besar sampel. Sedangkan pengertian dari jumlah sampel adalah

*“sample size is the number of sampling units in to which as aggregate is divided for the purpose sampling. each unit being regarded as individual an indivisible when the selection is made”*

Besar sampel adalah jumlah dari unit sampel yang kemudian dikelompokkan berdasarkan katagori katagori sesuai dengan kebutuhan penelitian. Masing masing unit tersebut tidak dapat terpisah antara satu dengan yang lain .

Di dalam penelitian penentuan besar sampel sebenarnya merupakan sesuatu yang urgent. Hal tersebut terkait dengan ketentuan dan kepercayaan dari pembaca atau pengguna penelitian yang dihasilkan. Tentu di dalam penentuan besar sampel mengacu pada beberapa aspek. Termasuk kemudahan akses sampel, waktu pengadaan penelitian, biaya dimiliki, dan ketersediaan jumlah sampel. Kadang-kadang sampel yang susah dicari menjadi pertimbangan sendiri untuk mennentukan jumlah sampe; misalkan penelitian mengenai penderita diabetes tentu berbeda dengan penelitian mengenai Sindrom Down. Penentuan jumlah sample menjadi kunci utama mengingat apabila sampel yang ada di lapangan sangat sedikit maka penelitian tersebut tidak akan selesai dengan cepat.

Penelitian lain bahkan mungkin hanya menggunakan satu sampel saja karena kesulitan yang amat sangat di dalam mendapatkan sampel tersebut. Contoh penelitian mengenai keadaan kromosom pada penderita sindrom klinefelter. Melihat dari aspek yang digunakan maka untuk menentukan sindrom klinefelter merupakan suatu permasalahan sendiri dan Butuh Waktu serta kemampuan di dalam negosiasi kepada pasien atau sampel. Belum masih mengingat online tersebut membutuhkan waktu yang tidak sedikit maka dapat dimungkinkan penelitian disebut Impossible untuk dilakukan.

## a. Populasi dan Sampel

Seringkali kita mendengar istilah sampel dan populasi. Populasi dapat dikatakan sebagai keseluruhan objek yang menjadi target penelitian. Sedangkan sampel merupakan cuplikan atau bagian atau contoh dari populasi tersebut yang digunakan untuk pengukuran dan interpretasi populasi. Di dalam penelitian alasan penggunaan sampel diantaranya adalah populasi penelitian yang terlalu besar keterbatasan sumber daya untuk menjadikan populasi sebagai keseluruhan objek penelitian atau bahkan karena populasi dinilai terlalu besar dikawatirkan terjadi error pada ada peneliti itu sendiri. Meskipun sampel merupakan cuplikan atau bagian dari populasi diharapkan sampel dapat memberikan gambaran autentik dari sebuah hasil untuk mendeskripsikan sebuah populasi.

Penentuan sampel mengikuti kaidah-kaidah namun demikian seringkali menjadi prioritas peneliti. Akan tetapi kepercayaan pengguna menjadi penting untuk diperhatikan sehingga seringkali penelitian mengacu pada kaidah-kaidah yang sudah berlaku atau mengacu kepada ilmuwan. Bisa saja peneliti menggunakan sampel sebanyak 10 orang namun yang perlu diperhatikan Apakah dengan 10 orang tersebut pembaca atau pembelinya sudah yakin dengan hasil yang dikeluarkan dan apakah tidak ada eror dari sampel yang digunakan.

## b. Besar sampel

Penentuan besar sampel sebenarnya banyak sekali teori yang menyatakan hal tersebut. Taman pada buku kali ini kita akan membahas mengenai beberapa yang paling sering digunakan di dunia kesehatan khususnya penelitian yang bersifat sains ataupun survei.

### I. Penelitian Laboratorik

Penelitian laboratory biasanya digunakan ketika penelitian membutuhkan uji coba di dalam laboratorium. Sebagai contoh peneliti ingin

mengetahui pengaruh dari pemberian rebusan daun ubi jalar terhadap peningkatan trombosit pada tikus galur wistar. Jalan uji coba ini maka diperlukan sebuah ulangan untuk menjadikan penelitian dapat dipercaya. Ulangan tersebut digunakan untuk meminimalisir error yang terjadi ketika penelitian. Adapun rumus yang digunakan untuk penelitian ini adalah rumus Federer.

$$(t - 1)(r - 1) \geq 15$$

di mana:

t (treatment) = jumlah perlakuan;

r (repeat) = jumlah ulangan.

R dan T adalah data **diskrit**

Sebagai contoh apabila penelitian yang dilakukan menggunakan 4 konsentrasi maka aplikasi dari rumus tersebut adalah sebagai berikut

$$(t - 1)(r - 1) \geq 15$$

$$(3-1)(r-1) > 15$$

$$2(r-1) > 15$$

$$2r - 2 > 15$$

$$2r > 15 + 2$$

$$R > 17/2$$

5.5. . . . Karena data diskrit menjadi 9

## II. Penelitian Laboratorium secara umum

- a. Ukuran sampel lebih dari 30 dan kurang dari 500 adalah tepat untuk kebanyakan penelitian.

- b. Jika sampel dipecah ke dalam sub-sampel (pria/wanita, junior/senior, dsb), ukuran sampel minimum 30 untuk tiap kategori adalah tepat.
- c. Dalam penelitian multivariate (termasuk analisis regresi berganda), ukuran sampel sebaiknya 10 kali lebih besar dari jumlah variabel dalam penelitian.
- d. Untuk penelitian eksperimental sederhana dengan control eksperimen yang ketat, penelitian yang sukses adalah mungkin dengan ukuran sampel kecil antara 10 sampai dengan 20 buah.

Contoh judul penelitian yang dapat digunakan penelitian dengan menggunakan ketentuan ini adalah

1. Perbedaan kadar nikotin Antara perokok dan perokok
2. Analisis kandungan timbah darah pada pekerja bangunan yang bekerja lebih dari 5 tahun
3. Perbedaan SGOT dan SGPT pada pekerja malam dan pekerja siang pabrik

### **III. Penelitian Survei**

Penelitian survei seringkali digunakan untuk mendeteksi jumlah besar sampel sebagai contoh Apabila Anda ingin mengetahui respon masyarakat terhadap kehidupan PHBS maka salah satu rumus yang digunakan adalah rumus slovin.

■ Rumus Slovin

$$n = \frac{N}{1 + Ne^2}$$

Keterangan:

n= Ukuran sampel,

N= Ukuran populasi,

e= Persen kelonggaran ketidaktelitian

karena kesalahan pengambilan sampel yang

masih dapat ditolerir atau diinginkan

(misalnya, 1 %, 5 %, 10 % ).

Berdasarkan rumus tersebut sebenarnya dapat diketahui dengan mudah nilai besar sample menggunakan table di bawah ini :

POP	Sig		
	0.05	0.03	0.01
50	45	48	49
55	50	53	55
60	53	57	59
65	59	63	65
70	59	67	69
75	64	72	75
80	68	76	79
85	71	82	85
90	75	84	89
95	79	89	95
100	79	93	99
110	90	100	108
120	93	110	118
130	99	116	128
140	106	125	138
150	110	133	148
160	117	140	157
170	119	150	167
180	126	155	177

POP	Sig		
	0.05	0.03	0.01
320	179	250	310
340	184	265	330
360	190	272	350
380	199	284	367
400	199	299	387
420	210	302	402
440	210	316	422
460	210	327	442
480	219	333	457
500	219	350	477
550	230	367	522
600	239	385	567
650	250	414	610
700	250	433	653
750	259	450	699
800	259	467	744
850	270	484	787
900	270	499	828
950	279	516	867

POP	Sig		
	0.05	0.03	0.01
2400	339	750	1932
2600	339	767	2050
2800	339	784	2183
3000	339	799	2299
3500	350	833	2581
4000	359	850	2850
4500	359	867	3089
5000	359	884	3299
6000	359	916	3699
7000	370	933	4050
8000	370	950	4399
9000	370	967	4650
10000	370	967	4899
15000	379	999	5899
20000	379	1016	6499
30000	379	1033	7299
40000	379	1050	7750
50000	379	1050	8099
75000	390	1067	8550

190	130	165	187
200	135	168	197
210	139	180	206
220	140	184	216
230	150	191	226
240	150	199	236
250	155	204	246
260	159	214	253
270	159	216	263
280	170	225	273
290	170	233	283
300	170	234	293

1000	279	516	906
1100	290	550	995
1200	299	567	1071
1300	299	599	1150
1400	310	616	1228
1500	310	633	1299
1600	310	650	1379
1700	319	667	1450
1800	319	684	1522
1900	319	699	1599
2000	330	699	1657
2200	330	733	1799

100000	399	1084	8799
200000	399	1084	9199
300000	399	1084	9350
400000	399	1084	9399
500000	399	1084	9499

Selain menggunakan rumus slovin dapat pula menggunakan rumus yang lebih reliable yaitu *Krejcie, R. & Morgan D. (1970)*.

$$S(n) = \frac{X^2NP(1-P)}{d^2(N-1)+(X^2P(1-P))}$$

s = ukuran sampel yang diperlukan

N = ukuran populasi yang dikenal pasti

P = perkadaran populasi diandaikan 0.5 kerana magnitud ini akan menghasilkan ukuran sampel yang maksimum.

d = darjah ketepatan yang maksimum (0.05)

x<sup>2</sup> = nilai khi square, db 1 yaitu 3.841

#### IV. Untuk populasi besar

Penelitian dengan populasi besar biasanya populasi berasal dari karakteristik yang sudah biasa ada di masyarakat. Seperti kebiasaan merokok atau petani di pedesaan atau pengunjung rumah sakit. Penelitian yang demikian dapat menggunakan rumus

$$N = \frac{Z^2 \cdot p \cdot q}{d^2}$$

untuk populasi tak terbatas atau lebih besar dari 10.000

Keterangan :

N = besarnya sampel

Z = nilai standar normal yang berkaitan dengan nilai  $\alpha$

p = probabilitas terjadinya kejadian (bila tidak ada data dibuat 50%)

q = probabilitas tidak terjadinya kejadian (bila tidak ada data dibuat 50%)

q = 1 - p

d = tingkat kecermatan yang diinginkan

sedangkan untuk populasi kurang dari 10.000

$$n_f = \frac{n}{1 + \frac{n}{N}}$$

Keterangan :

- N : besarnya sampel untuk populasi lebih dari 10.000
- n : besarnya populasi
- $n_f$  : besarnya sampel dengan populasi kurang dari 10.000

Atau ketika ingin membandingkan suatu parameter dapat menggunakan rumus

Keterangan :

- N : besarnya populasi
- Z : nilai standar normal yang berkaitan dengan nilai  $\alpha$
- p : probabilitas terjadinya kejadian (bila tidak ada data dibuat 50%)
- d : tingkat kecermatan yang diinginkan
- n : besarnya sampel

### c. Metode sampling

Metode pengambilan sampel juga menjadi kunci utama selain besar sampel untuk menjamin penelitian tersebut dapat dipercaya dan meminimalisir error yang dihasilkan. Sebagai contoh Ketika anda ingin mengetahui gambaran histologi yang diberi pengaruh daun tapak dara kemudian fokus peneliti ingin mengetahui pengaruh terhadap gambaran histologi kanker liver untuk penelitian pada tikus. Dengan menggunakan rumus Federer

dapat diketahui bahwa besar sampel yang harus digunakan untuk 4 perlakuan adalah 9 tikus. Pada penelitian tersebut dapat disimpulkan bahwa minimal tikus yang digunakan berjumlah 36. Dalam penelitian ini tentu tidak menggunakan atau mengambil tikus dengan jumlah yang pas akan tetapi perlu dilebihkan sedikit untuk menghindari tikus yang benar-benar tidak memenuhi kriteria. Lalu Bagaimana cara pengambilan sampel?, apakah anda hanya membagi pada setiap kotak atau kelompok sampel atau unit sampel dengan besar sampel secara begitu saja?.

Tidakkah anda berpikir bahwa yang satu apabila diambil di wadah A. kemudian tikus kedua diambil lagi di wadah A. Kemudian tikus yang ketiga diambil dan ditaruh di wadah dan seterusnya sampai wadah a memiliki 9 tikus kemudian dilanjut dengan an-najah dengan cara yang sama. Mungkin cara demikian adalah cara yang paling sering digunakan namun Tidakkah anda berpikir bahwa ketika anda menyelesaikan pertama kali pada wadah a apakah terdapat kemungkinan bahwa tikus pada wadah tersebut merupakan tikus tikus yang mudah yang notabene merupakan tikus yang kurang aktif. Untuk meminimalisir kesalahan tersebut maka digunakan metode tertentu sehingga penelitian dapat lebih reliabel atau dapat dipercaya. Maka untuk menentukan metode sampling yang tepat peneliti mengkategorikan bahwa sampling minimal dapat dikategorikan menjadi Probability sampling dan nonprobability sampling. Probability sampling digunakan ketika sampel atau populasi pada penelitian tersebut diketahui jumlahnya secara keseluruhan. Sedangkan non probability sampling diketahui ketika sampel jumlahnya belum dapat ditentukan atau terbatas.

## **a. Probability Sampling**

### **i. Pengambilan sampel acak sederhana (Simple Random Sampling)**

Pengambilan sampel acak sederhana menekankan sistem pengambilan sampel yang didasarkan pada angka (bilangan) yang muncul. Langkah-langkah; Menentukan nomer untuk setiap individu dalam populasi; Melakukan proses acak (dapat dilakukan dengan tabel bilangan acak) untuk mendapatkan  $n$  angka antara 1 dan  $N$ .

- a). Tentukan besarnya populasi
- b). Tentukan besarnya sampel yang diinginkan
- c). Buat sampling frame / kerangka sampling / list / daftar unit populasi lengkap dengan nomor urutnya.
- d). Tulis nomor pada kertas undian, kemudian dilinting atau tulis nomor pada bola atau yang lainnya.
- e). Ambil / lakukan undian untuk mengambil lintingan kertas sebanyak sampel yang diinginkan dengan mata tertutup.
- f). Buat daftar nomor sampel yang terpilih.

Selain metode diatas dapat pula digunakan kertas random. Caranya mirip dengan cara menentukan berapa jumlah sample terlebih dahulu kemudian diteruskan dengan menjatuhkan bolpoint kedalam area kertas acak. Secara detail dapat diikuti langkah berikut

- a) Tentukan besarnya populasi
- b) Tentukan besarnya sampel yang diinginkan

- c) Buat sampling frame / kerangka sampling / list / daftar unit populasi lengkap dengan nomor urutnya.
- d) Ambil tabel bilangan random. Bila tabel bilangannya 5 digit sedangkan nomor sampel terbesar hanya 2 digit, maka nomor tabel bilangan random yang dipakai 2 digit nomor yang di belakang.
- e) Tanpa melihat nomor pada tabel bilangan random, jatuhkan pensil atau penunjuk pada sembarang nomor pada tabel bilangan random.
- f) Nomor yang tepat dengan pensil sebagai nomor sampel yang pertama.
- g) Sampel kedua, ketiga, keempat dan seterusnya, didapatkan dengan menggeser pensil ke atas atau ke bawah atau ke kanan atau ke kiri. Bila sampai ujung nomor habis, pensil geser ke atas atau ke bawah, atau ke kanan atau ke kiri kemudian pensil geser lagi kebalikan semula. Yang penting ada konsistensinya dalam pergerakan pensil. Keadaan tersebut terus bergerak sampai jumlah kebutuhan sampel terpenuhi.
- h) Tulis nomor yang menjadi sampel

## ii. Pengambilan sampel Sistematis Random Sampling (systematic random sampling)

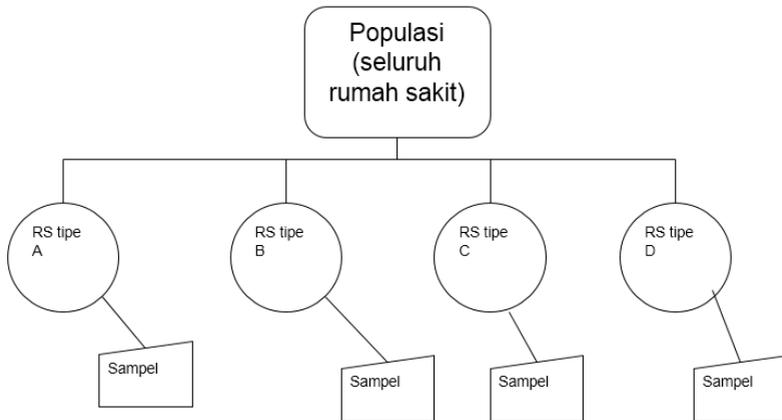
Pengertian sistematis random sampling pengambilan sampel menekankan pada sistem interval dimana interval ditentukan dari pembagian antara jumlah populasi dibagi dengan jumlah sampel. Adapun langkah langkah adalah sebagai berikut

- a) Pengambilan sampel ini lebih menekankan pada sistem interval
- b) Langkah-langkah;
- c) Memberi angka (nomer) untuk seluruh populasi.
- d) penentuan angka didasarkan proporsi terbanyak-terkecil.
- e) Interval sampel.
- f) Melakukan proses acak untuk interval pertama.
- g) Hasil acak interval pertama sebagai sampel no 1.

## iii. Pengambilan sampel berstrata (*Stratified sampling*)

Pengambilan sampel dengan stratifikasi lebih menekankan dan memperhatikan sub-klaster yang ada. Pembagian sub-klaster dapat didasarkan pada karakteristik atau tipe dari populasi. Sebagai contoh seorang ATLM ingin melakukan penelitian mengenai kepuasan pelanggan rumah sakit yang ada di Jawa Timur. Maka dilakukan pengelompokan pengelompokan sesuai dengan

tipenya. Kemudian masing masing di turunkan sesuai dengan proporsi kriteria sampel. Metode seperti itu dapat dilihat melalui gambar berikut



Berdasarkan gambar diatas maka langkah langkah yang dapat dilakukan untuk melakukan model sampling tersebut adalah sebagai berikut:

- a) Menentukan populasi sasaran.
- b) Menentukan sub-klaster yang dapat didasarkan pada karakteristik populasi. Ini lebih sering dikenal dengan alokasi sampling.
- c) Melakukan proses random (acak) untuk setiap sub klaster.
- d) jumlah Sampel yang terambil untuk setiap sub-klaster adalah sama.

#### **iv. Pengambilan sampel Klaster**

suatu rangka yang terdiri dari klaster-klaster unit pencacahan, dibagi menjadi beberapa klaster yang saling pisah klaster tidak harus sehomogen mungkin.

## **b. Non Probability Sampling**

(Pengambilan sampel non random. Pengambilan sampel non random dapat dilakukan menggunakan beberapa cara diantaranya sampling kuota, insidental, sampling purposive penuh, dan Snowball sampling.

### **i. Sampling kuota**

Sampling kuota dilakukan secara acak sesuai dengan batasan jumlah yang telah ditentukan oleh peneliti sehingga ketika kuota sedang memenuhi maka responden selanjutnya tidak memiliki kemungkinan untuk menjadi sampel.

### **ii. Sampling insidental**

Sampling insidental dilakukan dengan memperhatikan rentang waktu yang diberikan oleh peneliti sehingga berapapun jumlah sampel yang masuk akan menjadi objek penelitian. Perbedaan utama sampling kuota dengan insidental adalah sampling kuota menitikberatkan pada jatah dari jumlah sampel sedangkan sampling insidental menitikberatkan pada rentang waktu yang diberikan oleh peneliti.

### **iii. Purposive sampling**

Jenis sampling yang satu ini seringkali digunakan di dunia kesehatan karena sampel memiliki karakteristik khusus yang dilakukan atau ditentukan oleh peneliti itu sendiri atau berdasarkan pendapat ahli. Sebagai contoh seseorang ingin meneliti perbedaan kadar gliserol pada orang yang memiliki berat badan lebih dan dan berat badan normal. Maka untuk menentukan seseorang tersebut memiliki berat badan lebih harus mempertimbangkan rumus indeks massa tubuh.

Seperti halnya Ketika seseorang peneliti ingin mengetahui perbedaan kadar leukosit pada penderita diabetes yang mengalami genreïn. Pada penelitian jenis tersebut peneliti harus sikapmu dalam menyikapi Apakah seseorang tersebut benar-benar diabetes. pengungkapan seseorang yang terkena diabetes harus memperhatikan diagnosa dokter atau berdasarkan tes laboratorium yang valid.

Pada purposive sampling seringkali di lakukan identifikasi kriteria yang sering dibuat sebagai kriteria inklusi dan kriteria eksklusï. Kriteria inklusi syarat kan bahwa sampel harus memiliki Kriteria kriteria tersebut sehingga dapat digunakan menjadi sampel yang valid. Sebagai contoh penelitian mengenai diabetes mellitus yang digunakan adalah memiliki nilai pemeriksaan hba1c lebih dari 7 atas sudah pernah dilakukan diagnosa oleh dokter, bersedia menjadi pasien dan memberikan informasi mengenai penyakitnya, rumor antara 30 sampai 65. Rangkaian kriteria eksklusï merupakan kebalikan dari kriteria inklusi. Dengan bahasa sederhana serta kreteria inklusi merupakan suatu kriteria yang membuat suatu sampel tidak bisa digunakan dan harus mencari sampel yang lain.

#### **iv. Sampling jenuh**

sampling jenuh masih menuai perdebatan apa yang bener masih menuai Apakah tersebut merupakan jenis penelitian yang menggunakan sampel atau tidak karena pada sampling jenuh keseluruhan populasi yang ada digunakan sebagai sampel.

#### **v. Sampling Snowball**

Sampling Snowball lebih cocok digunakan ketika kita tidak mengetahui suatu keadaan mengenai penelitian yang dilakukan. Sebagai contoh seorang peneliti ingin memiliki kelainan aneh yang khas yang ada pada daerah tertentu. Pada dasarnya peneliti tidak tahu-menahu ingin mendaftarkan kelainan tersebut misalkan kebiasaan masyarakat pedalaman untuk menggunakan jenis obat tradisional yang terkenal mujarab. Dalam tujuannya peneliti ingin mengetahui kira-kira isi dari obat tersebut dan karakter obat didalam menyembuhkan suatu penyakit seperti apa. Maka peneliti dapat menanyakan pada populasi tersebut Siapa yang mengerti mengenai asal-usul maupun metode di dalam penggunaan obat tersebut. Sistem yang digunakan adalah semacam tanya jawab dan petunjuk dari orang-orang yang ahli dibidangnya.

#### **d. Ringkasan Materi**

Populasi merupakan keseluruhan objek dari penelitian sedangkan sampel merupakan bagian dari populasi yang digunakan untuk menginterpretasikan sebuah populasi tersebut. Tanpa sendiri merupakan bagian-bagian yang memiliki karakter khusus. Sedangkan besar sampel merupakan bagian dari unit sampel yang menyatakan jumlah dari unit sampel tersebut. Proses untuk melakukan identifikasi atau pemilihan sampel dari populasi disebut dengan sampling. Metode sampling dilakukan menggunakan dua cara yaitu probability sampling dan nonprobability sampling. Probability sampling dapat dilakukan dengan menggunakan beberapa metode diantaranya simple random sampling, systematic random sampling, stratified sampling, atau Cluster sampling. Non

probability sampling dapat dilakukan dengan menggunakan kuota sampling , insidental sampling, purposive sampling, sampling jenuh atau snowball sampling.

### e. Latihan soal

Jawablah pertanyaan berikut secara singkat dan tepat

1. Jelaskan perbedaan Antara populasi, sample dan besar sample!
2. Seorang peneliti ingin mengetahui perbedaan trombosit pada pasien DBD yang di treatment menggunakan rebusan pucuk daun singkong dan yang tidak di treatment. Jelaskan pandangan anda mengenai populasi pada penelitian tersebut, sampel pada penelitian tersebut, metode sampling, & besar sampel?
3. Seorang peneliti ingin mengetahui pengaruh pemberian jus kurma terdapat peningkatan leukosit pada tikus yang di injeksi dengan virus HIV. Jelaskan pandangan anda mengenai populasi pada penelitian tersebut, sampel pada penelitian tersebut, metode sampling, & besar sampel?

### f. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Dahlan M.S., 2009. Besar sampel dan cara pengambilan sampel: dalam penelitian kedokteran dan kesehatan
3. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
4. Lemeshow, Stanley et all, 1997. Besar Sampel penelitian Kesehatan, UGM Press,

5. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
6. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
7. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
8. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito

## Bab V. Hipotesis & Statistika Inferensial

---

### Tujuan

**Mahasiswa dapat memahami dan dapat menentukan jenis jenis hipotesis dan menjabarkan cara kerja statistika Inferensial**

---

### Materi

---

1. Tingkat kepercayaan
  2. Hipotesis
  3. Derajad Kebebasan
  4. Rumusan dan arah hipotesis
  5. Statistika inferensial
- 

**H**ipotesis berasal dari 2 kalimat yaitu Hipo yang berarti dugaan dan taksis yang berarti sementara. Apabila kita Terjemahkan menjadi kalimat maka hipotesis merupakan dugaan hasil sementara pada sebuah penelitian. Meskipun demikian hipotesis hendaknya berdasarkan fakta-fakta ilmiah sehingga kesimpulan secara. Munculnya hipotesis yang kuat untuk berasal dari kajian latar belakang sehingga harapan dari tujuan pelatihan dapat tercapai. Hipotesis dapat dibagi menjadi dua yaitu hipotesis nihil ( $h_0$ ) dan hipotesis alternatif ( $h_a/h_1$ ).  $H_0$  merupakan hipotesis yang bersifat tidak ada hubungan atau tidak ada perbedaan atau tidak ada korelasi atau tidak adanya kesesuaian. Semua itu tergantung dari jenis dan sifat penelitian itu sendiri.

Sedangkan hipotesis alternatif merupakan kebalikan dari hipotesis nihil yaitu terdapat perbedaan atau terdapat hubungan atau terdapat kesesuaian.

### **a. Tingkat Kepercayaan**

Taraf signifikansi atau tingkat kepercayaan ( $\alpha$ ) menunjukkan probabilitas kesalahan yang ditetapkan oleh peneliti yang diakibatkan oleh kesalahan dalam pengambilan sampel (sampling error). Biasanya dinyatakan juga dengan  $p$  ( $\rho$ ) atau Sig (signifikansi) atau  $p$  value . Tingkat kepercayaan yang digunakan penelitian biasanya tergantung dengan jenis penelitian sendiri sebagai contoh penelitian yang bersifat laboratorium dan membutuhkan ketelitian yang sangat tinggi seperti daya hambat bakteri maupun respon suatu obat terhadap sel kultur secara in Vitro taraf signifikan yang digunakan sebesar 1%. Sedangkan penelitian yang bersifat lebih ke metode kasar seperti penelitian yang membedakan antara pengaruh penggunaan antikoagulan untuk meneliti kadar glukosa pada kelainan diabetes melitus maka taraf kepercayaan yang digunakan lebih besar yaitu 5%. Penelitian yang bersifat lapangan seperti analisa gaya hidup dan pengaruh lingkungan sosial Maka dapat dimungkinkan taraf kepercayaan yang digunakan lebih dari 5% (ex. 10%). Untuk penelitian juga dapat menggunakan karakter ikan yang lebih besar lagi seperti untuk meneliti kebiasaan hidup makhluk buas maupun tingkah laku pada jenis atau ras hewan yang berbeda.

### **b. Hipotesis**

Di dalam menentukan kesimpulan dan hipotesis seringkali ditemukan kendala atau kesalahan penarikan data kesalahan tersebut dapat dikategorikan menjadi kesalahan atau kekeliruan Tipe 1 atau tipe a dan kesalahan atau kekeliruan tipe 2 atau tipe B.

Kesalahan penggunaan Tipe 1 atau tipe a Yaitu penggunaan hipotesis benar namun demikian kesimpulan yang diambil hipotesis ditolak sehingga tentu hasilnya pun tidak sesuai dengan kenyataan. Kesalahan penggunaan hipotesis tipe 2 atau sering disebut sebagai

tipe B merupakan penggunaan hipotesis yang salah Namun demikian hipotesis tersebut diterima. Untuk mempermudah pemahaman tersebut dapat dilihat melalui gambar berikut

KESIMPULAN	KEADAAN SEBENARNYA	
	HIPOTESIS BENAR	HIPOTESIS SALAH
MENERIMA HIPOTESIS	BENAR	KEKELIRUAN TIPE II ( B )
MENOLAK HIPOTESIS	KEKELIRUAN TIPE I ( A )	BENAR

### c. Derajat kebebasan / df (degree of freedom)

Adalah jumlah pengamatan dalam sampel (N) di kurangi banyaknya kendali (linier) atau pembatas (restriksi) yang diletakkan atas pengamatan tadi. Secara lebih sederhana degree of freedom menunjukkan suatu kemungkinan untuk pemilihan secara acak pada beberapa jenis individu misalkan ada individu a b dan c maka individu a berkemungkinan memilih b, c, d, dan e tidak diperkenankan untuk memilih dirinya sendiri sehingga rumus cara mudah dapat dirumuskan jumlah individu dikurangi 1.

### d. Langkah-langkah uji hipotesis

Di dalam Pengujian Hipotesis dibutuhkan langkah-langkah sehingga hasil yang didapatkan dapat dipertanggungjawabkan langkah tersebut adalah

#### i. Penentuan rumusan masalah

Penentuan rumusan masalah menjadi sangat penting karena merupakan ide pokok penelitian dapat

dilaksanakan. Penentuan rumusan masalah hendaknya terpaku pada latar belakang yang menjadi masalah itu ada.

## ii. Menyatakan hipotesis nol dan hipotesis alternatif

Seringkali peneliti membutuhkan kira-kira menjadi tolak ukur keberhasilan penelitian dikehendaki adalah  $H_0$  atau  $H_a$ . Bantu yang mengetahui kemungkinan atau hipotesis yang diharapkan yaitu peneliti itu sendiri.

## iii. Penentuan taraf signifikansi

Terang signifikansi ditentukan berdasarkan pertimbangan metodologi penelitian yang diambil.

## iv. Uji statistik yang sesuai

Di dalam menentukan uji statistik harus memperhatikan beberapa aspek yaitu jenis data, homogenitas data maupun tingkat krusial datanya. Data yang bersifat parametrik hendaknya tetap dijaga dalam keadaan parametrik sehingga nilai-nilai setiap data dihargai di dalam uji statistik.

## e. Rumusan dan arah hipotesis.

Kebutuhan hipotesis untuk menentukan hasil akhir penelitian menjadi sangat krusial. Di dalam rumusan nya harus memperhatikan dari rumusan masalah contoh hipotesis pada penelitian pengaruh pemberian ekstrak tapak dara terhadap pertumbuhan *Staphylococcus aureus*. Penelitian tersebut tentu bertujuan untuk memberikan gambaran pemberian ekstrak yang mampu menahan laju pertumbuhan *Staphylococcus aureus*. Sehingga ada 2 hipotesis yang bisa diajukan yaitu:

$H_0$ : Tidak terdapat pengaruh pemberian ekstrak terhadap pertumbuhan *Staphylococcus aureus*.

$H_a/H_1$ : Terdapat pengaruh pemberian ekstrak terhadap pertumbuhan *Staphylococcus aureus*.

Di dalam pernikahan tentu kita harus memilih salah satu diantara hipotesis tersebut. Undangan untuk memilih berdasarkan latar

belakang. Mungkin dalam hal ini peneliti mencoba untuk memberikan gambaran bahwa yang mampu menahan pertumbuhan *Staphylococcus aureus*. Sehingga dengan alasan tersebut pula peneliti berharap bahwa terdapat efek positif terhadap penghambatan bakteri tersebut. Dalam penulisan naskah tidak diperkenankan menulis H0 atau Ha Namun langsung berkenaan dengan pokok materi hipotesis yang digunakan pakai contoh

*Dalam pernikahan ini " terdapat pengaruh pemberian ekstrak tapak dara terhadap pertumbuhan Staphylococcus aureus ".*

Arah hipotesis hipotesis yang dapat dikelompokkan menjadi dua kelompok yaitu satu arah dan dua arah. Hipotesis satu arah dapat dikelompokkan lagi menjadi batas yang condong kekanan dan kekiri. Sedangkan pada sebuah arah biasanya membagi satu nilai kritis menjadi dua.

Contoh penggunaan hipotesis satu arah dapat di terapkan pada penelitian berikut:

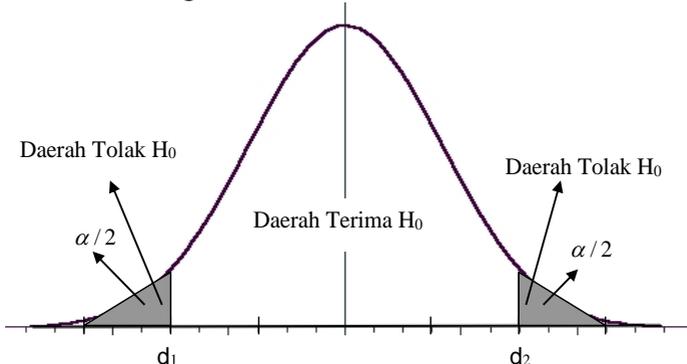
*Seorang ATLM ingin meneliti apakah benar bahwa kadar gula darah pada penderita diabetes adalah lebih dari 200 ng/mL dengan ketentuan uji HBA1C menunjukkan nilai lebih dari 7.*

Kenapa tulisan tersebut peneliti sentuh asumsi apabila nilai jauh dari 200 maka tidak akan dipermasalahkan sebagai contoh ternyata yang dapat nilainya mata 250 lalu bagaimana apabila hasil penelitian di bawah nilai maka 200. Maka statistik harus digunakan untuk menentukan apakah nilai yang diperoleh merupakan error atau menunjukkan nilai yang sebenarnya sehingga apabila rata-rata 190 yang signifikan menunjukkan nilai kurang dari 5% maka bisa

dianggap bahwa nilai HBA1C > 7 tidak identik dengan lebih dari 200 ng/mL.

Contoh lain yang dapat diterapkan dalam metode ini adalah ada asumsi bahwa rata-rata ATLM yang lulus pada uji kompetensi memiliki IPK dari 3,5. Maka untuk membuktikan asumsi tersebut diperlukan sebuah analisa hipotesis satu arah. Bunyi hipotesis nihil dari penelitian ini adalah " rata-rata ATLM yang lulus pada uji kompetensi memiliki IPK lebih dari 3,5". sedangkan hipotesis alternatif pada penelitian ini adalah rata-rata ATLM yang lulus pada objek uji kompetensi tidak memiliki IPK lebih dari 3,5. Hipotesis satu arah biasanya diperlakukan pada uji non parametrik dan Anova.

Hipotesis dua arah biasanya lebih sering digunakan dibandingkan hipotesis satu arah. Hipotesis dua arah seringkali juga digunakan untuk membandingkan dua variabel. Sebagai contoh seorang peneliti ingin membandingkan perbedaan kadar nikotin pada laki-laki perokok dan laki-laki tidak Pro. Ada hal ini cuplikan suatu perbedaan baik itu perbedaan ke arah bawah maupun perbedaan ke arah atas. Gambar dari rata-rata kurva tersebut dapat dilihat melalui gambar berikut:



Nilai Alfa atau error pada kurva jika dibagi menjadi dua bagian yaitu bagian kanan dan bagian kiri.

## f. Statistika Inferensial

**S**ebuah kendaraan bermotor. kita ambil contoh saja "motor matic". Motor keluaran terbaru sekarang membutuhkan bahan bakar yang beroktan relative tinggi. beberapa tidak bisa diberi ikan bakar biasa pakai premium. Demikian apabila kita masukkan bahan bakar tadi tentu dapat dilakukan misalnya motor matic bisa di isi premium bisa diisi pertalite, akan bisa diisi dengan bahan bakar yang saat ini bukan seperti minyak tanah, minyak goreng. Namun demikian ketika sedang menjalankan motor tersebut dengan menggunakan bahan bakar yang tidak sesuai. Motor yang membutuhkan oktan tinggi akan kemungkinan tersandar jalannya ketika diberi dengan kopi premium. Adakalanya motor besar menunjukkan gejala yang sangat terlihat yaitu tidak dapat berjalan dengan baik atau bahkan mogok seperti apabila motor matic tersebut diisi dengan Premium. Motor akan menunjukkan respon mogok apabila diisi air.

Ilustrasi motor di atas merupakan gambaran dari statistik. Pepes ikan terdiri atas yang mana kata tersebut merupakan maka dapat dimasukkan ke dalam setiap yang ada di statistika. namun hal yang berbeda adalah apabila motor menunjukkan respon, analisis statistika tetap dapat berjalan dan menunjukkan hasil yang lebih parahnya lagi adalah peneliti tidak akan tahu Apakah hasil tersebut sangat baik sesuai dengan norma-norma statistika atau bahkan Salah kaprah. Sehingga yang dapat mengetahui kesalahan seseorang menggunakan uji statistik adalah orang yang paham statistika juga. Inilah bahayanya ketika kita salah dalam melakukan uji statistik.

## g. Tahapan Uji Statistika

Setelah kita mengetahui jenis-jenis kata-kata seperti pada bab 1 yaitu saya baca tadi pagi menjadi nominal, ordinal, interval, dan rasio. Maka pengelompokan data tersebut dapat digunakan

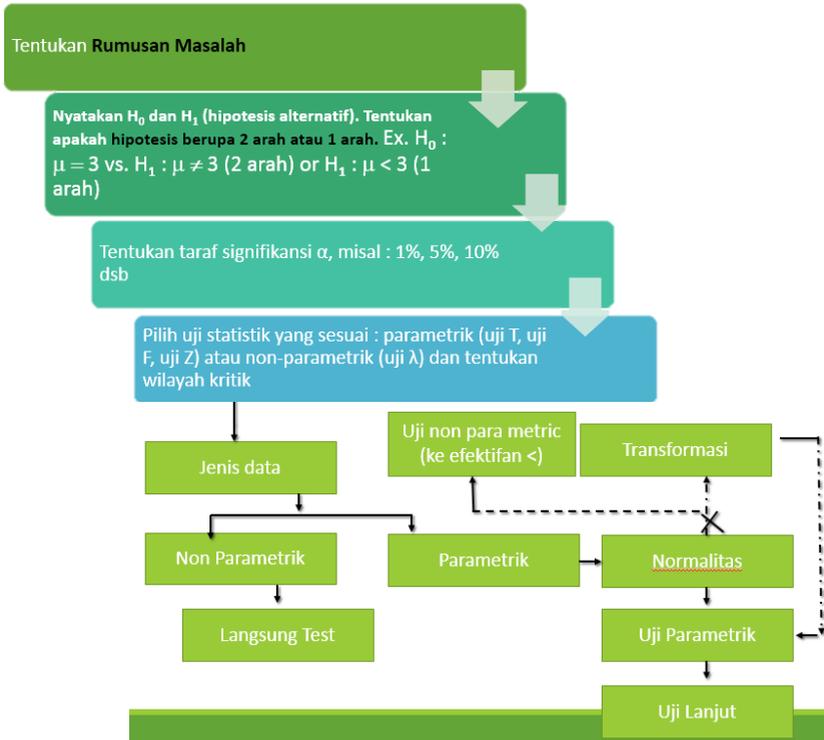
untuk menentukan apakah data yang didapat diolah menggunakan uji yang sesuai.

Untuk itu di dalam pengujian statistik terutama di statistika inferensial dibagi menjadi yang bersifat komparatif dan yang bersifat korelatif. Komparatif merupakan membandingkan dua lebih beberapa parameter untuk mengetahui perbedaan, di dalam uji tersebut harus diusahakan untuk diuji secara parametrik karena dengan uji parametrik nilai-nilai bisa terhilang kan begitu saja. Kita masih ingat bahwa tahapan hipotesis dimulai dari Menentukan rumusan masalah, menyatakan hak nol atau  $H_0$  yang mendukung, menentukan taraf signifikansi, kemudian memilih statistika yang sesuai.

Telah ditentukan ujian sesuai hal yang pertama untuk melakukan tahapan tersebut yaitu menentukan jenis data apakah data bersifat non parametrik atau data bersifat parametrik. Apabila data bersifat non parametrik maka kita bisa langsung melakukan uji statistik tanpa memperhatikan data berdistribusi normal atau tidak. Namun apabila data bersifat parametrik maka harus dilakukan uji normalitas untuk mengetahui data terdistribusi Normal atau tidak akan dijelaskan lebih lanjut pada subbab dibawah ini.

Ketika data sudah berdiskusi normal maka uji parametrik dapat di lakukan. Namun demikian apabila data tidak berdistribusi normal dan memperhatikan ketidak normalan tersebut Apakah jauh menyimpang atau hanya kecil saja dapat diusahakan melalui dua cara yang pertama apabila distribusi mendekati normal maka dapat dilakukan metode transformasi data. namun apabila data sudah tidak memungkinkan untuk dilakukan normalitas dan tidak dapat di transformasi maka dengan sangat terpaksa pengujian dialihkan ke uji non parametrik. untuk uji non Parametrik sangat tidak disarankan dengan tipe data yang bersifat interval atau rasio. Hal ini dikarenakan uji non parametrik menghilangkan informasi karakteristik data itu sendiri. Misal kita memiliki data kolesterol total 170,128, 200 mg/dl. Di dalam uttaran ada angka-angka tak terjaga 170,128, 200 namun apabila kita alihkan maka angka akan diurutkan berdasarkan nilai terendah ke nilai tertinggi 128, 170, 200

dan akan menjadi kategori 1, 2, dan 3. Apabila data sudah selesai dilakukan transformasi dan ternyata data tersebut Setelah diuji berubah menjadi normal maka dapat dilakukan uji lanjut uji parametrik. Pada beberapa kasus uji melibatkan lebih dari dua parameter maka harus dilakukan uji lanjut untuk mengetahui perbedaan pada tiap-tiap variabel.



### h. Pemilihan uji statistik univariat atau bivariat

Sesuai dengan yang diuraikan sub bab sebelumnya, statistik dapat dikelompokkan menjadi ini univariat dan bivariat. Selain itu dapat dikelompokkan juga menjadi comparative dan korelative.

Kata yang lebih detail comparative yaitu uji yang berusaha membedakan dua sampel atau lebih. Sedangkan uji korelasi berusaha untuk mencari hubungan antar beberapa parameter. Uji yang dilakukan untuk membedakan dua sampel berbeda dengan uji yang dilakukan untuk membedakan lebih dari 2 sampel.

Uji dua sampel atau lebih dari 2 sampel dapat dikategorikan menjadi sampel yang bebas dan sampel yang tidak bebas. Sampel bebas/ independen dengan karakteristik populasi berdistribusi normal dapat menggunakan uji T dependen, sedangkan apabila tidak terdistribusi normal maka dapat dilakukan uji Mann Whitney atau uji jumlah peringkat dari wilcoxon. Data yang bersifat nominal atau kategori maka dapat dilakukan uji Chi square atau Fisher. Uji dua sampel yang bersifat pasangan atau related dapat dilakukan menggunakan uji t sampel berpasangan (data terdistribusi normal) atau uji wilcoxon (data tidak berdistribusi normal) atau data yang bersifat nomina kata-kata zakat menggunakan uji mcnemar.

Sampel yang berjumlah lebih dari 2 dengan kategori bebas maka dapat dicuci menggunakan Anova satu arah atau Anova dua arah (data terdistribusi normal) sedangkan data tidak berdistribusi normal dapat dilakukan uji kruskal Wallis. Data yang bersifat non parametrik dapat diuji menggunakan Chi square. Sedangkan sampel yang lebih dari 2 yang bersifat berpasangan dapat diuji dengan Anova untuk subjek yang sama (data terdistribusi normal) atau uji Friedman (data tidak terdistribusi normal) dan untuk data yang bersifat katagori/nominal dapat dilakukan uji Cochran's Q.

Uji korelasi dapat dilakukan menggunakan korelasi pearson untuk data yang normal dan para lagi Superman asosiasi untuk data yang tidak terdistribusi secara normal. Tangan nominal Apakah dapat dilakukan menggunakan uji koefisien kontingensi atau uji koefisien phi. Untuk dapat memahami lebih lanjut dapat dilihat pada tabel berikut

PEMILIHAN UJI STATISTIK UNIVARIAT / BIVARIAT					
Tujuan Uji	Jumlah sampel / pasangan	Macam sampel (bebas / berpasangan)	Jenis variabel		
			Rasio-Interval pop. berdistribusi normal	Ordinal / Rasio-Interval distrib. tak normal	Nominal / kategorik
Komparasi (perbedaan)	2	Bebas ( <i>independent</i> )	Uji t 2 sampel bebas	~ Uji Mann-Whitney ~ Uji jumlah peringkat dari Wilcoxon	~ Uji khi-kuadrat ~ Uji eksak dari Fisher
		Berpasangan ( <i>related/paired</i> )	Uji t sampel berpasangan	Uji peringkat bertanda dari Wilcoxon	Uji McNemar ( <i>u/ kategori dikotomik</i> )
	> 2	Bebas ( <i>independent</i> )	Anava 1 arah	Uji Kruskal-Wallis	Uji khi-kuadrat
		Berpasangan ( <i>related/paired</i> )	Anava u/ subyek yg sama	Uji Friedman	Uji Cochran's Q ( <i>u/ kategori dikotomik</i> )
Korelasi			~ Korelasi dari Pearson (r) ~ (Regresi)	~ Korelasi dari Spearman (r <sub>s</sub> ) ~ Asosiasi Kappa (κ)	~ Koefisien Kontingensi (C) ~ Koefisien Phi

## i. Ringkasan Materi

Tingkat kepercayaan merupakan penjamin penelitian kita dapat dipercaya oleh peneliti lain. Didalam penentuan tingkat kepercayaan memang tidak ada patokan yang mutlak tetapi ada tingkat kewajaran yang menjadi standart penelitian. Hipotesis merupakan dugaan sementara berdasarkan dasar teori ilmiah yang kuat. Hipotesis di bagi menjadi Hipotesis nol dan hipotesis alfa sedangkan dalam penerapan kesimpulan kesalahan dapat dikategorikan kesalahan alfa dan kesalahan beta.

Statistika inferensial merupakan metode pengolahan data menggunakan analisis spesifik. Pengolahan data secara parametric membutuhkan homogenitas data. Sedangkan pengolahan data non parametric tidak memerlukan syarat homogenitas data

## j. Latihan soal

Jawablah pertanyaan berikut secara singkat dan tepat

1. Seorang peneliti ingin mengetahui perbedaan trombosit pada pasien DBD yang di treatment menggunakan rebusan pucuk daun singkong dan yang tidak di treatment. Diskusikan
  1. Tingkat kepercayaan yang di perlukan
  2. Bunyi dari hipotesis nol dan hipotesis alfa
  3. Uji yang butuhkan dan tahapan dalam uji ketika pasien DBD yang di treatment sebanyak 35 orang dan pasien DBD yang tidak di treatment sebanyak 39 orang
2. Seorang peneliti ingin mengetahui pengaruh pemberian jus kurma terdapat peningkatan leukosit pada tikus yang di injeksi dengan virus HIV.
  1. Tingkat kepercayaan yang di perlukan
  2. Bunyi dari hipotesis nol dan hipotesis alfa
  3. Uji yang butuhkan dan tahapan dalam uji apabila kosentrasi yang digunakan adalah 10 100%, 80%, 60%, 40% dan control.

## k. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
2. Hadi, S, 1997. *Statistik jilid I : cetakan kedua puluh*. Yogyakarta: Penerbit Andi
3. Hasan, Iqbal. 2002. *Pokok-pokok Materi Statistik 2 (Statistik Inferensif)*. Jakarta: Bumi Aksara.
4. Ott R. L. & Longnecker M. 2010. *An Introduction to statistical Methods and Data Analysis: Sixth Edition*. Cengage Learning
5. Purnawan, J., 1997 *Teknik Analisis Data*, Rineka Cipta,

6. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
7. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
8. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## **BAB VI. Uji perbedaan mean 2 sampel dengan statistik parametrik (Uji T)**

---

### **Tujuan**

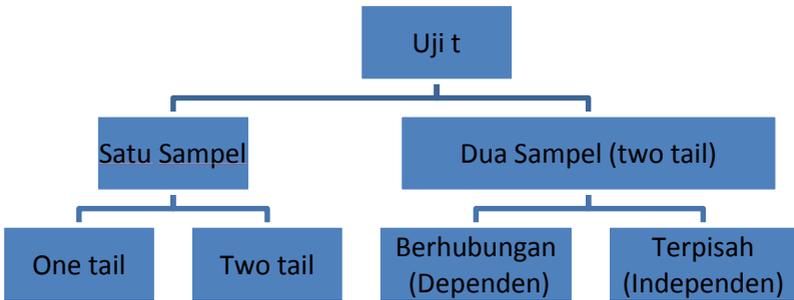
**Mahasiswa memahami dan dapat melakukan uji Perbedaan mean 2 sample dengan statistic parametric dengan tepat**

---

### **Materi**

- 1. Uji T Satu Sample One**
  - 2. Uji T Satu Sample two tail**
- 

**P**ada tahun 1908, W S Gosset dengan nama samaran student mempublikasin karyanya yang berupa distribusi student (sebaran student) atau distribusi t Pengujian rata-rata satu sampel dimaksudkan untuk menguji nilai tengah atau rata-rata populasi  $\mu$  sama dengan nilai tertentu  $\mu_0$ , lawan hipotesis alternatifnya bahwa nilai tengah atau rata-rata populasi  $\mu$  tidak sama dengan  $\mu_0$ . Uji T di bedakan menjadi beberapa Kelompok seperti pada diagram berikut:



Gambar Pembagian Uji T

Bentuk kurva normal pada uji T mirip dengan kurva normal pada umumnya. Sehingga didalam memutuskan suatu hipotesis maka perlu dibandingkan antara T hitung dengan T table. Df atau *Degree of freedom* ditentukan dengan rumus =  $n-1$ . Dan yang perlu diperhatikan apakah uji tersebut bersifat one tail atau two tail. Apabila one tail maka nilai *alfa* akan menjadi pemiliki sepenuhnya pada satu sisi saja, akan tetapi apabila bersifat two tail maka nilai *alfa* dibagi menjadi dua bagian yang sama yaitu kanan dan kiri. Untuk menarik simpulan maka kita perlu memperhatikan nilai T hitung. Apabila T hitung  $>$  T table maka dapat diputuskan  $H_a$  diterima dan apabila T hitung  $<$  T table maka dapat disimpulkan  $H_0$  diterima.

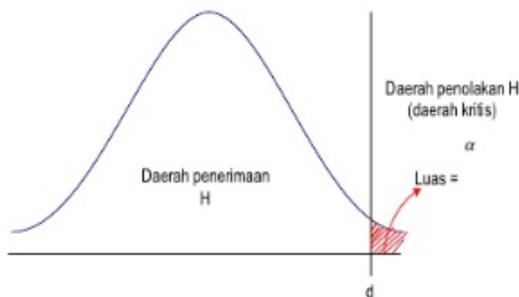


Gambar Pembagian nilai *alfa* 5% berdasarkan arah daerah penolakan

### a. Uji T Satu Sample (One/ two tail)

Pengujian satu sampel pada prinsipnya ingin menguji apakah suatu nilai tertentu (yang diberikan sebagai pembanding) berbeda secara nyata ataukah tidak dengan rata-rata sebuah sampel. Nilai tertentu di sini pada umumnya adalah sebuah nilai parameter untuk mengukur suatu populasi atau sebuah teori. Sebagai contoh seorang analis kesehatan ingin menganalisa waktu pelayanan laboratorium pada rumah sakit X apakah sesuai dengan ketentuan pemerintah atau tidak (ketentuan pemerintah < 140 menit). Maka dalam hal ini kita membutuhkan analisa kurva normal eror sebelah kanan saja (one tail) (Gambar 7.3). Kenapa demikian karena bunyi dari analisa tersebut adalah kurang dari 140. Maka dapat kita Tarik kesimpulan ternyata apabila ada seorang pasien yang berobat ternyata pemeriksaannya hanya 5 menit tentu tidak ada yang memperlmasalahkan bukan, akan tetapi coba bayangkan apabila pemeriksaan berlangsung 145 menit tentu si pasien tersebut akan mengklaim bahwa rumah sakit tesebut tidak professional atau tidak bermutu.

Berbeda apabila ada seorang ahli gizi di minta untuk menganalisa kadar air buah kemasan, karena perusahaan hanya ingin kadar air dalam buah adalah 10 %. Maka pada kasus ini ahli gizi tersebut harus benar benar memastikan kadar air tidak berbeda jauh lebih rendah atau lebih tinggi dari angka 10 %.



Gambar 7.3. Daerah Penolakan satu arah (kanan) pada Uji T one sample (one tail)

Untuk menyelesaikan uji hioptesis menggunakan uji T one sample digunakan rumus:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- t = nilai dari t hitung
- $\bar{x}$  = rata rata dari data pengamatan
- $\mu$  = acuan dari rerata populasi atau teori
- S = standart deviasi
- n = jumlah sampel

**Contoh 1 :**

Seorang analis kesehatan ingin membuktikan bahwasannya kadar glukosa penderita diabetes yang melakukan senam Program Pengelolaan Penyakit Kronis (Prolanis) adalah tidak lebih dari 200 mg/dl. Data dari 20 peserta adalah sebagai berikut:

171	172	190	159	207
206	154	159	198	166
179	176	156	189	190
168	180	156	156	173

Tentukan apakah data diatas memenuhi asumsi yang di inginkan oleh peneliti atau tidak!

**Penyelesaian Contoh 1 menggunakan cara manual:**

Dari data tesebut dapat di ketahui

1. *Rumusan masalah*

Apakah kadar glukosa penderita diabetes yang melakukan senam Program Pengelolaan Penyakit Kronis (Prolanis) adalah tidak lebih dari 200 mg/dl?. ( $\alpha=5\%$ )

2. *Hipotesis*

- a.  $H_0$  = kadar glukosa penderita diabetes yang melakukan senam Program Pengelolaan Penyakit Kronis (Prolanis) adalah tidak lebih dari 200 mg/dl
- b.  $H_a$  = kadar glukosa penderita diabetes yang melakukan senam Program Pengelolaan Penyakit Kronis (Prolanis) adalah lebih dari 200 mg/dl

3. *Uji Hipotesis dan penentuan confident level*

Uji T One Sample (one tail), confident level yang digunakan 95%

4. *Perhitungan*

Perhitungan dilakukan dengan menggunakan rumus:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$\bar{x} = 175.5 \text{ (rata rata)}$$

$$\mu = 200$$

$$S = 16.85$$

$$t = \frac{175.5 - 200}{16.85/\sqrt{30}}$$

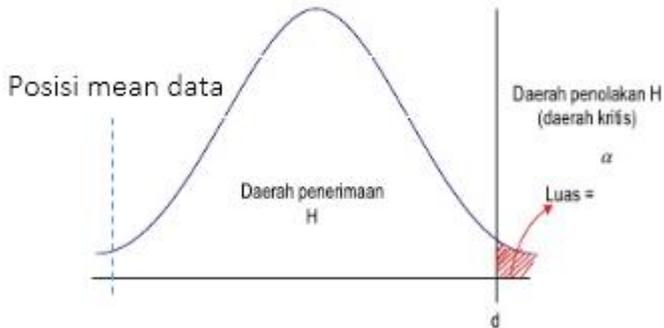
$$T \text{ hitung} = -6.568$$

Simpanan T hitung lebih besar dibandingkan dengan t table maka secara teori perbandingan t maka dapat disimpulkan  $H_a$  diterima,

**t Table**

cum. prob	$t_{.50}$	$t_{.25}$	$t_{.20}$	$t_{.15}$	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460

akan tetapi apabila kita melihat nilai kurva normal yang menunjukkan rata rata posisi tidak pada daerah penolakan  $H_a$  Maka  $H_0$  diterima



**Penyelesaian Contoh 1 menggunakan cara SPSS:**

**1. Langkah Pertama;**

Masukkan data pada sheet data. Data disusun secara vertical. Nomor yang ada menunjukkan nomor data

\*Untitled1 [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-on

1 : VAR00001 171

	VAR00001	var	var	var
1	171.00			
2	206.00			
3	179.00			
4	168.00			
5	172.00			
6	154.00			
7	176.00			
8	180.00			
9	190.00			
10	159.00			
11	156.00			
12	156.00			
13	159.00			
14	198.00			
15	189.00			
16	156.00			
17	207.00			
18	166.00			
19	190.00			
20	173.00			

## 2. Langkah kedua

Pastikan bentuk data terpilih dengan pilihan skala dan type datan dalam bentuk numeric

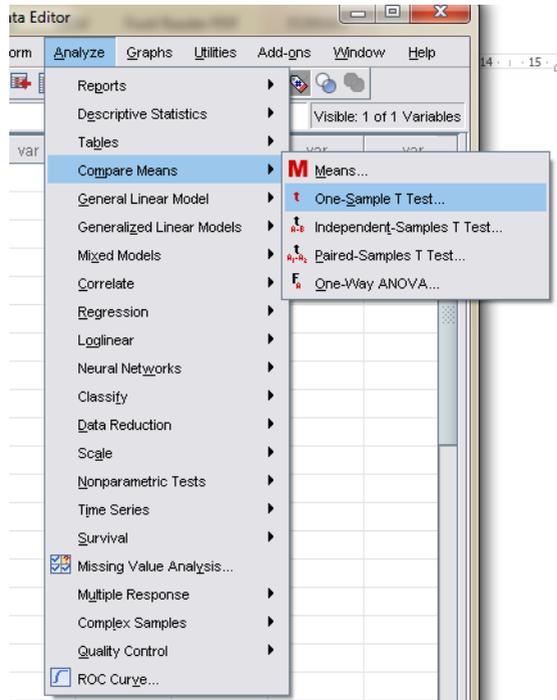
\*Untitled2 [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	VAR00001	Numeric	2	0		None	None	2	Left	Scale
2										

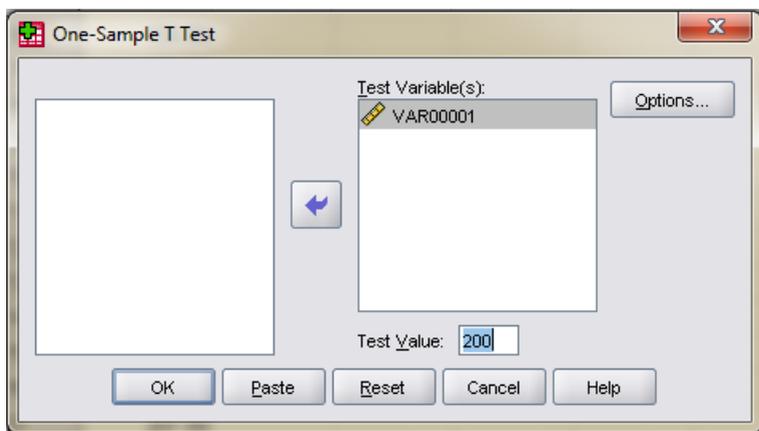
### 3. Langkah ketiga

Pilih menu pada bagian atas analyze> Compare Mean> One-Sample T Test.



### 4. Langkah keempat

Masukkan variable yang digunakan kepada test variable dan masukkan test value. Test value merupakan nilai patokan yang digunakan



5. Langkah ke empat klik ok

### ▶ T-Test

[DataSet0]

**One-Sample Statistics**

	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	20	1.7525E2	16.85191	3.76820

**One-Sample Test**

	Test Value = 200					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of Difference	
					Lower	Upper
VAR00001	-6.568	19	.000	-24.75000	-32.6369	-16.8631

6. Cara membaca

Lihat nilai t, nilai t akan sama dengan nilai t hitung, apabila kita mengambil keputusan dari sini maka kita harus membandingkan dengan nilai t table,

kemudian lihat nilai sig (2-tailed) maka kita harus membagi dua, ingat kasus yang dihadapi adalah one sample t test (one tail) (SPSS tidak menyediakan menu one tail untuk uji T test). Sehingga  $0.000/2 = 0.000$ . Karena  $0.000 < 0.05$  maka  $H_a$  diterima, akan tetapi apabila kita melihat nilai kurva normal yang menunjukkan rata-rata posisi tidak pada daerah penolakan  $H_a$  Maka  $H_0$  diterima.

### **b. Uji T Dua Sample**

Uji-t 2 sampel adalah metode yang digunakan untuk menguji kesamaan rata-rata dari 2 populasi, dimana peneliti tidak memiliki informasi mengenai ragam populasi. Yang membedakan antara independent dengan dependent adalah sumber sample yang digunakan, ingat sampel bukan berarti jumlah ulangan pada tiap perlakuan atau berapa banyak orang yang digunakan, sampel didalam statistika merupakan keragaman yang menjadi perbandingan. Sebagai contoh seorang peneliti meneliti perbedaan trombosit menggunakan cara manual dan menggunakan hematology analyzer, peneliti tersebut memanfaatkan pasien yang datang ke puskesmas sebanyak 34 orang. Jumlah sampel tersebut adalah 2, yaitu darah yang di analisis secara manual dan darah yang dianalisis menggunakan hematology analyzer, lalu 34 disini apa?. 34 adalah besar sample. Lalu uji apa yang tepat untuk digunakan? Tentu disini adalah uji T dependent. Alasan utama adalah sumber sampel yang digunakan adalah dari subjek yang sama.

Penelitian diatas bisa saja dianalisis menggunakan independent t test apabila unit sampel yang digunakan berasal dari sumber yang berbeda, contoh si peneliti meneliti 30 orang untuk dianalisis secara manual dan 30 orang dianalisis menggunakan hematology analyzer.

Pada uji 2 sampel, maka anggapan  $H_0$  pada penelitian ini adalah  $\mu_1 = \mu_2$  secara kasar dapat kita artikan antara rerata sampel 1 dengan rerata sampel 2. Sehingga bunyi dari Hipotesis nol adalah "tidak ada perbedaan antara ..... (sampel 1) dan ..... (sampel 2). Melihan dari uraian tersebut maka  $H_a/H_1$  memiliki konsekuensi yang berbeda yaitu  $\mu_1 \neq \mu_2$  dapat kita artikan antara rerata sampel 1 tidak sama dengan rerata sampel 2. Sehingga bunyi dari Hipotesis nol adalah " ada perbedaan antara ..... (sampel 1) dan ..... (sampel 2).

Adapun rumus untuk mencari nilai t adalah

**1. Dependent t test**

$$t = \frac{\sum D}{\frac{s}{\sqrt{n}}}$$

$$D = (x_1 - x_2)$$

$$s = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

## 2. Independent T test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2} \left( \frac{n_1 + n_2}{n_1 \cdot n_2} \right)}}$$

### Contoh 1

Seorang analis kesehatan ingin membandingkan apakah terdapat perbedaan kadar glukosa darah yang disimpan menggunakan Natrium sitrat dengan darah yang disimpan menggunakan EDTA. Data di ambil dari 45 orang: (CI =95%)

#### Hasil darah yang disimpan menggunakan Natrium Sitrat

117	177	188	182	129
176	185	129	122	133
149	192	115	171	183
154	113	168	111	134
126	152	112	155	188
115	152	147	145	152
157	196	132	178	110
164	164	117	144	142

#### Hasil darah yang disimpan menggunakan EDTA

168	171	116	175	192
112	182	184	114	152
115	174	183	124	136
132	148	174	117	155

125	189	139	184	200
189	144	120	132	141
171	139	118	161	180
186	157	175	159	110

$$t = \frac{\frac{\sum D}{n}}{\frac{s}{\sqrt{n}}} \quad D = (x_1 - x_2) \quad s = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

Natrium sitrat	EDTA	d	d2
117	168	-51	2601
176	112	64	4096
149	115	34	1156
154	132	22	484
126	125	1	1
115	189	-74	5476
157	171	-14	196
164	186	-22	484
177	171	6	36
185	182	3	9
192	174	18	324
113	148	-35	1225
152	189	-37	1369
152	144	8	64
196	139	57	3249
164	157	7	49
188	116	72	5184
129	184	-55	3025

115	183	-68	4624
168	174	-6	36
112	139	-27	729
147	120	27	729
132	118	14	196
117	175	-58	3364
182	175	7	49
122	114	8	64
171	124	47	2209
111	117	-6	36
155	184	-29	841
145	132	13	169
178	161	17	289
144	159	-15	225
129	192	-63	3969
133	152	-19	361
183	136	47	2209
134	155	-21	441
188	200	-12	144
152	141	11	121
110	180	-70	4900
142	110	32	1024
jumlah		-167	55757
(jumlah d) <sup>2</sup>		27889	

***Rumusan masalah***

Apakah terdapat perbedaan kadar glukosa darah yang disimpan menggunakan Natrium sitrat dengan darah yang disimpan menggunakan EDTA?

### **Tujuan**

Untuk mengetahui perbedaan kadar glukosa darah yang disimpan menggunakan Natrium sitrat dengan darah yang disimpan menggunakan EDTA.

### **Tingkat kepercayaan**

Confident level yang digunakan 95%

### **Rumusan masalah**

Ho = tidak terdapat perbedaan kadar glukosa darah yang disimpan menggunakan Natrium sitrat dengan darah yang disimpan menggunakan EDTA.

Ha = terdapat perbedaan kadar glukosa darah yang disimpan menggunakan Natrium sitrat dengan darah yang disimpan menggunakan EDTA.

### **Uji yang digunakan**

Dependent T test

$$t = \frac{\frac{\sum D}{n}}{\frac{s}{\sqrt{n}}} \quad D = (x_1 - x_2) \quad s = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

$$D = (x_1 - x_2) = -167$$

$$s = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}} = s = \sqrt{\frac{1}{45-1} \left\{ 55757 - \frac{278889}{45} \right\}} = 35.39$$

$$T = -167/45/35.39/6.7$$

T= - 0.703

Dengan memperhatikan nilai T hitung < dari t table (2. 014) maka dapat disimpulkan Ho diterima, df= n-1

df \ Pr	0.25	0.10	0.05	0.025	0.01	0.005	0.001
	0.50	0.20	0.10	0.050	0.02	0.010	0.002
41	0.68052	1.30254	1.68288	2.01954	2.42080	2.70118	3.30127
42	0.68038	1.30204	1.68195	2.01808	2.41847	2.69807	3.29595
43	0.68024	1.30155	1.68107	2.01669	2.41625	2.69510	3.29089
44	0.68011	1.30109	1.68023	2.01537	2.41413	2.69228	3.28607
45	0.67998	1.30065	1.67943	2.01410	2.41212	2.68959	3.28148
46	0.67986	1.30023	1.67866	2.01290	2.41019	2.68701	3.27710
47	0.67975	1.29982	1.67793	2.01174	2.40835	2.68456	3.27291
48	0.67964	1.29944	1.67722	2.01063	2.40658	2.68220	3.26891
49	0.67953	1.29907	1.67655	2.00958	2.40489	2.67995	3.26508

### Cara menggunakan SPSS

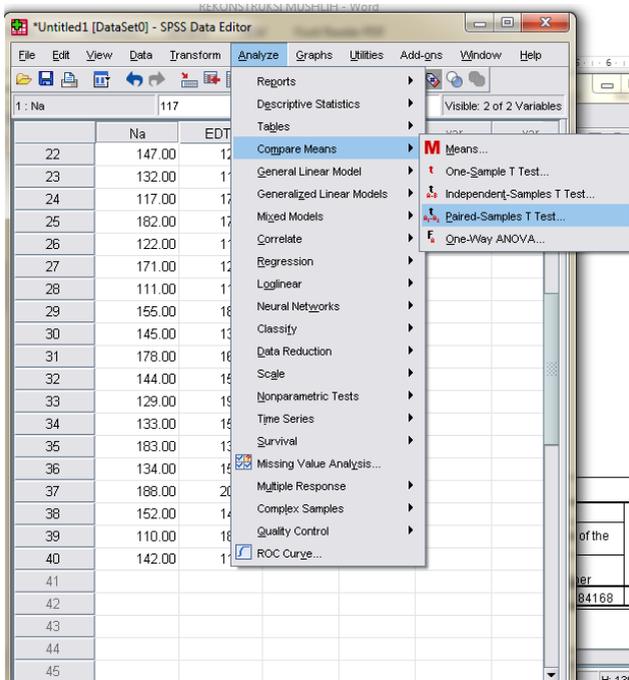
#### 1. Tahap Petama

Masukkan data data tersebut kedalam excel. Pengaturan penataletaan data dilakukan bersebelahan. Hal tersebut diasumsikan data diolah dari sampel yang sama

	Na	EDTA
1	117.00	168.00
2	176.00	112.00
3	149.00	115.00
4	154.00	132.00
5	126.00	125.00
6	115.00	189.00
7	157.00	171.00
8	164.00	106.00
9	177.00	171.00
10	185.00	182.00
11	192.00	174.00
12	113.00	148.00
13	152.00	189.00
36	134.00	155.00
37	108.00	200.00
38	152.00	141.00
39	110.00	180.00
40	142.00	110.00

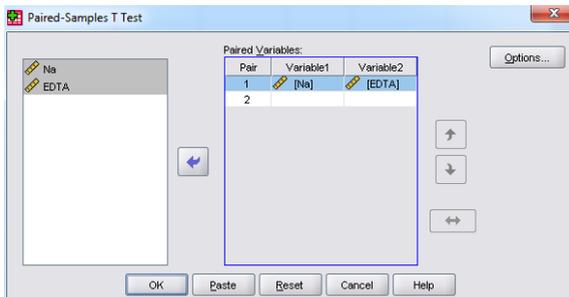
## 2. Tahap Kedua

Pilih Analyze > Compare mean > Paired-Sample T test



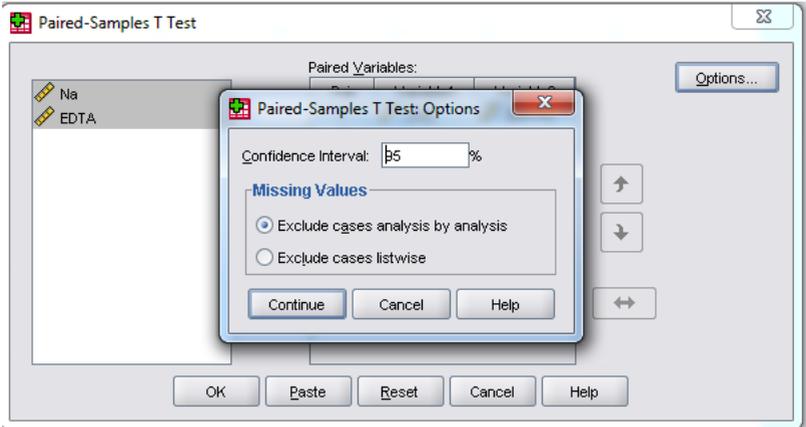
### 3. Tahap Ketiga

Maka akan muncul seperti table berikut kemudian masukkan pada masing masing variable



4. Tahap Ke empat

Pilih bagian option untuk memilih nilai confident level (confident interval)



5. Tahap ke lima. Klik continu dan kemudian klik ok

**Paired Samples Correlations**

		N	Correlation	Sig.
Pair 1	Na & EDTA	40	.032	.843

**Paired Samples Test**

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% CI				
				Lower	Upper			
Pair 1 Na - EDTA	-4.17500	37.573	5.94094	-16.19168	7.84168	-.703	39	.486

### Cara membaca

Lihat nilai t, nilai t akan sama dengan nilai t hitung, apabila kita mengambil keputusan dari sini maka kita harus membandingkan dengan nilai t table, kemudian lihat nilai sig (2-tailed)  $0.486 > 0.05$   $H_0$  diterima.

### Contoh 2.

Seorang ahli gizi ingin membandingkan jumlah vitamin c (tiap 100 gr) pada buah mangga manalagi dan mangga arumanis, menggunakan masing masing 25 buah mangga:

mangga manalagi		mangga arumanis	
51	81	124	131
100	64	74	74
119	86	81	71
60	132	100	97
59	76	74	80
136	108	81	142
56	67	75	165
104	90	170	95
149	150	134	153
117	65	135	82
65	70	116	111
144	146	118	80
70		128	

## Cara penyelesaian

### 1. Rumusan Masalah

Apakah ada perbedaan kandungan vitamin C (tiap 100 gr) pada buah mangga manalagi dan mangga arumanis?

### 2. Hipotesis

**Ho**= Tidak ada perbedaan kandungan vitamin c (tiap 100 gr) pada buah mangga manalagi dan mangga arumanis

**Ha**= Ada perbedaan kandungan vitamin c (tiap 100 gr) pada buah mangga manalagi dan mangga arumanis

### 3. Uji Hipotesis

Uji Independent T Test

### 4. Simpulan

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2} \left( \frac{n_1 + n_2}{n_1 \cdot n_2} \right)}}$$

mangga manalagi	mangga arumanis	$(x_1 - \bar{x}_1)$	$(x_2 - \bar{x}_2)$	$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$
51	124	-95	44	9025	1936
100	74	-46	-6	2116	36
119	81	-27	1	729	1

60	100	-86	20	7396	400
59	74	-87	-6	7569	36
136	81	-10	1	100	1
56	75	-90	-5	8100	25
104	170	-42	90	1764	8100
149	134	3	54	9	2916
117	135	-29	55	841	3025
65	116	-81	36	6561	1296
144	118	-2	38	4	1444
70	128	-76	48	5776	2304
81	131	-65	51	4225	2601
64	74	-82	-6	6724	36
86	71	-60	-9	3600	81
132	97	-14	17	196	289
76	80	-70	0	4900	0
108	142	-38	62	1444	3844
67	165	-79	85	6241	7225
90	95	-56	15	3136	225
150	153	4	73	16	5329
65	82	-81	2	6561	4
70	111	-76	31	5776	961
146	80	0	0	0	0
94.6	107.6 4	jumlah		92809	42115

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}} \left( \frac{n_1 + n_2}{n_1 \cdot n_2} \right)}$$

$$t = \frac{94.6 - 107.64}{\sqrt{\frac{92809 + 42115}{25 + 25 - 2} \left( \frac{25 + 25}{25 \cdot 25} \right)}}$$

$$T = 1.43$$

Dengan memperhatikan nilai T hitung < dari t table (2. 01) maka dapat disimpulkan Ho diterima, rumus df yang digunakan adalah  $n_1+n_2-2$

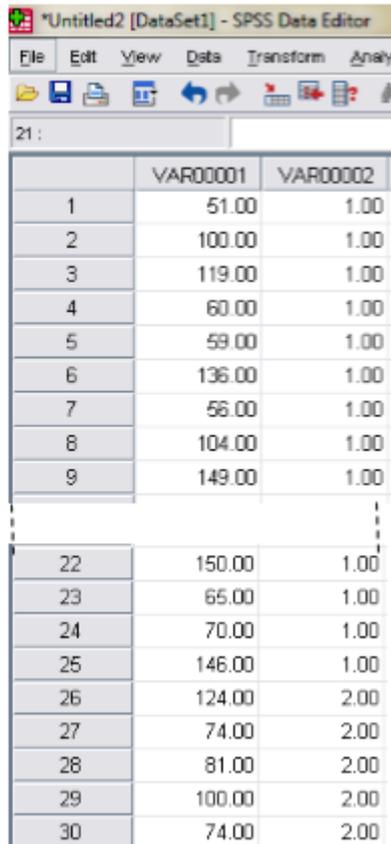
#### Titik Persentase Distribusi t (df = 41 – 80)

Pr df	0.25	0.10	0.05	0.025	0.01	0.005	0.001
	0.50	0.20	0.10	0.050	0.02	0.010	0.002
41	0.68052	1.30254	1.68288	2.01954	2.42080	2.70118	3.30127
42	0.68038	1.30204	1.68195	2.01808	2.41847	2.69807	3.29595
43	0.68024	1.30155	1.68107	2.01669	2.41625	2.69510	3.29089
44	0.68011	1.30109	1.68023	2.01537	2.41413	2.69228	3.28607
45	0.67998	1.30065	1.67943	2.01410	2.41212	2.68959	3.28148
46	0.67986	1.30023	1.67866	2.01290	2.41019	2.68701	3.27710
47	0.67975	1.29982	1.67793	2.01174	2.40835	2.68456	3.27291
48	0.67964	1.29944	1.67722	2.01063	2.40658	2.68220	3.26891
49	0.67953	1.29907	1.67655	2.00958	2.40489	2.67995	3.26508
50	0.67943	1.29871	1.67591	2.00856	2.40327	2.67779	3.26141
51	0.67933	1.29837	1.67528	2.00758	2.40172	2.67572	3.25789
52	0.67924	1.29805	1.67469	2.00665	2.40022	2.67373	3.25451
53	0.67915	1.29773	1.67412	2.00575	2.39879	2.67182	3.25127

### Penyelesaian masalah menggunakan SPSS

1. **Tahap 1.** Masukkan data berdasarkan valuenya. Cara memasukkan didalam kolom berbeda dengan cara memasukkan dependent T test. Posisi di dilakukan

secara berkelanjutan yang kemudian di beri tanda atau label pada variable berikutnya

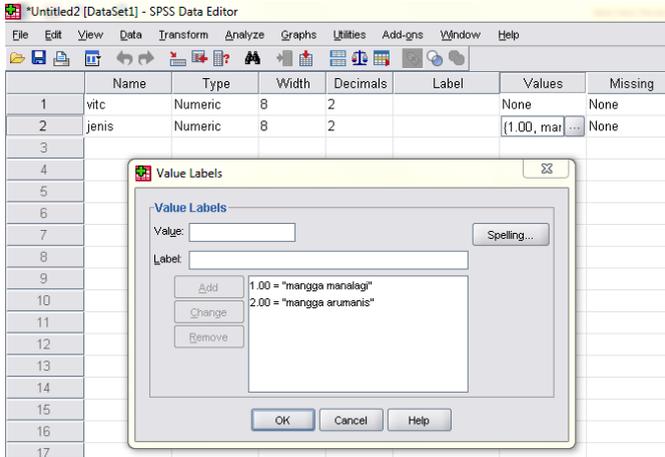


The screenshot shows the SPSS Data Editor window titled '\*Untitled2 [DataSet1] - SPSS Data Editor'. The menu bar includes File, Edit, View, Data, Transform, and Analyze. Below the menu bar is a toolbar with various icons. The main window displays a data table with the following structure:

	VAR00001	VAR00002
1	51.00	1.00
2	100.00	1.00
3	119.00	1.00
4	60.00	1.00
5	59.00	1.00
6	136.00	1.00
7	56.00	1.00
8	104.00	1.00
9	149.00	1.00
...	...	...
22	150.00	1.00
23	65.00	1.00
24	70.00	1.00
25	146.00	1.00
26	124.00	2.00
27	74.00	2.00
28	81.00	2.00
29	100.00	2.00
30	74.00	2.00

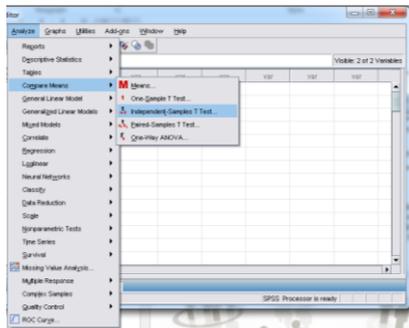
## 2. Tahap Ke dua.

Untuk memberi keterangan pada var label bias dilakukan pengeditan pada value. Masukkan value 1 dan terakan label manga manalagi pada kolom label dan tekan add. Begitu seterusnya.



### 3. Tahap Ke tiga

Tahapan selanjutnya pilih analyze> compare mean> independent sample T Test



### 4. Tahap Ke Empat

Klik ok maka akan muncul table berikut

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error	95% Confidence	
									Lower	Upper
vitc	Equal variances assumed	.308	.581	-1.432	48	.159	-13.04000	9.10822	-31.35330	5.27330
	Equal variances not assumed			-1.432	47.730	.159	-13.04000	9.10822	-31.35597	5.27597

#### Cara membaca

Lihat nilai t, nilai t akan sama dengan nilai t hitung, apabila kita mengambil keputusan dari sini maka kita harus membandingkan dengan nilai t table, kemudian lihat nilai sig (2-tailed)  $0.159 > 0.05$  Ho diterima.

### c. Ringkasan materi

Uji T atau uji student merupakan ujian digunakan untuk membedakan satu parameter atau 2 parameter. Pada uji satu parameter dapat bersifat satu arah atau dua arah tergantung dari hipotesa atau permasalahannya diselesaikan. Selain itu uji satu parameter membandingkan satu parameter dengan ketentuan atau teori yang sudah ada. Uji T pada dua parameter dapat bersifat independen atau independen. Independent t test merupakan uji yang membedakan dua parameter yang mutlak berbeda satu sama lain. Thanks and the painted menguji satu sampel yang diuji dua kali dalam artian satu parameter akan

dilakukan pengukuran dua kali yang hasilnya selanjutnya akan dibandingkan.

#### d. Latihan soal

Jawablah soal berikut dengan tepat!

3. Seorang peneliti ingin mengetahui Pengaruh pemberian jus semangka terhadap peningkatan tekanan darah ibu-ibu lansia, jus semangka yang digunakan adalah jus murni tanpa tambahan apapun, hasil dari analisa kasus adalah sebagai berikut:

sebelum di beri jus	130	144	145	120	150	130	120	130
	144	134	143	125	134	123	143	146
sesudah di beri jus	125	130	140	115	150	129	119	130
	120	130	140	121	131	120	143	144

dari kasus tersebut tentukan!

- rumusan masalah
  - hipotesis
  - uji yang cocok
  - kesimpulan dari permasalahan tersebut ( sertakan tahapan hitungannya)
4. seorang peneliti ingin membuktikan bahwasannya kebutuhan seseorang apakah sesuai dengan teori yang menyatakan kebutuhan seseorang adalah 2000 kkal/ hari, buktikan secara statistik apakah teori tersebut benar atau tidak

no sampel	jumlah kal yang dibutuhkan
1	1705
2	1781
3	1648
4	2230
5	1674
6	2230
7	2219
8	1978
9	1634
10	2042
11	1804
12	1658
13	2181
14	2153
15	1669
16	2076

tentukan :

- 1 . Rumusan Masalah
- 2 . Hipotesis ( $H_0$  dan  $H_a$ )  
Cara Menganalisis dan
- 3 . Kesimpulan

### e. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Lemeshow, Stanley et all, 1997. *Besar Sampel penelitian Kesehatan*, UGM Press,

4. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
6. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
7. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
8. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
9. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
10. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## **BAB VII. Uji perbedaan mean 2 sampel dengan statistic non parametrik**

---

### **Tujuan**

**Mahasiswa memahami dan dapat melakukan uji Perbedaan mean 2 sample dengan statistic non parametric dengan tepat**

---

### **Materi**

- 1. Uji Wilcoxon**
  - 2. Uji man whitney**
- 

**U**ji beda dua sampel dengan jenis sampel non parametrik atau pada sampel yang tidak berdistribusi normal tidak dapat menggunakan uji beban seperti pada penjelasan bab sebelumnya. Lalu apakah terjadi kegagalan apabila uji tersebut tetap dijalankan pada alat bantu seperti SPSS?. Seperti yang telah dijelaskan sebelumnya bahwa sebuah motor akan dapat dimasuki segala jenis bahan bakar maupun yang tidak bahan bakar jenis material tersebut sama. Artinya sepeda motor akan dapat diisi bahan bakar meskipun air kali pun. Sama dengan uji statistik ketika kita ingin menguji yang mana sebenarnya data tidak bisa diuji secara non parametrik maka angka tersebut tetap akan dapat

menunjukkan hasil. Namun yang menjadi perhatian adalah hasil yang tidak reliabel Atau mungkin bahkan perbincangan oleh orang yang paham statistik. Oleh sebab itu lebih sulit mendeskripsikan kesalahan di bidang daripada kesalahan memasukkan jenis bahan bakar pada sepeda motor. Apabila terjadi kesalahan pada pemasukan bahan bakar maka efek yang ditimbulkan akan kelihatan namun berbeda dengan statistik apabila terjadi kesalahan operasi atau analisis data akan tetap berjalan Namun demikian tidak akan terlihat sampai seseorang yang paham politik analisa dari penelitian tersebut.

Uji yang sering digunakan untuk mendeteksi adanya perbedaan pada sampel yang bersifat non parametrik adalah uji mann-witney dan uji wilcoxon.

### a. Uji wilcoxon

Uji wilcoxon atau sign YouTube merupakan uji statistik non parametrik yang digunakan untuk mengetahui ada atau tidaknya perbedaan pada dua sampel yang bersifat dependen berpasangan. Pada Uji T atas hujan kita mengenal istilah independen dan dependen maka dapat diartikan bahwa uji wilcoxon merupakan uji alternatif pada uji t yang bersifat dependen. Maka dengan asumsi tersebut jumlah dari masing-masing harus memiliki jumlah yang sama. Sifat dari uji wilcoxon biasanya menggunakan data yang bersifat ordinal atau interval yang tidak berdistribusi normal. Untuk menentukan nilai dari uji beda ini dapat menggunakan rumus sebagai berikut:

$$Z = \frac{T - \left[ \frac{1}{4N(N+1)} \right]}{\sqrt{\frac{1}{24N(N+1)(2N+1)}}$$

Dimana :  
 N = banyak data yang berubah setelah diberi perlakuan berbeda  
 T = jumlah ranking dari nilai selisih yang negative (apabila banyaknya

selisih yang positif lebih banyak dari banyaknya selisih negatif)

Apa contoh untuk menganalisa dapat di beri contoh bahwa Apakah terdapat perbedaan tekanan darah sebelum setelah pemberian jus melon pada pasien darah tinggi

Initial Pasien	sebelum treatment	Setelah Perlakuan
SD	177	131
AD	171	142
X	151	146
Z	155	126
DA	156	150
WE	135	124
TR	144	147
FGH	173	140
UJ	110	146
YT	151	144
REW	169	139
DS	141	125
VB	145	134
XV	144	136
ER	179	138
FF	175	123
MM	171	138
FS	157	122
GSF	157	123
YTG	189	144
OYU	190	80

Penyelesaian menggunakan SPSS

**Rumusan masalah**

Apakah terdapat perbedaan tekanan darah sebelum dan setelah pemberian jus melon pada pasien darah tinggi?

**Hipotesis**

H<sub>0</sub> = tidak terdapat perbedaan tekanan darah sebelum dan setelah pemberian jus melon pada pasien darah tinggi?

H<sub>a</sub> = terdapat perbedaan tekanan darah sebelum dan setelah pemberian jus melon pada pasien darah tinggi?

**Taraf Kepercayaan**

Confident level yang digunakan adalah 5%

**Uji yang digunakan**

Dependent T Test

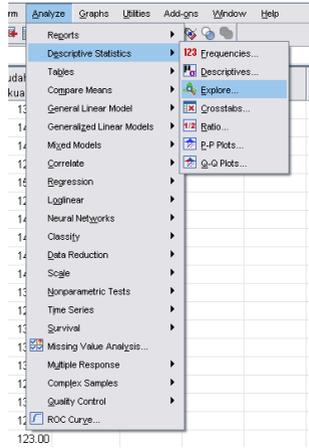
Masukkan data ke SPSS seperti gambar berikut

	Sebelumperla kuan	Sesudahperla kuan	var	
1	177.00	131.00		
2	171.00	142.00		
3	151.00	146.00		
4	155.00	126.00		
5	156.00	150.00		
6	135.00	124.00		
7	144.00	147.00		
8	173.00	140.00		
9	110.00	146.00		
10	151.00	144.00		
11	169.00	139.00		
12	141.00	125.00		
13	145.00	134.00		
14	144.00	136.00		
15	179.00	138.00		
16	175.00	123.00		
17	171.00	138.00		
18	157.00	122.00		
19	157.00	123.00		
20	189.00	144.00		
21	190.00	80.00		
22				

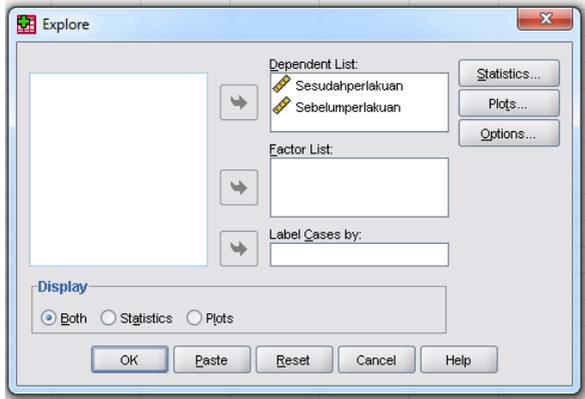
Untuk merubah nilai variabel dapat mengganti pada sheet ini

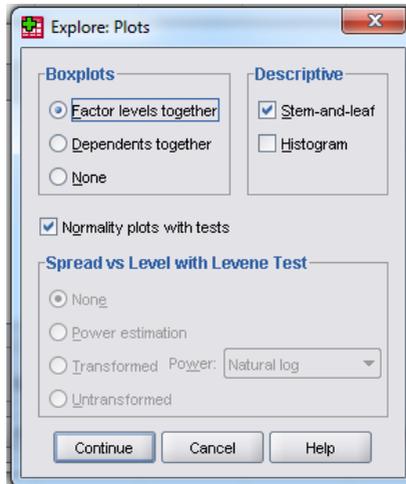
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Sebelumperlakuan	Numeric	8	2		None	None	8	Right	Scale
2	Sesudahperlakuan	Numeric	8	2		None	None	8	Right	Scale
3										

Kita uji cobat dengan uji normalitas data dengan cara masuk ke analyze > Deskriptive Statistik > Explore seperti gambar di bawah ini.



Kemudian masukkan masing masing variable pada collom yang tersedia dan jangan lupa mencentang normality with test seperti pada gambar berikut





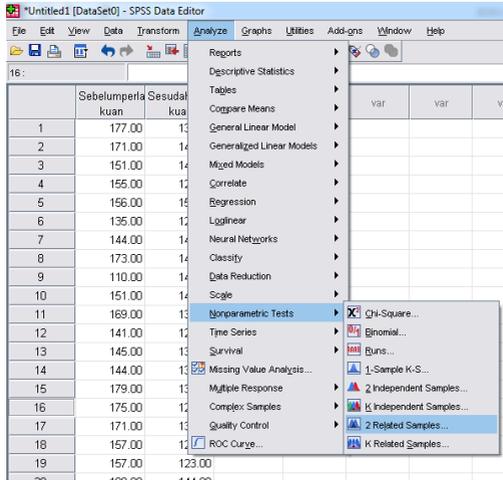
Setelah di klik ok maka keluar hasil dan cari pada bagian ini  
**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Sesudahperlakuan	.182	21	.067	.785	21	.000
Sebelumperlakuan	.125	21	.200 <sup>*</sup>	.959	21	.498

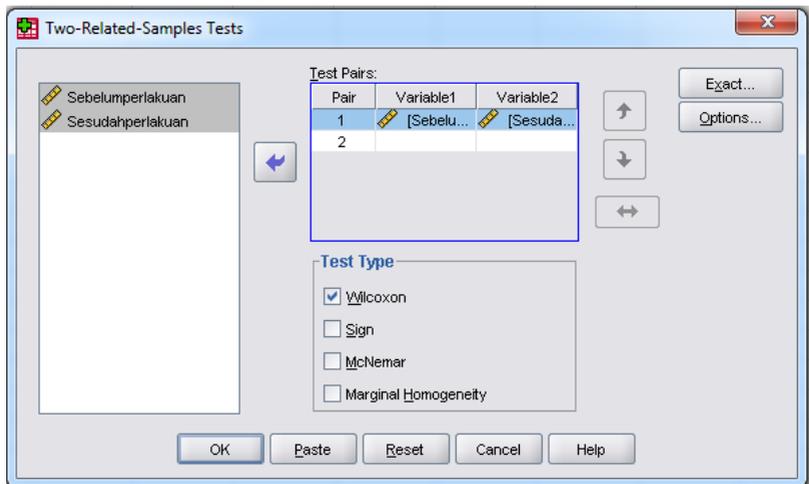
a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Dari hasil yang didapatkan maka keluar uji kolmogorv smirnov dan Shapiro wilk. Karena besar sampel yang kata gunakan relative sedikit yaitu 21 maka yang digunakan untuk pengujian adalah Shapiro wilk. Memperhatikan hasil yang keluar maka data sebelum perlakuan menunjukkan nilai yang cukup rendah yaitu 0.000 maka karena nila p value < 0.05 maka kita berasumsi bahwa data tersebut tidak tersebar secara normal. Sehingga yang dilakukan untuk analisis tidak diperkenankan menggunakan uji t seperti biasa. Ada dua pilihan ketika data berstatus demikian yaitu melakukan tranformasi dan yang kedua kita melakukan uji non parametric yaitu Wilcoxon test. Pada kesempatan kali ini kita langsung melakukan uji Wilcoxon



Pindahkan variable ke kolom variable



Tekan ok kemudian akan muncul kolom sebagai berikut

### Ranks

		N	Mean Rank	Sum of Ranks
Sesudahperlakuan - Sebelumperlakuan	Negative Ranks	19 <sup>a</sup>	11.26	214.00
	Positive Ranks	2 <sup>b</sup>	8.50	17.00
	Ties	0 <sup>c</sup>		
	Total	21		

- a. Sesudahperlakuan < Sebelumperlakuan
- b. Sesudahperlakuan > Sebelumperlakuan
- c. Sesudahperlakuan = Sebelumperlakuan

### Test Statistics<sup>b</sup>

	Sesudahperla kuan - Sebelumperla kuan
Z	-3.424 <sup>a</sup>
Asymp. Sig. (2-tailed)	.001

- a. Based on positive ranks.
- b. Wilcoxon Signed Ranks Test

Sehingga dapat disimpulkan  $H_a$  diterima dan  $H_0$  di tolak. Sehingga kesimpulan yang dapat ditarik adalah terdapat perbedaan tekanan darah sebelum dan setelah pemberian jus melon pada pasien darah tinggi

## b. uji mann-withney

Uji Mann Whitney lebih lazim digunakan untuk sampel yang bersifat bebas atau nama lain apabila kita menggunakan uji parametrik yaitu independen tes. Sehingga apabila kita menggunakan uji ini besar sampel pada masing-masing tidak harus sama besar atau bisa berbeda.

$$U_1 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

$$U_2 = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - \sum R_1$$

$$U_1 = n_1 \cdot n_2 - U_2$$

$$U_2 = n_1 \cdot n_2 - U_1$$

Keterangan :

- U<sub>1</sub> = Penguji U<sub>1</sub>
- U<sub>2</sub> = Penguji U<sub>2</sub>
- R<sub>1</sub> = Jumlah rank sampel 1
- R<sub>2</sub> = Jumlah rank sampel 2
- n<sub>1</sub> = Banyaknya anggota sampel 1
- n<sub>2</sub> = Banyaknya anggota sampel 2

Sebagai contoh penggunaan uji mann-withney maka berikut gambaran permasalahan. Seorang ATLM ingin meneliti perbedaan Antara denyut jantung laki laki dengan denyut jantung perempuan pada keadaan setelah lari 50 meter. Data adalah sebagai berikut

Initial Pasien	laki laki	Perempuan
SD	177	131
AD	171	142
X	151	146
Z	155	126
DA	156	150
WE	135	124
TR	144	147
FGH	173	140

UJ	110	146
YT	151	144
REW	169	139
DS	141	125
VB	145	134
XV	144	136
ER	179	138
FF	175	123
MM	171	138
FS	157	122
GSF	157	123
YTG	189	144
OYU	190	80

---

Karena data sama dengan contoh diatas maka kita langsung saja menganalisa permasalahan tersebut dengan menggunakan uji mann Whitney.

### **Rumusan masalah**

Apakah terdapat perbedaan Antara denyut jantung laki laki dengan denyut jantung perempuan pada keadaan setelah lari 50 meter?.

### **Tujuan**

Untuk mengetahui perbedaan Antara denyut jantung laki laki dengan denyut jantung perempuan pada keadaan setelah lari 50 meter.

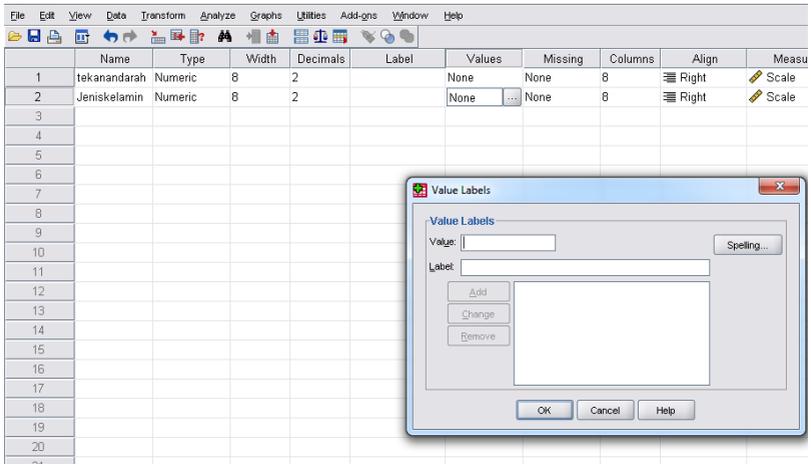
### **Hipotesis**

Ho = tidak terdapat perbedaan Antara denyut jantung laki laki dengan denyut jantung perempuan pada keadaan setelah lari 50 meter.

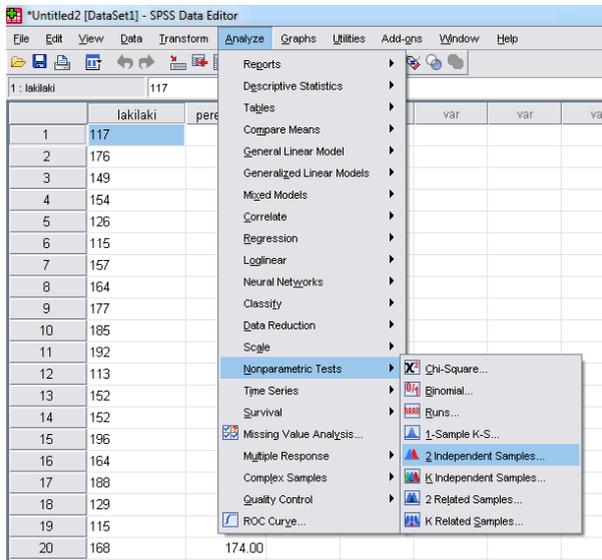
Ha = terdapat perbedaan Antara denyut jantung laki laki dengan denyut jantung perempuan pada keadaan setelah lari 50 meter.

Untuk menjawab soal yang ada maka data di masukkan seperti menganalisis dengan menggunakan independent T Test .

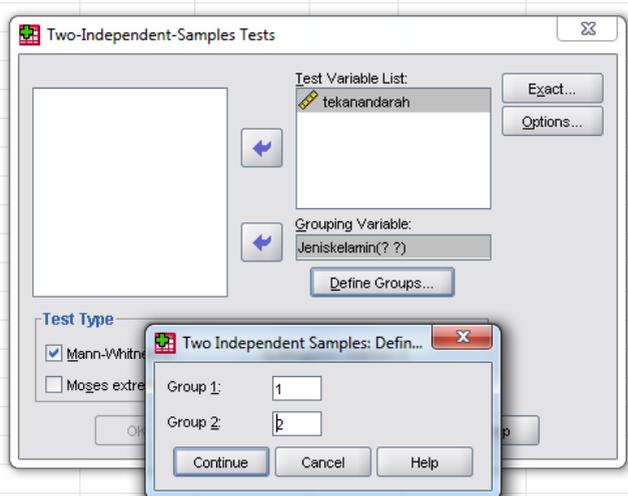
	VAR00001	VAR00002
1	177.00	1.00
2	171.00	1.00
3	151.00	1.00
4	155.00	1.00
5	156.00	1.00
6	135.00	1.00
7	144.00	1.00
8	173.00	1.00
9	110.00	1.00
10	151.00	1.00
11	169.00	1.00
12	141.00	1.00
13	145.00	1.00
14	144.00	1.00
⋮		
22	131.00	2.00
23	142.00	2.00
24	146.00	2.00
25	126.00	2.00
26	150.00	2.00
27	124.00	2.00
28	147.00	2.00



Untuk mengganti jenis label dapat dilakukan seperti mengganti pada uji T independent kemudian pilih Analyze > Non Parametric Test > 2 Independent Sample.



Kemudian masukkan variable sesuai dengan katagori seperti pada Gambar berikut



Sehingga hasilnya akan keluar seperti berikut

## Mann-Whitney

### Ranks

		N	Mean Rank	Sum of Ranks
tekanandarah	Jeniskel... laki laki	21	29.48	619.00
	perempuan	21	13.52	284.00
	Total	42		

### Test Statistics<sup>a</sup>

	tekanandarah
Mann-Whitney U	53.000
Wilcoxon W	284.000
Z	-4.216
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Jeniskelamin

Berdasarkan hasil analisis menggunakan Mann Whitney maka dapat disimpulkan  $H_0$  diterima berdasarkan nilai P value lebih bandingkan 0,05. Sehingga bunyi dari Kesimpulan tersebut adalah

### c. Ringkasan materi

Uji non parametrik merupakan alternatif ketika uji parametrik tidak dilakukan karena terkendala persyaratan. Sehingga apabila parametrik dipaksakan maka hasil yang dihasilkan untuk tidak akan menunjukkan hasil yang reliable. Uji Mann Whitney dan Wilcoxon merupakan pengganti uji parametrik Uji T. Uji Mann Whitney lebih tepat digunakan untuk menggantikan independent t-test sedangkan uji wilcoxon lebih tepat digunakan untuk menggantikan dependen T test.

### d. Latihan soal

Jawablah soal berikut dengna tepat!

5. Seorang peneliti ingin mengetahui Pengaruh pemberian jus semangka terhadap peningkatan tekanan darah ibu-ibu lansia, jus semangka yang digunakan adalah jus murni tanpa tambahan apapun, hasil dari analisa kasus adalah sebagai berikut:

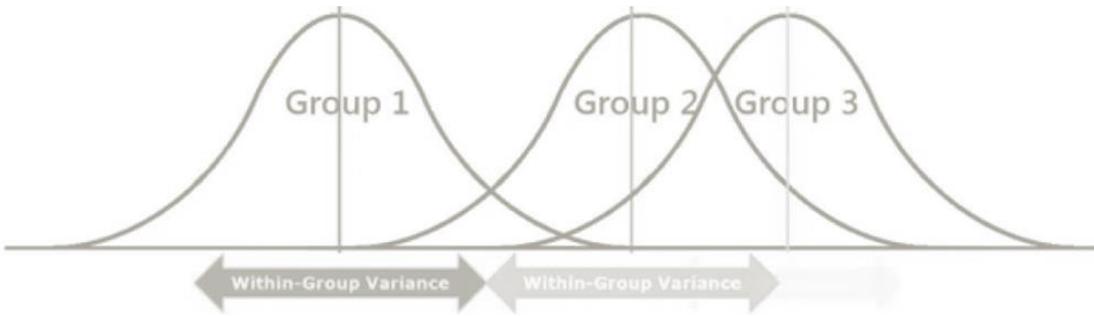
sebelum di beri jus	130	144	145	120	150	130	120	130
	144	134	143	125	134	123	80	90
sesudah di beri jus	125	130	140	115	150	129	119	130
	120	130	140	121	131	120	143	144

dari kasus tersebut tentukan!

- a. rumusan masalah
- b. hipotesis
- c. uji yang cocok
- d. kesimpulan dari permasalahan tersebut ( sertakan tahapan hitungannya)

## f. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Lemeshow, Stanley et all, 1997. *Besar Sampel penelitian Kesehatan*, UGM Press,
4. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
6. Ott R. L. & Longnecker M. 2010. *An Introduction to statistical Methods and Data Analysis: Sixth Edition*. Cengage Learning
7. Purnawan, J., 1997 *Teknik Analisis Data*, Rineka Cipta,
8. Sudjana, 1996. *Metoda Statistika Edisi ke 6*. Bandung: Penerbit Tarsito
9. Santoso, S. 2014. *Panduan lengkap spss versi 20 edisi revisi*, Jakarta:gramedia
10. Widyato, J. 2012. *Spss for windows*. Surakarta:FKIP UMS



## Bab VIII. Uji perbedaan mean lebih dari 2 sampel dengan statistic parametrik

---

### Tujuan

Mahasiswa memahami dan dapat melakukan uji Perbedaan lebih dari 2 sample dengan statistic parametric dengan tepat

---

### Materi

1. Pengertian Anova
  2. Uji Anova one way dan anova two way
  3. RAK dan RAL
- 

**A**nalysis of variance atau seringkali disebut dengan Anova merupakan uji yang digunakan untuk mendeteksi atau mengidentifikasi perbedaan sampel lebih dari 2 . Anova ditemukan oleh Ronald A. Fisher (1890-1962) pada tahun 1920-an. Apakah azan azan maulah mainkan kamu nakaleliau merupakan ahli di bidang pertanian. uji anova (F) daerah penolakan selalu pada posisi sebelah kanan. Angka nilai tabel merupakan batas-batas daerah penolakan. Sisi sebelah kiri secara otomatis nilai selalu negatif walaupun pada nilai tabel tidak ada nilai negatif.

Uji anova mengacu pada kurva normal yang condong ke kanan. Pada kurva normal posisi ditengah yang membagi kurva simetris merupakan titik angka nol, sedangkan kurva menceng ke kanan posisi nol pada ujung sebelah kiri. Asumsi pada penelitian anova adalah Populasi-populasi yang akan diuji berdistribusi

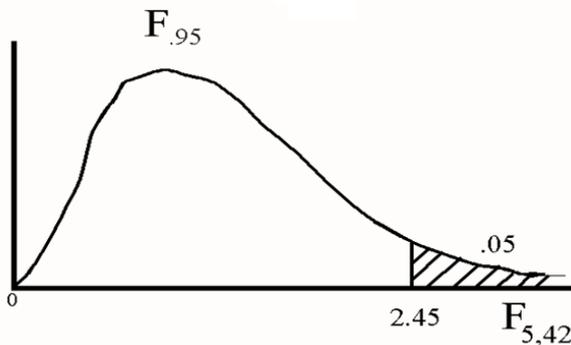
normal, Varians untuk masing-masing populasi adalah sama, Sampel tidak berhubungan satu sama lain.

Untuk menganalisa suatu permasalahan khususnya didunia kesehatan, ada kalanya di perlukan pengujian terhadap kumpulan hasil pengamatan mengenai suatu hal, misalnya seorang ahli analis kesehatan ingin membandingkan daya hambat ekstrak tapak data terhadap bakteri *Staphylococcus aureus* dengan menggunakan kosentrasi 0%, 25%, dan 50%. Maka apabila berkesimpulan bahwa uji yang sebelumnya (Uji T) bisa digunakan tentu pemikiran tersebut tidak dapat dibenarkan. Uji T dependent maupun uji T independent hanya terbatas pada 2 sampel dan tidak mampu menganalisa lebih dari 2 sampel (Pada kasus di atas yang kita adanalisa adalah tiga sampel yaitu 0%, 25%, dan 50%). Sehingga diperlukan uji yang mampu menangani permasalahan tersebut yaitu ANOVA (Analyses of Variances). Uji dalam anova menggunakan uji F karena dipakai untuk pengujian lebih dari 2 sampel. Dalam praktik, analisis varians dapat merupakan uji hipotesis (lebih sering dipakai) maupun pendugaan (estimation, khususnya di bidang genetika terapan).

Kasus lain yang dapat di terapkan misalnya, seorang ahli gizi ingin menganalisa apakah terdapat perbedaan kadar vitamin C pada Buah naga yang telah diberi beberapa perlakuan yaitu, 1. Diawetkan didalam kaleng, 2. Dibekukan didalam lemari pendingin., 3. Buah segar. Maka dibutuhkan uji statistika ANOVA dalam penelitian ini, sekali lagi uji T tidak dapat digunakan karena Uji T hanya mampu menganalisa jumlah sampel paling banyak 2.

Anova (Analysis of variances) digunakan untuk melakukan analisis komparasi multivariabel. Teknik analisis komparatif dengan menggunakan tes "t" yakni dengan mencari perbedaan yang signifikan dari dua buah mean hanya efektif bila jumlah variabelnya dua. Jenis data pada variable independent yang tepat untuk anova adalah nominal dan ordinal, jika data pada variabel bebasnya dalam bentuk interval atau ratio maka harus diubah dulu dalam bentuk ordinal atau nominal. Sedangkan variabel terikatnya adalah data interval atau ratio. Adapun asumsi dasar yang harus terpenuhi dalam analisis varian adalah :

1. Kenormalan Distribusi data harus normal, agar data berdistribusi normal dapat ditempuh dengan cara memperbanyak jumlah sampel dalam kelompok.
2. Kesamaan variansi Setiap kelompok hendaknya berasal dari populasi yang sama dengan variansi yang sama pula. Bila banyaknya sampel sama pada setiap kelompok maka kesamaan variansinya dapat diabaikan. Tapi bila banyak sampel pada masing masing kelompok tidak sama maka kesamaan variansi populasi sangat diperlukan.
3. Pengamatan bebas Sampel hendaknya diambil secara acak (random), sehingga setiap pengamatan merupakan informasi yang bebas. Anova lebih akurat digunakan untuk sejumlah sampel yang sama pada setiap kelompoknya, misalnya masing masing variabel setiap kelompok jumlah sampel atau responden nya sama sama 250 orang



Anova dapat digolongkan kedalam beberapa kriteria, yaitu :

1. Klasifikasi 1 arah (One Way ANOVA) Anova klasifikasi 1 arah merupakan ANOVA yang didasarkan pada pengamatan 1 kriteria atau satu faktor yang menimbulkan variasi.

2. Klasifikasi 2 arah (Two Way ANOVA) ANOVA klasifikasi 2 arah merupakan ANOVA yang didasarkan pada pengamatan 2 kriteria atau 2 faktor yang menimbulkan variasi.
3. Klasifikasi banyak arah (MANOVA) ANOVA banyak arah merupakan ANOVA yang didasarkan pada pengamatan banyak kriteria.

### a. Anova satu arah dan dua arah

Anova satu arah (one way anova) digunakan apabila yang akan dianalisis terdiri dari satu variabel terikat dan satu variabel bebas. Interaksi suatu kebersamaan antar faktor dalam mempengaruhi variabel bebas, dengan sendirinya pengaruh faktor-faktor secara mandiri telah dihilangkan. Jika terdapat interaksi berarti efek faktor satu terhadap variabel terikat memiliki garis yang tidak sejajar dengan efek faktor lain terhadap variabel terikat sejajar (saling berpotongan), maka antara faktor tidak mempunyai interaksi.

Ada tiga bagian pengukuran variabilitas pada data yang akan dianalisis dengan anova, yaitu :

1. Variabilitas antar kelompok (between treatments variability)  
Variabilitas antar kelompok adalah variansi mean kelompok sampel terhadap rata-rata total, sehingga variansi lebih terpengaruh oleh adanya perbedaan perlakuan antar kelompok, atau Jumlah Kuadrat antar kelompok.
2. Variabilitas dalam kelompok (within treatments variability)  
Variabilitas dalam kelompok adalah variansi yang ada dalam masing-masing kelompok. Banyaknya variansi akan tergantung pada banyaknya kelompok. Variansi tidak terpengaruh oleh perbedaan perlakuan antar kelompok, atau Jumlah Kuadrat dalam.
3. Jumlah kuadrat penyimpangan total (total sum of squares)  
Jumlah kuadrat penyimpangan total adalah jumlah kuadrat selisih antara skor individual dengan mean totalnya, atau JKT.

Sehingga didalam menyelesaikan uji hipotesis yang menggunakan ANOVA maka kita membutuhkan table

Langkah – langkah menghitung ANAVA 1 jalur :

- 1) Hipotesis
  - a. Hipotesis anava 1 jalur  
 $H_0 : \mu_1 = \mu_2 = \mu_3$   
 $H_1 : \text{paling sedikit satu tanda sama dengan tidak berlaku.}$
- 2) Menghitung jumlah (sum of square) total (JKt), antar kelompok (JKa), dan dalam kelompok (JKd). berikut rumus yang digunakan untuk menghitung jumlah kuadrat:
  - a.  $Jkt = \sum x^2 - (\sum x)^2/N$
  - b.  $Jka = [(\sum x_1)^2/n_1 + (\sum x_2)^2/n_2 + \dots + (\sum x_k)^2/n_k]$
  - c.  $Jkd = Jkt - Jka$Keterangan :
  - $(\sum x)^2/n = sk = cf = \text{suku koreksi (correction factor).}$
  - $N = \text{jumlah subyek}$
- 3) Menghitung derajat kebebasan (degree of freedom) total (dbt), antar kelompok (dba) dan dalam kelompok (dbd). berikut rumus yang digunakan untuk menghitung derajat kebebasan:
  - a.  $Dbt = N-1$
  - b.  $Dbk = K-1$
  - c.  $Dbt = N - K$Dimana :
  - $Db = \text{derajat kebebasan}$
  - $N = \text{jumlah subyek}$
  - $K = \text{jumlah kelompok}$
- 4) Menghitung rata-rata kuadrat (mean of square ) antar kelompok (Rka), dan dalam kelompok (Rkd). berikut rumus yang digunakan untuk menghitung rata-rata kuadrat:
  - $Rka = Jka /dba$
  - $Rkd = Jkd /dbd$

- 5) Menghitung rasio F dimana F rasio itu adalah perbandingan antara rata – rata kuadrat antar kelompok dengan rata-rata kuadrat dalam kelompok.
- $F = Jka / Rkd$
- 6) Melakukan interpretasi dan uji signifikansi pada rasio F. ada dua F yang digunakan untuk melakukan interpretasi dan uji signifikansi yaitu: F empirik dan F teoritik. Dimana F empirik adalah rasio F atau F hasil hitung dan F teoritik adalah F yang diperoleh dari tabel F. untuk menentukan nilai F teoritik pada tabel sama seperti menentukan nilai t tabel. Namun yang perlu diperhatikan dalam menentukan F tabel adalah :
- Derajat kebebasan pembilang = derajat kebebasan antar kelompok, dan
  - Derajat kebebasan penyebut = derajat kebebasan dalam kelompok Setelah menentukan F teoritik maka langkah selanjutnya membandingkan F empirik dan F teoritik. Jika:
  - $F \text{ empirik} \geq F \text{ teoritik}$  ; maka diinterpretasikan signifikan yang berarti ada perbedaan.
  - $F \text{ empiric} < F \text{ teoritik}$ ; maka diinterpretasikan tidak signifikan yang berarti tidak terdapat perbedaan.

Sebagai gambaran uji one way dengan judul percobaan pengaruh pemberian kafein pertumbuhan peningkatan tekanan darah pada tikus wistar jantan dengan menggunakan beberapa kosentrasi. Pada kasus tersebut merupakan uji anova one way. Sedangkan apabila peneliti ingin membedakan pengaruh pemberian kafein pertumbuhan peningkatan tekanan darah pada tikus wistar jantan dan cafein yang digunakan berasal dari kopi dan teh maka pada penelitian tersebut termasuk uji anova two way. Berikut gambaran mengenai anova one way dan two way

**Anova One way**

asal kafein	ulangan	Peningkatan tekanan darah pada konsentrasi			
		20%	40%	60%	80%
kafein dari kopi	1	7.1	11.5	19.1	22.7
	2	6.6	12	16.1	22.4
	3	9.1	11.4	18.9	22.2
	4	7.4	13.3	18.4	22.5
	5	8.7	10.3	18.4	21.6

**Anova two way**

asal kafein	ulangan	Peningkatan tekanan darah pada konsentrasi			
		20%	40%	60%	80%
kafein dari kopi	1	7.1	11.5	19.1	22.7
	2	6.6	12	16.1	22.4
	3	9.1	11.4	18.9	22.2
	4	7.4	13.3	18.4	22.5
	5	8.7	10.3	18.4	21.6
	6	12	14.8	18.5	18.6
kafein dari teh	1	10.3	15.6	18.7	18.9
	2	7.4	16.4	20.2	16.9
	3	12.1	16.7	19.6	22.2
	4	9.9	13.1	22	24.4
	5	13	12.7	21.7	19.3
	6	12.2	12.3	19.6	22.1

Untuk mengaplikasikan langsung marilah kita praktikan saja menggunakan data yang ada. Untuk memulai pada kasus one way anova . Pertama masukkan data secara kebawah kemudian pada

variable berikutnya beri penanda 1 untuk 20%, angka 2 untuk 40% dan seterusnya.

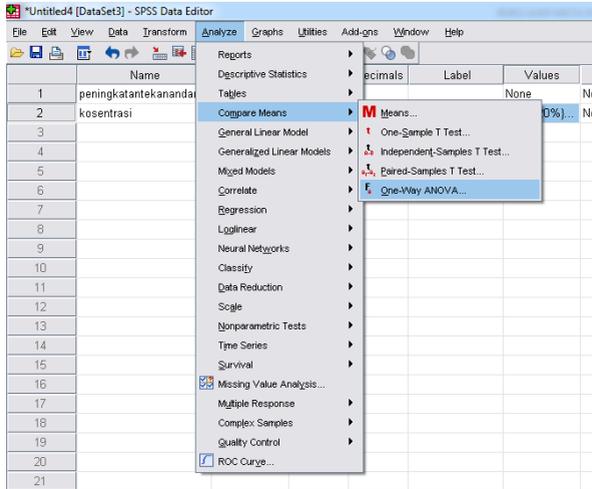
	VAR00001	VAR00002
1	7.10	1.00
2	6.60	1.00
3	9.10	1.00
4	7.40	1.00
5	8.70	1.00
6	11.50	2.00
7	12.00	2.00
8	11.40	2.00
9	13.30	2.00
10	10.30	2.00
11	19.10	3.00
12	16.10	3.00
13	18.90	3.00
14	18.40	3.00
15	18.40	3.00
16	22.70	4.00
17	22.40	4.00
18	22.20	4.00
19	22.50	4.00
20	21.60	4.00

Untuk mengganti penanda tersebut gantilah label yang terletak pada sheet variable view seperti pada gambar berikut

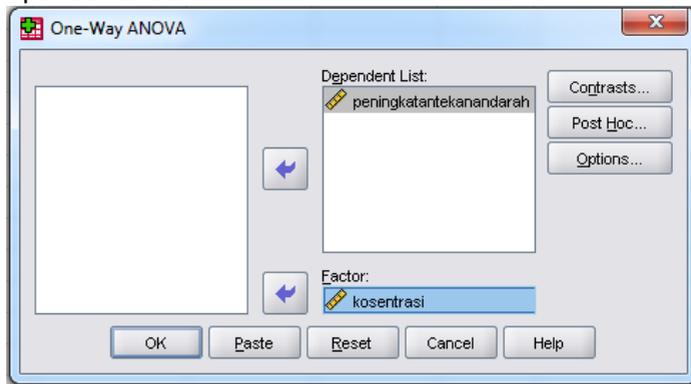
Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
peningkatantekananarah	Numeric	8	2		None	None	8	Right	Scale
kosentrasi	Numeric	8	2		{1.00, 2.00} ...	None	8	Right	Scale

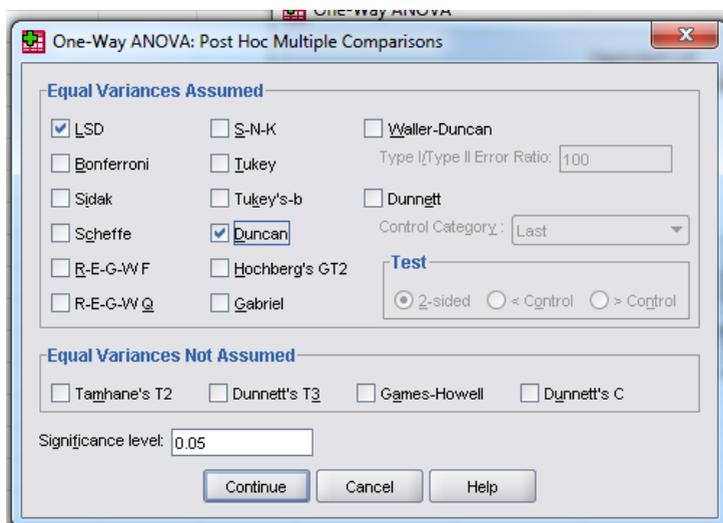
Kemudian lanjutkan dengan analyze > cormpare mean > one way anova



Masukkan hasil penghitungan kenaikan prosentasi pada kolom dependent list



Kemudian untuk mengatur uji lanjut maka klik post hoc



## ANOVA

peningkatantekanan darah

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	630.642	3	210.214	212.552	.000
Within Groups	15.824	16	.989		
Total	646.466	19			

## Post Hoc

### Multiple Comparisons

Dependent Variable: peningkatantekanandarah

	(I) kose ntras i	(J) kose ntras j	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	20%	40%	-3.92000 <sup>a</sup>	.62897	.000	-5.2534	-2.5866
		60%	-10.40000 <sup>a</sup>	.62897	.000	-11.7334	-9.0666
		80%	-14.50000 <sup>a</sup>	.62897	.000	-15.8334	-13.1666
	40%	20%	3.92000 <sup>a</sup>	.62897	.000	2.5866	5.2534
		60%	-6.48000 <sup>a</sup>	.62897	.000	-7.8134	-5.1466
		80%	-10.58000 <sup>a</sup>	.62897	.000	-11.9134	-9.2466
	60%	20%	10.40000 <sup>a</sup>	.62897	.000	9.0666	11.7334
		40%	6.48000 <sup>a</sup>	.62897	.000	5.1466	7.8134
		80%	-4.10000 <sup>a</sup>	.62897	.000	-5.4334	-2.7666
	80%	20%	14.50000 <sup>a</sup>	.62897	.000	13.1666	15.8334
		40%	10.58000 <sup>a</sup>	.62897	.000	9.2466	11.9134
		60%	4.10000 <sup>a</sup>	.62897	.000	2.7666	5.4334

.<sup>a</sup>. The mean difference is significant at the 0.05 level.

## Homogeneous

### peningkatantekananarah

	kose ntras i	N	Subset for alpha = 0.05			
			1	2	3	4
Duncan <sup>a</sup>	20%	5	7.7800			
	40%	5		11.7000		
	60%	5			18.1800	
	80%	5				22.2800
	Sig.			1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

## Cara membaca hasil

Berdasarkan hasil yang didapat bangun menunjukkan nilai P value adalah 0,000. Banyak sekali mahasiswa beranggapan kalau hasil seperti ini adalah keliru atau salah. Sehingga dari hasil tersebut menunjukkan nilai P value kurang dari 0.05. Sehingga kesimpulan dari penelitian kelas terdapat perbedaan pengaruh dengan detak jantung menciit. Berdasarkan hasil yang didapat bangun

menunjukkan nilai P value adalah 0,000. Banyak sekali mahasiswa beranggapan kalau hasil seperti ini adalah keliru atau salah. Sehingga dari hasil tersebut menunjukkan nilai P value kurang dari 0.05. Sehingga kesimpulan dari penelitian kelas terdapat perbedaan pengaruh dengan detak jantung mencet. Uji Anova Memiliki keterbatasan yaitu hanya dapat membedakan tanpa memberi konsep yang mana yang berbeda dan mana yang sama. Uji post hoc menjadi pelengkap data dari anova. post hock mampu memisahkan antara variabel yang satu dengan variabel lainnya. Hasil uji LSD dan duncan sebagai perwakilan uji sampel menunjukkan adanya perbedaan pada masing masing konsentrasi

### Contoh Uji Anova two way

Seorang ATLM, melakukan percobaan pengaruh pemberian kafein terhadap peningkatan tekanan darah pada tikus wistar jantan. Kafein didapat dari dua bahan yaitu teh dan kopi, konsentrasi yang digunakan adalah 20 ppm, 40 ppm, 60 ppm dan 80 ppm, data diperoleh di rubah ke dalam bentuk prosentasi sesuai dengan data awal pengamatan, berikut rangkuman data.

asal kafein	ulangan	Peningkatan tekanan darah pada konsentrasi			
		20%	40%	60%	80%
kafein dari kopi	1	7.1	11.5	19.1	22.7
	2	6.6	12	16.1	22.4
	3	9.1	11.4	18.9	22.2
	4	7.4	13.3	18.4	22.5
	5	8.7	10.3	18.4	21.6
	6	12	14.8	18.5	18.6
kafein dari teh	1	10.3	15.6	18.7	18.9
	2	7.4	16.4	20.2	16.9
	3	12.1	16.7	19.6	22.2

4	9.9	13.1	22	24.4
5	13	12.7	21.7	19.3
6	12.2	12.3	19.6	22.1

## Penyelesaian

Seperti uji uji sebelumnya untuk menyelesaikan permasalahan ini maka diperlukan beberapa tahapan:

### Tahap 1. Perumusan Masalah

Rumusan masalah dalam penelitian ini dapat dipisahkan berdasarkan kebutuhan analisis

- Apakah terdapat perbedaan kemampuan kafein dari kopi dan teh untuk meningkatkan tekanan darah pada tikus wistar jantan?
- Berapakah kosentrasi yang optimum yang dapat meningkatkan tekanan darah pada tikus wistar jantan ?

### Tahap 2. Perumusan tujuan

Tujuan pada penelitian ini adalah untuk:

- Mengetahui apakah terdapat perbedaan kemampuan kafein dari kopi dan teh untuk meningkatkan tekanan darah pada tikus wistar jantan.
- Mengetahui Berapakah kosentrasi yang optimum yang dapat meningkatkan tekanan darah pada tikus wistar jantan.

### Tahap 3. Menentukan hipotesis

Hipotesis pada penelitian ini membutuhkan dua hipotesis, hipotesis dari rumusan masalah pertama adalah

**Ho:** tidak terdapat perbedaan kemampuan kafein dari kopi dan teh untuk meningkatkan tekanan darah pada tikus wistar jantan.

**Ha :** terdapat perbedaan kemampuan kafein dari kopi dan teh untuk meningkatkan tekanan darah pada tikus wistar jantan.

Sedangkan hipotesis pada rumusan masalah yang kedua adalah

**Ho:** tidak terdapat kosentrasi yang optimum yang dapat meningkatkan tekanan darah pada tikus wistar jantan.

**Ha:** terdapat kosentrasi yang optimum yang dapat meningkatkan tekanan darah pada tikus wistar jantan.

### **Tahap 3. Menentukan area kritis**

Tingkat area kritis ditentukan kisaran 5% atau dengan confident level 95%.

### **Tahap 4. Menentukan jenis data**

Jenis data pada penelitian ini secara keseluruhan adalah rasio, sehingga saran yang dilakukan adalah uji parametric, namun demikian syarat uji parametric adalah terdistribusi normal.

#### **a. Menentukan normalitas data**

Seelum melakukan analisis normalitas data maka data sebaiknya di atur di exel sesuai dengan gambar di bawah ini

	kenaikan	kosentrasi	jenis_kafein	var
1	7.1	1.00	1.00	
2	6.6	1.00	1.00	
3	9.1	1.00	1.00	
4	7.4	1.00	1.00	
5	8.7	1.00	1.00	
6	12.0	1.00	1.00	
7	11.5	2.00	1.00	
8	12.0	2.00	1.00	
9	11.4	2.00	1.00	
10	13.3	2.00	1.00	
11	10.3	2.00	1.00	
12	14.8	2.00	1.00	
13	19.1	3.00	1.00	
14	16.1	3.00	1.00	
15	18.9	3.00	1.00	
16	18.4	3.00	1.00	
17	18.4	3.00	1.00	
18	18.5	3.00	1.00	
19	22.7	4.00	1.00	
20	22.4	4.00	1.00	
21	22.2	4.00	1.00	
22	22.5	4.00	1.00	

ketentuan		
kosentrasi	1	20%
	2	40%
	3	60%
	4	80%
asal kefein	1	kopi
	2	t eh

Kemudian kopi paste ke kolom spss dan di atur pada sheet variable view.

1		None	None	6	Left
2		0, 20%}...	None	8	Right
2		None	None	8	Right

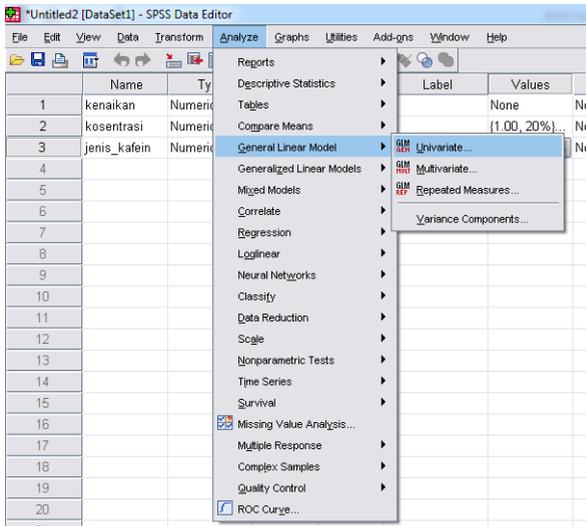


Ubah label pada variable jenis kafein yang digunakan

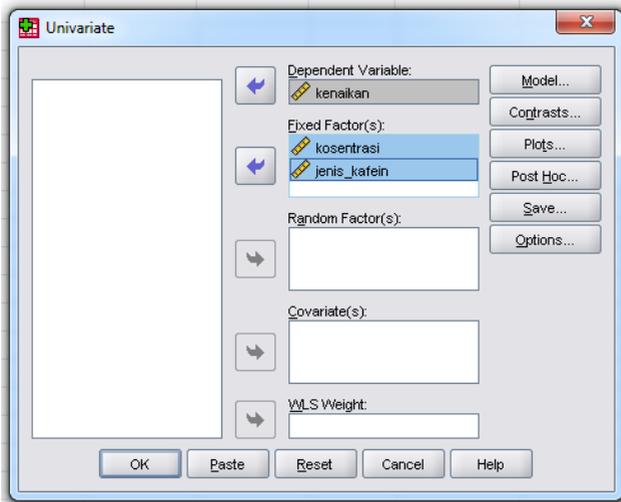
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values
1	kenaikan	Numeric	2	1		None
2	kosentrasi	Numeric	8	2		{1.00, 20%}...
3	jenis_kafein	Numeric	8	2		None ...
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						

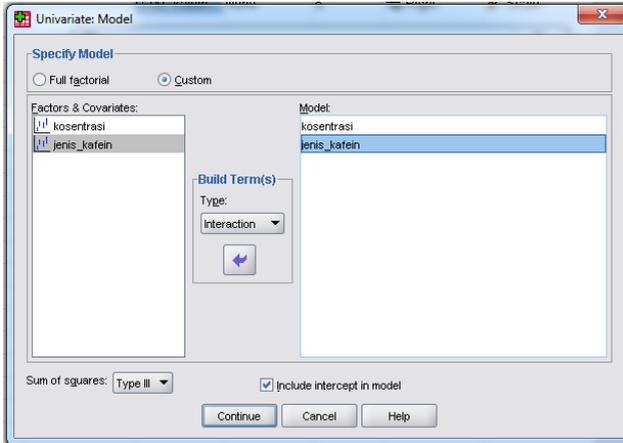
Kemudian masuk ke analyze >General Linear Model> Univariat



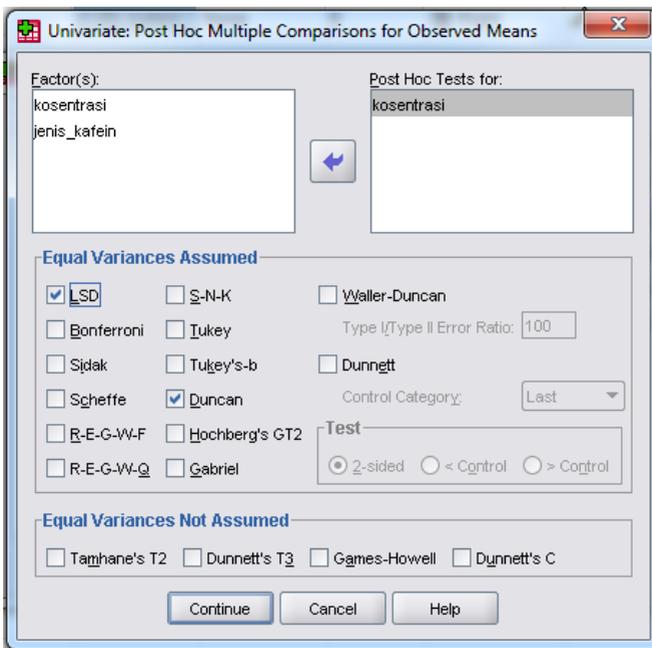
Setting seperti gambar berikut



Tekan model kemudian arahkan seperti gambar berikut



Klik tombol post hock dan arahkan kosentrasi untuk dianalisis menggunakan LSD dan Duncan



Setelah itu lanjut di ok. Kemudian akan keluar hasil seperti ini

### Tests of Between-Subjects Effects

Dependent Variable: kenaikan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1037.604 <sup>a</sup>	4	259.401	69.438	.000
Intercept	12061.850	1	12061.850	3.229E3	.000
kosentrasi	1013.944	3	337.981	90.473	.000
jenis_kafein	23.660	1	23.660	6.334	.016
Error	160.636	43	3.736		
Total	13260.090	48			
Corrected Total	1198.240	47			

a. R Squared = .866 (Adjusted R Squared = .853)

### Multiple Comparisons

Dependent Variable: kenaikan

	(I) kose ntras	(J) kose ntras	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	20%	40%	-3.692'	.7891	.000	-5.283	-2.100
		60%	-9.617'	.7891	.000	-11.208	-8.025
		80%	-11.500'	.7891	.000	-13.091	-9.909
	40%	20%	3.692'	.7891	.000	2.100	5.283
		60%	-5.925'	.7891	.000	-7.516	-4.334
		80%	-7.808'	.7891	.000	-9.400	-6.217
	60%	20%	9.617'	.7891	.000	8.025	11.208
		40%	5.925'	.7891	.000	4.334	7.516
		80%	-1.883'	.7891	.021	-3.475	-.292
	80%	20%	11.500'	.7891	.000	9.909	13.091
		40%	7.808'	.7891	.000	6.217	9.400
		60%	1.883'	.7891	.021	.292	3.475

Based on observed means.

The error term is Mean Square(Error) = 3.736.

\*. The mean difference is significant at the 0.05 level.

### kenaikan

	kone ntras i	N	Subset			
			1	2	3	4
Duncan <sup>a</sup>	20%	12	9.650			
	40%	12		13.342		
	60%	12			19.267	
	80%	12				21.150
	Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

Based on observed means.

The error term is Mean Square(Error) = 3.736.

a. Uses Harmonic Mean Sample Size = 12.000.

### Intrepretasi hasil

Berdasarkan hasil yang didapat menggunakan anu batu maka dapat disimpulkan bahwa di antara penggunaan dua jenis kafein menunjukkan perbedaan yang nyata yang dibuktikan dengan nilai p-value sebesar 0,000 apa kurang dari 0,05. Hal ini menunjukkan  $H_a$  pada penelitian ini diterima dan  $H_o$  ditolak. Analisa mengenai konsentrasi menunjukkan perbedaan yang signifikan pada tiap konsentrasi artinya Tidak diragukan lagi bahwa konsentrasi tertinggi merupakan konsentrasi terbaik yang dapat meningkatkan detak jantung pada tikus.

### b. Penelitian RAK dan RAL

Anova dapat dibagi menjadi dua yaitu anova one way dan anova two way. selain itu anova dapat bersifat rancangan acak kelompok (RAK) dan rancangan acak lengkap (RAL). Rancangan acak kelompok didasarkan keterbatasan kemampuan maupun sumberdaya yang ada sehingga penelitian terpaksa dilakukan dengan cara bertahap. Sedangkan rancangan acak lengkap merupakan penelitian yang dilakukan secara keseluruhan sekaligus tanpa adanya gelombang perlakuan. Ketika penelitian bersifat rak

maka harus dipertimbangkan pula nilai ulangan yang dilakukan dan harus tidak berbeda nyata.

### c. Ringkasan materi

Uji Anova merupakan ujian digunakan untuk mengidentifikasi perbedaan sampel yang berjumlah lebih dari 2. Uji Anova yaitu distribusi dan data tersebar secara homogen. Uji Anova dikelompokkan menjadi dua yaitu uji Anova one way dan uji Anova two way. Apa yang dilakukan dengan cara rancangan acak kelompok harus ikut menganalisis nilai ulangan yang ada.

### d. Latihan soal

1. Seorang ATLM meneliti perbedaan perasan daun sirih dengan perasan daun pepaya untuk membunuh jentik nyamuk *aedes aegypti*. Penelitian menggunakan 4 konsentrasi dan 6 ulangan, hasil penelitian adalah

Bahan	ulangan	Mortalitas nyamuk			
		20%	40%	60%	80%
perasan daun sirih	1	10	14	17	15
	2	11	20	12	13
	3	14	12	13	16
	4	13	18	16	15
	5	5	11	14	13
	6	10	10	12	13
perasan daun pepaya	1	15	12	15	22
	2	8	10	17	21
	3	15	10	18	12
	4	8	12	13	15
	5	6	13	19	14
	6	14	14	12	20

Tentukan

1. Apakah ada perbedaan tingkat mortalitas aedest aegepty pada kedua perasan daun tersebut?
  2. Berapakah kosentrasi optimal tingkat mortalitas aedest aegepty pada penelitian tersebut?
2. Peneliti melakukan uji coba penelitian dengan judul “Pengaruh Pemberian N–Acetylcysteine terhadap Kadar SGOT pada Tikus Wistar yang Diberi Parasetamol dengan beberapa kosentrasi” hasil adalah sebagai berikut

ulangan	Nilai SGOT			
	20%	40%	60%	80%
1	44	20	38	23
2	10	17	16	33
3	49	6	15	45
4	19	7	27	13
5	32	17	8	43
6	44	15	19	29

Tentukan

1. Apakah ada perbedaan tingkat nilai SGOT pada perlakuan tersebut?
2. Berapakah kosentrasi terbaik pada perlakuan tersebut

### e. Daftar Pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. *Pokok-pokok Materi Statistik 2 (Statistik Inferensif)*. Jakarta: Bumi Aksara.
3. Lemeshow, Stanley et all, 1997. *Besar Sampel penelitian Kesehatan*, UGM Press,

4. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
6. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
7. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
8. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
9. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
10. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## Bab IX. Uji perbedaan mean lebih dari 2 sampel dengan statistic non parametrik

---

### Tujuan

Mahasiswa memahami dan dapat melakukan uji Perbedaan lebih dari 2 sample dengan statistic non parametric dengan tepat

---

### Materi

---

#### 1. Clustal wallis

---

##### a. Clustal Wallis

Syarat dilakukannya uji parametrik terutama untuk sampel lebih dari 3 adalah data terdistribusi secara normal. Namun demikian seringkali kita temukan data tidak berdistribusi normal dan tidak memungkinkan untuk dilakukan transformasi maka dengan terpaksa uji data tersebut dilakukan menggunakan uji non parametrik. Alternatif apabila data seharusnya dilakukan menggunakan Anova akan tetapi data tidak terdistribusi secara normal maka uji yang dapat digunakan adalah clustal wallis. Rumus yang digunakan dalam uji clustal wallis adalah

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

Di mana

N = jumlah sampel

R<sub>i</sub> = jumlah peringkat pada kelompok i

n<sub>i</sub> = jumlah sampel pada kelompok i

untuk mamahamkan uji beda Seorang peneliti ingin mengetahui apakah terdapat perbedaan kandungan timbal darah pada penduduk yang dekat dengan pabrik semen dengan yang yang jauh. Kategori dibedakan menjadi dekat (<100 meter), sedang (<500 meter), dan jauh (>500 meter dan <1 km). Hasil analisis tinggal terdapat pada tabel berikut:

dekat	sedang	jauh
12.3	6.1	1.8
13.6	10.9	5.5
12.5	6	3.1
13.8	8.1	3.6
9	8.6	1.9
9.2	7.8	2.3
9	8.2	1.5
10.3	7.8	1.4
11	7.2	2.9
14.9	12.2	1.4
11.9	9.7	4.4
13.1	7.7	5.9
12.3	11.9	4
2	1.6	5.2
14.6		2.8
9.9		4.2
15.1		
13.6		

Bedasarkan data tersebut dapat dianalisis sebagai berikut

### **Rumusan Masalah**

Apakah terdapat perbedaan kandungan timbal darah pada penduduk yang dekat dengan pabrik semen dengan yang yang jauh. Kategori dibedakan menjadi dekat (<100 meter), sedang (<500 meter), dan jauh (>500 meter dan <1 km)?

### **Tujuan**

Untuk mengetahui Apakah terdapat perbedaan kandungan timbal darah pada penduduk yang dekat dengan pabrik semen dengan yang yang jauh. Kategori dibedakan menjadi dekat (<100 meter), sedang (<500 meter), dan jauh (>500 meter dan <1 km)

### **Hipotesis**

Ho = tidak terdapat perbedaan kandungan timbal darah pada penduduk yang dekat dengan pabrik semen dengan yang yang jauh. Kategori dibedakan menjadi dekat (<100 meter), sedang (<500 meter), dan jauh (>500 meter dan <1 km)

Ha = Terdapat perbedaan kandungan timbal darah pada penduduk yang dekat dengan pabrik semen dengan yang yang jauh. Kategori dibedakan menjadi dekat (<100 meter), sedang (<500 meter), dan jauh (>500 meter dan <1 km)

### **Taraf kepercayaan**

Error 5%

Untuk menjawab soal yang ada maka data di masukkan seperti menganalisis dengan menggunakan independent T Test

	VAR00001	VAR00002
1	177.00	1.00
2	171.00	1.00
3	151.00	1.00
4	155.00	1.00
5	156.00	1.00
6	135.00	1.00
7	144.00	1.00
8	173.00	1.00
9	110.00	1.00
10	151.00	1.00
11	169.00	1.00
12	141.00	1.00
13	145.00	1.00
14	144.00	1.00
⋮		
22	131.00	2.00
23	142.00	2.00
24	146.00	2.00
25	126.00	2.00
26	150.00	2.00
27	124.00	2.00
28	147.00	2.00

Sehingga dalam penghitungan manual didapatkan

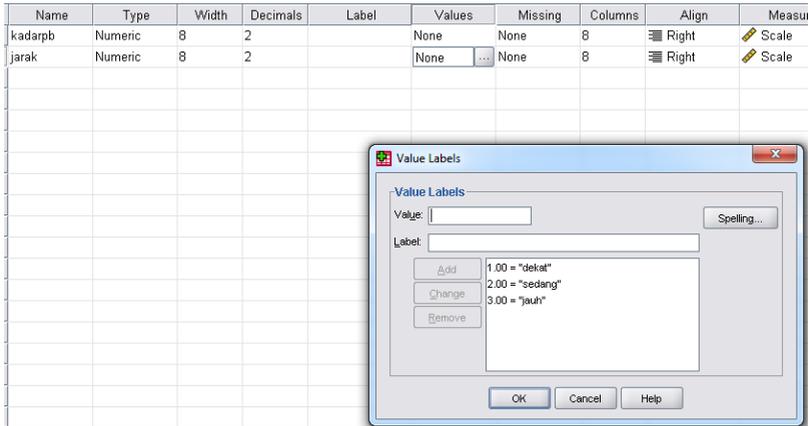
no	kadar timbal			peringkat		
	dekat	sedang	jauh	dekat	sedang	jauh
1	12.3	6.1	1.8	39	20	5
2	13.6	10.9	5.5	43	34	17
3	12.5	6	3.1	41	19	11
4	13.8	8.1	3.6	45	25	12
5	9	8.6	1.9	28	27	6
6	9.2	7.8	2.3	30	23	8
7	9	8.2	1.5	29	26	3
8	10.3	7.8	1.4	33	24	1
9	11	7.2	2.9	35	21	10
10	14.9	12.2	1.4	47	38	2
11	11.9	9.7	4.4	36	31	15
12	13.1	7.7	5.9	42	22	18
13	12.3	11.9	4	40	37	13
14	2	1.6	5.2	7	4	16
15	14.6		2.8	46		9
16	9.9		4.2	32		14
17	15.1			48		
18	13.6			44		
			Ri	665	351	160
			Ri <sup>2</sup>	442225	123201	25600

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

$$H = \frac{12}{45(45+1)} \times \left( \frac{442225}{18} + \frac{123201}{14} + \frac{25600}{16} \right) - 3(45+1) =$$

H = 26.71378. nilai H ini selanjutnya di bandingkan dengan H table

Sedangkan penyelesaian dengan menggunakan SPSS maka dapat dilakukan dengan mengisi kolom seperti biasanya



Kemudian dilakukan uji normalitas dan hasil dari normalitas tersebut menunjukkan ada yang tidak terdistribusi normal

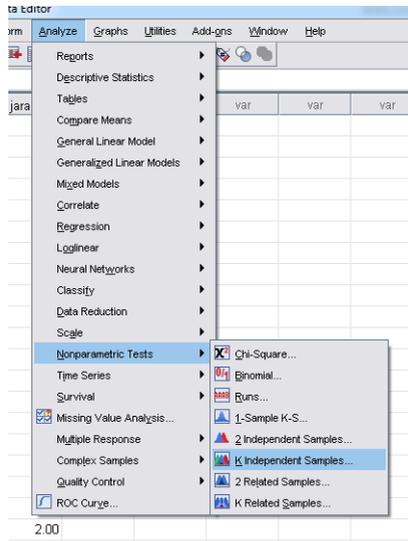
#### Tests of Normality

jarak	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
kadarpb dekat	.154	18	.200 <sup>*</sup>	.851	18	.009
sedang	.151	14	.200 <sup>*</sup>	.923	14	.245
jauh	.127	16	.200 <sup>*</sup>	.931	16	.252

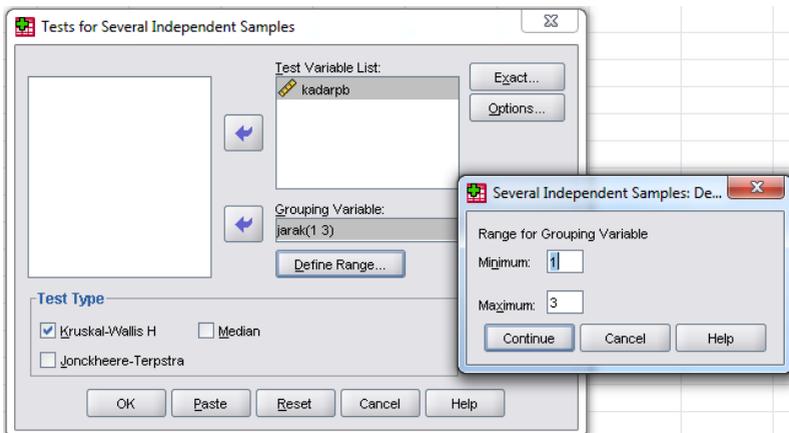
a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Kemudian klik analyses > Non parametric test > K independent sample



Masukkan pengaturan pada test value dan Grouping variable



## Kruskal-Wallis

Ranks

	jarak	N	Mean Rank
kadarpb	dekat	18	36.97
	sedang	14	25.04
	jauh	16	10.00
	Total	48	

Test Statistics<sup>a,b</sup>

	kadarpb
Chi-Square	31.480
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: jarak

Berdasarkan hasil yang didapatkan bahwa terdapat perbedaan yang signifikan pada kadar timbal darah yang di dapat.

### b. Ringkasan materi

Uji clustal wallis digunakan untuk membedakan lebih dari dua sampel pada sampel yang tidak terdistribusi normal. Secara sederhana Clustal wallis merupakan uji alternative ketika uji anova tidak memungkinkan untuk dilakukan

### c. Latihan Soal

Berikut adalah kebutuhan kalori pada masing masing pekerja.

no	jenis kelamin	jumlah kal yang dibutuhkan
1	petani	1020
2	pns	1056
3	wiraswasta	1693
4	petani	1839
5	pns	1765
6	wiraswasta	2106
7	petani	2039
8	pns	1696
9	wiraswasta	2133
10	petani	1898
11	pns	2218
12	wiraswasta	1869
13	petani	1869
14	pns	1801
15	wiraswasta	2031
16	petani	1772

no	jenis kelamin	jumlah kal yang dibutuhkan
17	pns	1657
18	wiraswasta	2103
19	petani	1745
20	pns	1858
21	wiraswasta	1601
22	petani	2108
23	pns	1626
24	wiraswasta	1949
25	petani	2231
26	pns	1974
27	wiraswasta	1589
28	petani	1859
29	pns	1744
30	wiraswasta	2047
31	petani	2239
32	pns	1808

Tentukan kebenaran hipotesis  $H_0$ !

### d. Daftar pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.

3. Lemeshow, Stanley et al, 1997. Besar Sampel penelitian Kesehatan, UGM Press,
4. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
5. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
6. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
7. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
8. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
9. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
10. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## Bab X. Uji Korelasi dan Regresi

---

### Tujuan

Mahasiswa memahami dan dapat melakukan uji Perbedaan lebih dari 2 sample dengan statistic parametric dengan tepat

---

### Materi

---

1. Koefisien Korelasi
  2. Regresi
- 

Di dalam penelitian kesehatan kita sering saling jumpai suatu peubah atau parameter yang saling berkaitan antara satu dengan yang lain. sebagai contoh, dalam keadaan normal nilai dari leukosit akan berbanding lurus dengan jumlah virus yang menyerang. Kata-kata bahwa semakin tinggi nilai  $x$  maka akan diikuti oleh kenaikan nilai variabel  $y$ . Atau mungkin kita pernah menemukan bahwa suatu variabel apabila semakin naik maka akan diikuti oleh variabel yang lain yang menunjukkan angka semakin turun. Perubahan kedua variabel tersebut menunjukkan adanya suatu hubungan dan dapat dikatakan berkorelasi. Korelasi positif menunjukkan bahwa suatu variabel akan diikuti oleh kenaikan variabel yang lain. Terangkan korelasi negatif menunjukkan bahwa

ketika variabel yang satu menunjukkan kenaikan maka variabel yang lain akan semakin menurun.

Kesalahan yang biasanya umum terjadi ketika menanggapi fenomena tersebut yaitu variabel yang satu dianggap sebagai penyebab kenaikan atau penurunan variabel yang lain. Sebagai contoh korelasi yang erat kaitannya dengan banyaknya perokok yang menderita kanker paru-paru. Namun yang perlu dicermati adalah mungkin faktor tersebut merupakan salah satu dari faktor dan mungkin tidak banyak faktor yang mempengaruhi kejadian yang lain korelasi tidak harus merupakan suatu sebab akibat. Banyak faktor yang harus dipertimbangkan untuk memutuskan bahwa hal tersebut merupakan satu-satunya faktor yang lain.

Sampai saat ini pasti Kalian sering mendengar istilah korelasi dan regresi. Hal yang menjadikan korelasi dan regresi ini sangat erat adalah memang pada penelitian Penelitian terdahulu perhitungan matematik diantara keduanya menunjukkan kemiripan yang tinggi. An-nur yang menyebabkan 2 kata ini menjadi rancu dan seringkali mahasiswa baru belajar statistik menjadi bingung adalah memang bahwa buku-buku memisahkan antara korelasi dan regresi. Salah satu tujuan dari topik Pada bab ini adalah untuk menjelaskan perbedaan korelasi dan regresi. Beberapa konsep penting yang harus diperhatikan didalam uji korelasi

- Nilai korelasi berkisar  $-1$  s.d.  $1$
- Korelasi bersifat simetrik
- Korelasi bebas dari origin dan skala

$$P = a_1 + b_1X_1$$

$$Q = a_2 + b_2X_2$$

Dimana  $b_1 > 1$ ,  $b_2 > 1$ ,  $a_1$  dan  $a_2$  konstanta maka korelasi P dgn Q akan sama dengan korelasi  $X_1$  dgn  $X_2$

- Jika X dan Y saling bebas maka korelasi akan bernilai 0

- Meskipun korelasi mengukur derajat hubungan, tetapi bukan alat uji kausal.

Segala sesuatu analisis yang mengidentifikasi "keeratan hubungan" maka akan diidentifikasi sebagai sebuah korelasi. Pangkat di dalam rekreasi di situ menghitung ketergantungan suatu variabel dengan variabel yang lain variabel  $x$  gantung dengan variabel  $y$ .

### a. Koefisien Korelasi

Korelasi dinyatakan dengan koefisien ( $r$ ) dan merentang  $-1$  sampai  $+1$ . Koefisien  $1$ , dengan tanda  $+$  atau  $-$ , menunjukkan korelasi sempurna antara dua variabel. Sebaliknya, koefisien nol berarti tidak ada korelasi sama sekali. keragaman dalam derajat korelasi dinyatakan oleh koefisien yang merentang dari  $0$  sampai  $1$  dan dari  $-1$  sampai  $0$ .

Kita dapat merajah deretan titik yang masing-masing mewakili besaran variabel  $X$  dan  $Y$  (koordinat, yang terdiri atas pasangan nilai variabel  $X$  dan  $Y$ ) dan memperoleh berbagai pola; beberapa diperlihatkan oleh Gambar 13.1. Diagram (a) memperlihatkan deretan titik yang tidak terletak pada satu garis lurus. Ini menyatakan korelasi yang kurang sempurna, tetapi memperlihatkan kecenderungan umum dalam arah positif. Jika tinggi badan dan bobot dalam sampel yang berupa orang dewasa dikorelasikan, mungkin sekali diperoleh pola serupa dengan pola a. Pola korelasi tidak sempurna yang serupa diperlihatkan oleh pola b, tetapi dalam hal ini kemiringan itu negatif, menunjukkan korelasi negatif. Pola serupa dengan b dapat diperoleh dari hasil perajahan kejadian 'banyaknya gigi berlubang' lawan 'taraf fluorisasi' pada anak-anak. Diagram (c) memperlihatkan korelasi negatif yang sempurna ( $-1$ ) yang dapat diperoleh dari perajahan percepatan

lawan massa dalam menggunakan hukum kedua Newton,  $F = ma$ . Diagram (d) memperlihatkan hamburan acak titik yang dapat diharapkan dari perajahan Y lawan X bila kedua variabel itu sama sekali tidak berkaitan.

Apakah korelasi yang 'baik' itu? Ini jelas bergantung pada apa yang dilakukan peneliti atau pada apa yang diharapkannya dari uji kajinya. Jika ia berharap uji kajinya menunjukkan tidak terdapat kaitan antara dua variabel, maka koefisien sebesar nol akan sangat menggembirakan. Jika sebaliknya, ia mengharapkan akan terlihat kaitan yang erat, maka harga  $r$  yang mendekati  $\pm 1$  akan di pandang sebagai hasil yang optimum. Sebagaimana lazimnya, sistem yang hidup tidak membantu dengan menghasilkan bilangan bulat; keragaman lebih mungkin terjadi daripada perkecualian. Karena itu, dalam kebanyakan hal, kemungkinan besar biologiwan harus puas dengan koefisien yang kurang sempurna.

Jika koefisien korelasi( $r$ ) dikuadratkan, didapat koefisien penentu( $r^2$ ). Nilai ini boleh dipakai sebagai taksiran untuk kekuatan kaitan antara dua variabel yang berkorelasi. Koefisien penentu khususnya menaksirkan persentase keragaman X yang berkaitan dengan ( atau "diterangkan oleh") keragaman Y – atau sebaliknya. Sebagai contoh, jika korelasi terok antara dua variabel (seperti klorofil a dan biomassa) besarnya 0,50, koefisien penentu sebesar 0,25 diperoleh dari mengkuadratkan korelasi 0,50 tersebut. Ini memberi petunjuk bahwa 25% dari keragaman salah satu dari kedua variabel berkaitan dengan atau 'diterangkan oleh' keragaman variabel yang lain. Kita tidak dapat mengatakan bahwa yang mana 'menerangkan' yang mana karena keduanya dipandang sebagai variabel tak terbatas.

Tentu saja, bila dikatakan bahwa 25% keragaman salah satu variabel 'diterangkan' oleh variabel lainnya, kita mengambil titik pandang optimis bahwa satu gelas berisi seperempat penuh.

Seorang pesimis sebaliknya mungkin mengatakan bahwa gelas itu tiga perempat kosong! Hal kita ini nadanya tidak sekeras bila kita mengatakan bahwa 75% keragaman salah satu variabel tidak dijelaskan oleh keragaman variabel lainnya.

Jika kita menyimak perhitungannya, tampak bahwa nilai  $r^2$  bertambah dengan cepat jika  $r$  bertambah. Misalnya, bila  $r = 0,10$ ,  $r^2 = 0,01$ ; bila  $r = 0,20$ ,  $r^2 = 0,04$ ; bila  $r = 0,60$ ,  $r^2 = 0,36$ ; bila  $r = 0,90$ ,  $r^2 = 0,81$ , dan seterusnya. Karena itu kita harus memperoleh nilai  $r$  yang cukup besar untuk membenarkan bahwa kedua variabel itu mempunyai derajat kaitan yang tinggi.

### **b. Pengetesan Signifikansi**

Bilamana nilai  $r$  sudah ditemukan, kita dapat menarik beberapa kesimpulan dari nilai  $r$  yang kita kenakan pada populasi. Hal ini tentu sangat tergantung pada sifat sampel data. Bilamana sampel yang diambil untuk penelitian adalah sampel random, maka kita dapat menarik kesimpulan tentang populasi atas dasar bahan-bahan dari sampel itu. Akan tetapi bilamana keadaannya adalah sebaliknya, yaitu bilamana kita meneliti suatu sampel yang tidak diambil secara random, maka kita harus puas dengan kesimpulan-kesimpulan yang hanya terbatas pada sampel itu sendiri, tanpa ada maksud untuk mengadakan generalisasi pada populasi. Dalam hal semacam ini seorang peneliti tidak ubahnya seorang wartawan, yaitu menyajikan fakta-fakta. Bilamana hendak diadakan simpulan secara ilmiah, simpulan itu lebih bersifat indikatif daripada konklusif.

Suatu pertanyaan yang ingin dijawab bilamana kita ingin menarik kesimpulan dari sampel untuk populasi adalah apakah mungkin kita memperoleh  $r$  sebesar sekian atau sekian dari suatu sampel random kalau  $r$  dalam populasi adalah nol. Ini adalah semacam mengetes hipotesa nihil, yaitu bahwa tidak ada

perbedaan antara nilai  $r$  yang diperoleh dari sampel dengan nilai  $r$  dari populasi, atas dasar pokok pikiran bahwa nilai  $r$  dari populasi adalah nol.

Untuk menyelesaikan soal itu tidaklah sukar, karena di sini kita tidak memerlukan rumus-rumus lain. Dengan nilai  $r$  yang telah kita peroleh, kita dapat secara langsung melihat tabel korelasi yang disediakan dalam apendiks di bagian belakang buku ini untuk mengetes apakah nilai  $r$  yang kita peroleh itu berarti atau tidak (signifikan atau nonsignifikan) atas dasar taraf kepercayaan sekian atau sekian persen. Tabel korelasi itu mencantumkan batas-batas nilai  $r$  yang signifikan (berarti) pada taraf-taraf signifikansi tertentu a). Bilamana nilai  $r$  yang kita peroleh sama dengan atau lebih besar daripada nilai dalam tabel  $r$  itu, maka nilai  $r$  yang kita peroleh itu signifikan. Dengan nilai  $r$  yang signifikan kita akan menolak hipotesa yang mengatakan bahwa korelasi antara  $X$  dan  $Y$  dalam populasi adalah nol, atas dasar taraf signifikansi yang kita gunakan (yaitu 5% atau 1%). Ini sama halnya dengan mengatakan bahwa tidak mungkin kita memperoleh  $r$  (sebesar sekian atau sekian) itu dari sampel yang kita ambil secara random kalau nilai  $r$  dalam populasi adalah nol. Sebabnya karena batas-batas nilai yang disebabkan oleh kesalahan sampling sudah dilewati. Perhatikan baik-baik tabel  $r$  itu. Bilangan-bilangan yang membatasi signifikan tidaknya sesuatu nilai  $r$  sangat tergantung kepada  $N$  yang diselidiki dalam sampel. Makin besar  $N$  makin rendah batas signifikansi.

Untuk memahami secara kongkret apa yang baru dibicarakan itu baiklah kita ambil sebagai contoh nilai  $r$  yang kita peroleh dari tabel 44. Nilai yang kita peroleh adalah 0,830. Dengan nilai  $r$  itu kita hendak mengadakan pengetesan apakah nilai itu signifikan ataukah tidak atas dasar taraf signifikansi 5%. Jumlah subyek atau  $N$  yang diselidiki ada 30. Dengan melihat  $N = 30$  dalam kolom  $N$  dan membacanya ke kanan dalam kolom taraf signifikansi

5% dalam tabel  $r$  akan kita temukan bilangan 0,361. Bilangan ini menunjukkan bilangan batas signifikansi. Oleh karena nilai  $r$  yang kita peroleh, yaitu 0,830 berada jauh di atas batas signifikansinya, yaitu 0,361, maka nilai  $r$  itu kita katakan signifikan. Dengan demikian kita menolak hipotesa (nihil) yang mengatakan bahwa nilai  $r$  dalam populasi adalah nol (tidak ada korelasi antara  $X$  dan  $Y$ , atau tegasnya tidak ada korelasi antara pengetahuan umum dan matematik), sekiranya kita mengambil 30 orang dalam sampel penyelidikan itu secara random.

### C. Syarat-Syarat untuk Pengetesan Nilai $r$

Agar kesimpulan yang ditarik tidak menyimpang dari kebenaran yang seharusnya ditarik, maka syarat-syarat berikut ini perlu dipenuhi :

- (1) Sampel yang digunakan dalam penyelidikan harus sampel yang diambil secara random dari populasi terhadap mana kesimpulan penyelidikan hendak kita kenakan.
- (2) Hubungan antara variabel  $x$  dengan variabel  $y$  merupakan hubungan garis lurus atau hubungan linier.
- (3) Bentuk distribusi variabel  $x$  dan variabel  $y$  dalam populasi adalah atau mendekati distribusi normal.

Dari ketiga syarat itu, karena syarat yang ketiga sangat sukar diketahui a), maka bilamana kedua syarat yang pertama telah dipenuhi biasanya orang dapat menarik kesimpulan-kesimpulan dari sampel populasi tanpa kesalahan-kesalahan yang berarti.

#### d. Apa Gunanya $r$ ?

Bilamana ada alat yang dapat digunakan untuk landasan mengadakan ramalan atau prediksi, salah satu dari alat itu adalah koefisien korelasi dari Pearson. Mengadakan prediksi ini pada hakekatnya selalu kita jalankan dalam kehidupan sehari-hari. Sekolah-sekolah menerima murid-murid baru atas dasar nilai-nilai sekolah. Apakah artinya ini? Ini berarti bahwa sekolah-sekolah tersebut membuat ramalan bahwa bilamana murid-murid memperoleh angka-angka yang baik dalam ujian, maka mereka akan sukses dalam sekolah itu. Ini tentu saja merupakan tindakan yang jauh lebih bijaksana daripada menyerahkan diri pada hal-hal yang terjadi secara kebetulan, atau meraba-raba tanpa dasar-dasar yang tegas.

Seperti telah dikatakan di muka, bilamanana kita memperoleh korelasi yang positif antara dua variabel, ini berarti bahwa kenaikan nilai-nilai dalam variabel yang satu secara proporsional akan diikuti oleh kenaikan pada nilai-nilai variabel lainnya, sehingga dari nilai variabel yang satu kita sampai pada taraf-taraf tertentu dapat meramalkan nilai variabel lainnya yang belum kita ketahui. Jadi kalau misalnya ada korelasi yang positif antara banyaknya minuman keras yang diminum oleh suami dengan jumlah perselisihan dalam rumah tangga, maka makin banyak suami minum alkohol dapat diramalkan akan terjadi makin banyak perselisihan dalam rumah tangganya.

Akan tetapi bilamana variabel  $X$  dan variabel  $Y$  mempunyai hubungan yang negatif, maka keadaan yang sebaliknya yang akan kita jumpai. Nilai  $r$  yang negatif menunjukkan bahwa kenaikan nilai variabel yang satu akan diikuti secara proporsional oleh turunnya nilai variabel lainnya, sehingga kita dapat meramalkan bahwa orang-orang yang mempunyai nilai yang besar dalam variabel  $X$

akan menunjukkan nilai yang kecil dalam variabel Y. Korelasi negatif yang hampir sempurna pernah diketemukan oleh TRAVIS antara variabel berat badan dan kemampuan mempertahankan keseimbangan badan. Dengan demikian kita dapat meramalkan bahwa makin berat badan orang, akan makin kecil kemampuannya untuk mempertahankan keseimbangan badannya. Dan sebaliknya, makin kurang berat badan orang akan makin besar kemampuannya untuk mempertahankan keseimbangan badannya.

Bilamana korelasi antara variabel X dan variabel Y sama dengan nol, maka tidak ada ramalan yang dapat dibuat. Seperti telah dikatakan di muka, koefisien korelasi sama dengan nol berarti tidak ada korelasi. Jika antara dua variabel X dan Y tidak ada korelasi maka sesuatu nilai dalam variabel X mungkin disertai oleh nilai variabel Y yang tinggi atau rendah, sehingga tidak dapat diramalkan apakah kenaikan dalam variabel X akan diikuti oleh kenaikan atau penurunan nilai variabel Y. rujukan lain mencoba untuk mengkatagoikan nilai r kepada beberapa kelompok

nilai r	interpretasi
0,00-< 0,20	Hubungan sangat lemah (diabaikan, dianggap tidak ada hubungan)
≥ 2,00 - < 4,00	Hubungan rendah
≥ 4,00 - < 6,00	Hubungan sedang atau cukup
≥ 6,00 - < 8,00	Hubungan kuat atau tingg
≥ 8,00 - 1	Hubungan sangat kuat atau tinggi

### e. Pembagian uji Korelasi

Korelasi pada pembagiannya secara sederhana dapat di bagi menjadi

- Korelasi Product Moment (*Pearson*) (*Interval/rasio*)
- Korelasi Rank Spearman (*Ordinal/nominal/Interval*)
- Korelasi Kendall tau b (NP dari individu yang sama)

## f. Korelasi Pearson

Korelasi pearson digunakan untuk mengidentifikasi hubungan dua variabel itu digunakan juga untuk mengetahui Kuat lemahnya hubungan kedua variabel tersebut. Korelasi pearson harus bersifat interval atau rasio tetap bersifat terdistribusi normal. Rumus dari korelasi pearson adalah

$$r_{xy} = \frac{N \cdot \sum X.Y - \sum X \cdot \sum Y}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

Untuk memahami makna dari korelasi maka di sini akan di kita bahas sebuah contoh Seorang ATLM ingin mengetahui apakah penundaan pemeriksaan gula darah berkorelasi dengan penurunan penghitungan gula darah berkorelasi dengan kadar leukosit . Oleh sebab itu peneliti menggunakan sampel darah yang dilakukan pemeriksaan secara berkala

kadar gula	kadar leukosit
165	9567
165	9223
169	9123
171	8689
174	8659
176	8621
177	8605
180	8375
180	8372
181	8033
182	8032
187	7857
191	7380
193	7254

194	7242
197	6582
198	6357
198	5819
200	5473
200	5121

---

### **Rumusan Masalah**

Apakah penundaan pemeriksaan gula darah berkorelasi dengan penurunan penghitungan gula darah berkorelasi dengan kadar leukosit?

### **Tujuan**

Tujuan dari penelitian ini adalah Apakah penundaan pemeriksaan gula darah berkorelasi dengan penurunan penghitungan gula darah berkorelasi dengan kadar leukosit.

### **Hipotesis**

Ho= penundaan pemeriksaan gula darah tidak berkorelasi dengan penurunan penghitungan gula darah berkorelasi dengan kadar leukosit

Ha= penundaan pemeriksaan gula darah tidak berkorelasi dengan penurunan penghitungan gula darah berkorelasi dengan kadar leukosit

## Analisis penghitungan

Maka dari data tersebut terlebih dahulu dilakukan

kadar gula	kadar leukosit	X <sup>2</sup>	Y <sup>2</sup>	XY	
165	9567	27225	91527489	1578555	
165	9223	27225	85063729	1521795	
169	9123	28561	83229129	1541787	
171	8689	29241	75498721	1485819	
174	8659	30276	74978281	1506666	
176	8621	30976	74321641	1517296	
177	8605	31329	74046025	1523085	
180	8375	32400	70140625	1507500	
180	8372	32400	70090384	1506960	
181	8033	32761	64529089	1453973	
182	8032	33124	64513024	1461824	
187	7857	34969	61732449	1469259	
191	7380	36481	54464400	1409580	
193	7254	37249	52620516	1400022	
194	7242	37636	52446564	1404948	
197	6582	38809	43322724	1296654	
198	6357	39204	40411449	1258686	
198	5819	39204	33860761	1152162	
200	5473	40000	29953729	1094600	
200	5121	40000	26224641	1024200	
sum	3678	154384	679070	1222975370	28115371

$$r_{xy} = \frac{N \cdot \sum XY - \sum X \cdot \sum Y}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

$$r_{xy} = \left[ \frac{20 \cdot 28115371 - 3678 \times 394}{\sqrt{(20 \times 679070 - 13527684) \times (20 \times 679070 - 23834419456)}} \right]$$

$$r_{xy} = -0.95$$

guladarah	leukosit
165.00	9567.00
165.00	9223.00
169.00	9123.00
171.00	8689.00
174.00	8659.00
176.00	8621.00
177.00	8605.00
180.00	8375.00
180.00	8372.00
181.00	8033.00
182.00	8032.00
187.00	7857.00
191.00	7380.00
193.00	7254.00
194.00	7242.00
197.00	6582.00
198.00	6357.00
198.00	5819.00
200.00	5473.00
200.00	5121.00

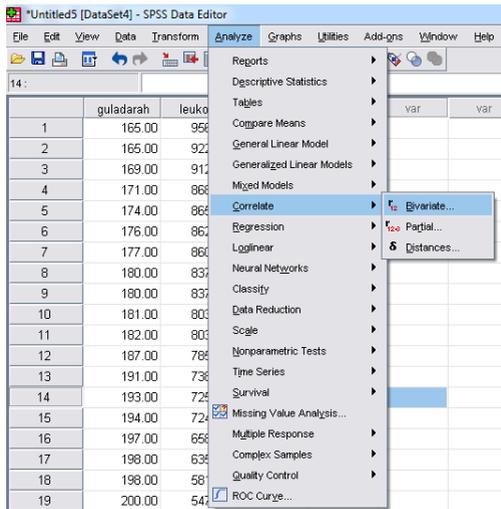
### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
guladarah	.128	20	.200 <sup>*</sup>	.926	20	.129
leukosit	.146	20	.200 <sup>*</sup>	.935	20	.193

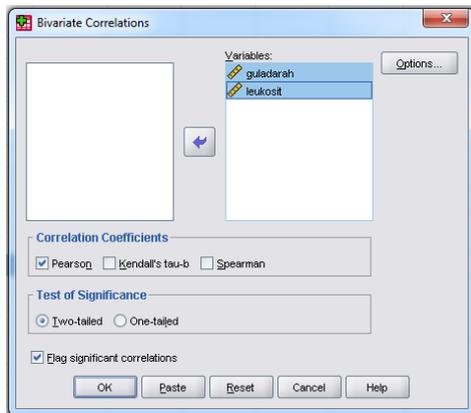
a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Kemudian masuk ke analyze > Correlate > Bivariate



Kemudian klik



### Correlations

		guladarah	leukosit
guladarah	Pearson Correlation	1	-.952**
	Sig. (2-tailed)		.000
	N	20	20
leukosit	Pearson Correlation	-.952**	1
	Sig. (2-tailed)	.000	
	N	20	20

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Intrepretasi hasil yang diperoleh maka nilai signifikan 0.000 artinya bahwa terdapat korelasi atau hubungan yang signifikan antara kedua parameter. Nilai pearson -0.95 artinya terdapat hubungan yang bertolak belakang Apabila suatu parameter menunjukkan kenaikan maka akan diikuti oleh penurunan pada parameter lain begitu pula dengan sebaliknya. Nilai 0.95 menunjukkan bahwa hubungan kedua Parameter tersebut sangat erat. Masih ingat kah anda bahwa semakin mendekati angka 1 atau min 1 maka hubungan kedua parameter semakin kuat.

### g. Korelasi Spearman

Korelasi spearman atau rho merupakan uji yang sering digunakan untuk identifikasi masalah penelitian di bidang kesehatan. Kegunaannya sama dengan pearson yaitu untuk menguji signifikansi hubungan dua variabel dan selain itu juga untuk mengetahui Kuat lemahnya hubungan variabel tersebut prasyarat dari uji ini adalah Data berskala ordinal, interval atau rasio yang tidak terdistribusi normal. Rumus yang digunakan untuk menentukan jenis korelasi ini adalah

$$rho_{xy} = 1 - \frac{6 \cdot \sum d^2}{N \cdot (N^2 - 1)}$$

Keterangan :

- $rho_{xy}$  = Koefisien Korelasi Tata Jenjang Spearman  
 $d$  = beda ranking variabel pertama dengan variabel kedua  
 $N$  = banyaknya sampel

Berdasar hasil penelitian hubungan tingkat pengetahuan kelompok masyarakat dengan insiden daire didapatkan data sebagai berikut:

DESA	RERATA PENGETAHUAN MASYARAKAT	INSIDEN DIARE (%)
Mojo	Baik	8
Koto	Sedang	13
Mrico	Sangat baik sekali	5
Sikep	Rendah	16
Rejo	Baik	10
Gedang	Sedang	14
Suka	Rendah	14
Ganting	Baik	8
Keboan	Sangat rendah sekali	23
Kliwon	Sangat baik	8
Paci	Sangat rendah	20
Soma	Baik	9
Alang	Rendah	14
Kriyo	Sangat rendah	20

Selidikilah dengan  $\alpha = 5\%$ , apakah ada hubungan negatif antara tingkat pengetahuan masyarakat dengan insiden daire?

Penyelesaian :

a. Hipotesis

Ho :  $\rho = 0$  ~ tidak ada hubungan antara pengetahuan dengan insiden diare

Ha :  $\rho < 0$  ~ semakin tinggi pengetahuan diikuti dengan semakin rendah insiden diare

b. Level signifikansi

$$\alpha = 5\% = 0,05$$

No	DESA	RERATA PENGETAHUAN MASYARAKAT	INSIDEN DIARE (%)	RANK		D	D <sup>2</sup>
				VAR. I	VAR. II		
1	Aryo	Baik	8	4,50	12,00	7,50	56,25
2	Koto	Sedang	13	7,50	8,00	0,50	0,25
3	Mrico	Sangat baik sekali	5	1,00	14,00	13,00	169,00
4	Sikep	Rendah	16	10,00	4,00	6,00	36,00
5	Rejo	Baik	10	4,50	9,00	4,50	20,25
6	Gedang	Sedang	14	7,50	6,00	1,50	2,25
7	Suka	Rendah	14	10,00	6,00	4,00	16,00
8	Ganting	Baik	8	4,50	12,00	7,50	56,25
9	Keboan	Sangat rendah sekali	23	14,00	1,00	13,00	169,00
10	Kliwon	Sangat baik	8	2,00	12,00	10,00	100,00
11	Paci	Sangat rendah	20	12,50	2,50	10,00	100,00
12	Soma	Baik	9	4,50	10,00	5,50	30,25
13	Alang	Rendah	14	10,00	6,00	4,00	16,00
14	Kriyo	Sangat rendah	20	12,50	2,50	10,00	100,00
JUMLAH							871,50

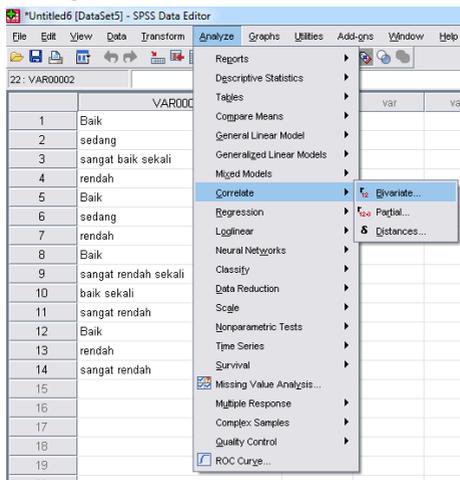
- Sangat baik sekali : 7
- Sangat baik : 6
- Baik : 5
- Sedang : 4
- Rendah : 3
- Sangat rendah : 2
- Sangat rendah sekali : 1

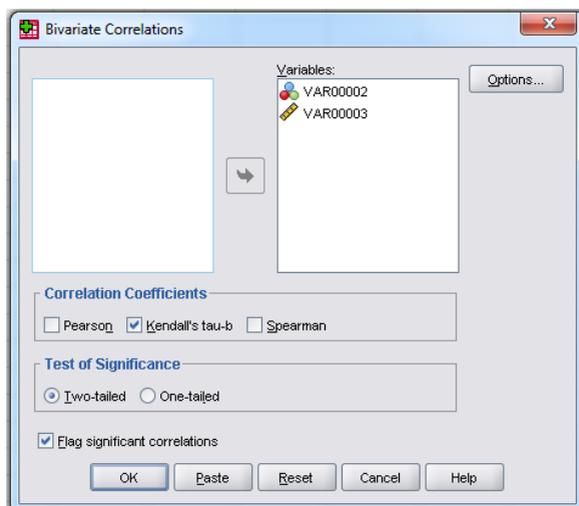
$$rho_{XY} = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

$$rho_{XY} = 1 - \frac{6.871,50}{14.(14^2 - 1)}$$

$$rho_{XY} = -0,92$$

Berdasarkan nilai yang dihasilkan dapat disimpulkan hubungan kedua variable dapat dikatakan sangat kuat. Dengan hubungan negative.





#### Correlations

			VAR00002	VAR00003
Kendall's tau_b	VAR00002	Correlation Coefficient	1.000	-.927**
		Sig. (2-tailed)	.	.000
		N	14	14
	VAR00003	Correlation Coefficient	-.927**	1.000
		Sig. (2-tailed)	.000	.
		N	14	14

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Berdasarkan hasil di atas maka dapat disimpulkan bahwa kedua parameter memiliki hubungan yang erat

## h. Korelasi Kendal tau b

Uji korelasi Kendal tau merupakan pilihan kedua apabila data yang dianalisis tidak memungkinkan untuk dilakukan uji pearson. Meskipun mirip dengan uji spearman namun uji Kendall tau B memiliki keunggulan yaitu apabila data merupakan data berpasangan atau berasal dari satu sampel. Sisi lain dari Kendall topi yaitu tidak membutuhkan normalitas data karena Kendal topi merupakan analisis non parametrik. Apabila di suruh memilih apakah yang lebih baik antara persen dan Kendall tau b. Jawaban yang paling menonjol dan nasional adalah tergantung dari normalitas data. Apabila data merupakan data normal maka uji kendall tau b tentu tidak disarankan sehingga lebih disarankan menggunakan uji parametrik yaitu pearson. Rumus yang digunakan adalah

$$T = \frac{2S}{N(N-1)}$$

S adalah total skor seluruhnya (grand total), yang merupakan jumlah skor urutan kewajaran pasangan data pada salah satu variabel. Jika urutan ranking wajar diberi skor +1, jika urutan ranking tdk wajar diberi skor -1. N adalah banyaknya pasangan ranking. Metode yang digunakan juga relative sederhana akan tetapi tidak mengurangi dari keberartian data. Kita ibaratkan memiliki data berikut

lama penundaan	rank lama penundaan	kadar glukosa darah	rank lama penundaan
30	1	235	4
60	2	230	2
90	3	231	3
120	4	229	1

	1	2	3	4	jumlah
rank lama penundaan	4	2	3	1	
	4	-	-	-	-3
		2	+	-	0
			3	-	1
				1	
					-2

Maka nilai Kendal korelasi adalah

$$T = 2(-2)/4(4-1)$$

$$T = -4/12 = -0.325$$

Sehingga nilai koefisien korelasi adalah -0.325

Sebagai contoh untuk memperdalam pemahaman maka kita identifikasi apakah terdapat hubungan lama penundaan pemeriksaan terhadap kadar glukosa

kadar gula	kadar leukosit
103	9567
165	9223
169	9123
171	8689
174	8659
176	8621
177	8605
180	8375
180	8372
181	8033
182	8032
187	7857
191	7380
193	7254

194	7242
197	6582
198	6357
198	5819
200	5473
200	5121

---

lama penundaan	kadar glukosa darah	lama penundaan	kadar glukosa darah
30	235	330	186
60	231	360	183
90	231	390	181
120	229	420	175
150	227	450	172
180	223	480	171
210	223	510	170
240	222	540	169
270	200	570	163
300	188	600	163

---

	VAR00001	VAR00002
1	30.00	235.00
2	60.00	231.00
3	90.00	231.00
4	120.00	229.00
5	150.00	227.00
6	180.00	223.00
7	210.00	223.00
8	240.00	222.00
9	270.00	200.00
10	300.00	188.00
11	330.00	186.00
12	360.00	183.00
13	390.00	181.00
14	420.00	175.00
15	450.00	172.00
16	480.00	171.00
17	510.00	170.00
18	540.00	169.00
19	570.00	163.00
20	600.00	163.00
21		

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
lamapenundaan	Numeric	8	2		None	None	8	Right	Scale
glukosa	Numeric	8	2		None	None	8	Right	Scale

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
lamapenundaan	.077	20	.200 <sup>*</sup>	.960	20	.551
glukosa	.221	20	.011	.851	20	.006

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

## i. Ringkasan materi

Korelasi dan regresi memiliki hubungan yang sangat erat. Relasi lebih mengedepankan analisis mengenai hubungan dua variabel. Relasi dapat dibagi menjadi parametrik dan nonparametrik. analisa parametrik dapat dilakukan dengan cara korelasi pearson. Korelasi non parametrik dapat dilakukan dengan korelasi spearman dan Kendall tau b .

## j. Latihan Soal

Jawablah pertanyaan berikut dengan benar

1. Seorang ATLM ingin mengetahui hubungan antara penurunan kualitas pemeriksaan proposal apabila pemeriksaan tersebut ditunda pada waktu tertentu. Hasil dan waktu penginderaan tertera pada tabel berikut:

No	lama penundaan	kadar glukosa	No	lama penundaan	kadar glukosa
1	30	222	11	330	186
2	60	222	12	360	183
3	90	220	13	390	181
4	120	229	14	420	175
5	150	227	15	450	172
6	180	223	16	480	171
7	210	223	17	510	170
8	240	222	18	540	169
9	270	200	19	570	169
10	300	188	20	600	163

Tentukan hubungan kedua varabel tersebut dengan tepat!

2. Mahasiswa akhir yang sedang mengerjakan skripsi melakukan uji coba karakter fungsi ginjal yaitu ureum, kreatinin dan asam urat

<b>fungsi ginjal</b>		
ureum	kreatinin	asam urat
13	0.83	2.6
27	0.89	9.7
26	0.87	8.9
17	0.83	8.3
37	0.6	5.6
34	0.93	8.2
32	0.73	8.9
25	0.79	8.4
29	0.95	6.7
40	0.86	10
27	0.67	7
20	0.84	5.9
21	0.91	7.5
14	0.9	10
35	0.68	8.1
30	0.69	9.4
39	0.62	9.7
12	0.99	6.3
40	0.57	8.8
37	0.97	5.1
12	0.71	9.9
18	0.96	7.7
17	0.94	7.4
27	0.98	7.4
17	0.72	8.3

27	0.52	9.2
25	0.52	8.2
26	0.58	7.3
27	0.57	8.6
14	0.85	8.3
30	0.5	8.8
30	0.76	6.8
38	0.69	6.7
31	0.73	9.8
24	0.72	9.5
15	0.63	8.4
36	0.51	6.1
18	0.91	9.7
37	0.82	6.3
21	0.68	8.8
21	0.82	8.7
29	0.92	8.8
38	0.55	6
14	0.95	6.6
19	0.66	7.7
20	0.57	5.2
12	0.57	9.9
12	0.88	8.1

---

Tentukan hubungan tiap variable tersebut dengan tepat!

### e. Daftar pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.

3. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
4. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
5. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
6. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
7. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
8. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
9. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## Bab XII. Uji Regresi

---

### Tujuan

---

Mahasiswa memahami dan dapat melakukan regresi dengan tepat

---

### Materi

---

#### 1. Analisa Regresi

---

##### a. Analisis Regresi

**P**ada pembahasan sebelumnya kita sudah menganalisa secara singkat mengenai regresi. Perbedaan utama antara korelasi dan regresi adalah menitikberatkan pada hubungan variabel dan menitikberatkan pada perkiraan suatu nilai berdasarkan data yang ada. Beberapa perbedaan antara korelasi dan regresi adalah analisa korelasi tidak ada variabel terikat maupun bebas sehingga semuanya dipukul rata yaitu sebagai sebuah data yang memiliki tingkatan yang sama. Analisa regresi memiliki variabel terikat dan variabel tidak terikat atau bebas.

Ekspresi dalam pengertian yang lebih mendalam dapat dikatakan bahwa variabel tertentu tergantung terhadap variabel yang lain yang ditujukan untuk mengestimasi atau memprediksi nilai rata yang ada pada populasi tersebut.

Kelayakan regresi di SPSS dapat didasarkan pada beberapa hal sebagai berikut

1. Model regresi dapat dikatakan dipercaya apabila nilai signifikan pada Anova menunjukkan perbedaan yang nyata
2. Efisien respirasi harus bersifat signifikan pengujian dapat dilakukan dengan menggunakan uji t regresi dikatakan signifikan apabila nilai T hitung lebih besar dengan nilai T tabel atau apabila menggunakan SPSS maka nilai signifikansi harus kurang dari 0,05
3. Linieritas atau keselarasan hasil modernisasi dapat diterangkan menggunakan kuadrat sehingga dapat disimpulkan bahwa apabila r kuadrat semakin tinggi maka keselarasan pada data tersebut semakin baik.
4. Data terdistribusi normal
5. Skala rasio atau interval
6. Terdapat hubungan yang erat Sehingga dalam beberapa kasus diharuskan untuk menganalisa terlebih dahulu sehingga hasil dari respirasi dapat dipercaya.

Rumus untuk regresi adalah sebagai berikut

$$Y = a + bX$$

$$a = \frac{\sum Y - b(\sum X)}{n}$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

Keterangan :

$a$  = nilai intercept

$b$  = slope  $aY/bX$  (koefisien arah regresi)

$X$  = nilai variabel 1 (variabel pengaruh/independent)

$Y$  = nilai variabel 2 (variabel dependent)

$N$  = banyaknya pasang data / pengukuran / sampel

Untuk memperdalam pemahaman mari kita ulas mengenai korelasi dengan contoh yang mirip dengan contoh sebelumnya

kadar gula	kadar leukosit
165	9567
165	9223
169	9123
171	8689
174	8659
176	8621
177	8605
180	8375
180	8372
181	8033
182	8032
187	7857
191	7380
193	7254
194	7242
197	6582
198	6357
198	5819
200	5473
200	5121

Penyelesain

no	kadar gula	kadar leukosit	XY	x2	y2
1	165	9567	1578555	27225	91527489
2	165	9223	1521795	27225	85063729
3	169	9123	1541787	28561	83229129
4	171	8689	1485819	29241	75498721
5	174	8659	1506666	30276	74978281
6	176	8621	1517296	30976	74321641
7	177	8605	1523085	31329	74046025
8	180	8375	1507500	32400	70140625
9	180	8372	1506960	32400	70090384
10	181	8033	1453973	32761	64529089
11	182	8032	1461824	33124	64513024
12	187	7857	1469259	34969	61732449
13	191	7380	1409580	36481	54464400
14	193	7254	1400022	37249	52620516
15	194	7242	1404948	37636	52446564
16	197	6582	1296654	38809	43322724
17	198	6357	1258686	39204	40411449
18	198	5819	1152162	39204	33860761
19	200	5473	1094600	40000	29953729
20	200	5121	1024200	40000	26224641
210	3678	154384	28115371	679070	1222975370
22	13527684				

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$b = \frac{20 (28115371) - (3678)(154384)}{20 (679070) - (13527684)}$$

$$B = -102.705$$

$$a = \frac{\sum Y - b(\sum X)}{n}$$

$$a = \frac{154384 - (-102.705 * 3678)}{20} = 26606$$

rumus untuk regresi adalah

$$Y = a + bX$$

$$Y = 26606 + (-102.705 * x)$$

Artinya untuk mencari nilai lain yang belum diketahui maka rumus tersebut dapat digunakan missal kita ingin mencari berapa nilai glukosa ketika kadar gula 300 maka tinggal memasukkan

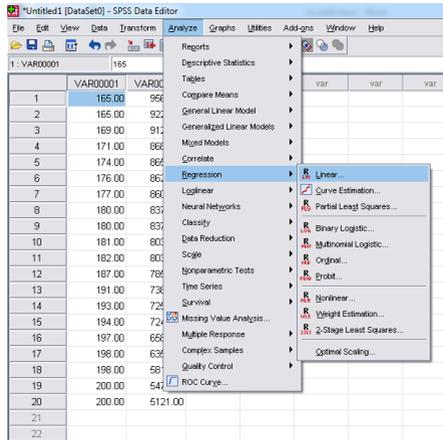
$$Y = 26606 + (-102.705 * x)$$

$$Y = 26606 + (-102.705 * 250)$$

$$Y = 1106$$

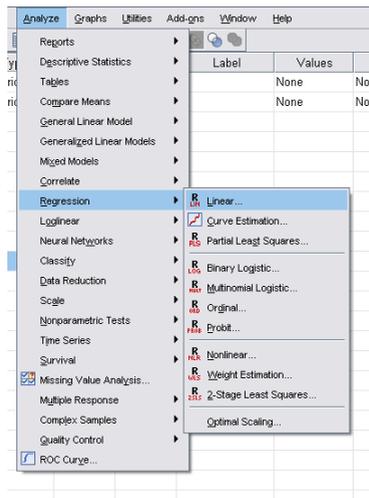
Dengan menggunakan SPSS

## Langkah 1. Input data



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	glukosa	Numeric	8	2		None	None	8	Right	Scale
2	leukosit	Numeric	8	2		None	None	8	Right	Scale

## Langkah 2. Masuk ke Analisis > Regresion > Linear



Langkah 3. Masukkan factor dependent (Y) dan factor independent (X) ke tempatnya



Langkah 4. Klik ok dan akan keluar gambar berikut

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2,833E7	1	2,833E7	174,439	,000 <sup>a</sup>
	Residual	2923416,959	18	162412,053		
	Total	3,125E7	19			

a. Predictors: (Constant), VAR00001

b. Dependent Variable: VAR00002

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	26606,753	1432,896		18,569	,000
	VAR00001	-102,706	7,776	-.952	-13,208	,000

a. Dependent Variable: VAR00002

Nilai 26606 adalah nilai a

Sedangkan -102,706 adalah nilai b. sehingga rumus di buat adalah

$$Y = 26606 + (-102,705 * x)$$

## b. Ringkasan materi

Regresi merupakan analisis yang digunakan untuk memprediksikan nilai yang belum ada, selain itu regresi juga digunakan untuk penyelarasan. Syarat dari regresi adalah memiliki nilai anova yang berbeda signifikan, data terdistribusi normal, dan berkorelasi kuat.

## c. Latihan soal

Berikut adalah gambaran karakter

sample	umur	hematologi		
		hemoglobin	leukosit	trombosit (rb)
1	55	14.6	8450	29
2	60	15.1	9488	30
3	29	11.9	4799	15
4	39	12.6	5689	19
5	52	14.1	7712	26
6	46	13.6	6878	22
7	56	14.6	8655	29
8	71	15.6	9777	34
9	18	11.1	3148	9
10	51	14	7563	24
11	33	12.4	4918	17
12	19	11.6	3986	12
13	19	11.3	3217	10
14	49	13.8	7463	23
15	69	15.5	9738	33
16	37	12.4	5442	18
17	46	13.7	7269	22

18	74	15.7	10136	34
19	53	14.3	8044	28
20	55	14.5	8423	28
21	67	15.3	9620	31
22	68	15.3	9699	32
23	33	12.1	4902	16
24	42	13.5	6474	21
25	48	13.7	7385	23
26	40	13.2	5714	19
27	80	15.9	10789	35
28	69	15.6	9764	33
29	19	11.5	3613	11
30	27	11.9	4417	12
31	31	12	4831	16
32	58	14.9	9348	29
33	52	14.3	7738	27
34	57	14.7	9255	29
35	37	12.6	5483	18
36	68	15.3	9686	32
37	74	15.8	10405	35
38	61	15.2	9604	30
39	20	11.6	4349	12
40	41	13.4	6389	20

---

Tentukan nilai persamaan pada masing masing karakter berdasarkan umur dan perkirakan karakter darah ketika seseorang berumur 80 tahun!

#### d. Daftar pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.
3. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
4. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
5. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
6. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
7. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
8. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
9. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS



## Bab XIII. Uji Chi-Square

---

### Tujuan

---

Mahasiswa memahami dan dapat melakukan uji chi square tepat

---

### Materi

---

1. Homogeneity test
  2. Goodness of fit tes
  3. Test of independence
- 

Uchiha Square merupakan uji yang digunakan untuk menguji hipotesa pada suatu proporsi atau lebih. Uji chi Square ditemukan oleh helmet pada tahun 1900. Dan digunakan untuk menguji kebebasan pada dua sampel atau variabel yang tersusun dalam baris dan kolom. Pada prakteknya seringkali digunakan untuk beberapa keperluan sehingga Uchiha server seringkali dibedakan menjadi beberapa jenis yaitu;

**Test of homogeneity:** Tes untuk membuktikan bahwa dalam populasi yang berbeda terdapat beberapa kesamaan proporsi karakteristik.

**Goodness of fit tests :** Tes yang digunakan untuk mengetes apakah suatu data yang telah kita peroleh ini sesuai ( fit) dengan distribusi/teori yang pilih

**Test of independence** Suatu tes yang bertujuan untuk membuktikan bahwa variabel di kolom dan baris saling tidak berhubungan

Membandingkan nilai Absolut atau nilai pengamatan dengan nilai nilai harapan semakin besar perbedaan antara nilai dengan cepat maka kemungkinan perbedaan antara proporsi yang diuji semakin besar pula. Contoh Apabila seseorang ibu memiliki 12 anak di mana ternyata dari ke-12 ciri dari 2 laki-laki dan 10 perempuan. Menurut perhitungan probabilita maka kelahiran laki-laki dan perempuan adalah 50% banding 50%. Sehingga seharusnya anak yang dimiliki dari ke-12 adalah 6 perempuan dan 6 laki-laki. Lalu dengan menggunakan chi square maka diharapkan dapat menguji apakah anak yang dimiliki masih sesuai atau tidak dengan teori probabilitas. Perbandingan nilai observe (Pengamatan) dengan nilai expected (Harapan) Makin besar perbedaan nilai observe dengan expected maka kemungkinan perbedaan antara proporsi yang diuji. Karena ada perbedaan jenis jenis tersebut maka bunyi dari hipotesis masing masing jenis tes juga berbeda. Adapun hipotesis adalah sebagai berikut

Ada 3 jenis kegunaan uji X2 yaitu:

Uji Goodness of fit (Kesesuaian)

Ho= tidak ada perbedaan antara. . . . .

Ho= frekwensi . . . . . Sesuai dengan . . . . .

Uji Homogeneity (Homogenitas)

Ho= tidak ada satu item yang keluar dari.. . . .

Ho= sampel. . . . . Adalah homogen

Uji Independency (Assosiasi= hubungan)

Ho= tidak ada hubungan

Ho= variabel-variabel saling bebas

### a. Uji Homogenity (Homogenitas) chi square test

Uji homogenitas pada chi square digunakan untuk mengetahui data yang di miliki apakah tersebar normal atau tidak. Untuk menguji ini semua nilai expected

Manajer Personalia ingin melihat apakah pola absensi terdistribusi secara merata sepanjang enam hari kerja. Hipotesis nol yang akan diuji adalah “Absensi terdistribusi secara merata selama enam hari kerja. Taraf nyata yang digunakan adalah 0,05. Hasil dari sampel ditunjukkan sebagai berikut :

Hari	Jumlah Absen
Senin	12
Selasa	9
Rabu	11
Kamis	10
Jum’at	9
Sabtu	9

Ujilah hipotesis tersebut !

	Observe	Expected	fo-fe	(fo-fe) <sup>2</sup>	(fo-fe) <sup>2</sup> /fe
Senin	12	10	-2	4	0.4
Selasa	9	10	1	1	0.1
Rabu	11	10	-1	1	0.1
Kamis	10	10	0	0	0
Jum’at	9	10	1	1	0.1
Sabtu	9	10	1	1	0.1
	10			8	0.8

Sehingga nilai chi square hitung =

$$X^2 = \sum \left[ \frac{(f_o - f_e)^2}{f_e} \right]$$

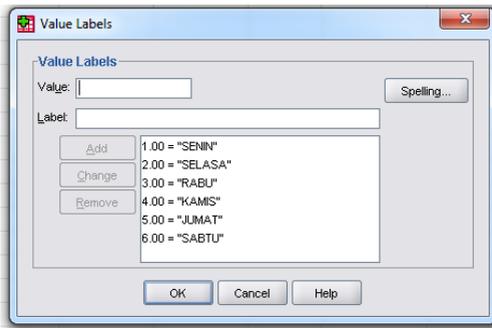
$$X^2 = 0.8$$

Untuk menentukan apakah uji tersebut berbeda nyata atau tidak maka harus di bandingkan dengan chi square table .

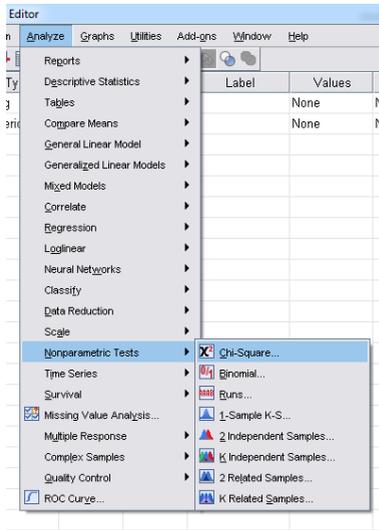
### Penyelesaian menggunakan SPSS

Untuk menyelesaikan permasalahan tersebut anda dapat menggunakan

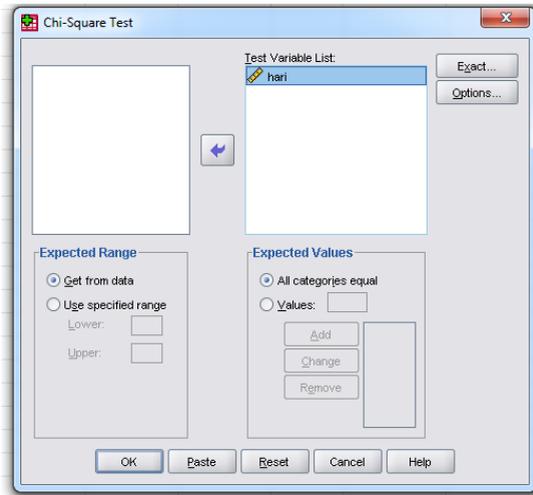
	hari
1	3.00
2	1.00
3	1.00
4	1.00
5	3.00
6	4.00
7	5.00
8	6.00
9	1.00
10	2.00
11	3.00
12	4.00
13	1.00
14	2.00
15	3.00
16	4.00
17	5.00
18	6.00
19	1.00
20	2.00
21	3.00
22	4.00
23	5.00



Kemudian klik analyze > Non parametric test > Chi- Square



Maka akan muncul seperti gambar berikut



Karena kita mengasumsikan semua homogeny maka centang all catagories equal

hari			
	Observed N	Expected N	Residual
BENIN	12	10.0	2.0
SELASA	9	10.0	-1.0
RABU	11	10.0	1.0
KAMIS	10	10.0	.0
JUMAT	9	10.0	-1.0
SABTU	9	10.0	-1.0
Total	60		

**Test Statistics**

	hari
Chi-Square	.800 <sup>a</sup>
df	5
Asymp. Sig.	.977

a. 0 cells (.0%) have expected frequencies less than 5. The frequency is 10.0.

Dari hasil analisis didapatkan nilai chi square adalah 0.8 dan nilai asymt. Sig sebesar 0.977 itu artinya hasil masih cukup tinggi untuk menerima Ha. Sehingga Ho diterima yang artinya tidak ada perbedaan antara data observasi dengan data harapan (expected).

## b. Uji Goodness of fit (Kesesuaian) chi square test

Uji Goodness of fit (Kesesuaian) chi square test digunakan untuk mengetahui kesesuaian teori yang ada dengan data hasil observasi. Secara peluang apabila kita melempar dadu maka ada 1/6 kemungkinan untuk dapat angka 1. Ternyata dalam kenyataannya setelah 60 kali lempar mungkin kecil kemungkinan anda mendapatkan 10 kali angka 1. Dalam hal ini anda dapat kenyataan bahwa mungkin lebih rendah atau bahkan lebih tinggi dari angka 10. Sama seperti anda melempar koin. Secara teori koin akan menghasilkan 50 ; 50 antara satu sisi dengan sisi yang lain. Dalam 100 kali lempar apakah anda yakin 1 sisi pasti menghasilkan 50 sisi A. tentu jawabannya adalah tidak. Sisi yang lain dapat menunjukkan nilai yang lebih tinggi atau lebih rendah. Apabila anda mendapatkan angka yang benar benar 50 kali pada sisi A. jawabannya mungkin anda sedang beruntung. Penyimpangan ini

yang digunakan oleh Uji Goodness of fit (Kesesuaian) chi square test untuk menentukan apakah data yang diobservasi masih sesuai dengan hukum atau tidak.

Untuk mempermudah pemahaman anda coba perhatikan contoh di bawah ini “Uji Saya ingin membuktikan bahwa pewarisan rambut keriting menunjukkan pewarisan sifat 9:3:3:1, pada pengamatan saya temukan keriting 46, ikal 24, gelombang 26, dan lurus 7, buktikan apakah bentuk rambut tersebut sesuai dengan hukum pewarisan 9:3:3:1?”

**Penyelesaian**

*Rumusan masalah*

Apakah hasil observasi pewarisan rambut sesuai dengan teori 9:3:3:1?

*Tujuan*

Untuk mengetahui apakah hasil observasi pewarisan rambut sesuai dengan teori 9:3:3:1.

*Hipotesis*

Ho= hasil observasi pewarisan rambut sesuai dengan teori 9:3:3:1.

Ha= hasil observasi pewarisan rambut tidak sesuai dengan teori 9:3:3:1.

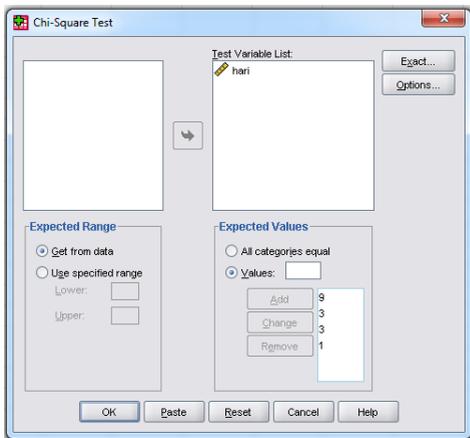
*Analaisis penghitungan*

kreteria	Observed	prop	presen prop	Fe	Fo-Fe	(fo-fe) <sup>2</sup>	
keriting	46	9	0.5	57.9	-11.9	142.	2.45
Ikal	24	3	0.1	19.3	4.6	21	1.13
Gelombang	26	3	0.1	19.3	6.6	44	2.31
Lurus	7	1	0.06	6.4	0.5	0	0.04
	103	16	1				5.96

Maka nilai  $\chi^2$  adalah 5.96

## Penyelesaian menggunakan SPSS

hari	1
1	keriting
2	keriting
3	keriting
4	keriting
5	keriting
6	keriting
7	keriting
8	keriting
9	keriting
10	keriting
11	keriting
12	keriting
13	keriting
14	keriting
15	keriting
16	keriting
17	keriting
18	keriting



Kemudian setelah itu klik ok akan muncul data seperti ini

## Frequencies

hari			
	Observed N	Expected N	Residual
keriting	46	57.9	-11.9
ikal	24	19.3	4.7
gelombang	26	19.3	6.7
lurus	7	6.4	.6
Total	103		

## Test Statistics

	hari
Chi-Square	5.962 <sup>a</sup>
df	3
Asymp. Sig.	.113

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.4.

Dari hasil tersebut

## c. Independence chi square test

Independence chi square test digunakan untuk menguji ketee keterkaitan sautu variable dengan variable lainnya. Sebagai contoh kita akan menganalisis pemasalahan berikut Seorang peneliti ingin mengetahui hubungan antara kebiasaayaan merokok dengan kejadian kanker prostat data tersaji pada table berikut

Karakteristik	B		Total
	+	-	
A1	a	B	a + b
A2	c	D	c + d
Total	a + c	b + d	a+b+c+d

Dengan nilai ekspected

$$Ea = \frac{(a+b)(a+c)}{n}$$

$$Eb = \frac{(a+b)(b+d)}{n}$$

$$Ec = \frac{(a+c)(c+d)}{n}$$

$$Ed = \frac{(b+d)(c+d)}{n}$$

untuk memperdalam pemahaman ini maka lebih baik menyimak contoh soal berikut seorang ATLM ingin mengetahui hubungan kebiasaan merokok dengan kejadian kanker prostat. Peneliti melibatkan 231 responden, data adalah sebagai berikut:

Paparan	Kanker prostat		Total
	+	-	
perokok	67 (a)	56 (b)	123
non perokok	32 (c)	76 (d)	108
Total	99	132	231

### Rumusan Masalah

Apakah terdapat hubungan kebiasaan merokok dengan kejadian kanker prostat. Peneliti melibatkan 231 responden?

### Tujuan

Terdapat hubungan kebiasaan merokok dengan kejadian kanker prostat.

### Hipotesis

Ho = tidak ada hubungan kebiasaan merokok dengan kejadian kanker prostat.

Ha= terdapat hubungan kebiasaan merokok dengan kejadian kanker prostat.

### Taraf kepercayaan

Taraf kepercayaan yang digunakan 95%

### Penghitungan

Untuk mencari nilai ekspekted maka digunakan rumus

$$Ea = \frac{(a+b) (a+c)}{N}$$

$$Ea = \frac{123 \times 99}{231} = 52.71$$

$$Eb = \frac{(a+b) (b+d)}{N}$$

$$Eb = \frac{123 \times 132}{231} = 70.28$$

$$Ec = \frac{(a+c) (c+d)}{N}$$

$$Ec = \frac{123 \times 99}{231} = 46,28$$

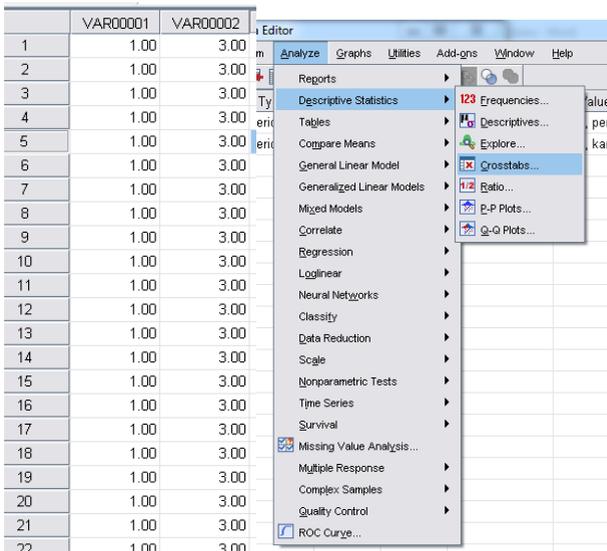
$$Ed = \frac{(b+d) (c+d)}{n}$$

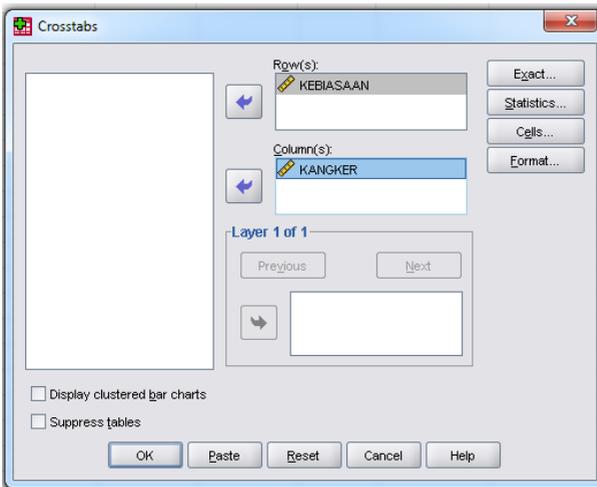
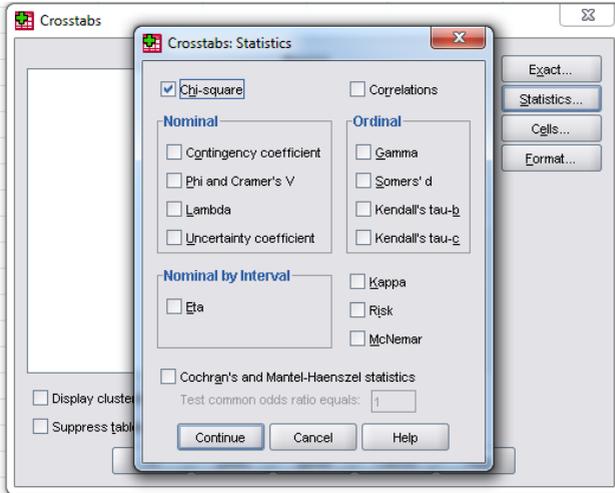
$$Ed = \frac{132 \times 108}{231} = 61.71$$

## Penghitungan menggunakan rumus

Karakteristik	Status kangker	o	e	fo-fe	(fo-fe) <sup>2</sup>	(fo-fe) <sup>2</sup> / Fe
perokok	positif	67	52.71	-14.28	204.08	3.87
perokok	negatif	56	70.28	14.28	204.08	2.90
non perokok	positif	32	46.28	14.28	204.08	4.40
non perokok	negatif	76	61.71	-14.28	204.08	3.30
$\chi^2$						14.49

## Penghitungan menggunakan SPSS





[DataSet0]

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
KEBIASAAN * KANGKER	232	100.0%	0	.0%	232	100.0%

**KEBIASAAN \* KANGKER Crosstabulation**

Count		KANGKER		Total
		3	4	
KEBIASAAN	perokok	67	57	124
	non perokok	32	76	108
Total		99	133	232

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	14.051 <sup>a</sup>	1	.000		
Continuity Correction <sup>b</sup>	13.072	1	.000		
Likelihood Ratio	14.265	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	13.991	1	.000		
N of Valid Cases <sup>a</sup>	232				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 46.09.

b. Computed only for a 2x2 table

## d. Ringkasan Materi

Uji Chi square termasuk uji non parametric. Berdasarkan jenisnya dapat di bagi menjadi 3 yaitu goodness of fit chi square test, homogeneity chi-square test dan independence chi-square test. Masing masing memiliki fungsi yang khas dan spesifik. Data yang digunakan pada uji ini adalah nominal atau ordinal.

### e. Latihan soal

Seorang peneliti ingin mengetahui hubungan kebiasaan makan sayur dengan anemia

No	Kebiasaan makan sayur	Anemia
1	Ya	Tidak
2	Tidak	Ya
3	Tidak	Ya
4	Tidak	Ya
5	Ya	Tidak
6	Tidak	Ya
7	Ya	Tidak
8	Ya	Tidak
9	Ya	Tidak
10	Ya	Tidak
11	Tidak	Ya
12	Tidak	Ya
13	Tidak	Tidak
14	Ya	Tidak
15	Tidak	Ya
16	Tidak	Tidak
17	Tidak	Ya
18	Tidak	Ya
19	Tidak	Ya

Tentukan apakah terdapat ada hubungan kebiasaan makan sayur dengan anemia !

#### a. Daftar pustaka

1. Abdurrahman, M., Muhidin, S.A., Somantri. 2005. Dasar-dasar metode statistika untuk penelitian. Bandung: Pustaka Setia
2. Hasan, Iqbal. 2002. Pokok-pokok Materi Statistik 2 (Statistik Inferensif). Jakarta: Bumi Aksara.

3. Murti, Bhisma, Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
4. Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung
5. Ott R. L. & Longnecker M. 2010. An Introduction to statistical Methods and Data Analysis: Sixth Edition. Cengage Learning
6. Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
7. Sudjana, 1996. Metoda Statistika Edisi ke 6. Bandung: Penerbit Tarsito
8. Santoso, S. 2014. Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia
9. Widyato, J. 2012. Spss for windows. Surakarta:FKIP UMS

# Daftar Pustaka

- Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
- Dahlan M.S., 2009. Besar sampel dan cara pengambilan sampel: dalam penelitian kedokteran dan kesehatan
- Hadi, S, 1997. *Statistik jilid I* : cetakan kedua puluh. Yogyakarta: Penerbit Andi
- Hasan, Iqbal. 2002. *Pokok-pokok Materi Statistik 2 (Statistik Inferensif)*. Jakarta: Bumi Aksara.
- Kustianto, B.& Badrudin, R. 1994, *Statistika 1 (Deskriptif)*, Gunadarma
- Lemeshow, Stanley et all, 1997. Besar Sampel penelitian Kesehatan, UGM Press,
- Murti, Bhisma, *Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan*, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
- Nurgiyantoro, Gunawan, Marzuki. 2000. *Statistik Terapan untuk Penelitian Ilmu-ilmu sosial*. Gadjah Mada University Press. Yogyakarta.
- Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung

- Ott R. L. & Longnecker M. 2010. *An Introduction to statistical Methods and Data Analysis: Sixth Edition*. Cengage Learning
- Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
- Siegel, S. 1994. *Statistik Non parametrik untuk ilmu-ilmu sosial. PT Gramedia pustaka utama. Jakarta.*
- Sudjana, 1996. *Metoda Statistika Edisi ke 6*. Bandung: Penerbit Tarsito
- Santoso, S. 2014. *Panduan lengkap spss versi 20 edisi revisi, Jakarta:gramedia*
- Widyato, J. 2012. *Spss for windows. Surakarta:FKIP UMS*

Lampiran -Lampiran

Table for Determining Sample Size from a Given Population

$N$	$S$	$N$	$S$	$N$	$S$
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	100000	384

$N$  is population size.

$S$  is sample size.

# Daftar Pustaka

- Abdurrahman, M., Muhidin, S.A., Somantri. 2005. *Dasar-dasar metode statistika untuk penelitian*. Bandung: Pustaka Setia
- Dahlan M.S., 2009. Besar sampel dan cara pengambilan sampel: dalam penelitian kedokteran dan kesehatan
- Hadi, S, 1997. *Statistik jilid I* : cetakan kedua puluh. Yogyakarta: Penerbit Andi
- Hasan, Iqbal. 2002. *Pokok-pokok Materi Statistik 2 (Statistik Inferensif)*. Jakarta: Bumi Aksara.
- Kustianto, B.& Badrudin, R. 1994, *Statistika 1 (Deskriptif)*, Gunadarma
- Lemeshow, Stanley et all, 1997. Besar Sampel penelitian Kesehatan, UGM Press,
- Murti, Bhisma, *Penerapan Metode Statistik Non-Parametrik Dalam Ilmu-ilmu Kesehatan*, Penerbit PT Gramedia Pustaka Utama, Jakarta. 1996.
- Nurgiyantoro, Gunawan, Marzuki. 2000. *Statistik Terapan untuk Penelitian Ilmu-ilmu sosial*. Gadjah Mada University Press. Yogyakarta.
- Gasperz, Vincent, Teknik Penarikan Contoh, Tarsito, Bandung

- Ott R. L. & Longnecker M. 2010. *An Introduction to statistical Methods and Data Analysis: Sixth Edition*. Cengage Learning
- Purnawan, J., 1997 Teknik Analisis Data, Rineka Cipta,
- Siegel, S. 1994. *Statistik Non parametrik untuk ilmu-ilmu sosial*. PT Gramedia pustaka utama. Jakarta.
- Sudjana, 1996. *Metoda Statistika Edisi ke 6*. Bandung: Penerbit Tarsito
- Santoso, S. 2014. *Panduan lengkap spss versi 20 edisi revisi*, Jakarta:gramedia
- Widyato, J. 2012. *Spss for windows*. Surakarta:FKIP UMS

## BIODATA



lulus dari Universitas Negeri Malang (UM) pada tahun 2011, Pada tahun 2014 lulus Program Pascasarjana Biologi, Universitas Gadjah Mada (UGM) dengan fokus penelitian Genetika dan Biologi Molekular. Pada tahun 2015-2017, menjadi Staff pengajar di STIKES Perintis Sumatera Barat Prodi Analis Kesehatan (sekarang Teknologi Laboratorium Medis (TLM)), dan mengembangkan Laboratorium Biologi Molekular. Pada tahun 2018- sekarang dipercaya mengembangkan Lab Biologi Molekular prodi TLM di Universitas Muhammadiyah Sidoarjo dan menjadi Dosen tetap di Kampus tersebut.



**Rafhani Rosyidah., S.Keb., Bd., M.Sc.** dilahirkan di Magetan, 12 Juni 1987. Lulus Sarjana kebidanan pada tahun 2011, dan profesi bidan pada tahun 2012 di Universitas Airlangga Surabaya. Penulis melanjutkan studi S2 di Prodi Ilmu Kedokteran Klinik peminatan maternal perinatal di Universitas Gadjah mada Yogyakarta. Tahun 2015, penulis secara resmi mendamatkan gelar M.Sc. Saat ini penulis aktif bekerja di prodi kebidanan Fakultas Ilmu Kesehatan Universitas Muhammadiyah Sidoarjo sebagai dosen, selain mengajar penulis juga aktif dalam melakukan penelitian dalam bidang kebidanan.