

Desember 2021

MODUL AJAR

# Workshop Statistik



**HENDRYADI, SE., MM**



**Sekolah Tinggi Ilmu Ekonomi  
Indonesia Jakarta  
2021**

## **KATA PENGANTAR**

Puji syukur kehadirat Allah SWT atas rahmat dan hidayahNya saya dapat menyelesaikan Modul Workshop Statistik SPSS. Adapun tujuan dari pembuatan modul ini adalah sebagai bahan ajar dan referensi bagi para pembaca, khususnya mahasiswa di Sekolah Tinggi Ilmu Ekonomi Indonesia Jakarta.

Mudah-mudahan buku ini dapat membantu para pembaca yang berminat untuk mengembangkan diri, memperkaya wawasan dan menambah khasanah ilmu pengetahuan. Kami menyadari bahwa penyelesaian buku ini tidak terlepas dari bantuan berbagai pihak, dan masih banyak terdapat kekurangan dalam penulisan buku ini. Oleh karena itu, kami mengharapkan kritik dan saran yang membangun dari pembaca.

Jakarta, 1 Oktober 2021

Hendryadi

## DAFTAR ISI

KATA PENGANTAR .....	1
DAFTAR ISI .....	2
PERTEMUAN 1 PENGENALAN STATISTIKA .....	3
PERTEMUAN 2 PENGENALAN SPSS.....	13
PERTEMUAN 3 DESKRIPSI DATA.....	22
PERTEMUAN 4-5 ANALISIS KORELASI.....	37
PERTEMUAN 6-7-8 ANALISIS REGRESI .....	46
PERTEMUAN 9-10-11 UJI BEDA RATA-RATA.....	62
PERTEMUAN 12 UJI RELIABILITAS DAN VALIDITAS .....	77

## PERTEMUAN 1 PENGENALAN STATISTIKA

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami dan menjelaskan tentang makna statistika dan penggunaannya dalam kehidupan sehari-hari.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami dan menjelaskan tentang makna statistika, jenis data dan penggunaannya dalam kehidupan sehari-hari.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Pengertian statistika dan statistik
  - b. Jenis statistika
  - c. Variabel dan jenisnya
  - d. Pengujian hipotesis

### 5. Materi

#### Definisi Statistika

- *Statistika* adalah ilmu mengumpulkan, menata, menyajikan, menganalisis dan menginterpretasikan data menjadi informasi untuk membantu pengambilan keputusan yang efektif.
- *Statistik* adalah suatu kumpulan angka yang tersusun lebih dari satu angka.

#### Jenis Statistika

- *Statistika Deskriptif* adalah metode statistika yang digunakan untuk menggambarkan atau mendeskripsikan data yang telah dikumpulkan menjadi sebuah informasi.
- *Statistika Induktif* adalah metode yang digunakan untuk mengetahui tentang sebuah populasi berdasarkan suatu sampel atau contoh dengan menganalisis dan menginterpretasikan data menjadi sebuah kesimpulan.

#### Data

Secara umum, data dikelompokkan menjadi dua yaitu kualitatif dan kuantitatif. Data kualitatif (non metric) adalah seperti jenis kelamin, tempat lahir, pendidikan, warna favorit, suku, dan sebagainya. Sementara data kuantitatif adalah data yang berbentuk angka (metric) dan dapat dikelompokkan menjadi dua yaitu data diskrit dan data kontinyu.

Data diskrit yakni data dalam bentuk bilangan bulat. Nilai ini merupakan bilangan asli dan tidak mungkin dalam bentuk pecahan atau desimal, kecuali jika memberikan pengertian rata-rata. Misalnya ukuran sepatu, jumlah anggota keluarga, jumlah siswa di kelas dan lain sebagainya. Data kontinyu, yaitu data yang dapat mempunyai nilai yang terletak dalam seluruh interval. Misalnya tinggi badan, berat badan, jumlah penjualan, dan sebagainya.

Berdasarkan level pengukurannya, data juga dapat dikelompokkan menjadi empat tingkatan yaitu nominal, ordinal, interval dan rasio. Skala nominal digunakan

untuk memberikan label pada sebuah variabel, dan bukan nilai kuantitatif. Skala ordinal, merupakan urutan nilai-nilai apa yang dianggap penting dan signifikan, tetapi perbedaan antara masing-masing nilai tidak benar-benar diketahui. Skala interval adalah skala numerik yang merupakan urutan dan memiliki perbedaan yang tepat antara nilai-nilainya. Contoh dari skala interval adalah suhu Celcius karena perbedaan antara setiap nilai adalah sama. Ke-empat yaitu skala rasio adalah skala level tertinggi yang dapat memberikan informasi sebuah nilai yang tepat pada sebuah unit, dan memiliki nilai nol mutlak. Baik skala rasio dan interval dapat menggunakan statistik deskriptif dan inferensial dalam proses analisisnya.

Tabel 1. Skala Pengukuran

Tipe Skala	Contoh	Ciri Khusus
Nominal	Jenis Kelamin 1. Laki-Laki 2. Wanita Pekerjaan 1. Wiraswasta 2. PNS 3. Lain Suku 1. Batak 2. Padang 3. Jawa 4. Dan seterusnya	Penomoran hanya digunakan sebagai label dan tidak untuk operasi matematika
Ordinal	Likert Scale 1. Sangat tidak setuju 2. Tidak setuju 3. Ragu 4. Setuju 5. Sangat Setuju Kepuasan 1. Sangat tidak puas 2. Tidak puas 3. Cukup puas 4. Puas 5. Sangat puas	Tidak diketahui secara jelas mengenai jarak antar skor.
Interval	Suhu : perbedaan antara 40 dan 50 derajat adalah 10 derajat, seperti perbedaan antara 70 dan 60 derajat.	Jarak antar nilai diketahui secara jelas dan nol tidak bernilai mutlak
Rasio	Penjualan : Rp. 100 Juta Harga Saham : Rp. 1500 Jumlah mobil : 200 unit	Jarak antar nilai diketahui secara jelas dan nol bernilai mutlak

Perbedaan ketiga adalah jenis data berdasarkan waktu pengumpulannya. Jenis ini terdiri dari tiga yaitu data cross-sectional, data time-series dan data panel.

Data cross-sectional adalah data yang dikumpulkan dalam waktu tertentu dan dapat memberikan gambaran mengenai kegiatan atau nilai pada waktu itu sendiri. Kedua adalah data time-series yaitu data yang dikumpulkan dari waktu ke waktu untuk memberikan gambaran tentang perkembangan suatu kegiatan selama periode spesifik yang diamati. Data berkala sering kali disebut pula sebagai data historis. Ketiga adalah data panel, yaitu data gabungan time series dan cross-section.

Tabel 2. Klasifikasi data berdasarkan waktu pengumpulan

<b>Type Data</b>	<b>Contoh</b>														
Cross-sectional	<p>Peneliti melakukan pengumpulan data mengenai harga saham pada 6 perusahaan pada bulan Januari 2016.</p> <table border="1"> <thead> <tr> <th><b>Perusahaan</b></th> <th><b>Harga Saham</b></th> </tr> </thead> <tbody> <tr> <td>A</td> <td>1,500</td> </tr> <tr> <td>B</td> <td>3,000</td> </tr> <tr> <td>C</td> <td>45,000</td> </tr> <tr> <td>D</td> <td>2,500</td> </tr> <tr> <td>E</td> <td>3,500</td> </tr> <tr> <td>F</td> <td>5,500</td> </tr> </tbody> </table>	<b>Perusahaan</b>	<b>Harga Saham</b>	A	1,500	B	3,000	C	45,000	D	2,500	E	3,500	F	5,500
<b>Perusahaan</b>	<b>Harga Saham</b>														
A	1,500														
B	3,000														
C	45,000														
D	2,500														
E	3,500														
F	5,500														
Time Series	<p>Peneliti melakukan pengumpulan data mengenai harga saham pada perusahaan A pada periode bulan Januari - Juni 2016.</p> <p>Harga Saham PT. A</p> <table border="1"> <thead> <tr> <th><b>Bulan</b></th> <th><b>Harga Saham</b></th> </tr> </thead> <tbody> <tr> <td>Januari</td> <td>1,500</td> </tr> <tr> <td>Februari</td> <td>1,350</td> </tr> <tr> <td>Maret</td> <td>2,000</td> </tr> <tr> <td>April</td> <td>2,100</td> </tr> <tr> <td>Mei</td> <td>1,750</td> </tr> <tr> <td>Juni</td> <td>1,800</td> </tr> </tbody> </table>	<b>Bulan</b>	<b>Harga Saham</b>	Januari	1,500	Februari	1,350	Maret	2,000	April	2,100	Mei	1,750	Juni	1,800
<b>Bulan</b>	<b>Harga Saham</b>														
Januari	1,500														
Februari	1,350														
Maret	2,000														
April	2,100														
Mei	1,750														
Juni	1,800														
Data Panel	<p>Peneliti melakukan pengumpulan data mengenai harga saham pada perusahaan A, B, dan C pada periode bulan Januari - Juni 2016</p>														

Harga Saham PT. A, B, dan C		
Perusahaan	Bulan	Harga Saham
A	Januari	1,500
	Februari	1,350
	Maret	2,000
	April	2,100
	Mei	1,750
	Juni	1,800
B	Januari	3,000
	Februari	3,250
	Maret	2,800
	April	2,350
	Mei	1,900
	Juni	1,450
C	Januari	4,500
	Februari	4,800
	Maret	4,750
	April	4,550
	Mei	4,200
	Juni	5,200

### Variabel penelitian

Dalam penelitian kuantitatif, teknik analisis terbagi menjadi dua yaitu statistik deskriptif dan statistik inferensial. Statistik deskriptif adalah statistik yang menggambarkan fenomena atau data sebagaimana dalam bentuk tabel, grafik, frekuensi, rata-rata ataupun bentuk lainnya. Statistik deskriptif umumnya hanya memberikan gambaran (deskripsi) mengenai keadaan data sebenarnya tanpa bermaksud membuat generalisasi dari data tersebut. Dalam statistik deskriptif dilakukan analisis dalam bentuk tabel, kolom, grafik, perhitungan frekuensi, ukuran tendensi pusat (mean, median, modus), ukuran disperse (kisaran, varian, standar deviasi), dan lain sebagainya. Statistik inferensial digunakan untuk penelitian sampel, dimana peneliti bermaksud melakukan generalisasi atau menaksir populasi berdasarkan data sampel. Statistik inferensi ini terbagi menjadi dua yaitu parametris dan nonparametris. Beberapa teknik dalam statistik inferensial dapat dikelompokkan menjadi model asosiatif, komparatif, dan interdependen.

### Hipotesis Penelitian

Hipotesis merupakan jawaban sementara terhadap permasalahan yang diajukan. Hipotesis merupakan pernyataan sementara berupa dugaan mengenai apa saja yang sedang kita amati dalam usaha untuk memahaminya. Hipotesis merupakan kebenaran sementara yang masih harus diuji. Oleh karena itu hipotesis berfungsi sebagai cara untuk menguji kebenaran.<sup>1</sup>

Hipotesis merupakan pernyataan spesifik yang masih bersifat prediksi atau dugaan peneliti, atau menjelaskan secara konkrit (bukan teoritis) apa yang diharapkan oleh peneliti dari rumusan masalah yang sudah diajukan sebelumnya. Dengan

<sup>1</sup> Suryani dan Hendryadi. (2015). Metode Riset Kuantitatif. (Jakarta, Prenadamedia Group), hlm.

demikian, pernyataan hipotesis merupakan jawaban sementara dari rumusan masalah yang diajukan. Hipotesis diajukan didasarkan pada teori dan bukti empiris yang memadai. Namun, teori umumnya hanya menjelaskan mengenai sebuah hubungan sehingga dapat dianggap bersifat tentatif untuk diuji ulang. Dengan melakukan pengujian, maka sebuah penelitian bisa memberikan kesimpulan apakah temuan penelitian berhasil mendukung hipotesis atau menolaknya.

### **Tipe Hipotesis**

Dalam praktik pengujiannya, hipotesis dikelompokkan menjadi tiga, yaitu hipotesis penelitian/kerja, hipotesis operasional, dan hipotesis statistik. Hipotesis penelitian/hipotesis kerja merupakan dugaan peneliti terhadap suatu masalah yang sedang dikaji. Peneliti menganggap hipotesis ini benar untuk kemudian dilakukan pengujian secara empiris dengan menggunakan data-data hasil penelitian. Hipotesis penelitian ini biasanya disimbolkan dengan  $H_1$ .

Hipotesis operasional merupakan lawan dari hipotesis penelitian yang berperan sebagai hipotesis pembanding untuk memberikan keseimbangan pada hipotesis penelitian. Karena peneliti mempunyai keyakinan bahwa dalam pengujian yang dilakukan nanti akan menemukan kebenaran atau kekeliruan hipotesis penelitian, hal ini sangat tergantung pada data yang diperoleh selama melakukan penelitian. Secara teknis hipotesis operasional disebut hipotesis nol ( $H_0$ ).

Sebagai contoh, jika peneliti mengajukan hipotesis penelitian bahwa "Indeks Prestasi Kumulatif (IPK) mahasiswa di Universitas XYZ lebih besar atau sama dengan 3.00", maka sebagai pembanding akan diajukan hipotesis nol yaitu "Indeks Prestasi Kumulatif (IPK) mahasiswa di Universitas XYZ lebih kecil dari 3.00". Jika peneliti dapat menunjukkan bahwa nilai IPK ini melebihi 3.00 maka kesimpulan yang diambil adalah menolak hipotesis nol, artinya dugaan "Indeks Prestasi Kumulatif (IPK) mahasiswa di Universitas swasta di Jakarta lebih besari dari 3.00" terdukung.

Hipotesis statistik merupakan jenis hipotesis yang dirumuskan dalam bentuk notasi statistik, hipotesis ini didasarkan pada pengamatan peneliti terhadap populasi angka-angka (kuantitatif).

### **Pengujian Hipotesis**

Pengujian hipotesis berarti pembuktian secara empiris untuk menentukan apakah hipotesis penelitian yang diajukan berhasil didukung atau ditolak. Dalam pengujian ini, peneliti melakukan pengumpulan data kemudian melakukan serangkaian pengujian untuk dapat memberikan kesimpulan tersebut. Untuk menghindari, atau setidaknya mengurangi jumlah subjektivitas yang ada saat peneliti menafsirkan data yang dikumpulkan maka terdapat aturan keputusan yang secara spesifik mengatur untuk menolak atau menerima hipotesis. Aturan ini disebut sebagai tingkat signifikansi.

### **Tingkat signifikansi dan p value**

Tingkat signifikansi atau level signifikansi, mengacu pada kriteria keputusan di mana keputusan dibuat mengenai nilai dinyatakan dalam hipotesis nol. Kriteria ini didasarkan pada statistik probabilitas yang diukur dalam data sampel. Tingkat signifikansi, juga dilambangkan sebagai alpha, adalah probabilitas menolak hipotesis nol. Sebagai contoh, tingkat signifikansi 0,05 menunjukkan risiko 5% dari kesimpulan ada perbedaan ketika tidak ada perbedaan yang sebenarnya.



### P value

Dalam penelitian sosial, kriteria atau tingkat signifikansi biasanya ditetapkan pada 1-5%. Ketika probabilitas rata-rata sampel kurang dari 5% (misalnya) maka hipotesis nol terbukti salah sehingga peneliti dapat menolak nilai dinyatakan dalam hipotesis nol. Untuk menaksir populasi berdasarkan sampel, ada dua kemungkinan kesalahan antara lain :

1. Kesalahan tipe I. Kesalahan ini terjadi jika peneliti menolak hipotesis nol ( $H_0$ ) yang benar (seharusnya diterima). Kemungkinan melakukan kesalahan Tipe I disebut tingkat signifikansi, dan sering dilambangkan dengan  $\alpha$  /  $\alpha$ .
2. Kesalahan tipe II adalah kesalahan yang terjadi ketika menerima hipotesis yang salah (seharusnya ditolak). Kemungkinan melakukan kesalahan tipe II disebut Beta, dan sering dilambangkan dengan  $\beta$ .

### Signifikansi Statistik

Probabilitas (p-value) dalam penelitian sosial merupakan kemungkinan hasil yang mungkin dari sebuah eksperimen, yaitu, kemungkinan bahwa pengamatan seseorang memberikan hasil sesuai yang diharapkan atau tidak diharapkan. P-value yang rendah menunjukkan bahwa sampel yang digunakan dapat memberikan cukup bukti menolak hipotesis nol untuk seluruh populasi. Sedangkan signifikansi, atau signifikansi statistik, menggambarkan keputusan yang dibuat mengenai nilai dinyatakan dalam hipotesis nol. Ketika hipotesis nol ditolak, hasilnya adalah signifikan. Ketika hipotesis nol ditolak, kita mencapai signifikansi, dan sebaliknya ketika hipotesis nol dipertahankan, kita gagal untuk mencapai signifikansi.

Sebagai contoh, misalnya kita melempar koin sebanyak 100 kali, kita akan mengharapkan hasil mendapatkan 50 kepala dan 50 ekor. Hasil yang diharapkan ini mengasumsikan bahwa koin tidak bias. Dalam pelemparan koin, hanya dua kemungkinan hasil, kepala atau ekor, dan jika koin adalah berisi, maka kita harapkan kepala 50% dan ekor 50% dari 100 kali pelemparan koin. Dengan kata lain, probabilitas mendapatkan kepala adalah  $P = 0,5$ , dan kemampuan mendapatkan ekor adalah  $P = 0,5$ .

Namun dalam penelitian sosial, probabilitas atau tingkat signifikansi berkisar antara 0.01 sampai dengan 0.05 atau  $p \text{ value} \leq 0.05$  /  $p \text{ value} \leq 0.01$ . Ketika nilai  $p$  kurang dari 5% ( $p < 0,05$ ), maka hipotesis nol ditolak. Kita dapat mengacu  $p < 0,05$  sebagai kriteria untuk memutuskan penolakan hipotesis nol, meskipun dicatat bahwa ketika  $p = 0,05$ , keputusan ini juga dapat menolak hipotesis nol. Ketika  $p \text{ value}$  lebih besar dari 5% ( $p > 0,05$ ), maka kita mempertahankan hipotesis nol. Keputusan untuk menolak atau mempertahankan hipotesis nol disebut signifikansi. Ketika nilai  $p$  kurang dari 0,05, kita mencapai signifikansi; keputusannya adalah hipotesis nol ditolak. Ketika nilai  $p$  lebih besar dari 0,05, kita gagal untuk mencapai signifikansi; keputusannya adalah hipotesis nol dipertahankan.

Menggunakan contoh pelemparan koin di atas, kita mengharapkan kepala muncul sebanyak 95 kali atau lebih dari 100 kali melempar atau dalam istilah statistik, kita katakan bahwa hasil pengamatan secara statistik signifikan jika probabilitas perbedaan terjadinya kesalahan kurang dari lima kali dari seratus kali pengamatan ( $p \text{ value} < 0.05$ ).

Tingkat signifikansi dengan demikian dapat dinyatakan sebagai kemungkinan melakukan kesalahan tipe I, yaitu kesalahan menolak hipotesis nol ketika hipotesis tersebut benar. Tingkat signifikansi ini tergantung pada tingkat kepentingan penelitian

itu sendiri. Sementara arti dari tingkat kepercayaan sebesar 95% adalah tingkat dimana 95% nilai sample akan mewakili nilai populasi dimana sample berasal.

### Critical Value

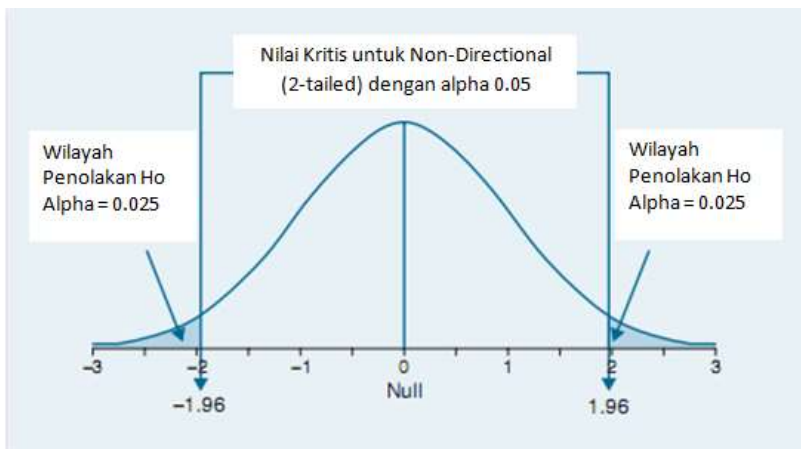
Nilai kritis adalah nilai cut-off yang mendefinisikan batas-batas luar yang kurang dari 5% dari sampel sehingga dapat disimpulkan jika hipotesis nol adalah benar. Dengan demikian, sampel yang diperoleh di luar nilai kritis akan menghasilkan keputusan untuk menolak hipotesis nol.

Dalam pengujian hipotesis non-directional (tidak diputuskan apakah hipotesis berada di sisi kanan atau kiri), maka nilai alpha dapat dibagi dua sehingga memberikan proporsi yang sama dari daerah ekor atas dan bawah (upper and lower tail). Misalnya  $\alpha = 0,05$ , sehingga probabilitas ini dapat dibagi dua menjadi  $0,05 / 2 = 0,025$  untuk masing-masing ekor.

### Power

Kekuatan dari pengujian hipotesis statistik merupakan kemampuan tes untuk menolak hipotesis nol ketika itu sebenarnya. Dengan kata lain, kekuatan uji hipotesis adalah probabilitas untuk tidak melakukan kesalahan tipe II. Hal ini dihitung dengan mengurangi probabilitas kesalahan tipe II dengan 1, biasanya dinyatakan sebagai :  
 $\text{Power} = 1 - P(\text{error tipe II}) = (1 - \beta)$

Maksimum power adalah 1, minimum adalah 0. Idealnya dalam penelitian ingin memiliki kekuatan tinggi yaitu mendekati 1.



Gambar 1. Nilai-nilai kritis ( $\pm 1,96$ ) untuk uji non-directional (dua-tailed) dengan tingkat signifikansi 0,05 (z test)

Empat tahapan dalam pengujian hipotesis :

- Langkah 1: Merumuskan hipotesis
- Langkah 2: Menetapkan kriteria untuk keputusan.
- Langkah 3: Melakukan uji statistik.
- Langkah 4: Membuat keputusan.

Contoh Uji Beda Rata-rata 1 Kelompok:

Seorang dosen mengklaim bahwa "Indeks Prestasi Kumulatif (IPK) mahasiswa di Universitas XYZ lebih besar dari 3.00". Untuk membuktikan hal ini kemudian diambil sampel acak sebanyak 30 mahasiswa ( $n = 30$ ), dan degree of freedom ( $df = n-1 = 30 - 1 = 29$ ). Selanjutnya tingkat signifikansi alpha ditetapkan sebesar 0,05, sehingga hanya memiliki 5% kesempatan untuk membuat kesalahan Tipe I. Dengan tingkat signifikansi sebesar 5% atau 0.05 maka akan diperoleh nilai t hitung sebesar 2.045<sup>2</sup>.

### Rumusan Hipotesis

$H_0 : IPK \leq 3.00$

$H_1 : IPK > 3.00$

### Hipotesis statistik

$H_0 : \mu_0 \leq 3.00$

$H_1 : \mu_0 > 3.00$

### Kriteria uji :

Jika - t tabel  $\leq$  t hitung maka  $H_0$  diterima, dan  $H_1$  di tolak

Jika - t tabel  $>$  t hitung maka  $H_0$  ditolak, dan  $H_1$  diterima atau,

Jika - p value  $\leq$  0.05 maka  $H_0$  diterima, dan  $H_1$  di tolak

Jika - p value  $>$  0.05 maka  $H_0$  ditolak, dan  $H_1$  diterima

### Hasil analisis dan keputusan

Hasil analisis menghasilkan nilai rata-rata sebesar 3.0643, dengan t statistik sebesar 2.144. Karena t statistic  $>$  t tabel ( $2.144 > 2.045$ ) dan p value  $<$  a ( $0.041 < 0,05$ ), kita dapat menolak hipotesis nol bahwa IPK mahasiswa kurang dari atau sama dengan 3.00.<sup>3</sup>

### Teknik Analisis Data

Menentukan teknik analisis merupakan sebuah proses yang terintegrasi dalam prosedur penelitian. Analisis data dilakukan untuk menjawab rumusan masalah dan hipotesis yang sudah diajukan. Hasil analisis data selanjutnya diinterpretasikan dan dibuatkan kesimpulannya. Pilihan teknik statistik adalah masalah yang kompleks, didasarkan pada pertimbangan berbagai hal. Oleh karena itu, pada dua tabel berikut akan diberikan pedoman penentuan teknik analisis yang tepat sebagai pedoman umum.

### Metode interdependen

Metode ini bertujuan untuk mengelompokkan data/faktor atau variabel yang ada menjadi variabel independen atau dependen. Beberapa teknik dalam metode interdependen meliputi cluster analysis, analisis faktor, Multidimensional Scaling (MDS) dan analisis korespondensi

<sup>2</sup> Untuk menghasilkan t tabel dengan Excel dapat menggunakan formula dengan format " $=TIN(p,d)$ ", p adalah 0.05 dan d adalah 29, sehingga formula ditulis " $=TINV(0.05,29)$ "

<sup>3</sup> Contoh Perhitungan bisa di lihat pada File "01 Nilai IPK"

Tabel 3. Pedoman Penentuan Teknik Analisis Metode Interdependen

Jumlah Variabel	Level Pengukuran	Teknik Analisis
1	Nominal	Frequencies, Proportions
1	Ordinal	Median, Mode
1	Interval atau ratio scale	Mean, Median, Mode, Variance, Skewness, Kurtosis
2	Dikotomi	Cross-tabulation Chi-square
2	Nominal	Cross tabulation, Chi-square, Correspondence Analysis
2	Ordinal	Kendall's <i>Tau</i> , Spearman's <i>Rho</i> , Gamma
2	Interval-scale	Scatter plot, Pearson's Correlation Coefficient
> 2	Interval-scale	Principal Components Analysis, Factor Analysis, Cluster Analysis Multidimensional Scaling

### Metode Asosiatif (Hubungan dan Pengaruh)

Beberapa teknik analisis yang umum digunakan pada pendekatan metode asosiatif ditampilkan berikut ini :

Tabel 4. Pedoman Penentuan Teknik Analisis Metode Asosiatif

Jumlah variabel		Level Pengukuran		Teknik Analisis
<i>Var. Terikat</i>	<i>Var. Bebas</i>	<i>Var. Terikat</i>	<i>Var. Bebas</i>	
1	1	Nominal	Nominal	Non-parametric tests, Chi-square
1	1	Nominal	Nominal	Wilcoxon's two sample test, Chi-square, Kolmogorov-Smirnov Test
1	1	Interval-scale	Nominal (dikotomi)	t-test, Analysis of Variance
1	1	Interval-scale	Interval-scale	Simple Regression Analysis; Pearson Product Moment Correlation
1	1	Interval-scale	Nominal	Analysis of Variance
1	> 1	Nominal	Interval-scale atau campuran	Discriminant Analysis; Logistic regression
1	> 1	Interval-scale	Nominal	Analysis of Variance, Multiple Regression Analysis
1	> 1	Interval scale	Dummy	Analysis of Variance, Multiple Regression Analysis
1	> 1	Interval-scale	Interval-scale	Multiple Regression Analysis

### Metode Komparatif (perbandingan)

Selain menguji adanya hubungan / korelasi pada variabel penelitian, beberapa teknik statistik inferensial juga dapat dilakukan pada model penelitian komparatif seperti ditampilkan pada Tabel berikut ini

Tabel 5. Pedoman Penentuan Teknik Analisis Metode Komparasi

Jumlah kelompok sampel	Level Pengukuran	Teknik Analisis
1 sampel	Interval-Rasio	One-sample t test
	Nominal; Ordinal	Binominal test; Chi-Square Goodness-of-fit Test; One sample Runs Test of Randomness; Kolmogorov-Smirnov one sample test
2-related Sample	Interval-Rasio	Paired- t test
	Nominal; Ordinal	McNemar Change Test; Sign Test; Wilcoxon Signed Rank Test
k-related sample	Interval-Rasio	Repeated Measure ANOVA
	Nominal; Ordinal	Cohran Q Test; Friedman Two-Way Analysis of variance
2-Independen Sample	Interval-Rasio	Independent t test
	Nominal; Ordinal	Median test; Wilcoxon - Mann-Whitney Test
k-Independen Sample	Interval-Rasio	Analysis of Variance (ANOVA, MANOVA)
	Nominal; Ordinal	Perluasan Median Test; Kruskal-Wallis One-Way Analysis of Variance by Rank

#### Referensi :

- Bartlett, J.E., Kotrlík, J.W., Higgins, C.C. (2001). Organizational Research: Determining Appropriate Sample Size in Survey Research. *Information Technology, Learning, and Performance Journal*, Vol. 19, No. 1, Spring 2001
- Sedarmayanti, dan Hidayat, S. (2011). *Metode Penelitian*. Bandung : Mandar Maju
- Sekaran, U. (2003). *Research Methods for Business. A Skill Building Approach*. Fourth Ed. NY : John Willey & Son
- Suryani dan Hendryadi. (2015). *Metode Riset Kuantitatif*. Jakarta, Prenadamedia Group.

## PERTEMUAN 2 PENGENALAN SPSS

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami dan menjelaskan tentang menu dasar SPSS dan menggunakannya dalam menganalisis data.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Lembar kerja dan Menu SPSS
  - b. Manajemen data SPSS

### 5. Materi

IBM SPSS, atau Statistics Package For Social Science (yang kemudian juga diganti dengan istilah Statistical Product and Service Solutions) merupakan salah satu program yang paling banyak digunakan dan tersedia untuk personal computer (PC). Program aplikasi ini dapat mengambil data dari hampir semua jenis file dan menggunakannya untuk menghasilkan laporan ringkasan, grafik, statistik deskriptif, dan analisis statistik yang kompleks.

SPSS saat ini dipergunakan berbagai disiplin ilmu, seperti psikologi, kesehatan, pendidikan, dan lain sebagainya. SPSS membantu pengguna dalam menggambarkan data, pengujian hipotesis dan mencari korelasi atau hubungan antara satu atau lebih variabel. Menurut pengamatan di lapangan, buku-buku yang membahas mengenai analisis data sekarang ini di Indonesia mayoritas menggunakan alat bantu perangkat lunak IBM SPSS yang menunjukkan eksistensi perangkat lunak ini di dunia pendidikan dan penelitian di Indonesia. Versi terakhir IBM SPSS tahun 2014 adalah versi 22.

Beberapa fitur IBM SPSS antara lain IBM SPSS Data Collection (untuk menangkap pandangan yang akurat tentang sikap masyarakat, preferensi dan pendapat); IBM SPSS Statistic (Memprediksi: analisis statistik canggih untuk memahami data, mengidentifikasi tren dan menghasilkan perkiraan yang akurat); IBM SPSS Modeler (Prediksi: menemukan pola dan tren dalam data terstruktur dan tidak terstruktur melalui antarmuka visual yang intuitif); IBM Analytical Decision Management (digunakan untuk pengambilan keputusan)<sup>4</sup>. Dalam buku ini, program SPSS yang dipergunakan adalah versi 19.0

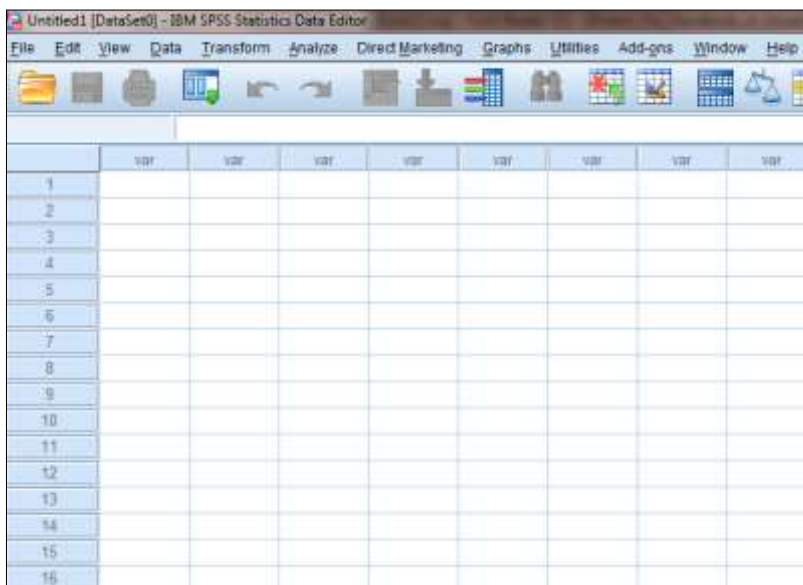
Ada banyak cara untuk memulai SPSS. Cara termudah adalah dengan memulainya dari tombol Start yang terletak di bawah desktop Windows. Klik tombol Start, kemudian klik All Programs, pilih SPSS Inc, dan akhirnya IBM SPSS Statistics 19. Berikut jendela awal dialog terbuka untuk ketika memulai SPSS.

---

<sup>4</sup> Penjelasan lengkap mengenai fitur SPSS ini dapat dilihat pada website resmi SPSS di <http://www-01.ibm.com/software/analytics/spss/>



Karena tujuan dari latihan ini adalah untuk membuat data baru, maka dapat mengklik tombol “Cancel” sehingga akan menghasilkan jendela Untitled1 [DataSet0] sebagai berikut :



## Manajemen Data

### Mempersiapkan Codebook

Sebelum entri data, terlebih dahulu dilakukan persiapan codebook yang berisi nama-nama variabel dalam kuesioner, untuk disesuaikan dengan nama variabel pada menu “variable view” yang terletak di sisi kiri bawah lembar kerja. Informasi ini berisi nama variabel, dan instruksi coding. Tujuan penting codebook tersebut adalah untuk memungkinkan peneliti untuk melacak semua variabel dalam kuesioner survei dan menyusunnya ke dalam lembar kerja SPSS. Contoh berikut ini merupakan codebook pada kuesioner survey kepuasan pengunjung supermarket.

Tabel 2.1. Codebook

variabel	SPSS Variable Name	Kode
Jenis Kelamin	Gender	0 = laki-laki 1 = wanita
Kategori Usia	Age	1 = 18 - 24 th 2 = 25 - 30 th 3 = 31 - 35 th 4 = 36 - 40 th 5 = > 40 tahun
Pendidikan	Educ	0 = tidak berpendidikan 1 = SMA 2 = DIII/S1 3 = S2 4 = S3
harga	Harga	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju
kelengkapan produk	Produk	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju
kemudahan akses	Akses	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju
pembayaran	Bayar	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju
Keramahan petugas	Ramah	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju
Keamanan	aman	1 = Sangat tidak setuju 2 = Tidak setuju 3 = Netral 4 = Setuju 5 = Sangat Setuju



Raw data pada program Excel

	A	B	C	D	E	F	G	H	I	J
1	no	gender	age	educ	harga	produk	akses	bayar	ramah	aman
2	1	0	3	1	4	5	1	2	4	1
3	2	0	1	2	4	5	4	1	2	4
4	3	1	3	1	4	2	2	2	3	3
5	4	0	3	1	3	2	2	1	3	2
6	5	1	4	2	2	1	3	2	4	2
7	6	0	3	2	2	4	4	2	1	3
8	7	1	4	3	2	3	3	3	2	4
9	8	0	3	3	5	5	5	4	4	4
10	9	1	4	3	3	2	3	4	2	2
11	10	0	3	1	3	4	4	5	5	4
12	11	0	3	0	3	5	5	5	4	4
13	12	0	4	2	5	5	3	5	3	5
14	13	1	4	1	1	2	3	1	1	1
15	14	0	3	2	4	1	3	1	3	2

### Setting Data

Sebelum entri data, variabel di set data harus diberi nama dan didefinisikan. Pada layar Untitled1 [DataSet0] - IBM SPSS Statistik Data Editor. Untuk melakukan setting data, klik pada pilihan "variable view" pada bagian kiri bawah lembar kerja SPSS.

	Name	Type	Width	Decimals	Label	Values	Missing
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							

Pada tampilan awal akan tampil kolom-kolom pilihan (Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, Role).

Tabel 2.2. Keterangan pada Tampilan "Variable View"

Nama Kolom	Keterangan
Name	<ul style="list-style-type: none"> <li>• Berfungsi untuk memberikan nama unik variabel.</li> <li>• Jumlah karakter/huruf maksimal adalah 64 karakter.</li> <li>• Tidak boleh ada tanda aritmatik (+, -, /, *)</li> <li>• Tidak boleh ada spasi, jika nama variabel terdiri dari dua kata maka gunakan tanda " _ " (underscore)</li> <li>• Karakter tidak boleh diakhiri dengan titik (.)</li> <li>• Hindari penamaan variabel dengan istilah-istilah yang digunakan oleh SPSS seperti : ALL, BY, EQ, NOT, OR, TO, WITH, dan lainnya</li> </ul>
Type	<ul style="list-style-type: none"> <li>• Numeric : digunakan data dalam bentuk angka.</li> <li>• Comma : Data ini berbentuk numerik, dianggap valid bila dalam bentuk angka, bisa bertanda plus maupun minus di depan angka, ditandai koma sebagai pembeda ribuan.</li> <li>• Dot : Data ini berbentuk numerik, dianggap valid bila dalam bentuk angka, bisa bertanda plus maupun minus di depan angka, ditandai titik sebagai pembeda ribuan.</li> <li>• Scientific Notation : berbentuk numerik, dianggap valid bila dalam bentuk angka, bisa bertanda plus maupun minus, ditandai dengan simbol E.</li> <li>• Date : digunakan untuk memasukkan data tanggal.</li> <li>• Dollar : Data ini berbentuk numerik, dianggap valid bila dalam bentuk angka, yang ditandai (\$) dengan tanda koma sebagai pemisah tanda ribuan.</li> <li>• Custom Currency : Bentuk tipe ini digunakan untuk menampilkan format mata uang yang dibuat melalui kotak dialog Options dari menu edit.</li> <li>• String : data dalam bentuk karakter/huruf</li> </ul>
Width	Adalah banyaknya karakter dari data yang akan dimasukkan.
Decimal	Decimals adalah untuk menentukan berapa angka desimal setelah tanda koma
Label	Label merupakan penjelasan pada kolom Name.
Values	Merupakan kode yang diberikan jika variabel yang digunakan kategorial. Dalam contoh ini, kita memberikan kode 0 untuk laki-laki dan kode 1 untuk wanita
Missing	Data yang dientri yang dianggap hilang yang tidak diikuti dalam analisis. Ada tiga pilihan :No missing values (tidak ada data yang dianggap hilang); Discrete missing values (angka tertentu yang kita anggap hilang: dan Range plus one optional discrete missing value ( jika ada data hilang dengan range tertentu).
Columns	Adalah lebar tempat nama karakter dari NAME Besarnya adalah minimal sama dengan nilai di WIDTH
Align	Sama seperti Align di Microsoft Word/Excel. Ada berbagai pilihan tampilan yaitu Left/Centre/Right
Measure	Adalah skala pengukuran dari variabel yang di entri. Ada tiga pilihan pengukuran yaitu nominal, ordinal dan scale (untuk data rasio dan interval).

Berdasarkan codebook pada Tabel 2.1, maka kita dapat melakukan setting pada data seperti tampilan berikut<sup>5</sup> :

	Name	Type	Width	Decimals	Label	Values
1	gender	Numeric	4	0	jenis kelamin	{0, laki-laki}...
2	age	Numeric	4	0	ketegori usia	{1, 18-24}...
3	educ	Numeric	4	0	pendidikan	{1, SMA}...
4	harga	Numeric	4	0	kepuasan pada harga	{1, Sangat ti...
5	produk	Numeric	4	0	kepuasan pada kelengkapan produk	{1, Sangat ti...
6	akses	Numeric	4	0	kepuasan pada kemudahan akses	{1, Sangat ti...
7	bayar	Numeric	4	0	kepuasan pada metode pembayaran	{1, Sangat ti...
8	ramah	Numeric	4	0	kepuasan pada keramahan	{1, Sangat ti...
9	aman	Numeric	4	0	kepuasan pada keamanan	{1, Sangat ti...

### Data Entry

Untuk meng-entry data, kita dapat aktifkan pilihan “Data Views” pada bagian bawah kiri lembar kerja SPSS. Entry data dapat dilakukan secara manual atau langsung dicopy dari file data Excel yang sebelumnya sudah dientry.

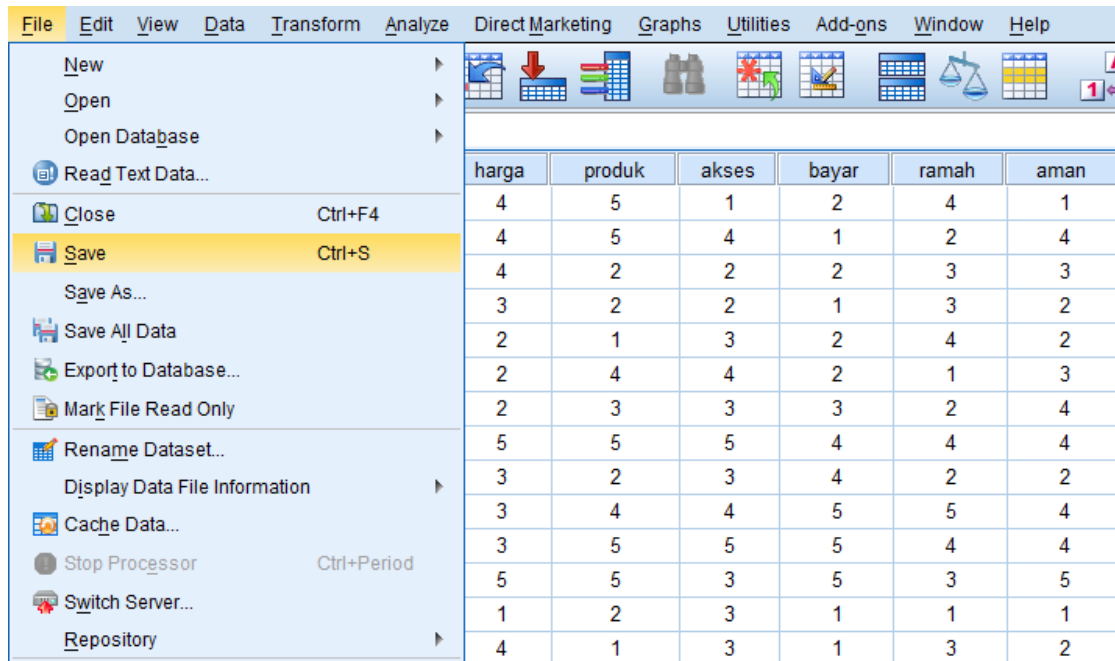
Hasil Copy data dari Excel

	gender	age	educ	harga	produk	akses	bayar	ramah	aman
1	0	3	1	4	5	1	2	4	1
2	0	1	2	4	5	4	1	2	4
3	1	3	1	4	2	2	2	3	3
4	0	3	1	3	2	2	1	3	2
5	1	4	2	2	1	3	2	4	2
6	0	3	2	2	4	4	2	1	3
7	1	4	3	2	3	3	3	2	4
8	0	3	3	5	5	5	4	4	4
9	1	4	3	3	2	3	4	2	2
10	0	3	1	3	4	4	5	5	4
11	0	3	0	3	5	5	5	4	4
12	0	4	2	5	5	3	5	3	5
13	1	4	1	1	2	3	1	1	1
14	0	3	2	4	1	3	1	3	2
15	1	3	1	2	1	3	2	1	1
16	0	3	3	2	1	1	1	1	1
17	1	3	4	5	5	1	5	5	4

<sup>5</sup> buka file “02 satis.SAV” untuk melihat setting data secara lengkap

### Menyimpan Data

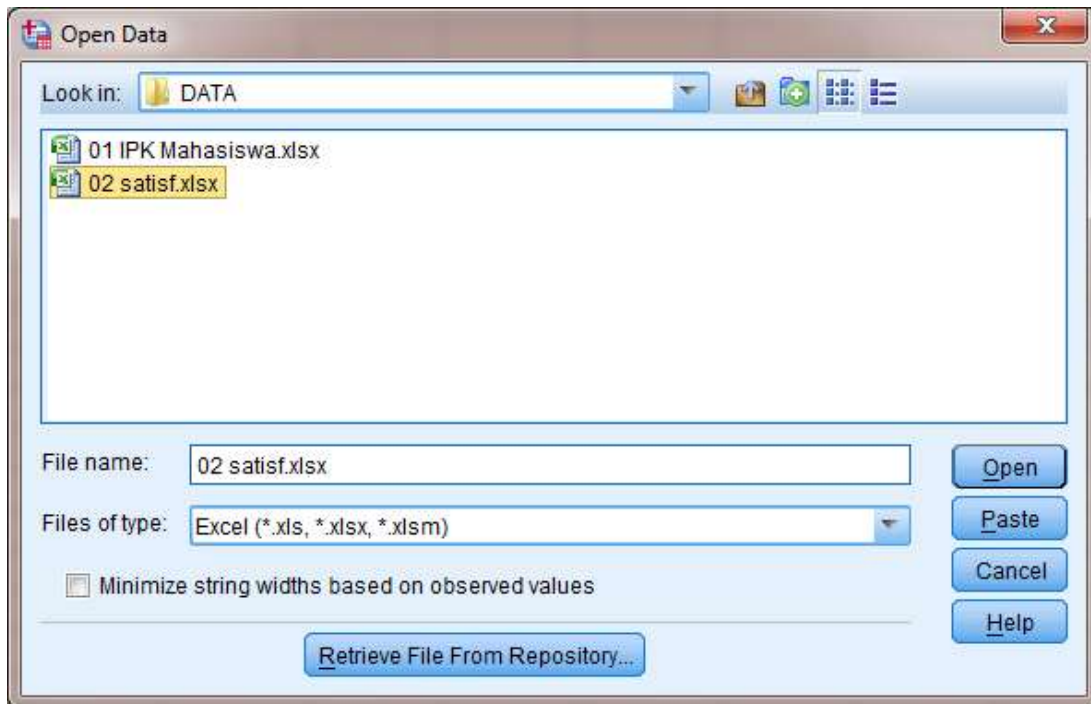
Untuk menyimpan data, klik pada pilihan File – Save. Tentukan lokasi penyimpanan kemudian ketik nama file pada opsi “File Name”, abaikan yang lain kemudian klik “SAVE”.



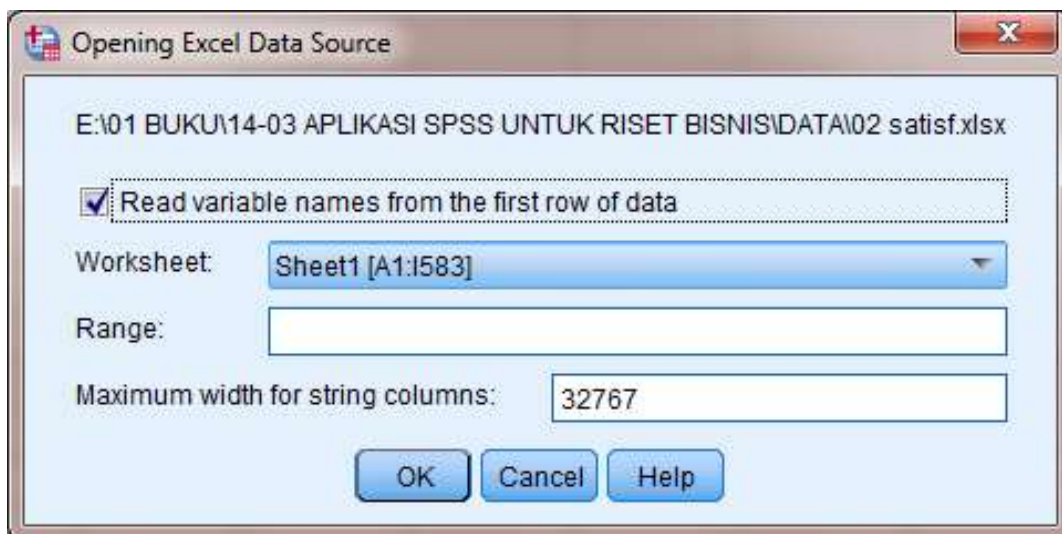
### Import Data

SPSS secara umum dapat mengimport data dari berbagai jenis file, termasuk Excel (.xls dan .xlsx); csv, dbase, SAS dan lainnya. Dalam contoh kali ini, import data dilakukan pada dari file excel versi 2007.

1. Klik File > Open > Data.
2. Dalam Tipe file pilih "Excel (\*.xls, \*.xlsx, \*.xlsm)" untuk menentukan bahwa data yang digunakan adalah data file Excel. Cari dan klik pada tipe file. Nama file akan muncul di bidang File name. Klik “open”.

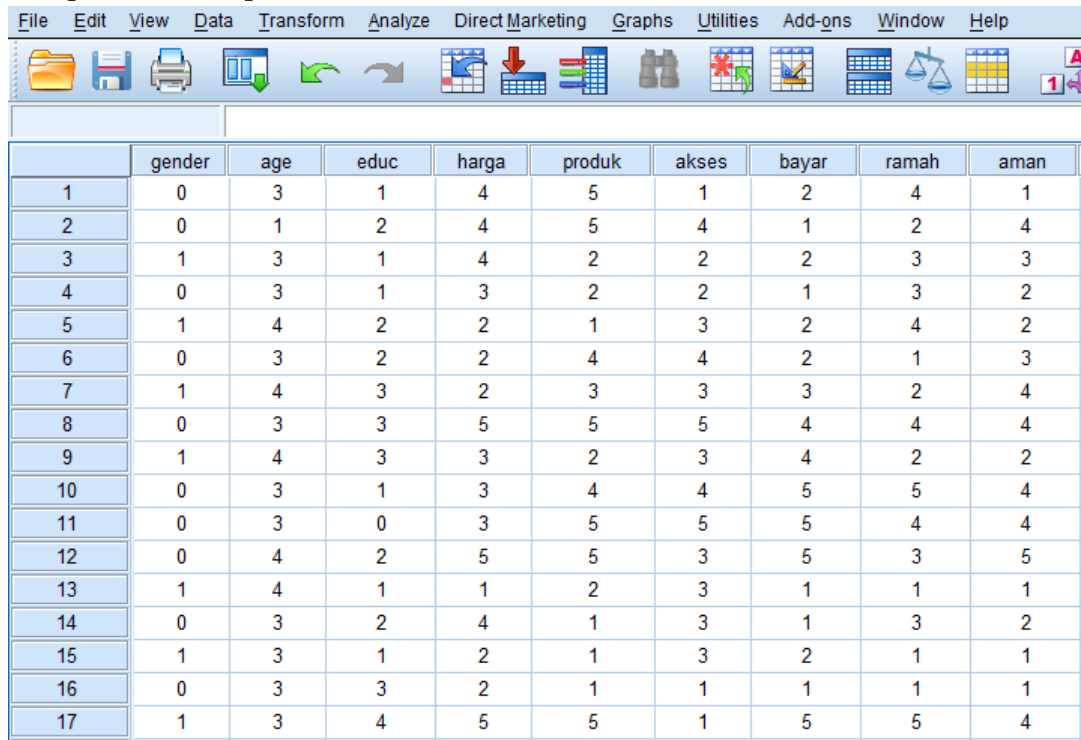


3. Pembukaan "opening excel data source" akan tampil.



Klik OK

## 4. Tampilan hasil import



	gender	age	educ	harga	produk	akses	bayar	ramah	aman
1	0	3	1	4	5	1	2	4	1
2	0	1	2	4	5	4	1	2	4
3	1	3	1	4	2	2	2	3	3
4	0	3	1	3	2	2	1	3	2
5	1	4	2	2	1	3	2	4	2
6	0	3	2	2	4	4	2	1	3
7	1	4	3	2	3	3	3	2	4
8	0	3	3	5	5	5	4	4	4
9	1	4	3	3	2	3	4	2	2
10	0	3	1	3	4	4	5	5	4
11	0	3	0	3	5	5	5	4	4
12	0	4	2	5	5	3	5	3	5
13	1	4	1	1	2	3	1	1	1
14	0	3	2	4	1	3	1	3	2
15	1	3	1	2	1	3	2	1	1
16	0	3	3	2	1	1	1	1	1
17	1	3	4	5	5	1	5	5	4

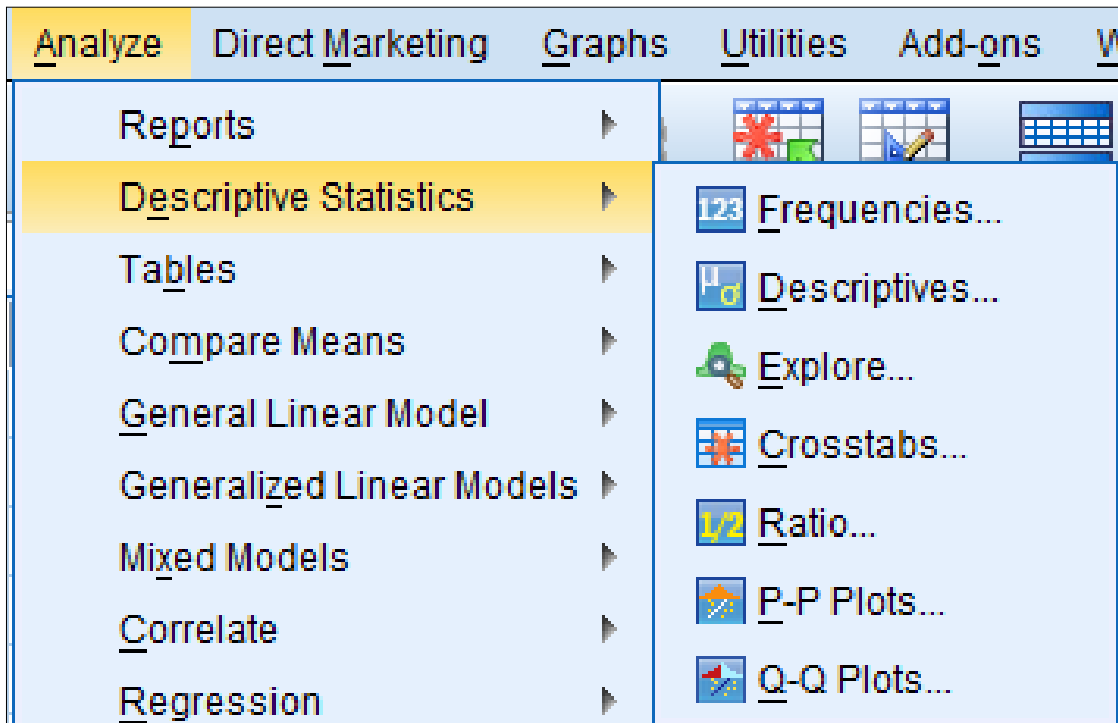
## 5. Simpan data

### PERTEMUAN 3 DESKRIPSI DATA

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami dan menjelaskan deskripsi data.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data deskriptif menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Analisis Frekuensi
  - b. Analisis Deskriptif
  - c. Analisis Explore

#### 5. Materi

Statistik deskriptif secara umum digunakan untuk menggambarkan fitur dasar dari data dalam sebuah penelitian. Teknik ini menyediakan ringkasan sederhana tentang sampel dan pengukurannya. Statistik deskriptif berbeda dari statistik inferensial. Dengan statistik deskriptif peneliti hanya menggambarkan apa yang data tunjukkan, sedangkan pada statistik inferensial, peneliti mencoba untuk mengambil kesimpulan berdasarkan data yang dikumpulkan.



Gambar 3.1. Menu Analisis Deskriptif

SPSS memberikan beberapa prosedur dalam statistik deskriptif seperti frekuensi, deskriptif, explore, crosstab, ratio, P-P Plots dan Q-Q Plots. Uraian mengenai prosedur analisis tersebut akan dijelaskan satu persatu dalam sub bab 4 ini.

### A. Analisis Frequencies

Analisis frekuensi menyediakan data statistik dan tampilan grafis yang berguna untuk menggambarkan berbagai jenis variabel. Frekuensi adalah prosedur awal yang berguna untuk melihat data yang digunakan dalam penelitian. Contoh berikut ini akan menampilkan prosedur analisis frekuensi pada data karyawan di beberapa perusahaan di tiga sektor yaitu perbankan, manufaktur, dan pertambangan. Total responden pada survey ini adalah sebanyak 400 orang pegawai. File latihan tersimpan di CD latihan dengan nama "employee.SAV".

	id	gender	usia	sektor	educ	salary
1	1	.00	30.0000	1	1	7475
2	2	.00	30.0000	1	1	7088
3	3	1.00	35.0000	1	2	6938
4	4	.00	47.0000	2	1	8775
5	5	1.00	33.0000	2	1	6825
6	6	.00	31.0000	2	2	10200
7	7	1.00	33.0000	2	2	11500
8	8	.00	40.0000	3	3	24688
9	9	1.00	33.0000	1	3	11825
10	10	.00	33.0000	2	3	10963
11	11	.00	55.0000	2	1	5438
12	12	.00	31.0000	1	0	6000
13	13	1.00	53.0000	1	2	7288
14	14	.00	29.0000	1	1	7763
15	15	1.00	42.0000	3	2	20094
16	16	.00	33.0000	2	1	8138
17	17	1.00	34.0000	3	3	15050
18	18	1.00	32.0000	3	4	26538

Gambar 3.2. Data Karyawan

### Analisis Frekuensi pada Data Kategori

Analisis frekuensi untuk data kategori digunakan untuk menggambarkan data dalam bentuk jumlah atau frekuensi dan proporsi. Contoh berikut ini adalah data frekuensi pada gender, sektor, dan pendidikan.

Tahapan Penyelesaian

- Buka file pada CD Latihan, kemudian pilih file "employee.SAV"

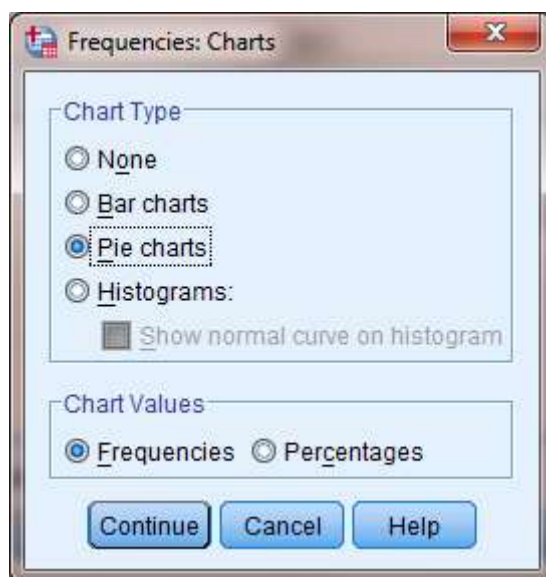


- Klik Analyze → Descriptive Statistics → Frequencies pada menu sehingga akan muncul kotak dialog frequencies
- Masukkan variabel gender, sektor dan pendidikan ke kotak “variables”.



Gambar 3.3. Kotak Dialog Frequencies

- Klik tombol “Charts” sehingga akan tampil menu pilihan “Charts”
- Pilih “Pie Charts” pada pilihan Chart Type dan pilih Frequencies pada kotak Chart Values.



Gambar 4.4. Pilihan Chart

- Klik Continue

- Klik OK, sehingga akan tampil output SPSS berikut ini :

**Statistics**

		gender	sektor	educ
N	Valid	400	400	400
	Missing	0	0	0

Tabel statistics memberikan informasi mengenai jumlah data secara umum yaitu sebanyak 400 dan seluruhnya valid (tidak terdapat data missing).

**gender**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Wanita	172	43.0	43.0	43.0
	Laki-Laki	228	57.0	57.0	100.0
Total		400	100.0	100.0	

Pada tabel Gender memperlihatkan informasi mengenai jumlah responden berdasarkan gender yaitu wanita sebanyak 172 orang (43%) dan laki-laki sebanyak 228 orang (57%).

**sektor**

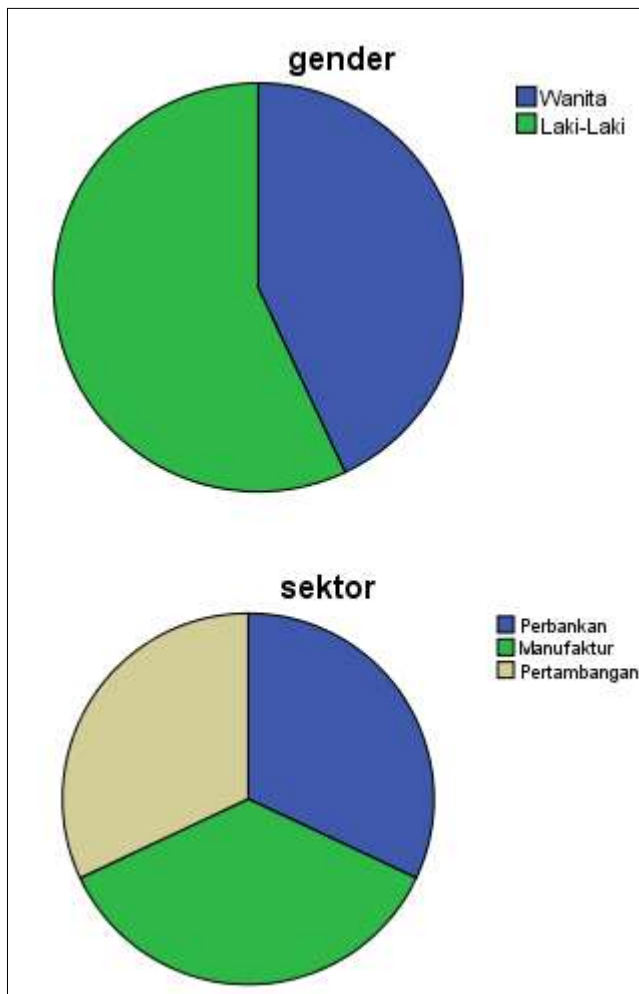
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Perbankan	128	32.0	32.0	32.0
	Manufaktur	144	36.0	36.0	68.0
	Pertambangan	128	32.0	32.0	100.0
Total		400	100.0	100.0	

Tabel sektor menampilkan informasi mengenai pengelompokan responden berdasarkan sektor pekerjaan, yaitu 128 orang (32%) bekerja di sektor perbankan, 144 orang (36%) bekerja di sektor manufaktur, dan 128 orang (32%) bekerja di sektor pertambangan.

**educ**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	SMA/SMK	37	9.3	9.3	9.3
	Diploma	104	26.0	26.0	35.3
	S1	140	35.0	35.0	70.3
	S2	98	24.5	24.5	94.8
	S3	21	5.3	5.3	100.0
	Total	400	100.0	100.0	

Klasifikasi responden berdasarkan pendidikan memperlihatkan sebanyak 37 orang (9.3%) masih berpendidikan SMA sederajat, 104 orang (26%) berpendidikan Diploma, 140 orang (35%) berpendidikan S1, 98 orang (24.5%) berpendidikan S2, dan 21 orang (5.3%) berpendidikan S3.



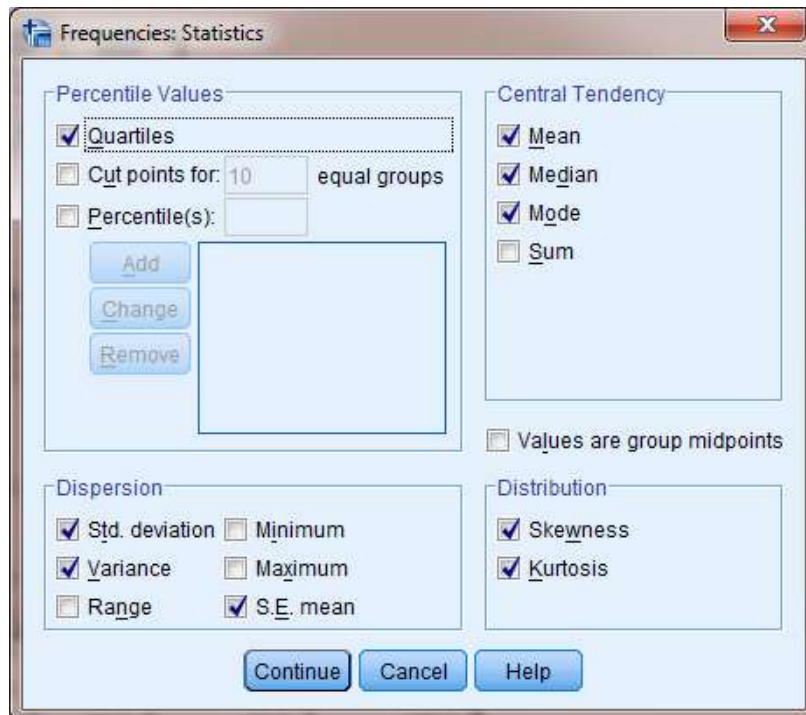
Gambar 3.5. Tampilan Pie Chart pada Frekuensi Data Kategori

### Analisis Frekuensi pada Data Skala

Analisis frekuensi untuk data skala digunakan untuk menggambarkan data dalam bentuk sentral tendensi, dispersi dan distribusi. Contoh berikut ini adalah data frekuensi pada salary.

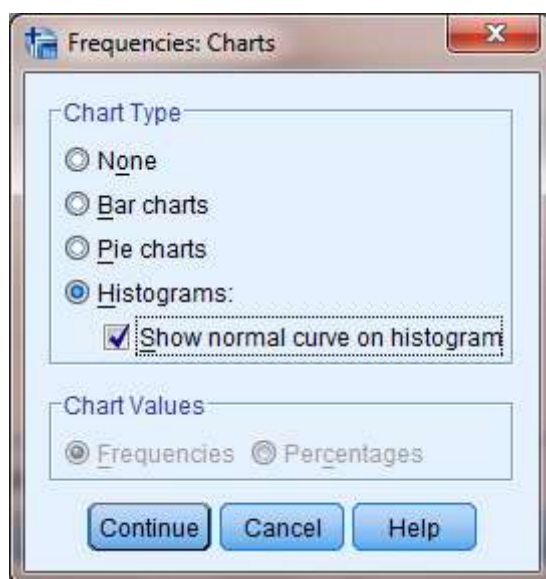
#### Tahapan Penyelesaian

- Buka file pada CD Latihan, kemudian pilih file employee.SAV
- Klik Analyze → Descriptive Statistics → Frequencies pada menu sehingga akan muncul kotak dialog frequencies
- Masukkan variabel salary ke kotak "variables".
- Klik tombol "statistics" sehingga akan muncul kotak dialog sebagai berikut :



Gambar 3.6. Kotak Dialog Frequencies : statistics

- Pilih nilai-nilai pada “Percentile Values, Central Tendency, Dispersion, dan Distribution”
- Klik Tombol Continue
- Klik Tombol Chart, kemudian pilih “Histogram with normal curve” pada kotak Chart Type



Gambar 3.7. Kotak Dialog Frequencies : Charts

- Klik Continue
- Klik OK

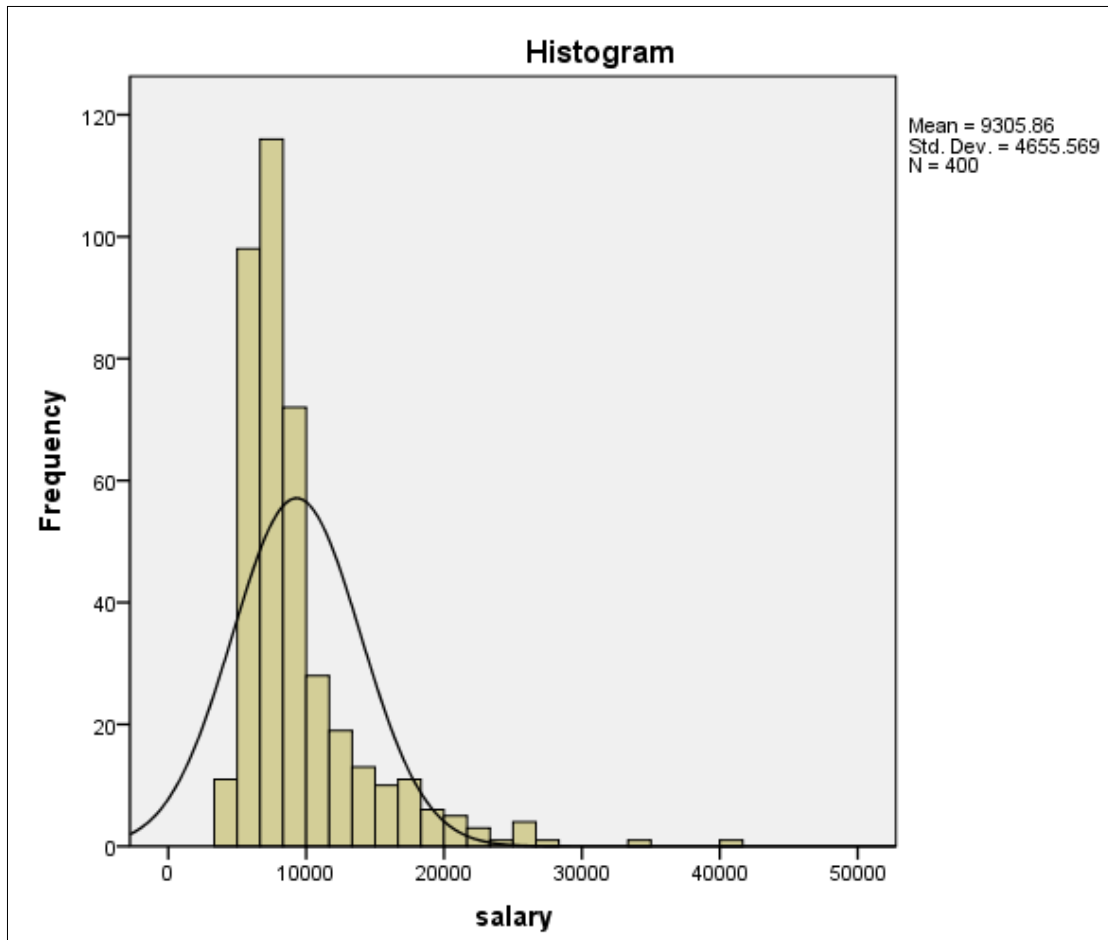
## Frequencies

Statistics		
salary		
N	Valid	400
	Missing	0
Mean		9305.86
Std. Error of Mean		232.778
Median		7838.00
Mode		6000 <sup>a</sup>
Std. Deviation		4655.569
Variance		21674318.51
Skewness		2.498
Std. Error of Skewness		.122
Kurtosis		8.859
Std. Error of Kurtosis		.243
Percentiles	25	6563.00
	50	7838.00
	75	10050.00

a. Multiple modes exist. The smallest value is shown

Tampilan tabel di atas memberikan informasi penting mengenai :

- N, jumlah data pengamatan = 400 (tanpa data missing)
- Mean, rata-rata salary = 9,305 Juta Rupiah
- Median, nilai tengah = 7,838 Juta Rupiah
- Mode (Modus), nilai yang sering muncul = 6,000 Juta Rupiah
- Standar Deviasi = 4,655 Juta Rupiah
- Skewness = 2,498 Juta Rupiah (skewness positif mengindikasikan data memiliki kemencengan ke kiri)
- Percentiles 25, tingkat salary 25% di bawah 6,563 juta rupiah
- Percentiles 50, tingkat salary 50% di bawah 7,838 juta rupiah
- Percentiles 75, tingkat salary 75% di bawah 10,050 juta rupiah



Gambar 3.8. Hasil analisis frequencies : Histogram

Tampilan grafik di atas menunjukkan data tidak normal, atau menceng ke kiri (positive skewness)

## B. Descriptives

Analisis deskriptif merupakan analisis yang dapat dipergunakan untuk memberikan rangkuman data dalam berbagai bentuk ukuran. Prosedur ini menampilkan ringkasan statistik univariat untuk beberapa variabel dalam satu tabel dan menghitung nilai standar (z skor). Ketika skor z masukkan, mereka ditambahkan ke data dalam data Editor dan tersedia untuk grafik, daftar data, dan analisa.

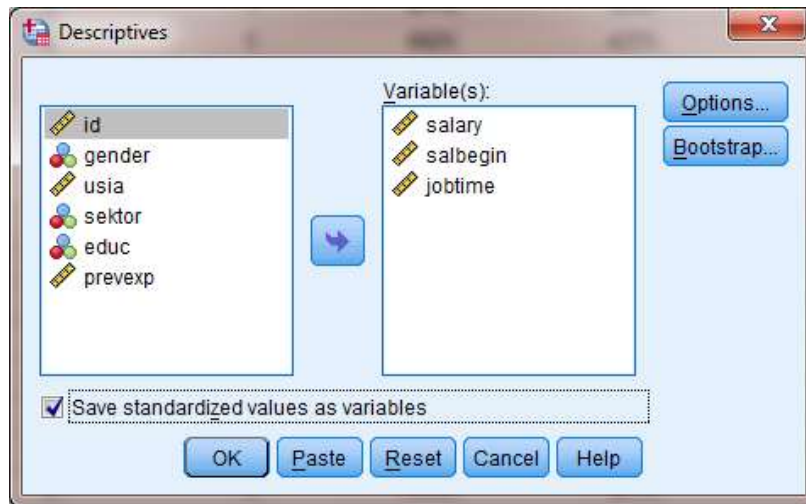
### Pilihan Menu Descriptives

Statistics. Ukuran sampel, rata-rata, minimum, maksimum, standar deviasi, varians, range, jumlah, standar error dari mean, dan kurtosis dan skewness dengan kesalahan standar mereka. Contoh berikut ini adalah data deskripsi data dari salary, salbegin (salary awal bekerja), dan job time (masa kerja dalam bulan).

Tahapan Penyelesaian

- Buka file pada CD Latihan, kemudian pilih file employee.SAV

- Klik Analyze → Descriptive Statistics → Descriptives pada menu sehingga akan muncul kotak dialog Descriptive



Gambar 3.9. Kotak Dialog Descriptives

- Masukkan variabel salary, salbegin dan jobtime ke kotak “variables”.
- Aktifkan pilihan “save standardized value as variables”
- Klik tombol OK sehingga akan muncul output sebagai berikut :

prevexp	Zsalary	Zsalbegin	Zjobtime
0	-.39326	-.49220	1.76679
26	-.47639	-.30245	1.76679
34	-.50861	-.78762	1.76679
137	-.11403	-.57255	1.76679
66	-.53288	-.85109	1.65600
24	.19206	-.72448	1.65600
48	.47129	-.78762	1.65600
70	3.30403	1.76649	1.65600
103	.54110	.64728	1.65600
17	.35595	-.30245	1.65600
315	-.83080	-.91423	1.65600
75	-.71009	-1.05366	1.65600
171	-.43343	-.89329	1.65600
14	-.33140	-.92706	1.54521
96	2.31726	1.34177	1.54521
43	-.25085	-.78762	1.54521

Gambar 3.10. Tampilan Penambahan Variabel Baru “Z Score”

Pada data view, akan tampil tiga variabel baru yaitu Zsalary, Zsalbegin, dan Zjobtime. Tiga variabel ini merupakan hasil penyimpanan dari “save standardized value as variables” yang diaktifkan. Tiga variabel ini menunjukkan penyimpangan data (outlier) dari rata-rata. Nilai kritis untuk nilai Z berdasarkan tingkat kepercayaan 95% atau

tingkat signifikansi 5% (2-tailed) adalah sebesar 1.96 (masing-masing sisi adalah 2.5%). Dengan demikian, konversi nilai Z score akan dikatakan menyimpang jika berada di luar  $-1.96 s/d + 1.96$ .

## Descriptives

	N	Minimum	Maximum	Mean	Std. Deviation
salary	400	3975	41219	9305.86	4655.569
salbegin	400	2125	27344	6895.84	2961.890
jobtime	400	66	98	82.05	9.026
Valid N (listwise)	400				

Tampilan tabel di atas memperlihatkan bahwa besaran gaji minimum adalah 3,975 juta rupiah dan maksimum sebesar 41,219 juta rupiah. Untuk gaji awal, minimum adalah sebesar 2,125 juta rupiah dan maksimum sebesar 27,344 juta rupiah. Masa kerja minimum adalah selama 66 bulan (5.5 tahun) dan maksimum 98 bulan (8.17 tahun).

### C. Explore

Prosedur explore menghasilkan ringkasan statistik dan tampilan grafis, baik untuk kasus terpisah maupun untuk kelompok kasus. Ada banyak alasan untuk menggunakan prosedur explore yaitu digunakan untuk screening data, identifikasi outlier, deskripsi, memeriksa asumsi dan mencirikan perbedaan antara sub-populasi (kelompok kasus).

Screening data dapat menunjukkan nilai-nilai yang tidak biasa, nilai-nilai ekstrim, kesenjangan dalam data, atau keanehan lainnya. Explore data dapat membantu untuk menentukan apakah teknik statistik yang digunakan sesuai dengan karakter data. Eksplorasi dapat menunjukkan bahwa apakah perlu untuk mengubah data jika teknik membutuhkan distribusi normal. Atau mungkin memutuskan metode alternatif menggunakan pendekatan nonparametrik.

### Pilihan Menu Explore

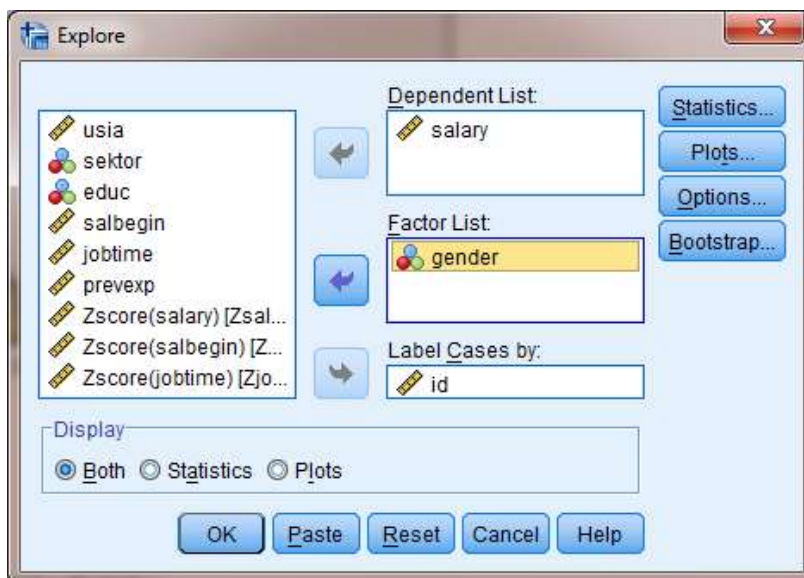
- Descriptives. Terdiri dari sentral tendensi dan dispersi ditampilkan secara default. Ukuran sentral tendensi menunjukkan posisi distribusi data; termasuk mean, median, dan mode. Ukuran dispersi menunjukkan perbedaan dari nilai rata-rata; ini termasuk kesalahan standar, varians, standar deviasi, minimum, maksimum, jangkauan, dan kisaran interkuartil. Statistik deskriptif juga mencakup langkah-langkah dari bentuk distribusi; skewness dan kurtosis ditampilkan dengan kesalahan standar data.
- M-estimator. Robust alternatives untuk mean sampel dan median untuk memperkirakan posisi. Estimator dihitung berbeda dalam bobot mereka berdasarkan kasus tertentu. Huber M-estimator, Andrews "wave estimator", Hampel ini redescending M-estimator, dan estimator biweight Tukey merupakan beberapa pilihan yang dapat digunakan.
- Outliers. Menampilkan lima terbesar dan lima nilai terkecil data dengan label kasus.
- Persentil. Menampilkan nilai-nilai untuk persentil 5, 10, 25, 50, 75, 90, dan 95.



Contoh berikut ini adalah data deskripsi data dari salary dan gender.

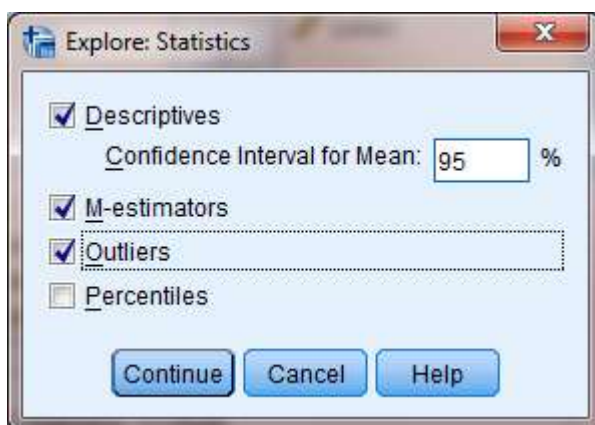
Tahapan Penyelesaian

- Buka file pada CD Latihan, kemudian pilih file employee.SAV
- Klik Analyze → Descriptive Statistics → Explore pada menu sehingga akan muncul kotak dialog Explore



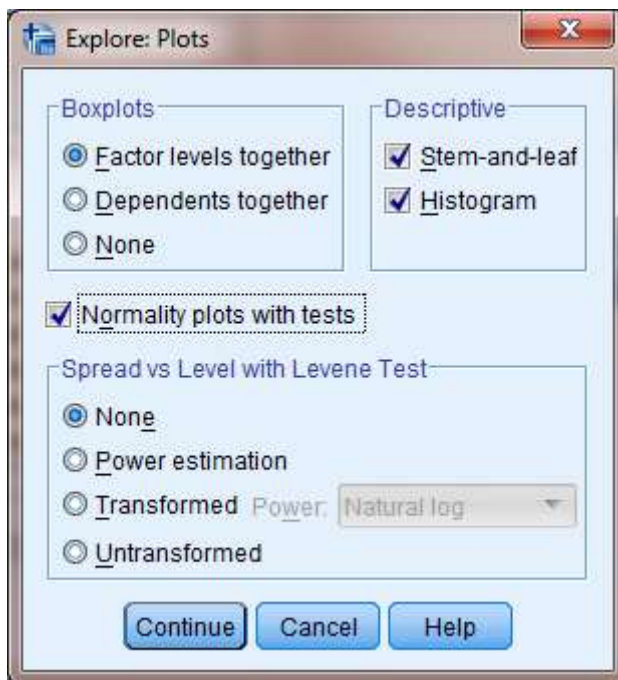
Gambar 4.11. Kotak Dialog Explore

- Masukkan variabel "salary", ke kotak "Dependent List", kemudian variabel "gender" pada kotak "factor List" dan "ID" pada kotak "List cases by".
- Klik Statistics, kemudian aktifkan pilihan Descriptives, M-estimator, dan outliers. Kemudian Klik Continue



Gambar 3.12. Kotak Dialog Explorer : statistics

- Klik "Plots" sehingga akan menampilkan kotak dialog Plots. Aktifkan pilihan "Histogram" dan "Normal plots with tests", kemudian klik "Continue"



Gambar 3.13. Kotak Dialog Explore : Plots

- Klik tombol OK sehingga akan muncul output sebagai berikut :

### gender

		Case Processing Summary					
		Cases					
		Valid		Missing		Total	
gender	N	Percent	N	Percent	N	Percent	
salary Wanita	172	100.0%	0	.0%	172	100.0%	
Laki-Laki	228	100.0%	0	.0%	228	100.0%	

Tabel pertama adalah Case Processing Summary menunjukkan jumlah responden berdasarkan jenis kelamin (gender).

Descriptives					
gender				Statistic	Std. Error
salary	Wanita	Mean		8885.69	329.972
		95% Confidence Interval for Mean	Lower Bound	8234.35	
			Upper Bound	9537.04	
		5% Trimmed Mean		8338.04	
		5% Trimmed Mean		7600.00	
		Variance		18727640.47	
		Std. Deviation		4327.544	
		Minimum		4688	
		Maximum		41219	
		Range		36531	
		Interquartile Range		3453	
		Skewness		3.650	.185
		Kurtosis		20.250	.368
		Laki-Laki	Laki-Laki	Mean	
95% Confidence Interval for Mean	Lower Bound			8986.79	
	Upper Bound			10258.86	
5% Trimmed Mean				9104.45	
5% Trimmed Mean				7925.00	
Variance				23754868.27	
Std. Deviation				4873.897	
Minimum				3975	
Maximum				33650	

Tabel Descriptives memberikan informasi mengenai nilai statistik deskriptif salary berdasarkan jenis kelamin.

- 95% Confidence interval for mean-lower Bound and Upper merupakan batas atas dan batas bawah interval kepercayaan 95% dari data-rata. Data di luar itu dianggap sebagai outlier.
- 5% Trimmed Mean adalah mengeliminasi 5% data ekstrem bawah dan atas, kemudian dilakukan perhitungan rata-ratanya. Tujuan dari perhitungan ini adalah mengurangi penyimpangan data. Dari output di atas terlihat bahwa nilai mean pada kelompok wanita adalah 8,885 dibandingkan dengan Trimmed Mean sebesar 8,338.
- Interquartile Range merupakan selisih nilai persentil 25 dengan nilai persentile 75.
- Interpretasi data yang lain sama dengan contoh yang diberikan sebelumnya yaitu pada analisis frekuensi data skala.

M-Estimators					
gender		Huber's M-Estimator <sup>a</sup>	Tukey's Biweight <sup>b</sup>	Hampel's M-Estimator <sup>c</sup>	Andrews' WWave <sup>d</sup>
salary	Wanita	7863.86	7501.20	7740.23	7496.28
	Laki-Laki	8195.69	7638.60	7981.70	7628.68

a. The weighting constant is 1.339.

b. The weighting constant is 4.685.

c. The weighting constants are 1.700, 3.400, and 8.500.

d. The weighting constant is  $1.340 \cdot \pi$ .

Tabel M-Estimators memberikan informasi alternatif tendensi sentral yaitu M-Estimator Huber bernilai 7,863.86, Tukey's bernilai 7,501.00 dan Hampel sebesar 7,740.23.

**Extreme Values**

gender			Case Number	id	Value	
salary	Wanita	Highest	1	95	95	41219
			2	112	112	27344
			3	8	8	24688
			4	88	88	22813
			5	245	245	17969
	Laki-Laki	Lowest	1	103	103	4688
			2	232	232	4913
			3	220	220	4913
			4	229	229	5031
			5	400	400	5100
Laki-Laki	Highest	1	78	78	33650	
		2	18	18	26538	
		3	313	313	25875	
		4	85	85	25500	
		5	50	50	25475	
	Laki-Laki	Lowest	1	310	310	3975
			2	373	373	4050
			3	74	74	4050
			4	328	328	4238
			5	226	226	4275

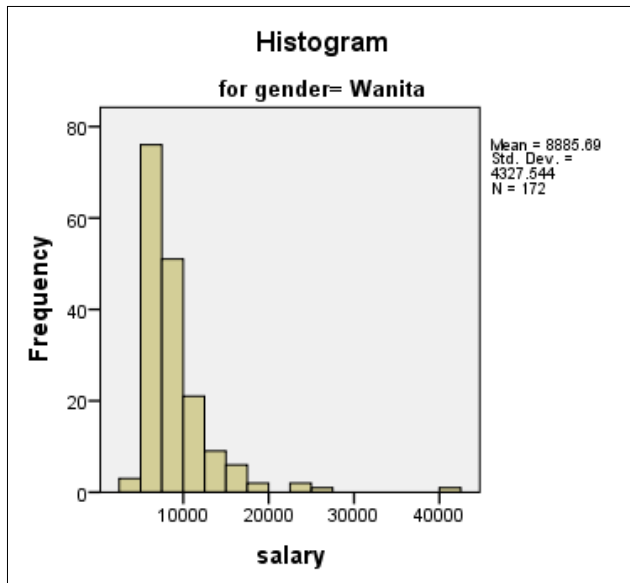
Tabel Extreme Values memberikan informasi mengenai salary terbesar dan terkecil. Dari kolom Case Number dan ID kita dapat mengetahui posisi data yang extreme, baik terendah maupun tertinggi.

**Tests of Normality**

gender	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
salary Wanita	.196	172	.000	.678	172	.000
Laki-Laki	.213	228	.000	.796	228	.000

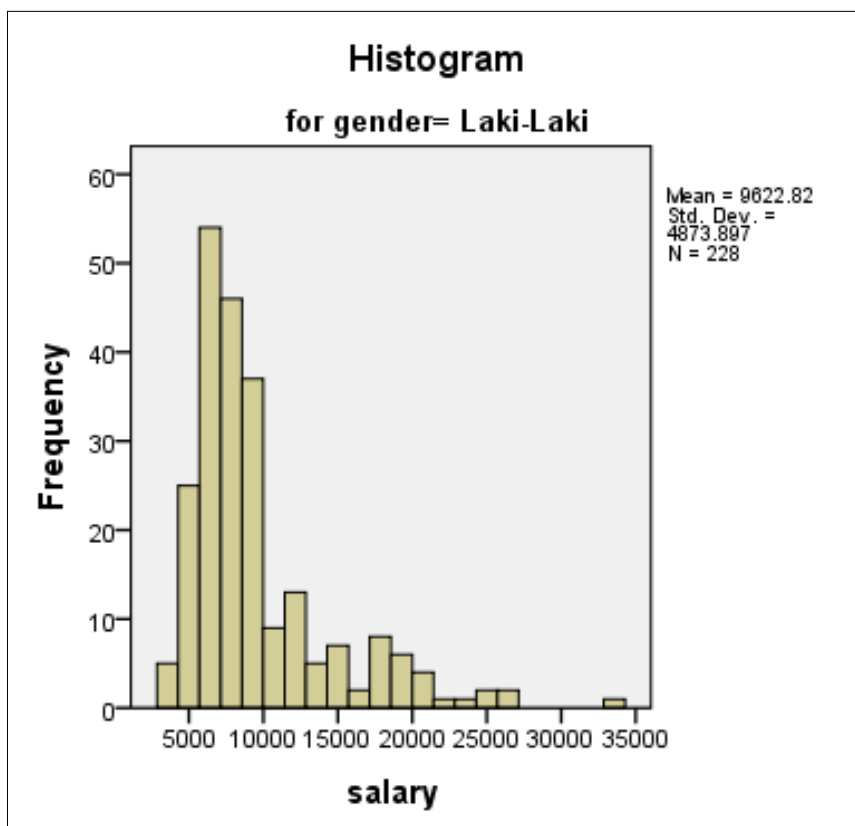
a. Lilliefors Significance Correction

Tabel test of Normality menampilkan hasil uji normalitas dengan dua teknik yaitu Kolmogorov-Smirnov dan Shapiro-Wilk. Hasil menunjukkan bahwa data salary terbukti tidak normal ( $\text{sig} < 0.05$ ) baik untuk kelompok laki-laki maupun wanita.



Gambar 3.14. Histogram Salary Kelompok wanita

Melalui Histogram kita dapat melihat bahwa data tidak berdistribusi normal (menceng ke kiri, atau positive skewness) baik untuk kelompok laki-laki dan wanita. Tampilan Histogram ini relevan dengan uji normalitas dengan Kolmogorov-Smirnov dan Shapiro-Wilk yang dilakukan sebelumnya.



Gambar 3.14. Histogram Salary Kelompok wanita

## PERTEMUAN 4-5 ANALISIS KORELASI

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami dan menjelaskan korelasi.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data korelasi menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Bivariate correlation
  - b. Partial Correlation

### 5. Materi

Analisis korelasi merupakan prosedur untuk mengukur hubungan antara dua variabel atau lebih. Nilai yang dihasilkan disebut koefisien korelasi yang memperlihatkan jika perubahan satu item akan mengakibatkan perubahan pada item lain. Bab ini kita akan secara khusus membahas mengenai korelasi bivariate, yaitu korelasi antara dua variabel menggunakan SPSS. Sifat korelasi ini adalah setara, yaitu tidak ada variabel yang diposisikan sebagai independen dan dependen.

#### Korelasi bivariate dan korelasi multivariate

Korelasi bivariate adalah korelasi antar dua variabel, sedangkan korelasi multivariate adalah korelasi lebih dari satu variabel dengan variabel lainnya.

#### Koefisien Korelasi dan Kekuatan Hubungan

Korelasi statistik diukur dengan apa yang disebut koefisien korelasi ( $r$ ). Nilai numerik yang berkisar dari positif 1,0 ke -1,0. Nilai ini memberikan indikasi kekuatan hubungan. Secara umum,  $r > 0$  menunjukkan hubungan positif,  $r < 0$  menunjukkan hubungan negatif sementara  $r = 0$  menunjukkan tidak ada hubungan. Selanjutnya,  $r = 1,0$  menggambarkan korelasi positif sempurna dan  $r = -1,0$  menjelaskan korelasi negatif sempurna. Semakin nilai koefisien mendekati angka 1,0 atau -1,0, maka hubungan antara dua variabel tersebut dinyatakan semakin besar.

#### Kriteria penafsiran korelasi adalah :

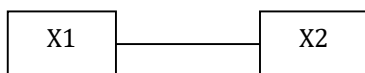
Nilai r positif	Nilai r negatif	Interpretasi
0,70 s/d 1,0	-0,70 s/d -1,0	Kuat
0,40 s/d 0,69	-0,40 s/d -0,69	Moderat
0,10 s/d 0,39	-0,10 s/d -0,39	Lemah
< 0.10	< -0.10	Sangatlemah / tidakadakorelasi

Sebuah koefisien korelasi rendah (misalnya, kurang dari  $\pm 0,10$ ) menunjukkan bahwa hubungan antara dua variabel lemah atau tidak ada. Sebuah koefisien korelasi yang tinggi (mendekati angka plus atau minus 1) menunjukkan bahwa perubahan pada salah satu variabel berubah “secara kebetulan” variabel yang lain juga akan berubah. Karena sifat korelasi adalah non-causal (bukan sebab akibat, maka tidak dapat diputuskan mana yang mempengaruhi dan mana yang dipengaruhi).

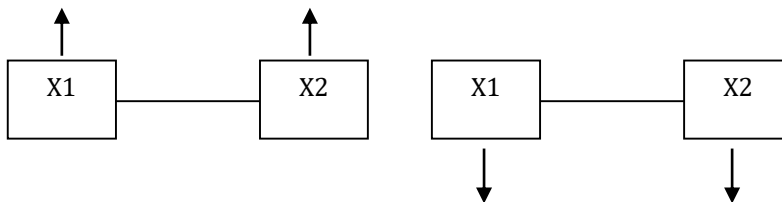
### Arah hubungan

Arah perubahan variabel dependen tergantung pada tanda koefisien. Jika koefisien adalah angka positif, maka jika variabel X1 tinggi maka X2 juga akan tinggi, dan dapat diinterpretasikan sebaliknya yaitu jika X2 tinggi maka X1 juga akan tinggi. Sedangkan angka koefisien negatif menandakan hubungan berbanding terbalik yaitu jika variabel X1 akan bergerak ke atas maka X2 akan bergerak ke arah yang berlawanan, dan atau sebaliknya.

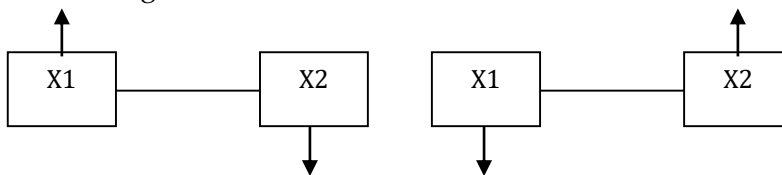
Korelasi non-causal



Korelasi Positif (Jika X1 naik maka X2 naik)

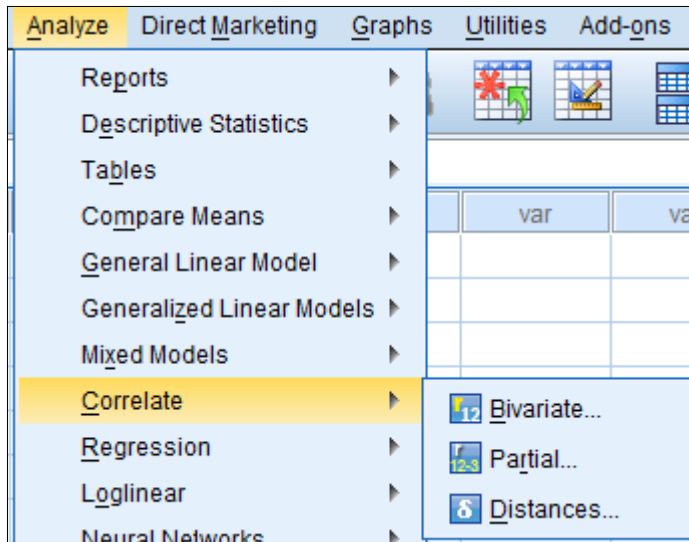


Korelasi Negatif



Gambar 4.1. Korelasi dan Arah hubungan

SPSS memberikan tiga prosedur dalam analisis korelasi yaitu korelasi bivariate, korelasi partial dan Distances. Tiga prosedur korelasi ini akan dibahas satu-persatu dalam bab ini.



Gambar 4.2. Menu Pilihan Korelasi pada SPSS

### A. Bivariate Correlation

Korelasi bivariat merupakan prosedur untuk menghitung koefisien korelasi Pearson, Spearmanrho, dan Kendall tau-b dengan tingkat signifikansi tertentu. Sebelum menghitung koefisien korelasi, dilakukan pengecekan outlier (yang dapat menyebabkan hasil yang menyesatkan) dan bukti hubungan linear. Koefisien korelasi Pearson adalah ukuran asosiasi linear. Dua variabel bisa berhubungan kuat, tetapi jika hubungan ini tidak linear, koefisien korelasi Pearson bukan statistik yang tepat untuk mengukur hubungan mereka.<sup>6</sup>

### Korelasi Pearson Product Moment

Koefisien korelasi Pearson mensyaratkan data bertipe interval/rasio, memiliki distribusi normal dan linier.

#### Format Data

Nama	Kolom	Keterangan
No	1	ID Responden
gender	2	0 = laki-laki 1 = wanita
Skortest	3	Skor test awalmasukkerja
Penjualan	4	Jumlah unit mobilterjual
Rank_Skor	5	Ranking Skor Test 1 = skor 70 - 75 2 = skor 76 - 80 3 = skor 81 - 90 4 = skorlebihdari 90
Rank_Jual	6	Ranking Penjualan

<sup>6</sup>[http://www-01.ibm.com/support/knowledgecenter/SSLVMB\\_20.0.0/com.ibm.spss.statistics.help/idh\\_corr.htm](http://www-01.ibm.com/support/knowledgecenter/SSLVMB_20.0.0/com.ibm.spss.statistics.help/idh_corr.htm)



---

1 = 1 s/d 3 unit  
 2 = 4 s/d 5 unit  
 3 = 6 s/d 8 unit  
 4 = lebih dari 8 unit

---

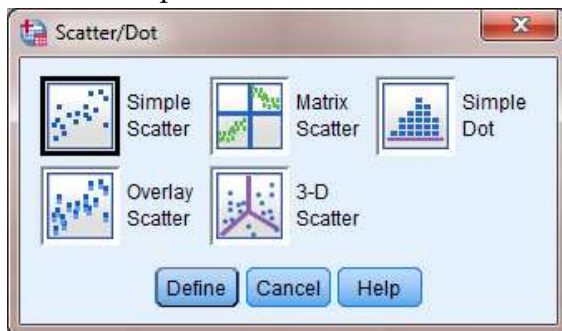
#### Contoh kasus

Seorang manager pemasaran ingin mengetahui hubungan antara skor test awal masuk kerja dengan unit mobil terjual pada 30 orang karyawan bagian penjualan. Jumlah penjualan mobil dihitung selama tiga bulan setelah karyawan tersebut diterima bekerja.

#### Penyelesaian Kasus

Sebelum melakukan analisis korelasi Pearson Product Moment, dilakukan pengujian Linearitas menggunakan Scatterplots dan normalitas menggunakan Kolmogorov-Smirnov Test.

1. Buka file "salesman.SAV"
2. Dari menu awal, klik Graphs, pilih Legacy Dialogs, dan kemudian pilih Scatter/Dot. Pilih Simple Scatter

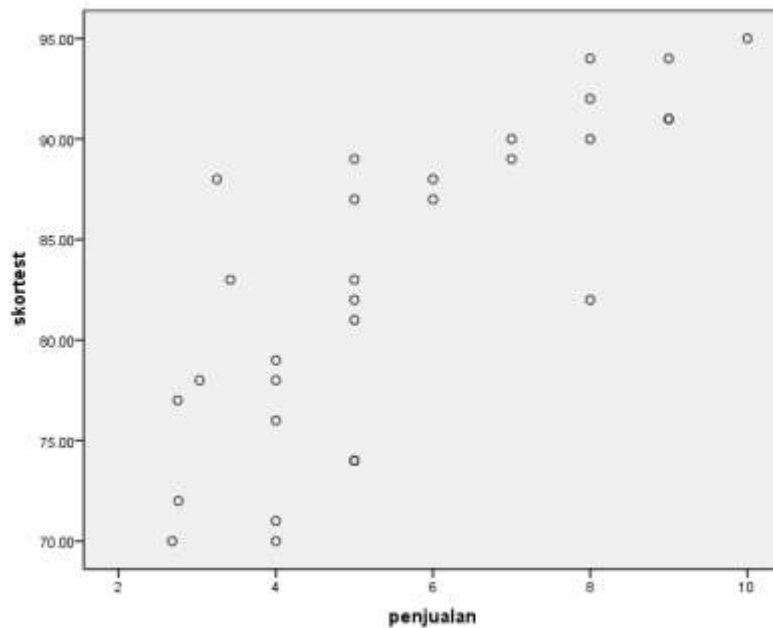


Gambar 4.1. Pilihan Grafik

3. Masukkan variabel scortest ke **Y axis** dan variabel penjualan ke X axis. Abaikan pilihan yang lain kemudian klik OK

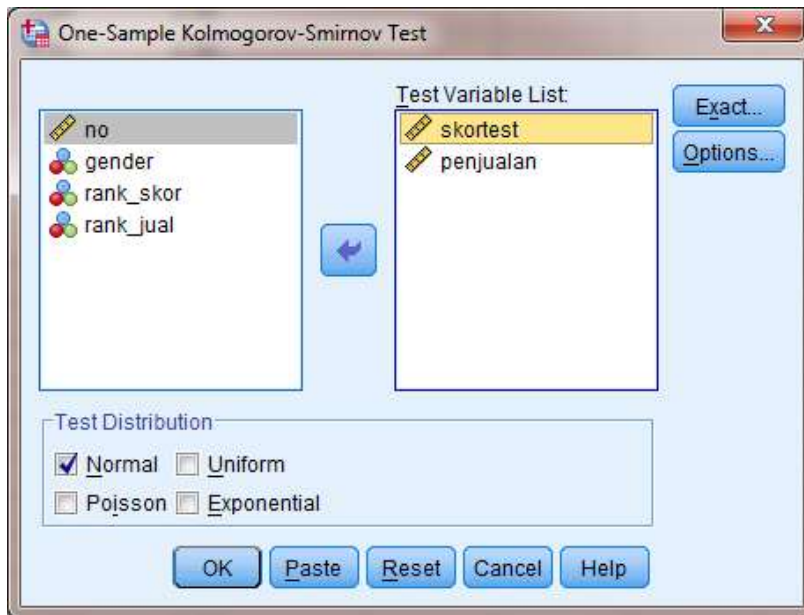


Gambar 4.2. Simple Scatterplot



Seperti dapat dilihat dari tampilan output di atas dapat disimpulkan bahwa hubungan antara skortest dan penjualan adalah linier. Sehingga dapat dinyatakan bahwa jika skor test awal masuk kerja tinggi, begitu pula penjualan.

4. Selanjutnya adalah melakukan uji normalitas. Pertama-tama klik Analyze, kemudian pilih Nonparametric Test, lalu Legacy Dialogs dan pilih "1 sample-K-S



Gambar 4.3. One-Sample Kolmogorov Smirnov Test

5. Pada kotak dialog, masukkan variabel skortest dan penjualan. Abaikan yang lain, lalu klik OK

		skortest	penjualan
N		30	30
Normal Parameters <sup>a,b</sup>	Mean	83.1667	5.60
	Std. Deviation	7.80841	2.183
Most Extreme Differences	Absolute	.155	.208
	Positive	.080	.208
	Negative	-.155	-.131
Kolmogorov-Smirnov Z		.849	1.137
Asymp. Sig. (2-tailed)		.467	.150

a. Test distribution is Normal.  
b. Calculated from data.

Tampilan output One-Sample Kolmogorov Smirnov Test menunjukkan nilai Asymp.Sig (2-tailed) untuk kedua variabel lebih besar dari 0.05 sehingga dapat disimpulkan kedua data telah memenuhi asumsi normalitas.

6. Tahap terakhir adalah melakukan uji korelasi. Klik Analyze kemudian pilih Correlate lalu Bivariate Correlation. Masukkan variabel skortest dan penjualan ke kotak "Test variable List".



Gambar 4.4. Bivariate Correlation: Pearson

7. Klik OK

		skortest	penjualan
skortest	Pearson Correlation	1	.786**
	Sig. (2-tailed)		.000
	N	30	30
penjualan	Pearson Correlation	.786**	1
	Sig. (2-tailed)	.000	
	N	30	30

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

Output pada tabel Correlation menampilkan besaran korelasi antara penjualan dan skortest sebesar 0.786 dengan sig 0.000. Ini berarti bahwa skor test awal masuk kerja yang tinggi sejalan dengan penjualan yang dihasilkan dalam tiga bulan pertama bekerja. Hubungan kedua variabel ini signifikan berdasarkan nilai p value  $0.000 < 0.01$ .

### Spearman Rank Order Correlation Coefficient

Spearman's rank-order correlation adalah versi nonparametrik dari korelasi Pearson Product moment. Koefisien korelasi Spearman, ( $\rho$  atau  $r_s$ ) mengukur kekuatan hubungan antara dua variabel yang memiliki karakteristik peringkat. Karena korelasi Spearman merupakan teknik statistik non-parametrik maka teknik ini tidak

memerlukan asumsi normalitas, dan biasa dipergunakan sebagai alternatif uji jika asumsi yang dibutuhkan untuk korelasi Pearson Product Moment tidak terpenuhi.

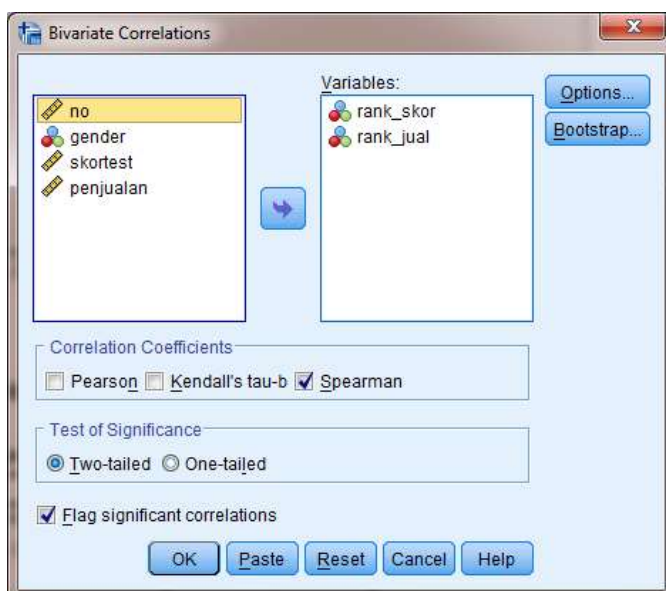
Skala yang dipergunakan adalah skala peringkat (Ordinal) dan hubungan antar variabel diasumsikan merupakan Monotonic relationship<sup>7</sup>

### Contoh Kasus

Masih menggunakan data yang sama yaitu "06 salesman.SAV", kita akan menganalisis hubungan antara ranking skor test (rank\_skor) dengan ranking penjualan (rank\_jual) karyawan bagian penjualan.

### Penyelesaian

1. Pada menu awal, klik Analyze, pilih Correlate, dan kemudian pilih Bivariate.
2. Masukkan variabel rank\_skor dan rank\_jual ke kotak "Test variable List".



Gambar 4.5. Bivariate Correlation : Spearman Rank

3. Aktifkan pilihan "Spearman" pada kotak "Correlation Coefficients"
4. Klik OK

### Hasil dan Interpretasi

<sup>7</sup>Monotonic relationship adalah hubungan terjadi: (1) jika nilai variabel satu meningkat, demikian juga nilai dari variabel lainnya; atau (2) jika nilai variabel satu meningkat, nilai variabel lainnya menurun

## Nonparametric Correlations

			rank_skor	rank_jual
Spearman's rho	rank_skor	Correlation Coefficient	1.000	.796**
		Sig. (2-tailed)	.	.000
		N	30	30
	rank_jual	Correlation Coefficient	.796**	1.000
		Sig. (2-tailed)	.000	.
		N	30	30

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

Koefisien yang diperoleh Spearman rank-order ( $\rho = 0,796$ ,  $p = 0,000 < 0,001$ ) menunjukkan adanya hubungan yang kuat antara ranking skor test dengan ranking penjualan. Hasil ini sejalan dengan korelasi Pearson yang memperlihatkan bahwa ketika skor test tinggi maka penjualan yang dihasilkan oleh karyawan tiga bulan setelah diterima bekerja juga tinggi.

## PERTEMUAN 6-7-8 ANALISIS REGRESI

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami dan menjelaskan regresi liner sederhana dan berganda.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data regresi menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Regresi Sederhana
  - b. Regresi Berganda

### 5. Materi

#### A. Regresi Sederhana

Regresi linier menghitung koefisien dari persamaan linear, yang melibatkan satu atau lebih variabel independen. Sebagai contoh, misalnya seorang manager ingin memprediksi total penjualan tahunan tenaga penjual (variabel dependen) dari variabel independen seperti usia, pendidikan, dan pengalaman kerja karyawan.

Regresi menggunakan hubungan historis pada variabel independen dan variabel dependen untuk memprediksi nilai masa depan dari variabel dependen. Dalam dunia bisnis, analisis regresi digunakan untuk memprediksi hal-hal seperti penjualan masa depan, harga saham, nilai tukar mata uang, dan peningkatan produktivitas yang dihasilkan dari program pelatihan.

Dalam persamaan tunggal atau hanya satu variabel bebas yang digunakan, maka regresi yang digunakan adalah regresi linear sederhana dengan persamaan  $Y = a + \beta X + \varepsilon$

Dimana,

$Y$  = nilai prediksi variabel dependen / terikat

$a$  = konstanta / intercept

$\beta$  = koefisien regresi / slope

$X$  = variabel bebas

Nilai konstanta/intercept merupakan nilai  $Y$  jika  $X$  bernilai konstan atau nol. Sedangkan  $\beta$  atau slope merupakan kenaikan (jika positif) atau penurunan (jika negatif) pada  $Y$  pada setiap kenaikan atau penurunan nilai  $X$ .

Persamaan untuk menghitung intercept ( $a$ ) adalah

$$a = \bar{y} - b\bar{x}$$

di mana slope atau dihitung sebagai berikut :

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

Dalam regresi linier sederhana ini, kita menghitung dampak satu variabel independen pada variabel dependen. Jika tinggi adalah satu-satunya penentu berat badan, kita akan mengharapkan bahwa poin untuk mata pelajaran individu akan berada dekat dengan garis. Namun, jika ada faktor-faktor lain (variabel independen) yang mempengaruhi berat badan selain tinggi (misalnya, usia, asupan kalori, dan tingkat latihan), kita mungkin berharap bahwa poin untuk mata pelajaran individu akan lebih longgar tersebar di sekitar garis, karena kita hanya mengambil tinggi ke rekening.

### **Asumsi Regresi Linier Sederhana**

Teknik regresi merupakan pendekatan parametrik / inferensial sehingga membutuhkan beberapa asumsi meliputi :

1. Pertama, variabel yang digunakan memiliki skala kontinyu (interval atau rasio).
2. Jumlah observasi direkomendasikan 20 kasus per variabel independen dalam analisis.
3. Regresi linier membutuhkan hubungan antara variabel independen dan dependen merupakan hubungan linear. Uji ini dapat dilakukan menggunakan scatter plot
4. Residual (kesalahan) dari garis regresi berdistribusi normal. Dua metode umum untuk memeriksa asumsi dapat digunakan grafik histogram atau p-plot normal

### **Contoh Kasus**

Seorang manager penjualan ingin mengetahui pengaruh biaya promosi yang dikeluarkan setiap bulannya dengan total unit motor terjual. Data biaya promosi dan penjualan dikumpulkan selama 3 tahun dengan total observasi 36 kasus.

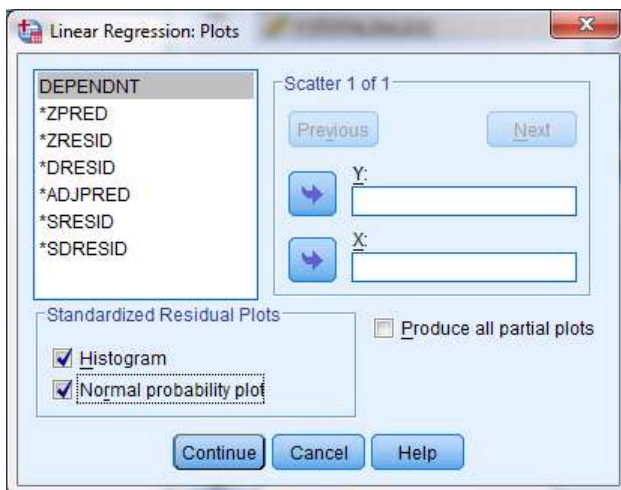
### **Penyelesaian**

5. Buka file "08 promo.SAV"
6. Pada menu awal, klik Analyze- Regression - Linier
7. Masukkan variabel Promo ke box independent, dan Jual (unit terjual) ke box "dependent"





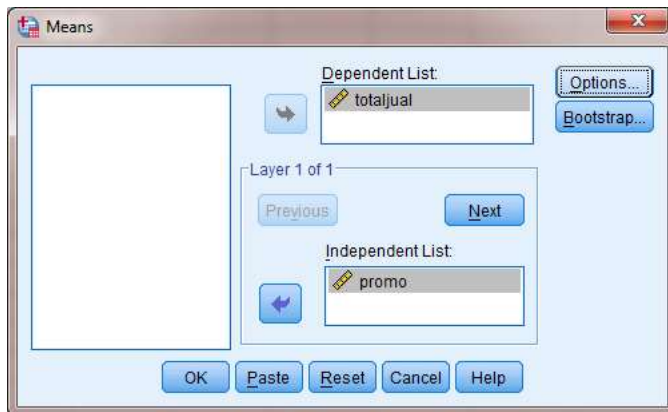
8. Klik Plots, lalu Tick pada pilihan “Histogram” dan “Normal Probability Plot”, laluKlik Continue



9. Klik OK

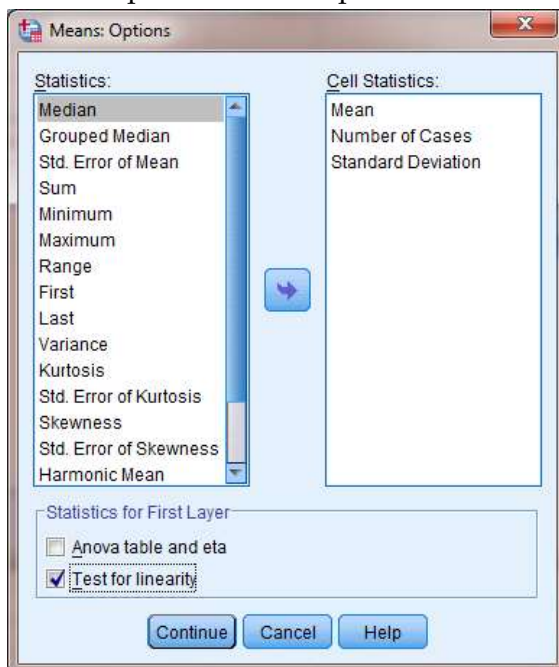
Uji Linieritas

- 10.Klik Compare Means > Means



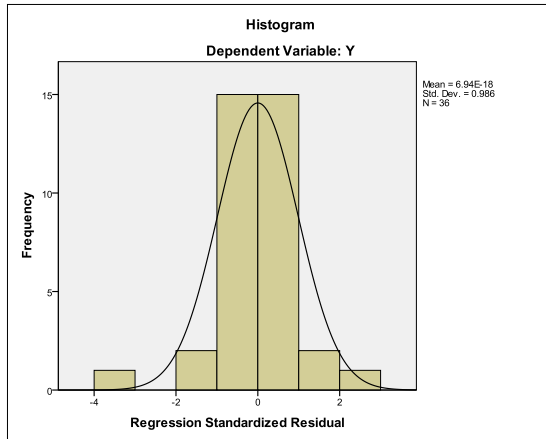
11. Masukkan variabel "promo" ke "independent list" dan "totalsales" ke "dependent list".

12. Klik Options. Aktifkan pilihan "test for linearity"



13. Klik Continue, lalu Klik OK

### Hasil dan Interpretasi



Berdasarkan hasil uji di atas terlihat bahwa menyebar ke kanan dan ke kiri secara merata sehingga dapat dinyatakan bahwa residual model regresi sudah berdistribusi normal sehingga analisis dapat dilanjutkan.

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Y*X	Between Groups	(Combined)	7730578.083	12	644214.840	2.494	.029
		Linearity	5774219.382	1	5774219.382	22.356	.000
		Deviation from Linearity	1956358.701	11	177850.791	.689	.736
	Within Groups		5940482.889	23	258281.865		
	Total		13671060.97	35			

Hasil di atas dapat diinterpretasikan dengan melihat nilai "Deviation from linearity". Angka sig pada deviation from linearity > 0.05 menandakan bahwa hubungan antara variable bebas dengan terikat adalah linier. Melalui tampilan di atas terlihat bahwa hubungan antara Promo dengan Total sales adalah linier (sig deviation from linearity > 0.05).

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650 <sup>a</sup>	.422	.405	481.93365

a. Predictors: (Constant), X  
b. Dependent Variable: Y

ANOVA <sup>b</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5774219.382	1	5774219.382	24.861	.000 <sup>a</sup>
	Residual	7896841.590	34	232260.047		
	Total	13671060.97	35			

a. Predictors: (Constant), X  
b. Dependent Variable: Y

Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5909.021	394.430		14.981	.000
	X	75.426	15.127	.650	4.986	.000

a. Dependent Variable: Y

Korelasi antara biaya promosi dengan penjualan (unit terjual) adalah sebesar 0.650, dengan koefisien determinasi 0.422 (R Square). Dengan demikian dapat dinyatakan bahwa variasi penjualan mampu dijelaskan oleh biaya promosisebesar 42.20%, dan sisanya dipengaruhi faktor lain di luar model.

### Pengujian Hipotesis

Ho : Tidak ada pengaruh biaya promosi terhadap Penjualan

Ha : Ada pengaruh positif dan signifikan biaya promosi terhadap Penjualan

Dari hasil uji signifikansi terlihat dari tabel "Coefficients" bahwa nilai probabilitas adalah sebesar 0,00 (< 0,01) sehingga Ho ditolak. Artinya, biaya promosi terbukti secara signifikan mempengaruhi penjualan sehingga hipotesis alternatif (Ha) diterima.

### Persamaan Regresi

Persamaan regresi : Penjualan = 5909.021 + 75.426 (biaya promosi). Persamaan regresi tersebut mempunyai makna sebagai berikut:

Konstanta sebesar 5909.021 menandakan tanpa adanya biaya yang dikeluarkan untuk promosi (maka penjualan sepeda motor adalah sebesar 5909 unit. Koefisien b sebesar

75.426 dapat diinterpretasikan bahwa setiap kenaikan 1 juta pada biaya promosi maka akan menaikkan penjualan sebesar 75 (pembulatan) unit motor.

### **B. Regresi Linier Berganda**

Regresi berganda merupakan perluasan dari regresi linier sederhana. Teknik ini digunakan ketika kita ingin memprediksi nilai suatu variabel berdasarkan nilai dari dua atau lebih variabel lainnya. Variabel yang ingin diprediksi disebut variabel dependen (atau kadang-kadang disebut sebagai variabel outcome, target atau variabel kriteria), sedangkan variabel yang digunakan untuk memprediksi nilai variabel dependen disebut variabel independen (atau kadang-kadang, prediktor, penjelas atau variabel regressor).

Sebagai perluasan dari regresi sederhana, regresi berganda dipergunakan dalam berbagai bidang, seperti ekonomi, psikologi, pendidikan, pemasaran, kesehatan, dan berbagai bidang keilmuan lainnya. Di bidang ekonomi, peneliti terkadang tertarik untuk memprediksi harga saham berdasarkan nilai tukar, suku bunga, dan faktor-faktor fundamental dalam perusahaan. Di bidang psikologi, peneliti tertarik untuk memprediksi prestasi seorang siswa berdasarkan IQ. Di bidang pendidikan, peneliti misalnya tertarik untuk memprediksi prestasi belajar siswa berdasarkan metode pembelajaran yang digunakan. Dan di bidang kesehatan misalnya, peneliti tertarik untuk menyelidiki apakah konsumsi rokok setiap hari dapat diprediksi berdasarkan durasi merokok, usia saat mulai merokok, jenis rokok, pendapatan dan lainnya.

Berbagai kegunaan analisis regresi berganda itu kemudian menjadikan teknik analisis ini merupakan teknik yang paling populer dipergunakan oleh peneliti dalam berbagai jurnal publikasi selama 20 tahun terakhir.

### **Persamaan Regresi Berganda**

Analisis regresi berganda digunakan untuk menjelaskan suatu variabel respon (variabel terikat / dependen) menggunakan lebih dari satu variabel input (variabel bebas, independent variable / eksogen). Persamaan regresi ganda dinotasikan sebagai berikut :

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 \dots B_kX_k + e$$

dimana Y adalah variabel respon, a = konstanta, b = parameter regresi)

### **Asumsi Regresi Berganda**

Sebagai teknik analisis yang tergolong dalam statistika inferensial, analisis regresi memiliki beberapa syarat atau asumsi yang harus dipenuhi. Asumsi tersebut meliputi asumsi klasik (multikolinieritas, heterokedastisitas, autokorelasi), dan asumsi lainnya adalah normalitas galat (residual), linieritas, ukuran kasus yang memadai dan jenis data bertipe kontinyu.

1. Variabel dependen yang digunakan harus memiliki skala kontinyu (interval dan rasio). Jika variabel dependen diukur dalam skala ordinal, maka teknik analisis yang digunakan adalah regresi ordinal dibanding regresi berganda.
2. Jumlah observasi direkomendasikan minimal 20 kali jumlah variabel independen. Misalnya jumlah variabel independen sebanyak empat, maka dibutuhkan minimal 80 kasus dalam analisis.
3. Linearitas : analisis regresi didasarkan pada konsep hubungan linear antara dependen dan variabel independen. Linearitas dapat dengan mudah diperiksa melalui plot residual. Teknik kedua adalah menggunakan nilai deviation from linearity dari uji F liniernya. Jika nilai sig pada deviation from linearity  $> 0.05$  maka dapat dikatakan hubungan antara variabel adalah linier.
4. Homokedastisitas, yaitu varians sama antara pasangan variabel. Scatter plot adalah salah satu cara untuk memeriksa apakah model regresi terkena masalah heterokedastisitas. Jika sebaran data tidak mengumpul di satu sudut/bagian maka disimpulkan bahwa tidak terjadi masalah heterokedastisitas, sehingga data dikatakan adalah homogen.
5. Multikolinieritas. Hal ini terjadi ketika variabel independen tidak saling bebas swatu sama lain. Teknik yang digunakan untuk mendeteksi multikolinieritas dapat dilakukan melalui empat cara, meliputi :
  - a. Matriks korelasi : digunakan dengan menghitung matrik matriks menggunakan korelasi bivariat Pearson antara semua variabel independen. Koefisien korelasi yang lebih besar dari 0.90 dianggap memiliki masalah multikolinieritas yang serius.
  - b. Tolerance yaitu toleransi untuk mengukur pengaruh satu variabel independen pada semua variabel independen lainnya. Tolerance dihitung dengan analisis regresi linier awal. Toleransi didefinisikan sebagai  $T = 1 - R^2$  untuk analisis regresi. Jika Tolerance  $< 0,2$  ada kemungkinan data terkena masalah multikolinieritas dalam data dan jika Tolerance  $< 0,01$  dipastikan data terkena masalah multikolinieritas
  - c. Variance Inflation Factor (VIF) - faktor inflasi varians dari regresi linier didefinisikan sebagai  $VIF = 1 / T$ .  $VIF > 10$  merupakan indikasi adanya multikolinieritas
6. Autokolerasi atau independence of error terms, yaitu independensi kesalahan dalam pengukuran dalam regresi. Autokorelasi terjadi ketika residu tidak independen dari satu sama lain. Dengan kata lain ketika nilai  $y$  ( $x + 1$ ) tidak independen dari nilai  $y$  ( $x$ ). Ini misalnya terjadi pada harga saham, dimana harga saat ini tidak independen dari harga saham kemarin. Jika residual independen, pola akan muncul secara acak. Pelanggaran akan ditandai dengan pola yang konsisten pada residual. Pelanggaran asumsi ini juga dapat dideteksi oleh Durbin Watson test. Jika Durbin-Watson d statistic berada diantara dua nilai kritis  $1,5 < d < 2,5$ , maka dapat diasumsikan bahwa tidak ada kolinearity atau autokorelasi dalam data.

7. Normalitas, yaitu diasumsikan bahwa residual memiliki distribusi normal. Pelanggaran asumsi ini dapat dideteksi oleh pemeriksaan plot residual melalui Histogram atau p-p plot, atau menggunakan teknik Kolmogorov Smirnov.

### Pilihan Tahapan

Ada beberapa pilihan metode dalam proses analisis dengan SPSS yaitu metode Enter, Stepwise, Backward dan Forward.

1. Metode enter berarti semua prediktor dimasukkan dalam model yang dianalisis, dan peneliti sendiri yang akan melakukan pengambilan keputusan mana variabel bebas yang signifikan dan mana yang tidak.
2. Metode Stepwise digunakan ketika peneliti ingin memasukkan variabel dalam model regresi yang dapat meningkatkan F setidaknya 0,05 dan peneliti ingin mengeliminasi variabel tambahan jika kenaikan F kurang dari 0,1.
3. Metode Backward yaitu peneliti akan diberikan output SPSS berdasarkan tahapan pengujian pengaruh secara individual. Program akan mengeliminasi prediktor-prediktor yang tidak signifikan satu per satu sampai ditemukannya prediktor yang secara individual signifikan.
4. Metode Forward digunakan untuk mencari pengaruh antar prediktor dimana program akan mengeliminasi prediktor yang tidak signifikan. Perbedaannya dengan metode Backward adalah pada metode Forward, prediktor yang tidak signifikan akan dieliminasi dari awal sehingga akan langsung dieliminasi dari model.

### Kasus Pemasaran

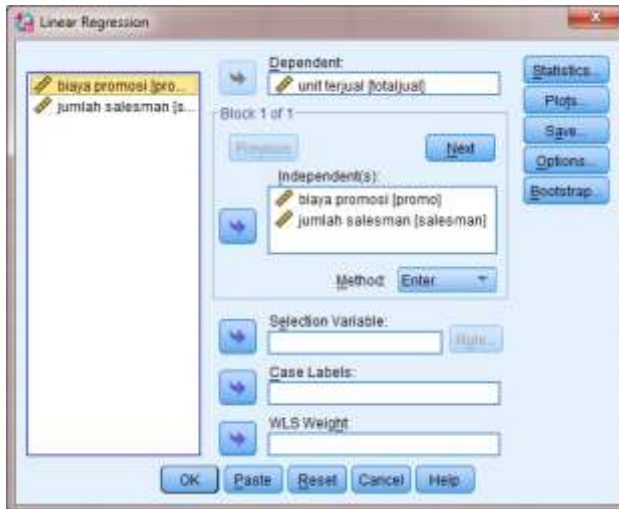
Seorang manajer penjualan salah satu agen sepeda motor ingin mengetahui pengaruh biaya promosi dengan biaya pelatihan yang dikeluarkan terhadap jumlah unit motor yang terjual dalam 4 tahun terakhir. Data diambil dari bulan Januari 2011 sampai dengan Desember 2014.

#### Format Data

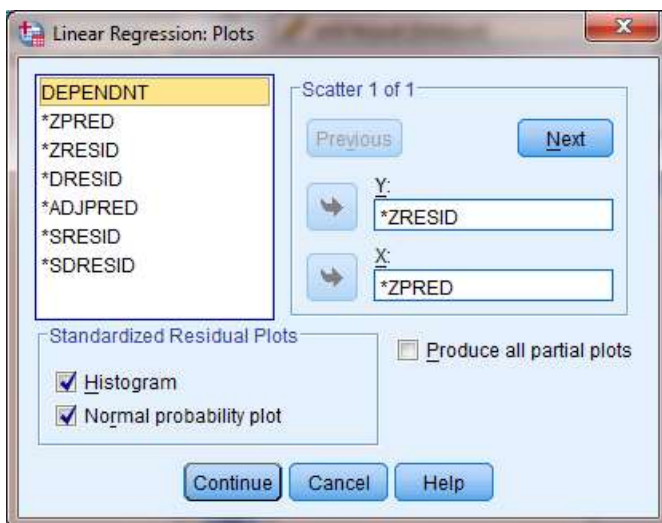
Nama	Kolom	Keterangan
promo	1	Biaya promosi dalam juta Rp
salesman	2	Jumlah salesman
totalsales	3	Jumlah unit motor terjual

#### Penyelesaian

1. Buka file "07 promo2.SAV"
2. Pada menu awal, klik Analyze- Regression - Linier
3. Masukkan variabel Promo dan salesman ke box independent, dan totaljual (unit terjual) ke box "dependent"

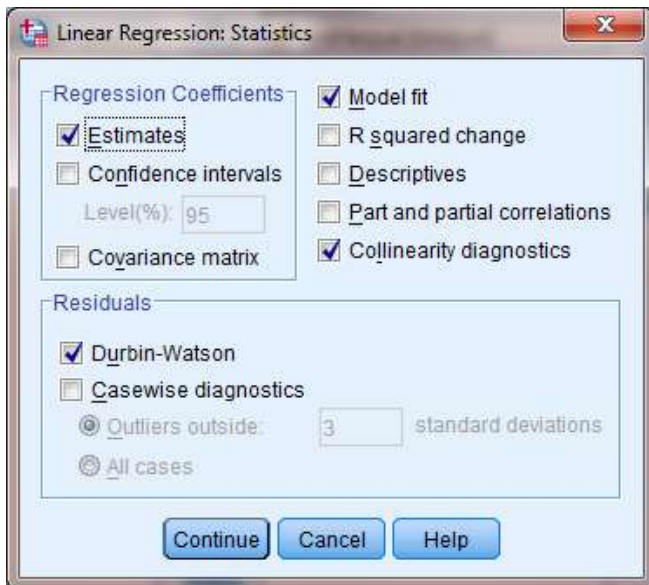


4. Klik Plots, lalu Tick pada pilihan “Histogram” dan “Normal Probability Plot” (ini dilakukan untuk mengeluarkan grafik histogram dan p-p plot untuk menguji normalitas residual). Masukkan Zresid ke Y dan Zpred ke X pada pilihan Scatter (ini dilakukan untuk mengeluarkan grafik Scatter Plot pada uji Heterokedastisitas)



5. lalu Klik Continue
6. Klik Statistics. Pada kotak dialog statistics, aktifkan “Collinearity diagnostic” untuk menguji multikolinieritas, dan “Durbin-Watson” pada box Residual untuk menguji Autokorelasi.

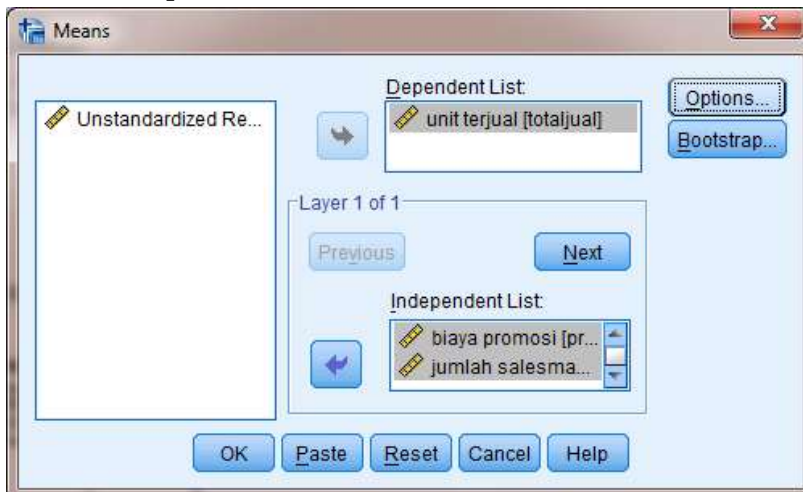




7. Klik Save. Kemudian aktifkan “unstandardized” untuk menghasilkan nilai residual. Klik Continue, kemudian OK
8. Output SPSS akan ditampilkan, namun kita abaikan terlebih dahulu untuk melanjutkan ke tahap selanjutnya yaitu uji linieritas.

#### Uji Linieritas

9. Klik Compare Means > Means



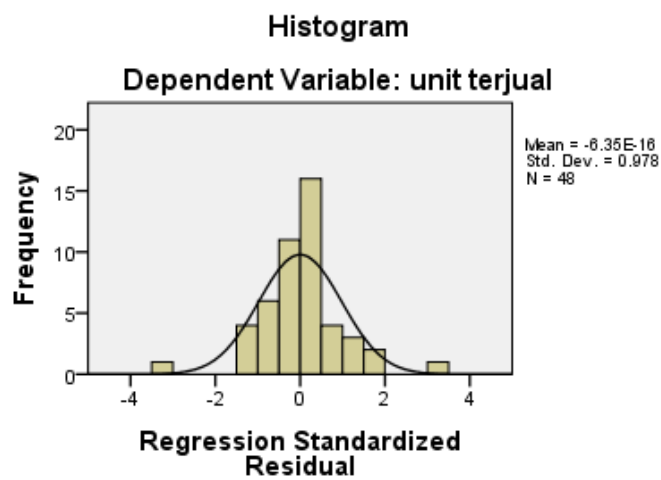
10. Masukkan variabel “promo” dan “salesman” ke “independent list” dan “totaljual” ke “depedent list”.
11. Klik Options. Aktifkan pilihan “test for linearity”
12. Klik Continue, kemudian Klik OK

#### Hasil dan Interpretasi

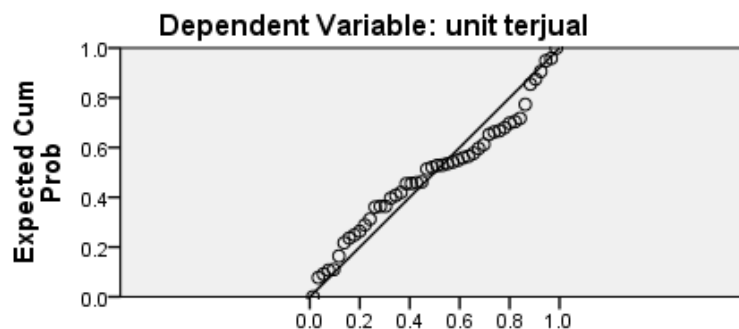
Sebelum membahas hasil regresi, terlebih dahulu akan dibahas mengenai uji asumsi pada regresi.

#### Uji Normalitas

Dua grafik yang ditampilkan memperlihatkan bahwa residual sudah berdistribusi normal. Pertama, melalui histogram terlihat bahwa sebaran data sudah membentuk kurva normal dan berada di tengah-tengah wilayah positif dan negatif. Kedua, melalui Normal P-P Plot terlihat bahwa titik-titik data berada di sekitar garis diagonal dan mengikuti garis tersebut, sehingga dapat disimpulkan bahwa model regresi sudah memenuhi asumsi normalitas residual.

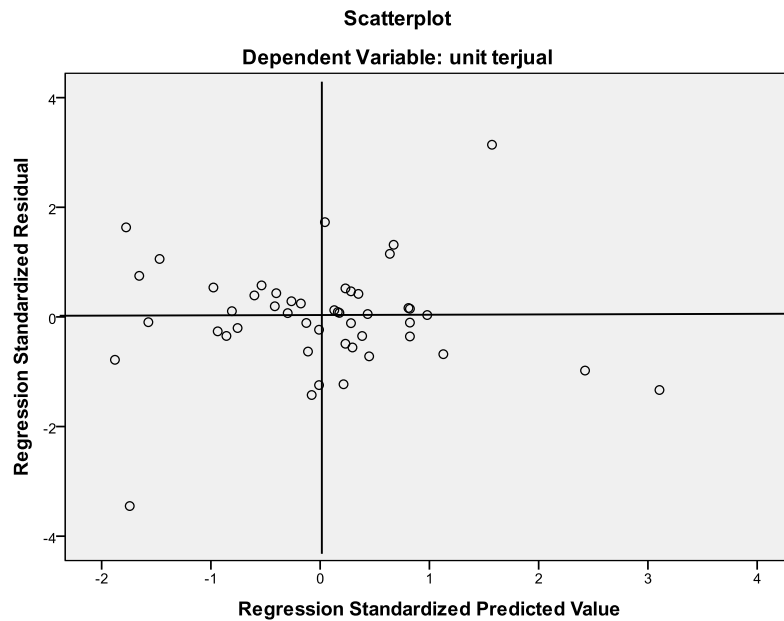


#### Normal P-P Plot of Regression Standardized Residual



#### Uji Heterokedastisitas

Uji Heterokedastisitas dengan grafik Scatter Plot memperlihatkan bahwa titik-titik menyebar ke kiri dan ke kanan, atas dan ke bawah pada sumbu 0. Selain itu, tidak terbentuk pola tertentu sehingga dapat disimpulkan model regresi terbebas dari masalah heterokedastisitas.



### Uji Multikolinieritas

Model	Standardized Coefficients	t	Sig.	Collinearity Statistics	
	Beta			Tolerance	VIF
1		11.427	.000		
2	.328	2.841	.007	.769	1.301
3	.518	4.483	.000	.769	1.301

Pada tabel Coefficient, dapat dilihat nilai VIF dan Tolerance untuk menguji multikolinieritas. Dengan melihat Nilai VIF (Varian Inflation Factor) diketahui bahwa tidak ada variabel yang memiliki nilai VIF lebih dari 10, serta nilai tolerance yang kurang dari 0.10. Dengan demikian dapat disimpulkan bahwa variabel independen terbebas dari masalah Multikolinieritas

### Uji Autokorelasi

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.734 <sup>a</sup>	.539	.518	396.13460	1.921

a. Predictors: (Constant), jumlah salesman, biaya promosi  
b. Dependent Variable: unit terjual

Hasil uji autokorelasi dapat dilihat dari nilai Durbin\_Watson sebesar 1.921. Untuk  $n=48$ , dan  $k=2$  diperoleh nilai DW tabel  $D_1$  1.354 dan  $D_u$  1.584. Nilai DW hitung 1.874 > dari

batas atas ( $d_u$ ) yaitu 1.584 dan kurang dari  $4 - d_u$ , sehingga dapat disimpulkan tidak ada autokorelasi positif maupun negative pada model.

Sebagai pedoman, untuk menentukan ada tidaknya autokorelasi, berikut ditampilkan tabel pengambilan keputusan pada uji autokorelasi :

Hipotesis Nol	Jika	Keputusan
No positive autocorrelation	$0 < d < d_l$	Tolak
No positive autocorrelation	$d_l \leq d \leq d_u$	Tidak ada keputusan
No negative autocorrelation	$4 - d_l < d < 4$	Tolak
No negative autocorrelation	$4 - d_u \leq d \leq 4 - d_l$	Tidak ada keputusan
No positive/negative autocorrelation	$Du < d < 4 - du$	Tidak ditolak

### Uji Linieritas

ANOVA Table					
	Sum of Squares	df	Mean Square	F	Sig.
(Combined)	10736754.00	22	488034.273	2.668	.010
Linearity	5094745.276	1	5094745.276	27.848	.000
Deviation from Linearity	5642008.724	21	268667.082	1.469	.178
	4573780.000	25	182951.200		
	15310534.00	47			

ANOVA Table					
	Sum of Squares	df	Mean Square	F	Sig.
(Combined)	8640973.905	12	720081.159	3.779	.001
Linearity	6982401.188	1	6982401.188	36.642	.000
Deviation from Linearity	1658572.716	11	150779.338	.791	.647
	6669560.095	35	190558.860		
	15310534.00	47			

Hasil di atas dapat diinterpretasikan dengan melihat nilai "Deviation from linearity". Angka sig pada deviation from linierity  $> 0.05$  menandakan bahwa hubungan antara variabel bebas dengan terikat adalah linier. Melalui tampilan di atas terlihat bahwa hubungan antara Promo dengan Totalsales maupun jumlah salesman dengan totalsales seluruhnya linier (sig deviation from linearity  $> 0.05$ ) yaitu sig 0.178 (promo-total jual) dan sig 0.647 (salesman - total jual).

## Pembahasan Korelasi dan Regresi

**Variables Entered/Removed<sup>b</sup>**

Model	Variables Entered	Variables Removed	Method
1	jumlah salesman, biaya promosi	.	Enter

a. All requested variables entered.  
b. Dependent Variable: unit terjual

**Model Summary<sup>b</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.734 <sup>a</sup>	.539	.518	396.13460	1.921

a. Predictors: (Constant), jumlah salesman, biaya promosi  
b. Dependent Variable: unit terjual

Korelasi antara biaya promosi dan jumlah salesman dengan penjualan (unit terjual) adalah sebesar 0.734, dengan koefisien determinasi 0.518 (adjusted R Square). Dengan demikian dapat dinyatakan bahwa variasi penjualan mampu dijelaskan oleh biaya promosi dan jumlah tenaga penjual sebesar 51.80%, dan sisanya dipengaruhi faktor lain di luar model.

### Pengujian Hipotesis Simultan

$H_0$  : Tidak ada pengaruh biaya promosi dan jumlah salesman terhadap Penjualan

$H_1$  : Ada pengaruh positif dan signifikan biaya promosi dan jumlah salesman terhadap Penjualan

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8249016.090	2	4124508.045	26.284	.000 <sup>a</sup>
	Residual	7061517.910	45	156922.620		
	Total	15310534.00	47			

a. Predictors: (Constant), jumlah salesman, biaya promosi  
b. Dependent Variable: unit terjual

Dari hasil uji signifikansi terlihat bahwa nilai probabilitas adalah sebesar 0,00 ( $< 0,01$ ) sehingga  $H_0$  ditolak. Artinya, pengaruh biaya promosi (iklan TV dan Radio) dan jumlah salesman secara simultan terbukti secara signifikan mempengaruhi penjualan sehingga hipotesis alternatif ( $H_1$ ) diterima.

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4099.777	684.277		5.991	.000
	PROMO	27.124	8.791	.407	3.085	.004
	TRAIN	53.294	15.311	.459	3.481	.001

a. Dependent Variable: TOTALSALES

### Pengujian Hipotesis Parsial

Hasil uji model parsial dengan memperhatikan nilai probabilitas pada uji t memperoleh nilai t hitung untuk promo sebesar 0.004 dan jumlah salesman sebesar 0.001. Karena probabilitas < 0.05, maka dapat disimpulkan bahwa secara parsial dua variabel ini terbukti berpengaruh signifikan terhadap penjualan.

### Persamaan Regresi

Persamaan regresi : Penjualan = 4099.777 + 27.124 (biaya promosi) + 53.294 (jumlah tenaga penjual), Persamaan regresi tersebut mempunyai makna sebagai berikut:

Konstanta sebesar 4099 menandakan tanpa adanya biaya yang dikeluarkan untuk promosi (melalui TV dan Radio) dan tenaga penjual, maka penjualan sepeda motor adalah sebesar 4099 unit.

Jika variabel biaya promosi naik satu juta maka akan menyebabkan kenaikan (karena tanda positif) sebesar 27.124 unit pada penjualan sepeda motor, dengan asumsi tenaga penjual bernilai konstan. Sedangkan jika jumlah tenaga penjual (salesman) naik 1 orang, maka akan menyebabkan kenaikan pada penjualan sebesar 53.294 unit sepeda motor, dengan asumsi biaya promosi bernilai konstan.

Simulasi angka dari persamaan regresi

Berdasarkan persamaan yang sudah dibentuk, yaitu :

$$\text{Penjualan} = 4099.777 + 27.124 (\text{biaya promosi}) + 53.294 (\text{jumlah tenaga penjual})$$

Dengan asumsi biaya promosi yang dikeluarkan sebesar 150 Juta Rupiah, dan jumlah tenaga penjual sebanyak 50 orang, maka dapat dihitung perkiraan penjualan sebesar :

$$Y = 4,099.777 + 27.124 (150) + 53.294 (50)$$

$$Y = 4,099.777 + 4,068.60 + 2,664.70$$

$$Y = 10,833.08$$

Perkiraan penjualan unit sepeda motor dengan simulasi angka promosi sebesar 150 Juta Rp dan 50 tenaga penjual adalah sebanyak 10,833.08 unit (sepuluh ribu delapan ratus tiga puluh tiga unit motor).

## PERTEMUAN 9-10-11 UJI BEDA RATA-RATA

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami, menjelaskan, dan mengaplikasikan uji beda rata-rata.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data uji beda rata-rata menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Uji Beda (Uji beda 1 Sampel)
  - b. Uji Beda 2 Sampel Berhubungan,
  - c. Uji Beda 2 Sampel Independen,
  - d. Uji Beda > 2 Sampel Berhubungan, dan
  - e. Uji Beda > 2 Sampel Independen

### 5. Materi

#### Pendahuluan

Bab ini memberikan simulasi mengenai uji perbedaan (komparasi) pada 5 desain uji yaitu uji beda satu sampel, uji beda 2 sampel berhubungan, uji beda dua sampel bebas (independen), uji beda k sampel berhubungan, dan uji beda k sampel bebas. Masing-masing uji diberikan contoh dalam dua pendekatan yaitu parametric dan nonparametric sehingga dapat dijadikan pedoman dan alternative pemilihan teknik analisis jika asumsi yang harus dipenuhi pada statistic parametric tidak dapat dipenuhi

#### A. Uji Beda 1 Sampel

##### 1. One sample t test

Uji Beda 1 Sampel atau *one sample t test* merupakan teknik analisis untuk membandingkan perbedaan rata-rata satu variabel dengan nilai yang sudah ditentukan sebelumnya. Uji t sebagai teknik pengujian hipotesis deskriptif memiliki tiga criteria yaitu uji pihak kanan, kiri dan dua pihak. One sample t test merupakan prosedur untuk menguji apakah ada perbedaan rata-rata dari satu variabel tunggal dengan nilai konstan yang sudah ditetapkan. Contoh. Seorang penelitimungkin ingin menguji apakahskortes rata-rata untuksekelompok mahasiswaberbeda dari 80, apakah ada rata-rata tabungan karyawan berbeda dari 10 juta pada tingkat kepercayaan 95%.

#### Karakteristik Uji :

**Jenis Data :** Untuk menguji nilai dari variabel kuantitatif terhadap nilai uji hipotesis, memilih variabel kuantitatif dan masukkan nilai uji hipotesis.

**Asumsi :** Tes ini mengasumsikan bahwa data terdistribusi secara normal

Seorang guru menduga bahwa rata-rata nilai test matematika 25 siswa kelas IV-A berbeda dari nilai rata-rata kelas IV-B yang sebelumnya sudah diketahui yaitu sebesar 75.

#### DATA PENELITIAN

no	Nilai
1	67.00
2	75.00
3	81.00
4	60.00
5	80.00
6	75.00
7	71.00
8	68.00
9	80.00
10	78.00
11	71.00
12	80.00
13	65.00
14	57.00
15	78.00
16	63.00
17	76.00
18	73.00
19	63.00
20	65.00
21	72.00
22	80.00
23	75.00
24	67.00
25	72.00

#### **Pengujian Hipotesis :**

##### **Rumusan masalah**

Hipotesis kalimat

Ha : Nilai matematika siswa kelas IVA tidak berbeda dari 75

Ho : Nilai matematika siswa kelas IVA berbeda dengan 75

##### **Hipotesis statistik**

Ha :  $\mu_0 \neq 75$

Ho :  $\mu_0 = 75$



**Parameter uji :**

Jika sig < 0.05, maka tolak Ho

Jika sig > 0.05, maka terima Ho

**Penyelesaian Kasus****Penyelesaian Kasus Uji Beda Satu Sampel One Sample t test**

Analyze

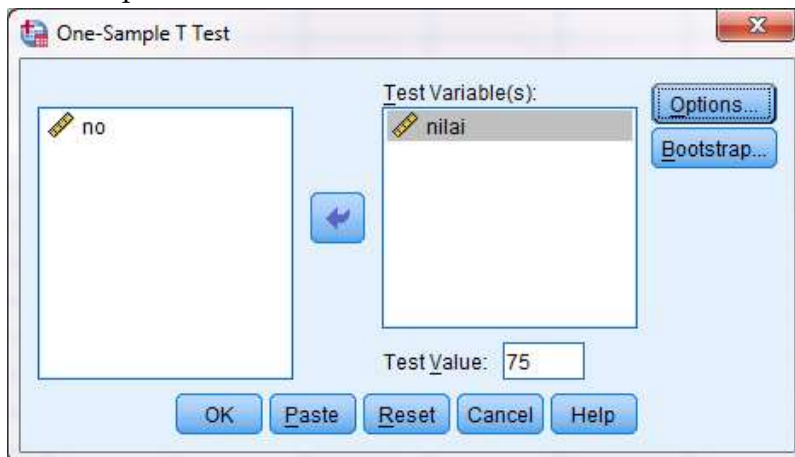
Compare Means

One Sample t test

Pada "variable" masukkan variabel "nilai"

Pada Test Value, masukkan angka 75

Klik OK

**One Sample t Test****Hasil Dan Pembahasan****One Sample t Test****One-Sample Statistics**

	N	Mean	Std. Deviation	Std. Error Mean
nilai	25	71.6800	6.91448	1.38290

**One-Sample Test**

	Test Value = 75					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
nilai	-2.401	24	.024	-3.32000	-6.1742	-.4658

Hasil uji di atas menunjukkan bahwa t hitung = -2.401 (artinya nilai kelas IV-A lebih kecil dibandingkan dengan nilai rata-rata kelas IV-B), dengan sig sebesar 0.024.

Karena sig (p value)  $0.024 < 0.05$ , maka dapat disimpulkan  $H_0$  ditolak, dan  $H_a$  diterima. Artinya, nilai rata-rata kelas IV-A terbukti berbeda dari 75 (dan terbukti lebih kecil yaitu sebesar 71.68).

## B. Uji Beda 2 Sampel Berhubungan

Tesdi bagian ini membandingkan distribusi dari dua variabel yang saling berhubungan. Teknik analisis yang digunakan ditentukan berdasarkan jenis data penelitian yaitu jika data yang digunakan adalah data kontinyu maka dapat menggunakan uji paired samples t test (parametric) atau uji tanda (sign test) dan Wilcoxon signed-rank untuk kelompok nonparametric. Jika data bertipe kategorial, maka teknik analisis yang dapat digunakan adalah McNemar test.

### Paired Sample t test

Paired sample t test merupakan uji beda dua sampel berpasangan. Sampel berpasangan merupakan subjek yang sama namun mengalami perlakuan yang berbeda. Karena paired sample t test merupakan salah satu teknik statistik parametric, maka asumsi yang harus dipenuhi adalah data berdistribusi normal.

### Karakteristik Uji

Data. Untuk setiap pasangan variabel, data minimal berbentuk interval atau rasio  
 Asumsi. Pengamatan untuk setiap pasangan harus dibuat dalam kondisi yang sama. Perbedaan rata-rata harus terdistribusi normal, dan varians masing-masing variabel dapat seimbang atau tidak seimbang.

### Contoh Kasus 3B

Akan diteliti mengenai perbedaan penjualan sepeda motor merk A di sebuah Kabupaten sebelum dan sesudah kenaikan harga BBM. Data diambil dari 15 dealer pada satu bulan sebelum kenaikan BBM dan 1 bulan setelah kenaikan BBM.

Dealer	Sebelum	Sesudah
1	67.00	68.00
2	75.00	76.00
3	81.00	80.00
4	60.00	63.00
5	80.00	82.00
6	75.00	74.00
7	71.00	70.00
8	68.00	71.00
9	80.00	82.00
10	78.00	79.00
11	71.00	78.00
12	80.00	77.00

13	65.00	69.00
14	57.00	67.00
15	78.00	68.00

### Penyelesaian Kasus Uji Beda 2 Sampel berhubungan

#### Penyelesaian Kasus Uji Beda 2 Sampel berhubungan

Buka Kasus10.xls

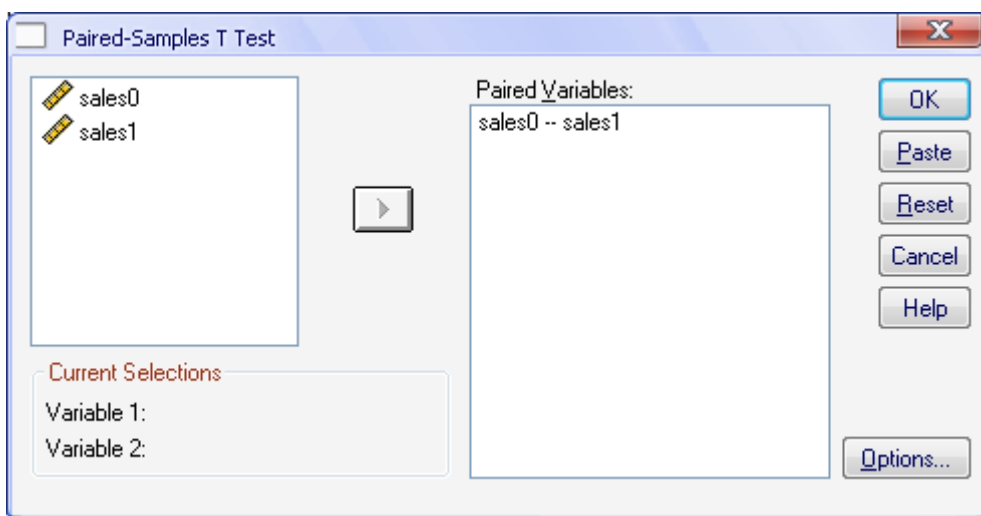
Analyze

Compare Means

Paired Sample t test

Pada "variable" masukkan variabel "sales0 dan sales01"

Klik OK



### Paired Samples Statistics

#### Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	sales0	72.4000	15	7.66066	1.97797
	sales1	73.6000	15	6.02139	1.55472

Tabel pertama berisi informasi mengenai rata-rata, jumlah observasi, standar deviasi dan standar error rata-rata. Rata-rata pada sales0 (sebelum kenaikan BBM) adalah sebesar 72.40 unit, sedangkan rata-rata penjualan pada sales1 (sesudah kenaikan BBM) adalah sebesar 73.96 unit.

### Paired Samples Correlations

#### Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	sales0 & sales1	15	.809	.000

Tabel kedua berisi informasi mengenai korelasi antara sales0 (sebelum) dan sales1 (sesudah kenaikan BBM). Angka korelasi diperoleh sebesar 0.809 dengan sig 0.000 yang menunjukkan kuatnya hubungan antara data penjualan sebelum dan sesudah kenaikan BBM. Tabel ketiga terlihat bahwa nilai sig (2-tailed) adalah sebesar 0.320 dan t hitung sebesar -1.031. Karena sig atau p value > 0.05 maka  $H_0$  diterima dan  $H_a$  ditolak. Atau dengan kata tidak ada perbedaan penjualan sepeda motor merk A antara sebelum kenaikan BBM dengan sesudah kenaikan BBM.

### C. Uji Beda 2 Sampel Independen

Seperti halnya uji beda dua sampel berhubungan, uji beda dua sampel bebas juga memiliki dua pendekatan yaitu parametric dengan independent sample t test dan non-parametrik dengan Median test, Wilcoxon-Mann-Whitney Test atau Kolmogorov-Smirnov Two Sample Test. Dua teknik yang akan dicontohkan pada bagian ini adalah independent sample t test dan Wilcoxon-Mann-Whitney test.

#### Independent Sample t tests

Independent-Sampel t test merupakan prosedur analisis untuk membandingkan perbedaan rata-rata pada dua kelompok data yang tidak berhubungan. Idealnya, untuk tes ini, subyek harus secara acak dipilih untuk dua kelompok, sehingga setiap perbedaan dalam respon ini disebabkan oleh "adanya perlakuan/ treatment" atau "tanpa perlakuan" dan karena faktor-faktor lain. Kondisi ini tidak akan secara maksimal terpenuhi jika dilakukan perbandingan rata-rata pendapatan pada karyawan pria dan wanita. Hal ini karena seseorang tidak dapat secara acak ditempatkan sebagai pria atau wanita. Dalam kondisi ini, perbedaan rata-rata pendapatan dapat dipengaruhi oleh faktor-faktor seperti pendidikan, pengalaman kerja, jabatan dan faktor lainnya.

#### Karakteristik Uji

Data : metric (interval – rasio)

Asumsi : sampel tidak berhubungan, dipilih secara acak, berdistribusi normal dan adanya kesamaan varians (menggunakan levene test)

#### Contoh kasus 3C

Seorang guru ingin mengetahui apakah ada perbedaan nilai matematika siswa melalui penggunaan metode pembelajaran diskusi kelompok. Untuk melakukan studi ini, Guru membagi kelas ke dalam dua kelompok yaitu kelompok sampel dan kelompok kontrol yang dipilih secara acak dengan masing-masing kelompok berjumlah 15 siswa.

## DATA PENELITIAN

nilai	grup
65.00	0.00
75.00	0.00
80.00	0.00
82.00	0.00
80.00	0.00
60.00	0.00
62.00	0.00
76.00	0.00
80.00	0.00
78.00	0.00
71.00	0.00
80.00	0.00
65.00	0.00
56.00	0.00
80.00	0.00
78.00	1.00
82.00	1.00
80.00	1.00
83.00	1.00
78.00	1.00
74.00	1.00
89.00	1.00
71.00	1.00
82.00	1.00
79.00	1.00
78.00	1.00
77.00	1.00
85.00	1.00
67.00	1.00
97.00	1.00

## Penyelesaian Kasus

Penyelesaian Kasus Uji Beda 2 Sampel Bebas Independent sample t test

Analyze

Compare Means

Independent Sample t test

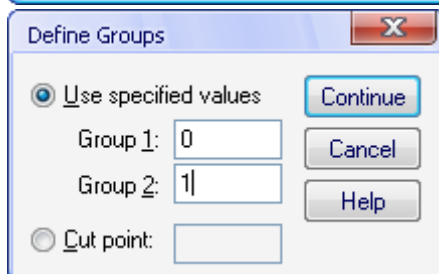
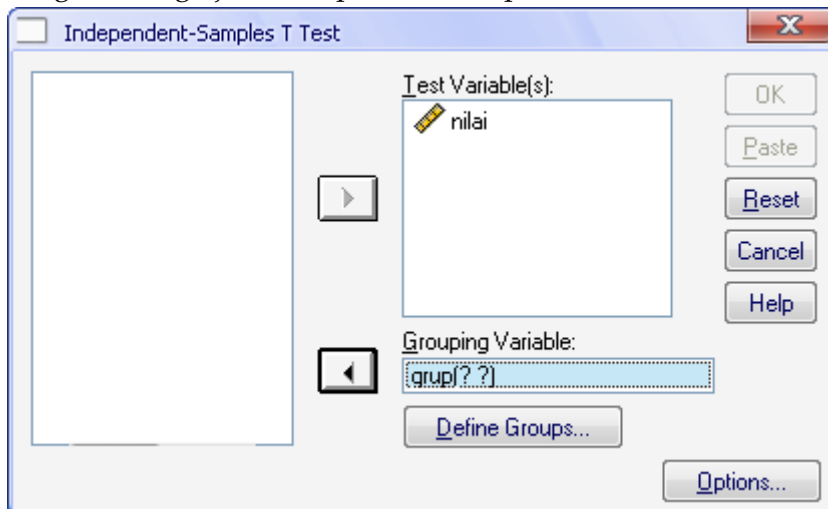
Pada "Test Variables" masukkan variabel "nilai"

Masukkan variabel grup ke "Grouping variables" dan berikan kode 0 untuk "Group 1" dan 1 untuk "Group 2" \*\*  
Klik OK

Keterangan :

\*\*Kode grup dapat diberikan sesuai dengan kode yang diinginkan. Dalam contoh ini, kelompok kontrol diberikan kode 0, dan kelompok uji diberikan kode 1

Langkah Pengerjaan Independent sample t test



Keterangan :

\*\*Kode grup dapat diberikan sesuai dengan kode yang diinginkan. Dalam contoh ini, kelompok kontrol diberikan kode 0, dan kelompok uji diberikan kode 1

Independent samples t Test

**Group Statistics**

	grup	N	Mean	Std. Deviation	Std. Error Mean
nilai	.00	15	72.6667	8.75595	2.26078
	1.00	15	80.0000	7.17137	1.85164

		Levene's Test for Equality of Variances	
		F	Sig.
nilai	Equal variances assumed	2.487	.126
	Equal variances not assumed		

t-test for Equality of Means						
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
					Lower	Upper
-2.509	28	.018	-7.33333	2.92227	-13.3193	-1.34733
-2.509	26.954	.018	-7.33333	2.92227	-13.3298	-1.33684

**Group Statistics** : memberikan informasi mengenai kelompok yang diidentifikasi, ukuran sampel (N), nilai rata-rata, deviasi standar mereka, dan standard error of estimate (SEM). Melalui Tampilan tabel Group Statistics terlihat bahwa rata-rata nilai siswa pada kelompok 0 (kelompok control) adalah sebesar 72.66 sedangkan nilai pada kelompok 1 (kelompok uji) adalah 80.00.

#### Levene Test

Melalui Uji Levene untuk kesamaan varians pada dua kelompok skor dapat dilihat bahwa nilai sig adalah sebesar 0.126. Karena  $\text{sig} > 0.05$  ( $0.126 > 0.05$ ), maka dapat dinyatakan bahwa asumsi homogenitas varians terpenuhi.

#### Independent Samples Test

Karena asumsi homogenitas dapat dipenuhi, maka selanjutnya adalah menguji apakah ada perbedaan nilai pada dua kelompok. Melalui tampilan tabel di atas terlihat bahwa nilai t hitung (two-tailed) adalah sebesar  $-2,509$ ,  $p=0,018$ . Dengan demikian,  $H_0$  ditolak dan  $H_1$  diterima. Perbedaan rata-rata adalah sebesar 7.33 adalah signifikan.

#### Kesimpulan

Dapat disimpulkan bahwa ada perbedaan hasil belajar siswa melalui penggunaan metode diskusi kelompok dengan selisih nilai rata-rata sebesar 7.33. Atau, melalui pemberian metode pembelajaran diskusi kelompok, ada kenaikan rata-rata nilai belajar siswa sebesar 7.33 dan terjadi secara signifikan.

### D. Uji Beda k Sampel Berhubungan

#### Analysis of variance (ANOVA)

Analysis of variance atau ANOVA merupakan salah satu teknik analisis multivariate yang berfungsi untuk membedakan rerata lebih dari dua kelompok data dengan cara membandingkan variansinya. Analisis varian termasuk dalam kategori statistik parametric. Sebagai alat statistika parametric, maka untuk dapat menggunakan

rumus ANOVA harus terlebih dahulu perlu dilakukan uji asumsi meliputi normalitas, heterokedastisitas dan random sampling (Ghozali, 2011).<sup>8</sup>

Analisis varian dapat dilakukan untuk menganalisis data yang berasal dari berbagai macam jenis dan desain penelitian. Analisis varian banyak dipergunakan pada penelitian-penelitian yang banyak melibatkan pengujian komparatif yaitu menguji variabel terikat dengan caramembandingkannya pada kelompok2 sampel independen yang diamati. Analisis varian saat ini banyak digunakan dalam penelitian survey dan penelitian eksperimen.

### One-Way ANOVA

ANOVA satu jalur (One-Way ANOVA)digunakan untuk mengujiperbedaan antaradua atau lebih kelompok.Contohnya adalah menguji perbedaan rata-rata gaji pada kelompok karyawan berdasarkan jenis pekerjaan.

#### Asumsi ONE-WAY ANOVA

Pengamatan diperoleh secara independen dan acak dari populasi

Populasi pada setiap kelompok berdistribusi normal

Populasi pada setiap kelompok memiliki variansi yang sama

#### Contoh Kasus 3D : One-Way ANOVA

Seorang peneliti ingin menguji apakah ada perbedaan rata-rata gaji pegawai level supervisor pada tiga kelompok industry yaitu Manufaktur, Perbankan, dan Pertambangan. Data diambil dari 15 perusahaan yang diambil secara acak.

id	educ	sector	gender	salary
1	16	1	0.00	5,000
2	12	2	0.00	6,400
3	8	2	0.00	7,400
4	15	3	0.00	4,400
5	15	2	1.00	5,000
6	12	1	1.00	7,600
7	15	3	0.00	4,400
8	12	3	1.00	4,200
9	16	3	1.00	3,600
10	8	1	1.00	7,000
11	15	2	0.00	8,400
12	15	2	0.00	7,200
13	12	1	0.00	7,000
14	12	2	0.00	6,000
15	12	3	0.00	3,600

<sup>8</sup>Ghozali, I. (2011). Aplikasi Analisis Multivarite dengan Program IBM-SPSS 19. (Semarang, BP-UNDIP), hlm. 71



16	16	3	1.00	4,000
17	15	3	1.00	4,000
18	15	2	0.00	6,800
19	15	1	0.00	7,200
20	15	1	0.00	7,600
21	12	2	0.00	5,600
22	19	2	0.00	4,800
23	15	1	0.00	8,200
24	8	3	1.00	7,000
25	12	1	0.00	7,600
26	15	1	0.00	7,400
27	16	2	1.00	7,600
28	15	1	0.00	6,800
29	12	3	1.00	5,000
30	15	1	1.00	6,400
31	12	1	1.00	7,000
32	8	2	1.00	6,400
33	15	1	0.00	6,600
34	12	1	0.00	7,000
35	15	2	1.00	5,800
36	12	3	1.00	5,200
37	12	2	1.00	7,000
38	12	1	1.00	7,600
39	15	2	1.00	5,400
40	15	2	1.00	5,800
41	15	3	1.00	4,400
42	15	3	1.00	4,800
43	12	3	0.00	4,200
44	8	3	1.00	3,600
45	12	3	1.00	4,400

### Langkah Penyelesaian

#### ONE-WAY ANOVA

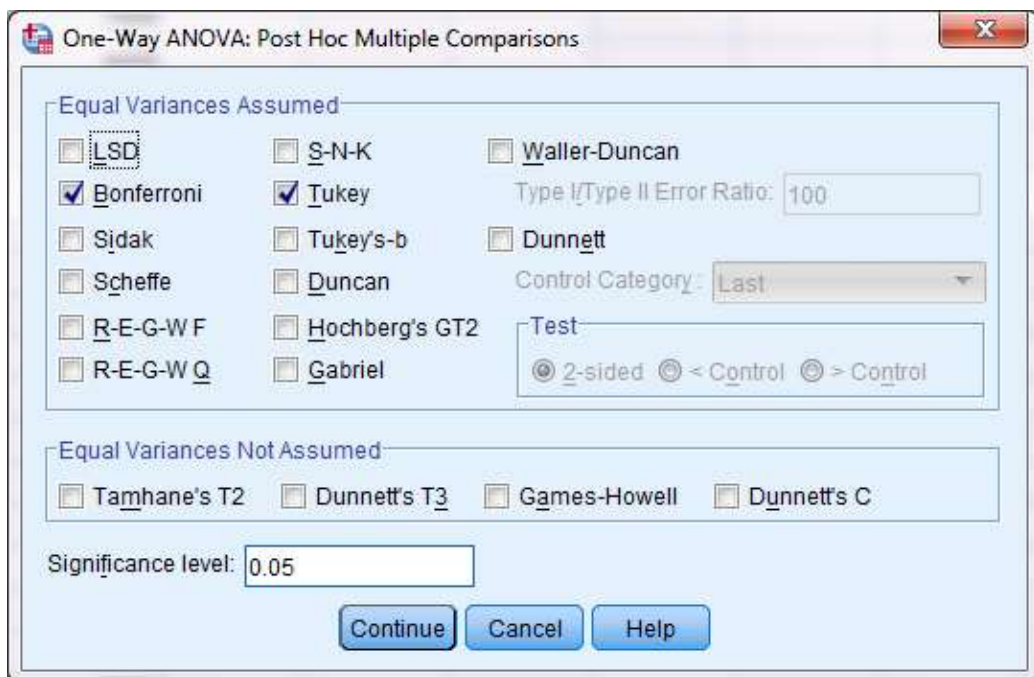
Step 1. Analyze > Compare Means > One-Way ANOVA

Masukkan Variabel "Salary" ke "Dependent List" dan "sector pekerjaan" ke "Factor".

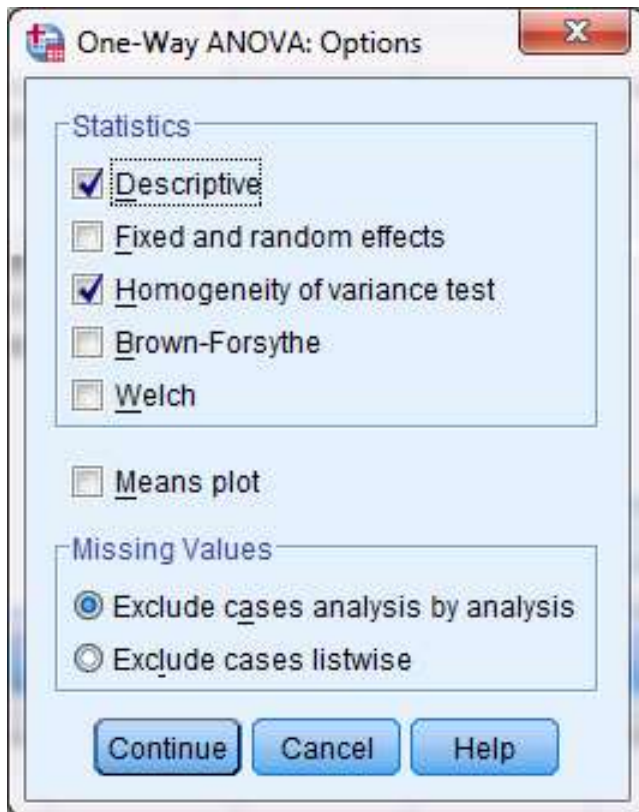
Klik "Post Hoc"



Step 2. Pilih "Bonferroni" dan "Tukey" kemudian "continue"  
Klik "Options"



Step 3. Pada pilihan Options, aktifkan "Descriptive" dan "homogeneity of variance test"  
Klik Continue, kemudian OK



### Hasil dan Pembahasan

Hasil One-Way ANOVA

### Descriptives

Salary (dalam ribuan)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Manufaktur	15	7066.67	735.495	189.904	6659.36	7473.97	5000	8200
Perbankan	15	6373.33	1022.229	263.938	5807.24	6939.42	4800	8400
Pertambangan	15	4453.33	853.452	220.360	3980.71	4925.96	3600	7000
Total	45	5964.44	1409.248	210.078	5541.06	6387.83	3600	8400

### Test of Homogeneity of Variances

Salary (dalam ribuan)

Levene Statistic	df1	df2	Sig.
1.424	2	42	.252

### ANOVA

Salary (dalam ribuan)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	54983111.11	2	27491555.55	35.637	.000
Within Groups	32400000.00	42	771428.571		
Total	87383111.11	44			

Berdasarkan output Deskriptif diperoleh rata-rata gaji untuk pekerja sector manufaktur adalah 7.066 juta, perbankan 6.373 juta dan sector pertambangan 4.453 juta rupiah. Gaji minimal sector manufaktur adalah 5 juta dan maksimal 8.2 juta. Sedangkan gaji minimal pekerja di sector perbankan adalah sebesar 4.8 juta dan maksimal 8.4 juta. Gaji minimal pekerja di sector pertambangan adalah sebesar 3.6 juta dan tertinggi sebesar 7 juta.

Standar deviasi terendah adalah di sector manufaktur, sementara yang tertinggi adalah di sector perbankan. Nilai ini menunjukkan keseragaman data, sehingga semakin besar nilai standar deviasi menunjukkan semakin besarnya ketidakseragaman data. Dengan demikian, dapat disimpulkan bahwa sector perbankan menunjukkan variasi yang tinggi dalam hal gaji.

### Test Homogenitas Varians

Berdasarkan hasil pengujian dengan Levene test diperoleh nilai sebesar 1.424 dengan sig 0.252. Nilai sig > 0.05 menandakan bahwa ada variance adalah sama. Dengan demikian dapat dinyatakan bahwa data memenuhi asumsi homogenitas varians.

### Pengujian Hipotesis

Output Anova menunjukkan bahwa nilai F hitung adalah sebesar 35.637 dengan sig 0.000. Dengan demikian dapat disimpulkan bahwa  $H_0$  ditolak, atau terdapat perbedaan yang signifikan antara gaji kelompok sector manufaktur, perbankan dan pertambangan. Tabel 13.23. Post Hoc Test

Multiple Comparisons							
Dependent Variable: Salary (dalam ribuan)							
(i) sektor pekerjaan	(j) sektor pekerjaan	Mean Difference (i-j)	Std. Error	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
Tukey HSD	Manufaktur	Perbankan	.693.333	320.713	.090	-85.84	1472.50
		Pertambangan	-.2613.333 <sup>a</sup>	320.713	.000	1834.16	3392.50
	Perbankan	Manufaktur	-.693.333	320.713	.090	-1472.50	85.84
		Pertambangan	1.920.000 <sup>a</sup>	320.713	.000	1140.63	2699.17
	Pertambangan	Manufaktur	-.2613.333 <sup>a</sup>	320.713	.000	-3392.50	-1834.16
		Perbankan	-1.920.000 <sup>a</sup>	320.713	.000	-2699.17	-1140.63
Bonferroni	Manufaktur	Perbankan	.693.333	320.713	.109	-106.42	1493.09
		Pertambangan	-.2613.333 <sup>a</sup>	320.713	.000	1813.58	3413.09
	Perbankan	Manufaktur	-.693.333	320.713	.109	-1493.09	106.42
		Pertambangan	1.920.000 <sup>a</sup>	320.713	.000	1120.25	2719.75
	Pertambangan	Manufaktur	-.2613.333 <sup>a</sup>	320.713	.000	-3413.09	-1813.58
		Perbankan	-1.920.000 <sup>a</sup>	320.713	.000	-2719.75	-1120.25

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

Uji Post Hoc Tukey menunjukkan perbedaan antar kelompok yang dirangkum sebagai berikut :

- Manufaktur dengan Perbankan : Tidak berbeda (sig 0.090 > 0.05)
- Manufaktur dengan Pertambangan : Berbeda (sig 0.000 < 0.05)
- Perbankan dengan Pertambangan : Berbeda (sig 0.000 < 0.05)

#### Homogeneous Subsets

Salary (dalam ribuan)				
sektor pekerjaan	N	Subset for alpha = 0.05		
		1	2	
Tukey HSD <sup>a</sup>				
Pertambangan	15	4453.33		
Perbankan	15		6373.33	
Manufaktur	15		7066.67	
Sig.		1.000	.090	

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 15.000.

Lalu pada bagian uji Post Hoc Tukey, perbedaan tiap kelompok dapat dilihat dari nilai harmonic mean yang dihasilkan tiap kelompok berada dalam kolom subset yang sama atau berbeda. Pada hasil uji menunjukkan ketiga kelompok sampel berada pada dua kolom subset. Kelompok pertambangan ke dalam kolom 1 sedangkan perbankan dan manufaktur masuk ke kolom 2. Ini mengindikasikan bahwa sector perbankan dan manufaktur memiliki rata-rata gaji yang relative sama, sedangkan sector pertambangan berbeda dibandingkan dua kelompok sampel lainnya.

## PERTEMUAN 12 UJI RELIABILITAS DAN VALIDITAS

1. **Capaian Pembelajaran:** Setelah mengikuti praktikum, mahasiswa diharapkan mampu memahami, menjelaskan, dan mengaplikasikan uji reliabilitas dan validitas.
2. **Kemampuan Akhir yang diharapkan:** Setelah mengikuti praktikum, Mahasiswa diharapkan mampu memahami, menjelaskan, dan mempraktikkan analisis data uji reliabilitas dan validitas menggunakan SPSS.
3. **Pokok Bahasan : Pengenalan Statistika**
4. **Sub Pokok Bahasan :**
  - a. Reliabilitas
  - b. Validitas

### 5. Materi

#### Definisi Validitas dan Reliabilitas

Validitas mengacu pada aspek ketepatan dan kecermatan hasil pengukuran. Pengukuran sendiri dilakukan untuk mengetahui seberapa banyak aspek (dalam arti kuantitatif) suatu aspek psikologis terdapat dalam diri seseorang, yang dinyatakan oleh skornya pada instrument pengukur yang bersangkutan. Dalam konteks Teori Skor-murni Klasik, Azwar (2012) lebih lanjut menjelaskan bahwa makna validitas dapat dinyatakan sebagai sejauhmana besaran skor-tampak  $X$  mampu mendekati besaran skor-murni  $T$ . Semakin skor-tampak mendekati skor-murni berarti semakin tinggi validitas dan sebaliknya sebaliknya, semakin rendah validitas hasil pengukuran berarti semakin besar perbedaan skor-tampak dari skor-murni.<sup>9</sup>

Jika reliabilitas menyangkut isu-isu konsistensi, maka validitas mengakses akurasi test. Validitas adalah sejauh alat ukur (tes) benar-benar menggambarkan apa yang hendak diukur. Menetapkan validitas sebuah test atau instrument test sangat sulit, terutama karena variabel-variabel psikologi biasanya adalah konsep-konsep abstrak, seperti inteligensi, kecemasan, dan kepribadian. Konsep-konsep ini tidak memiliki realitas konkret sehingga eksistensinya harus diinferensi melalui sarana yang tidak langsung. Ada tiga jenis validitas yaitu content validity (validitas isi), criterion validity (validitas kriteria), dan construct validity (validitas konsep).

Reliabilitas suatu test merujuk pada derajat stabilitas, konsistensi, daya prediksi, dan akurasi. Uji ini dilakukan untuk melihat seberapa skor-skor yang diperoleh seseorang itu akan menjadi sama jika orang itu diperiksa ulang dengan tes yang sama pada kesempatan berbeda. Beberapa teknik yang sering digunakan untuk menguji reliabilitas instrument adalah stabilitas pengukuran yang dapat diperoleh melalui test-retest, dan parallel form reliability, dan konsistensi ukuran yang diperoleh melalui reliabilitas belah dua (split-half), koefisien alpha, reliabilitas antar skor. Menurut Sekaran (2006), reliabilitas atau keandalan suatu pengukuran menunjukkan sejauh mana pengukuran tersebut tanpa bias (bebas dari kesalahan) dan karena itu menjamin pengukuran yang konsisten lintas waktu dan lintas beragam item dalam instrument.

---

<sup>9</sup> Azwar, S..., hal 40 - 41

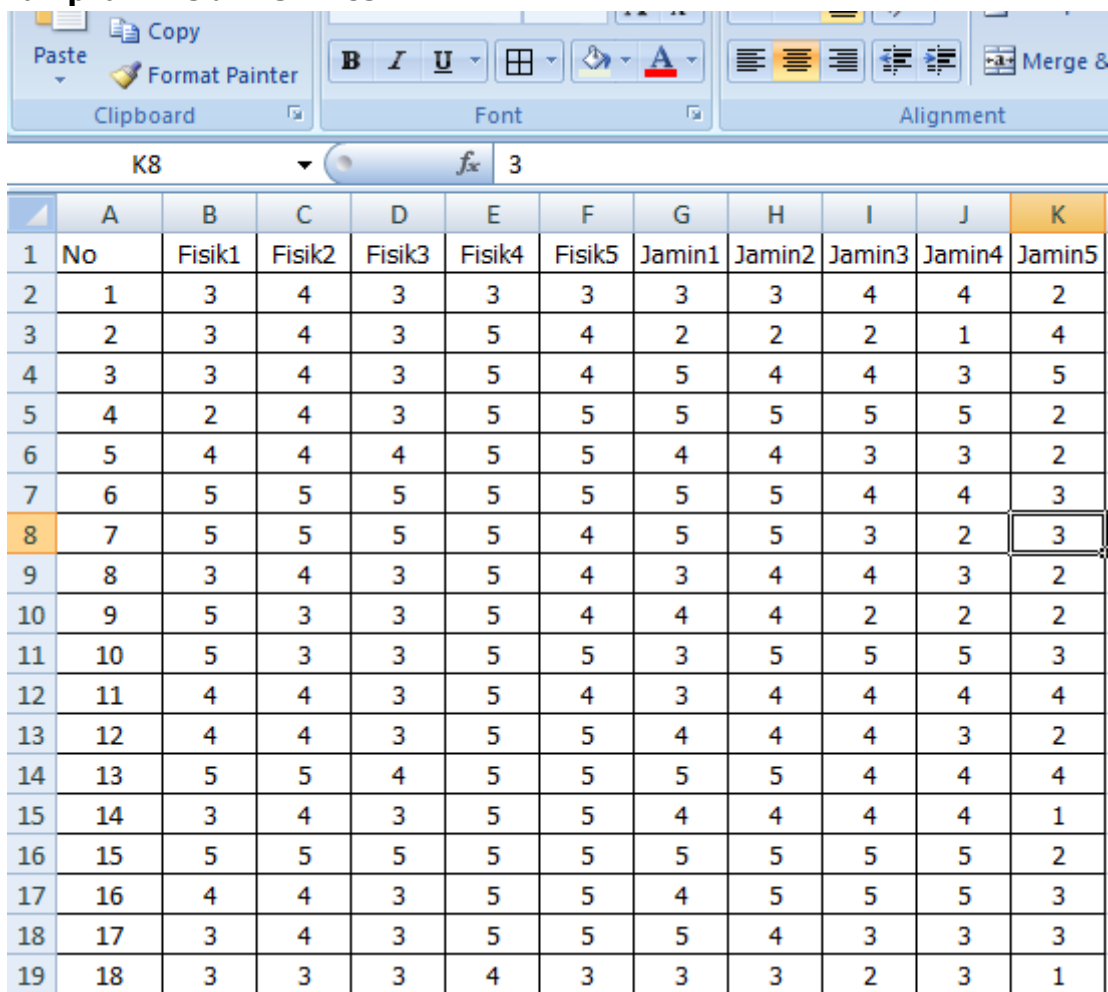
Dengan kata lain, keandalan suatu pengukuran merupakan indikasi mengenai stabilitas dan konsistensi di mana instrument mengukur konsep dan membantu menilai “ketepatan” sebuah pengukuran.<sup>10</sup>

#### A. Kasus 1 : Uji Validitas dan Reliabilitas Kualitas Pelayanan

Konsep yang diukur adalah Kualitas pelayanan. Contoh uji ini hanya menggunakan 2 faktor yaitu faktor fisik (5 item) dan aspek jaminan & keandalan (5 item). Total item yang dipergunakan adalah sebanyak 10 item pertanyaan.

Data dikumpulkan dari 50 orang responden melalui penyebaran kuesioner yang sebagian datanya ditampilkan sebagai berikut :

#### Tampilan File di MS. Excel



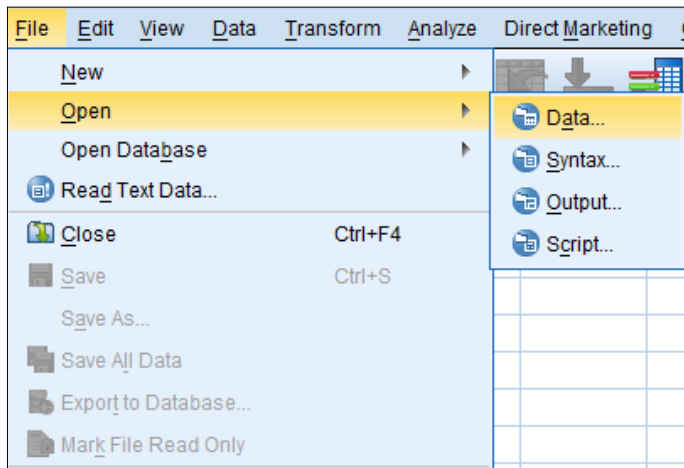
	A	B	C	D	E	F	G	H	I	J	K
1	No	Fisik1	Fisik2	Fisik3	Fisik4	Fisik5	Jamin1	Jamin2	Jamin3	Jamin4	Jamin5
2	1	3	4	3	3	3	3	3	4	4	2
3	2	3	4	3	5	4	2	2	2	1	4
4	3	3	4	3	5	4	5	4	4	3	5
5	4	2	4	3	5	5	5	5	5	5	2
6	5	4	4	4	5	5	4	4	3	3	2
7	6	5	5	5	5	5	5	5	4	4	3
8	7	5	5	5	5	4	5	5	3	2	3
9	8	3	4	3	5	4	3	4	4	3	2
10	9	5	3	3	5	4	4	4	2	2	2
11	10	5	3	3	5	5	3	5	5	5	3
12	11	4	4	3	5	4	3	4	4	4	4
13	12	4	4	3	5	5	4	4	4	3	2
14	13	5	5	4	5	5	5	5	4	4	4
15	14	3	4	3	5	5	4	4	4	4	1
16	15	5	5	5	5	5	5	5	5	5	2
17	16	4	4	3	5	5	4	5	5	5	3
18	17	3	4	3	5	5	5	4	3	3	3
19	18	3	3	3	4	3	3	3	2	3	1

#### IMPORT DATA DARI EXCEL KE SPSS

<sup>10</sup> Sekaran, U..., *Metodologi Penelitian...*, hlm. 40

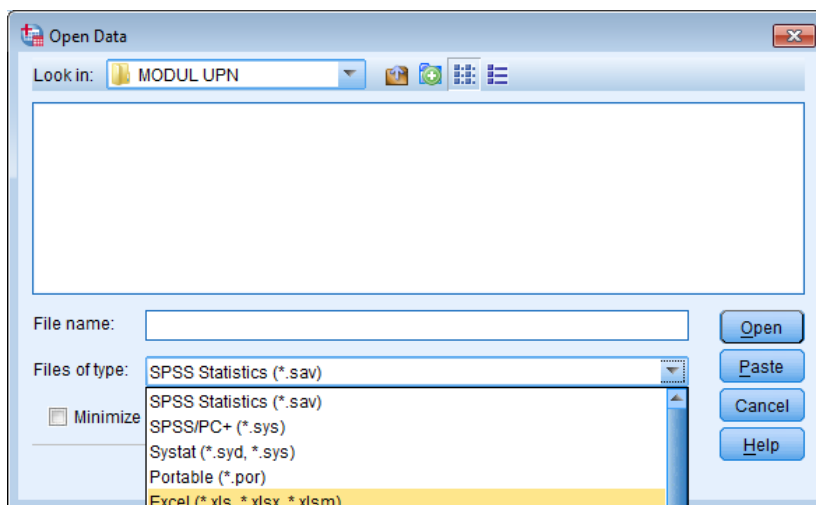
### Tahap 1. Membuka File

Klik File → Open → Data



### Tahap 2. Memilih File yang akan digunakan

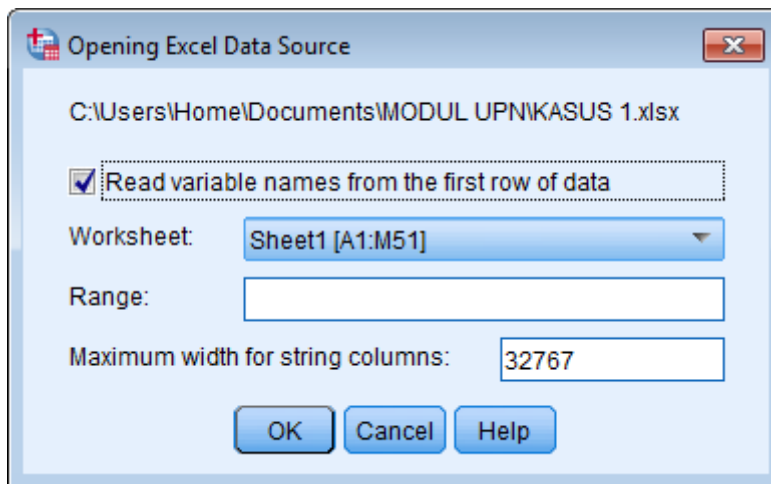
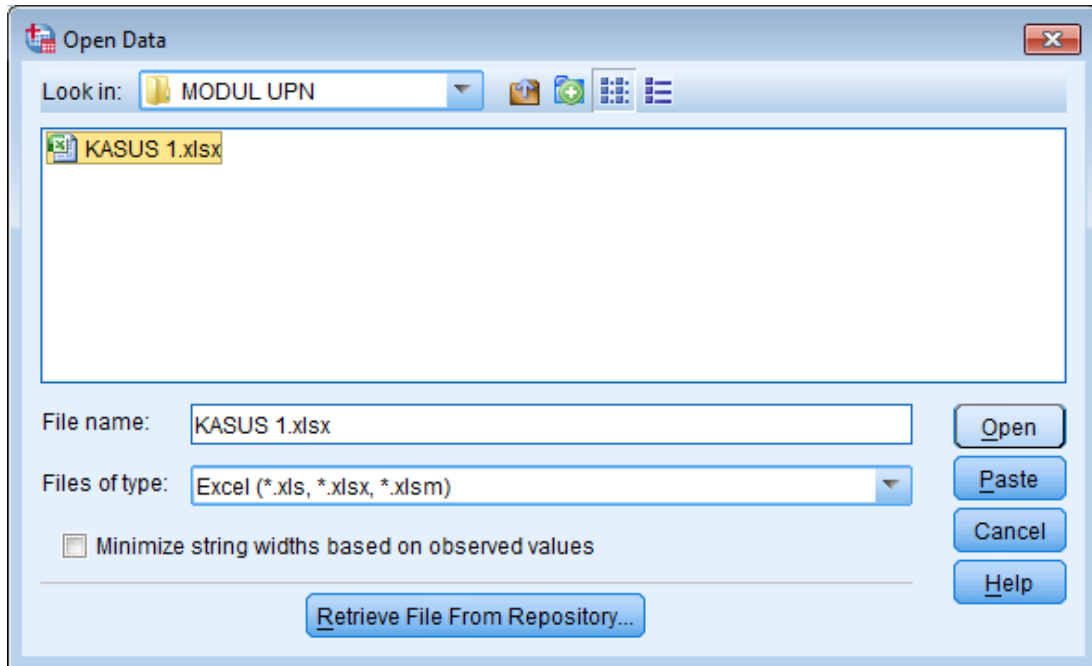
Pada Files of Type..pilih Excel (\*.XLS, \*XLSX)



### Tahap 3. Memilih File yang akan dibuka

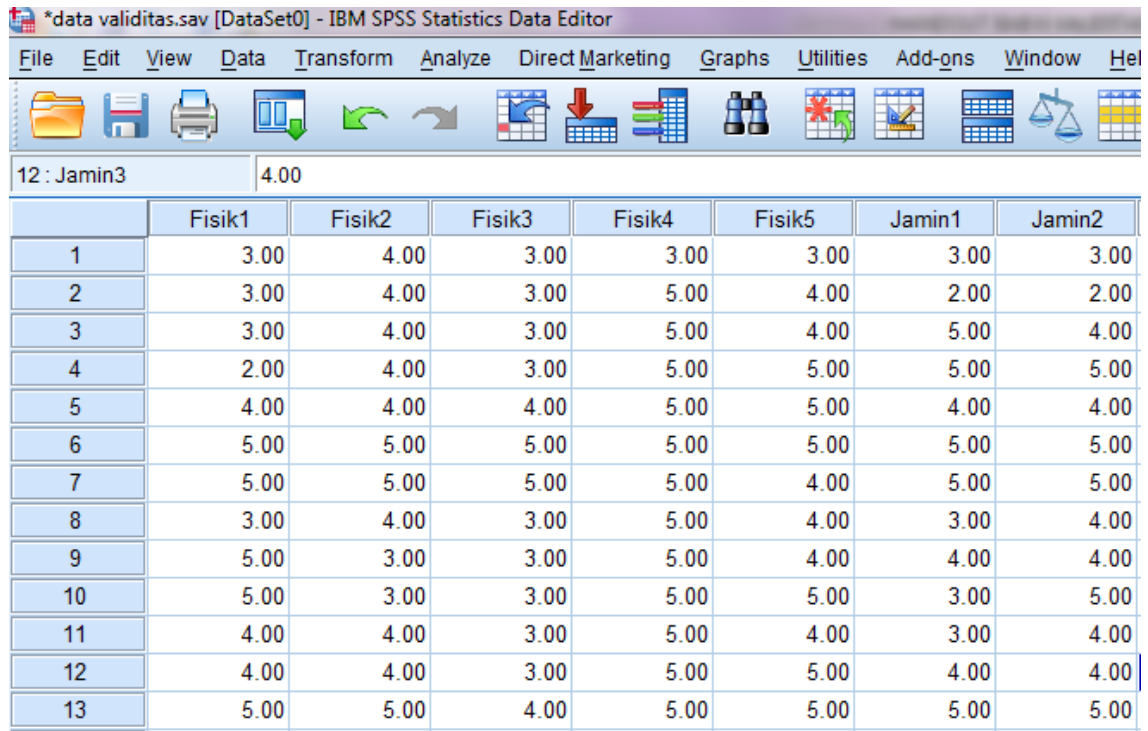
Contoh adalah file KASUS1.xlsx





Klik OK

#### Tahap 4. Tampilan Lembar Kerja SPSS



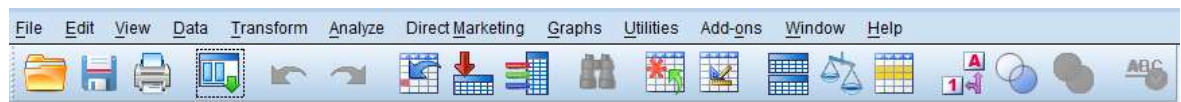
The screenshot shows the IBM SPSS Statistics Data Editor interface. The title bar reads '\*data validitas.sav [DataSet0] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations, editing, and analysis. The main window displays a data table with 13 rows and 8 columns. The columns are labeled Fisik1, Fisik2, Fisik3, Fisik4, Fisik5, Jamin1, and Jamin2. The first column contains row numbers from 1 to 13. The data values are as follows:

	Fisik1	Fisik2	Fisik3	Fisik4	Fisik5	Jamin1	Jamin2
1	3.00	4.00	3.00	3.00	3.00	3.00	3.00
2	3.00	4.00	3.00	5.00	4.00	2.00	2.00
3	3.00	4.00	3.00	5.00	4.00	5.00	4.00
4	2.00	4.00	3.00	5.00	5.00	5.00	5.00
5	4.00	4.00	4.00	5.00	5.00	4.00	4.00
6	5.00	5.00	5.00	5.00	5.00	5.00	5.00
7	5.00	5.00	5.00	5.00	4.00	5.00	5.00
8	3.00	4.00	3.00	5.00	4.00	3.00	4.00
9	5.00	3.00	3.00	5.00	4.00	4.00	4.00
10	5.00	3.00	3.00	5.00	5.00	3.00	5.00
11	4.00	4.00	3.00	5.00	4.00	3.00	4.00
12	4.00	4.00	3.00	5.00	5.00	4.00	4.00
13	5.00	5.00	4.00	5.00	5.00	5.00	5.00

Cara yang lain juga dapat dilakukan dengan mengcopy file dari Excel ke SPSS dengan Metode Copy → Paste

#### SETTING DATA

##### Tampilan Berbagai Menu



##### Data View dan Variable View

10	3.00	3.00
19	3.00	3.00
20	5.00	5.00
21	5.00	5.00
22	5.00	5.00
23	4.00	4.00
1		
Data View	Variable View	

Data view adalah tab yang berisi data, sedangkan variable view adalah tab yang berisi tentang variabel seperti tampilan dibawah ini

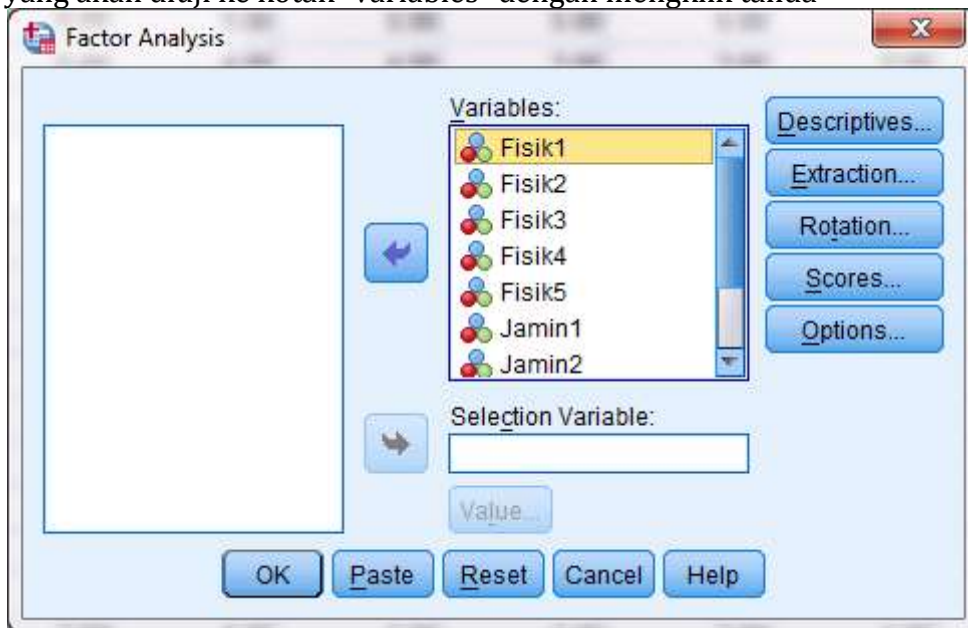
### Tampilan Variable View

	Name	Type	Width	Decimals	Label	Values	Measure
1	Fisik1	Numeric	8	2		None	None
2	Fisik2	Numeric	8	2		None	None
3	Fisik3	Numeric	8	2		None	None
4	Fisik4	Numeric	8	2		None	None
5	Fisik5	Numeric	8	2		None	None
6	Jamin1	Numeric	8	2		None	None
7	Jamin2	Numeric	8	2		None	None
8	Jamin3	Numeric	8	2		None	None
9	Jamin4	Numeric	8	2		None	None
10	Jamin5	Numeric	8	2		None	None

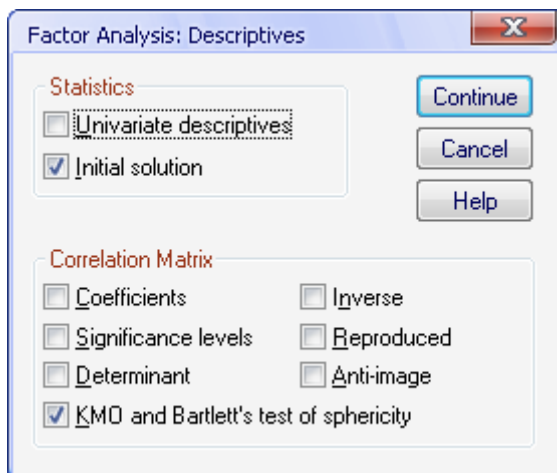
Keterangan :

- Name adalah nama variabel
- Type adalah tipe data yang digunakan (contoh numeric, string, etc)
- Width adalah lebar kolom
- Label adalah alternative nama variabel
- Values dapat digunakan untuk memberikan keterangan atau nilai tertentu pada skor
- Measure adalah tipe / jenis data yang digunakan
- Dst

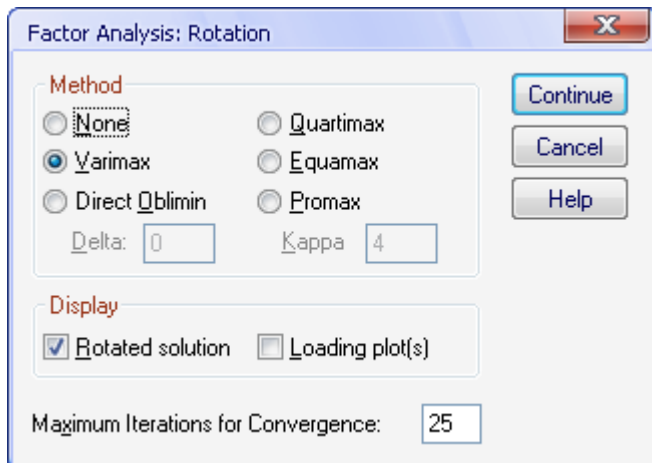
Step 2 : Klik Analyze, Pilih Dimension Reduction. Selanjutnya, masukkan variabel yang akan diuji ke kotak “variables” dengan mengklik tanda “>”



Step 3. Klik Descriptive, lalu beri tick pada pilihan “KMO and Bartlett’s Test of Sphericity”. Klik Continue



Step 4. Klik Rotations, lalu pilih “Varimax”, dan klik “Continue”



Klik OK

### Hasil Dan Pembahasan Uji Validitas

Prinsip yang digunakan analisis factor adalah mengelompokkan data berdasarkan interkorelasi antar butir. Parameter pertama adalah dengan memperhatikan nilai *measure of sampling adequacy* (MSA). Menurut Hair, dkk (2006), nilai MSA berkisar antara 0 – 1 ketika setiap variabel diprediksi secara sempurna tanpa error dari variabel lain. Nilai 0.80 atau lebih dianggap pantas, 0.70 atau lebih dianggap lumayan, 0.60 atau lebih dianggap sedang, 0.50 atau lebih dianggap menyedihkan/buruk, dan di bawah 0.50 tidak dapat diterima. MSA dapat meningkat ketika : (1) sampel meningkat; (2) rata-rata korelasi meningkat; (3) jumlah variabel meningkat; atau jumlah factor berkurang.

MSA merupakan keputusan awal untuk mengevaluasi apakah analisis factor dapat diteruskan atau tidak. Selanjutnya memperhatikan nilai Bartlett test of sphericity, yaitu uji statistic untuk menguji korelasi antar butir yang dapat membentuk satu factor (Hair, dkk, 2006). Ketiga adalah memperhatikan nilai total variance explained yaitu nilai eigenvalues > 1 dianggap dapat membentuk faktor, dan keempat adalah dengan memperhatikan nilai muatan factor (loading factor), yaitu sebuah butir / item dinyatakan merupakan pembentuk factor jika nilai muatan factor lebih besar sama dengan ( $\geq$ ) 0,50.

Tabel 1. KMO-MSA kasus 1

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.833
Bartlett's Test of Sphericity	Approx. Chi-Square	410.377
	df	45
	Sig.	.000

Nilai *Kaiser-Meyer-Olkin MSA (KMO-MSA)* sebesar 0,833 sudah menunjukkan bahwa hasil yang baik ( $> 0.60$ ). Dengan demikian, syarat analisis faktor pertama dapat dipenuhi sehingga dapat dilanjutkan untuk pengujian selanjutnya. Kedua, nilai *Bartlett's Test of Sphericity* sebesar 410.377 pada signifikansi 0,000. Nilai ini menandakan bahwa faktor pembentuk variabel sudah baik.

Tabel 2. Total Variance Explained

Componen t	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	5.828	58.284	58.284
2	1.535	15.351	73.636
3	.808	8.079	81.715
4	.662	6.617	88.332
5	.397	3.967	92.299
6	.264	2.635	94.934
7	.191	1.913	96.847
8	.120	1.199	98.046
9	.115	1.155	99.200
10	.080	.800	100.000

Berdasarkan hasil uji total variance explained diketahui bahwa 2 faktor yang digunakan mampu menjelaskan 73.636% dari keseluruhan konstruk kualitas pelayanan. Dapat disimpulkan bahwa melalui analisis faktor berhasil dibuktikan bahwa 10 item yang digunakan untuk mengukur variabel kualitas pelayanan dikelompokkan menjadi 2 faktor yaitu yaitu faktor 1 yang menjelaskan 58.284% varian, faktor 2 menjelaskan 15.351% varian konstruk.

Tabel 3. Rotasi Komponen

**Rotated Component Matrix<sup>a</sup>**

	Component	
	1	2
Fisik1	.157	.797
Fisik2	.380	.835
Fisik3	.365	.762
Fisik4	.055	.824
Fisik5	.424	.732
Jamin1	.810	.368
Jamin2	.840	.358
Jamin3	.857	.234
Jamin4	.899	.224
Jamin5	.686	.092

Koefisien bobot faktor harus > 0.50 untuk dinyatakan valid.

Cth :

Fisik1 0.797 > 0.50 = valid

Fisik2 0.835 > 0.50 = valid

dst

Hasil *Rotated component matrix* menunjukkan bahwa item Fisik1 – Fisik5 mengelompok pada faktor 2, item Jamin1 – Jamin5 mengelompok pada faktor 1. Dari hasil ini jelas dapat disimpulkan bahwa kontrak kualitas pelayanan merupakan kontrak multidimensi yang terdiri dari 2 faktor yaitu Faktor Fisik dan Faktor Jaminan. Seluruh item dinyatakan valid dengan memperhatikan nilai koefisien bobot faktor > 0.50.

Jika ada item yang memiliki nilai koefisien bobot faktor < 0.50, maka item tersebut dapat dieliminasi dan kemudian item diuji ulang (kembali ke step 2) dengan mengeluarkan item yang dinyatakan tidak valid.

### Uji Reliabilitas

Cronbach-Alpha digunakan jika sudah ada bukti bahwa item yang akan diuji memiliki unidimensional. Dalam hal ini, hasil uji analisis factor menunjukkan bahwa item Fisik1 s/d Fisik5 menjadi 2 faktor, dan item Jamin1 s/d Jamin5 menjadi Faktor 1

Langkah Pengujian :

#### Penyelesaian Reliabilitas Analyze

Scale

Reliability Analysis

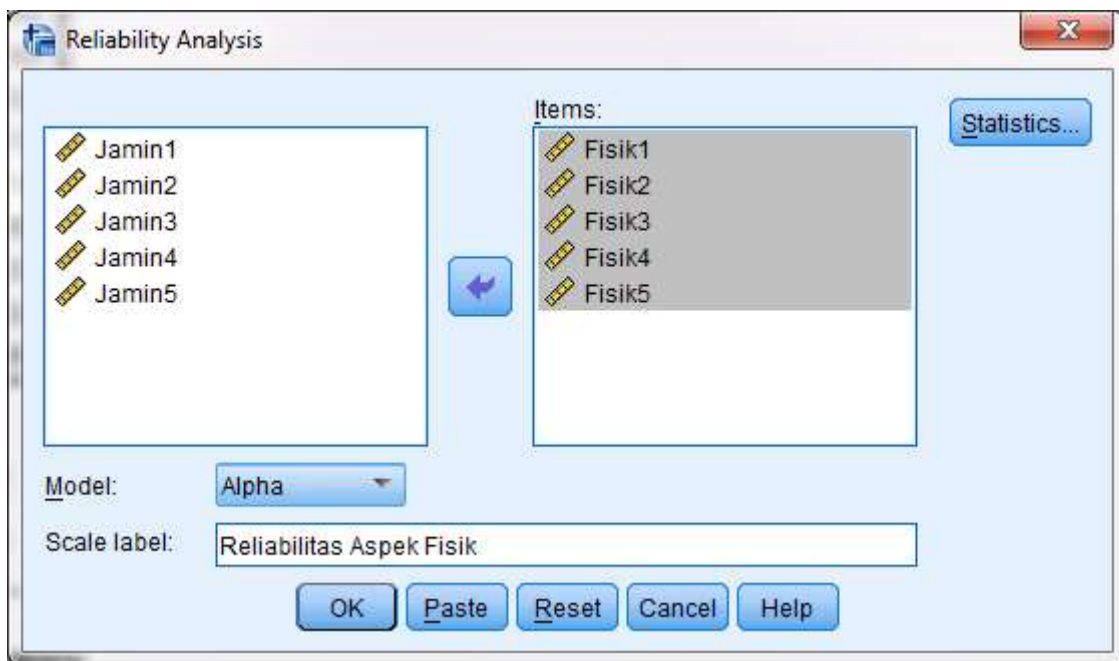
**Variables** : masukkan aitem AC1 – AC4

OK

Masih menggunakan data yang sama.

Klik Analyze, kemudian pilih Scale → Reliability Analysis, kemudian Masukkan Fisik1 - Fisik5 ke dalam Kotak "Items"

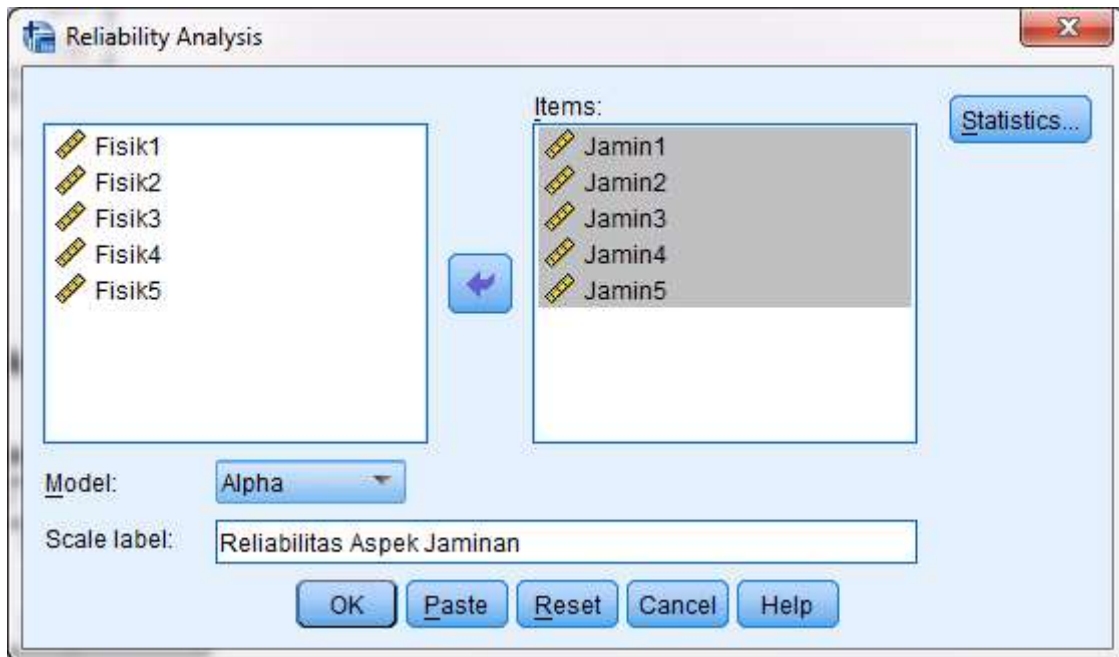
Ketik nama Reliabilitas Aspek Fisik pada Kotak "scale test"



Klik OK

Ulangi langkah di atas dengan mengganti Jamin1 - Jamin5 ke Kotak "Items"





Klik OK

## Hasil dan Penjelasan

### Scale: Reliabilitas Aspek Fisik

#### Case Processing Summary

		N	%
Cases	Valid	50	100.0
	Excluded <sup>a</sup>	0	.0
	Total	50	100.0

a. Listwise deletion based on all variables in the procedure.

#### Reliability Statistics

Cronbach's Alpha	N of Items
.891	5

## Scale: Reliabilitas Aspek Jaminan

### Case Processing Summary

		N	%
Cases	Valid	50	100.0
	Excluded <sup>a</sup>	0	.0
	Total	50	100.0

a. Listwise deletion based on all variables in the procedure.

### Reliability Statistics

Cronbach's Alpha	N of Items
.910	5

Hasil Cronbach Alpha Koefisien memperlihatkan angka sebesar 0.891 untuk Aspek Fisik dan 0.910 untuk Aspek Jaminan. Karena nilai Cronbach Alpha > 0.70 maka dapat disimpulkan bahwa dua faktor yang diuji telah memenuhi syarat keandalan (reliabilitas) yang baik.

## Referensi

- Anastasi, A., & Urbina, S. (1997). *Psychological testing (7th ed.)*. Upper Saddle River, NJ: Prentice-Hall.
- Azwar, S. (2012). *Reliabilitas dan Validitas*. Edisi 4. Yogyakarta : Pustaka Pelajar.
- Hair, dkk. (2006). *Multivariate Data Analysis*. Sixth Edition. New Jersey : Pearson Education
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.)*. New York: McGraw-Hill

**Data Penelitian**

No	Fisik 1	Fisik 2	Fisik 3	Fisik 4	Fisik 5	Jamin 1	Jamin 2	Jamin 3	Jamin 4	Jamin 5
1	3	4	3	3	3	3	3	4	4	2
2	3	4	3	5	4	2	2	2	1	2
3	3	4	3	5	4	5	4	4	3	4
4	2	4	3	5	5	5	5	5	5	2
5	4	4	4	5	5	4	4	3	3	2
6	5	5	5	5	5	5	5	4	4	3
7	5	5	5	5	4	5	5	3	2	3
8	3	4	3	5	4	3	4	4	3	2
9	5	3	3	5	4	2	2	2	2	2
10	5	3	3	5	5	3	5	5	5	3
11	4	4	3	5	4	3	4	4	4	4
12	4	4	3	5	5	4	4	4	3	2
13	5	5	4	5	5	5	5	4	4	4
14	3	4	3	5	5	4	4	4	4	4
15	5	5	5	5	5	5	5	5	5	2
16	4	4	3	5	5	3	3	3	3	3
17	3	4	3	5	5	5	4	3	3	3
18	3	3	3	4	3	1	1	2	1	1
19	3	3	3	5	3	3	3	1	3	4
20	5	5	5	5	5	5	5	5	5	3
21	5	5	5	5	4	2	2	2	2	2
22	5	5	5	5	5	5	5	5	5	5
23	4	4	3	5	4	4	4	3	3	5
24	5	5	5	5	5	4	4	4	4	5
25	2	1	2	3	2	2	3	2	2	3
26	4	4	4	4	4	4	4	3	3	4
27	3	3	3	5	4	3	4	3	3	4
28	4	4	4	3	3	3	4	4	4	4
29	5	5	5	5	5	5	5	5	5	5
30	3	4	3	5	5	4	4	3	3	3
31	5	5	5	5	5	5	5	5	4	3
32	5	4	5	5	5	4	5	3	3	5
33	5	5	5	5	5	5	5	4	4	4

34	4	4	4	4	5	5	5	3	3	3
35	5	5	5	5	5	5	5	5	5	5
36	4	4	3	4	5	4	5	3	3	4
37	5	5	5	5	5	5	5	5	5	5
38	3	4	3	3	4	3	4	3	3	3
39	5	4	5	5	5	5	5	5	5	5
40	2	3	2	4	5	4	3	4	4	5
41	5	5	5	5	5	5	5	5	5	5
42	5	5	5	5	5	5	4	4	4	4
43	5	5	5	5	5	4	5	5	5	3
44	1	4	4	5	5	5	5	4	5	5
45	4	4	2	5	5	3	3	3	3	3
46	5	5	5	5	5	5	5	5	5	5
47	5	5	5	5	5	5	5	5	5	5
48	3	4	4	4	4	4	5	5	4	4
49	3	3	2	2	3	5	5	5	5	4
50	2	1	2	1	1	2	1	2	1	2

### JENIS RELIABILITAS

Jenis Reliabilitas	Metode / Formula	Penjelasan
Tes Ulang	Test-Retest	Dilakukan dengan melakukan korelasi antara hasil tes 1 dengan tes 2. Teknik korelasi yang digunakan adalah Product Moment Pearson. Antara tes 1 dan tes 2 dipisahkan oleh waktu tertentu sehingga test-retest menggambarkan konsistensi antar waktu
	Paralel	Sama dengan Test-Retest, namun memiliki perbedaan dimana tes dilakukan secara bersamaan. Secara teknis, reliabilitas bentuk parallel merupakan upaya untuk melihat kesetaraan dua hasil tes yang diukur dalam waktu bersamaan.
Konsistensi Internal	Split-Half Spearman-Brown	Reliabilitas belah dua (split-half) dilakukan dengan mengkorelasikan dua kelompok belahan test (misalnya, ganjil - genap) kemudian menghitungnya dengan formula Spearman-Brown. Teknik ini dapat dilakukan jika peneliti memiliki keyakinan bahwa dua kelompok belahan test adalah parallel.
	Alpha-Cronbach	Alpha Cronbach adalah formula yang paling umum digunakan untuk mengestimasi reliabilitas pada skor

		yang menggunakan skala Likert. Alpha Cronbach merupakan bentuk konsistensi internal sehingga sebelum menguji reliabilitas dengan Alpha Cronbach, item-item yang digunakan harus memiliki unidimensional yang dapat diuji dengan analisis factor.
	Kuder-Richardson	Kuder-Richardson (KR) merupakan teknik untuk mengestimasi reliabilitas skor dikotomo (0 - 1). KR 21 mengasumsikan bahwa semua pertanyaan memiliki tingkat kesulitan yang sama, sedangkan K-R 20 tidak mengasumsikan hal ini.
	Rulon	Formula Rulon digunakan untuk mengestimasi reliabilitas skor dengan pendekatan belah dua tanpa perlu berasumsi bahwa kedua belahan tersebut mempunyai kesetaraan varian skor diantara dua belahan.

### **JENIS VALIDITAS**

- Validitas Permukaan (*face validity*) → sangat sederhana (dimensi dan item tunggal)
- Validitas Kriteria (*criterion validity*) → apakah sudah mencakup semua aspek
- Validitas Isi (*content validity*) → sejauh mana isi alat ukur telah mewakili semua aspek suatu konsep.
- Validitas konstruk (*construct validity*) → konsep yang paling abstrak
- Validitas prediktif (*predictive validity*) → Instrumen yang dipergunakan dapat memprediksi hasil di masa depan.
- Validitas eksternal (Instrumen dibandingkan dengan alat ukur sejenis)