Kebanyakan dari mahasiswa Ekonomi dan Bisnis memerlukan mengolah data secara statistik. Hal itu menuntut mahasiswa memahami baik secara dasar hingga pada tahap *expert.* Akan tetapi mereka juga banyak mengalami kesulitan dalam mencerna materi-materi yang disampaikan, baik dalam buku maupun materi ketika proses pembelajaran pada mata kuliah statistika.

Maka dari itu, Buku ini disusun untuk memudahkan mahasiswa dalam memahami statistika baik statistika deskriptif maupun inferensial. Pada Bab pertama dan kedua, dijelaskan secara singkat jelas dan padat mengenai statistika beserta penuajian datanya. Kemudian pada Bab-Bab selanjutnya dijelaskan dengan mudah dasar-dasar analisis yang sering digunakan dalam pengujian data statistik ekonomi dan bisnis.

Penulisan pada buku ini cukup ringkas dalam memberikan pemahaman, karena tidak bertele-tele dalam penyampaiannya. Tidak hanya itu, di dalamnya juga disertai contoh-contoh perhitungan software *Statistical Product and Service Solution* (SPSS). Sehingga memberikan cara mudah memahami dan mempraktekan materi statistika.

Akademia Pustaka
Perum. BMW Madani Kavling 16, Tulungagung
https://akademiapustaka.com/
redaksi.akademia.pustaka@gmail.com
@redaksi.akademia.pustaka
@akademiapustaka
081216178398

ISBN 978-623-6364-06-2
9 786236 364062

Statistika Untuk Ekonomi dan Bisnis: Analisis Deskriptif dan Inferensial

STATISTIKA *Untuk* Ekonomi dan Bisnis

Analisis Deskriptif dan Inferensial

Pendahuluan Statistika
Data dan Penyajiannya
Pengukuran Deskriptif

Analisis Korelasi dan Regresi Linear Sederhana
Analisis Regresi Linear Berganda
Analisis Deret Berkala

Angka Indeks
Pengujian Hipotesis
Analisis Regresi Nonlinear

ELOK FITRIANI RAFIKASARI

STATISTIKA UNTUK EKONOMI DAN BISNIS:

Analisis Deskriptif dan Inferensial

Elok Fitriani Rafikasari, M.Si

AKADEMIA PUSTAKA

Statistika Untuk Ekonomi dan Bisnis:
Analisis Deskriptif dan Inferensial

Layouter: Saiful Mustofa
Desain cover: Dicky M. Fauzi
vii + 153 hlm: 14 x 21 cm
Cetakan: Pertama, Juni 2021
ISBN: 978-623-6364-06-2

Anggota IKAPI

Diterbitkan oleh:
Akademia Pustaka
Perum. BMW Madani Kavling 16, Tulungagung
Telp: 081216178398
Email: redaksi.akademia.pustaka@gmail.com
Website: www.akademiapustaka.com

KATA PENGANTAR

Puji syukur Alhamdulillahirobbil Alamin penulis sampaikan kehadirat Allah SWT yang telah memberikan rahmat, taufiq, dan hidayah-Nya kepada penulis sehingga dapat menyelesaikan buku *"Statistika Untuk Ekonomi dan Bisnis: Analisis Deskriptif dan Inferensial"* ini dengan baik dan tepat waktu. Buku ini disusun untuk memudahkan mahasiswa dalam memahami statistika baik statistika deskriptif maupun inferensial.

Penulis menyadari bahwa buku ini masih jauh dari sempurna. Oleh karena itu kritik dan saran yang bersifat membangun sangat penulis harapkan demi sempurnanya penulisan selanjutnya. Akhirnya penulis berharap semoga buku ini bermanfaat bagi kita semua, Aamiin.

Tulungagung, Juni, 2021

Penulis

DAFTAR ISI

BAB 1
PENDAHULUAN

A. Pengertian Statistik dan Statistika

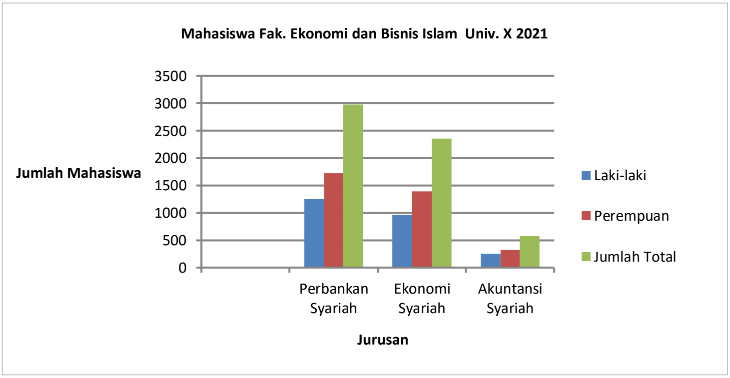
Statistik dan statistika adalah dua hal yang berbeda akan tetapi pada umumnya sebagian orang belum mengerti perbedaan antara keduanya. Berikut akan dijelaskan perbedaan antara statistik dan statistika. Statistik dalam praktiknya menyangkut perhitungan numerik dengan menggunakan rumus tertentu. Hasil perhitungan dengan rumus tersebut akan mewakili sebagian dari sekumpulan data tertentu. Sebagian dari sekumpulan data tersebut disebut sampel dan keseluruhan himpunan data disebut populasi. Wakil dari sebagian dari sekumpulan data itulah yang disebut dengan statistik. Contoh dari statistik adalah mean, median, modus, kuartil, desil, persentil, variansi, standar deviasi, dan lain-lain. Sedangkan statistika adalah sekumpulan metode yang digunakan untuk mengumpulkan, menganalisis, menjelaskan dan menginterpretasikan data dan untuk membuat suatu kesimpulan.

B. Jenis-jenis Statistika

Statistika dapat dibedakan menjadi dua bagian yaitu, statistika deskriptif dan statistika inferensia. Jenis statistika yang pertama adalah statistika deskriptif. Statistika deskriptif merupakan metode untuk mengorganisir, menggambarkan dan mendeskripsikan data dengan menggunakan tabel, diagram dan ringkasan pengukuran yang ada (mean, median, modus, rentang dan simpangan baku). Contoh dari statistika deskriptif adalah sebagai berikut:

Tabel 1.1 Jumlah Mahasiswa Fakultas Ekonomi dan Bisnis Islam di Universitas X Tahun 2021

Jurusan	Banyak Mahasiswa		Jumlah Total
	Laki-laki	Perempuan	
Perbankan Syariah	1257	1723	2980
Ekonomi Syariah	967	1388	2355
Akuntansi Syariah	251	321	572



Gambar 1.1 Jumlah Mahasiswa Fakultas Ekonomi dan Bisnis Islam Universitas X 2021

Jenis statistika yang kedua adalah statistika inferensia. **Statistika inferensia** merupakan metode yang menggunakan hasil dari sampel untuk membuat kesimpulan atau prediksi mengenai populasi. Contoh dari statistika inferensia adalah, misalkan suatu perusahaan tekstil memproduksi pakaian. Untuk melihat kualitas pakaian yang diproduksi perusahaan tersebut memilih beberapa pakaian untuk diperiksa kualitasnya dan membuat kesimpulan tentang kualitas pakaian yang diproduksi.

C. Populasi dan Sampel

Pengertian populasi dan sampel secara umum sudah dijelaskan pada sub bab 1.1. Bagian ini akan menjelaskan populasi dan sampel beserta contoh dalam suatu kasus penelitian. **Populasi** didefinisikan sebagai keseluruhan pengamatan yang menjadi perhatian dalam suatu penelitian baik sejumlah terhingga maupun tak terhingga. Populasi biasa dilambangkan dengan huruf **N** ("N" besar). Sebagai contoh, misalkan seorang peneliti ingin mengetahui tentang tingkat pendapatan semua orang tua mahasiswa IAIN Tulungagung 2021. Pada kasus ini seorang peneliti fokus pada semua orang tua mahasiswa IAIN Tulungagung, **semua orang tua mahasiswa** inilah yang disebut sebagai **populasi** atau biasanya juga disebut sebagai **target populasi**.

Sampel didefinisikan sebagai bagian dari populasi atau sebagian anggota dari populasi yang telah dipilih. Sampel dilambangkan dengan huruf **n** ("n" kecil). Pemilihan anggota populasi menjadi sampel harus representatif, artinya sampel harus dapat mewakili atau dapat menjelaskan keragaman yang ada pada populasi. Sampel dapat berupa sampel acak (*random sample*) atau tidak acak (*nonrandom sample*). Pada *random sample* setiap anggota populasi mempunyai peluang atau kesempatan yang sama untuk menjadi sampel sedangkan pada *nonrandom sample* anggota populasi tidak mempunyai kesempatan yang sama untuk menjadi sampel. Salah satu cara memilih sampel acek yang paling mudah adalah dengan membuat undian atau lotre. Contohnya, misalkan seorang pemimpin perusahan akan memilih 10 orang dari 100 pegawai untuk ditugaskan sementara ke luar kota, pemimpin perusahaan tersebut harus menuliskan masing-masing nama-nama 100 pegawai pada kertas yang berbeda kemudian menggulungnya satu per satu dan mengambil secara acak 10 kertas, nama yang keluar pada 10 kertas itulah yang merupakan pegawai yang terpilih untuk ditugaskan sementara ke luar kota.

D. Variabel

Variabel adalah segala sesuatu yang berupa apa saja yang ditetapkan oleh seorang peneliti untuk dipelajari sehingga mendapatkan informasi tentang hal tersebut dan kemudian menyimpulkannya. Variabel juga didefinisikan sebagai atribut atau karakteristik dari "orang" maupun "objek" yang mempunyai variasi antara satu dengan yang lainnya. Tinggi badan, berat badan, usia, dan motivasi belajar merupakan sebagian contoh variabel yang merupakan atribut-atribut dari orang. Warna, ukuran, volume, luas, dan bentuk merupakan sebagian contoh variabel yang merupakan atribut-atribut dari objek atau benda. Bahan baku produksi, metode pemasaran, periklanan, keuntungan, dan teknologi produksi merupakan sebagian contoh variabel dalam bidang ekonomi bisnis.

Variabel dapat diklasifikasikan menjadi 2 kelompok yaitu, variabel kuantitatif dan variabel kualitatif. **Variabel kuantitatif** adalah variabel yang dapat diukur secara numerik sedangkan data dari variabel kuantitatif disebut **data kuantitatif**. Variabel kuantitatif diklasifikasikan menjadi 2 yaitu:

1. Variabel Diskrit

   Variabel diskrit adalah variabel kuantitatif yang nilainya dapat dihitung atau dengan kata lain hanya mempunyai nilai pasti tanpa ada "nilai antara". Contohnya, banyaknya mobil yang terjual pada suatu *showroom* adalah 0 (tidak ada mobil yang terjual), 1 (ada 1 mobil yang terjual), 2 (ada 2 mobil yang terjual) dan seterusnya. Banyaknya mobil yang terjual adalah $0,1,2,3,\cdots$ dan tidak mungkin banyaknya mobil yang terjual adalah antara 0 dan 1 atau antara 1 dan 2.

2. Variabel Kontinu

   Variabel kontinu adalah variabel kuantitatif yang nilainya tidak dapat dihitung. Contoh variabel kontinu antara lain total aset perusahaan asuransi, berat barang, hasil panen jagung pada setiap hektar lahan serta setiap variabel yang melibatkan uang misalnya total pendapatan.

Klasifikasi variabel yang kedua adalah variabel kualitatif. **Variabel kualitatif** atau disebut dengan **variabel kategorik** adalah variabel yang tidak dapat diukur secara numerik tetapi dapat diklasifikan dalam dua atau lebih kategori yang berbeda. Data dari variabel kualitatif disebut **data kualitatif**. Misalnya jenis pekerjaan dapat dikelompokkan dalam 3 bidang yaitu, bidang pendidikan, perdagangan dan pertanian. Contoh lain dari variabel kualitatif antara lain jenis kelamin seseorang, warna rambut dan tingkat pendidikan.

Berdasarkan hubungan antar variabel satu dengan variabel lain, variabel dapat diklasifikasikan menjadi 5 yaitu,

1. **Variabel Independen** atau variabel bebas adalah variabel yang mempengaruhi atau menjadi sebab timbulnya variabel lain.
2. **Variabel Dependen** atau variabel terikat adalah variabel yang dipengaruhi atau menjadi akibat adanya variabel bebas.
3. **Variabel Moderator** adalah variabel yang mempengaruhi (memperkuat atau memperlemah) hubungan antara variabel independen dengan variabel dependen.
4. **Variabel Intervening** adalah variabel yang secara teoritis mempengaruhi hubungan antara variabel independen dengan variabel dependen tetapi pengaruhnya tidak dapat diamatai atau diukur.
5. **Variabel Kontrol** adalah variabel yang dikendalikan atau dianggap konstan sehingga hubungan antara variabel independen dengan variabel dependen tidak dipengaruhi oleh faktor lain yang tidak masuk dalam fokus penelitian.

E. Skala Pengukuran

Data dapat diklasifikasikan berdasarkan skala pengukuran yang terdiri dari 4 macam skala, yaitu:

1. **Skala Nominal** adalah skala yang membedakan benda atau peristiwa berdasarkan sifatnya. Contohnya antara lain jenis kelamin (laki-laki dan perempuan), warna (merah, kuning, hijau, biru), pilihan ya atau tidak. Angka yang diberikan untuk skala nominal hanya digunakan sebagai label saja misalnya dalam kasus jenis kelamin 1 untuk laki-laki dan 0 untuk perempuan.

2. **Skala Ordinal** adalah skala yang bersifat membedakan berdasarkan peringkat atau tingkatannya. Contohnya tingkat pendidikan dibedakan menjadi SD/MI, SMP/MTs, SMA/MA dan Sarjana. Angka yang diberikan untuk skala ini menunjukkan peringkat dari objek, misalnya 1 untuk SD/MI, 2 untuk SMP/MTs, 3 untuk SMA/MA dan 4 untuk Sarjana.

3. **Skala Interval** adalah skala pada suatu objek yang dapat diurutkan dan dapat dibedakan antara yang satu dengan yang lain serta perbedaan antara dua nilai dapat diperhitungkan dan diinterpretasikan. Skala ini memiliki nilai nol tetapi bukan nol yang mutlak. Contohnya adalah pengukuran suhu ruangan kelas menggunakan derajat Celcius, misal diperoleh nilai $0^0, 21^0, 26^0$. Nilai 0^0 bukan berarti sebuah ruangan tidak mempunyai suhu akan tetapi ruangan tersebut memiliki suhu 0^0.

4. **Skala Rasio** adalah skala yang diperoleh dari proses pengukuran. Skala ini memiliki nilai nol mutlak dan

nilai yang diperoleh merupakan nilai sebenarnya dari objek yang diukur. Nilai pada skala ini dapat dilakukan operasi aritmatika seperti penjumlahan, pengurangan, perkalian dan pembagian. Contohnya antara lain data berat badan, dan tinggi badan.

BAB 2

DATA DAN PENYAJIANNYA

A. Penyajian Data Kualitatif

Data kualitatif dapat disajikan dalam bentuk tabel biasa atau dalam bentuk berikut:

1. Tabel Distribusi Frekuensi

   Tabel distribusi frekuensi untuk data kualitatif terdiri dari semua kategori dan banyaknya anggota pada setiap kategori.

   Contoh, diambil 20 sampel siswa SMA "XYZ" Kediri yang berencana melanjutkan pendidikan ke jenjang perguruan tinggi. Setiap siswa diminta untuk memilih rencana jurusan/program studi yang akan diambilnya dengan diberikan pilihan sebagai berikut: 1) Matematika, 2) Ekonomi, 3) Manajemen 4) Biologi, dan 5) lainnya.

   Pilihan dari setiap siswa adalah sebagai berikut:

Matematika	*Manajemen*	*Lainnya*	*Manajemen*
Lainnya	*Manajemen*	*Biologi*	*Matematika*
Biologi	*Ekonomi*	*Matematika*	*Biologi*
Matematika	*Matematika*	*Ekonomi*	*Matematika*
Ekonomi	*Biologi*	*Manajemen*	*Ekonomi*

Penyelesaian:

Variabel pilihan jurusan/program studi dapat diklasifikasikan ke dalam 5 kategori yaitu, 1) Matematika, 2) Ekonomi, 3) Manajemen 4) Biologi, dan 5) lainnya. Kemudian dihitung banyaknya siswa yang memilih setiap kategori dan selanjutnya dibuat tabel distribusi frekuensinya.

Tabel 2.1 Rencana Pilihan Jurusan/Program Studi Siswa SMA "XYZ" Kediri

Jurusan/Program Studi	Tally	Frekuensi
Matematika	~~||||~~ |	6
Ekonomi	||||	4
Manajemen	||||	4
Biologi	||||	4
Lainnya	||	2

2. Tabel Distribusi Frekuensi Relatif dan Persentase

Frekuensi relatif dari setiap kategori diperoleh dengan membagi frekuensi setiap kategori dengan jumlah semua frekuensinya. Sedangkan **persentase** diperoleh dengan mengalikan frekuensi relatif dengan 100.

$$\text{frekuensi relatif} = \frac{\text{frekuensi kategori}}{\text{jumlah semua frekuensi}}$$

$$\text{persentase} = \text{frekuensi relatif} \times 100$$

Contoh: (berdasarkan Tabel 2.1)

Tabel 2.2 Rencana Pilihan Jurusan/Program Studi Siswa SMA "XYZ" Kediri

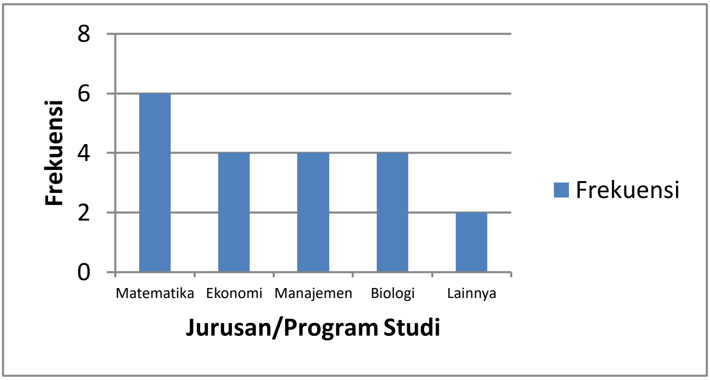
Jurusan/Program Studi	Frekuensi Relatif	Persentase
Matematika	6/20 = 0,3	0,3 x 100 = 30
Ekonomi	4/20 = 0,2	0,2 x 100 = 20
Manajemen	4/20 = 0,2	0,2 x 100 = 20
Biologi	4/20 = 0,2	0,2 x 100 = 20
Lainnya	2/20 = 0,1	0,1 x 100 = 10
Jumlah	1,00	100

3. Grafik/Diagram

a. Diagram Batang

Diagram batang adalah diagram yang menggambarkan data dengan menggunakan sumbu vertikal untuk menunjukkan frekuensi dan sumbu horizontal untuk menunjukkan kategori.

Contoh: (berdasarkan **Tabel 2.1**)



Gambar 2.1 Diagram Batang Rencana Pilihan Jurusan/Program Studi Siswa SMA "XYZ" Kediri

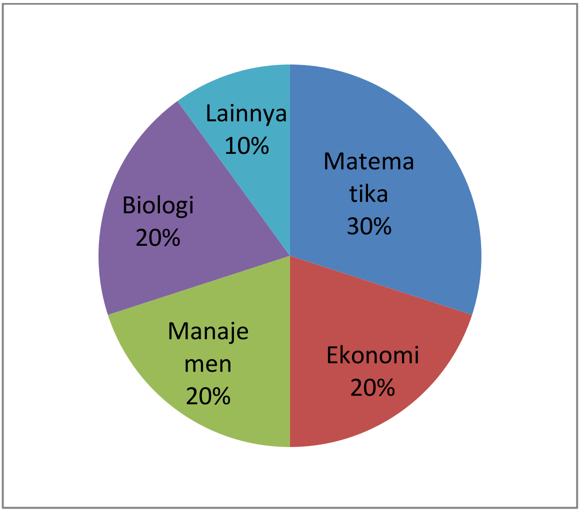
b. Diagram Lingkaran

Diagram lingkaran adalah suatu lingkaran yang dibagi berdasarkan persentase frekuensi dari setiap kategori distribusi.

Contoh: (berdasarkan **Tabel 2.2**)

Tabel 2.3 Cara Menghitung Ukuran Sudut pada Diagram Lingkaran

Jurusan/Program Studi	Frekuensi Relatif	Persentase
Matematika	0,3	0,3 x 360 = 108
Ekonomi	0,2	0,2 x 360 = 72
Manajemen	0,2	0,2 x 360 = 72
Biologi	0,2	0,2 x 360 = 72
Lainnya	0,1	0,1 x 360 = 36
Jumlah	1,00	360



Gambar 2.2 Diagram Lingkaran Rencana Pilihan Jurusan/Program Studi Siswa SMA "XYZ" Kediri

B. Penyajian Data Kuantitatif

1. Tabel Distribusi Frekuensi

Tabel distribusi frekuensi untuk data kuantitatif terdiri dari semua kelas dan banyaknya nilai atau frekensi pada setiap kelas tersebut. Data yang disajikan dalam bentuk tabel distribusi frekuensi ini disebut **data berkelompok**.

Contoh:

Tabel 2.4 Nilai Statistika Mahasiswa Ekonomi Syari'ah

Nilai	Batas Kelas	Panjang Kelas	Nilai Tengah Kelas	Frekuensi
65 – 68	64,5 – 68,5	4	66,5	1
69 – 72	68,5 – 72,5	4	70,5	4
73 – 76	72,5 – 76,5	4	74,5	13
77 – 80	76,5 – 80,5	4	78,5	15
81 – 84	80,5 – 84,5	4	82,5	4
85 – 88	84,5 – 88,5	4	86,5	3

Cara Membuat Tabel Distribusi Frekuensi

Langkah-langkah menyajikan data dalam bentuk tabel distribusi frekuensi adalah sebagai berikut:

a. Mengurutkan data dari yang terkecil ke yang terbesar untuk memudahkan dalam perhitungan frekuensi setiap kelas atau interval

b. Menghitung **banyaknya kelas** dengan menggunakan rumus Sturges $K = 1 + 3{,}3 \log n$

   dengan:

   K : banyaknya kelas interval

   n : banyaknya data

   log : logaritma

c. Menghitung **rentang data** yaitu data terbesar dikurangi data terkecil atau $X_{max} - X_{min}$

d. Menghitung **panjang kelas** yaitu rentang dibagi jumlah kelas atau $\text{panjang kelas} = \dfrac{\text{rentang}}{\text{jumlah kelas}}$

e. Menyusun interval kelas dimulai dari data yang paling kecil
f. Masukkan data pada setiap intervalnya dan hitung frekuensi masing-masing kelas
g. Menyuusun kelas interval dan frekuensi dalam tabel distribusi frekuensi

Contoh:

Berikut ini merupakan nilai UAS matakuliah Statistika I sebanyak 100 mahasiswa:

62	71	31	44	36	79	61	70	91	41
94	32	30	90	33	32	79	45	34	48
59	62	62	48	53	30	71	51	77	33
50	44	47	78	58	93	65	39	85	31
40	77	53	70	45	84	81	40	62	37
41	60	45	71	41	67	65	51	87	46
56	80	87	43	79	79	51	67	32	59
41	59	59	69	53	39	92	71	36	63
89	50	67	40	80	83	63	66	86	72
73	58	41	69	93	64	37	33	80	57

Data yang telah diurutkan:

30	37	42	49	59	63	71	75	82	90
31	38	43	51	59	63	72	75	82	90
31	38	43	51	60	66	72	75	83	90
31	38	45	52	60	69	72	75	83	91

32	38	45	53	60	69	72	76	85	92
32	39	46	54	60	70	73	76	85	93
33	40	46	55	62	70	73	78	86	94
33	40	47	57	62	70	73	78	86	94
35	40	48	58	62	70	74	81	87	94
36	40	48	58	63	71	75	81	88	94

Banyak Kelas:

$$
\begin{aligned}
K &= 1 + 3{,}3 \log n \\
&= 1 + 3{,}3 \log(100 \\
&= 1 + 3{,}3(2) \\
&= 1 + 6{,}6 \\
&= 7{,}6 \approx 8
\end{aligned}
$$

Rentang Data = 94 – 30 = 64

Panjang Kelas = 64 / 8 = 8

Tabel 2.5 Penyusunan Tabel Distribusi Frekuensi

Nilai	Tally	Frekuensi
30 – 38	|||/ |/|| /|||	15
39 – 46	||/ |/|| ||	12
47 – 55	|/| /||	10
56 – 63	|||/ |/|| /|||	15
64 – 71	|/| ||||	9
72 – 79	|||/ |/|| /||| ||	17
80 – 87	|||| |||| |	11
88 – 96	|||| |||| |	11

2. Tabel Distribusi Frekuensi Relatif dan Persentase

Frekuensi relatif dari setiap kelas diperoleh dengan membagi frekuensi setiap kelas dengan jumlah semua frekuensinya. Sedangkan **persentase** diperoleh dengan mengalikan frekuensi relatif dengan 100.

$$\text{frekuensi relatif kelas} = \frac{\text{frekuensi kelas}}{\text{jumlah semua frekuensi}}$$

$$\text{persentase} = \text{frekuensi relatif } \times 100$$

Contoh: (berdasarkan Tabel 2.5)

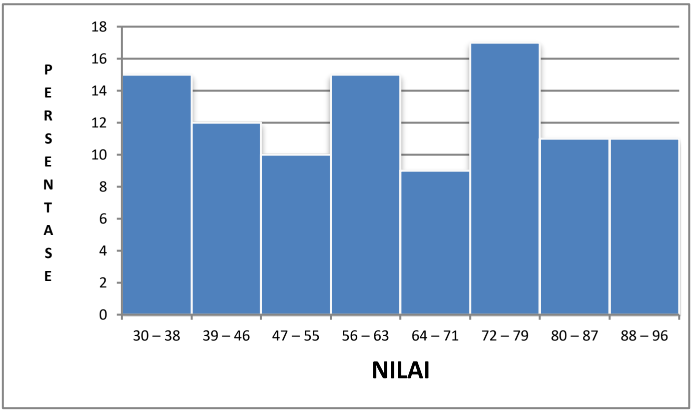
Tabel 2.6 Tabel Distribusi Frekuensi Relatif dan Persentase untuk Tabel 2.5

Nilai	Batas Kelas	Frekuensi Relatif	Persentase
30 – 38	29,5 – 38,5	15 / 100 = 0,15	15
39 – 46	38,5 – 46,5	12 / 100 = 0,12	12
47 – 55	46,5 – 55,5	10 / 100 = 0,1	10
56 – 63	55,5 – 63,5	15 / 100 = 0,15	15
64 – 71	63,5 – 71,5	9 / 100 = 0,09	9
72 – 79	71,5 – 79,5	17 / 100 = 0,17	17
80 – 87	79,5 – 87,5	11 / 100 = 0,11	11
88 – 96	87,5 – 96,5	11 / 100 = 0,11	11
Jumlah		1,00	100

3. Grafik

a. Histogram

Histogram adalah grafik dengan kelas sebagai sumbu horizontal dan frekuensi/frekuensi relatif/persentase sebagai sumbu vertikal.

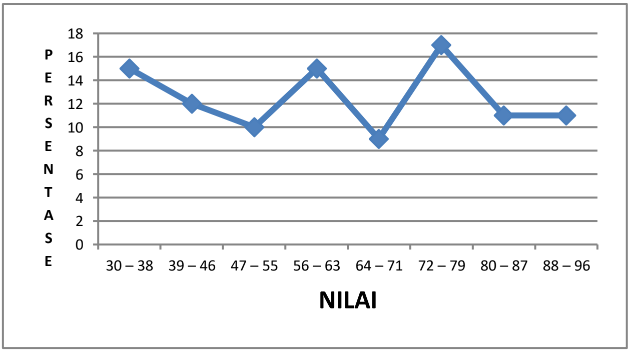


Gambar 2.3 Histogram dari Tabel 2.6

b. Poligon

Poligon adalah grafik yang dibentuk dengan menghubungkan titik tengah histogram bagian atas dengan garis lurus.

Contoh:



Gambar 2.4 Poligon dari Tabel 2.6

BAB 3

PENGUKURAN DESKRIPTIF

A. Ukuran Pemusatan untuk Data Tidak Berkelompok

1. Mean

Mean untuk data sampel dinotasikan dengan $\bar{x}$ (x bar) dan untuk data populasi dinotasikan dengan μ (mu). Mean pada data tidak berkelompok diperoleh dengan cara membagi jumlah semua nilai dengan banyaknya jumlah nilai pada data, sehingga:

Mean untuk data populasi:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Mean untuk data sampel:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Contoh:

Skor IQ dari 10 mahasiswa Fakultas Ekonomi adalah 101, 105, 120, 113, 112, 115, 110, 117, 101, 103.

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10}$$

$$= \frac{101+105+120+113+112+115+110+117+101+103}{10}$$

$$= 109{,}7$$

2. Median

Median adalah nilai tengah data setelah diurutkan dari yang terkecil sampai yang terbesar. Letak median adalah pada data ke- $\frac{n+1}{2}$

Contoh:

Skor IQ dari 10 mahasiswa Fakultas Ekonomi yang telah diurutkan adalah 101, 101, 103, 105, 110, 112, 113, 115, 117, 120.

Median data di atas terletak pada data ke- $\frac{10+1}{2} = 5{,}5$

Nilai median data di atas adalah $\frac{110+112}{2} = 111$

3. Modus

Modus adalah nilai yang sering muncul atau nilai yang memiliki frekuensi paling banyak. Data yang memiliki satu modus disebut ***unimodal*** sedangkan yang memiliki dua modus disebut ***bimodal*** dan yang memiliki lebih dari dua modus disebut ***multimodal***.

Contoh:

Skor IQ dari 10 mahasiswa Fakultas Ekonomi yang telah diurutkan adalah 101, 101, 103, 105, 110, 112, 113, 115, 117, 120.

Modus untuk data di atas adalah 101.

B. Ukuran Dispersi untuk Data Tidak Berkelompok

1. Range

Range adalah ukuran dispersi yang paling mudah untuk dihitung karena hanya memperhatikan nilai yang paling besar dan paling kecil.

$$Range = Data\ terbesar - Data\ terkecil$$

Contoh:

Banyaknya lampu yang diproduksi oleh perusahaan "Terang Benderang" dalam tujuh hari terakhir adalah sebanyak 240, 210, 310, 241, 220, 230, 200.

Range dari data tersebut adalah $Range = 310 - 200 = 110$

2. Varian dan Standar Deviasi

Standar deviasi adalah ukuran dispersi yang paling sering digunakan. Nilai standar deviasi menjelaskan seberapa dekat nilai data dengan mean. Standar deviasi adalah nilai akar dari varian.

Varian untuk populasi:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Varian untuk sampel:

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

Contoh:

Banyaknya lampu yang diproduksi oleh perusahaan “Terang Benderang” dalam tujuh hari terakhir adalah sebanyak 240, 210, 310, 241, 220, 230, 200.

$$\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{7} x_i}{7} \\
&= \frac{240 + 210 + 310 + 241 + 220 + 230 + 200}{7} \\
&= 235{,}86
\end{aligned}$$

Sehingga varian untuk sampel tersebut adalah:

$$\begin{aligned}
s^2 &= \frac{\sum_{i=1}^{7} (x_i - \bar{x})^2}{7-1} \\
&= \frac{(240 - 235{,}86)^2 + (210 - 235{,}86)^2 + \cdots + (200 - 235{,}86)^2}{7-1} \\
&= \frac{7780{,}86}{6} \\
&= 1296.77
\end{aligned}$$

3. Koefisien Variasi

Koefisien variasi dinotasikan dengan CV yang menunjukkan standar deviasi sebagai persentase dari mean. Koefisien variasi bukanlah suatu unit pengukuran melainkan suatu persentase.

Koefisien variasi untuk data populasi:

$$CV = \frac{\sigma}{\mu} \times 100\%$$

Koefisien variasi untuk data sampel:

$$CV = \frac{s}{\bar{x}} \times 100\%$$

Contoh:

Banyaknya lampu yang diproduksi oleh perusahaan "Terang Benderang" dalam tujuh hari terakhir adalah sebanyak 240, 210, 310, 241, 220, 230, 200 dengan $\bar{x} = 235{,}86$ dan $s = \sqrt{s^2} = \sqrt{1296.,7} = 36{,}01$.

$$\begin{aligned} CV &= \frac{s}{\bar{x}} \times 100\% \\ &= \frac{36{,}01}{235{,}86} \times 100\% \\ &= 15{,}27\% \end{aligned}$$

4. Parameter Populasi dan Statistik Sampel

Parameter populasi atau parameter merupakan ukuran secara numerik seperti mean, median, modus, range, varian atau standar deviasi yang dihitung dari data populasi. Sedangkan **statistik sampel** merupakan ukuran numerik yang dihitung dari data sampel. μ dan σ adalah parameter populasi sedangkan $\bar{x}$ dan s adalah statistik sampel.

C. Mean, Varian dan Standar Deviasi untuk Data Berkelompok

Data yang berbentuk tabel frekuensi atau disebut dengan data berkelompok mempunyai formula tersendiri dalam perhitungan mean, varian dan standar deviasi seperti penjelasan berikut:

1. Mean

Mean untuk data populasi:

$$\mu = \frac{\sum_{i=1}^{N} x_i f_i}{N}$$

Mean untuk data sampel:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{n}$$

dimana x_i adalah nilai tengah kelas ke-i dan

f_i frekuensi kelas ke-i..

Langkah pertama untuk menghitung mean pada data berkelompok adalah dengan mencari nilai tengah pada setiap kelas kemudian mengalikannya dengan frekuensi pada setiap kelas yang bersesuaian. Langkah selanjutnya adalah menjumlahkan setiap hasil perkalian pada langkah pertama dan kemudian membaginya dengan banyaknya data.

Contoh:

Rasio perbandingan harga saham dengan laba bersih perusahaan (PER) dari 20 perusahaan adalah sebagai berikut:

Tabel 3.1 Price-Earnings Ratio (PER)

PER	Frekuensi	Nilai Tengah
10 – 15	4	12,5
16 – 21	8	18,5
22 – 27	3	24,5
28 – 33	5	30,5

Rata-rata dari data di atas adalah:

$$\bar{x} = \frac{(12{,}5x4) + (18{,}5x8) + (24{,}5x3) + (30{,}5x5)}{20}$$

$$= 21{,}2$$

2. Varian dan Standar Deviasi

Varian untuk populasi:

$$\sigma^2 = \frac{\sum_{i=1}^{N} f_i (x_i - \mu)^2}{N}$$

Varian untuk sampel:

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

dimana x_i adalah nilai tengah kelas ke - i dan f_i frekuensi kelas ke - i.

Contoh:

Varian untuk data pada Tabel 2.7 adalah:

$$s^2 = \frac{(12{,}5 - 21{,}2)^2 + (18{,}5 - 21{,}2)^2 + (24{,}5 - 21{,}2)^2 + (30{,}5 - 21{,}2)^2}{20 - 1}$$

$$= 9{,}49$$

D. Ukuran Letak

1. Kuartil

Suatu nilai yang membagi data yang sudah terurut dari terkecil sampai terbesar ke dalam empat bagian yang sama. Ada tiga macam kuartil, kuartil pertama ialah nilai dalam distribusi yang membatasi 25% frekuensi di bagian atas dan 75% frekuensi di bagian bawah distribusi, kuartil kedua ialah nilai dalam distribusi yang membatasi 50% frekuensi di bagian atas dan 50% di bawahnya, sedangkan kuartil tiga ialah nilai dalam distribusi yang membatasi 75% frekuensi di bagian atas dan 25% frekuensi di bagian bawah distribusi.

Letak kuartil adalah pada data ke:

$$K_1 = {}^1\!/_4(n+1)$$

$$K_2 = {}^1\!/_2(n+1)$$

$$K_3 = {}^3\!/_4(n+1)$$

Nilai kuartil diperoleh dengan perhitungan berikut:

$$K_1 = Bb + p\,\frac{\left({}^1\!/_4\,n - F\right)}{f}$$

$$K_2 = Bb + p\,\frac{\left({}^1\!/_2\,n - F\right)}{f}$$

$$K_3 = Bb + p\,\frac{\left({}^3\!/_4\,n - F\right)}{f}$$

dengan:

K_i : nilai kuartil

Bb : batas bawah kelas kuartil

p : panjang kelas kuartil

n : banyaknya data

f : frekuensi kelas kuartil

F : frekuensi kumulatif di atas kelas kuartil

Contoh (data dalam tabel distribusi frekuensi):

Tabel 3.2 Data Ketersediaan Pipa di Toko Bangunan "XX" Tulungagung

Diameter Pipa (mm)	Frekuensi	Frekuensi Kumulatif
65 – 67	2	2
68 – 70	5	7
71 – 73	13	20
74 – 76	14	34
77 – 79	4	38
80 – 82	2	40

Penyelesaian:

Letak Kuartil:

$$K_1 = \tfrac{1}{4}(40) = 10, \text{Bb kelas interval } 71 - 0{,}5 = 70{,}5$$

$$K_2 = \tfrac{1}{2}(40) = 20, \text{Bb kelas interval } 71 - 0{,}5 = 70{,}5$$

$$K_3 = \tfrac{3}{4}(40) = 30, \text{Bb kelas interval } 74 - 0{,}5 = 73{,}5$$

Nilai Kuartil:

$$K_1 = 70{,}5 + 3\frac{\left(\tfrac{1}{4}\bullet 40 - 7\right)}{13}$$
$$= 70{,}5 + 0{,}69$$
$$= 71{,}19$$

$$K_2 = 70{,}5 + 3\frac{\left(\tfrac{1}{2}\bullet 40 - 7\right)}{13}$$
$$= 70{,}5 + 3$$
$$= 73{,}5$$

$$K_3 = 73{,}5 + 3\frac{\left(\tfrac{3}{4}\bullet 40 - 7\right)}{13}$$
$$= 73{,}5 + 2{,}14$$
$$= 75{,}64$$

2. Desil

Suatu nilai yang membagi data yang sudah terurut dari terkecil sampai terbesar ke dalam sepuluh bagian yang sama.

Letak desil adalah pada data ke:

$$D_1 = \tfrac{1}{10}(n+1) \qquad D_6 = \tfrac{6}{10}(n+1)$$
$$D_2 = \tfrac{2}{10}(n+1) \qquad D_7 = \tfrac{7}{10}(n+1)$$
$$D_3 = \tfrac{3}{10}(n+1) \qquad D_8 = \tfrac{8}{10}(n+1)$$
$$D_4 = \tfrac{4}{10}(n+1) \qquad D_9 = \tfrac{9}{10}(n+1)$$
$$D_5 = \tfrac{5}{10}(n+1)$$

Nilai desil diperoleh dengan perhitungan berikut:

$$D_i = Bb + p\,\frac{\left(\tfrac{i}{10}n - F\right)}{f}$$

dengan:

D_i : nilai desil ke - i

Bb : batas bawah kelas desil

p : panjang kelas desil

n : banyaknya data

f : frekuensi kelas desil

F : frekuensi kumulatif di atas kelas desil

Contoh (data sama dengan pada contoh Kuartil):

Penyelesaian:

Letak Desil:

$D_1 = \tfrac{1}{10}(40) = 4$, Bb kelas interval $68 - 0{,}5 = 67{,}5$

$D_9 = \tfrac{9}{10}(40) = 36$, Bb kelas interval $77 - 0{,}5 = 76{,}5$

Nilai Desil:

$$\begin{aligned} D_1 &= 67{,}5 + 3\,\frac{\left(\tfrac{1}{10} \bullet 40 - 2\right)}{5} \\ &= 67{,}5 + 1{,}2 \\ &= 68{,}7 \end{aligned}$$

$$D_9 = 76{,}5 + 3\frac{\left(\frac{9}{10} \bullet 40 - 34\right)}{4}$$

$$= 76{,}5 + 1{,}5$$

$$= 78$$

3. Persentil

Suatu nilai yang membagi data yang sudah terurut dari terkecil sampai terbesar ke dalam seratus bagian yang sama.

Letak persentil adalah pada data ke:

$$P_i = \frac{i}{100}(n+1), i = 1,2,3,\cdots,99$$

Nilai persentil diperoleh dengan perhitungan berikut:

$$P_i = Bb + p\frac{\left(\frac{i}{100}n - F\right)}{f}$$

dengan:

P_i : nilai persentil ke - i

Bb : batas bawah kelas persentil

p : panjang kelas persentil

n : banyaknya data

f : frekuensi kelas persentil

F : frekuensi kumulatif di atas kelas persentil

Contoh (data sama dengan pada contoh Kuartil):

Penyelesaian:

Letak Persentil:

$$P_{88} = {}^{88}\!/_{100}(40) = 35{,}2, \text{ Bb kelas interval } 77 - 0{,}5 = 76{,}5$$

Nilai Persentil:

$$P_{88} = 76{,}5 + 3\frac{\left({}^{88}\!/_{100} \bullet 40 - 34\right)}{4}$$

$$= 76{,}5 + 0{,}9$$

$$= 77{,}4$$

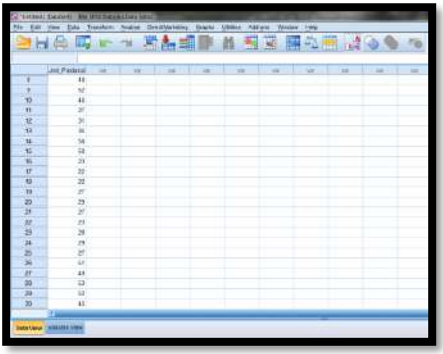
E. Tutorial SPSS

Berikut adalah data mengenai jumlah barang yang diproduksi oleh suatu *home industry* selama bulan Nopember 2020:

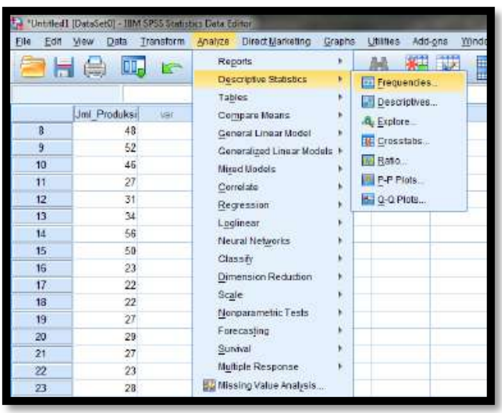
Tanggal	Jumlah	Tanggal	Jumlah	Tanggal	Jumlah
1	24	11	27	21	27
2	21	12	31	22	23
3	31	13	34	23	28
4	32	14	56	24	29
5	26	15	50	25	27
6	35	16	23	26	51
7	45	17	22	27	49
8	48	18	22	28	53
9	52	19	27	29	52
10	46	20	29	30	45

Langkah-langkah:

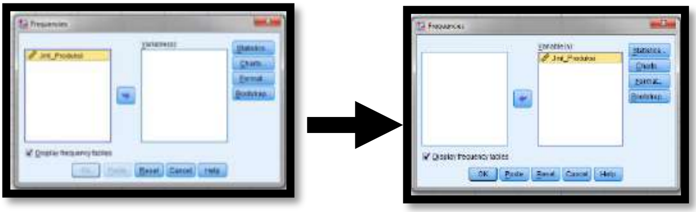
1. Inputkan data jumlah produksi ke SPSS



2. Klik **Analyze →Descriptive Statistics →Frequencies**



3. Pindahkan variabel jml_produksi ke kolom ***variables*** dengan klik tanda panah kemudian klik ***statistics***



4. Setelah muncul kotak dialog berikut, centang sesuai kebutuhan

5. Output

Statistics		
Jml_Produksi		
N	Valid	30
	Missing	0
Mean		35.50
Median		31.00
Mode		27
Std. Deviation		11.664
Variance		136.052
Range		35
Percentiles	25	26.75
	30	27.00
	45	29.00
	50	31.00
	70	45.70
	75	48.25

Jml_Produksi					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	21	1	3.3	3.3	3.3
	22	2	6.7	6.7	10.0
	23	2	6.7	6.7	16.7
	24	1	3.3	3.3	20.0
	26	1	3.3	3.3	23.3
	27	4	13.3	13.3	36.7
	28	1	3.3	3.3	40.0
	29	2	6.7	6.7	46.7
	31	2	6.7	6.7	53.3
	32	1	3.3	3.3	56.7
	34	1	3.3	3.3	60.0
	35	1	3.3	3.3	63.3
	45	2	6.7	6.7	70.0
	46	1	3.3	3.3	73.3
	48	1	3.3	3.3	76.7
	49	1	3.3	3.3	80.0
	50	1	3.3	3.3	83.3
	51	1	3.3	3.3	86.7
	52	2	6.7	6.7	93.3
	53	1	3.3	3.3	96.7
	56	1	3.3	3.3	100.0
	Total	30	100.0	100.0	

BAB 4

ANALISIS KORELASI DAN REGRESI LINIER SEDERHANA

A. Analisis Korelasi

Analisis korelasi adalah analisis statistika yang digunakan untuk melihat ada atau tidaknya hubungan antar variabel. Nilai dari korelasi berkisar antara -1 dan 1 atau secara matematis dapat ditulis $-1 \le r \le 1$.

r = -1 artinya terdapat hubungan negatif sempurna antara dua variabel

r = 0 artinya tidak terdapat hubungan antara dua variabel

r = 1 artinya terdapat hubungan positif sempurna antara dua variabel

Perlu diperhatikan bahwa apabila nilai r = 0, tidak selalu antara dua variabel tersebut tidak berhubungan karena ada kemungkinan hubungan tersebut nonlinier. Pembagian dari analisis korelasi adalah sebagai berikut:

1. Korelasi Sederhana (Pearson Product Moment)

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left\{n\sum X^2 - \left(\sum X\right)^2\right\}\left\{n\sum Y^2 - \left(\sum Y\right)^2\right\}}}$$

Pengujian signifikansi korelasi dapat dilakukan dengan menggunakan t-test dengan hipotesis sebagai berikut:

H_0 : r = 0 (tidak ada hubungan yang signifikan antara variabel X dan Y

H_1 : $r \neq 0$ (ada hubungan yang signifikan antara variabel X dan Y

Statistik uji:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Dengan pengambilan keputusan tolak H_0 jika $t \geq t_{\frac{\alpha}{2};n-2}$.

Contoh:

Tabel 4.1 Tabel Penolong Perhitungan Korelasi Sederhana

No	Matematika (X)	Statistika (Y)	X^2	Y^2	XY
1	90	80	8100	6400	7200
2	60	70	3600	4900	4200
3	50	60	2500	3600	3000
4	70	80	4900	6400	5600
5	40	50	1600	2500	2000

6	30	40	900	1600	1200
7	20	20	400	400	400
8	80	90	6400	8100	7200
9	70	80	4900	6400	5600
10	60	60	3600	3600	3600
Jumlah	570	630	36900	43900	40000

$$r = \frac{10x40000 - 570x630}{\sqrt{\left\{10x36900 - (570)^2\right\}\left\{10x43900 - (630)^2\right\}}} = 0.949$$

$$t = \frac{0.949\sqrt{10-2}}{\sqrt{1-(0.949)^2}} = 8.514$$

Karena $t = 9.697 \geq 1.86 = t_{0.05/2;8}$ maka disimpulkan bahwa terdapat korelasi positif antara nilai mahasiswa di bidang matematika dan statistika.

2. Korelasi Berganda

Korelasi berganda adalah perluasan dari korelasi sederhana dimana variabel yang digunakan berjumlah lebih dari dua. Koefisien korelasi ganda mengukur seberapa besar keragaman variabel X_1, X_2, ..., X_p menjelaskan variabel respon Y. Pengujian signifikansi koefisien korelasi berganda dilakukan dengan menggunakan uji F.

Statistik uji:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)};$$

Dengan n: banyaknya pengamatan dan k: banyaknya variabel bebas. Pengambilan keputusan tolak H_0 jika $F > F_{\alpha;k,(n-k-1)}$.

3. Korelasi Parsial

Korelasi parsial adalah koefisien korelasi yang menyatakan hubungan antara variabel Y dan X_1 dengan variabel X_2 tetap. Koefisien korelasi parsial dihitung dengan formula sebagai berikut:

$$r_{yx_1.x_2} = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{\left(1 - r^2_{yx_2}\right)\left(1 - r^2_{x_1x_2}\right)}}$$

Pengujian koefisien korelasi parsial menggunakan statistik uji t sebagai berikut:

$$t = \frac{r_{yx_1.x_2}\sqrt{n-k-1}}{\sqrt{1 - r^2_{yx_1.x_2}}}$$

Dengan pengambilan keputusan tolak H_0 jika $t \geq t_{\frac{\alpha}{2};n-k-1}$.

4. Korelasi Rank Spearman

Korelasi *Rank Spearman* digunakan untuk menentukan besarnya korelasi antara dua variabel yang datanya berskala ordinal. Koefisien korelasi *Rank Spearman* dihitung dengan formula sebagai berikut:

$$r_{xy} = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

Pengujian signifikansi korelasi dapat dilakukan dengan menggunakan t-test dengan hipotesis sebagai berikut:

H_0 : r = 0 (tidak ada hubungan yang signifikan antara variabel X dan Y

H_1 : r ≠ 0 (ada hubungan yang signifikan antara variabel X dan Y

Statistik uji:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Dengan pengambilan keputusan tolak H_0 jika $t \geq t_{\frac{\alpha}{2};n-2}$.

Contoh:

Tabel 4.2 Tabel Penolong Perhitungan Korelasi *Rank-Spearman*

Matematika (X)	Statistika (Y)	Rank X	Rank Y	d	d^2
90	80	1	3	-2	4
60	70	5.5	5	0.5	0.25
50	60	7	6.5	0.5	0.25
70	80	3.5	3	0.5	0.25
40	50	8	8	0	0
30	40	9	9	0	0

20	20	10	10	0	0
80	90	2	1	1	1
70	80	3.5	3	0.5	0.25
60	60	5.5	6.5	-1	1
Jumlah	[redacted]				7

$$r_{xy} = 1 - \frac{6x7}{10\left(10^2 - 1\right)} = 0.96$$

$$t = \frac{0.96\sqrt{10-2}}{\sqrt{1-(0.96)^2}} = 9.697$$

Karena $t = 9.697 \geq 1.86 = t_{0.05/2;8}$ maka disimpulkan bahwa terdapat korelasi positif antara nilai mahasiswa di bidang matematika dan statistika.

B. Analisis Regresi Linier Sederhana

Analisis regresi adalah suatu analisis statistika yang digunakan untuk melihat hubungan dan pengaruh antara variabel dependen (terikat) dengan variabel independen (bebas). Regresi linier sederhana digunakan untuk melihat hubungan variabel yang digambarkan dengan garis lurus dimana hanya memuat dua variabel saja, satu independen dan satu dependen. Persamaan regresi linier sederhana didefinisikan sebagai berikut:

$$Y = a + bX$$

dimana

a : konstanta

b : koefisien regresi

Y : variabel dependen

X : variabel independen

Nilai a dan b diperoleh dengan menggunakan metode kuadrat terkecil (*Least Square*) sebagai berikut:

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

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

Pengujian hipotesis:

H_0 : tidak ada pengaruh variabel X terhadap Y

H_1 : terdapat pengaruh variabel X terhadap Y

Berikut adalah tabel ANOVA yang digunakan dalam pengujian hipotesis regresi linier berganda dengan dua variabel bebas:

Tabel 4.3 *Analysis of Variance* (ANOVA) Regresi Linier Sederhana

Sumber Variasi	db	JK	RJK	F
Total	n	$\sum Y^2$	JK_T/n	$\dfrac{RJK_{b\|a}}{RJK_S}$
Koefisien (a)	1	$\left(\sum Y\right)^2/n$	JK_a	
Regresi (b|a)	1	$b\left\{\sum XY - \dfrac{\sum X \sum Y}{n}\right\}$	$JK_{b\|a}$	
Sisa	n-2	$JK_T - JK_a - JK_{b\|a}$	$JK_S/(n-2)$	

Contoh:

Tabel 4.4 Tabel Penolong Perhitungan Koefisien Regresi Linier Sederhana

No	X	Y	X^2	Y^2	XY
1	3	9	9	81	27
2	4	7	16	49	28
3	3	9	9	81	27
4	2	8	4	64	16
5	1	9	1	81	9
$\sum$	13	42	39	356	107

$$b = \frac{5x107 - 13x42}{5x39 - 13^2} = -0.423$$

$$a = \frac{42 - (-0.423)x13}{5} = 9.499$$

Sehingga diperoleh persamaan regresi $Y = 9.499 - 0.423X$.

Tabel 4.5 ANOVA Model Regresi

Sumber Variasi	db	JK	RJK	F
Total	5	356	$356/5 = 71.2$	$\frac{0.9306}{0.756} = 1.2309$
Koefisien (a)	1	$(42)^2/5 = 352.8$	352.8	
Regresi (b|a)	1	$(-0.423)\left\{107 - \frac{13x42}{5}\right\} = 0.9306$	0.9306	
Sisa	3	$356 - 352.8 - 0.9306 = 2.2694$	$2.2694/3 = 0.756$	

Dari hasil tabel ANOVA di atas dapat disimpulkan bahwa gagal tolak H_0 karena $F = 1.2309 < 10.13 = F_{0.05;1,3}$ atau dengan kata lain tidak ada pengaruh yang signifikan antara variabel X terhadap Y.

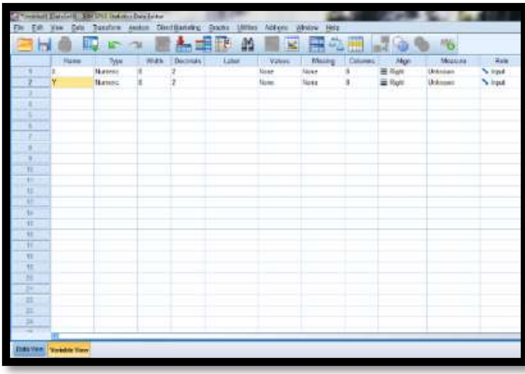
C. Contoh Kasus dan Tutorial SPSS

Suatu *home industry* pembuatan batu bata ingin melihat kualitas ketahanan batu bata yang diproduksi. Dalam suatu pengujian produk batu bata, tekanan normal (X) atas batu bata diketahui berkaitan secara fungsional dengan ketahanan (Y) batu bata tersebut. Berikut ini adalah data percobaan mengenai kedua variabel tersebut setelah disamakan satuannya:

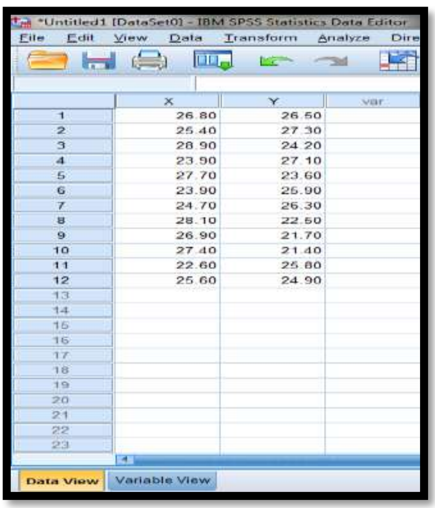
X	Y
26,8	26,5
25,4	27,3
28,9	24,2
23,9	27,1
27,7	23,6
23,9	25,9
24,7	26,3
28,1	22,5
26,9	21,7
27,4	21,4
22,6	25,8
25,6	24,9

1. Analisis Regresi

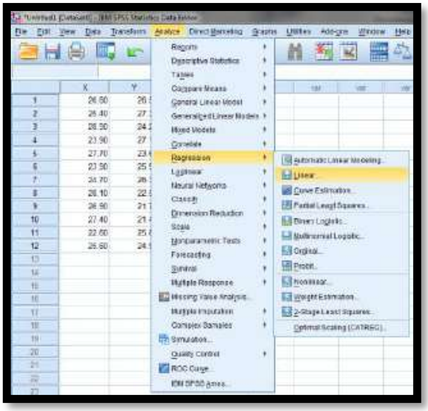
a. Definisikan nama variabel yang digunakan pada "**Variable View**"



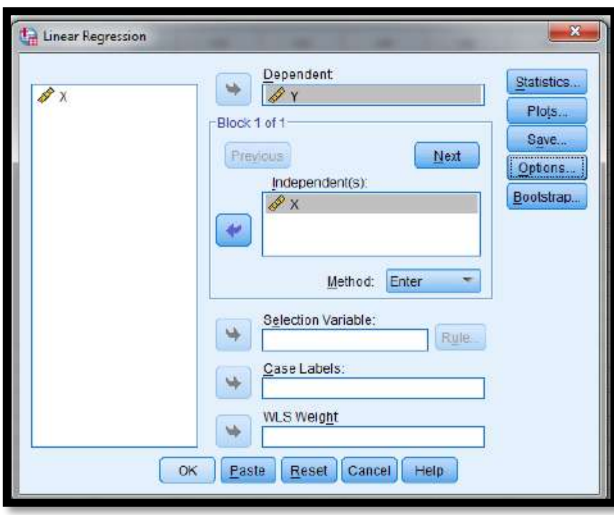
b. Inputkan data pada "**Data View**"



c. Klik **Analyze→Regression→Linear**



d. Masukkan variabel **Y** ke kolom **Dependent** dan **X** ke kolom **Independent** dengan klik panah di sebelah kolom **→OK**



e. Output:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	X[b]	.	Enter

a. Dependent Variable: Y

b. All requested variables entered.

Variables entered → menunjukkan variabel prediktor yang digunakan, yaitu X.

Dependent variable → menunjukkan variabel respon yang digunakan, yaitu Y.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650[a]	.423	.365	1.64968

a. Predictors: (Constant), X

R Square → menunjukkan besarnya pengaruh variabel prediktor (X) terhadap variabel respon (Y). Diperoleh R square sebesar 0.423=42.3%, artinya variabel X berpengaruh terhadap variabel Y sebesar 42.3%.

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19.932	1	19.932	7.324	.022[b]
	Residual	27.214	10	2.721		
	Total	47.147	11			
a. Dependent Variable: Y						
b. Predictors: (Constant), X						

F → F hitung yang nantinya akan dibandingkan dengan F tabel (dengan derajat bebas 1 dan 10)

Sig → alternatif lain selain uji F untuk melihat signifikansi dari regresi yang dihasilkan

Regresi dikatakan signifikan atau terdapat pengaruh yang signifikan antara variabel prediktor terhadap variabel respon apabila nilai sig < taraf signifikansi.

Misal dipakai taraf signifikansi 5% atau 0.05, karena 0.022 < 0.05 maka regresi yang dihasilkan adalah signifikan. Atau secara serentak, variabel X berpengaruh signifikan terhadap Y.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	42.735	6.656		6.420	.000
	X	-.691	.255	-.650	-2.706	.022

a. Dependent Variable: Y

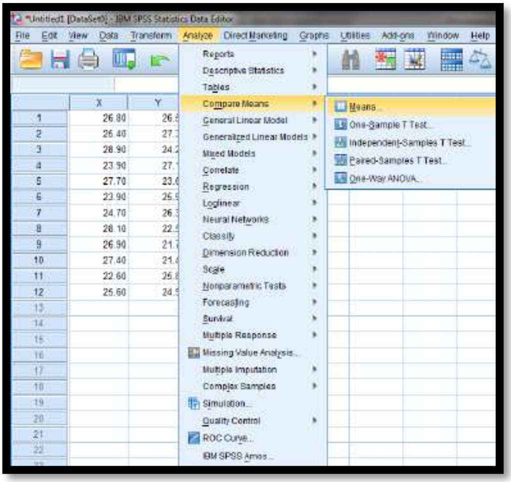
Unstandardized Coefficients → menunjukkan persamaan regresi yang dihasilkan, yaitu

$$Y = 42.735 - 0.691$$

Sig → menunjukkan signifikansi dari pengaruh setiap variabel prediktor (X) terhadap respon (Y). Karena dalam kasus ini hanya ada 1 variabel prediktor maka hanya dianalisis satu saja, yaitu 0.022 < 0.05 artinya X berpengaruh signifikan terhadap Y.

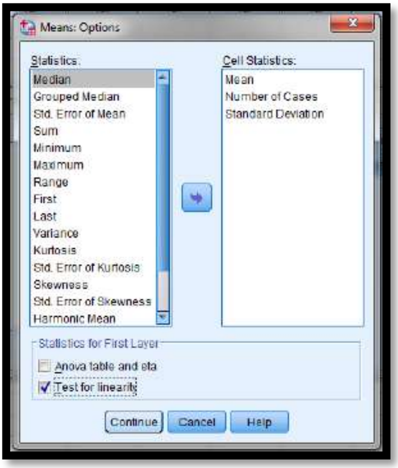
2. Pengujian Asumsi Linieritas Regresi

a. Klik **Analyze → Compare means → means**



b. Masukkan variabel **Y** ke kolom **Dependent** dan **X** ke kolom **Independent** dengan klik panah di sebelah kolom → klik **Options**

c. Pada Options, centang bagian test for linearity → continue → OK



d. Output:

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Y * X	12	100.0%	0	0.0%	12	100.0%

Tabel ini menunjukkan banyaknya data yang dianalisis, yaitu sebanyak 12 pasang data variabel X dan variabel Y.

ANOVA Table							
			Sum of Squares	df	Mean Square	F	Sig.
Y * X	Between Groups	(Combined)	46.427	10	4.643	6.448	.298
		Linearity	19.932	1	19.932	27.684	.120
		Deviation from Linearity	**26.494**	**9**	**2.944**	**4.089**	**.367**
	Within Groups		.720	1	.720		
	Total		47.147	11			

Selain menggunakan uji F seperti pada cara manual (F hitung = 4.089), bisa menggunakan nilai sig untuk melihat hubungan antara variabel prediktor dan respon (linier atau tidak).

Hubungan tersebut dikatakan linier apabila nilai sig > taraf signifikansi.

Diperoleh sig sebesar 0.367 > 0.05 → hubungan antara variabel respon dan prediktor bersifat linier.

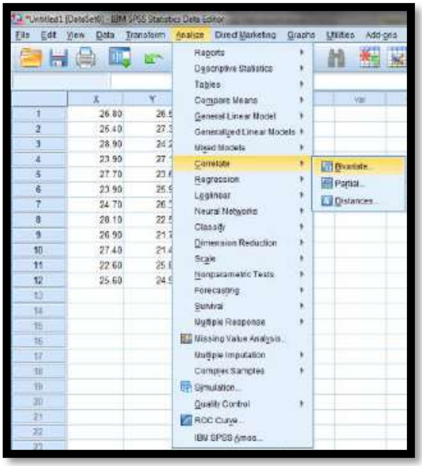
Measures of Association

	R	R Squared	Eta	Eta Squared
Y * X	-.650	.423	.992	.985

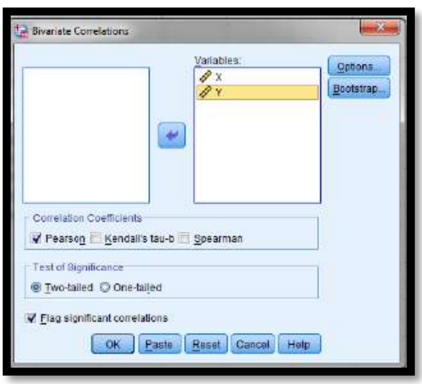
Nilai R menunjukkan nilai korelasi antara variabel X dan Y, yaitu sebesar -0.65

3. Analisis Korelasi Sederhana

a. Klik **Analyze → correlate → bivariate**



b. Masukkan variabel yang akan dicari nilai korelasinya ke kolom **Variables → OK**



c. Output

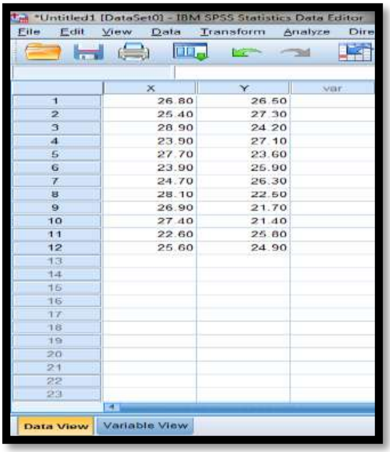
Correlations			
		X	Y
X	Pearson Correlation	1	-.650*
	Sig. (2-tailed)		.022
	N	12	12
Y	Pearson Correlation	-.650*	1
	Sig. (2-tailed)	.022	
	N	12	12

*. Correlation is significant at the 0.05 level (2-tailed).

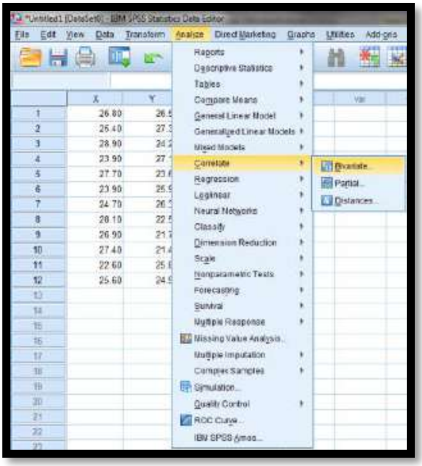
Korelasi antara X dan Y sebesar -0.65. Korelasi tersebut adalah signifikan karena sig sebesar 0.022 < 0.05.

4. Analisis Korelasi Spearman dan Tau Kendall

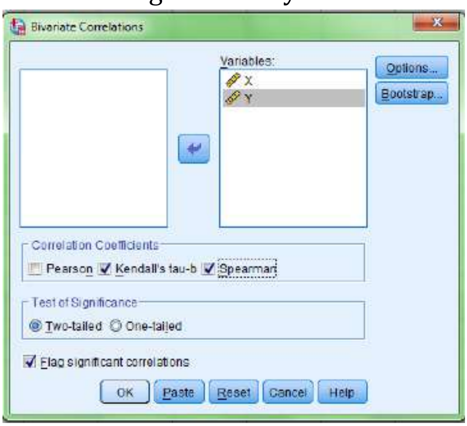
a. Inputkan data pada "**Data View**"



b. Klik **Analyze → Correlate → Bivariate**



c. Masukkan variabel yang akan dicari nilai korelasinya ke kolom **Variables → Kendall's** (*untuk korelasi Kendall*) atau **Spearman** (*untuk korelasi Spearman*) atau centang dua-duanya **→ OK**



d. Output

Correlations				
			X	Y
Kendall's tau_b	X	Correlation Coefficient	1.000	-.382

		Sig. (2-tailed)	.	.086
		N	12	12
	Y	Correlation Coefficient	-.382	1.000
		Sig. (2-tailed)	.086	.
		N	12	12
Spearman's rho	X	Correlation Coefficient	1.000	-.648*
		Sig. (2-tailed)	.	.023
		N	12	12
	Y	Correlation Coefficient	-.648*	1.000
		Sig. (2-tailed)	.023	.
		N	12	12
*. Correlation is significant at the 0.05 level (2-tailed).				

a. Korelasi Kendall

$$H_0 : \tau = 0$$

$$H_1 : \tau \neq 0$$

Korelasi antara X dan Y sebesar -0,382.

Jika taraf signifikansi yang digunakan 5%, maka nilai sig dibandingkan dengan 0,05

Nilai sig sebesar **0,086 > 0,05 → H_0 diterima**

b. Korelasi Spearman

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Korelasi antara X dan Y sebesar -0,648.

Jika taraf signifikansi yang digunakan 5%, maka nilai sig dibandingkan dengan 0,05

Nilai sig sebesar **0,023 < 0,05 → H_0 ditolak**

BAB 5

ANALISIS REGRESI LINIER BERGANDA (*MULTIPLE REGRESSION*)

A. Model Regresi Linier Berganda

Analisis regresi linier berganda adalah perluasan dari analisis regresi linier sederhana dimana pada regresi linier berganda menggunakan lebih dari satu variabel bebas. Bentuk umum dari analisis regresi linier berganda adalah:

$$Y = a + b_1 X_1 + b_2 X_2 + \cdots + b_p X_p$$

Untuk perhitungan nilai koefisien regresi pada regresi linier berganda dengan dua variabel bebas digunakan formula sebagai berikut:

$$b_1 = \frac{\sum x_2^2 \sum x_1 y - \sum x_2 y \sum x_1 x_2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

$$b_2 = \frac{\sum x_1^2 \sum x_2 y - \sum x_1 y \sum x_1 x_2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

$$a = \overline{Y} - b_1 \overline{X_1} - b_2 \overline{X_2}$$

Dengan:

$$\sum x_1^2 = \sum X_1^2 - n\overline{X_1}^2$$

$$\sum x_2^2 = \sum X_2^2 - n\overline{X_2}^2$$

$$\sum x_1 y = \sum X_1 Y - n\overline{X_1}\,\overline{Y}$$

$$\sum x_2 y = \sum X_2 Y - n\overline{X_2}\,\overline{Y}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - n\overline{X_1}\,\overline{X_2}$$

$$\sum y^2 = \sum Y^2 - n\overline{Y}^2$$

$$\overline{Y} = \sum Y \big/ n$$

$$\overline{X_1} = \sum X_1 \big/ n$$

$$\overline{X_2} = \sum X_2 \big/ n$$

Pengujian hipotesis:

H_0 : tidak ada pengaruh variabel X_1 dan X_2 terhadap Y

H_1 : terdapat pengaruh variabel X_1 dan X_2 terhadap Y

Berikut adalah tabel ANOVA yang digunakan dalam pengujian hipotesis regresi linier berganda dengan dua variabel bebas:

Tabel 5.1 Analysis of Variance (ANOVA)

Sumber Variasi	db	JK	RJK	F
Regresi	k-1	$b_1\left\{\sum X_1Y - n\overline{X_1}\,\overline{Y}\right\} + b_2\left\{\sum X_2Y - n\overline{X_2}\,\overline{Y}\right\}$	$JK_R \big/ k-1$	$\frac{RJK_R}{RJK_E}$
Error	n-k	$JK_T - JK_R$	$JK_E \big/ n-k$	
Total	n-1	$\sum Y^2 - n\overline{Y}^2$		

Dengan pengambilan keputusan tolak H_0 jika $F \geq F_{\alpha;k-1,n-k}$.

Contoh:

Tabel 5.2 Tabel Penolong dalam Analisis Regresi Linier Berganda

No	X_1	X_2	Y	X_1^2	X_2^2	Y^2	X_1Y	X_2Y	X_1X_2
1	3	5	9	9	25	81	27	45	15
2	4	6	7	16	36	49	28	42	24
3	3	7	9	9	49	81	27	63	21
4	2	8	8	4	64	64	16	64	16
5	1	9	9	1	81	81	9	81	9
Σ	13	35	42	39	255	356	107	295	85

$$\overline{Y} = {}^{42}\!/_{5} = 8.4$$

$$\overline{X_1} = {}^{13}\!/_{5} = 2.6$$

$$\overline{X_2} = {}^{35}\!/_{5} = 7$$

$$\sum x_1^2 = 39 - 5(2.6)^2 = 5.2$$

$$\sum x_2^2 = 255 - 5(7)^2 = 10$$

$$\sum x_1 y = 107 - 5(2.6x8.4) = -2.2$$

$$\sum x_2 y = 295 - 5(7x8.4) = 1$$

$$\sum x_1 x_2 = 85 - 5(2.6x7) = -6$$

$$\sum y^2 = 356 - 5(8.4)^2 = 3.2$$

$$b_1 = \frac{10x(-2.2) - 1x(-6)}{5.2x10 - (-6)^2} = -1$$

$$b_2 = \frac{5.2x1 - (-2.2)x(-6)}{5.2x10 - (-6)^2} = -0.5$$

$$a = 8.4 - (-1)x2.6 - (-0.5)x7 = 14.5$$

Sehingga diperoleh persamaan regresi $Y = 14.5 - X_1 - 0.5X_2$

.

Tabel 5.3 ANOVA

Sumber Variasi	db	JK	RJK	F
Regresi	1	$(-1)\{107 - 5x2.6x8.4\}$ $+(-0.5)\{295 - 5x7x8.4\} = 1.7$	1.7	0.016
Error	3	$314 - 1.7 = 312.3$	$312.3/3 = 104.1$	
Total	4	$356 - 5x8.4 = 314$		

Dari hasil tabel ANOVA di atas dapat disimpulkan bahwa gagal tolak H_0 karena $F = 0.016 < 10.13 = F_{0.05;1,3}$ atau dengan kata lain tidak ada pengaruh yang signifikan antara variabel X_1 dan X_2 terhadap Y.

B. Asumsi-asumsi dalam Regresi Linier Berganda

Asumsi-asumsi yang harus terpenuhi dalam analisis regresi linier berganda adalah:

1. Normalitas

   Pengujian asumsi normalitas bertujuan untuk melihat apakah residual dalam regresi berdistribusi normal. Asumsi ini tidak boleh dilanggar karena uji t dan F mengasumsikan bahwa residual berdistribusi normal sehingga uji statistik akan menjadi tidak valid apabila asumsi ini dilanggar. Normalitas residual dapat diketahui dari cara berikut:

   a. Grafik

      Grafik adalah cara yang paling mudah untuk mendeteksi normalitas residual yaitu dengan membandingkan distribusi kumulatif data dengan distribusi kumulatif distribusi normal. Grafik yang menggambarkan perbandingan ini disebut *normal probability plot*.

b. Uji Statistik

Uji statistik yang sering dipakai dalam mendeteksi normalitas residual adalah uji Kolmogorov-Smirnov (KS). Uji KS dapat dengan mudah dilakukan dengan bantuan *software* statistik SPSS maupun Minitab. Hipotesis yang digunakan adalah:

$$H_0: residual\ berdistribusi\ normal$$
$$H_1: residual\ tidak\ berdistribusi\ normal$$

H_0 diterima atau residual dinyatakan berdistribusi normal apabila nilai signifikansi hasil pengujian di atas 0,05.

2. Heteroskedastisitas

Model regresi harus terbebas dari masalah heteroskedastisitas yang berarti bahwa terdapat *variance* yang sama dalam model. Adanya heteroskedastisitas dapat dideteksi dengan melihat plot antara nilai prediksi variabel respon dengan nilai residualnya. Jika plot menyebar dan tidak membentuk pola tertentu maka dapat disimpulkan tidak terjadi kasus heteroskedastisitas. Adanya heteroskedastisitas juga dapat dideteksi dengan menggunakan uji *Glejser* yaitu dengan meregresikan harga mutlak *error* dengan variabel independen. Terdapat kasus heteroskedastisitas apabila nilai signifikansi hasil pengujian di bawah 0,05

3. Multikolinieritas

Uji multikolinieritas digunakan untuk mendeteksi adanya hubungan atau korelasi yang tinggi antar variabel bebas. Model regresi yang baik adalah model yang bebas dari multikolinieritas. Multikolinieritas dapat dideteksi dengan melihat nilai korelasi antar variabel bebas. Jika terdapat korelasi yang kuat dan signifikan antar variabel bebas maka terindikasi adanya kasus multikolinieritas. Cara lain yang dapat digunakan adalah dengan melihat nilai *Variance Inflation Factor* (VIF), jika nilai VIF > 10 maka terindikasi adanya multikolinieritas.

4. Autokorelasi

Uji autokorelasi digunakan untuk melihat apakah terdapat korelasi yang linier antar residual pada periode t dengan periode t-1 pada model regresi. Model regresi yang baik adalah regresi yang bebas dari masalah autokorelasi. Autokorelasi dapat dideteksi dengan menggunakan nilai *Durbin-Watson* (DW) dengan kriteria:

a) Ada autokorelasi positif jika $0 < d < d_L$
b) Ada autokorelasi negatif jika $4 - d_L < d < 4$
c) Tidak ada autokorelasi jika $d_U \leq d \leq 4 - d_U$
d) Tidak dapat disimpulkan jika $d_L \leq d \leq d_U$ atau $4 - d_U \leq d \leq 4 - d_L$

C. Contoh Kasus

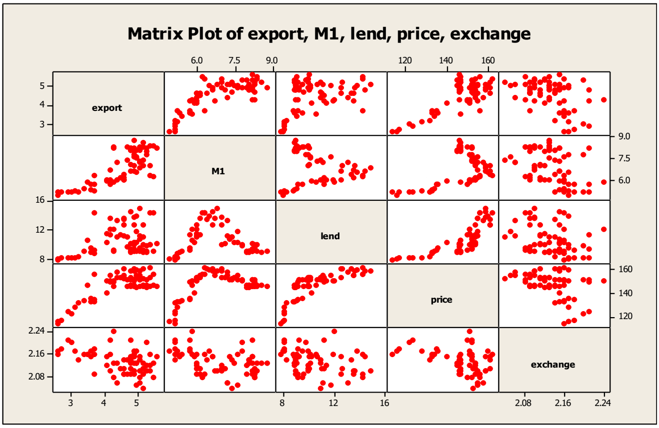
Understanding the relationship between U.S. export to Singapore and certain variables affecting the economy of the country.

Data

export	M1	lend	price	exchange	export	M1	lend	price	exchange
2.6	5.1	7.8	114	2.16	4.4	7.2	13.2	158	2.06
2.6	4.9	8.0	116	2.17	5.0	7.6	11.8	155	2.05
2.7	5.1	8.1	117	2.18	5.1	7.2	11.2	155	2.06
3.0	5.1	8.1	122	2.20	4.8	7.1	10.1	154	2.11

2.9	5.1	8.1	124	2.21	5.4	7.0	10.0	154	2.12
3.1	5.2	8.1	128	2.17	5.0	7.5	10.2	154	2.13
3.2	5.1	8.3	132	2.14	5.2	7.4	11.0	153	2.04
3.7	5.2	8.8	133	2.16	4.7	7.4	11.0	152	2.14
3.6	5.3	8.9	133	2.15	5.1	7.3	10.7	152	2.15
3.4	5.4	9.1	134	2.16	4.9	7.6	10.2	152	2.16
3.7	5.7	9.2	135	2.18	4.9	7.8	10.0	151	2.17
3.6	5.7	9.5	136	2.17	5.3	7.8	9.8	152	2.20
4.1	5.9	10.3	140	2.15	4.8	8.2	9.3	152	2.21
3.5	5.8	10.6	147	2.16	4.9	8.2	9.3	152	2.15
4.2	5.7	11.3	150	2.21	5.1	8.3	9.5	152	2.08
4.3	5.8	12.1	151	2.24	4.3	8.3	9.2	150	2.08
4.2	6.0	12.0	151	2.16	4.9	8.0	9.1	147	2.09
4.1	6.0	11.4	151	2.12	5.3	8.2	9.0	147	2.10
4.6	6.0	11.1	153	2.11	4.8	8.2	9.0	146	2.09
4.4	6.0	11.0	154	2.13	5.3	8.0	8.9	145	2.12
4.5	6.1	11.3	154	2.11	5.0	8.1	9.0	145	2.13
4.6	6.0	12.6	154	2.09	5.1	8.1	9.0	146	2.14
4.6	6.1	13.6	155	2.09	4.8	8.1	9.0	147	2.14
4.2	6.7	13.6	155	2.10	4.8	8.1	8.9	147	2.13
5.5	6.2	14.3	156	2.08	5.2	8.6	8.9	147	2.13
3.7	6.3	14.3	156	2.09	4.9	8.8	9.0	146	2.13
4.9	7.0	13.7	159	2.10	5.5	8.4	9.1	147	2.13
5.2	7.0	12.7	161	2.11	4.3	8.2	9.0	146	2.13
4.9	6.6	12.6	161	2.15	5.2	8.3	9.2	146	2.09
4.6	6.4	13.4	161	2.14	4.7	8.3	9.6	146	2.09
5.4	6.3	14.3	162	2.16	5.4	8.4	10.0	146	2.10
5.0	6.5	13.9	160	2.17	5.2	8.3	10.0	147	2.11
4.8	6.6	14.5	159	2.15	5.6	8.2	10.1	146	2.15
5.1	6.8	15.0	159	2.10					

1. Scatter plot



2. Korelasi

Correlations: export, M1, lend, price, exchange

	export	M1	lend	price
M1	0.775			
	0.000			
lend	0.335	-0.112		
	0.006	0.367		
price	0.770	0.447	0.745	
	0.000	0.000	0.000	
exchange	-0.433	-0.410	-0.279	-0.420
	0.000	0.001	0.022	0.000

Cell Contents: Pearson correlation
P-Value

Dari hasil analisis korelasi di atas, menunjukkan adanya korelasi yang signifikan antara variabel respon dengan semua variabel prediktor. Hal ini mengindikasikan adanya **hubungan linier** antara variabel respon dengan variabel prediktor.

Selain itu ada korelasi yang signifikan antar variabel prediktor, yaitu antara M1 dengan *price*, M1 dengan *exchange*, *lend* dengan *price*, *lend* dengan *exchange*, dan *price* dengan *exchange*. Hal ini mengindikasikan adanya kasus **multikolinieritas**.

D. Analisis Multikolinieritas

1. Regresi Serentak

Regression Analysis: export versus M1, lend, price, exchange

```
The regression equation is
export = - 4.02 + 0.368 M1 + 0.0047 lend + 0.0365
price + 0.27 exchange

Predictor        Coef    SE Coef      T      P
VIF
Constant       -4.015      2.766  -1.45  0.152
M1            0.36846    0.06385   5.77  0.000
3.207
lend          0.00470    0.04922   0.10  0.924
5.354
price        0.036511   0.009326   3.91  0.000
6.289
exchange        0.268      1.175   0.23  0.820
1.386

S = 0.335765   R-Sq = 82.5%   R-Sq(adj) = 81.4%
```

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	32.9463	8.2366	73.06	0.000
Residual Error	62	6.9898	0.1127		
Total	66	39.9361			

Source	DF	Seq SS
M1	1	23.9953
lend	1	7.1933
price	1	1.7519
exchange	1	0.0059

Unusual Observations

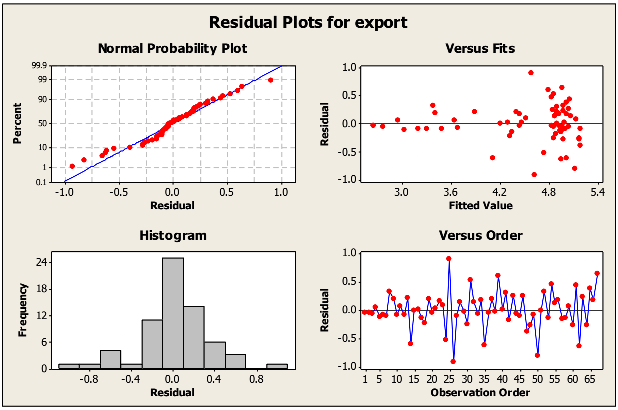
Obs	M1	export	Fit	SE Fit	Residual	St Resid
25	6.20	5.5000	4.5891	0.1107	0.9109	2.87R
26	6.30	3.7000	4.6286	0.1081	-0.9286	-2.92R
50	8.30	4.3000	5.1198	0.0846	-0.8198	-2.52R

R denotes an observation with a large standardized residual.

Dari hasil regresi diperoleh model regresi:

export=(-4.02)+0.368M1 + 0.0047 lend + 0.0365 price + 0.27 exchange dengan R^2 sebesar 82.5% akan tetapi tidak semua variabel prediktor berkontribusi signifikan terhadap variabel respon (export), yaitu variabel lend dan exchange yang tidak berkontribusi signifikan.

Dari nilai VIF terlihat bahwa tidak terjadi multikolinieritas antar variabel prediktor. Hal ini kontradiksi dengan yang terlihat pada *scatterplot* dan hasil korelasi sehingga perlu dilakukan regresi secara parsial untuk mengetahui ada tidaknya kasus multikolinieritas antar variabel prediktor dengan membandingkan tanda dari koefisien hasil regresi secara parsial dan serentak.



Dari *Normal Probability Plot* terlihat bahwa residual sudah memenuhi asumsi kenormalan akan tetapi plot *Versus Fits* tampak membentuk pola seperti corong yang mengindikasikan adanya kasus heteroskesdastisitas.

2. Regresi Parsial

- **Export dengan M1**

Regression Analysis: export versus M1

```
The regression equation is
export = 0.935 + 0.520 M1

Predictor      Coef   SE Coef     T      P
Constant     0.9348    0.3683  2.54  0.014
M1          0.52013   0.05258  9.89  0.000

S = 0.495221   R-Sq = 60.1%   R-Sq(adj) =
59.5%

Analysis of Variance

Source           DF       SS       MS       F
P
```

```
Regression        1  23.995  23.995  97.84
0.000
Residual Error   65  15.941   0.245
Total            66  39.936

Unusual Observations

Obs    M1  export     Fit  SE Fit  Residual  St Resid
  1  5.10  2.6000  3.5875  0.1127   -0.9875        -
2.05R
 25  6.20  5.5000  4.1596  0.0711    1.3404
2.73R
 31  6.30  5.4000  4.2116  0.0685    1.1884
2.42R

R denotes an observation with a large
standardized residual.
```

Koefisien M1 pada regresi secara parsial bertanda **positif** (sama dengan tanda pada regresi secara serentak).

- **Export dengan Land**

Regression Analysis: export versus lend

```
The regression equation is
export = 3.12 + 0.134 lend

Predictor      Coef   SE Coef     T      P
Constant     3.1174    0.5005  6.23  0.000
lend        0.13411   0.04679  2.87  0.006

S = 0.738559   R-Sq = 11.2%   R-Sq(adj) = 9.9%
Analysis of Variance
Source           DF        SS      MS     F
P
Regression        1    4.4806  4.4806  8.21
0.006
Residual Error   65   35.4555  0.5455
Total            66   39.9361

Unusual Observations

Obs  lend  export     Fit  SE Fit  Residual  St Resid
  1   7.8  2.6000  4.1635  0.1560   -1.5635        -
2.17R
  2   8.0  2.6000  4.1903  0.1485   -1.5903        -
2.20R
```

```
  3   8.1  2.7000  4.2037  0.1448   -1.5037     -
2.08R
 34  15.0  5.1000  5.1290  0.2282   -0.0290     -0.04
X

R denotes an observation with a large
standardized residual.
X denotes an observation whose X value gives
it large leverage.
```

Koefisien *lend* pada regresi secara parsial bertanda **positif** (sama dengan tanda pada regresi secara serentak).

- **Export dengan Price**

Regression Analysis: export versus price

```
The regression equation is
export = - 3.41 + 0.0539 price

Predictor       Coef   SE Coef      T      P
Constant     -3.4082    0.8183  -4.16  0.000
price       0.053886  0.005541   9.73  0.000

S = 0.500245   R-Sq = 59.3%   R-Sq(adj) =
58.6%

Analysis of Variance

Source          DF      SS      MS      F
P
Regression       1  23.670  23.670  94.59
0.000
Residual Error  65  16.266   0.250
Total           66  39.936

Unusual Observations

Obs  price  export     Fit  SE Fit  Residual  St Resid
  1    114  2.6000  2.7348  0.1943   -0.1348     -0.29 X
  2    116  2.6000  2.8426  0.1838   -0.2426     -0.52 X
  3    117  2.7000  2.8965  0.1786   -0.1965     -0.42 X
  4    122  3.0000  3.1659  0.1528   -0.1659     -0.35 X
 14    147  3.5000  4.5131  0.0611   -1.0131     -2.04R
 26    156  3.7000  4.9981  0.0779   -1.2981     -2.63R
 67    146  5.6000  4.4592  0.0615    1.1408      2.30R

R denotes an observation with a large
standardized residual.
```

```
X denotes an observation whose X value gives
it large leverage.
```

Koefisien *price* pada regresi secara parsial bertanda **positif** (sama dengan tanda pada regresi secara serentak).

- **Export dengan Exchange**

Regression Analysis: export versus exchange

```
The regression equation is
export = 21.9 - 8.14 exchange


Predictor     Coef   SE Coef      T      P
Constant    21.877     4.482   4.88  0.000
exchange    -8.135     2.101  -3.87  0.000

S = 0.706599   R-Sq = 18.7%   R-Sq(adj) =
17.5%

Analysis of Variance

Source          DF       SS      MS      F
P
Regression       1   7.4828  7.4828  14.99
0.000
Residual Error  65  32.4533  0.4993
Total           66  39.9361

Unusual Observations

Obs  exchange  export     Fit  SE Fit  Residual  St Resid
  1      2.16  2.6000  4.3049  0.1038   -1.7049    -2.44R
  2      2.17  2.6000  4.2236  0.1168   -1.6236    -2.33R
  3      2.18  2.7000  4.1422  0.1319   -1.4422    -2.08R
 16      2.24  4.3000  3.6541  0.2418    0.6459     0.97 X
 41      2.04  5.2000  5.2812  0.2128   -0.0812    -0.12 X

R denotes an observation with a large
standardized residual.
X denotes an observation whose X value gives
it large leverage.
```

Koefisien *price* pada regresi secara parsial bertanda **negatif** (berbeda dengan tanda pada regresi secara serentak).

Variabel	Tanda Koefisien Regresi	
	Serentak	Parsial
M1	Positif	Positif
Lend	Positif	Positif
Price	Positif	Positif
Exchange	Positif	Negatif

Adanya perbedaan tanda pada koefisien regresi mengindikasikan adanya kasus multikolinieritas sehingga perlu dilakukan penanganan lebih lanjut untuk menanggulangi adanya kasus ini. Salah satu cara menangani kasus multikolinieritas adalah dengan menggunakan **Regresi Stepwise**.

3. Regresi Stepwise

Stepwise Regression: export versus M1, lend, price, exchange

```
  Alpha-to-Enter: 0.15  Alpha-to-Remove:
0.15
Response is export on 4 predictors, with N =
67
Step               1        2
Constant      0.9348  -3.4230

M1             0.520    0.361
T-Value         9.89     9.21
P-Value        0.000    0.000

price                  0.0370
T-Value                  9.05
P-Value                 0.000

S              0.495    0.331
R-Sq           60.08    82.48
```

```
R-Sq(adj)     59.47    81.93
Mallows Cp     78.4      1.1
```

Dari hasil regresi stepwise diperoleh persamaan regresi: **`export = - 3.4230 + 0.361 M1 + 0.0370 price`** dengan R^2 sebesar 82.48% dan mengeliminasi variabel *lend* dan *exchange*. Selanjutnya akan dilakukan regresi antara *export* dengan variabel prediktor M1 dan *Price* untuk mendapatkan residualnya yang akan diuji asumsi IIDN (Identik, Independen dan Distribusi Normal).

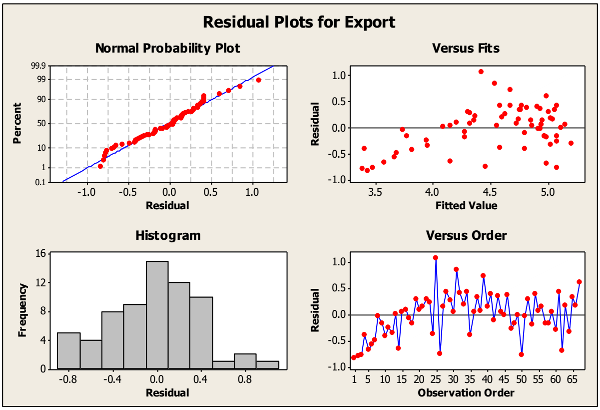
4. Regresi *Export* dengan M1 dan *Price*

   a. Hasil Regresi *Export* dengan M1 dan *Price*

Regression Analysis: export versus M1, price

```
The regression equation is
export = - 3.42 + 0.361 M1 + 0.0370 price

Predictor       Coef   SE Coef      T      P
Constant     -3.4230    0.5409  -6.33  0.000
M1           0.36142   0.03925   9.21  0.000
price       0.037033  0.004094   9.05  0.000

S = 0.330621   R-Sq = 82.5%   R-Sq(adj) =
81.9%


Analysis of Variance

Source          DF      SS      MS       F      P
Regression       2  32.940  16.470  150.67  0.000
Residual Error  64   6.996   0.109
Total           66  39.936

Source  DF  Seq SS
M1       1  23.995
price    1   8.945

Unusual Observations
Obs     M1   export     Fit   SE Fit   Residual   St Resid
```

```
 1   5.10   2.6000   2.6420   0.1288    -0.0420      -0.14 X
 2   4.90   2.6000   2.6438   0.1234    -0.0438      -0.14 X
25   6.20   5.5000   4.5949   0.0676     0.9051       2.80R
26   6.30   3.7000   4.6311   0.0651    -0.9311      -2.87R
50   8.30   4.3000   5.1317   0.0648    -0.8317      -2.57R
67   8.20   5.6000   4.9474   0.0668     0.6526       2.02R

R denotes an observation with a large
standardized residual.
X denotes an observation whose X value gives
it large leverage.
```



Dari hasil regresi stepwise diperoleh persamaan regresi: **`export = - 3.42 + 0.361 M1 + 0.0370 price`** dengan R^2 sebesar 82.5% yang berarti bahwa:

- Setiap M1 naik satu satuan, *export* akan meningkat sebesar 0.361
- Setiap *Price* naik satu satuan, *export* akan meningkat sebesar 0.0370
- Variabel M1 dan *price* mampu menjelaskan nilai *export* sebesar 82.5%

Dari gambar *Residual Plots for Export*:

- *Normal Probability Plot* menunjukkan residual tidak mengikuti distribusi normal
- *Versus Fits* menunjukkan plot residual menyebar secara acak dan homogen, hal ini mengindikasikan adanya homoskesdastisitas atau tidak terjadinya kasus heteroskesdastisitas.

5. Pengujian Asumsi Residual

   a. Identik

   Uji identik dilakukan dengan menggunakan uji *Glejser* yaitu dengan meregresikan nilai mutlak dari residual dengan variabel prediktor (M1 dan *Price*). Hasil regresi yang diperoleh adalah sebagai berikut:

Regression Analysis: ABS(RESI) versus M1, price

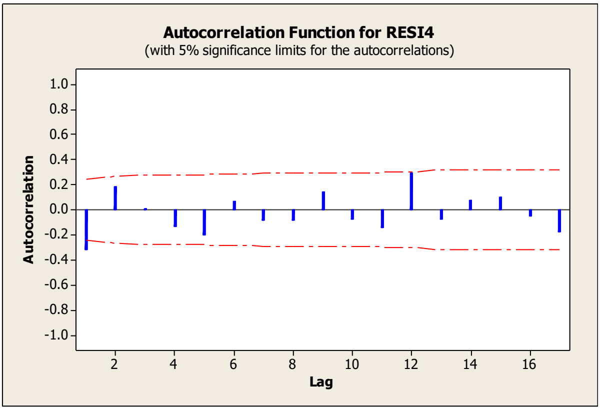
```
The regression equation is
ABS(RESI) = - 0.480 + 0.0280 M1 + 0.00353
price

Predictor        Coef    SE Coef       T
P
Constant      -0.4796     0.3586   -1.34
0.186
M1            0.02801    0.02602    1.08
0.286
price        0.003532   0.002715    1.30
0.198

S = 0.219235   R-Sq = 7.4%   R-Sq(adj) =
4.5%

Analysis of Variance

Source          DF        SS        MS      F       P
Regression       2   0.24651   0.12325   2.56   0.085
Residual Error  64   3.07609   0.04806
Total           66   3.32259

Source   DF    Seq SS
```

```
M1         1  0.16515
price      1  0.08136

Durbin-Watson statistic = 2.57648
```

Dari hasil regresi di atas, diketahui bahwa semua parameter regresi tidak signifikan dan regresinya juga tidak signifikan. Hal ini menunjukkan bahwa tidak adanya kasus heteroskedastisitas sehingga asumsi identik atau homoskedastisitas telah terpenuhi.

b. Independen

Uji asumsi independensi dilakukan dengan melihat autokorelasi residual pada lag-1. Beikut adalah plot ACF dari residual:



Dari plot ACF di atas, terlihat bahwa plot ACF residual keluar pada lag-1. Hal ini mengindikasikan belum terpenuhinya asumsi independensi (terjadi autokorelasi). Untuk mengatasi adanya autokorelasi digunakan metode *The Cochrane-Orcutt Two Step Procedure* yaitu dengan membuat model lag-1: $Y_{t-1} = \beta_0 + \beta_1 X_{1(t-1)} + \beta_2 X_{2(t-1)} + e_{t-1}$.

Selanjutnya kedua ruas dari model tersebut dikalikan dengan ρ sehingga menjadi:
$\rho Y_{t-1} = \rho\beta_0 + \rho\beta_1 X_{1t-1} + \rho\beta_2 X_{2t-1} + \rho e_{t-1}$
Kemudian kurangkan persamaan tersebut dengan persamaan regresi awal sehingga diperoleh:

$$Y_t - \rho Y_{t-1} = \beta_0(1-\rho) + \beta_1(X_{1(t)} - \rho X_{1(t-1)}) + \beta_2(X_{2(t)} - \rho X_{2(t-1)}) + e_t - \rho e_{t-1}$$

yang selanjutnya dinyatakan sebagai

$$Y_t* = \beta_0 * + \beta_1 * X_{1(t)} * + \beta_2 * X_{2(t)} * + v_t$$

Nilai ρ diestimasi berdasarkan Durbin-Watson Statistik dengan perhitungan:

$$\hat{\rho} = 1 - \frac{d}{2} = 1 - \frac{2.57648}{2} = -0.28824$$

Regresikan

$Y_t* = Y_t - \rho Y_{t-1}$ dengan $X_{1(t)}* = X_{1(t)} - \rho X_{1(t-1)}$ dan $X_{2(t)}* = X_{2(t)} - \rho X_{2(t-1)}$

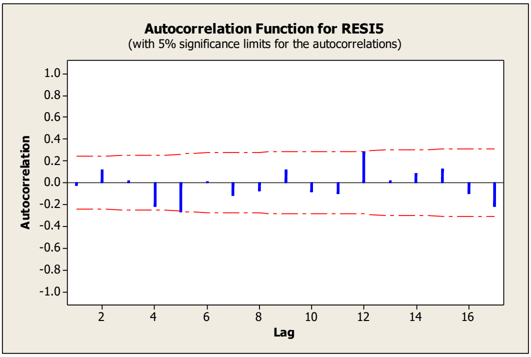
diperoleh hasil:

Regression Analysis: exp - p (lag versus M1 - p(lag M, price - p(lag price)

```
The regression equation is
exp - p (lag exp) = - 4.37 + 0.363 M1 - p(lag
M1) + 0.0367 price - p(lag price)
66 cases used, 1 cases contain missing values

Predictor                   Coef   SE Coef      T       P
Constant                 -4.3699    0.5480  -7.97
0.000
M1 - p(lag M1)           0.36346   0.02927  12.42
0.000
price - p(lag price)    0.036717  0.003173  11.57
0.000
S = 0.315041   R-Sq = 88.8%   R-Sq(adj) =
88.5%

Analysis of Variance
Source            DF       SS       MS        F
P
Regression         2   49.636   24.818   250.05  0.000
Residual Error    63    6.253    0.099
Total             65   55.889
```

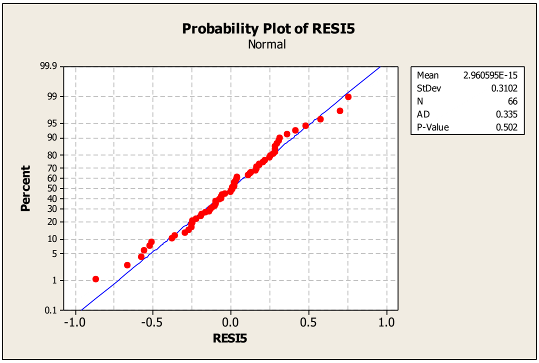
```
Source                     DF   Seq SS
M1 - p(lag M1)              1   36.346
price - p(lag price)        1   13.291
Durbin-Watson statistic = 1.97190
```

Untuk melihat masih ada atau tidaknya autokorelasi, dilihat plot ACF dari residual yang dihasilkan regresi di atas:



Dari plot ACF terlihat bahwa residual tidak keluar pada lag 1 sehingga disimpulkan sudah tidak ada kasus autokorelasi. Selanjutnya akan dilakukan pengujian asumsi berikutnya yaitu residual berdistribusi normal.

c. Distribusi Normal



Hipotesis:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Dari *Probability Plot* di atas diperoleh nilai *p-value* = 0.502 > 0.05. Hal ini berarti bahwa H_0 diterima dan diperoleh kesimpulan bahwa residual berdistribusi normal.

BAB 6
ANALISIS DERET BERKALA

A. Analisis Deret Berkala (*Time Series*)

Times series adalah himpunan nilai-nilai hasil pengamatan X_t yang diamati pada suatu waktu spesifik t. Jika dipunyai pengamatan *series* X_t pada himpunan waktu $T_0 = \{1,2,3,\ldots,n\}$ dan tertarik dengan prakiraan (*forecasting*) *series* untuk t=n+1, n+2, … maka X_t disebut *tines series* diskrit. Sedangkan jika *series* tersebut secara kontinuu atas suatu interval waktu $T_0=[0,1]$ maka X_t disebut *times series* kontinu. Langkah pertama dalam analisis *times series* adalah membuat grafik data melawan waktu. Langkah ini akan menunjukkan adanya trend, perilaku *seasonal* atau musiman, dan gambaran-gambaran sistematis lainnya dari data yang ada. Hal ini perlu dilakukan identifikasi agar perilaku-perilaku tersebut tidak terbawa ke dalam model matematika. Banyak model *times series* yang menghendaki data yang mempunyai histogram simetris sehingga perlu dilakukan transformasi data untuk membuat data tersebut mendekati simetris.

B. Metode Peramalan

1. Metode Peramalan Deret Waktu

   Peramalan adalah aktivitas untuk menghitung atau memprediksi beberapa kejadian atau kondisi yang akan datang yang umumnya sebagai hasil dari studi atau analisis dari sebagian data yang sudah ada. Peramalan diperlukan untuk mengetahui kapan dan bagaimana suatu peristiwa akan terjadi di masa mendatang sehingga dapat dipersiapkan tindakan yang lebih tepat yang dapat dilakukan.

Kecenderungan untuk dapat mengetahui peristiwa secara lebih tepat sebagai dasar untuk perencanaan di masa mendatang sangat dibantu dengan adanya metode peramalan yang merupakan cara untuk perencanaan ke depan.

Peramalan kuantitatif dapat diterapkan bila terdapat tiga kondisi berikut:

- Tersedia informasi tentang masa lalu
- Informasi masa lalu dapat dikuantitatifkan dalam bentuk data numerik
- Dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut di masa yang akan datang

Pengklasifikasian metode peramalan adalah dengan memperhatikan model yang mendasarinya. Terdapat dua jenis model peramalan yang utama, yaitu model deret berkala dan model deret kausal/eksplanatoris. Model deret berkala adalah pendugaan masa depan yang dilakukan berdasarkan nilai masa lalu dari suatu variabel dan/atau kesalahan masa lalu. Tujuan dari metode ini adalah untuk menemukan pola dalam deret data historis yang mengekstrapolasikan pola tersebut ke masa depan. Sedangkan model kausal/eksplanatoris adalah peramalan dengan mengasumsikan bahwa faktor yang diramalkan menunjukkan suatu hubungan sebab-akibat dengan satu atau lebih variabel. Maksud dari model ini adalah untuk menemukan bentuk hubungan dari variabel-variabel yang ada dan menggunakannya untuk meramalkan nilai mendatang dari variabel tak bebas. Metode peramalan dengan model deret berkala terdiri dari:

- Metode *Smoothing*
- Metode ARIMA Box Jenkins

- Metode Proyeksi Trend dengan Regresi

Identifikasi merupakan langkah awal analisis deret berkala untuk menentukan metode analisis yang tepat. Identifikasi pola dilakukan dengan membuat plot data untuk mendapatkan gambaran kecenderungan dari data secara grafis. Dari plot data tersebut akan diketahui apakah data bersifat acak, mempunyai pola trend, mempunyai pola musiman, atau mempunyai pola siklis. Adanya suatu trend dalam data berarti bahwa nilai-nilai yang berturut-turut sangat berkorelasi satu dengan yang lainnya. Dalam contoh yang demikian autokorelasi untuk satu *time lags* (r_1) adalah cukup besar. Autokorelasi untuk dua *time lags* (r_2) juga cukup besar, akan tetapi tidak sebesar yang pertama.

Musiman didefinisikan sebagai suatu pola yang berulang setelah interval waktu yang tetap. Misalnya penjualan es krim cukup besar pada musim panas, dan sebaliknya penjualan cukup kecil pada musim hujan. Pola penjualan tersebut akan berulang kembali selama masa dua belas bulan (satu tahun). Jika polanya adalah konsisten, maka koefisien autokorelasi dari lags 12 bulan (r_{12}) akan mempunyai nilai positif yang besar, yang menunjukkan adanya musiman. Jika tidak terdapat perbedaan-perbedaan yang nyata dari nol, maka hal ini menunjukkan bulan yang merupakan bagian dari satu tahun bersifat acak dengan pola yang tidak konsisten, yang timbul dari satu tahun ke tahun berikutnya. Data yang demikian adalah tidak musiman.

2. Uji Ketepatan Prakiraan

Untuk mengukur ketepatan prakiraan, dibutuhkan uji-uji ketepatan prakiraan. Beberapa uji ketepatan prakiraan yang sering digunakan antara lain sebagai berikut:

a. Deviasi Kuadrat Rata-rata/*Mean Squared Deviation* (MSD)

$$MSD = \frac{\sum_{1}^{n}(y_t - \hat{y}_t)^2}{n}$$

Dimana y_t adalah banyaknya data aktual, $\hat{y}_t$ adalah nilai prakiraan, dan n adalah banyaknya prakiraan.

b. Kesalahan Persentase Absolut Rata-rata/*Mean Absolute Percentage Error* (MAPE)

$$MAPE = \frac{\sum \left| \frac{(y_t - \hat{y}_t)}{y_t} \right|}{n} x100, (y_t \neq 0)$$

Dimana y_t adalah banyaknya data aktual, $\hat{y}_t$ adalah nilai prakiraan, dan n adalah banyaknya prakiraan.

c. Deviasi Absolut Rata-rata/*Mean Absolute Deviation* (MAD)

$$MAD = \frac{\sum |(y_t - \hat{y}_t)|}{n}$$

Dimana y_t adalah banyaknya data aktual, $\hat{y}_t$ adalah nilai prakiraan, dan n adalah banyaknya prakiraan.

Metode prakiraan yang paling sesuai pada umumnya adalah metode yang memiliki kesalahan rata-rata (MSD) dan kesalahan absolute (MAPE) yang paling kecil.

C. Prinsip dalam Peramalan

Metode peramalan dilakukan dengan cara mengekstrapolasi kondisi masa lalu untuk masa kondisi yang akan dating. Hal ini akan didasarkan pada asumsi bahwa kondisi masa lalu sama dengan kondisi masa mendatang. Atas dasar logika ini, langkah dalam metode peramalan secara umum adalah mengumpulkan data, menyeleksi dan memilih data, memilih model peramalan, menggunakan metode terpilih untuk melakukan peramalan, dan yang terakhir adalah evaluasi hasil akhir. Hal yang terpenting dalam peramalan adalah dapat meminimumkan kesalahan dalam peramalan.

D. Metode Peramalan Deret Waktu dengan Metode ARIMA

Model *Autoregressive Integrated Moving Average* (ARIMA) merupakan model statistik yang menganalisis sifat-sifat dari data runtun waktu yang telah ada di masa lalu sehingga didapat persamaan model yang menggambarkan hubungan dari data runtun waktu tersebut. Tahapan analisis dalam model ARIMA adalah sebagai berikut:

1. Tahap Identifikasi

   a. Pengujian Stasioneritas

      Dari nilai-nilai autokorelasi data asli yang telah dihitung, apabila nilai tersebut turun dengan cepat ke atau mendekati nol sesudah nilai kedua atau ketiga menandakan bahwa data tersebut stasioner ke data asli.

2. Penentuan Model Sementara

   Jika kestasioneran telah didapat, maka dari nilai-nilai autokorelasi dapat dilakukan perhitungan untuk mengetahui adanya pola-pola yang lain yang terkadang berada di dalam data runtun waktu. Kemungkinan pola yang terjadi adalah:

- Faktor musiman, dimana nilai autokorelasi untuk *time lags* setiap kuartal atau setiap tahun yang besar dan secara nyata berbeda dari nol
- Adanya proses AR atau MA. Pola dari autokorelasi dan autokorelasi parsial akan menunjukkan suatu model yang mungkin
- Musiman dan proses AR dan MA mungkin akan terlihat dan umumnya adalah model ARIMA

3. Penaksiran Parameter

   Setelah melakukan identifikasi model sementara, selanjutnya parameter-parameter AR dan MA, musiman dan non-musiman harus ditetapkan.

4. Pemeriksaan Diagnostik

   Setelah model ARIMA sementara ditetapkan, selanjtnya dilakukan pemeriksaan diagnostik untuk membuktikan bahwa model tersebut sudah cukup baik.

5. Peramalan dengan Model ARIMA

   Metode ARIMA merupakan metode yang dikembangkan oleh George Box dan Gwilym Jenkins sehingga nama mereka sering disinonimkan dengan proses ARIMA yang diterapkan untuk analisis data dan peramalan data runtun waktu. Metode ARIMA berbeda dengan metode peramalan lain karena metode ini dapat dipakai untuk semua tipe pola data. Metode ARIMA akan bekerja dengan baik apabila data runtun waktu yang digunakan bersifat dependen atau berhubungan satu sama lain secara statistik.

Secara umum model ARIMA dirumuskan dengan notasi ARIMA(p,d,q), dimana

p = orde atau derajat autokorelasi (AR)

d = orde atau derajat pembeda

q = orde atau derajat MA

Model ARIMA musiman umumnya dinotasikan dengan ARIMA(p,d,q)(P,D,Q)^s, dimana

(p,d,q) = bagian dari model yang non-musiman

(P,D,Q) = bagian dari model yang musiman

s = jumlah periode musiman

a. Model Autoregresif (AR)

Untuk mengetahui besarnya pengaruh dan hubungan suatu nilai variabel yang telah terjadi pada suatu periode berikutnya sering digunakan autoregresif. Model AR adalah model yang menggambarkan bahwa variabel dependen dipengaruhi oleh variabel dependen itu sendiri. Secara umum model AR mempunyai bentuk sebagai berikut:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t$$

Dimana:

Y_t = nilai variabel dependen pada waktu t

φ_0= intersep/nilai konstan

Y_{t-p}= variabel dependen yang dalam hal ini merupakan lag (beda waktu) dari variabel dependen pada satu periode sebelumnya.

ε_t= residual pada waktu t

Orde dari model AR diberi notasi p yang ditentukan oleh jumlah periode variabel dependen yang masuk dalam model.

b. Model Moving Average (MA)

Proses *moving average* dapat menyatakan adanya ketergantungan antara kesalahan acak pada waktu t dengan q kesalahan sebelumnya serta mengidentifikasi adanya pengaruh luar data yang menyebabkan adanya fluktuasi di dalam data selama q periode. Bila suatu pengamatan pada urutan waktu t bergantung pada penyimpangan pada saat t dan sebelumnya sepanjang q periode, maka modelnya berupa rata-rata bergerak dengan periode q, yaitu model *moving average* yang dinotasikan dengan MA(q).

Secara umum model MA mempunyai bentuk umum sebagai berikut:

$$Y_t = W_0 + \varepsilon_t - W_1\varepsilon_{t-1} - W_2\varepsilon_{t-2} - \cdots - W_q\varepsilon_{t-q}$$

Dimana:

Y_t= nilai variabel dependen pada waktu t

W_0= intersep/nilai konstan

$\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-q}$= nilai residual sebelumnya

$W_1, W_2, \ldots, W_q$= koefisien model MA yang menunjukkan bobot

ε_t= residual pada waktu t

Perbedaan model MA dengan AR terletak pada jenis variabel independen. Jika variabel pada model AR adalah nilai sebelumnya dari variabel independen maka pada model MA yang menjadi variabel independen adalah nilai residual pada periode sebelumnya.

c. Model Autoregresive Integrated Moving Average (ARIMA)

Suatu proses dikatakan tidak stasioner apabila proses tersebut mempunyai *means* dan *varians* yang tidak konstan untuk sebarang waktu pengamatan. Model deret waktu yang tidak stasioner dapat dikatakan sebagai ARIMA(p,d,q) yang merupakan gabungan model non-musiman. Bentuk umum dari model ARIMA(p,d,q) adalah sebagai berikut:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + W_1 \varepsilon_{t-1} - W_2 \varepsilon_{t-2} - \cdots - W_q \varepsilon_{t-q} + \varepsilon_t$$

Indikasi model:

- Autoregresif (AR) orde p
  - Menyatakan adanya ketergantungan pengamatan ke-t dengan p pengamatan sebelumnya
  - Mengindikasikan adanya pengaruh dari luar deret waktu yang menyebabkan adanya fluktuasi pada data dan timbulya ketergantungan antar pengamatan
- *Integrated* orde d

  Mengindikasikan adanya trend stokastik dalam deret dengan pola berorde d sejalan dengan tingkat deret.

- *Moving Average* orde q
  - Menyatakan adanya ketergantungan antara kesalahan acak pada indeks waktu t dengan q kesalahan acak sebelumnya
  - Mengindikasikan adanya pengaruh luar data itu sendiri selama q satuan waktu
- Musiman

  Adanya pola musiman dalam data deret waktu terindikasikan pada salah satu pola ACF dan PACF yang berulang dengan pola yang bersesuaian dari ACF maupun PACF model non-musiman sebagai berikut:
  - ACF menurun secara eksponensial pada lag s,…,s+p dan berulang pada lag 2s,…,2s+p, lag 3s,…,3s+p, dan seterusnya.
  - ACF signifikan pada lag s dan tidak signifikan pada lag 2s, 3s, … .
  - PACF tidak signifikan pada lag s
- Melakukan Estimasi Parameter dari Model Sementara

  Setelah model sementara ditetapkan, selanjutnya parameter-parameter AR dan MA baik yang musiman maupun non-musiman harus ditetapkan dengan cara yang terbaik. Penaksiran parameter adalah langkah untuk mendapatkan nilai-nilai awal parameter yang telah ditetapkan dalam identifikasi model awal.
- Pengujian Parameter

Setelah dilakukan identifikasi dan estimasi pada parameter, selanjutnya dilakukan pengujian parameter untuk mengetahui apakah parameter model sudah signifikan atau tidak. Dalam pengujian parameter model, digunakan hipotesis berikut:

$H_0 : \theta_i = 0$ (parameter tidak berpengaruh pada model)

$H_0 : \theta_i \neq 0$ (parameter berpengaruh pada model)

Dengan i=1,2,3,…,n

Statistik uji:

$$T = T_{hitung} = \frac{parameter\ estimasi}{standar\ deviasi\ sampel}$$
$$= \frac{\hat{\theta}_1}{s/\sqrt{n}}$$

Kriteria penguji: Menolah H_0 jika $|t| > \left(df, \frac{\alpha}{2}\right)$ atau jika p-value $<\alpha$ dengan df=(n-(p+q+1)).

E. Contoh Kasus

1. Data

Berikut adalah data Inflasi Bulanan Indonesia yang diperoleh dar diambil dari buku Statistik Indonesia tahun 2005–2009 yang berada di perpustakaan BPS Kota Malang dan diterbitkan oleh BPS serta dari website resmi BPS:

Indeks Harga Konsumen dan Inflasi Bulanan Indonesia,

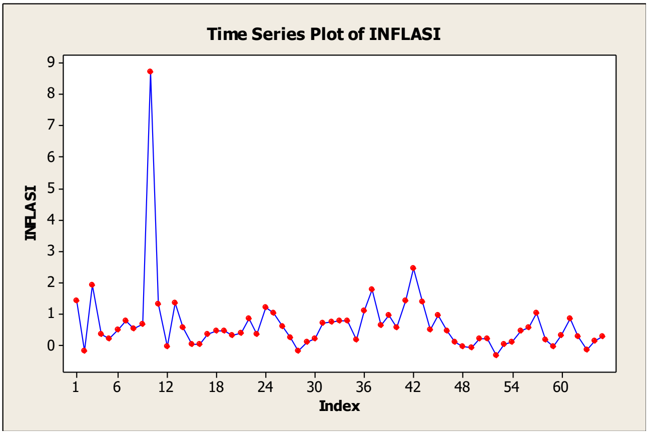
2005, 2006, 2007, Januari - Mei 2008 (2002=100), Juni - Desember 2008, 2009, Januari - Mei 2010 (2007 = 100)

Bulan	2005		2006		2007		2008		2009		2010	
	IHK	Inflasi	IHK	Inflasi	IHK	Inflasi	IHK	Inflasi	IHK	Inflasi	IHK	Inflasi
Januari	118.53	1.43	138.72	1.36	147.41	1.04	158.26	1.77	113.78	-0.07	118.01	0.84
Februari	118.33	-0.17	139.53	0.58	148.32	0.62	159.29	0.65	114.02	0.21	118.36	0.30
Maret	120.59	1.91	139.57	0.03	148.67	0.24	160.81	0.95	114.27	0.22	118.19	-0.14
April	121.00	0.34	139.64	0.05	148.43	-0.16	161.73	0.57	113.92	-0.31	118.37	0.15
Mei	121.25	0.21	140.16	0.37	148.58	0.10	164.01	1.41	113.97	0.04	118.71	0.29
Juni	121.86	0.50	140.79	0.45	148.92	0.23	110.08*)	2.46*)	114.10	0.11		
Juli	122.81	0.78	141.42	0.45	149.99	0.72	111.59	1.37	114.61	0.45		
Agustus	123.48	0.55	141.88	0.33	151.11	0.75	112.16	0.51	115.25	0.56		
September	124.33	0.69	142.42	0.38	152.32	0.80	113.25	0.97	116.46	1.05		
Oktober	135.15	8.70	143.65	0.86	153.53	0.79	113.76	0.45	116.68	0.19		
November	136.92	1.31	144.14	0.34	153.81	0.18	113.90	0.12	116.65	-0.03		
Desember	136.86	-0.04	145.89	1.21	155.50	1.10	113.86	-0.04	117.03	0.33		
Angka Inflasi		17.11		6.60		6.59		11.06		2.78		1.44

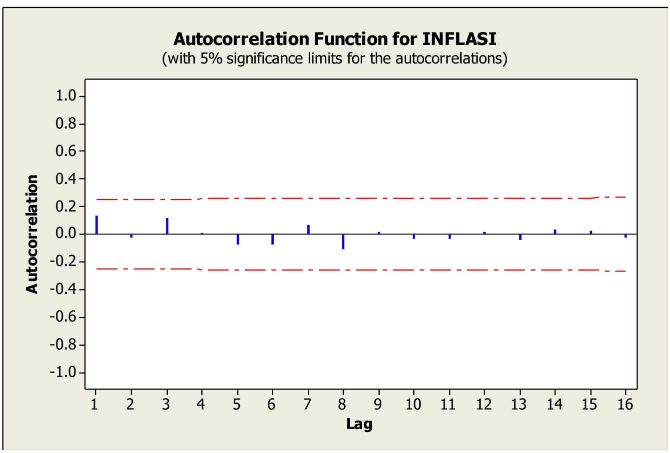
*) Sejak Juni 2008, IHK berdasarkan pola konsumsi yang didapat dari Survei Biaya Hidup di 66 Kota (2007=100)

2. Identifikasi Model

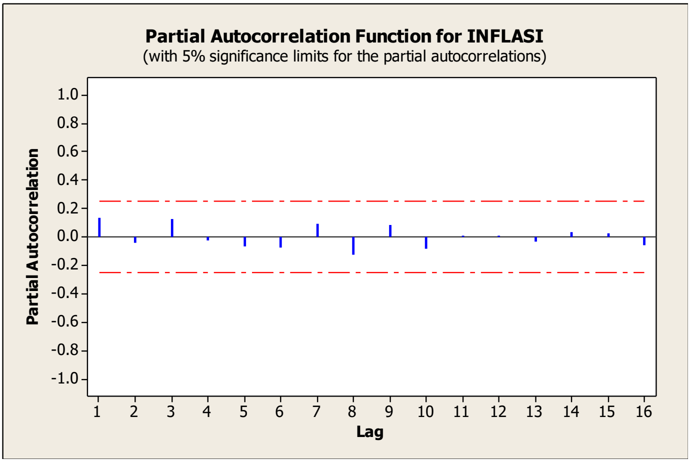
Identifikasi merupakan langkah awal dalam menentukan model *times series*. Ini bertujuan untuk mengetahui pola dan kestasioneran data. Plot data inflasi Indonesia periode Januari 2005 – Mei 2010 adalah sebagai berikut:



Selanjutnya dilakukan uji ACF dan PACF dari data asli inflasi Indonesia periode Januari 2005 – Mei 2010 dengan plot ACF dan PACF sebagai berikut:



Plot ACF terlihat turun secara eksponensial dan terjadi *cut off* pada lag 1 atau 2, sehingga dugaan awal untuk orde MA adalah orde 1 atau orde 2.



Plot PACF terlihat turun secara eksponensial dan terjadi *cut off* pada lag 1 atau 2, sehingga dugaan awal untuk orde AR adalah orde 1 atau orde 2. Dari plot ACF dan PACF di atas dapat diduga bahwa model yang sesuai adalah ARIMA (1,0,1), (1,0,2), (2,0,1), atau (2,0,2).

3. Penaksiran Model ARIMA

   a. ARIMA (1,0,1) atau ARMA (1,1)

      - Uji Signifikansi Parameter

        Dari output Minitab diperoleh:

```
Final Estimates of Parameters

Type          Coef   SE Coef       T      P
AR    1    -0.8537    0.0841  -10.15  0.000
MA    1    -0.9730    0.0426  -22.85  0.000
Constant    1.2649    0.2728    4.64  0.000
Mean        0.6823    0.1472

Number of observations:  65
Residuals:     SS =  79.6421
(backforecasts excluded)
```

```
MS =  1.2846  DF = 62
```

Dari output di atas terlihat bahwa untuk AR (1) memiliki $P - value = 0.000 < 0.05$, sehingga parameter untuk AR (1) signifikan. Untuk MA (1) memiliki $P - value = 0.000 < 0.05$, sehingga parameter untuk MA (1) signifikan.

- Uji Asumsi Independensi

  Uji ini dilakukan untuk mengetahui apakah deret sisaan atau residualnya bersifat *white noise* atau tidak. Dari output Minitab diperoleh:

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	2.7	4.8	12.7	17.1
DF	9	21	33	45
P-Value	0.975	1.000	0.999	1.000

Pada lag 12, $P - value = 0.975 > 0.05$ sehingga residual *white noise*.

Pada lag 24, $P - value = 1.000 > 0.05$ sehingga residual *white noise*.

Pada lag 36, $P - value = 0.999 > 0.05$ sehingga residual *white noise*.

Pada lag 48, $P - value = 1.000 > 0.05$ sehingga residual *white noise*.

Dari uji signifikansi parameter dan uji asumsi independensi di atas dapat disimpulkan bahwa model ARIMA (1,0,1) telah sesuai.

b. ARIMA (1,0,2) atau ARMA (1,2)

- Uji Signifikansi Parameter

  Dari output Minitab diperoleh:

Type	Coef	SE Coef	T	P
AR 1	-0.8646	0.0831	-10.41	0.000
MA 1	-1.0420	0.0239	-43.59	0.000
MA 2	-0.0656	0.0490	-1.34	0.186
Constant	1.2591	0.2951	4.27	0.000
Mean	0.6753	0.1582		

```
Number of observations:  65
Residuals:    SS =  78.6645 (backforecasts excluded)
              MS =  1.2896  DF = 61
```

  Dari output di atas terlihat bahwa untuk AR (1) memiliki $P - value = 0.000 < 0.05$, sehingga parameter untuk AR (1) signifikan. Untuk MA (1) memiliki $P - value = 0.000 < 0.05$, sehingga parameter untuk MA (1) signifikan. Untuk MA (2) memiliki $P - value = 0.186 \nless$

0.05, sehingga parameter untuk MA (2) tidak signifikan. Selanjutnya tidak perlu dilakukan uji asumsi independensi dan disimpulkan bahwa model ARIMA (1,0,2) tidak sesuai.

c. ARIMA (2,0,1) atau ARMA (2,1)

- Uji Signifikansi Parameter

  Dari output Minitab diperoleh:

```
Final Estimates of Parameters
Type        Coef  SE Coef     T     P
AR   1    0.6765  37.6987   0.02  0.986
AR   2   -0.0687   5.0474  -0.01  0.989
MA   1    0.5446  37.6919   0.01  0.989
Constant  0.26826  0.06591   4.07  0.000
Mean      0.6839   0.1680
Number of observations:  65
Residuals:    SS =  82.6294 (backforecasts excluded)
              MS =  1.3546  DF = 61
```

  Dari output di atas terlihat bahwa untuk AR (1) memiliki $P-value = 0.986 \nless 0.05$, sehingga parameter untuk AR (1) tidak signifikan. Untuk AR (2) memiliki $P-value = 0.989 \nless 0.05$, sehingga parameter untuk AR (2) tidak signifikan. Untuk MA (1) memiliki $P-value = 0.989 \nless 0.05$, sehingga parameter untuk MA (1) tidak signifikan. Selanjutnya tidak perlu dilakukan uji asumsi independensi dan disimpulkan bahwa model ARIMA (2,0,1) tidak sesuai.

d. ARIMA (2,0,2) atau ARMA (2,2)

- Uji Signifikansi Parameter

  Dari output Minitab diperoleh:

  Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-0.5440	0.1484	-3.67	0.001
AR 2	0.2809	0.1754	1.60	0.115
MA 1	-0.7334	0.0897	-8.18	0.000
MA 2	0.2459	0.1136	2.17	0.034
Constant	0.8492	0.2100	4.04	0.000
Mean	0.6723	0.1662		

  Number of observations: 65
  Residuals: SS = 77.9915 (backforecasts excluded)
  MS = 1.2999 DF = 60

  Dari output di atas terlihat bahwa untuk AR (1) memiliki $P - value = 0.001 < 0.05$, sehingga parameter untuk AR (1) signifikan. Untuk AR (2) memiliki $P - value = 0.115 \nless 0.05$, sehingga parameter untuk AR (2) tidak signifikan. Untuk MA (1) memiliki $P - value = 0.000 < 0.05$, sehingga parameter untuk MA (1) signifikan. Untuk MA (2) memiliki $P - value = 0.034 < 0.05$, sehingga parameter untuk MA (2) signifikan. Selanjutnya tidak perlu dilakukan uji asumsi independensi karena ada parameter yang tidak signifikan dan disimpulkan bahwa model ARIMA (2,0,2) tidak sesuai.

Dari keempat model di atas diperoleh model yang sesuai yaitu model ARIMA (1,0,1) atau ARMA (1,1) dan untuk selanjutnya model ini dapat digunakan untuk melakukan peramalan atau perkiraan nilai inflasi beberapa periode mendatang.

4. Prakiraan Model ARIMA

Dengan menggunakan Minitab, diperoleh hasil prakiraan atau peramalan dengan model ARIMA (1,0,1) sampai Desember 2010 sebagai berikut:

```
95 Percent Limits
Period  Forecast    Lower    Upper
  66    0.70674  -1.51513  2.92861
  67    0.66151  -1.57610  2.89912
  68    0.70013  -1.54888  2.94913
  69    0.66716  -1.59012  2.92444
  70    0.69530  -1.56799  2.95860
  71    0.67128  -1.59639  2.93894
  72    0.69179  -1.57905  2.96263
```

Bulan	Inflasi
Juni 2010	0.71
Juli 2010	0.66
Agustus 2010	0.70
September 2010	0.67
Oktober 2010	0.69
November 2010	0.67
Desember 2010	0.69

BAB 7

ANGKA INDEKS

A. Pengertian Angka Indeks

Angka indeks adalah suatu angka perbandingan yang dibuat dengan tujuan untuk mengukur perubahan nilai variabel berdasarkan tahun dasar tertentu. Terdapat dua istilah yang digunakan dalam angka indeks yaitu, tahun dasar dan tahun berjalan. Tahun dasar adalah tahun atau periode yang digunakan sebagai patokan dalam menentukan indeks selanjutnya, nilai indeks pada tahun dasar adalah 100. Tahun berjalan adalah tahun atau periode yang nilai variabelnya dibandingkan dengan tahun dasar. Angka indeks atau sering disebut dengan indeks pada umumnya digunakan untuk mengukur perubahan variabel sosial ekonomi seperti produksi, ekspor/impor, harga, jumlah uang beredar, dan lain-lain.

B. Macam-macam Angka Indeks

Macam-macam angka indeks adalah sebagai berikut:

1. Indeks Sederhana

Indeks sederhana digunakan untuk menghitung nilai indeks pada setiap komoditas saja.

a. Indeks Harga

$$I_P = \frac{P_t}{P_0} x100$$

Dengan:

P_t : harga pada tahun t

P_0 : harga pada tahun dasar

Contoh:

Tabel 7.1 Harga Jagung (Rp/kg) tahun 2013-2015

2013	2014	2015
3500	**3900**	**4500**

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{3900}{3500} x100 = 111.4$$

$$I_{2015} = \frac{4500}{3500} x100 = 128.57$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 harga jagung meningkat sebesar 11.4 % dan pada tahun 2015 meningkat sebesar 28.57%.

2. Indeks Kuantitas

$$I_Q = \frac{Q_t}{Q_0} x100$$

Dengan:

Q_t : kuantitas pada tahun t

Q_0 : kuantitas pada tahun dasar

Contoh:

Tabel 7.2 Produksi Jagung (ton) tahun 2013-2015

2013	2014	2015
117	**112**	**120**

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{112}{117} x100 = 95.73$$

$$I_{2015} = \frac{120}{117} x100 = 102.56$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 jumlah produksi jagung menurun sebesar 4.27 % dan pada tahun 2015 meningkat sebesar 2.56%.

3. Indeks Nilai

$$I_N = \frac{P_t \cdot Q_t}{P_0 \cdot Q_0} x100$$

Contoh:

Berdasarkan tabel 7.1 dan 7.2 Mmisal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{3900x112}{3500x117} x100 = 106.67$$

$$I_{2015} = \frac{4500x120}{3500x117} x100 = 131.87$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 jumlah nilai jagung meningkat sebesar 6.67 % dan pada tahun 2015 meningkat sebesar 31.87%.

4. Indeks Agregatif

Indeks agregatif digunakan untuk menghitung nilai indeks pada beberapa komoditas sekaligus.

a. Indeks Harga Agregatif

$$I_P = \frac{\sum P_t}{\sum P_0} x100$$

Dengan:

$\sum P_t$: jumlah harga keseluruhan komoditas pada tahun t

$\sum P_0$: jumlah harga keseluruhan komoditas pada tahun dasar

Contoh:

Tabel 7.3 Harga Komoditas Pertanian (Rp/kg) tahun 2013-2015

Komoditas	2013	2014	2015
Jagung	7500	7900	7500
Beras	10600	10800	11500
Kedelai	11000	11200	11800
Jumlah	29100	29900	30800

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{29900}{29100} x100 = 102.75$$

$$I_{2015} = \frac{30800}{29100} x100 = 105.84$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 harga komoditas pertanian meningkat sebesar 2.75 % dan pada tahun 2015 meningkat sebesar 5.84%.

b. Indeks Kuantitas Agregatif

$$I_P = \frac{\sum Q_t}{\sum Q_0} x100$$

Dengan:

$\sum Q_t$: jumlah kuantitas keseluruhan komoditas pada tahun t

$\sum Q_0$: jumlah kuantitas keseluruhan komoditas pada tahun dasar

Contoh:

Tabel 7.4 Jumlah Produksi Komoditas Pertanian (ton) tahun 2013-2015

Komoditas	2013	2014	2015
Jagung	112	115	118
Beras	130	128	122
Kedelai	100	98	180
Jumlah	342	341	420

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{341}{342} x100 = 99.71$$

$$I_{2015} = \frac{420}{342} x100 = 122.81$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 jumlah produksi komoditas pertanian turunt sebesar 0.29 % dan pada tahun 2015 meningkat sebesar 22.81%.

c. Indeks Nilai Agregatif

$$I_P = \frac{\sum P_t \cdot Q_t}{\sum P_0 \cdot Q_0} x100$$

Contoh: (Lihat Tabel 7.3 dan 7.4)

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{(7900x115 + 10800x128 + 11200x98)}{(7500x112 + 10600x130 + 11000x100)} x100 = 102.12$$

$$I_{2015} = \frac{(7500x118 + 11500x122 + 11800x180)}{(7500x112 + 10600x130 + 11000x100)} x100 = 132.97$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.12 % dan pada tahun 2015 meningkat sebesar 32.97%.

5. Indeks Agregatif Tertimbang

a. Formula Laspeyres

$$I_L = \frac{\sum P_t \cdot Q_0}{\sum P_0 \cdot Q_0} x100$$

Contoh: (Lihat Tabel 7.3 dan 7.4)

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{(7900x112 + 10800x130 + 11200x100)}{(7500x112 + 10600x130 + 11000x100)} x100 = 102.74$$

$$I_{2015} = \frac{(7500x112 + 11500x130 + 11800x100)}{(7500x112 + 10600x130 + 11000x100)} x100 = 105.94$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.74 % dan pada tahun 2015 meningkat sebesar 5.94%.

b. Formula *Paasche*

$$I_P = \frac{\sum P_t \cdot Q_t}{\sum P_0 \cdot Q_t} x100$$

Contoh: (Lihat Tabel 7.3 dan 7.4)

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{(7900x115 + 10800x128 + 11200x98)}{(7500x115 + 10600x128 + 11000x98)} x100 = 102.77$$

$$I_{2015} = \frac{(7500x118 + 11500x122 + 11800x180)}{(7500x118 + 10600x122 + 11000x180)} x100 = 106.1$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.77 % dan pada tahun 2015 meningkat sebesar 6.1%.

c. Formula *Fisher*

$$I_F = \sqrt{ILxIP}$$

Contoh: (perhatikan hasil dari contoh formula *laspeyres* dan *fisher*)

$$I_{2014} = \sqrt{102.74x102.77} = 102.76$$

$$I_{2015} = \sqrt{105.94x106.1} = 106.2$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.76 % dan pada tahun 2015 meningkat sebesar 6.2%.

d. Formula Marshal-Edgeworth

$$I_{ME} = \frac{\sum P_t \cdot \left(Q_t + Q_0\right)}{\sum P_0 \cdot \left(Q_t + Q_0\right)} x100$$

Contoh: (Lihat Tabel 7.3 dan 7.4)

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{7900x(115+112)+10800x(128+130)+11200x(98+100)}{7500x(115+112)+10600x(128+130)+11000x(98+100)}x100$$
$$= 102.75$$

$$I_{2015} = \frac{7500x(118+112)+11500x(122+130)+11800x(180+100)}{7500x(118+112)+10600x(122+130)+11000x(180+100)}x100$$
$$= 106.03$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.75 % dan pada tahun 2015 meningkat sebesar 6.03%.

e. Formula *Drobisch*

$$I_D = \frac{I_L + I_P}{2}$$

Contoh: (perhatikan hasil dari contoh formula *laspeyres* dan *fisher*)

$$I_{2014} = \frac{102.74 + 102.77}{2} = 102.76$$

$$I_{2015} = \frac{105.94 + 106.1}{2} = 106.02$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.76 % dan pada tahun 2015 meningkat sebesar 6.02%.

f. Formula *Walsh*

$$I_W = \frac{\sum P_t \cdot \sqrt{Q_t \cdot Q_0}}{\sum P_0 \cdot \sqrt{Q_t \cdot Q_0}} x100$$

Contoh: (Lihat Tabel 7.3 dan 7.4)

Misal menggunakan tahun 2013 sebagai tahun dasar, maka:

$$I_{2014} = \frac{7900x\sqrt{115\cdot 112} + 10800x\sqrt{128\cdot 130} + 11200x\sqrt{98\cdot 100}}{7500x\sqrt{115\cdot 112} + 10600x\sqrt{128\cdot 130} + 11000x\sqrt{98\cdot 100}}x100$$
$$= 102.75$$

$$I_{2015} = \frac{7500x\sqrt{118\cdot 112} + 11500x\sqrt{122\cdot 130} + 11800x\sqrt{180\cdot 100}}{7500x\sqrt{118\cdot 112} + 10600x\sqrt{122\cdot 130} + 11000x\sqrt{180\cdot 100}}x100$$
$$= 106.008$$

Dari indeks yang diperoleh menjelaskan bahwa dibandingkan tahun 2013, pada tahun 2014 nilai agregatif komoditas pertanian meningkat sebesar 2.75 % dan pada tahun 2015 meningkat sebesar 6.008%.

BAB 8
PENGUJIAN HIPOTESIS

A. Konsep Pengujian Hipotesis

Dalam penerapan ilmu statistika inferensia, hipotesis merupakan metode yang sering digunakan. Metode ini melakukan pendugaan suatu parameter dengan menggunakan nilai statistik. Sebagai contoh, jika seorang peneliti ingin menduga nilai rata-rata mahasiswa secara keseluruhan maka dia dapat melakukan pendugaan dengan hipotesis yang sesuai dari nilai sampel yang diambil.

Pada pengujian hipotesis maka akan dilakukan penerimaan atau penolakan dari hipotesis atau dugaan yang dilakukan. Penolakan suatu dugaan awal atau yang disebut dengan hipotesis nol (H_0) menjelaskan suatu dugaan awal yang salah sedangkan penerimaan H_0 mengindikasikan bahwa peneliti tidak memiliki nilai yang pasti untuk mempercayai kebalikan H_0. Sehingga yang diharapkan oleh peneliti adalah menolak H_0. Penolakan H_0 maka menyebabkan penerimaan hipotesis alternatif yang merupakan lawan hipotesis nol. Hipotesis alternatif dilambangkan dengan H_1.

Uji beda merupakan alat uji statistik komparatif yang digunakan untuk menguji perbandingan keadaan variabel (signifikansi hasil penelitian) dari dua sampel atau lebih. Uji statistik digunakan untuk menguji hipotesis statistik dan diambil keputusan apakah hipotesis yang diajukan benar atau salah. Berikut adalah metode pengujian hipotesis berdasarkan asumsi kenormalan data:

1. Parametrik

Statistika parametrik digunakan apabila asumsi kenormalan data terpenuhi. Statistik uji yang digunakan pada uji statistik komparatif parametrik untuk dua sampel independen adalah t-test dengan formula sebagai berikut:

Separated varians:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Polled varians:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{(n_1 - n_2)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Terdapat perbedaan formula yang dapat digunakan berdasarkan jumlah sampel dan varians data dengan kriteria berikut:

a. Apabila jumlah sampel dan varians homogen maka dapat menggunakan *separated varians* atau *polled varians* dengan derajat bebas $n_1 + n_2 - 2$
b. Apabila varians homogen tetapi jumlah sampel tidak sama maka menggunakan *polled varians* dengan derajat bebas $n_1 + n_2 - 2$
c. Apabila jumlah sampel sama tetapi varians tidak homogen maka dapat menggunakan *separated varians* atau *polled varians* dengan derajat bebas $n_1 - 1$ atau $n_2 - 1$
d. Apabila jumlah sampel tidak sama dan varians berbeda menggunakan *separated varians* nilai t

tabel diperoleh dari selisih antara nilai t tabel pada derajat bebas $n_1 - 1$ dan $n_2 - 1$.

2. Nonparametrik

Statistika nonparametrik digunakan apabila asumsi kenormalan data tidak terpenuhi. Statistik uji yang digunakan pada uji statistik komparatif nonparametrik untuk dua sampel independen adalah uji *mann whitney* dengan formula sebagai berikut:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

dan

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Dengan R_1 dan R_2 berturut-turut adalah jumlah ranking pada sampel n_1 dan n_2. Untuk selanjutnya yang digunakan untuk membandingkan dengan U tabel adalah nilai U yang lebih kecil.

B. Pengujian Hipotesis dengan SPSS

1. One Sample T-Test

a. Contoh soal:

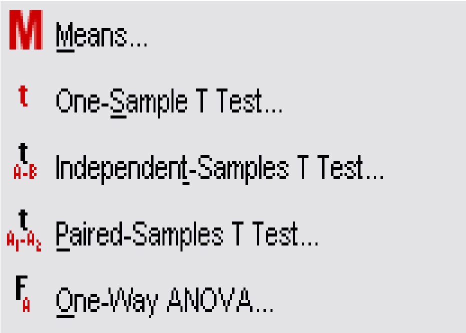
Data IPK mahasiswa yang diambil sampel 15 mahasiswa yang telah menempuh perkuliahan lebih dari 8 semester.

3,24	3,41	3,33	3,20	2,99
3,39	3,34	3,29	3,41	3,19
3,19	3,18	3,39	3,33	3,27

Dengan Taraf signifikansi 5%, dapatkah diyakini bahwa rata-rata IPK mahasiswa sebesar **3,22**?

b. Definisikan nama variabel (IPK) dan inputkan data

c. Klik **Analyze → Compare Means**



Means = pada bagian ini digunakan untuk uji linieritas dan untuk mengetahui nilia statistika deskriptif.

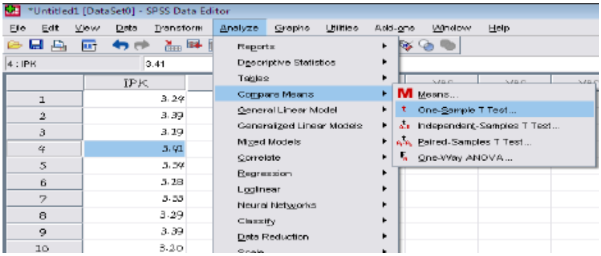
One-sample t test = untuk pengujian hipotesis parameter sampel kecil

Independent Sample T Test = pengujian hipotesis t untuk dua sampel independen

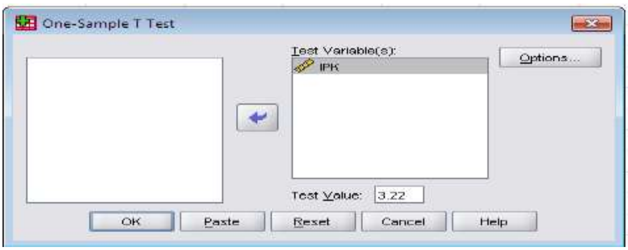
Paired Sample T Test = pengujian hipotesis t untuk data berpasangan

One-Way ANOVA = untuk melakukan pengujian lebih dari dua sampel

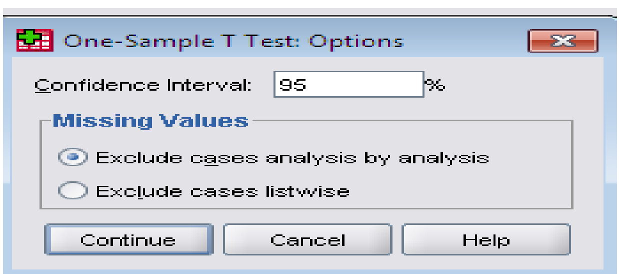
- Klik **One Sample T Test**



- Masukkan variabel pada kolom Test Variable(s) dan 3.22 pada Test Value



- Klik Option→ Continue → OK



Jika nilai taraf signifikansi 5% maka pada bagian **confidence interval** diisi 5%. Pengisian nilai pada **confidence interval** ditentukan dari nilai 100%-α

d. Output

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
IPK	15	3.2767	.11493	.02968

One-Sample Test

	Test Value = 3.22					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
IPK	1.910	14	.077	.05667	-.0070	.1203

$H_0 : \mu = 3.22$

$H_1 : \mu \neq 3.22$

$t_{hitung} = 1.910 < 2.145 = t_{0.05;14} \rightarrow H_0$ diterima

Atau $sig = 0.077 > 0.05 \rightarrow H_0$ diterima

2. Independent Sample T-Test

a. Contoh soal:

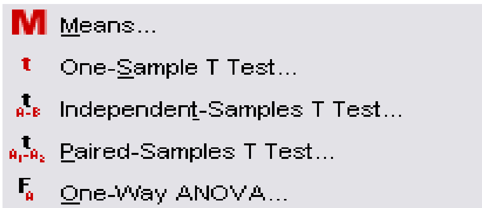
Kelas A	Nilai	Kelas B	Nilai
1	26.80	2	28.10
1	25.40	2	26.90
1	28.90	2	27.40
1	23.90	2	22.60

1	27.70	2	25.60
1	23.90		
1	24.70		

Dengan Taraf signifikansi 5% lakukan pengujian hipotesis untuk mengetahui apakah rata-rata nilai kelas A dengan kelas B berbeda atau tidak?

b. Definisikan nama variabel (Nilai dan Kelas) dan inputkan data

c. Klik **Analyze → Compare Means**



Means = pada bagian ini digunakan untuk uji linieritas dan untuk mengetahui nilia statistika deskriptif.

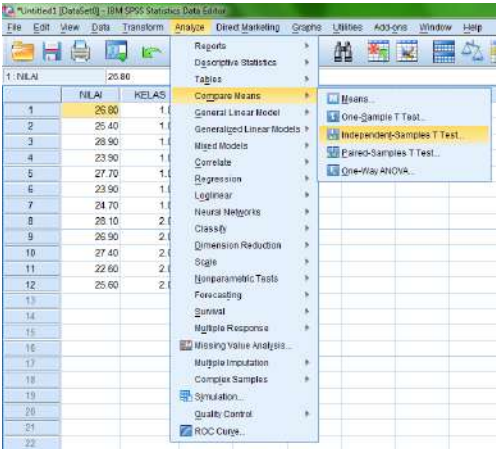
One-sample t test = untuk pengujian hipotesis parameter sampel kecil

Independent Sample T Test = pengujian hipotesis t untuk dua sampel independen

Paired Sample T Test = pengujian hipotesis t untuk data berpasangan

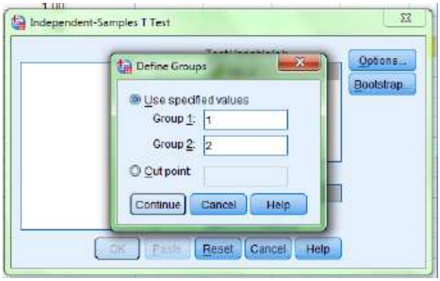
One-Way ANOVA = untuk melakukan pengujian lebih dari dua sampel

d. Klik Independent Sample T Test



Masukkan variabel **NILAI** pada kolom **Test Variable(s)** dan **KELAS** pada kolom **Grouping Variable**

e. Klik **Define Groups** → **1** untuk Group 1 dan **2** untuk Group 2 → **Continue** → **OK**



f. Output

Group Statistics

	KELAS	N	Mean	Std. Deviation	Std. Error Mean
NILAI	1.00	7	25.9000	1.94679	.73582
	2.00	5	26.1200	2.16956	.97026

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
NILAI	Equal variances assumed	.000	.983	-.184	10	.857	-.22000	1.19381	-2.87998	2.43998
	Equal variances not assumed			-.181	8.131	.861	-.22000	1.21771	-3.02019	2.58019

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$sig = 0.857 > 0.05 \rightarrow H_0 \text{ diterima}$$

3. Paired Sample T-Test

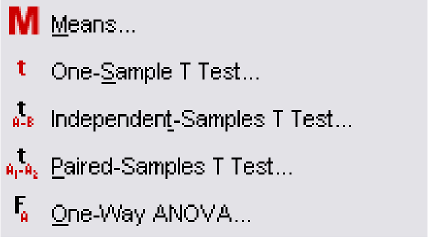
a. Contoh Soal

Pretest	Posttest
50.00	80.00
40.00	76.00
75.00	86.00
50.00	57.00
45.00	57.00
55.00	60.00
50.00	55.00
70.00	75.00
65.00	60.00
55.00	85.00
60.00	70.00
50.00	75.00
50.00	80.00
40.00	76.00

Dengan Taraf signifikansi 5% lakukan pengujian hipotesis untuk mengetahui apakah rata-rata nilai Pretest dengan Posttest berbeda atau tidak?

b. Definisikan nama variabel (Pretest dan Posttest) dan inputkan data

c. Klik Analyze → Compare Means



Means = pada bagian ini digunakan untuk uji linieritas dan untuk mengetahui nilia statistika deskriptif.

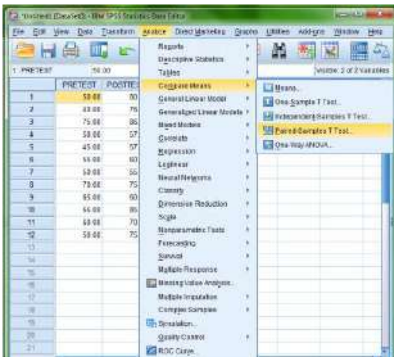
One-sample t test = untuk pengujian hipotesis parameter sampel kecil

Independent Sample T Test = pengujian hipotesis t untuk dua sampel independen

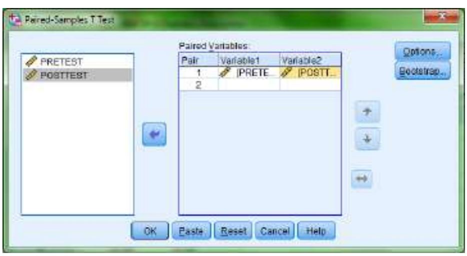
Paired Sample T Test = pengujian hipotesis t untuk data berpasangan

One-Way ANOVA = untuk melakukan pengujian lebih dari dua sampel

d. Klik **Paired Sample T Test**



e. Masukkan variabel **PRETEST** pada kolom **Variable 1** dan **POSTTEST** pada kolom **Variable 2 → OK**



f. Output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PRETEST	55.4167	12	10.32612	2.98089
	POSTTEST	69.6667	12	11.39644	3.28987

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	PRETEST & POSTTEST	12	.310	.326

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PRETEST - POSTTEST	-14.25000	12.78582	3.69095	-22.37372	-6.12628	-.861	11	.003

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$
$$sig = 0.003 < 0.05 \rightarrow H_0 \text{ ditolak}$$

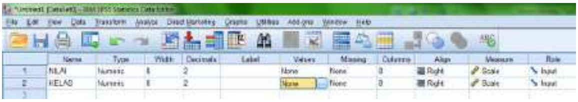
4. Mann Whitney-Test

a. Contoh soal:

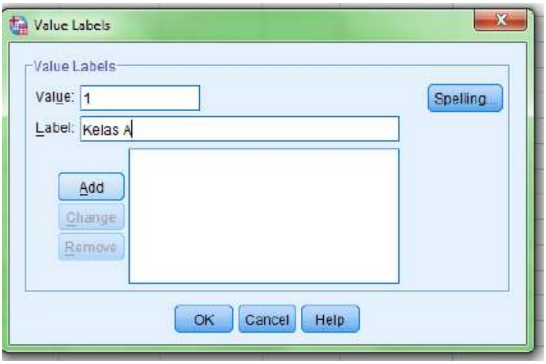
Kelas A	Nilai	Kelas B	Nilai
1	50.00	2	60.00
1	40.00	2	50.00
1	75.00	2	80.00
1	50.00	2	76.00
1	45.00	2	86.00
1	55.00	2	57.00
1	50.00	2	57.00
1	70.00	2	60.00
1	65.00	2	55.00
1	55.00	2	75.00

Dengan Taraf signifikansi 5% lakukan pengujian hipotesis untuk mengetahui apakah rata-rata nilai kelas A dengan kelas B berbeda atau tidak?

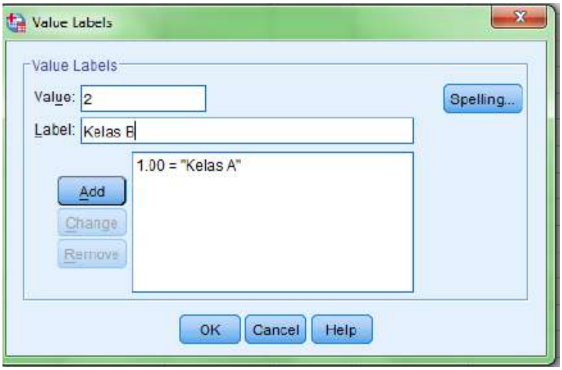
b. Definisikan nama variabel (Nilai dan Kelas) dan inputkan data
c. Pada variabel **KELAS** definisikan kelas yang dimaksud missal Kelas A dan Kelas B dengan klik Kotak Titik Tiga pada kolom Values Baris kelas



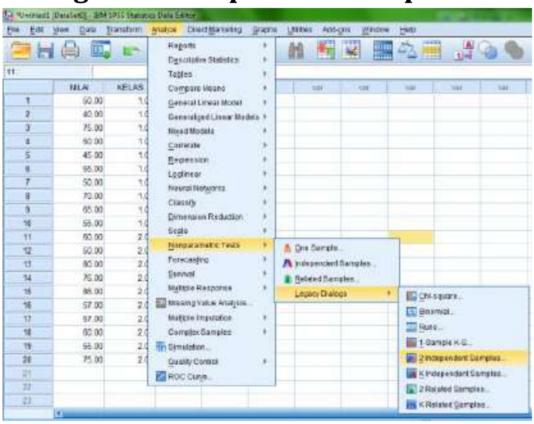
d. Masukkan 1 pada Value dan Kelas A pada Label → **klik Add**



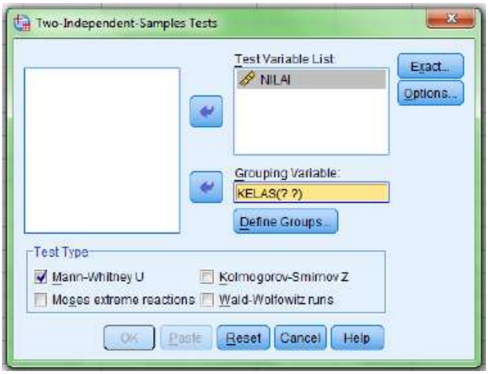
e. Masukkan 2 pada Value dan Kelas B pada Label → **klik Add → OK**



f. Klik **Analyze → Nonparametric Tests → Legacy Dialogs → 2 Independent Samples**



g. Masukkan variabel **NILAI** ke kolom **Test Variable List** dan **KELAS** ke kolom **Grouping Variable**



h. Centang **Mann Whitney U** lalu klik **Define Group** → masukkan 1 ke Group 1 dan 2 ke Group 2 → **Continue → OK**

i. Output

Ranks

	KELAS	N	Mean Rank	Sum of Ranks
NILAI	Kelas A	10	7.80	78.00
	Kelas B	10	13.20	132.00
	Total	20		

Test Statistics[a]

	NILAI
Mann-Whitney U	23.000
Wilcoxon W	78.000
Z	-2.054
Asymp. Sig. (2-tailed)	.040
Exact Sig. [2*(1-tailed Sig.)]	.043[b]

a. Grouping Variable: KELAS

b. Not corrected for ties.

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$sig = 0.04 < 0.05 \rightarrow H_0$ ditolak

BAB 9

ANALISIS REGRESI NONLINIER (METODE *GAUSS NEWTON* DALAM ESTIMASI PARAMETER REGRESI NONLINIER)

A. Regresi Nonlinier

Regresi nonlinier adalah analisis regresi dimana hubungan antara variabel terikat dan variabel bebas tidak memenuhi asumsi linieritas. Bentuk hubungan regresi nonlinier yang dikenal umum dan banyak digunakan adalah sebagai berikut:

1. Bentuk polinomial

    a) Polinomial pangkat k

    $y = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k$

    b) Polinomial pangkat 3 atau bentuk kubik

    $y = b_0 + b_1 x + b_2 x^2 + b_3 x^3$

    c) Polinomial pangkat dua atau bentuk parabola

    $y = b_0 + b_1 x + b_2 x^2$

2. Bentuk khusus

Bentuk khusus ini antara lain eksponen, eksponen pertumbuhan, geometri, power, *compound*, *sigmoid*, logistik, dan lain-lain, dimana setiap model dengan melakukan transformasi menjadi bentuk linier, maka dengan metode kuadrat terkecil koefisien-koefisien dari model regresi nonlinier dapat ditentukan.

Regresi nonlinier adalah suatu bentuk regresi yang memuat parameter nonlinier Parameter ini akan tetap mengandung parameter itu sendiri jika diturunkan terhadap parameter itu. Penentuan nilai parameter pada model regresi nonlinier dilakukan dengan meminimumkan jumlah kuadrat galat yang dalam penyelesaiannya dengan cara iterasi yang bergantung pada nilai awal. Model regresi nonlinier secara umum dapat dituliskan sebagai berikut:

$$Y = f\left(X_1, X_2, \ldots, X_k; \theta_1, \theta_2, \ldots, \theta_p\right) + \varepsilon$$

dengan:

Y : peubah respon

X : peubah bebas

θ : parameter

ε : galat

yang selanjutnya dapat diringkas menjadi:

$Y = f(X, \theta) + \varepsilon$ atau $E(Y) = f(X, \theta)$,

dengan asumsi bahwa $E(\varepsilon) = 0$, galat-galatnya tidak berkorelasi ($V(\varepsilon) = \sigma^2$), dan galat-galatnya berdistribusi normal serta saling bebas satu sama lain.

Jumlah kuadrat galat untuk model nonlinier didefinisikan sebagai berikut:

$S(\theta) = \sum_{i=1}^{n} \left[Y_i - f(X_i, \theta)\right]^2$ (Draper dan Smith, 1966).

B. Deret Taylor

Misalkan fungsi *f* adalah fungsi dari dua variabel X dan Y. Deret *Taylor* dapat digunakan untuk mencari nilai suatu fungsi di titik x dan y jika nilai pada titik x_0 dan y_0 diketahui. Ekspansi dari f(x,y) adalah sebagai berikut:

$$f(x,y) = f(x_0,y_0) + \left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right] f(x,y)\Big|_{(x_0,y_0)}$$
$$+ \frac{1}{4!}\left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right]^4 f(x,y)\Big|_{(x_0,y_0)}$$
$$+ \cdots + \frac{1}{n!}\left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right]^n f(x,y)\Big|_{(x_0,y_0)}$$
$$+ R_n(x),$$ (Spiegel dan Wrede, 2006 dalam Syilfi, Ispriyanti dan Safitri, 2012).

C. Metode Gauss Newton

Metode *Gauss Newton* diawali dengan inisialisasi nilai awal untuk parameter regresi yang akan diestimasi, yaitu parameter $\theta_0, \theta_1, \cdots, \theta_{p-1}$ yang selanjutnya dalam proses penaksiran diubah menjadi $g_0^{(0)}, g_1^{(0)}, \cdots, g_{p-1}^{(0)}$. Nilai-nilai awal yang diinisialisasikan merupakan nilai yang dimungkinkan berdasarkan dugaan kasar atau dimungkinkan pula berdasarkan informasi yang tersedia. Nilai-nilai awal tersebut nantinya akan digantikan dengan nilai yang baru dari proses iterasi.

Metode *Gauss Newton* merupakan suatu metode yang meminimumkan jumlah kuadrat galat dengan menggunakan ekspansi deret *Taylor*. Ekspansi deret *Taylor* digunakan untuk mjenyatakan persamaan nonlinier awal ke dalam bentuk hampiran yang linier. Misalkan suatu model dengan bentuk $Y_u = f(\xi,\theta) + \varepsilon_u$ dan $\theta_{10}, \theta_{20}, \ldots, \theta_{p0}$ adalah inisialisasi nilai awal untuk parameter $\theta_0, \theta_1, \ldots, \theta_p$. Ekspansi deret Taylor $f(\xi,\theta)$ di sekitar titik $\theta_0 = (\theta_{10}, \theta_{20}, \theta_{p0})$ sampai turunan pertama untuk θ dekat dengan θ_0 adalah:

$$f(\xi_u,\theta) = f(\xi_u,\theta_0) + \sum_{i=1}^{p}\left[\frac{\partial f(\xi_u,\theta)}{\partial \theta_i}\right]_{\theta=\hat{\theta}}(\theta_i - \theta_{i0})$$

Jika,

$$f_u^0 = f(\xi_u,\theta_0)$$

$$\beta_i^0 = \theta_i - \theta_{i0}$$

$$Z_{iu}^0 = \left[\frac{\partial f(\xi_u,\theta)}{\partial \theta_i}\right]_{\theta=\hat{\theta}}$$

maka akan menjadi bentuk linier sebagai berikut:

$$Y_u - f_u^0 \sum_{i-1}^{p} \beta_i^0 Z_{iu}^0 + \varepsilon_u$$

Sehingga parameter $\beta_i^0, i = 1,2,\ldots,p$ dapat diestimasi dengan metode kuadrat terkecil dengan menetapkan:

$$Z = \begin{bmatrix} Z_{11}^{0} & Z_{21}^{0} & \dots & Z_{p1}^{0} \\ Z_{12}^{0} & Z_{22}^{0} & \dots & Z_{p2}^{0} \\ \vdots & \vdots & \vdots & \vdots \\ Z_{1u}^{0} & Z_{2u}^{0} & \dots & Z_{pu}^{0} \\ \vdots & & & \\ Z_{1n}^{0} & Z_{2n}^{0} & \dots & Z_{pn}^{0} \end{bmatrix} = \{Z_{iu}^{0}, p \times n\}$$

$$y_0 = \begin{bmatrix} Y_1 - f_1^{0} \\ Y_2 - f_2^{0} \\ \vdots \\ Y_u - f_u^{0} \\ \vdots \\ Y_n - f_n^{0} \end{bmatrix} = Y - f^{0}$$

$$b_0 = \begin{bmatrix} b_1^{0} \\ b_2^{0} \\ \vdots \\ b_p^{0} \end{bmatrix}$$

Estimasi parameter $\beta_0 = \left(\beta_1^0, \beta_2^0, \dots, \beta_p^0\right)$ adalah $b_0 = \left(Z_0^{'} Z_0\right)^{-1} Z_0^{'} \left(Y - f^{0}\right)$ dengan meminimumkan jumlah kuadrat galat.

$$S(\theta) = \sum_{i=1}^{n} \{Y_i - f(\xi, \theta)\}^2$$

$$S(\theta)=[Y-f(\theta)][Y-f(\theta)]$$
$$f(\theta)=[f(\xi_1,\theta), f(\xi_2,\theta),..., f(\xi_n,\theta)] \quad Y=[Y_1,Y_2,...,Y_n]$$

Persamaan $S(\theta)=[Y-f(\theta)][Y-f(\theta)]$ diturunkan terhadap θ yang menghasilkan persamaan normal:

$$\sum_{i=1}^{n} Y_i\left[\frac{\partial f(\xi_i,\theta)}{\partial\theta}\right]_{\theta=\hat{\theta}} - \sum_{i=1}^{n} f(\xi,\theta)\left[\frac{\partial f(\xi_i,\theta)}{\partial\theta}\right]_{\theta=\hat{\theta}}$$

Iterasi metode *Gauss Newton* adalah sebagai berikut:

$$\hat{\theta}^{n+1}=\theta^n+\left(D(\theta^n)' D(\theta^n)\right)^{-1} D(\theta^n)'(Y_t-f(\xi,\theta)),$$

dengan:

θ^n : Nilai dugaan awal parameter

$\hat{\theta}^{n+1}$: Parameter yang akan diestimasi

$D(\theta^n)$: Matrik yang dihasilkan dari data

$Y_t-f(\xi,\theta)$: Vektor yang dihasilkan dari perbedaan antara pengukuran dan prediksi

Langkah-langkah proses iterasi metode *Gauss Newton* adalah sebagai berikut:

1) Inisialisasi $\hat{\theta}^{(0)}$ sebagai estimasi awal untuk θ,
2) Hitung $\hat{\theta}^{(i+1)}=\hat{\theta}^{(0)}+b_i$,
3) Nilai $\hat{\theta}^{(i+1)}$ digunakan sebagai nilai untuk menghampiri model linier,
4) Selanjutnya kembali ke langkah 1) dan menghitung nilai *b* untuk setiap iterasi, nilai *b* yang baru

ditambahkan ke penaksiran yang diperoleh dari iterasi sebelumnya,

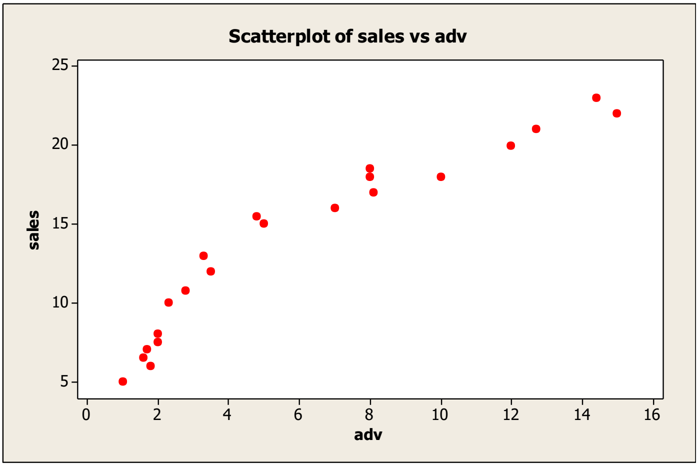
5) Iterasi dilanjutkan sampai diperoleh hasil penaksiran yang konvergen (Karim, 2009 dalam Nainggolan, 2010).

D. Contoh Kasus

Data

sales	adv		sales	adv		sales	adv
5.0	1.0		10.8	2.8		18.0	8.0
6.0	1.8		12.0	3.5		18.0	10.0
6.5	1.6		13.0	3.3		18.5	8.0
7.0	1.7		15.5	4.8		21.0	12.7
7.5	2.0		15.0	5.0		20.0	12.0
8.0	2.0		16.0	7.0		22.0	15.0
10.0	2.3		17.0	8.1		23.0	14.4

Scatterplot:



Hasil Regresi Nonlinier dengan Bentuk Kurva *Michaelis Menten*:

Nonlinear Regression: sales = Theta1 * adv / (Theta2 + adv)

```
Method
Algorithm        Gauss-Newton
Max iterations            200
Tolerance             0.00001
Starting Values for Parameters
Parameter  Value
Theta1         1
Theta2         1
Estimates at Each Iteration
Iteration      SSE   Theta1   Theta2
        0  4184.67   1.0000   1.0000
        1  2654.88  24.6310  38.0666
        2  1765.77  13.6478   7.7595
        3   180.80  20.7939   3.4452
        4    17.44  28.7633   5.1314
        5    15.24  29.4137   5.1359
        6    15.24  29.4137   5.1357
        7    15.24  29.4137   5.1357

Equation
sales = 29.4137 * adv / (5.13574 + adv)
Parameter Estimates

Parameter  Estimate  SE Estimate
Theta1      29.4137      1.13668
Theta2       5.1357      0.46402

sales = Theta1 * adv / (Theta2 + adv)
Lack of Fit
Source         DF        SS         MS      F      P
```

```
Error           19  15.2415  0.802184
  Lack of Fit  17  14.9915  0.881852  7.05  0.131
  Pure Error    2   0.2500  0.125000

Summary
Iterations          7
Final SSE     15.2415
DFE                19
MSE          0.802184
S            0.895647
```

Iterasi dengan Metode Gauss Newton:

$$f\left(X_i;\theta_1,\theta_2\right)=\frac{\theta_1 X_i}{\theta_2+X_i}$$

Fungsi di atas merupakan bentuk regresi nonlinier dari $Y_i = f\left(X_i,\theta\right)+\varepsilon$ dengan kuadrat terkecil *Q* terhadap θ_k :

$$Q=\sum_{i=1}^{n}\left[Y_i - f\left(X_i,\theta_k\right)\right]^2 \quad ; k=1,2,\ldots,p-1$$

Turunan parsial dari *Q* terhadap θ_k adalah:

$$\frac{\partial Q}{\partial \theta_k}=\sum_{i=1}^{n}-2\left[Y_i - f\left(X_i,\theta\right)\right]\left[\frac{\partial f\left(X_i,\theta\right)}{\partial \theta_k}\right]_{\hat{\theta}=g}=0$$

Sehingga

$$f\left(X_i;\theta_1,\theta_2\right)=\frac{\theta_1 X_i}{\theta_2+X_i}$$

dapat diturunkan terhadap parameter-parameternya sebagai berikut:

$$\frac{\partial f\left(X_i,\theta\right)}{\partial \theta_1}=\frac{X_i}{\theta_2+X_i}$$

$$\frac{\partial f\left(X_i,\theta\right)}{\partial \theta_2}=\frac{-\theta_1 X_i}{\left(\theta_2+X_i\right)^2}$$

Iterasi 1:

Nilai awal untuk $\theta_1 = 1 \,\text{dan}\, \theta_2 = 1$.

Hitung $\text{sales*} = \dfrac{1 * \text{adv}}{1+\text{adv}}$ kemudian hitung residual (sales-sales*) dan SSE nya. Selanjutnya hitung turunan untuk setiap parameter dengan memasukkan nilai awal, diperoleh:

sales	adv	yhat	resi	sse	df/d(teta1)	df/d(teta2)	
5.0	1.0	0.500000	45.000	4184.67	0.166667		-0.250000
6.0	1.8	0.642857	53.571		0.257143		-0.229592
6.5	1.6	0.615385	58.846		0.213333		-0.236686
7.0	1.7	0.629630	63.704		0.212500		-0.233196
7.5	2.0	0.666667	68.333		0.235294		-0.222222
8.0	2.0	0.666667	73.333		0.222222		-0.222222
10.0	2.3	0.696970	93.030		0.209091		-0.211203
10.8	2.8	0.736842	100.632		0.237288		-0.193906
12.0	3.5	0.777778	112.222		0.269231		-0.172840
13.0	3.3	0.767442	122.326		0.235714		-0.178475
15.5	4.8	0.827586	146.724		0.290909		-0.142687
15.0	5.0	0.833333	141.667		0.312500		-0.138889
16.0	7.0	0.875000	151.250		0.411765		-0.109375
17.0	8.1	0.890110	161.099		0.450000		-0.097814
18.0	8.0	0.888889	171.111		0.421053		-0.098765
18.0	10.0	0.909091	170.909		0.526316		-0.082645

18.5	8.0	0.888889	176.111	0.410256	-0.098765
21.0	12.7	0.927007	200.730	0.577273	-0.067665
20.0	12.0	0.923077	190.769	0.571429	-0.071006
22.0	15.0	0.937500	210.625	0.652174	-0.058594
23.0	14.4	0.935065	220.649	0.600000	-0.060719

$$\hat{\theta}^{n+1} = \theta^n + \left(D(\theta^n)' D(\theta^n)\right)^{-1} D(\theta^n)' (Y_t - f(\xi, \theta)),$$

Buat matriks **Y** dan **D** dari sales* dan hasil dari turunan terhadap parameternya:

Y	D	
0.5	0.166667	-0.25
0.642857	0.257143	-0.22959
0.615385	0.213333	-0.23669
0.62963	0.2125	-0.2332
0.666667	0.235294	-0.22222
0.666667	0.222222	-0.22222
0.69697	0.209091	-0.2112
0.736842	0.237288	-0.19391
0.777778	0.269231	-0.17284
0.767442	0.235714	-0.17847
0.827586	0.290909	-0.14269
0.833333	0.3125	-0.13889
0.875	0.411765	-0.10938
0.89011	0.45	-0.09781
0.888889	0.421053	-0.09877
0.909091	0.526316	-0.08264
0.888889	0.410256	-0.09877
0.927007	0.577273	-0.06766
0.923077	0.571429	-0.07101

0.9375	0.652174	-0.05859
0.935065	0.6	-0.06072

$$[y - f(\xi, \theta^{(0)})] = \begin{bmatrix} y_1 - f_1^{(0)} \\ y_2 - f_2^{(0)} \\ \vdots \\ y_{21} - f_{21}^{(0)} \\ \end{bmatrix}$$

$$= \begin{bmatrix} 4.5 \\ 5.35714 \\ \vdots \\ 22.0649 \\ \end{bmatrix}$$

$$SSE^{(0)} = \sum_{t=1}^{T} [y_t - f(X_t, \theta^{(0)})]^2 = \sum_{t=1}^{T} [y_t - f_t^{(0)}]^2$$

$$= 4.5^2 + 5{,}35714^2 + \ldots + 22{,}0649^2$$

$$= 4184.67$$

$$D\theta^{(0)} = \begin{bmatrix} \dfrac{\xi_1}{\theta_1^{(0)} + \xi_1} & -\theta_0^{(0)} \dfrac{\xi_1}{(\theta_1^{(0)} + \xi_1)^2} \\ \dfrac{\xi_2}{\theta_1^{(0)} + \xi_2} & -\theta_0^{(0)} \dfrac{\xi_2}{(\theta_1^{(0)} + \xi_2)^2} \\ \dfrac{\xi_3}{\theta_1^{(0)} + \xi_3} & -\theta_0^{(0)} \dfrac{\xi_3}{(\theta_1^{(0)} + \xi_3)^2} \\ \vdots & \vdots \\ \dfrac{\xi_{21}}{\theta_1^{(0)} + \xi_{21}} & -\theta_0^{(0)} \dfrac{\xi_{21}}{(\theta_1^{(0)} + \xi_{21})^2} \end{bmatrix} = \begin{bmatrix} 0.5 & -0.25 \\ 0.64286 & -0.22959 \\ 0.61538 & -0.23669 \\ \vdots & \vdots \\ 0.93506 & -0.06072 \end{bmatrix}$$

$$D\theta^{(0)}{}' D\theta^{(0)}$$

$$= \begin{bmatrix} 0.5 & 0.64286 & 0.61538 & \dots & 0.93506 \\ -0.25 & -0.22959 & -0.23669 & \dots & -0.06072 \end{bmatrix} \begin{bmatrix} 0.5 & -0.25 \\ 0.64286 & -0.22959 \\ 0.61538 & -0.23669 \\ \vdots & \vdots \\ 0.93506 & -0.06072 \end{bmatrix}$$

$$= \begin{bmatrix} 13.3585 & -2.3301 \\ -2.3301 & 0.5726 \end{bmatrix}$$

$$[D\theta^{(0)}{}' D\theta^{(0)}]^{-1} = \begin{bmatrix} 0.2580 & 1.0500 \\ 1.0500 & 6.0196 \end{bmatrix}$$

$$[D\theta^{(0)}{}' D\theta^{(0)}]^{-1} Z = D\theta^{(0)}{}'[y - f(\xi, \theta^{(0)})] = \begin{bmatrix} 23.6310 \\ 37.0667 \end{bmatrix}$$

$$\theta^{(1)} = \theta^{(0)} + [D\theta^{(0)}{}' D\theta^{(0)}]^{-1} D\theta^{(0)}{}'[y - f(\xi, \theta^{(0)})]$$

$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 23.6310 \\ 37.0667 \end{bmatrix} = \begin{bmatrix} 24.6310 \\ 38.0667 \end{bmatrix}$$

Iterasi 2:

$\theta_1^{(1)} = 24.6310$ dan $\theta_2^{(1)} = 38.0667$.

$$[y - f(\xi, \theta^{(1)})] = \begin{bmatrix} y_1 - f_1^{(1)} \\ y_2 - f_2^{(1)} \\ \vdots \\ y_{21} - f_{21}^{(1)} \end{bmatrix} = \begin{bmatrix} 4,36951 \\ 4,8879 \\ \vdots \\ 16,2398 \end{bmatrix}$$

$$SSE^{(1)} = \sum_{t=1}^{T} [y_t - f(\xi_t, \theta^{(1)})]^2 = \sum_{t=1}^{T} [y_t - f_t^{(1)}]^2$$
$$= 4.36951^2 + 4,8879^2 + \ldots + 16,2398^2 = 2654,88$$

$$D\theta^{(1)} = \begin{bmatrix} \frac{\xi_1}{\theta_1^{(1)} + \xi_1} & -\theta_0^{(1)} \frac{\xi_1}{(\theta_1^{(1)} + \xi_1)^2} \\ \frac{\xi_2}{\theta_1^{(1)} + \xi_2} & -\theta_0^{(1)} \frac{\xi_2}{(\theta_1^{(1)} + \xi_2)^2} \\ \frac{\xi_3}{\theta_1^{(1)} + \xi_3} & -\theta_0^{(1)} \frac{\xi_3}{(\theta_1^{(1)} + \xi_3)^2} \\ \vdots & \vdots \\ \frac{\xi_{21}}{\theta_1^{(1)} + \xi_{21}} & -\theta_0^{(1)} \frac{\xi_{21}}{(\theta_1^{(1)} + \xi_{21})^2} \end{bmatrix} = \begin{bmatrix} 0.0256 & -0.01614 \\ 0.04515 & -0.0279 \\ 0.04034 & -0.02505 \\ \vdots & \vdots \\ 0.27446 & -0.12885 \end{bmatrix}$$

$$D\theta^{(1)}{}' D\theta^{(1)} = \begin{bmatrix} 0,4923 & -0,24971 \\ -0,2497 & 0.1275 \end{bmatrix}$$

$$[D\theta^{(1)}{}' D\theta^{(1)}]^{-1} = \begin{bmatrix} 310,9 & 609 \\ 609 & 1200,6 \end{bmatrix}$$

$$[D\theta^{(1)}{}' D\theta^{(1)}]^{-1} D\theta^{(1)}{}'[y - f(\xi, \theta^{(1)})] = \begin{bmatrix} -175,7405 \\ -484,9322 \end{bmatrix}$$

$$\theta^{(2)} = \theta^{(1)} + [D\theta^{(1)}{}' D\theta^{(1)}]^{-1} D\theta^{(1)}{}'[y - f(\xi, \theta^{(1)})]$$
$$= \begin{bmatrix} 24,631 \\ 38,0667 \end{bmatrix} + \begin{bmatrix} -175,7405 \\ -484,9322 \end{bmatrix} = \begin{bmatrix} -151,1095 \\ -446,8655 \end{bmatrix}$$

BAB 10

ANALISIS REGRESI NONLINIER MODEL SIGMOID

A. Model Regresi Sigmoid

Fungsi sigmoid adalah fungsi yang menghasilkan kurva sigmoid yaitu kurva yang memiliki bentuk seperti huruf "S". Model dari fungsi sigmoid adalah $y = e^{a+\frac{b}{x}}$ sedangkan bentuk liniernya adalah $\ln y = a + \frac{b}{x}$. Dengan a dan b merupakan parameter yang harus diduga dari data. Sehingga akan didapatkan model linearnya, yaitu:

$$y = e^{a+\frac{b}{x}}$$

$$\ln y = \ln e^{a+\frac{b}{x}}$$

$$In\, y = a + \frac{b}{x}$$

Misalkan a adalah γ dan b adalah δ sehingga model regresi menjadi

$$y = e^{\gamma+\frac{\delta}{x}}$$

sehingga didapatkan kurva regresi model linier $\ln y = \gamma + \frac{\delta}{x}$, dan setiap data memenuhi hubungan:

$$\ln y_i = \gamma + \frac{\delta}{x} + e_i$$

$$\Leftrightarrow \ln y_i = a + \frac{b}{x_i} + e_i$$

Model regresi linier dari model regresi sigmoid didapat dengan cara:

$$y = e^{a+\frac{b}{x}}$$

$$\ln y = \ln e^{a+\frac{b}{x}}$$

$$In\ y = a + \frac{b}{x}$$

$y *= a + bx *$ adalah regresi linier yang terhadap (In y,$\frac{1}{x}$)

Pengujian untuk model Regresi digunakan uji F dalam ANOVA. Pengujian dilakukan dengan memperhatikan nilai F-hitung yang diperoleh dibandingkan dengan F-tabel dengan derajat bebas $(1, n-2)$ pada taraf kepercayaan $\alpha = 0.05$. Pada hasil di minitab atau SPSS disediakan nilai signifikansi F. Apabila nilai signifikasi F ini lebih kecil dari 0,05 (dalam hal ini taraf yang digunakan $\alpha = 0,05$), maka dapat disimpulkan F-hitung > F-tabel dan sebaliknya. Sedangkan untuk pengujian koefisien regresi linear digunakan uji T.

Pengujian hipotesisnya sebagai berikut :

1. H_0 = Model Regresi tidak berarti

   H_1 = Model Regresi berarti

2. $H_0 = b_1 = 0$

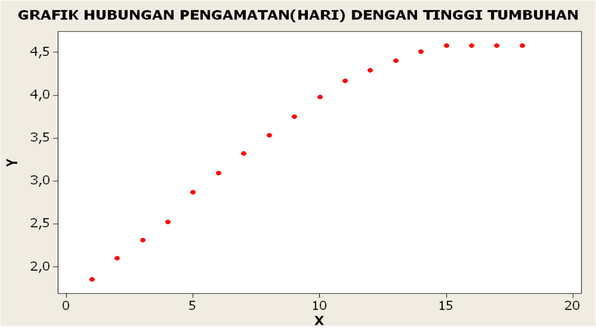
   $H_1 = b_1 \neq 0, b_1 > 0, b_1 < 0$

B. Contoh Kasus

Data diperoleh dari pengamatan pertumbuhan biji kacang tunggak yang diamati sebanyak 18 kali pengamatan dengan data hasil pengamatan sebagai berikut:

Pengamatan ke-(hari)	Tinggi Batang(cm)	Pengamatan ke-(hari)	Tinggi Batang(cm)
1	1,85	10	3,98
2	2,10	11	4,17
3	2,31	12	4,29
4	2,52	13	4,40
5	2,87	14	4,51
6	3,09	15	4,58
7	3,32	16	4,58
8	3,53	17	4,58
9	3,75	18	4,58

Scatter plot:



Keterangan :

X menyatakan pengamatan ke-(hari)

Y menyatakan panjang batang (cm)

Setelah dilakukan pengujian terhadap data di atas diperoleh persamaan regresi nonlinier bentuk sigmoid, oleh karena itu akan ditransformasikan dalam bentuk linier, yang ditampilkan dalam bentuk tabel di bawah ini.

X	Y	1/X=X*	In Y=Y*	X*Y*	X*X*
1	1,85	10,000	0,6152	0,6152	10,000
2	2,10	0,5000	0,7419	0,3710	0,2500
3	2,31	0,3333	0,8372	0,2791	0,1111
4	2,52	0,2500	0,9243	0,2311	0,0625
5	2,87	0,2000	10,543	0,2109	0,0400
6	3,09	0,1667	11,282	0,1880	0,0278
7	3,32	0,1429	12,000	0,1714	0,0204
8	3,53	0,1250	12,613	0,1577	0,0156
9	3,75	0,1111	13,218	0,1469	0,0123
10	3,98	0,1000	13,813	0,1381	0,0100
11	4,17	0,0909	14,279	0,1298	0,0083
12	4,29	0,0833	14,563	0,1214	0,0069
13	4,40	0,0769	14,816	0,1140	0,0059
14	4,51	0,0714	15,063	0,1076	0,0051
15	4,58	0,0667	15,217	0,1014	0,0044
16	4,58	0,0625	15,217	0,0951	0,0039
17	4,58	0,0588	15,217	0,0895	0,0035
18	4,58	0,0556	15,217	0,0845	0,0031
Σ		34,951	224,243	33,526	15,909

- **Dengan perhitungan manual diperoleh**

$$b = \frac{n\sum X_i^* \sum Y_i^* - (\sum X_i^*)(\sum Y_i^*)}{n(\sum X_i^{*2}) - (\sum X_i^*)^2}$$

$$b = \frac{18(3{,}3526) - (3{,}4951)(22{,}4243)}{18(1{,}5909) - (3{,}4951)^2}$$

$$b = \frac{60{,}3468 - 78{,}3751}{28{,}6362 - 12{,}2157}$$

$$b = \frac{-18{,}0283}{16{,}4205} = -1{,}0978$$

$$\overline{Y^*} = \frac{22{,}4243}{18} = 1{,}2458\ dan\ \overline{X^*} = \frac{3{,}4951}{18}$$

$$= 0{,}1942$$

$$a = \overline{Y^*} - b\overline{X^*}$$

$$a = 1{,}2458 - (-1{,}0978)(0{,}1942) = 1{,}4227$$

Sehingga persamaan yang diperoleh adalah $Y^* = 1{,}4227 - 1{,}0978X^*$ dimana $Y^* = \ln Y\ \ dan\ X^* = \frac{1}{X}$.

- **Pengolahan menggunakan minitab**

Regression Analysis: Y* versus X*

The regression equation is

Y* = 1,459 - 1,0979 X*

Predictor	Coef	SE Coef	T	P
Constant	1,45899	0,04949	29,48	0,000
X*	-1,0979	0,1665	-6,59	0,000

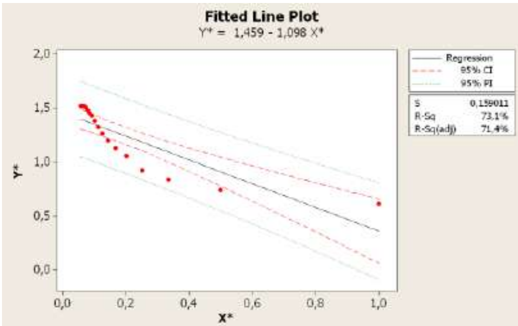
S = 0,159011   R-Sq = 73,1%   R-Sq(adj) = 71,4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1,0997	1,0997	43,49	0,000

Residual Error	16	0,4046	0,0253
Total	17	1,5042	

Persamaan yang diperoleh dari minitab adalah $Y^* = 1{,}459 - 1{,}0979X^*$dimana $Y^* = \ln Y \ \ dan\ X^* = \frac{1}{X}$.



1. Analisis data dari minitab

   a. Melihat taksiran parameter

   Dari hasil perhitungan minitab didapatkan persamaan lnY=1,459-1,0979(1/X). Persamaan ini memperlihatkan taksiran intersep b_0 sebesar 1,459 dan taksiran parameter b_1 sebesar -1,0979.

   b. Memeriksa mean square

   R-Sq atau koefisien determinasi menyatakan seberapa besar keragaman variable X* mempengaruhi Y*. Berdasarkan perhitungan minitab diperoleh R-Sq sebesar 73,1 % dalam model Y*=1,459-1,0979X* sisanya sebesar 26,9% dipengaruhi oleh variabel lain yang tidak masuk dalam model. R-Sq berkisar antara 0 sampai 1, dengan catatan semakn kecil nilai R-Sq, semakin lemah hubungan antara kedua variabel(begitu juga sebaliknya).

2. Pengujian koefisien regresi

   Hipotesis :

   $H_0: b_1 = 0$ artinya tidak ada pengaruh waktu dalam hari terhadap pertumbuhan batang.

   $H_1 : b_1 \neq 0$ artinya ada pengaruh waktu dalam hari terhadap pertumbuhan batang.

   Menggunakan uji T:

   T_{tabel} dengan $\alpha = 0.05$ diperoleh hasil 2,1098.

   T_{hitung} dari hasil minitab sebesar 29,48.

   Karena $T_{hit}>T_{tabel}$ sehingga menolak H_0. Hal ini berarti ada pengaruh lama hari terhadap pertumbuhan batang.

3. Pengujian model regresi

   Hipotesis:

   H_0: model yang diperoleh tidak bermakna.

   H_1: model yang diperoleh bermakna.

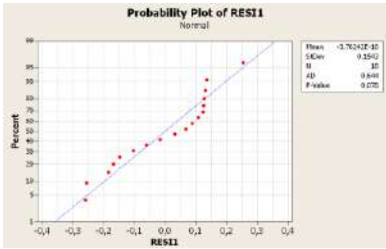
   Menggunakan uji F:

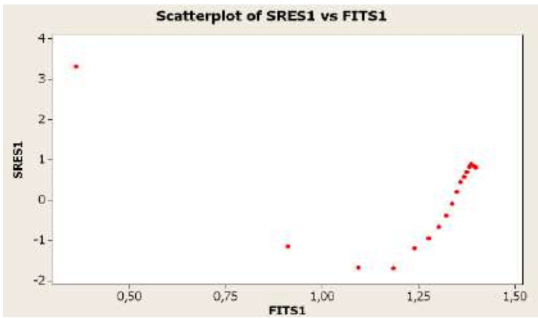
   F_{tabel} dengan derajat bebas (1,17) dengan $\alpha = 0.05$ sebesar 4,4513. sedangkan F hitung dari minitab 67.91. Karena $F_{hit}>F_{tabel}$ maka menolak H_0 dengan kata lain model yang diperoleh bermakna.

4. Pengujian asumsi

   a. Uji normalitas

Dari minitab diperoleh:



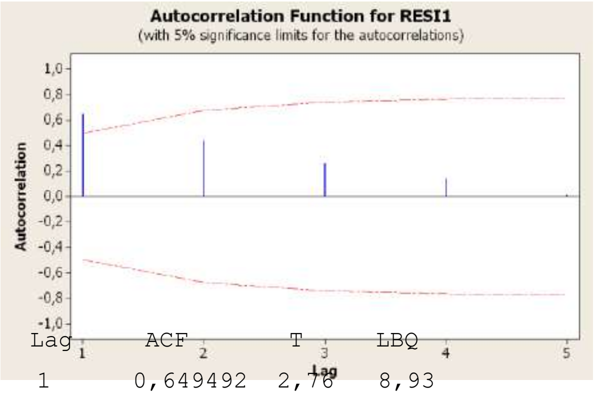


Karena P-value > α maka memenuhi asumsi kenormalan sisaan. Selain itu, dari gambar diatas dapat kita lihat bahwa plot yang terbentuk membentuk suatu garis lurus 45^0, maka dapat dikatakan sisaan mengikuti sebaran normal.

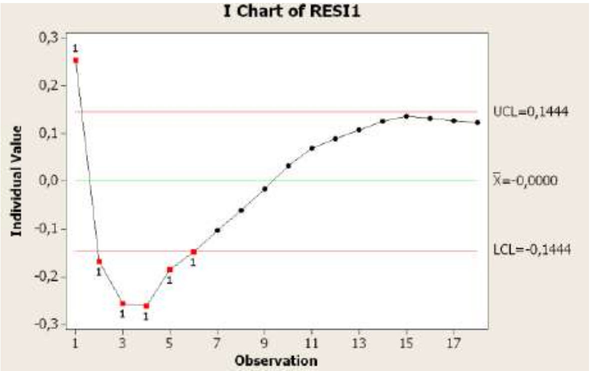
b. Uji homogenitas

Dari minitab diperoleh scatterplot hubungan antara sres1 dengan fits1. Berdasarkan gambar diketahui bahwa standart sisa 95% berada antara (-2,4) secara tidak merata. Dengan kata lain sisa dikatakan tidak berada dalam sebaran sehingga keragamannya tidak tetap.

c. Uji kebebasan



Lag	ACF	T	LBQ
1	0,649492	2,76	8,93
2	0,437403	1,37	13,24
3	0,259666	0,74	14,86
4	0,143239	0,40	15,38
5	0,008420	0,02	15,39



Dari gambar di atas dapat disimpulkan bahwa adanya hubungan antara sisaan. Hal ini dikarenakan ada garis atau data yang keluar dari garis yang berwarna merah. Jadi sisaan tidak bebas dengan kata lain saling berhubungan.

d. Analisis data dari SPSS

Model Summary

	Square	Adjusted R Square	Std. Error of the Estimate
855	731	.714	. 159

The independent variable is X.

ANOVA

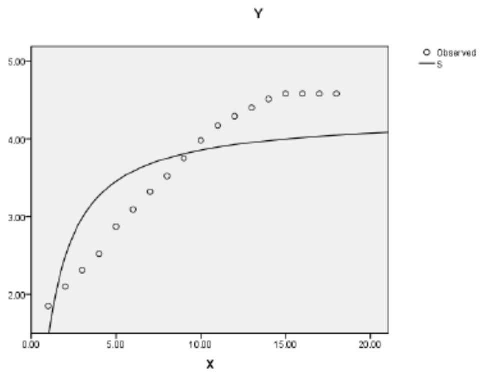
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.099	1	1.099	43.438	.000
Residual	.405	16	.025		
Total	1.504	17			

The independent variable is X.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 / X	-1.098	.167	-.855	-6.591	.000
(Constant)	1.459	.050		29.461	.000

The dependent variable is ln(Y).



Dari hasil pengolahan data menggunakan SPSS diperoleh parameter a=1,459 dan parameter b=-1,098 dan kurva berbentuk sigmoid. Sehingga diperoleh persamaan regresi non linier model sigmoid yaitu $y = e^{1,459\frac{(-1,098)}{x}}$. Selain itu diperoleh R-sq sebesar 73,1%. Hal ini menunjukkan bahwa variabel x mempengaruhi keragaman y sebesar 73,1% dalam model sigmoid $y = e^{1,459\frac{(-1,098)}{x}}$ sisanya sebesar 26,9% dipengaruhi variabel lain ang tidak masuk dalam model.

- **Uji F**

Dari table ANOVA diatas diperoleh F_{hitung}sebesar 43.438 dengan tingkat siknifikansi sebesar 0,000. Oleh karena probabilitas(0,000)<0,05(dalam kasus ini menggunakan taraf signifikansi atau α=5%), maka model regresi nonlinier sigmoid ini dapat digunakan untuk memprediksi tinggi batang. Biasanya output ini digunakan untuk menguji hipotisis. Hipotisisnya yaitu

H_0 : tidak ada hubungan antara pengamatan(hari) terhadap tinggi batang.

H_1 : ada hubungan antara pengamatan(hari) terhadap tinggi batang.

F_{tabel} = 4,4513

Karena statistik hitung(F_{hitung}) > statistik tabel(F_{tabel}) maka H_0 ditolak, dan probabilitas (0.0000) jauh lebih kecil dari 0.05 maka model regresi dapat dipakai untuk memprediksi tinggi batang.

- **Uji T**

Uji t digunakan untuk menguji signifikansi konstanta dan variabel independen (pengamatan).

Menguji signifikan konstanta pada model

Hipotisis:

H_0 : koefisien regresi a tidak signifikan.

H_1 : koefisien regresi a signifikan.

Dalam table coefficient diperoleh nilai signifikan sebesar 0,000 dibandingkan dengan taraf signifikan (α=5%) 0,05 maka

Sig<α maka disimpulkan untuk menolak H_0, yang berarti koefisien regresi a signifikan.

Menguji signifikan koefisien variabel

Hipotisis :

H_0 : koefisien regresi tinggi batang tidak signifikan.

H_1 : koefisien regresi tinggi batang signifikan.

Dari perhitungan SPSS diperoleh T_{hit}=29.461 sedangkan T_{tabel}=2,1098. karena T_{hit} >T_{tabel} , maka H_0 ditolak. Artinya koefisien regresi di atas sudah signifikan.

Karena signifikanT dan signifikanF adalah 0.0000 dan kurang dari 0.05, maka H_0 ditolak. Dengan kata lain koefisien regresi signifikan.

Dari hasil analisis regresi disimpulkan bahwa hubungan antara pengamatan dalam hari yang dengan tinggi batang (cm) dengan perhitungan secara manual, maupun menggunakan minitab dan SPSS menghasilkan persamaan regresi linier lnY=1,459-1,0979(1/X) sedangkan persamaan regresi non linier model sigmoidnya adalah $y = e^{1,459\frac{(-1,098)}{x}}$. Berdasarkan hasil analisis diperoleh nilai F-hitung (F =4.494) dengan signifikasi F =0,0000, dimana nilai signifikasi ini lebih kecil dari 0,05 yang berarti nilai F-hitung lebih besar dari F-tabel, dengan kata lain data sangat mendukung adanya hubungan antara waktu pengamatan dalam hari dengan tinggi batang (cm) kacang tunggak.

DAFTAR PUSTAKA

Arsyad, L. 1994. *Peramalan Bisnis Edisi Pertama*. Yogyakarta: BPFE.

Assauri, S. 1984. *Teknik dan Metode Peramalan Penerapannya dalam Ekonomi dan Dunia Usaha Edisi Pertama*. Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.

Cryer, J.D., 1986, *Time Series Analysis*, Boston: PWS-KENT Publishing Company.

Draper, N. R. Dan Smith, H., 1966, *Analisis Regresi Terapan*, Jakarta: PT Gramedia Pustaka.

Ghozali, Imam, 2016. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 23: Edisi 8*. Semarang: Badan Penerbit Universitas Diponegoro.

Ghozali, Imam, 2009, *Ekonometrika Teori, Konsep dan Aplikasi dengan SPSS 17*, Semarang: Badan Penerbit Universitas Diponegoro.

Gujarati, D. 2004. *Basic Econometrics 4th Edition*. New York: McGraw-Hill.

Hasan, M. Iqbal, 2005, *Pokok-pokok Materi Statistika 1*, Jakarta: PT Bumi Aksara.

Lipsey, R.G. dan P.O. Steyner. 1981. *Pengantar Ilmu Ekonomi III Edisi Keenam*. Jakarta: Bina Aksara.

Makridakis, S., S.C. Wheelwright dan V.E. McGee. 1988. *Metode dan Aplikasi Peramalan Edisi Kedua Jilid Satu*. Alih bahasa: Ir. Untung Sus Adriyanto, M.Sc dan Ir. Abdul Basith, M.Sc. Jakarta: Erlangga.

Makridakis, S., S.C. Wheelwright dan V.E. McGee. 1999. *Metode dan Aplikasi Peramalan Edisi Kedua Jilid Satu*. Alih bahasa: Hari Suminto. Jakarta: Binarupa Aksara.

Mann, Prem S., _, *Statistical for Business and Economics*, John Wilwy & Sons, Inc.

Nainggolan, S., 2010, *Perbandingan Metode Marquardt Compromise dan Metode Gauss Newton dalam Penaksiran Parameter Regresi Nonlinier*, Skripsi, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sumatera Utara.

Riduan dan Kuncoro, Engkos Achmad, 2008, *Cara Menggunakan dan Memaknai Analisis Jalur (Path Analysis)*, Bandung:Alfabeta.

Setiawan, & Kusrini, D. E. 2010. *Ekonometrika*. Yogyakarta: Penerbit ANDI.

Sudjana, 1996, *Metoda Statistika Edisi Ke 6*, Bandung: Tarsito

Sugiyono, 2010, *Statistika untuk Penelitian*, Bandung: Alfabeta.

Sylfi, Ispriyanti, D., dan Safitri, D., 2012, Analisis Regresi Linier *Piecewise* Dua Segmen, *Jurnal Gaussian*, Vol. 1, Nomor 1, Tahun 2012, 219-228.

Wei, W.W.S., 1990, *Time Series Analysis: Univariate and Multivariate Methods*, USA: Addison-Wesley Publishing Co.

Yamin, Sofyan dan Kurniawan, Heri, 2009, *Structural Equation Modeling Belajar Lebih Mudah Teknik Analisis Data Kuesioner dengan LISREL-PLS*, Jakarta: Salemba Infotek.