

**MODUL PRAKTIKUM**  
**SERTIFIKASI PRAKTIKUM ALAT ANALISIS**



Oleh

**Firsty Ramadhona Amalia Lubis , S.E., M.Ec.Dev**  
**60181150**

**PROGRAM STUDI EKONOMI PEMBANGUNAN**  
**FAKULTAS EKONOMI DAN BISNIS**  
**UNIVERSITAS AHMAD DAHLAN**  
**GASAL 2020-2021**

# **BAB I**

## **PENDAHULUAN**

Sertifikasi Praktikum Alat Analisis mempelajari berbagai dasar dan pengetahuan tentang pengolahan data statistika dan ekonometrika secara elektronis. Beberapa materi yang akan dibahas antara lain: konsep dasar Analisis dan Uji Regresi Data Panel, Fungsi Statistik, Pengertian, Jenis-Jenis Data, Sumber Data, dan Skala Pengukuran, Programisasi Linear dengan Microsoft Excel, Analisis Trend, Uji Validitas dan Reliabilitas, Uji Korelasi menggunakan program STATA, Regresi dengan Data Time Series, R-Squared, Uji Parsial dan Uji Serempak, dan Analisis Regresi Data Panel menggunakan program STATA. Sertifikasi Praktikum Alat Analisis tidak hanya menerapkan konsep teoritis namun juga mengenai penerapan berbagai macam alat analisis data statistika dalam bidang ekonomi dan bisnis. Tujuan analisis data adalah untuk mengidentifikasi, mengubah, dan mendukung pengambilan keputusan dan mengarah kepada kesimpulan hasil riset.

### **1.1 ANALISIS DATA**

Analisis data adalah metode menemukan fakta dan angka untuk memecahkan pertanyaan penelitian.. Sulit untuk menafsirkan berdasarkan data mentah saja, dalam hal ini data harus dianalisis dan disimpulkan. Data dari suatu penelitian dapat berupa numerik atau kuantitatif. Jika data tidak berbentuk numerik, dapat dilakukan analisis kualitatif. Menurut Ashirwadani (2014) analisis data penelitian dapat diklasifikasikan menjadi analisis data naratif, deskripsi dan statistik.

The diagram consists of three stacked, light gray rectangular boxes with rounded corners. Each box contains a title for a type of data analysis. The top box is labeled 'Analisis Data Naratif', the middle box is labeled 'Analisis Data Deskripsi', and the bottom box is labeled 'Analisis Data Statistik'. The boxes are arranged vertically, with the top box being the largest and the bottom box being the smallest, creating a descending staircase effect.

Analisis Data Naratif

Analisis Data Deskripsi

Analisis Data Statistik

**Gambar 1.1 Pengklasifikasian Analisis Data**

**a. Analisis Data Naratif**

Analisis data naratif yang muncul dari penelitian kualitatif menggunakan teks seperti cerita, wawancara, surat, percakapan, foto, buku harian, otobiografi, catatan lapangan, dll. Sebagai unit analisis untuk mendukung alasan pertanyaan penelitian. Jenis analisis ini digunakan sebagai alat untuk menganalisis bidang ilmu kognitif, penelitian organisasi, seni, sosiologi, dan pendidikan.

### **b. Analisis Data Deskripsi**

Analitik deskripsi unggul dalam menganalisis data menggunakan metode kuantitatif. Analisis deskripsi dimaksudkan untuk memberikan ringkasan sederhana dari sampel beserta pengamatan visual seperti grafik dan pengamatan laboratorium. Hasil dan kesimpulan dari jenis analisis data ini dianggap cukup untuk beberapa penyelidikan.

### **c. Analisis Data Statistik**

Analisis statistik sangat penting untuk setiap studi atau eksperimen yang menggunakan statistik sebagai metode penelitian. Sebagian besar ilmu-ilmu sosial menggunakan analisa statistika. Proses analisa ini merupakan alat yang sangat berguna untuk menghasilkan solusi perkiraan yang sangat kompleks.

Dalam setiap bidang, peneliti dapat menggunakan sejumlah pendekatan berbeda untuk dalam penelitian. Hal ini mencakup observasi, pengamatan, studi kasus, wawancara, kelompok fokus, survei, eksperimen laboratorium, dan eksperimen lapangan. Terlepas dari berbagai metode yang digunakan dan topik yang dipelajari, sebagian besar penelitian ilmu sosial memiliki beberapa fitur umum. Terlepas dari bidangnya, sebagian besar penelitian melibatkan peneliti mengumpulkan data dan melakukan analisis untuk menentukan makna data. (Lehman et al, 2005).

#### **- Langkah-langkah melakukan penelitian**

Menurut Lehman et al (2005), Langkah-langkah spesifik yang harus diikuti ketika melakukan penelitian tergantung pada topik penelitian. Namun, sebagian besar penelitian ilmu sosial mengikuti tindakan sistematis yang dimulai dengan perumusan pertanyaan penelitian dan diakhiri dengan kesimpulan dari hipotesis. Proses penelitian terdiri dari beberapa langkah berikut:

1. Mengembangkan pertanyaan penelitian
2. Mengembangkan pernyataan hipotesis penelitian
3. Mendefinisikan instrume riset
4. Data dikumpulkan
5. Analisa data

## 6. Menarik kesimpulan mengenai hipotesis

### - **Pertanyaan Penelitian**

Menurut Lehman et al (2005), Proses penelitian biasanya dimulai dengan upaya untuk sampai pada pernyataan yang jelas tentang pertanyaan penelitian. Pertanyaan penelitian adalah pernyataan tentang apa yang diharapkan ketika penelitian selesai. Jawaban atas pertanyaan-pertanyaan ini akan bermanfaat tergantung pada tujuan penelitian.

### - **Hipotesis**

Menurut Dayanand (2018), Hipotesis adalah pernyataan tentang harapan atau prediksi peneliti tentang hubungan antara variabel penelitian. Proses penelitian dimulai dan diakhiri dengan hipotesis. Pertanyaan peneliti mengidentifikasi konsep-konsep penelitian dan menanyakan bagaimana konsep-konsep ini berhubungan dengan hipotesis yang jawabannya harus diprediksi. Tiga aspek mendasar yang penting dalam hipotesis adalah: perbedaan yang kita coba pahami, hubungan dan interaksi. Dalam hubungan yang signifikan secara statistik berarti bahwa perbedaan hasil tidak terjadi secara kebetulan tetapi karena alasan statistik.

Jika perbedaan itu karena kebetulan, hasilnya tidak signifikan secara statistik. Biasanya ketika kita menyatakan hipotesis, akan ada variabel independen dan dependen. Variabel bebas adalah penyebab dan variabel terikat merupakan akibat. Sebuah hipotesis memastikan bahwa seluruh proses penelitian tetap ilmiah dan dapat diandalkan. Meskipun asumsi penting dalam proses penelitian, hipotesis dapat menyebabkan kompleksitas probabilitas, signifikansi, dan kesalahan. Hipotesis adalah dugaan atau prediksi berdasarkan pengamatan dalam suatu penelitian.

### - **Mendefinisikan Instrumen, Mengumpulkan, Menganalisis, dan Menarik kesimpulan pada data**

Menurut Lehman et al (2005), Dengan hipotesis yang dikemukakan, kita dapat melakukan uji penelitian dan menganalisis beberapa data yang relevan. Data dapat didefinisikan sebagai kumpulan skor yang diperoleh ketika karakteristik dan/atau kinerja subjek dievaluasi.

### - **Variabel, Nilai, dan Observasi**

Saat membahas sesuatu yang berhubungan dengan data, kita sering mendengar istilah variabel, nilai, dan observasi. Berikut dengan jelas dapat kita pahami istilah dari ketiga hal tersebut.

#### **a. Variabel**

Variabel adalah suatu konsep kemampuan untuk membagi ke dalam beberapa kategori dan jenis yang memungkinkan kita untuk mengetahui hubungannya dengan variabel lain (Abiodun, 2017).

#### **Tipe Variabel**

Menurut Abiodun (2017), Klasifikasi dari variabel dapat dilakukan berdasarkan tipologi yang berbeda. Berikut penjelasan klasifikasi variabel tersebut.

### **1. Variabel Kategori dan Kontinu**

#### **Variabel Kategori**

Variabel kategori dikenal sebagai variabel diskrit atau variabel kategoris atau variabel kualitatif. Variabel kategori yang memiliki sejumlah nilai yang berbeda dapat diklasifikasikan ke dalam beberapa jenis yang berbeda. Dalam variabel kategori, variasi variabel dikaitkan dengan kategori tertentu. Oleh karena itu, variasi dalam subset variabel kategori adalah fungsi dari posisi nilai pada posisi yang relatif berbeda terhadap skala.

Misalnya, gender dapat dibagi menjadi dua kategori yang berbeda, yaitu Pria dan Wanita. Pada hal ini seluruh anggota kategori atau subset diakui dan memiliki nilai yang sama. Variabel kategori dibagi menjadi nominal, ordinal, dikotomis, dan polikotomi.

(a) Variabel nominal: Ada dua atau lebih kategori tanpa nilai intrinsik. Pengklasifikasian nilai tipe variabel ini didasarkan pada persamaan atau perbedaan. Variabel nominal tidak mewakili ukuran. Sebagai gantinya, beri nama karakteristik subkelompok.

(b) Variabel ordinal: Variabel ini seperti variabel nominal, memiliki dua atau lebih kategori tetapi dapat diklasifikasikan. Oleh karena itu, satu jenis dapat dikatakan jumlah besar atau kecilnya dari yang lain.

Misalnya, tinggi dapat dianggap sebagai nilai rentang yang ditentukan. Contoh nilai ujian dan tingkat pendapatan.

(c) Variabel dikotomis: Variabel ini hanya memiliki dua kategori atau tingkatan. Variabel dikotomis memiliki fitur serupa dari variabel nominal dalam hal tidak memiliki nilai intrinsik. Contohnya adalah jenis kelamin. Kategorinya cenderung laki-laki dan perempuan.

(d) Variabel polikotomi: Variabel-variabel ini memiliki lebih dari dua kategori himpunan bagian. Contohnya adalah variabel seperti 'tingkat kemahiran komputer' yang dapat memiliki kelompok bagian seperti 'pemula', 'menengah' dan 'ahli'.

### **Variabel Kontinu**

Variabel dapat diukur dengan skala sedemikian rupa sehingga himpunan bagian atau pengelompokannya berbeda satu sama lain atas dasar kuantitas, derajat, tingkat atau jumlah. Variabel tersebut dapat diambil dengan jumlah nilai yang tidak terbatas. Variabel ini dikenal sebagai variabel kuantitatif atau terukur, dan memiliki fitur pengurutan peringkat. Variabel kontinu dapat dibagi lagi menjadi kelas-kelas seperti: variabel interval atau rasio.

## **2. Variabel Dependen dan Variabel Independen**

Dalam upaya memecahkan suatu permasalahan, peneliti dalam penelitian eksperimental berfokus pada:

Hubungan kausal yang dikenal sebagai hubungan fungsional, memanipulasi variabel untuk melihat efeknya pada variabel lain. Pekerjaan penelitian semacam ini dimulai dengan efek dan kemudian mencari penyebabnya. Variabel yang dimanipulasi dalam beberapa kasus bisa lebih dari satu yang disebut sebagai variabel bebas sedangkan variabel yang diharapkan akan terpengaruh oleh manipulasi disebut variabel terikat. Variabel independen dikenal sebagai variabel input dan melambangkan dugaan.

### **b. Nilai**

Nilai mengacu pada subjek yang relatif sama dengan variabel kuantitatif. Misalnya “skor”, Kata “skor” adalah pengganti yang tepat untuk kata "nilai" dalam kasus ini karena keduanya adalah variabel kuantitatif. Skor adalah variabel di mana angka berfungsi sebagai nilai (Abiodun, 2017).

### **c. Observasi**

Dalam membahas data, peneliti sering membuat referensi ke unit observasi (pengamatan). Pengamatan dapat diartikan sebagai subjek individu (objek lain) yang bertindak sebagai sumber data (Abiodun, 2017).

## **1.2 SAMPEL**

Sugiyono (2011) menyatakan bahwa sampel adalah sub dari total dan profil suatu populasi. Sampling didefinisikan sebagai proses memilih sampel dari sekelompok besar individu atau populasi untuk tujuan penelitian tertentu. Ada keuntungan dan kerugian yang berbeda dari pengambilan sampel. Mengapa perlu pengambilan sampel? Jawabannya adalah karena terlalu mahal dan terlalu memakan waktu untuk mensurvei seluruh populasi dalam studi penelitian (Bhardwaj, 2019).

Menurut Bhardwaj (2019), terdapat kelebihan dan kekurangan dalam menggunakan metode sampel.

### **Kelebihan**

- Menghemat waktu dan uang serta memberikan hasil yang lebih cepat karena ukuran sampel lebih kecil dari seluruh populasi.
- Pengambilan sampel memberikan hasil yang lebih akurat saat dilakukan oleh peneliti terlatih dan berpengalaman
- Bila terdapat jumlah populasi yang besar, pengambilan sampel adalah cara terbaik
- Pengambilan sampel memungkinkan untuk memperkirakan kesalahan pengambilan sampel. Hal ini karena dapat membantu dalam mendapatkan informasi mengenai beberapa karakteristik penduduk
- Studi yang menggunakan sampel membutuhkan lebih sedikit ruang dan peralatan karena ukurannya kecil
- Ketika sumber daya terbatas, pengambilan sampel adalah yang terbaik.

### **Kekurangan**

Kekurangan utama dari metode sampel adalah kemungkinan terjadinya bias. Namun, melihat begitu banyak keuntungan tersebut maka pengambilan sampel adalah cara terbaik dalam pengerjaan penelitian.

Menurut Retnawati (2017) dalam pengambilannya sampel terdiri dalam dua teknik pengambilan yaitu pertama, *Probability Sampling* adalah Teknik pengambilan sampel yang membuat setiap elemen (anggota) dari suatu populasi mempunyai kemungkinan yang sama untuk dipilih sebagai anggota sampel. Teknik *Probability Sampling* adalah teknik yang memungkinkan peneliti atau penilai untuk menggeneralisasi dari karakteristik sampel ke karakteristik populasi.

Ketika ukuran jumlah populasi sangat banyak, maka sampel merupakan alat penting yang bisa digunakan pada sebuah penelitian. Berdasarkan hal tersebut, Bhardwaj (2019) mengklasifikasikannya dengan 2 jenis yaitu *Probability Sampling* dan *Non-Probability Sampling*. Kedua jenis pengambilan sampel ini dibagi lagi menjadi sub tipe berikut:

### **Pengambilan sampel *simple random sampling***

Pada bagian di atas, kita telah membahas tentang metode melakukan *simple random sampling*. Ada dua jenis pengambilan sampel *simple random sampling*, yaitu:

#### **1. *Simple Random Sampling With Replacement (SRSWR)***

Memilih jumlah unit “n” dari unit “N” satu per satu sedemikian rupa sehingga pada setiap tahap seleksi, Sebuah sampel dari setiap unit memiliki kemungkinan yang sama untuk dipilih, yaitu  $1/N$ .

#### **2. *Simple Random Sampling Without Replacement (SRSWOR)***

Memilih jumlah unit “n” dari “N” satu per satu sehingga siapa pun dari unit yang tersisa memiliki probabilitas untuk dipilih sebagai sampel, yaitu  $1/N$ .

Sebagai contoh, jika kita ingin mengetahui jumlah kura-kura di kolam pada sebuah desa, maka jika kita menangkap kura-kura dari air, mengukurnya, dan mengembalikannya ke air, kemungkinan besar kita akan memilih kura-kura yang sama, ini adalah SRSWR. Namun, jika kita mengeluarkan penyu dari air dan tidak mengembalikannya tanpa mengambil berikutnya, hal ini menjadi SRSWOR.

**a. Stratified Random Sampling**

Dalam tipe ini, populasi pertama dibagi menjadi subkelompok yang disebut strata berdasarkan kesamaan dan kemudian dari setiap kelompok atau strata, anggotanya dipilih secara acak. Di sini, tujuannya adalah untuk mengatasi masalah homogenitas populasi yang kurang dan untuk membuat sampel benar-benar representatif.

Langkah-langkah rinci untuk memilih sampel acak bertingkat:

1. Pertama, kita menargetkan audiens
2. Kemudian kita mengenali variabel stratifikasi yang harus sesuai dengan tujuan penelitian dan kemudian mengetahui jumlah strata yang akan digunakan
3. Setelah mengumpulkan informasi tentang variabel stratifikasi, kita membuat kerangka untuk semua elemen kepada target audiens
4. Seluruh populasi kemudian dibagi menjadi strata yang berbeda dan dari masing-masing mencakup setiap elemen/anggota populasi. Tetapi, setiap anggota harus berada dalam satu strata saja
5. Kemudian kita menetapkan nomor unik dan acak untuk setiap elemen
6. Kemudian bagilah jumlah sampel yang dikumpulkan dari populasi dengan jumlah orang dalam kelompok.
7. Jumlah yang sekarang kita dapatkan adalah jumlah sampel untuk dipilih pada strata tertentu.

**b. Systematic sampling**

Langkah-langkah membentuk/memilih sampel menggunakan *Systematic sampling*:

- Mengembangkan populasi struktural yang terdefinisi dengan baik untuk memulai aspek pada sampel
- Tentukan ukuran sampel yang ideal
- Setelah menentukan ukuran sampel, tetapkan nomor untuk setiap anggota sampel
- Kemudian menentukan interval sampel

**c. Cluster sampling**

Dalam teknik pengambilan sampel *Cluster sampling*, berbagai segmen populasi diperlakukan sebagai *cluster*, dan anggota dari setiap *cluster* dipilih secara acak. Pengambilan sampel *Cluster sampling* dan pengambilan sampel *stratified sampling* berbeda antara dengan lainnya.

Dalam pengambilan sampel *stratified sampling*, peneliti membagi populasi menjadi subkelompok berdasarkan usia, jenis kelamin, profesi, dll., sedangkan dalam

pengambilan pada sampel *Cluster sampling* peneliti memilih secara acak dari kelompok/cluster yang sudah ada. Dalam metode ini, peneliti membuat cluster sesuai dengan kebutuhan dan kemudian memilih sampel sesuai dengan *simple random sampling/systematic sampling*.

**d. *Multistage sampling***

Dalam pengambilan sampel *multistage sampling* ini berisi banyak tahapan. Setiap cluster sampel dibagi lagi dari cluster yang kecil dan anggota dipilih dari setiap cluster yang kecil secara acak. Sampel *multistage sampling* adalah bentuk sampel cluster yang kompleks

**2. *Nonprobability sampling***

Pengambilan sampel *Nonprobability sampling* merupakan teknik sampling dengan tidak ada kemungkinan yang diketahui bahwa setiap anggota populasi akan dimasukkan dalam sampel. Misalnya, untuk mempelajari dampak pekerja anak dibawah umur, peneliti hanya akan mencari dan mewawancarai anak-anak yang menjadi sasaran pekerja anak. Berikut ini adalah jenis- jenis *Nonprobability sampling*:

**a. *Purposive sampling***

Dalam jenis pengambilan sampel ini, sesuai dengan tujuan penelitian maka anggota sampel dipilih. Hal Ini juga disebut dengan pengambilan sampel yang dilakukan secara disengaja.

**Waktu penggunaan pengambilan sampel *Purposive sampling***

- Ketika jumlah orang lebih sedikit dalam populasi dan peneliti mengetahui bahwa populasi sasaran telah memenuhi kriteria.
- Ketika terdapat kebutuhan untuk menyaring sampel yang dipilih oleh metode sampel lainnya, metode pengambilan sampel ini adalah yang terbaik namun pengetahuan dan pengalaman peneliti juga dapat menjadi faktor yang berpengaruh.

**Kelebihan menggunakan *Purposive sampling***

- Ketika pemilihan sampel dilakukan oleh peneliti yang berpengalaman maka tidak terdapat kesulitan dalam pengambilan sampel ini.

- Sampel yang dipilih menjadi responden merupakan sampel yang baik. Peneliti mendapatkan hasil real-time karena anggota sampel memiliki pengetahuan yang sesuai dan memahami subjek dengan baik.
- Seorang peneliti dapat menghasilkan hasil yang diinginkan dan dapat berkomunikasi langsung dengan sampel sasaran.

#### b. *Convenience sampling*

Dalam jenis pengambilan sampel ini memilih anggota sampel berdasarkan aksesibilitas.. Hanya anggota terpilih dan mudah diakses yang digunakan oleh peneliti. Data yang tersedia digunakan tanpa persyaratan tambahan. Peserta atau sampel dipilih mana yang lebih mudah dan didapatkan dalam pengambilan sampel.

#### **Kelebihan menggunakan *Convenience sampling***

- Sangat mudah diterapkan dan biaya yang dikeluarkan murah dalam proses pengambilan sampel.
- Berguna untuk percontohan dalam studi serta pengujian hipotesis
- Dalam waktu yang sangat singkat data dapat dikumpulkan.

### **1.4 DATA KUALITATIF DAN KUANTITATIF**

Penelitian kualitatif sangat baik dalam menjawab pertanyaan 'mengapa', 'apa' atau 'bagaimana'. Masing-masing pertanyaan ini dapat dijawab dengan menggunakan teknik kuantitatif seperti kuesioner terstruktur, penskalaan sikap, dan pengukuran standar. Data kuantitatif adalah pengukuran numerik yang diungkapkan bukan dengan deskripsi bahasa alami, melainkan dalam bentuk angka. Namun, tidak semua angka kontinu dan terukur. Data kuantitatif selalu dikaitkan dengan skala ukuran. Data ini dapat diwakili oleh skala ordinal, interval atau rasio dan cocok untuk sebagian besar manipulasi statistik.

### **1.5 APLIKASI KOMPUTER DALAM ANALISIS DATA**

Komputer sangat diperlukan selama proses penelitian. Peran komputer menjadi lebih ketika penelitian dilakukan dalam skala besar. Data yang dikumpulkan dapat disimpan di komputer untuk digunakan segera atau sebagai cadangan di perangkat penyimpanan tambahan. Selama analisis data, komputer membantu dalam bagian matematika yaitu perhitungan berbagai metode statistik. Ada berbagai perangkat lunak yang tersedia untuk keperluan perhitungan seperti Microsoft Excel, SPSS, E-Views, dan STATA yang populer digunakan untuk analisis data penelitian.

*Software* tersebut dapat digunakan dalam menghitung ukuran sampel penelitian yang diusulkan, pengujian hipotesis dan menghitung hasil penelitian. Komputer tidak hanya berguna untuk analisis statistik tetapi juga untuk memantau keakuratan dan kelengkapan data yang dikumpulkan. Seiring berkembangnya penelitian kuantitatif saat ini, penggunaan perangkat lunak statistik (*Statistical Software*) menjadi bagian yang penting dari analisis data. Para peneliti sedang mengalami transisi dari analisis manual dengan kertas ke analisis digital/elektronik yang lebih efisien dengan perangkat lunak statistik (Matthew dan Sunday, 2014).

- **Fitur *Statistical Software***

Menurut Matthew dan Sunday (2014), Perangkat lunak statistik memiliki beberapa karakteristik umum yang membuatnya dapat diandalkan dan cocok untuk analisis data:

1. Editor data pada baris dan kolom membuatnya sangat mudah untuk memasukkan data numerik.
2. Tersedia menu yang terdiri dari menu drop-down, analisis cepat serta panduan pengguna singkat.
3. Tingkat pengukuran statistik dipertimbangkan dalam entri data
4. Merekam langkah awal dalam proyek penelitian
  - (a) Menyiapkan data untuk dimasukkan ke dalam perangkat lunak.
  - (b) Pendefinisian dan pelabelan variabel
  - (c) Memasukkan data secara tepat dengan setiap baris berisi setiap kasus dan setiap kolom sebagai variabel.
  - (d) Pemeriksaan dan pembersihan data.
    - Semua data harus dalam bentuk numerik, meskipun terdapat tidak semua variabel yang diinginkan menggunakan huruf atau kata (variabel String) sebagai data. Hal ini dapat dicapai dengan pengkodean ulang huruf atau kata (data string) menjadi numerik yang diinginkan dan diberi label dengan tepat.
    - Eksplorasi data dapat dilakukan untuk memeriksa kesalahan dan akurasi lainnya.

## **BAB III**

### **STATISTIKA EKONOMI**

#### **3.1 PENGERTIAN STATISTIKA**

Menurut Sudjana (1989), statistika merupakan ilmu yang berkaitan dengan teknik dan sarana menganalisis, penyajian, dan penerbitan data dalam bentuk numerik. Aspek-aspek yang dibahas dalam statistika adalah bagaimana data dikumpulkan, diolah, disajikan, dianalisis, dan kesimpulan hasil analisa.

#### **3.2 JENIS-JENIS STATISTIKA**

Jenis-jenis statistika berdasarkan pembahasan adalah sebagai berikut:

**a. Statistik Matematika**

Topik yang dibahas meliputi pemahaman penggunaan uji-t, uji normalitas, analisis regresi, dll. Statistik matematika terinspirasi oleh statistik terapan dan mengembangkan banyak teknik untuk penerapannya.

**b. Statistik Terapan**

Memprioritaskan konsep dan teknik statistik dalam aplikasi dalam bidang tertentu, contoh statistik ekonomi, statistik sosial, statistik bisnis, dll.

Jenis-jenis statistika berdasarkan tahapan analisa adalah seperti berikut:

**a. Statistik deskriptif**

Statistik deskriptif merupakan metode mengumpulkan dan menampilkan data dalam menginformasikan hal yang berguna. Statistik deskriptif digunakan untuk memberikan gambaran tentang objek yang diteliti dengan menggunakan data sampel atau populasi. Meliputi tampilan data, pengukuran konsentrasi data, distribusi data, dan deret waktu, serta prakiraan.

**b. Statistik induktif**

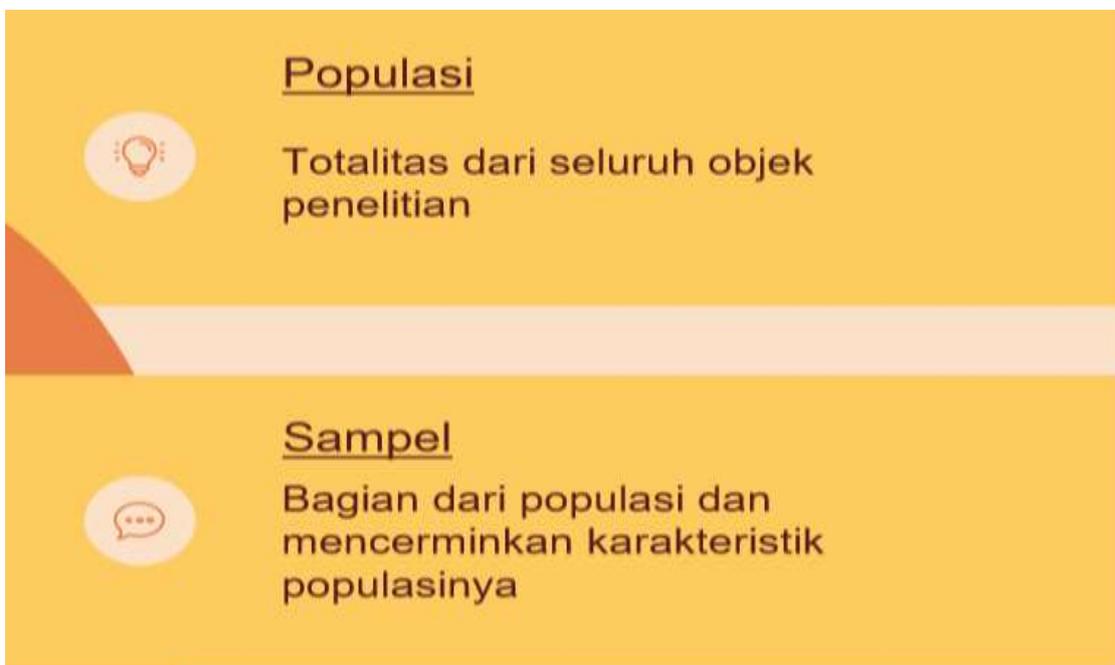
Statistik inferensi/induktif adalah statistik yang bertujuan untuk membuat estimasi umum populasi berdasarkan hasil sampel, termasuk estimasi teori dan pengujian teori. Mencakup probabilitas dan teori keputusan, sampling, pengujian hipotesis, regresi dan korelasi, dan statistik nonparametrik.

Penerapan konsep statistik sebagai metode ilmiah:

- Studi literatur
- Hipotesis
- Pengumpulan data
- Pemrosesan data
- Pengujian hipotesis
- Kesimpulan
- Publikasi hasil

### 3.3 POPULASI DAN SAMPEL

Dalam statistik, populasi adalah kumpulan data dengan karakteristik yang sama dan kesimpulan yang diambil darinya. Menurut Handayani (2020), populasi adalah seluruh subjek yang diteliti yang memiliki kesamaan karakteristik dan dapat berupa individu-individu dari kelompok, peristiwa atau apapun yang sedang diselidiki. Menurut Siyoto et al (2015), sampel merupakan sub dari total dan profil suatu populasi yang diperoleh dengan mengikuti tahapan tertentu untuk dapat mewakili populasi tersebut. Sampel adalah jumlah pengamatan tak bias yang diambil dari suatu himpunan.



**Gambar 4.1 Definisi Populasi dan Sampel**

### 3.4 SKALA PENGUKURAN

Skala Pengukuran merupakan dasar untuk mengukur jangka panjang atau pendek yang terdapat dalam suatu alat ukur, pada saat alat ukur tersebut digunakan dalam suatu pengukuran menghasilkan data kuantitatif (Sugiyono, 2006). Jenis skala pengukuran adalah sebagai berikut (Nurgiantoro et al, 2000):

**a. Skala nominal**

Skala nominal adalah skala terkecil/terrendah. Skala nominal memungkinkan untuk membedakan objek atau peristiwa hanya dengan nama. Penggunaan skala nominal untuk mengklasifikasikan objek, orang, atau grup ke dalam beberapa kategori. Tampilan numerik atau lambang pada skala nominal tidak dimaksudkan untuk diukur.

Contohnya, gender dikodekan sebagai 1 sebagai pria dan 2 sebagai wanita. Angka ini hanya berfungsi sebagai simbol kategoris dan tidak memiliki nilai atau makna intrinsik. Skala nominal merupakan sistem klasifikasi yang menempatkan orang, benda, atau entitas lain ke dalam kategori yang saling eksklusif. Variabel yang diukur dengan skala nominal adalah variabel klasifikasi yang menunjukkan kelompok tempat setiap subjek berada (Abiodun, 2017).

**b. Skala ordinal**

Nama lain skala ordinal adalah skala penilaian. Dalam skala ordinal, simbol hasil pengukuran tidak hanya mewakili perbedaan, tetapi juga urutan atau derajat benda yang diukur berdasarkan sifat-sifat tertentu. Contohnya, kepuasan individu terhadap suatu produk. Kita dapat melambangkan angka 5 sama dengan sangat puas, 4 sama dengan puas, 3 sama dengan netral, 2 sama dengan tidak puas, 1 sama dengan sangat tidak puas.

Nilai pada skala ordinal mewakili urutan peringkat subjek sehubungan dengan variabel yang dinilai. Misalnya, peringkat yang mewakili urutan peringkat subjek menurut keefektifannya secara keseluruhan. Nilai pada skala ordinal ini mewakili hierarki level sehubungan dengan efektivitas. Artinya, kita tahu bahwa misalnya agen peringkat "1" dianggap sebagai lebih efektif daripada agen peringkat "2," bahwa agen peringkat "2" lebih efektif daripada yang berperingkat "3", dan seterusnya (Abiodun, 2017).

**c. Skala interval**

Dengan skala interval, perbedaan yang sama antara nilai skala memiliki makna kuantitatif yang sama. Skala interval memberikan lebih banyak informasi kuantitatif daripada skala ordinal. Contoh skala interval, skala derajat Fahrenheit yang digunakan

untuk mengukur suhu. Dengan skala Fahrenheit, perbedaan antara 70 derajat dan 75 derajat sama dengan perbedaan antara 80 derajat dan 85 derajat. Satuan pengukuran sama di seluruh rentang skala (Abiodun, 2017).

**d. Skala rasio**

Skala rasio mirip dengan skala interval dalam perbedaan yang sama antara nilai skala memiliki arti kuantitatif yang sama. Namun, skala rasio juga memiliki titik nol yang sebenarnya. Misalnya, sistem inci yang digunakan dengan penggaris yang sama adalah contoh skala rasio. Ada titik nol yang benar karena nol inci sebenarnya menunjukkan tidak adanya panjang sama sekali. Dengan skala ini, memungkinkan untuk membuat pernyataan yang bermakna tentang rasio. Tepat untuk mengatakan bahwa panjang benda empat inci adalah dua kali panjang benda yang panjangnya dua inci. Usia yang diukur dalam tahun, juga dalam skala rasio — rumah berusia 10 tahun dua kali lebih tua dari rumah berusia 5 tahun (Abiodun, 2017).

## **BAB IV**

### **ANALISIS TREND**

#### **1.1 PENGERTIAN ANALISIS TREND**

Analisis trend adalah metode analisis yang dirancang untuk membuat perkiraan atau prakiraan di masa mendatang. Misalnya, pengembangan Pendapatan Asli Daerah, penerimaan pajak, royalti, harga, kinerja penjualan, dll. Analisis trend bertujuan untuk mengetahui trend atau kecenderungan keadaan perusahaan dimasa yang akan datang dan menunjukkan apakah trend tersebut naik, turun atau stabil. Tren dibagi menjadi tiga metode: tren linier, tren parabola, dan tren eksponensial. Keputusan untuk penggunaan metode adalah untuk menghasilkan plot sebar dari data yang diamati.

#### **1.2 KEGUNAAN ANALISIS TREND**

Analisis trend memiliki beberapa kegunaan, yaitu:

##### **a. Perencanaan**

Adanya rencana terpadu yang dapat digunakan sebagai alat untuk menyusun rencana dan mengelola berbagai kegiatan secara keseluruhan.

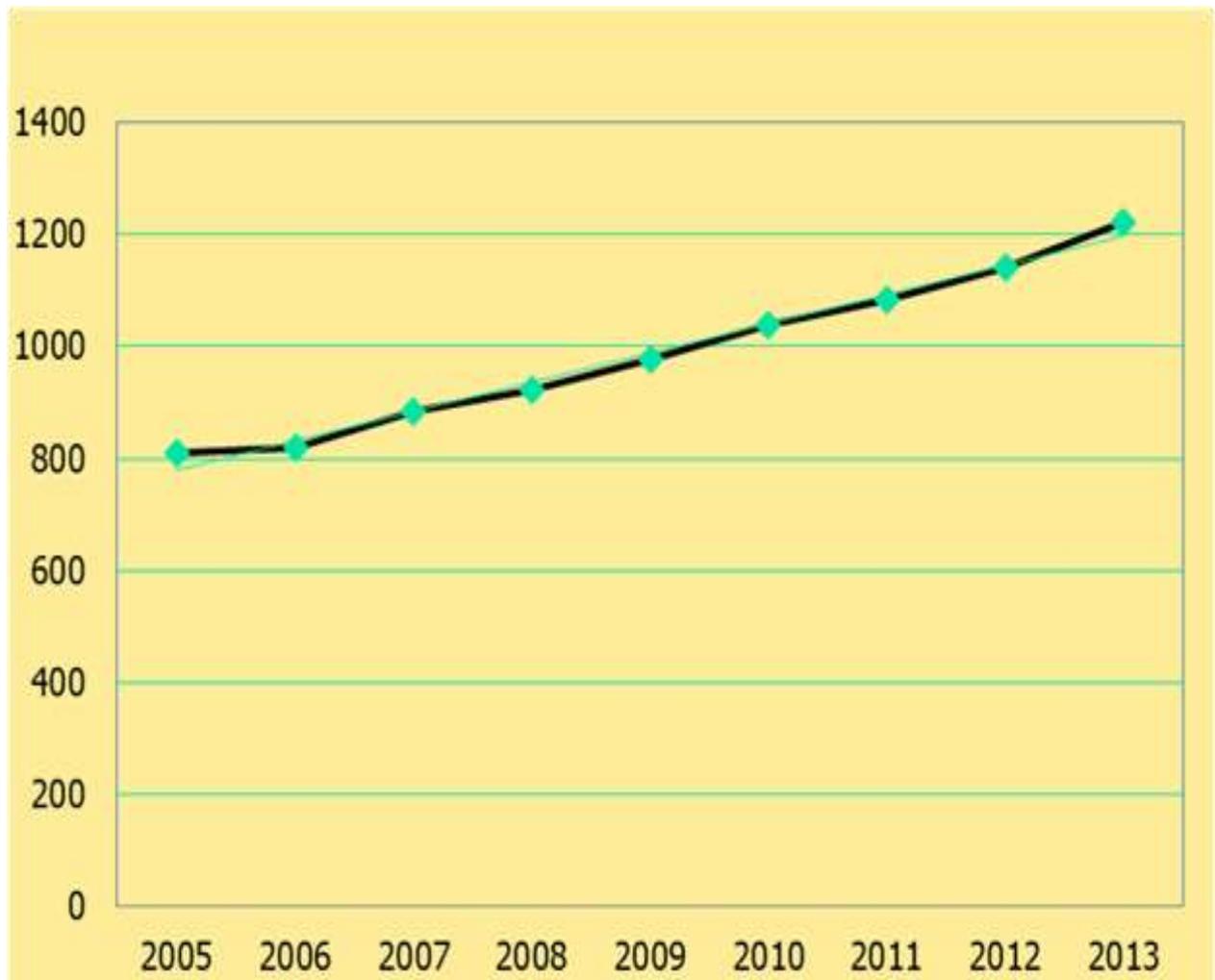
##### **b. Pengawasan**

Sebagai alat pemantauan kerja. Setiap anggaran harus memiliki seperangkat standar atau sasaran yang dapat dibandingkan dengan implementasinya sehingga kita dapat mengevaluasi kinerja setiap aktivitas. Menetapkan tolok ukur membutuhkan pemahaman yang realistis tentang aktivitas yang dilakukan dan analisis yang cermat. Pengaturan standar yang tanpa batas dasar pengetahuan dapat menyebabkan lebih banyak masalah daripada manfaat.

##### **c. Evaluasi**

Sebagai alat untuk mengevaluasi kegiatan. Anggaran disusun dengan baik dengan menerapkan kriteria yang memberikan arah perbaikan untuk memutuskan langkah-langkah yang harus diikuti sehingga dapat diimplementasikan dengan cara yang lebih tepat. Dengan kata lain, kita harus menggunakan sumber yang dianggap paling menguntungkan untuk setiap penyimpangan yang mungkin diderita perusahaan. Evaluasi dapat menentukan keputusan masa depan.

## Contoh garis trend



**Gambar 6.1 Produk Domestik Regional Bruto (PDRB) Provinsi Daerah Istimewa Yogyakarta Atas dasar harga konstan selama 9 tahun**

### 6.3 GERAKAN TREND JANGKA PANJANG

Terdapat beberapa variasi gerakan trend jangka panjang pada data berkala, yaitu sebagai berikut:

#### a. Fluktuasi periodik (efek periodik)

Fluktuasi yang dapat terjadi ketika data terdapat pengaruh oleh fluktuasi ekonomi pada jangka panjang. Fluktuasi periodik ini dapat berulang setelah periode waktu tertentu. Fluktuasi ini biasanya menjadi normal setiap 10 atau 20 tahun atau mungkin tidak terjadi kembali untuk jumlah waktu yang sama. Hal tersebut yang membedakan antara fluktuasi periodik dan musiman. Contoh yang menunjukkan fluktuasi periodik, seperti pada

industri konstruksi, memiliki pergerakan siklus 15 hingga 20 tahun, sedangkan industri otomotif dan pakaian jadi memiliki pergerakan siklus yang lebih pendek.

**b. Fluktuasi musiman (efek musiman)**

Fluktuasi musiman adalah perubahan yang beresilasi, seperti gelombang yang terjadi pada musim-musim tertentu. Fluktuasi sering terjadi setiap tahun, seringkali karena iklim dan kebiasaan (kadang-kadang secara konstan). Contoh fluktuasi musiman, seperti peningkatan penjualan pakaian selama liburan dan peningkatan penjualan buku dan tas sekolah di awal tahun ajaran.

**c. Fluktuasi tidak teratur (random fluktuasi)**

Fluktuasi yang terjadi akibat perubahan atau gerakan yang tidak teratur. Sulit untuk memprediksi fluktuasi ini. Contoh, perubahan data deret waktu karena perang, bencana alam, dan penyebab lainnya

## **6.4 METODE ANALISIS TREND**

Terdapat beberapa metode untuk menghitung analisis menggunakan tren, yaitu::

1) Metode Garis Trend Secara Bebas (Free Hand Method)

Menjelaskan tren menggunakan metode gratis ini cukup mudah. Hanya dengan mengamati distribusi data kita dapat mengenali garis tren dalam sampel data. Tentu saja, metode ini tidak dapat membenarkan hasil. Keuntungan dari metode ini merupakan metode sangat sederhana dan mudah diterapkan. Kelemahan dari metode ini adalah bahwa menggambar garis tren dari distribusi data sangat subjektif. Bahkan dengan data yang sama, tren garis akan bervariasi tergantung pada siapa yang menggambarinya

2) Metode Trend dengan Metode Setengah Rata-Rata (Semi Average Method)

Metode ini melibatkan usaha untuk menghilangkan subjektivitas seperti pada metode bebas. Langkah-langkah untuk mendapatkan trend linier menggunakan metode semi rata-rata sebagai berikut :

- a. Data dikelompokkan menjadi dua kelompok yang sama jika datanya ganjil maka data tengah diasumsikan dua atau rangkap dan dihitung dalam kelompok satu atau dua atau data dihilangkan.

- (i) Hitung mean grup pertama ( $\bar{X}_1$ ) dan mean grup kedua ( $\bar{X}_2$ ). Letakkan ( $\bar{X}_1$ ) di tengah tahun kelompok pertama dan ( $\bar{X}_2$ ) di tengah tahun kelompok kedua.
- (ii) Hitung selisih antara dua mean, dengan mengurangkan mean grup kedua dari mean grup pertama ( $\sum \bar{X}_2 - \sum \bar{X}_1$ ). Jika hasilnya + berarti trend naik, jika - berarti trend turun.

### 3) Metode Trend Kuadrat Terkecil (Least Square Method)

Oleh karena itu, garis tren diperoleh dengan menentukan persamaan garis perbedaan kuadrat terkecil antara data asli dan data pada garis tren. Metode kuadrat terkecil ini paling banyak digunakan dalam analisis deret rasio untuk peramalan perdagangan..

### 4) Trend Metode Moment

Gunakan perhitungan matematis dan statistik khusus untuk menemukan garis fitur garis lurus yang terbentuk dari data historis perusahaan. Oleh karena itu, faktor subjektif dapat dihindari.

## 6.5 PERSAMAAN TREND LINIER

Secara umum persamaan linear analisis runtut waktu adalah sebagai berikut:

$$Y' = a + bt'$$

Dimana:

$Y'$  = Estimasi nilai Y pada periode t

a = Intercept pada Y

b = Slope

t = Waktu, misalnya minggu, bulan, tahun.

Cara memperoleh persamaan tersebut:

1. Cari lewat observasi nilai t dan Y
2. Gunakan rumus berikut:

$$b = \frac{\sum (t_i - \bar{t})(Y_i - \bar{Y})}{\sum (t_i - \bar{t})^2}$$

$$a = \bar{Y} - b\bar{t}$$

**Contoh soal:**

Berikut data penjualan selama 8 tahun terakhir, proyeksikan penjualan untuk 5 tahun ke depan dengan mencari r nya terlebih dahulu untuk menentukan a dan b.

TAHUN	Penjualan
2000	Rp110
2001	Rp121
2002	Rp133
2003	Rp146
2004	Rp156
2005	Rp178
2006	Rp198
2007	Rp200

	A	B	C	D	E	F	G	H
1	t	r	$t_i - \bar{t}$	$r_i - \bar{r}$	$(t_i - \bar{t})^2$	$(t_i - \bar{t})(r_i - \bar{r})$		
2	0	Rp110	-4	-Rp45	12	Rp158		
3	1	Rp121	-3	-Rp34	6	Rp86		
4	2	Rp133	-2	-Rp22	2	Rp33		
5	3	Rp146	-1	-Rp9	0	Rp5		
6	4	Rp156	1	Rp1	0	Rp0		
7	5	Rp178	2	Rp23	2	Rp34		
8	6	Rp198	3	Rp43	6	Rp107		
9	7	Rp200	4	Rp45	12	Rp157		
10	28	Rp1.242	0	0	42	580		
11								
12	$\bar{t}$	$\bar{r}$						
13	3,5	155		tentukan				
14				b	14		a+b(t)	
15			$r - (b \cdot t)$	a	107		tahun	proyeksi
16							2008	217
17							2009	231
18							2010	245
19							2011	259
20							2012	273
21								

Gambar 6.2 Perhitungan Data Trend

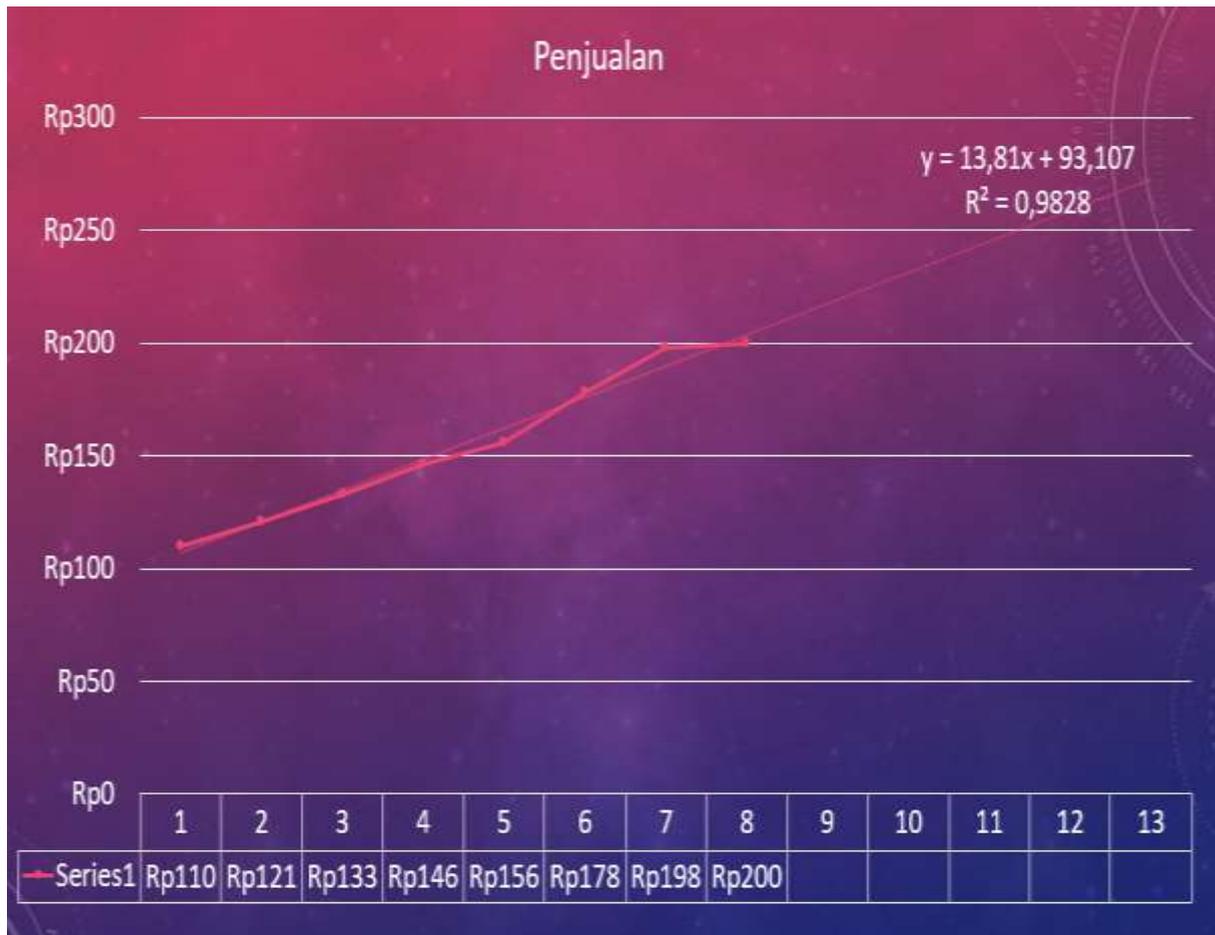
**Jawaban:**

t	r	ti-t̄	ri-r̄	ti-t̄ <sup>2</sup>	(ti-t̄)(ri-r̄)
0	Rp110	=A2-\$A\$13	=B2-\$B\$13	=C2^2	=C2*D2
1	Rp121	=A3-\$A\$13	=B3-\$B\$13	=C3^2	=C3*D3
2	Rp133	=A4-\$A\$13	=B4-\$B\$13	=C4^2	=C4*D4
3	Rp146	=A5-\$A\$13	=B5-\$B\$13	=C5^2	=C5*D5
4	Rp156	=A6-\$A\$13	=B6-\$B\$13	=C6^2	=C6*D6
5	Rp178	=A7-\$A\$13	=B7-\$B\$13	=C7^2	=C7*D7
6	Rp198	=A8-\$A\$13	=B8-\$B\$13	=C8^2	=C8*D8
7	Rp200	=A9-\$A\$13	=B9-\$B\$13	=C9^2	=C9*D9
=SUM(A3:A9)	=SUM(B2:B9)	=SUM(C2:C9)	=SUM(D2:D9)	=SUM(E2:E9)	=SUM(F2:F9)

t̄	r̄
=AVERAGE(A2:A9)	=AVERAGE(B2:B9)

Tentukan	
b	=F10/E10
a	=B13-(E14*A13)

Tahun	Proyeksi
2008	=\$E\$15+\$E\$14*8
2009	=\$E\$15+\$E\$14*9
2010	=\$E\$15+\$E\$14*10
2011	=\$E\$15+\$E\$14*11
2012	=\$E\$15+\$E\$14*12



**Gambar 6.3 Grafik Proyeksi Penjualan**

## LATIHAN

Data dibawah adalah data realisasi pajak reklame Kabupaten Nganjuk Jawa Timur.

- a. Hitunglah nilai a dan b nya
- b. Proyeksikan dalam 5 tahun kedepan

<b>Tahun</b>	<b>Realisasi</b>
<b>1</b>	<b>Rp6.443.950.466</b>
<b>2</b>	<b>Rp10.957.995.950</b>
<b>3</b>	<b>Rp13.261.397.470</b>
<b>4</b>	<b>Rp13.875.969.916</b>
<b>5</b>	<b>Rp16.317.201.997</b>
<b>6</b>	<b>Rp16.187.458.779</b>
<b>7</b>	<b>Rp24.316.259.790</b>
<b>8</b>	<b>Rp29.282.773.150</b>
<b>9</b>	<b>Rp33.028.045.600</b>

## BAB VI

### STATA DAN ANALISIS REGRESI

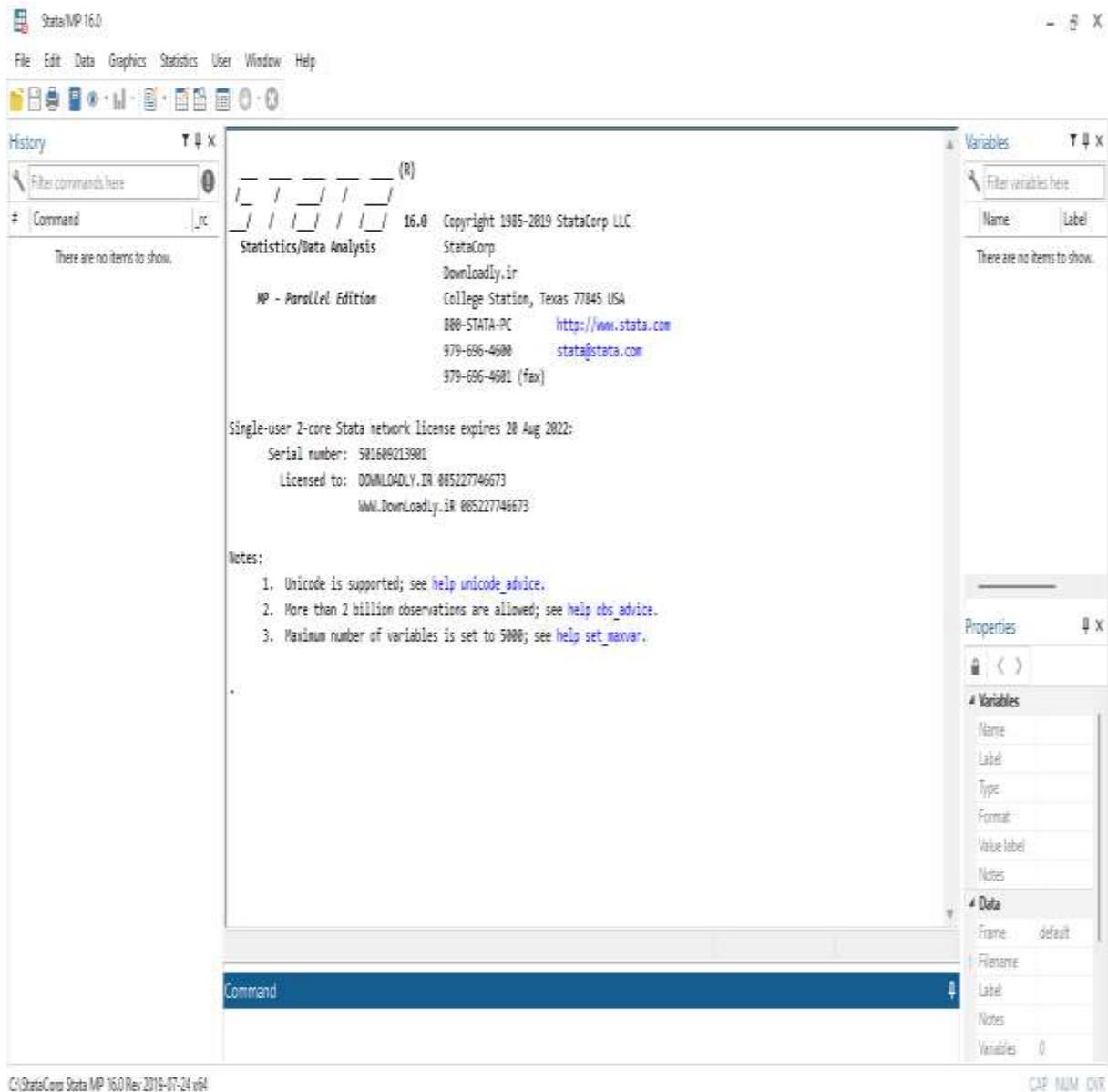
#### 6.1 PENGENALAN STATA

STATA merupakan *software* olah data statistik yang menyediakan berbagai fungsi untuk analisa dan pengelolaan data. STATA menyediakan berbagai fungsi analitis yang sangat luas untuk tujuan analitis. STATA menawarkan jangkauan yang sangat luas dari kemampuan analitis. Informasi lengkap mengenai STATA dapat diperoleh pada website: <http://www.stata.com/features/>.

##### 6.1.1 TAMPILAN STATA

Layar utama STATA memiliki lima panel, yaitu:

- a. **Result** menunjukkan hasil dari perintah pemrosesan yang dijalankan.
- b. **Command** digunakan untuk memasukkan perintah pemrosesan atau analisis.
- c. **Review** berisi daftar perintah yang dieksekusi sebelumnya. Pengguna tidak perlu memasukkan perintah yang sama lagi. Cukup klik perintah yang ingin Anda gunakan kembali di area ini.
- d. **Variabel** menunjukkan daftar variabel yang digunakan dalam kumpulan data yang diproses.
- e. **Properties** berisi informasi atau properti dari variabel yang dipilih.



**Gambar 6.1 Tampilan STATA**

### 6.1.2 PENGGUNAAN STATA

Dalam penggunaan STATA, Data yang dianalisa dapat diperoleh dari langkah-langkah berikut:

- a. Pengguna dapat menelusuri file yang telah disimpan menggunakan menu

**Klik File ---> Open File**

Kemudian cari di penyimpanan. Cara lain untuk menggunakan perintah panel control.

## Command: use namafilename.dta

b. Gunakan perintah impor,

### File-> Impor

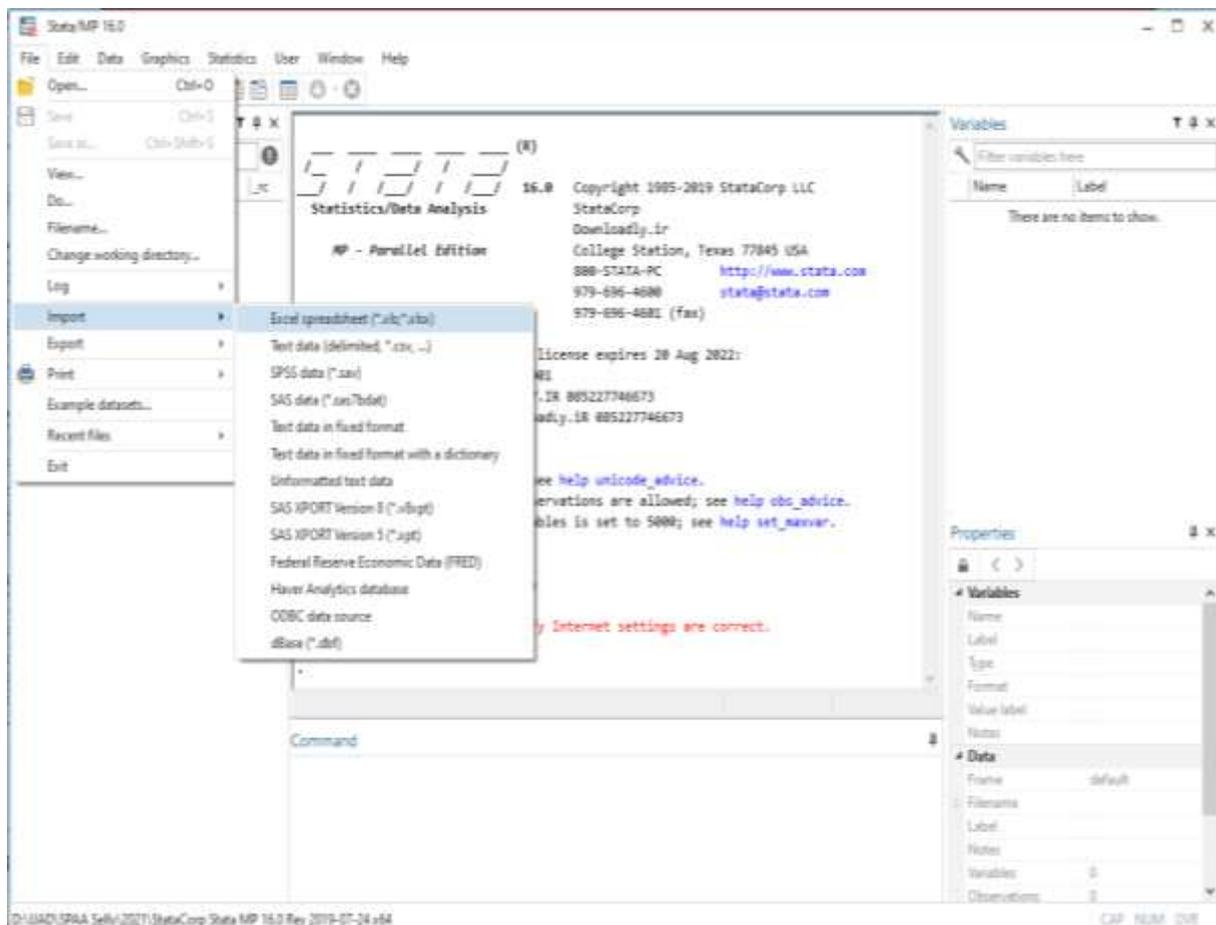
untuk membuka file dari database lain. STATA dapat mengimpor file dengan ekstensi .csv, .txt, .xls dan .xlsx.

c. Salin langsung dari data yang kita masukkan di spreadsheet dan tempel ke STATA (Contoh: Microsoft Excel).

Sebagai tambahan: Salin data yang digunakan oleh STATA. Kembali ke STATA dan buka

### Data Editor dari Data-> Data Editor

Tempelkan data yang telah disalin sebelumnya.



Gambar 6.2 Impor Data pada STATA

- d. **Describe**, describe merupakan perintah dalam Stata untuk melihat informasi umum mengenai data yang kita miliki.

**Command: describe**

Dari command ini, kita dapat melihat jumlah observasi, jumlah variabel, tipe variabel seperti numerik atau huruf (string), dan penjelasan mengenai setiap variabel.

Jika kita hanya memerlukan deskripsi singkat, maka kita bisa menggunakan:

**Command: describe, short**

Jika kita ingin melihat deskripsi beberapa variabel saja, maka kita bisa menaruh variabel yang diinginkan langsung setelah command describe atau des.

**Command: des Variabel1 Variabel2 Variabel3**

selain itu, jika kita hanya ingin melihat daftar variabel yang tersedia dalam data secara singkat, kita bisa menggunakan perintah ds.

**Command: ds**

**atau**

**ds Variabel1**

- e. **Browse**, browse merupakan perintah Stata untuk melihat nilai/isian variabel untuk semua observasi yang tersedia dalam data. Perintah ini dapat dimodifikasi juga sehingga hanya menunjukkan beberapa variabel. Hasil yang ditampilkan akan menyerupai tampilan spreadsheet excel.

**Command: browse**

**atau**

**br Variabel1 Variabel2 Variabel3**

- f. **List**, list merupakan perintah untuk mengeluarkan nilai/isian dari variabel di layar utama Stata. Kita bisa menggunakan list untuk menunjukkan observasi di sebuah variabel tertentu.

**Command: list Variabel1 Variabel2 Variabel3**

- g. **Tabulate**, tabulate atau tab merupakan perintah Stata untuk melihat frekuensi nilai dari sebuah variabel. Tabulasi bisa dilakukan secara *oneway* atau *twoway*.

**Command: tab Variabel1 Variabel2 Variabel3**

Untuk melihat persentase saat melakukan tabulasi, maka kita bisa menggunakan opsi tabulasi seperti col dan row.

**Command: tab Variabel1 Variabel2 Variabel3, col row**

- h. **Simbol operator pada STATA**, Dalam pengolahan data dan analisis data, akan banyak saat dimana kita perlu menggunakan operator untuk mengolah data. Berikut beberapa contoh yang dapat kita gunakan.

**Tabel 8.1 Contoh Simbol Operator pada STATA**

Simbol Operator	Keterangan
<=	Kurang dari sama dengan
>=	Lebih dari sama dengan
&	dan
<	Kurang dari
>	Lebih dari
	Atau

- i. **Count**, Perintah count menghitung jumlah observasi yang sesuai dengan kondisi yang diinginkan. Jika tidak ada kondisi, maka count akan menghitung jumlah observasi.

**Command: count**

**Atau**

**count Variabel1**

- j. **Summarize**, Perintah summarize atau sum memberikan informasi mengenai jumlah observasi yang tidak missing, mean, dan standard deviasi dari seluruh variabel. Untuk variabel huruf string, jumlah observasi ditulis 0 dan tidak ada rata-rata yang muncul dalam kolom ringkasan statistik yang tersedia. Kita juga harus berhati-hati menginterpretasikan ringkasan statistik untuk variabel numerik yang sifatnya kategoris.

**Command: Summarize**

**Atau**

**Sum Variabel1 Variabel2 Variabel3**

- k. **Logfiles**, adalah metode untuk menangkap semua perintah yang kita masukkan ke Stata dan output dari hasil olahan data kita. Untuk memulai sebuah logfile, kita menggunakan perintah log using.

**Command: log using nama\_folder\_simpan**

Semua perintah dan hasil yang dikeluarkan oleh Stata setelah perintah log using akan direkam dalam sebuah text file yang dapat dibuka menggunakan text editor seperti notepad. Sementara log close merupakan perintah untuk menutup log file yang sedang terbuka.

**Command: log close**

Untuk melihat hasil yang terekam dalam logfiles, silahkan membuka log file yang telah dibuat.

1. **Drop**, Dalam proses mengolah data, terdapat situasi dimana anda perlu menghapus variabel yang sudah dibuat. Semisal, kita ingin menggabungkan dua dataset yang berbeda, namun hanya memerlukan variabel-variabel tertentu saja dalam data. Ini bisa dilakukan dengan menggunakan perintah drop.

### **Command: drop Variabel1**

Hati-hati saat memutuskan untuk drop sebuah variabel. Variabel yang dihapus tidak dapat dikembalikan atau di-recover.

- m. **Do-file**, Do-file merupakan sebuah text file berisi perintah yang akan dijalankan oleh Stata. Do-file sangat berguna saat bekerja dengan data micro yang besar karena:
  - Membantu menyimpan semua perintah yang kita gunakan
  - Dapat melakukan penggantian perintah dengan mudah
  - Dapat menjalankan pengolahan data dan analisis di Stata tanpa harus ditunggu

## **6.1.3 MENYIMPAN DATA**

kumpulan data bekas ini dapat disimpan untuk digunakan di masa mendatang dengan klik **File** ---> **Save** atau **Ctrl + S** atau **Save As** atau **Ctrl + Shift + S**. Selain itu juga bisa menggunakan **save namafile** pada kolom **COMMAND**.

## **6.2 ANALISIS DATA: STATISTIK DESKRIPTIF MENGGUNAKAN STATA**

### **6.2.1 Analisis Statistik Deskriptif**

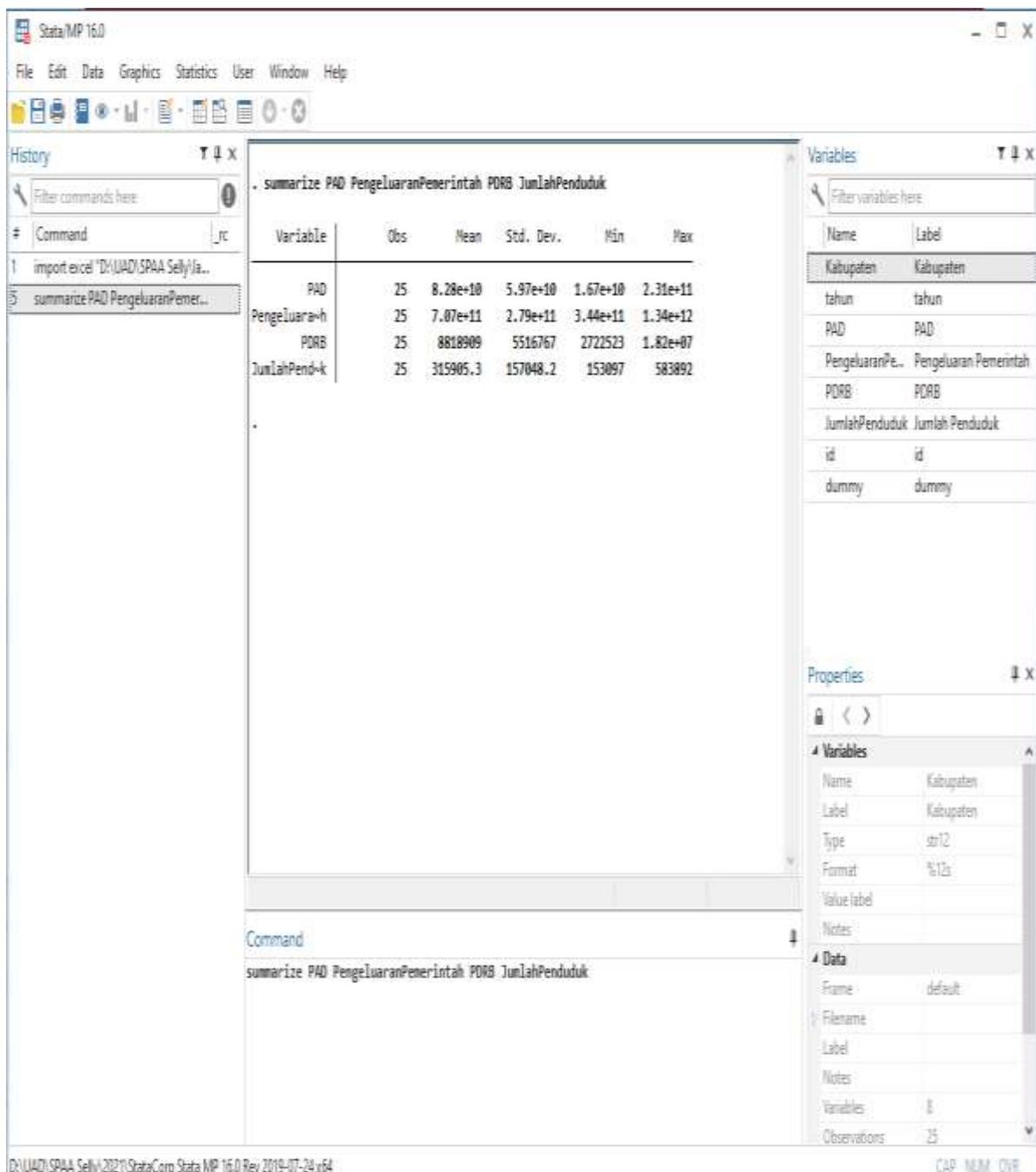
Pada analisa primer data, peneliti harus memahami data yang akan digunakan agar bisa menarik kesimpulan yang akurat dengan analisis statistik deskriptif. Dalam statistik deskriptif, hanya upaya yang akan dilakukan untuk menggambarkan, mempelajari dan menginformasikan tentang data. Umumnya, ada 2 kelompok utama, yaitu: Ukuran pemusatan dan Ukuran penyebaran

Sebagai perangkat lunak statistik, STATA tentunya mampu melakukan analisis deskriptif untuk membantu pengguna lebih memahami data. Perintah berikut digunakan untuk mengumpulkan statistik deskriptif pada STATA:

**Summarize**, digunakan dalam menampilkan nilai statistik deskriptif yang paling umum digunakan sebagai informasi konsentrasi awal dan distribusi data, termasuk mean, standar deviasi, nilai minimum yang diamati, dan nilai pengamatan maksimum.

**Langkah-langkah:**

1. Impor data file
2. Tulis di kolom command **Summarize Namavariabel** kemudian tekan enter



### Gambar 6.3 Penggunaan Summarize

Maka keluar hasil di kolom result seperti berikut. Copy dan kemudian paste ke dokumen yang akan digunakan sebagai lokasi hasil analisis.

```
. summarize PAD PengeluaranPemerintah PDRB JumlahPenduduk
```

Variable	Obs	Mean	Std. Dev.	Min	Max
PAD	25	8.28e+10	5.97e+10	1.67e+10	2.31e+11
Pengeluaran~h	25	7.07e+11	2.79e+11	3.44e+11	1.34e+12
PDRB	25	8818909	5516767	2722523	1.82e+07
JumlahPend~k	25	315905.3	157048.2	153097	583892

```
.
```

- b. Perintah untuk mengetahui detail deskriptif variabel maka gunakan perintah kolom command **summarize namavariabel, detail** kemudian tekan enter

. summarize PDRB , detail

PDRB				
	Percentiles	Smallest		
1%	2722523	2722523		
5%	2851123	2851123		
10%	2953612	2953612	Obs	25
25%	3245320	2981532	Sum of Wgt.	25
50%	9321038		Mean	8818909
		Largest	Std. Dev.	5516767
75%	1.21e+07	1.72e+07		
90%	1.78e+07	1.78e+07	Variance	3.04e+13
95%	1.79e+07	1.79e+07	Skewness	.3579977
99%	1.82e+07	1.82e+07	Kurtosis	1.8072

### 6.2.2 Penggunaan Tabel dan Grafik Statistika Deskriptif

Menggunakan tabel untuk menyajikan informasi memudahkan untuk membaca dan menarik kesimpulan. Grafik adalah gambaran mengenai proses naik turunnya suatu objek yang diperlihatkan melalui fluktuasi garis. Diagram adalah gambaran yang menerangkan sesuatu dan ditunjukkan dalam sebuah lingkaran. Histogram adalah representasi grafis dari tabel frekuensi yang ditampilkan sebagai diagram batang sebagai indikasi klasifikasi data. Setiap tampilan batang menampilkan proporsi frekuensi untuk setiap rangkaian kategori berurutan dalam interval yang tidak tumpang tindih. Dalam statistik, konsep histogram adalah representasi grafis (chart) yang mengatur dan menampilkan frekuensi data sampel dalam rentang tertentu.

Distribusi normal memiliki frekuensi nilai yang sama pada panel tengah di sisi kiri dan kanan. Keuntungan menggunakan histogram adalah menyediakan informasi tentang variabilitas proses dan membantu manajer membuat keputusan untuk terus meningkatkan proses (perbaikan proses berkelanjutan). Histogram digunakan untuk menunjukkan kepadatan

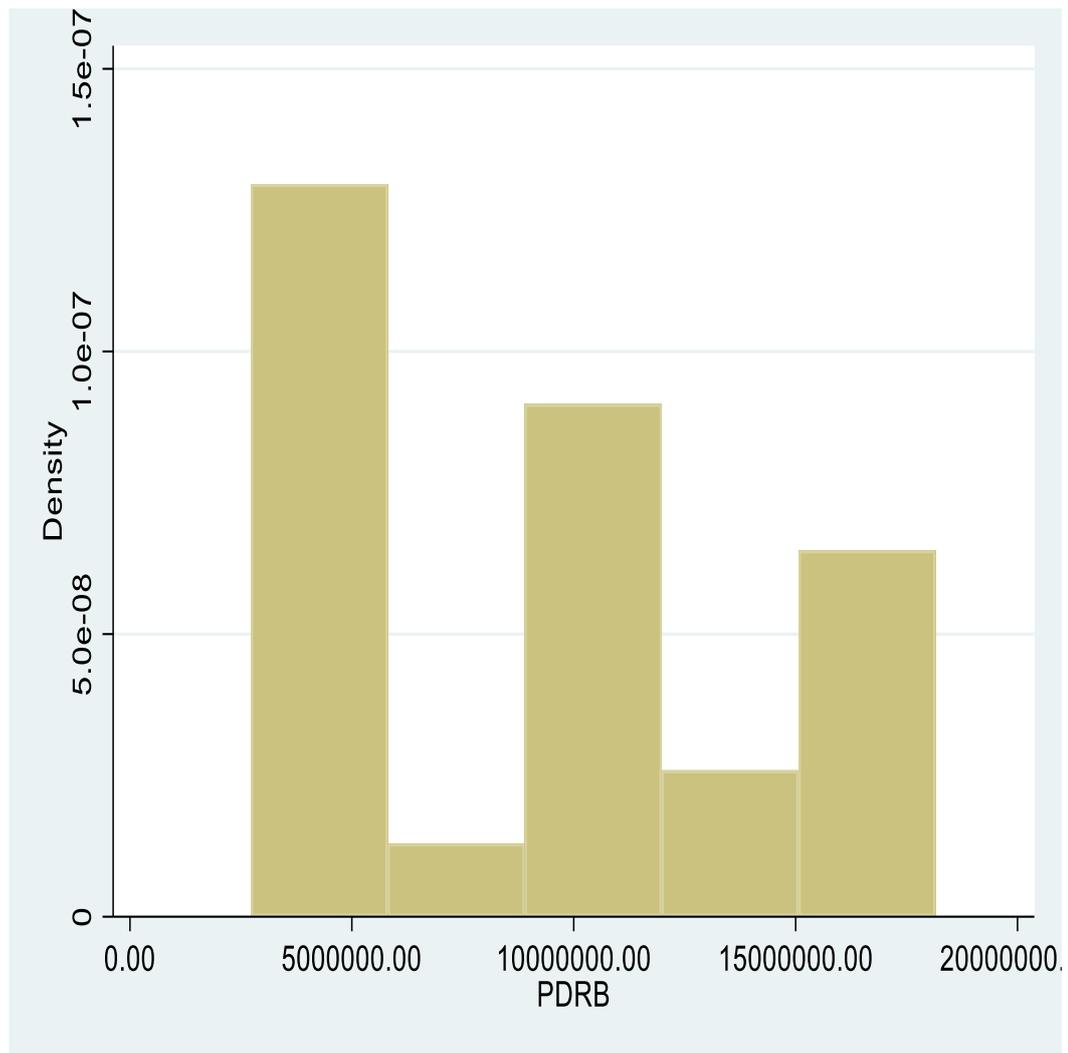
data dan sering digunakan untuk melakukan estimasi kepadatan (perkiraan fungsi kepadatan probabilitas dari variabel yang mendasarinya).

Kelemahan histogram adalah: Histogram hanya menampilkan data yang dikelompokkan berdasarkan kategori dan atribut. Histogram tidak mengizinkan pengguna untuk melihat tanggal setiap objek atau orang yang direkam.

Pengguna dapat menggunakan menu utama Graphic. Pada menu tersebut, terdapat banyak sekali pilihan jenis visualisasi statistika deskriptif sesuai kebutuhan pengguna, Selain menggunakan menu Graphics, tentunya STATA memiliki beberapa perintah yang dapat digunakan antara lain:

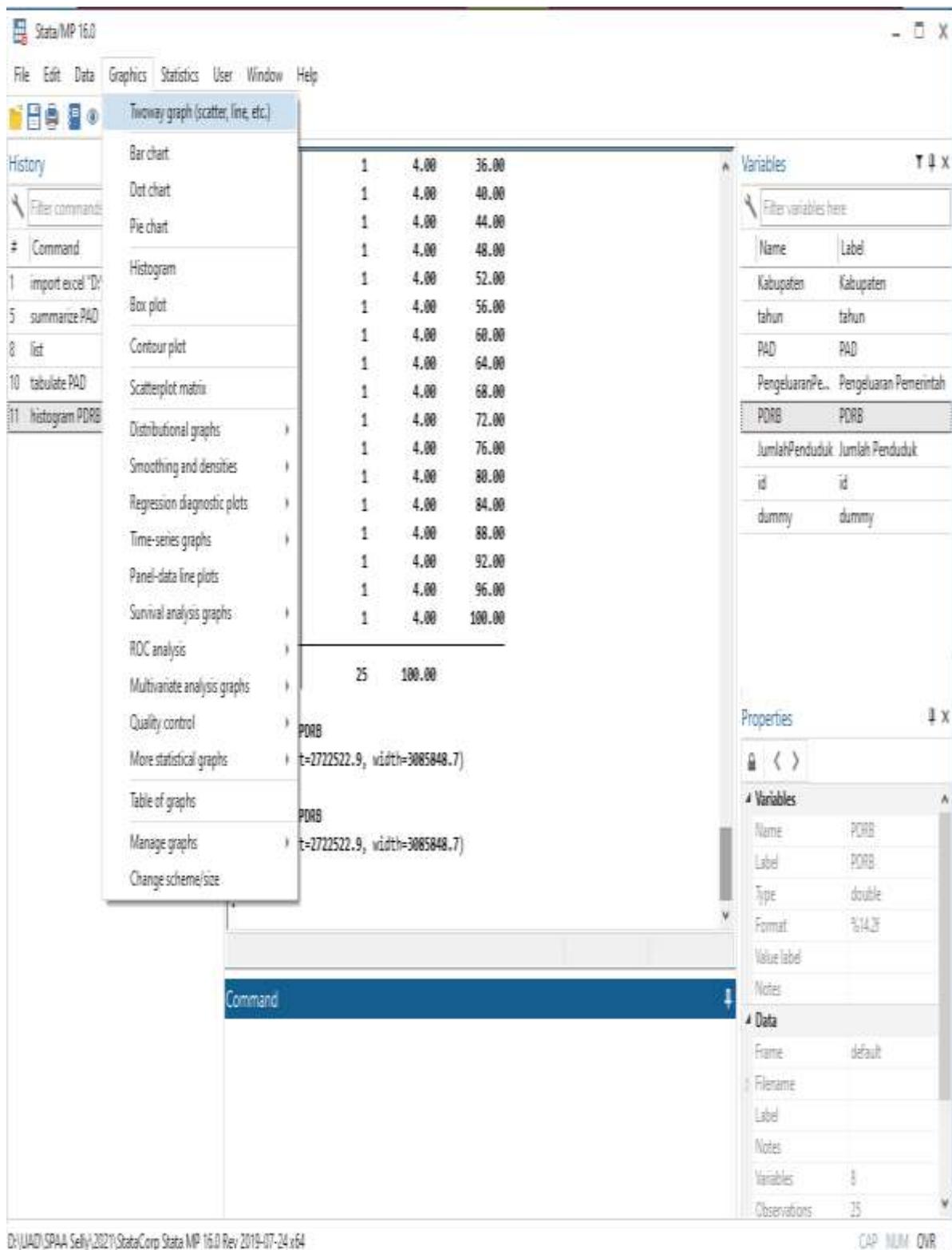
- a. Perintah/command histogram namavariabel dapat digunakan untuk menampilkan histogram suatu variabel. Opsi normal dapat dimasukkan bila menghendaki diperlihatkannya garis distribusi normal.

Sedangkan opsi frequency dapat digunakan untuk menunjukkan frekuensi setiap kelompok histogram. Perintah discrete dapat digunakan untuk mengetahui diagram batang untuk setiap nilai.



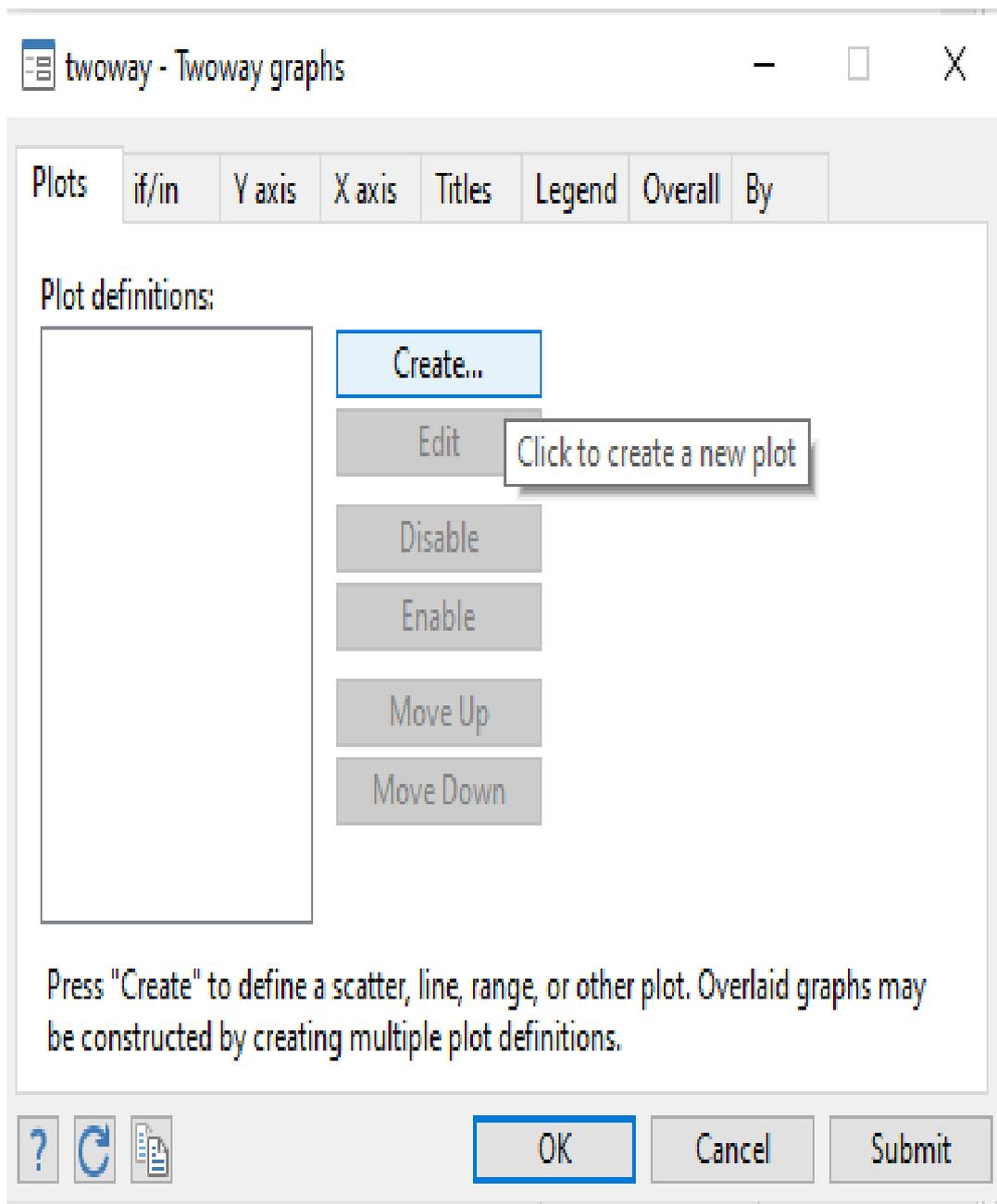
**Gambar 6.4 Contoh Histogram pada STATA**

- b.** Penggunaan grafik antarvariabel, dapat dengan perintah/command **twoway jenisplot namavariabel1 namavariabel2**. Atau klik **Graphic → twoway graph (scatter,line,etc)**.

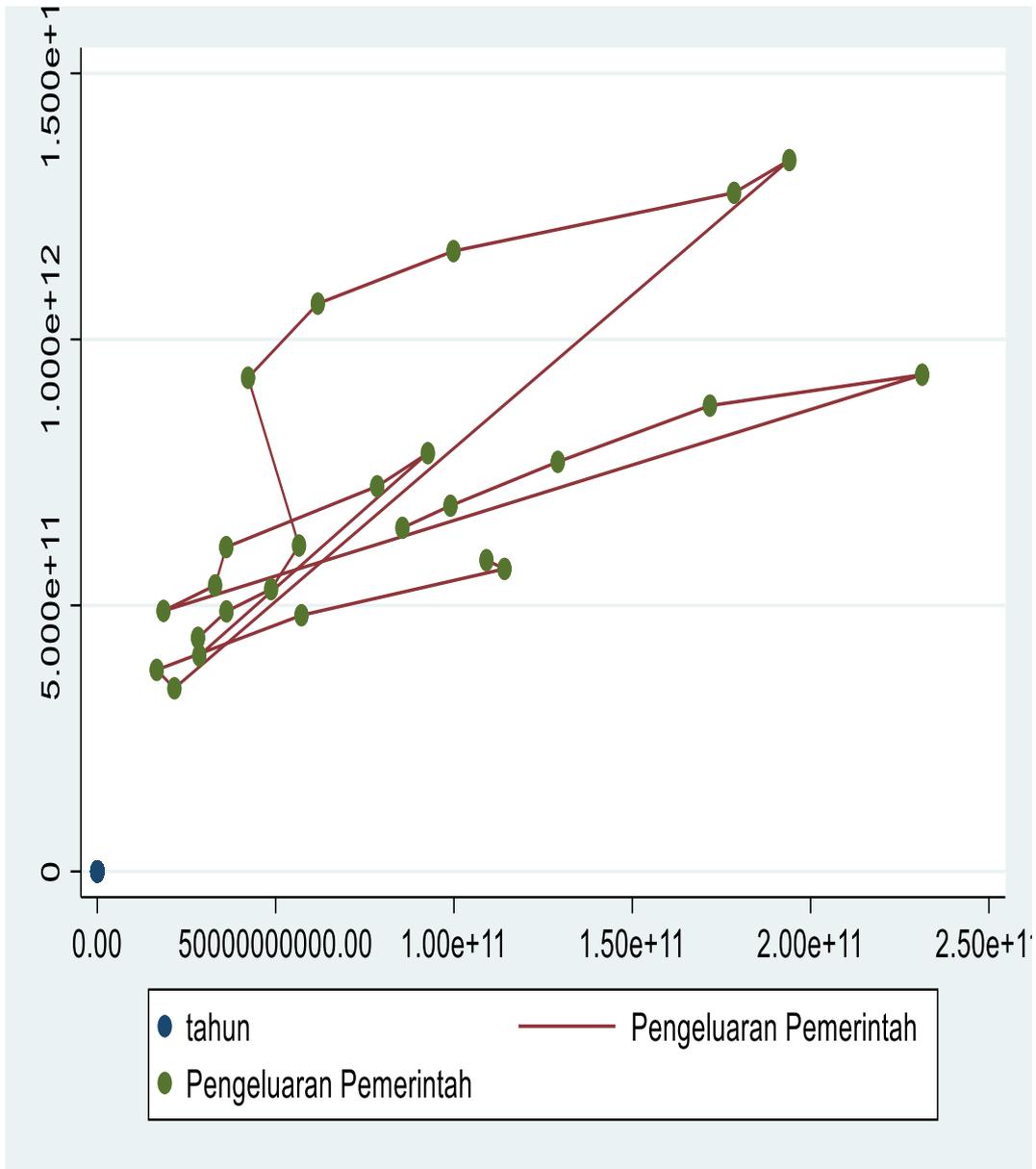


**Gambar 6.5 Langkah-langkah Pembuatan Grafik**

- c. Perintah grafik antarvariabel, **twoway jenisplot namavariabel1 namavariabel2**. Atau klik **Graphic** → **twoway graph (scatter,line,etc)**. Klik create lalu basic plots pilih (**scatter, line,dll**) dan tentukan Y dan X variabelnya. Klik **accept**.

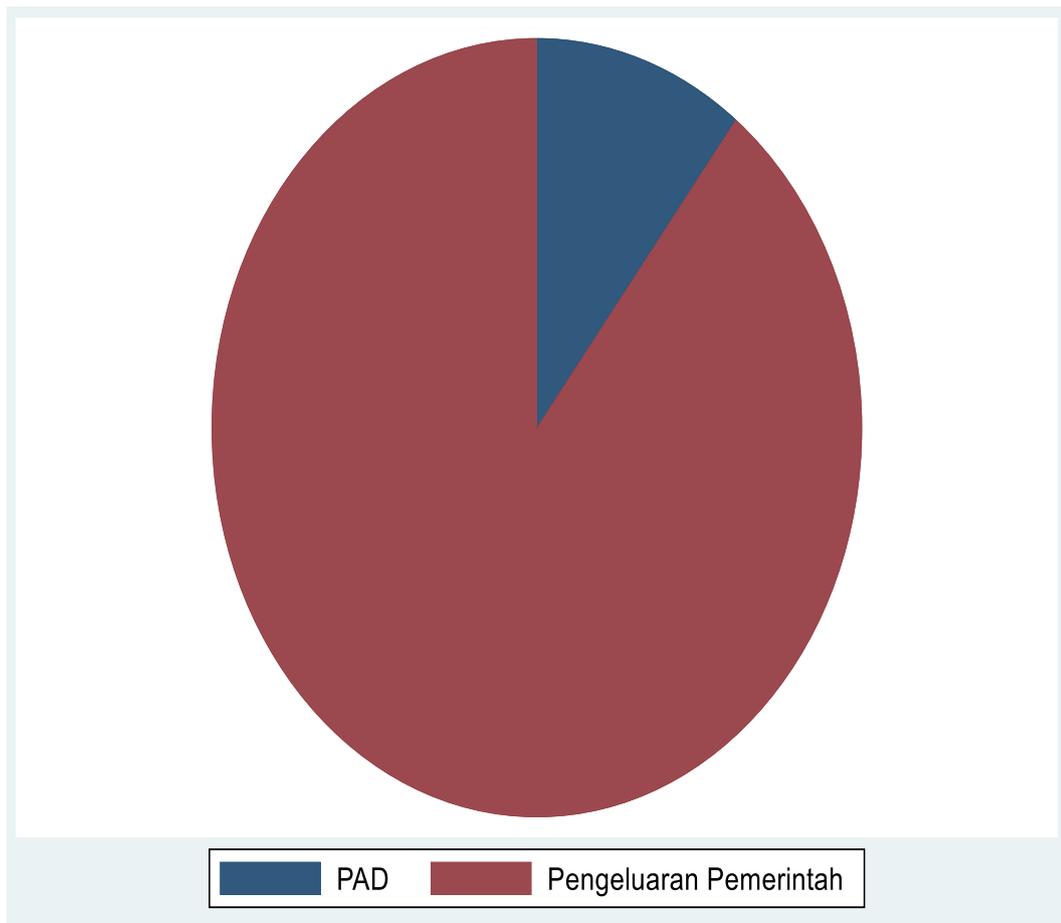


**Gambar 6.6 Twoway Graphs**



Gambar 6.7 Grafik Line

- d. Perintah memperlihatkan suatu grafik dengan **graph**. Misalnya untuk menunjukkan grafik pie. **Command graph pie namavariabel1 namavariabel2** atau **Klik Graphics → Pie chart.**



**Gambar 6.8 Grafik Pie**

### **6.3 ANALISIS REGRESI LINEAR MENGGUNAKAN STATA**

#### **6.3.1 Analisis Regresi**

Regresi adalah teknik statistik yang digunakan untuk memperkirakan hubungan antara variabel dependen dan satu atau lebih variabel independen. Analisis regresi dapat diterapkan di berbagai bidang seperti keuangan, investasi, dan pemasaran. Di sisi lain, hubungan antara variabel-variabel ini fungsional dan ditentukan oleh model matematika. Analisis regresi digunakan untuk membuat prediksi. Hair et al (1998) menyatakan bahwa jumlah sampel minimum untuk digunakan dalam teknik analisis regresi adalah 15-20 kali jumlah variabel.

#### **-Fungsi Regresi**

##### **1. Memprediksi masa depan**

Menganalisis apa yang akan terjadi di masa depan. Ini termasuk perkiraan risiko dan peluang. Fitur ini banyak digunakan dalam dunia bisnis. Sebagai contoh, katakanlah kita

ingin melakukan analisis permintaan produk di masa mendatang. Kira-kira berapa banyak produk yang dibeli konsumen? Hasilnya dapat ditentukan dengan metode regresi. Selain itu, perusahaan asuransi menggunakan regresi untuk memperkirakan kelayakan kredit pelanggan mereka dan memperkirakan jumlah klaim selama periode waktu tertentu.

## **2. Memperbaiki error**

Bahkan manajer yang paling berkomitmen dan kompeten pun dapat membuat kesalahan saat mengambil keputusan. Namun, masalah ini dapat diselesaikan dengan fungsi regresi yang dapat memperbaiki kesalahan tersebut. Misalnya, seorang manajer toko percaya bahwa memperpanjang jam buka toko akan meningkatkan penjualan. Setelah dilakukan perhitungan regresi, ternyata keputusan tersebut sangat membebani anggaran perusahaan. Oleh karena itu, fitur pencegahan kesalahan analisis regresi sangat berguna untuk administrator.

## **3. Memberikan wawasan baru**

Menjelajahi data memberikan informasi baru dan segar. Perusahaan perdagangan biasanya mengumpulkan data tentang pelanggannya. Namun, semua data tidak ada artinya tanpa analisis regresi yang tepat. Misalnya, jika kita ingin menguji data menggunakan analisis regresi, kita akan menemukan bahwa penjualan meningkat pada beberapa hari dalam seminggu dan menurun pada hari lainnya. Dalam hal ini, manajer dapat menyesuaikan remunerasi. Mulailah dengan menawarkan inventaris yang tepat setiap hari untuk mendapatkan bantuan lebih lanjut, atau pastikan memiliki pemasar dan produk terbaik yang tersedia hari itu.

## **4. Meningkatkan efisiensi operasional**

Perusahaan menggunakan regresi untuk mengoptimalkan operasi. Misalnya, manajer pabrik menghitung regresi untuk mempelajari pengaruh pemasaran pada pembelian produk klien.

### **- Kegunaan analisis regresi**

Kegunaan analisis regresi adalah untuk menentukan variabel kunci yang mempengaruhi variabel terikat, pemodelan, dan estimasi atau prediksi. Selain itu, ia memiliki beberapa kegunaan lain.

- Perkiraan mean dan nilai variabel dependen berdasarkan nilai variabel independen.
- Uji hipotesis untuk sifat-sifat dependen.
- Memprediksi nilai rata-rata variabel independen berdasarkan nilai variabel independen di luar rentang sampel.

Analisis regresi hanya berurusan dengan hubungan antara variabel terikat dan variabel bebas. Sedangkan analisis regresi berganda menggunakan lebih dari satu variabel bebas.

Persamaan regresi linier sederhana:

$$Y = a + bX + \varepsilon$$

Keterangan:

- Y = Variabel terikat yang akan diprediksi  
 X = Variabel bebas  
 a = konstanta  
 b = koefisien regresi.  
 $\varepsilon$  = error term

Berikut representasi matematis dari model regresi linear berganda:

$$Y = a + bX + cX + dX + \varepsilon$$

Keterangan:

- Y = Variabel terikat yang akan diprediksi  
 X = Variabel bebas  
 a = konstanta  
 b,c,d = koefisien regresi  
 $\varepsilon$  = error term

Gujarati dan Porter (2008) menyatakan penggunaan regresi linier sederhana wajib memenuhi kriteria sebagai berikut:

- a. Ukuran yang menghasilkan varians terendah (Best)

- b. Persamaan yang diperkirakan berbentuk Linear
- c. Estimasi ukuran yang Tidak Bias
- d. Ukuran yang dapat mengestimasi hasil akurat perhitungan populasi (Estimator)

Asumsi Model regresi linier sederhana:

1. Eksogenitas yang lemah
2. Linieritas
3. Varians error yang konstan
4. Autokorelasi untuk data time series

- **Asumsi klasik regresi linear berganda**

Seperti uji parametrik lainnya, regresi linier memiliki kondisi atau asumsi standar yang harus dipenuhi agar model prediksi yang dihasilkan adalah BLUE (Best Linear Unbiased Estimation). Asumsi klasik tentang regresi linier berganda meliputi:

**1. Data interval atau rasio**

Skala data untuk semua variabel, terutama variabel dependen, adalah interval atau rasio. Tidak perlu menguji hipotesis ini. Pastikan saja data yang digunakan adalah periode atau skala (numerik atau kuantitatif).

**2. Linearitas**

Variabel bebas berhubungan dengan variabel terikat. Hipotesis linier diuji dengan uji regresi linier, misalnya menggunakan kurva estimasi.

Untuk kurva estimasi, Anda dapat menguji nilai p linearitas untuk menentukan apakah ada hubungan linier. Jika  $p\text{-value} < 0,05$ , maka ada hubungan linier antara prediktor dan respon.

**3. Normalitas residual**

Residu adalah selisih antara prediksi  $y$  dan prediksi  $\hat{y}$ . Dalam hal ini,  $y$  adalah variabel dependen dan prediksi  $\hat{y}$  adalah hasil dari persamaan regresi yang dibangun. Oleh karena itu, residual dibentuk menggunakan rumus prediksi  $\hat{y} - y$ . Asumsi normalitas dalam regresi linier mengacu pada residual, bukan data untuk setiap variabel.

Pengujian normalitas regresi linier dapat diuji dengan menggunakan berbagai metode pengujian normalitas, antara lain: Uji Shapiro Wilk, Lilliefors atau Kolmogorov Smirnov,

Anderson Darling, Ryan Joiner, Shapiro Francia, Jarque Bera, uji kurtosis Skewness, dan berbagai jenis uji normalitas lainnya.

#### **4. Non outlier**

Outlier data dengan nilai yang sangat berbeda atau berbeda. Batas outlier tidak terlihat dari nilai absolut dari residual Studentized. Suatu sampel atau observasi dikatakan outlier jika nilai absolut dari residual Studentized melebihi 3.

#### **5. Homoskedastisitas**

Kebalikan dari homoskedastisitas adalah heteroskedastisitas. Model regresi linier berganda yang baik adalah model yang tidak memiliki kendala heteroskedastisitas.

Untuk menguji homoskedastisitas regresi linear berganda, dapat digunakan uji homoskedastisitas dari Glejser, uji Park, uji White, Spearman Heteroskedastisitas, dan masih banyak uji lainnya.

#### **6. Non Multikolinearitas**

Multikolinearitas adalah suatu kondisi ketika terdapat korelasi atau korelasi silang yang kuat antara variabel-variabel independen dalam model. Korelasi silang dikatakan ada jika korelasi antar variabel bebas dalam model regresi linier berganda  $> 0,8$ . Beberapa ahli menggunakan batas atas  $0,9$ . Metode lain yang lebih objektif adalah dengan menggunakan Variance Inflation Factor (VIF) dan Tolerance. Untuk nilai VIF  $> 10$  dan/atau toleransi  $< 10$ , atau multikolinearitas  $< 0,01$ . Berdasarkan pembahasan di atas, jelas bahwa asumsi multikolinearitas hanya ada pada regresi linier berganda dan tidak pada regresi linier sederhana. Hal ini karena regresi linier berganda memiliki banyak variabel bebas, sedangkan regresi linier sederhana hanya memiliki satu.

#### **7. Non Autokorelasi**

Autokorelasi dapat dipahami sebagai korelasi antar waktu. Oleh karena itu, mudah dipahami bahwa fenomena autokorelasi ini sering terjadi pada data time series atau regresi linier berganda dengan time series, dan jarang terjadi pada data cross-sectional. Uji autokorelasi ini dapat diuji dengan menggunakan nilai Durbin Watson (DW) dan uji *run*.

- **Kelebihan Regresi Berganda**

Keuntungan dari pendekatan regresi linier berganda antara lain generalisasi dan penggalan pola data tertentu, kemampuan untuk memperoleh pengetahuan bahkan ketika tidak ada yang pasti, dan kemampuan untuk melakukan perhitungan paralel, tetapi untuk membuat proses lebih singkat..

- **Kelemahan Analisis Regresi**

Keterbatasan yang paling menonjol adalah bahwa hasil prediksi dari analisis regresi adalah nilai estimasi, sehingga selalu ada kemungkinan tidak sesuai dengan data yang sebenarnya.

**-Syarat-Syarat Regresi**

Terdapat beberapa syarat dalam analisis regresi, yaitu:

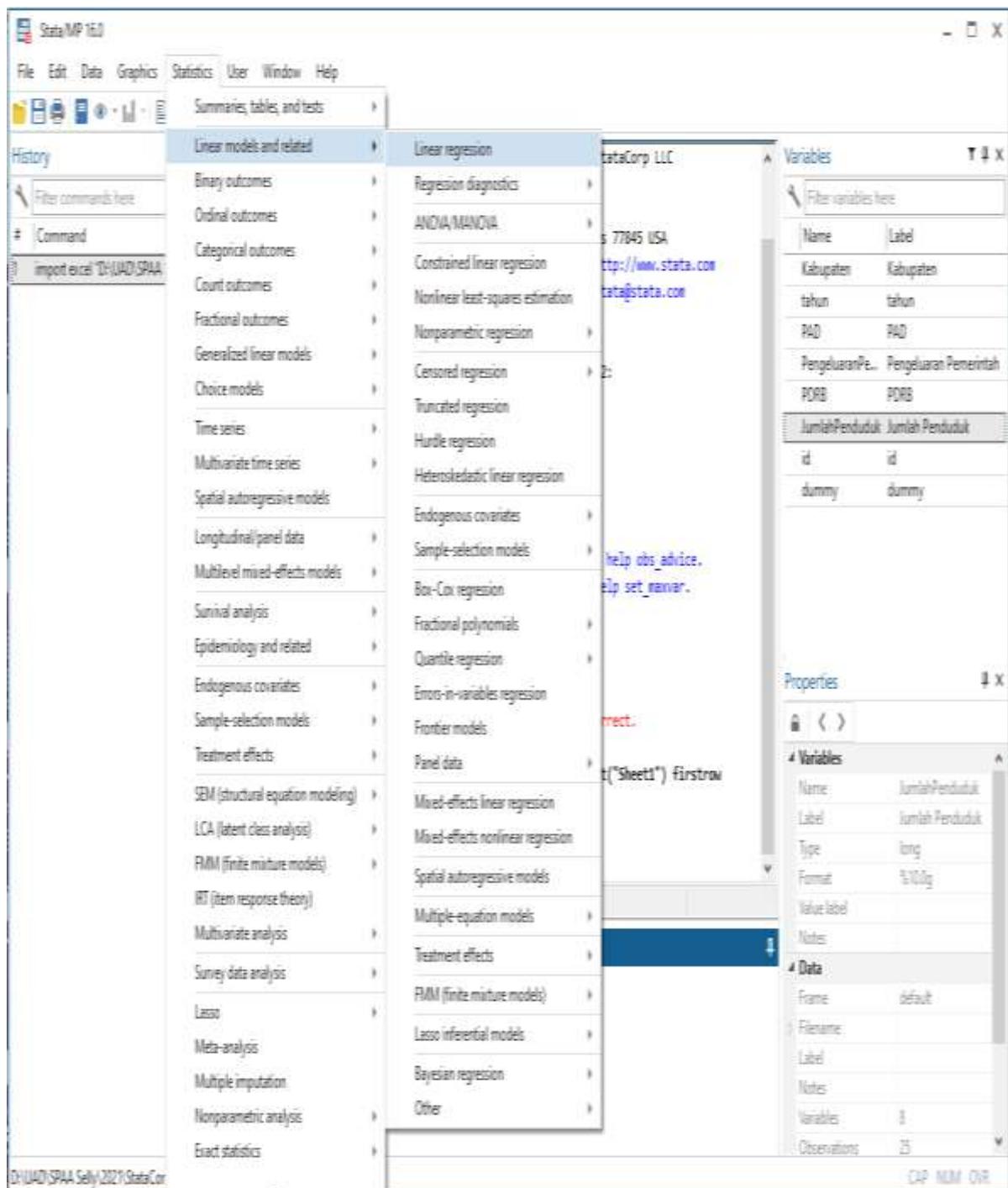
- Model regresi dikatakan layak jika jumlah signifikansi pada ANOVA adalah sebesar  $< 0.05$ .
- Prakiraan yang digunakan sebagai variabel bebas harus layak. Kelayakan ini diketahui jika angka Standard Error of Estimate  $<$  Standard Deviation.
- Koefisien regresi harus signifikan.

**6.3.2 Analisis Regresi Berganda Menggunakan STATA**

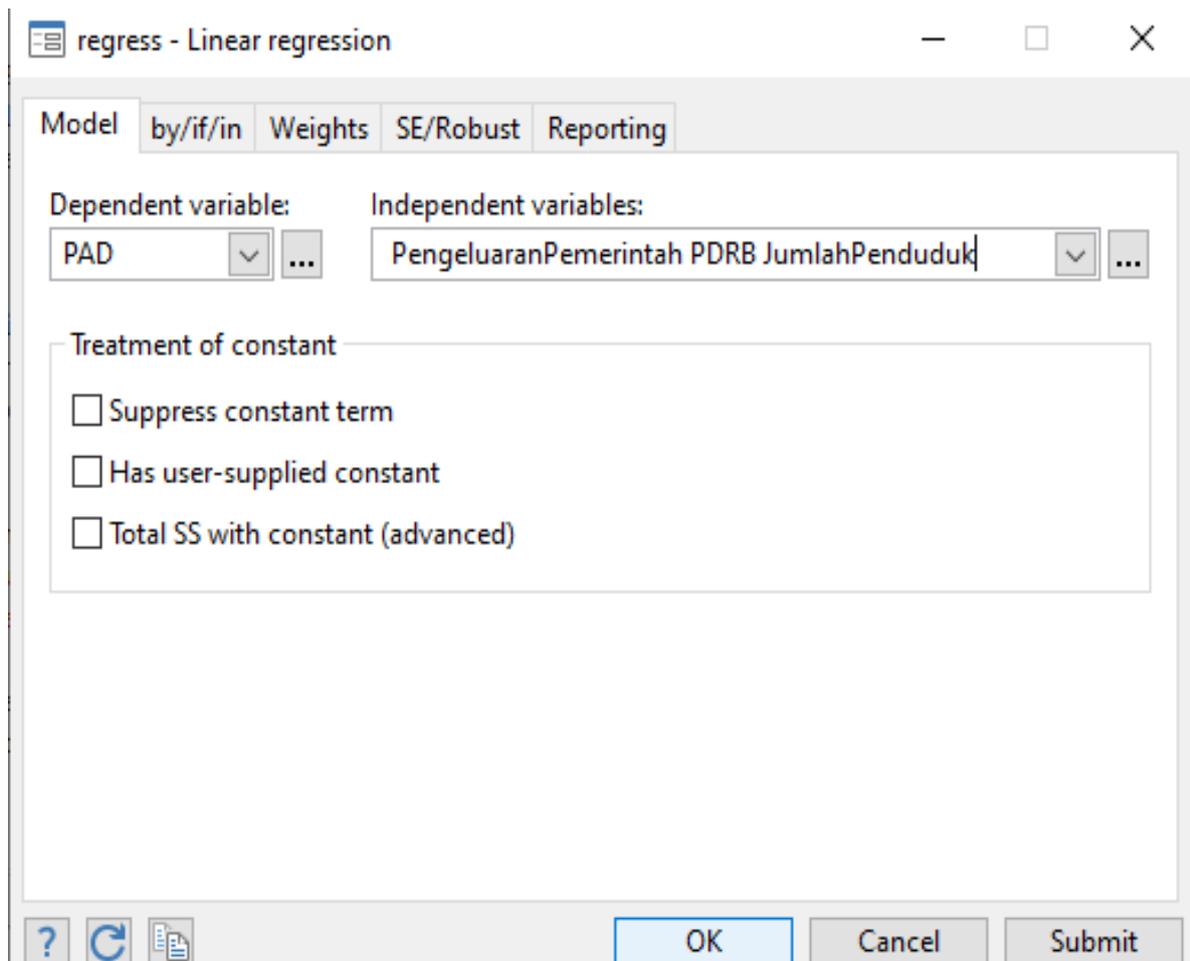
menu **STATISTICS ---> Linear Models and related ---> Linear regression.**

Penggunaan metode ini hanya perlu memilih dari daftar variabel milik masing-masing variabel terikat dan bebas untuk digunakan. Sementara itu, jika ingin menggunakan sintaks, gunakan sintaks umum/command untuk regresi pada stata yaitu:

**reg variabel\_dependen variabel\_independen**



**Gambar 6.9 Langkah-langkah Regresi linear STATA**



```
. regress PAD PengeluaranPemerintah PDRB JumlahPenduduk
```

Source	SS	df	MS	Number of obs	=	25
Model	7.4641e+22	3	2.4880e+22	F(3, 21)	=	47.80
Residual	1.0931e+22	21	5.2054e+20	Prob > F	=	0.0000
Total	8.5572e+22	24	3.5655e+21	R-squared	=	0.8723
				Adj R-squared	=	0.8540
				Root MSE	=	2.3e+10

PAD	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
PengeluaranPemer~h	.4078902	.0389426	10.47	0.000	.3269046 .4888758
PDRB	-5120.712	1373.895	-3.73	0.001	-7977.882 -2263.541
JumlahPenduduk	-383790.3	52058.72	-7.37	0.000	-492052.4 -275528.3
_cons	-3.91e+10	1.30e+10	-3.01	0.007	-6.61e+10 -1.21e+10

**Gambar 6.10 Hasil Output Regresi Linear STATA**

Hasil uji yang ditunjukkan pada gambar tersebut, informasi yang harus dipertimbangkan sebagai dasar:

- Menampilkan hasil uji global (F-test)
- Menunjukkan goodness of fit dari model
- Nilai  $\alpha$
- Menunjukkan koefisien estimasi ( untuk setiap variabel independen yang digunakan pada model.
- Menunjukkan standar deviasi dari koefisien estimasi
- Menunjukkan hasil pengujian individual (t-test).
- Alternatif untuk melakukan pengujian dalam mengetahui signifikansi suatu variabel.
- Menunjukkan jangkauan (range) pada setiap parameter estimasi.

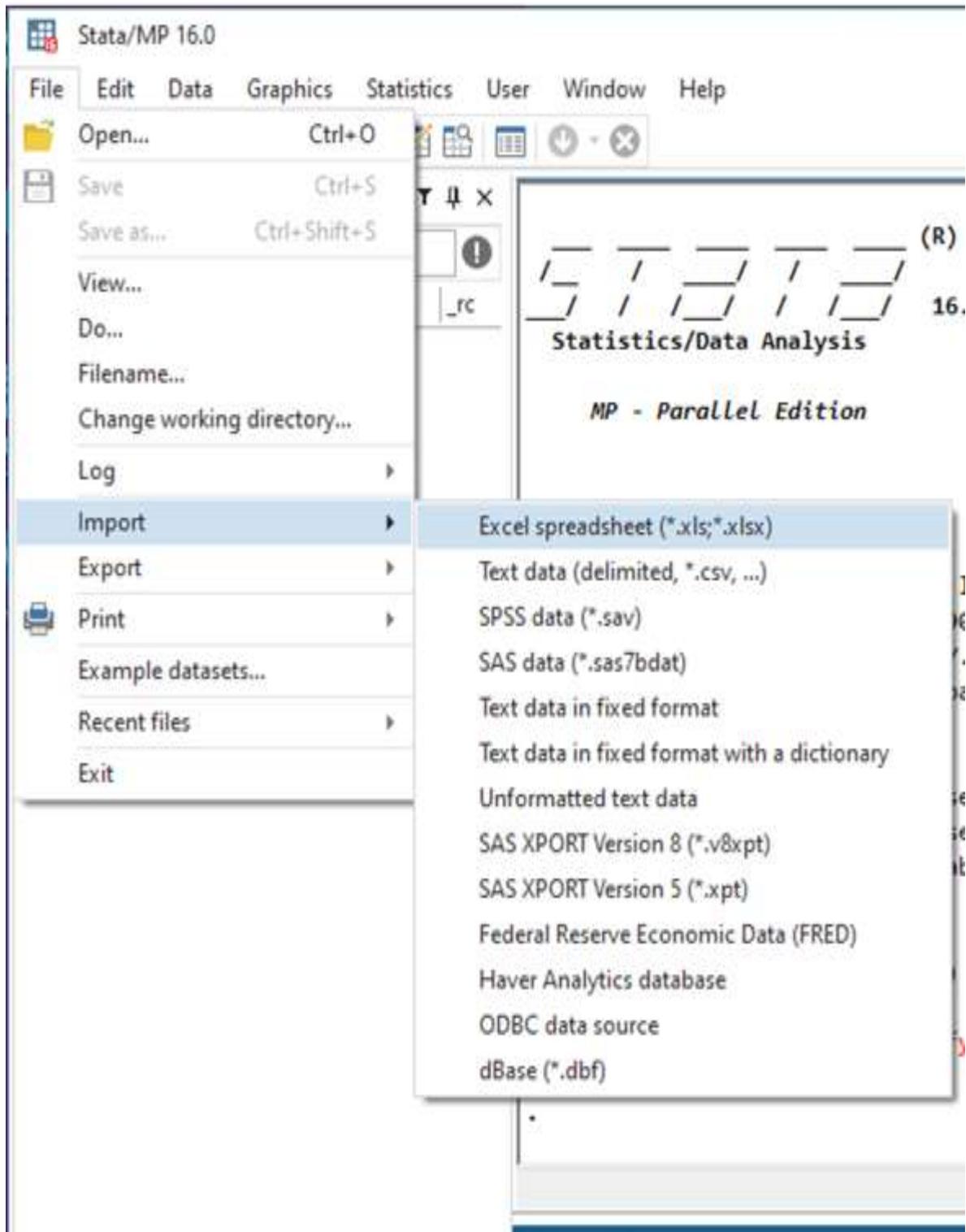
Dalam melakukan analisis regresi, hasil pengujian perlu memperhatikan tidak hanya makna variabel tetapi juga arah pengaruh masing-masing variabel.

#### **6.4 DUMMY VARIABLE**

Variabel dummy adalah variabel kategoris yang diasumsikan berpengaruh terhadap variabel kontinu. Variabel dummy juga dikenal sebagai variabel dummy, variabel biner, variabel kategoris, atau variabel dikotomis. Variabel dummy hanya memiliki dua (dua) nilai yaitu 1 dan 0, dan diberi simbol D. Variabel D memiliki nilai 1 ( $D = 1$ ) pada suatu kategori dan nilai 0 ( $D = 0$ ) pada direktori lain . Misalnya untuk kategori (pekerjaan),  $D = 1$ .  $D = 0$  (tidak bekerja) untuk item lainnya. Nilai 0 biasanya menunjukkan kelompok yang tidak diberi perlakuan dan 1 menunjukkan kelompok yang diberi perlakuan.

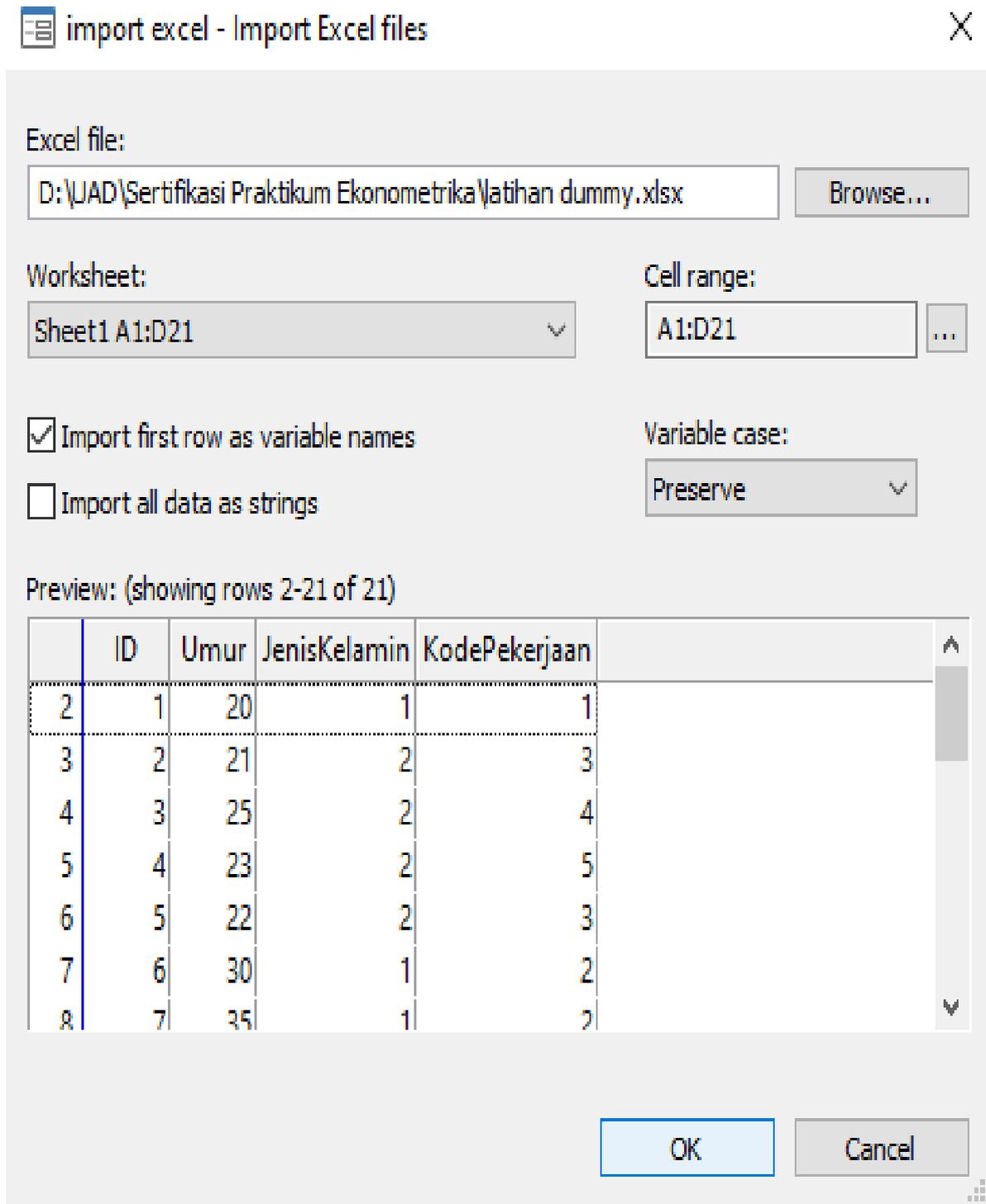
## LANGKAH-LANGKAH PADA STATA

### 1. Open Data / impor data latihan dummy



Gambar 6.11 Open Data

2. Klik Browse lalu pilih data yang akan di impor
3. Centang import first row as variables name
4. Klik ok



Gambar 6.12 Impor Data

Data sebelum dummy variable

	ID	Umur	JenisKelamin	KodePekerj~n
1	1	20	1	1
2	2	21	2	3
3	3	25	2	4
4	4	23	2	5
5	5	22	2	3
6	6	30	1	2
7	7	35	1	2
8	8	32	1	4
9	9	22	2	3
10	10	21	1	5
11	11	34	1	5
12	12	33	1	5
13	13	23	2	3
14	14	24	2	2
15	15	25	2	3
16	16	27	1	6
17	17	28	1	4
18	18	29	1	3
19	19	25	1	2
20	20	20	1	1

Gambar 6.13 Data Sebelum Dummy Variabel

Kita ingin membuat dummy pada variabel KodePekerjaan dimana:

**0= tidak bekerja; jika KodePekerjaan bernilai 1 dan 2**

**1= bekerja ; jika KodePekerjaan bernilai 3 hingga 6**

Langkah-langkah Command:

**gen D= KodePekerjaan**

**(membuat variabel baru bernama D)**

**recode D 1/2=0**

**recode D 3/6=1**

(menempatkan nilai 0 pada variabel D jika KodePekerjaan bernilai 1 atau KodePekerjaan bernilai 2)

Data setelah dummy variable ditambahkan

Data Editor (Edit) - [Untitled]

File Edit View Data Tools

ID[1] 1

	ID	Umur	JenisKelamin	KodePekerj~n	D
1	1	20	1	1	0
2	2	21	2	3	1
3	3	25	2	4	1
4	4	23	2	5	1
5	5	22	2	3	1
6	6	30	1	2	0
7	7	35	1	2	0
8	8	32	1	4	1
9	9	22	2	3	1
10	10	21	1	5	1
11	11	34	1	5	1
12	12	33	1	5	1
13	13	23	2	3	1
14	14	24	2	2	0
15	15	25	2	3	1
16	16	27	1	6	1
17	17	28	1	4	1
18	18	29	1	3	1
19	19	25	1	2	0
20	20	20	1	1	0

**Gambar 8.14 Data Setelah Dummy Variabel**

**8.5 Sebutkan fungsi regresi?**

1. Sebutkan kegunaan analisis regresi!
2. Apa saja kriteria regresi linear sederhana?
3. Sebutkan dan jelaskan kelebihan dan kelemahan analisis regresi!
4. Jelaskan mengenai variabel dummy?

