Research Data Analysis using SPSS

By
Dr.Anura Karunarathne
Senior Lecturer,
Department of Accountancy
University of Kelaniya

MBA 61013- Business Statistics and Research Methodology

Learning outcomes

At the end of the course unit, the student should be able to

- use statistics as a tool to collect, analyze and interpret data related to business,

- use statistical technique for business decisions,

- solve problems in the field of business with the use of **statistical software packages**

Why a Person/Manager Needs to Know about Statistics ?

- To know how to properly present information

- To know how to draw conclusions about populations based on sample information

- To know how to improve processes

- To know how to obtain reliable forecasts

Types of Statistics

Descriptive Statistics: Methods of organizing, summarizing, and presenting data in an informative way. This type of statistics allows researchers to summarize large quantities of data using measures that are easily understood by an observer.

○**EXAMPLE 1:** According to Consumer Reports, General Electric washing machine owners reported 9 problems per 100 machines during 2010. The statistic 9 describes the number of problems out of every 100 machines.

Types of Statistics......

Inferential Statistics

Inferential Statistics is body of methods used to draw conclusions or inferences about characteristics of populations based on information obtained from a sample taken from the same population.

Types of Statistics

Inferential Statistics: A decision, estimate, prediction, or generalization about a **population,** based on a **sample.**

A **population** is a **collection** of all possible individuals, objects, or measurements of interest.

A **sample** is a portion, or part, of the population of interest

SPSS Lecture Programme

Part-01

Introduction about the basics of SPSS and How to manage SPSS database

- Designing a SPSS data file based on questionnaire and data feeding.
- Data Presentation

Research Data Analysis using SPSS

- Measuring Basic Descriptive Statistical Measures
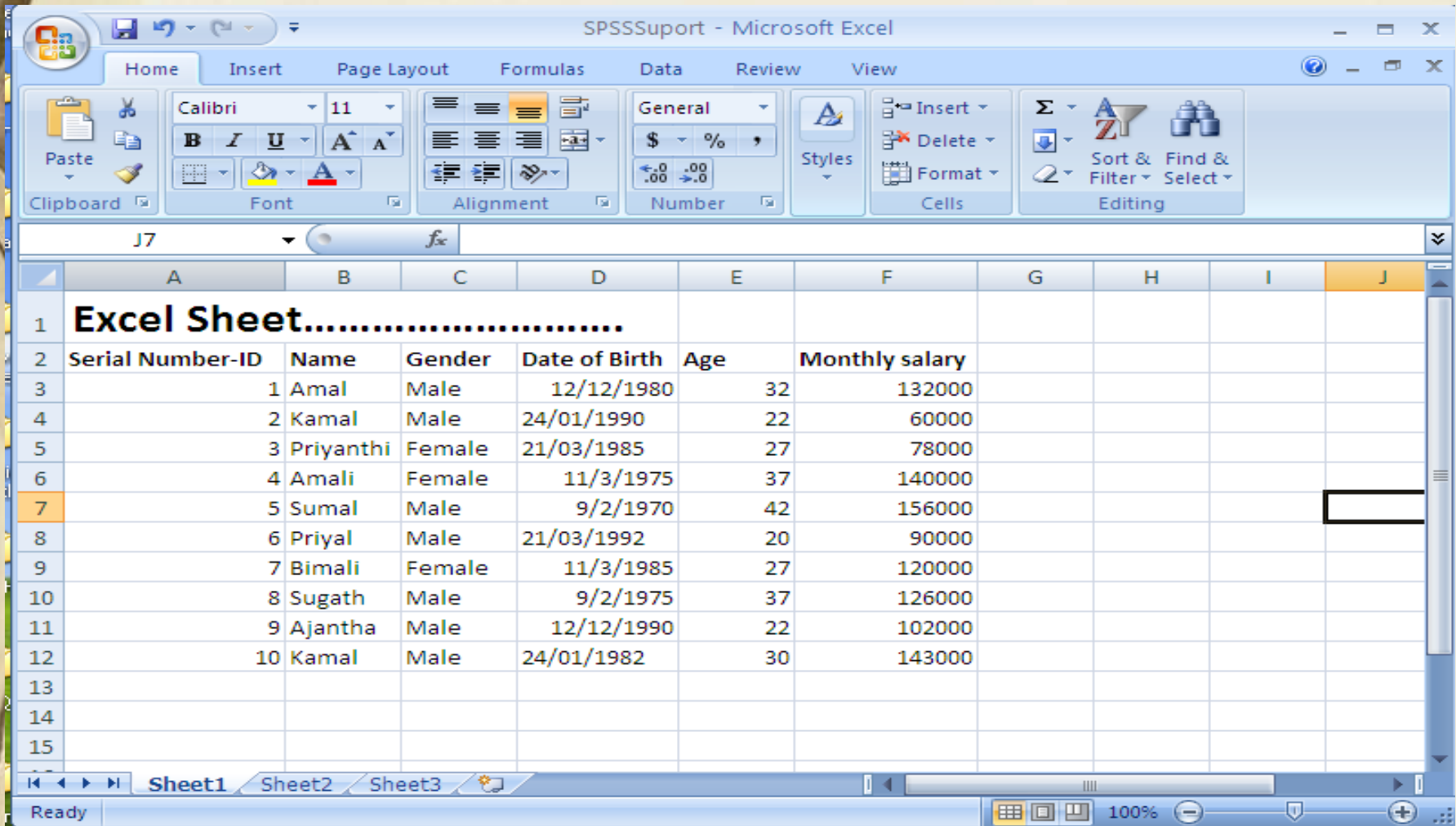- Regression and Correlation Analysis

Part-02

Statistical Tests

- Parametric Test
- Non-parametric Tests

Introduction: What is SPSS or PASW?

- One of the most popular **user friendly** statistical package which can perform highly complex data manipulation and analysis with simple instructions

- Originally it is an acronym of **S**tatistical **P**ackage for the **S**ocial **S**cience **(SPSS)** Then, it stands for **S**tatistical **P**roduct and **S**ervice **S**olutions **(SPSS)**

- Now…**P**redictive **A**nalytics **S**oftware **(PASW)**

Understanding the Difference between Excel and SPSS worksheet...............
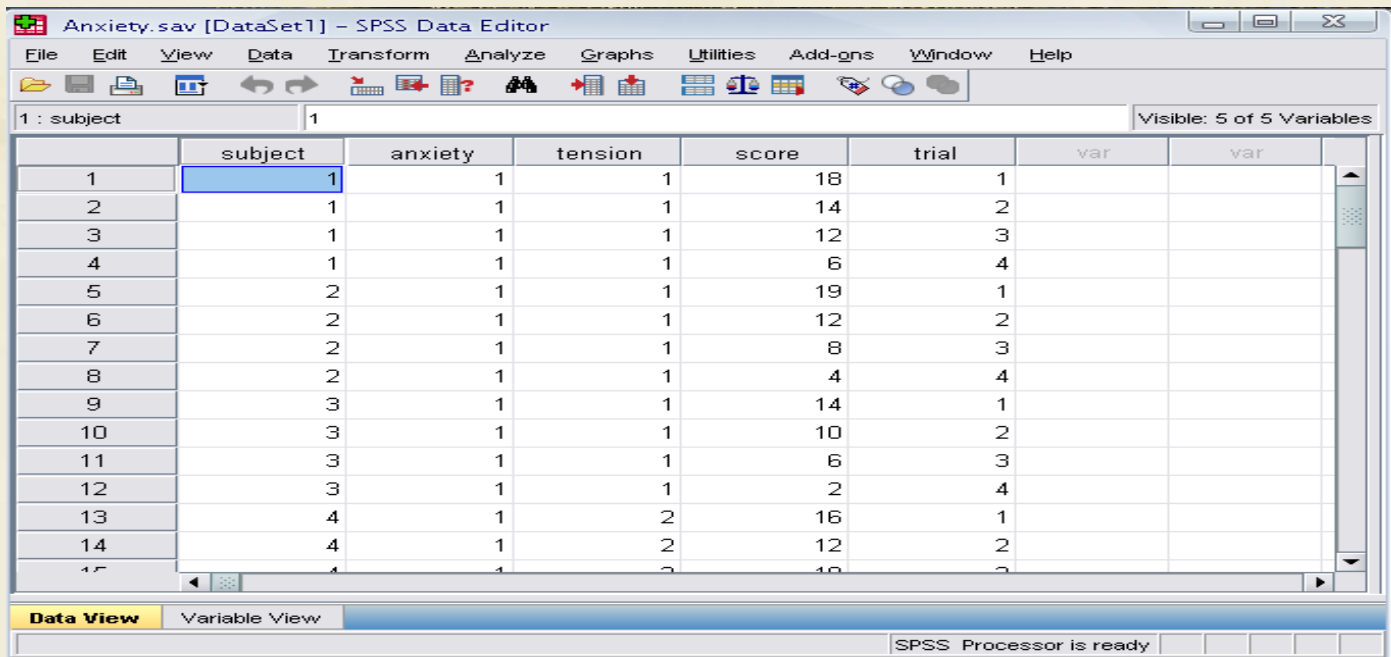
Excel Sheet...........................

Serial Number-ID	Name	Gender	Date of Birth	Age	Monthly salary
1	Amal	Male	12/12/1980	32	132000
2	Kamal	Male	24/01/1990	22	60000
3	Priyanthi	Female	21/03/1985	27	78000
4	Amali	Female	11/3/1975	37	140000
5	Sumal	Male	9/2/1970	42	156000
6	Priyal	Male	21/03/1992	20	90000
7	Bimali	Female	11/3/1985	27	120000
8	Sugath	Male	9/2/1975	37	126000
9	Ajantha	Male	12/12/1990	22	102000
10	Kamal	Male	24/01/1982	30	143000

The Four Windows: Data Editor

- **Data Editor**

  Spreadsheet-like system for defining, entering, editing, and displaying data. Extension of the saved file will be "sav."

Object of the Course

- About the four-windows in SPSS

- The basics of managing data files

- The basic analysis in SPSS

The Basics of Managing Data Files

The Four Windows

Data editor
Chart editor
Output viewer
Syntax editor

The Four Windows: Data Editor

- ## Data Editor

  Spreadsheet-like system for defining, entering, editing, and displaying data. Extension of the saved file will be "sav."

The Four Windows: Output Viewer

- ## Output Viewer
  Displays output and errors. Extension of the saved file will be "spv."
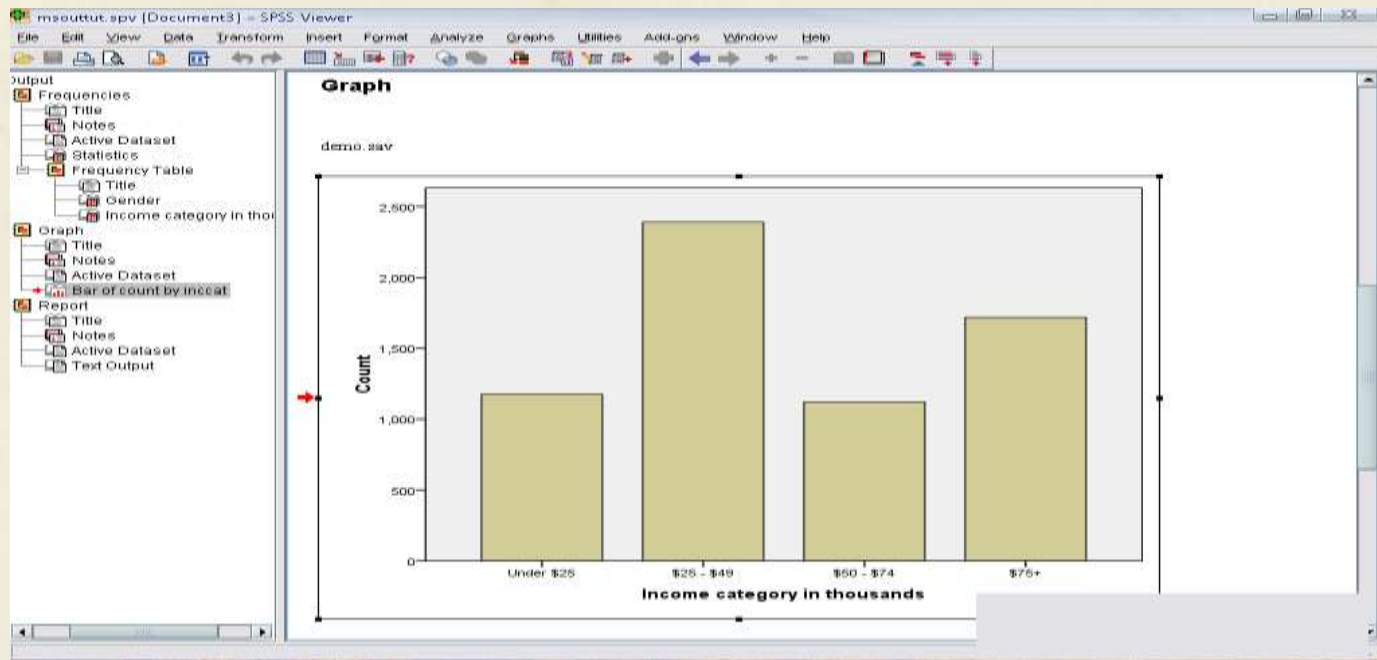
The Four Windows: Syntax editor

- ## Syntax Editor
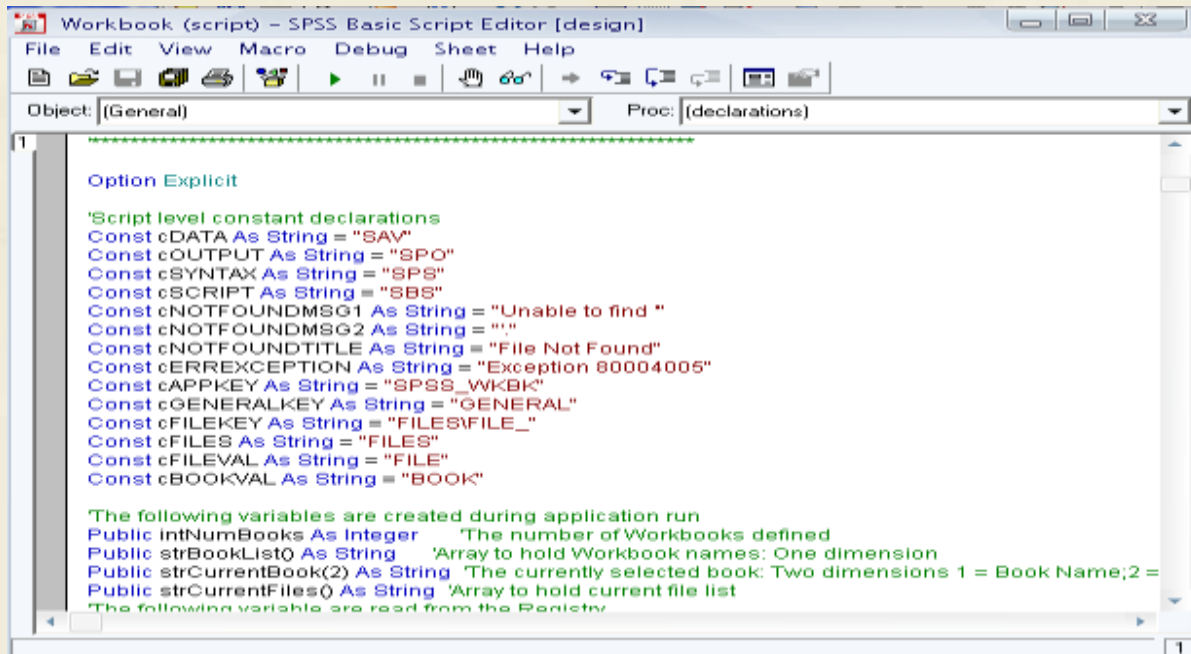
  Text editor for syntax composition. Extension of the saved file will be "sps."
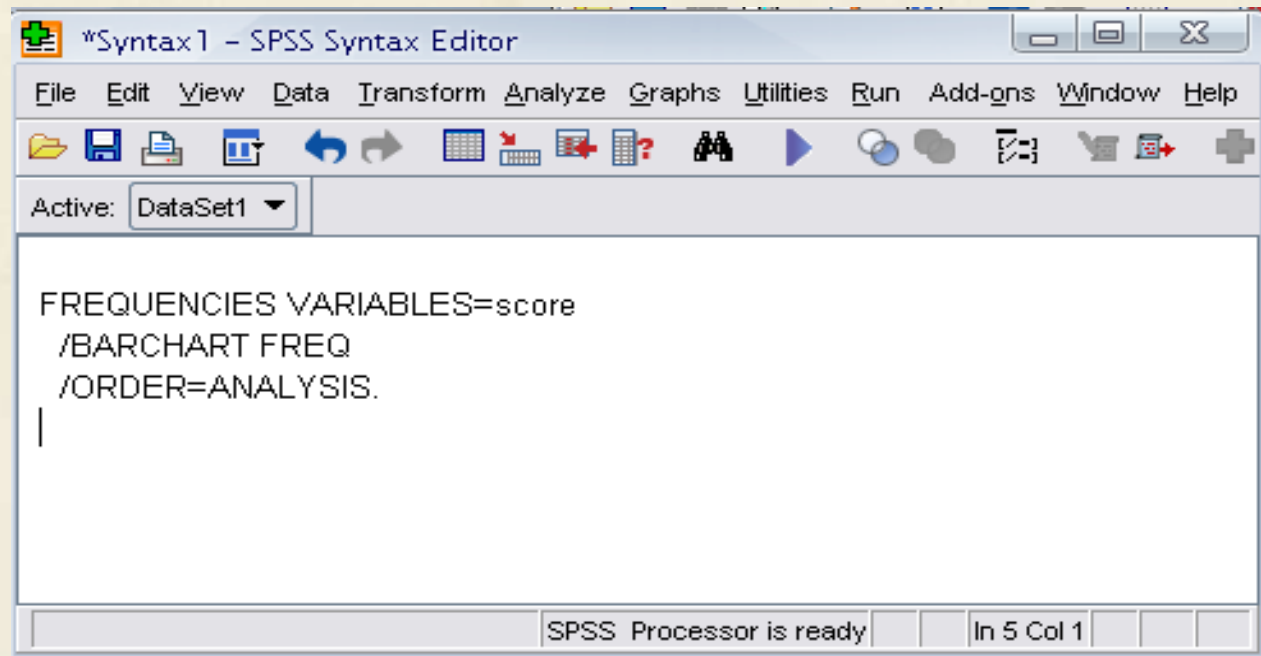
Script Window

- ## Script Window

  Provides the opportunity to write full-blown programs, in a BASIC-like language. Text editor for syntax composition. Extension of the saved file will be "sbs."

The Basics of Managing Data Files

SPSS Data Entry

- **SPSS data** can be entered manually.
  - The format is ready for analysis.
- **SAS, Excel, txt, etc.** data can be easily imported to SPSS.
- **SPSS data files** are saved as "SPSS data document **(.sav)**"
- **SPSS output files** are saved as "SPSS viewer document **(.spv)**"

Opening SPSS

- Start → All Programs → SPSS Inc→ SPSS 16.0 → SPSS 16.0

Opening SPSS

- The default window will have the data editor
- There are two sheets in the window:

  1. Data view   2. Variable view

Data View window

- The Data View window

  This sheet is visible when you first open the Data Editor and this sheet contains the data

- Click on the tab labeled Variable View
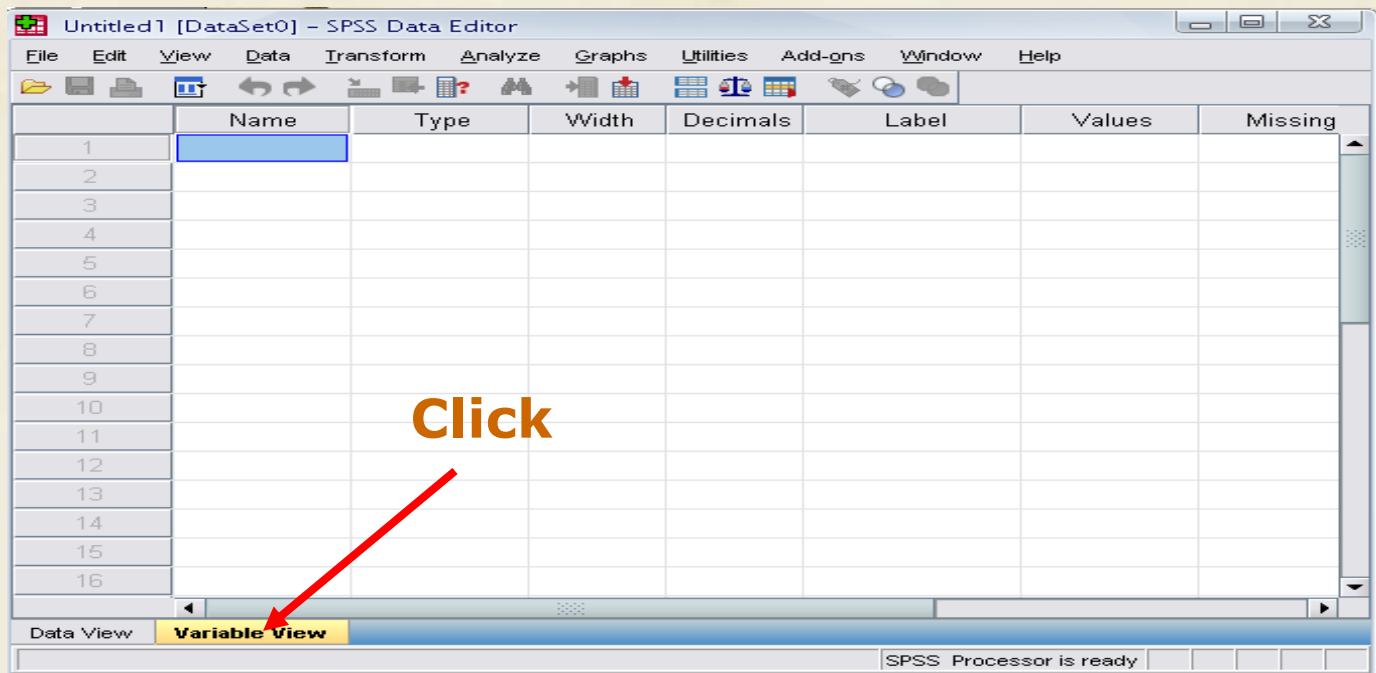
Variable View window
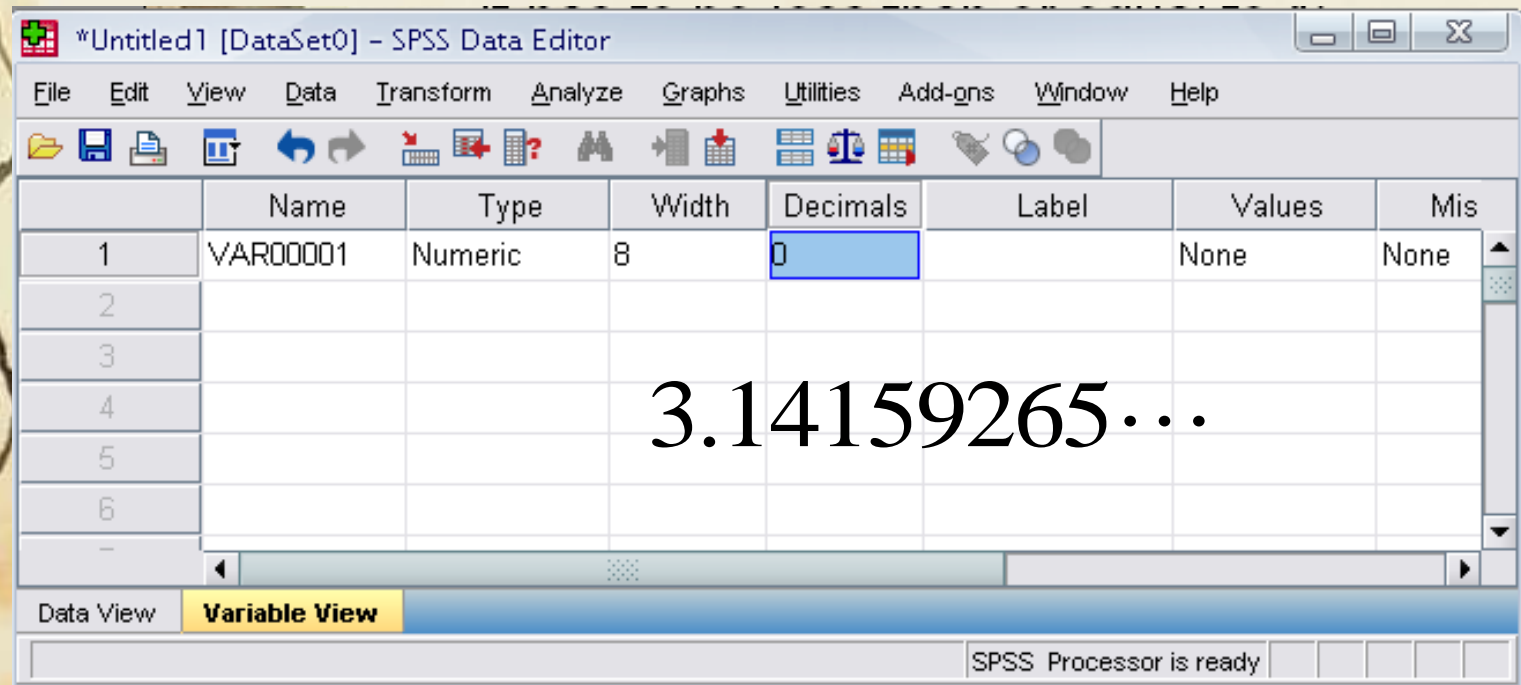
- This sheet contains information about the data set that is stored with the dataset

- **Name**
  - The first character of the variable name must be alphabetic
  - Variable names must be unique, and have to be less than 64 characters.
  - Spaces are NOT allowed.

Variable View window: Type

- **Type**
  - Click on the 'type' box. The two basic types of variables that you will use are numeric and string. This column enables you to specify the type of variable.

Variable View window: Width

- **Width**
  - Width allows you to determine the number of characters SPSS will allow to be entered for the variable
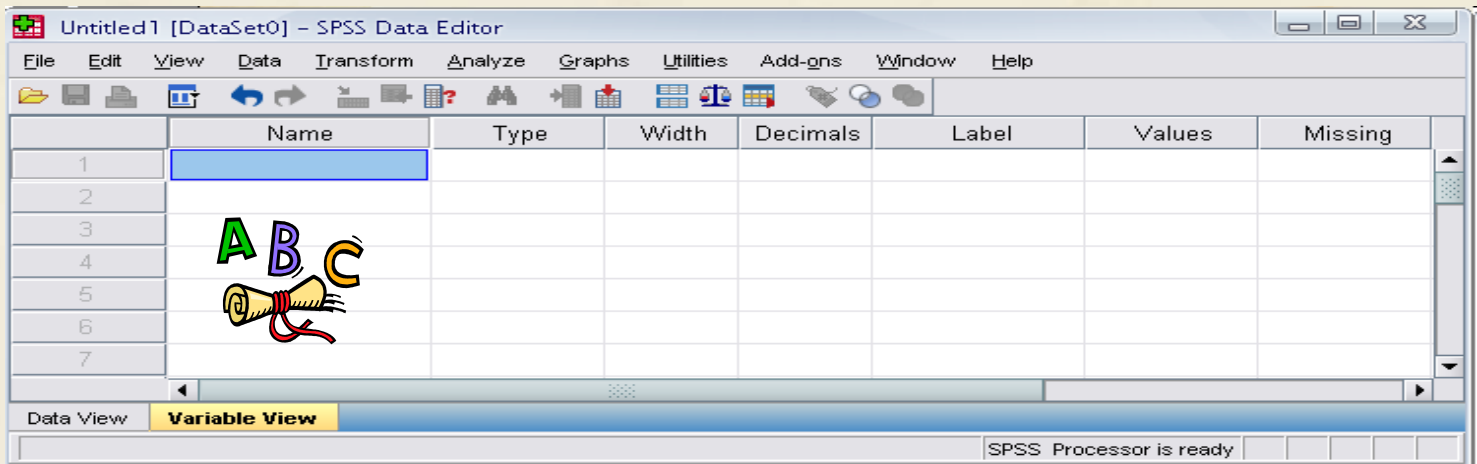
Variable View window: Decimals

- ## Decimals
  - Number of decimals
  - It has to be less than or equal to 16

Variable View window: Label

- **Label**
  - You can specify the details of the variable
  - You can write characters with spaces up to 256 characters
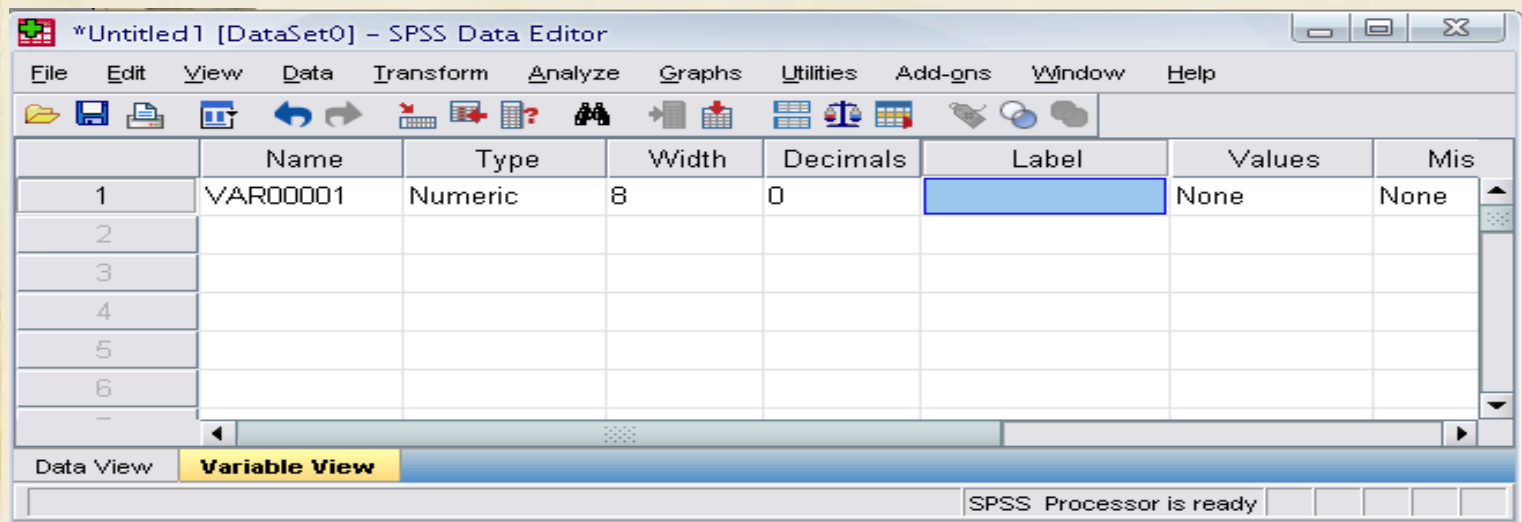
Variable View window: Values

- ## Values

  - This is used and to suggest which numbers represent which categories when the variable represents a category
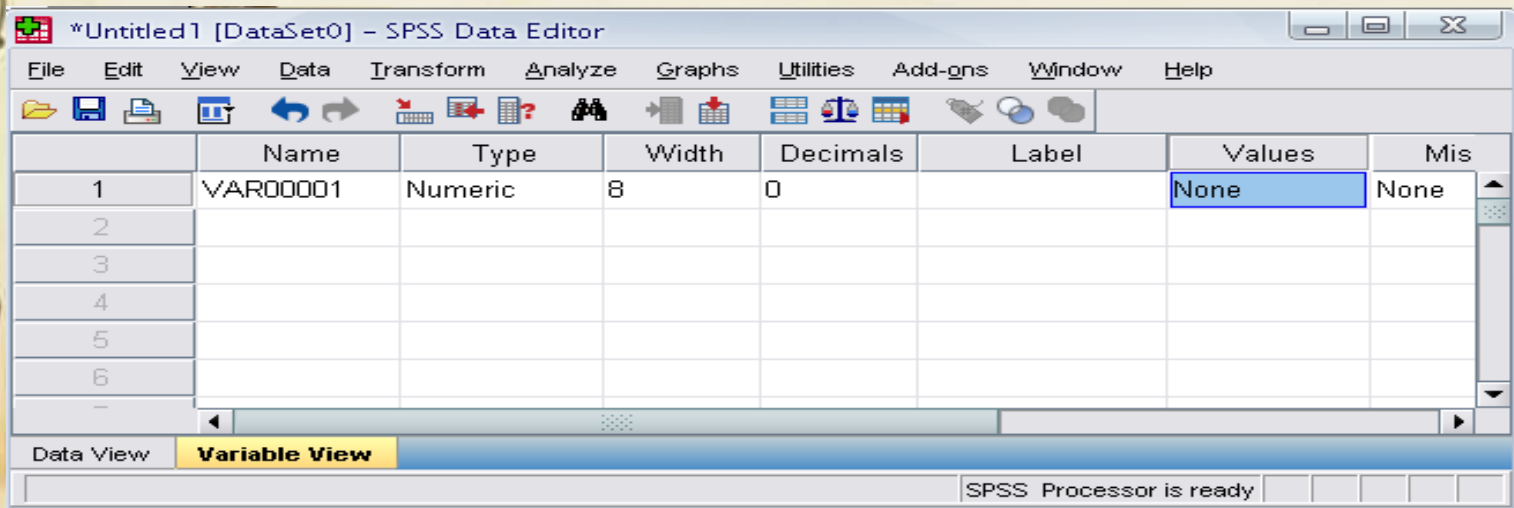
Defining the value labels

- Click the cell in the values column as shown below
- For the value, and the label, you can put up to 60 characters.
- After defining the values click add and then click OK.

Click

Data Types
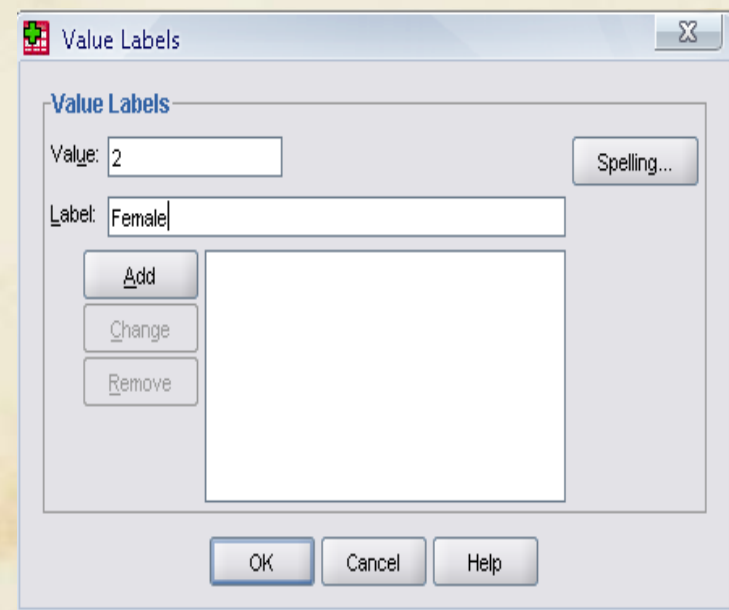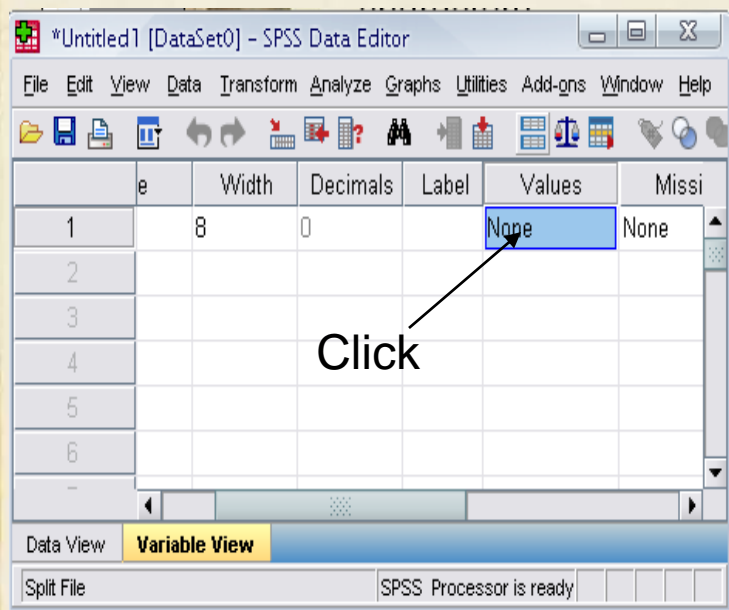
```
                          ┌─────────────┐
                          │    Data     │
                          └──────┬──────┘
              ┌──────────────────┴──────────────────┐
      ┌───────────────┐                    ┌──────────────────┐
      │  Qualitative  │                    │   Quantitative   │
      │ (Categorical) │                    │    (Numerical)   │
      └───────────────┘                    └────────┬─────────┘
                                         ┌───────────┴───────────┐
                                  ┌────────────┐          ┌──────────────┐
                                  │  Discrete  │          │  Continuous  │
                                  └────────────┘          └──────────────┘
```
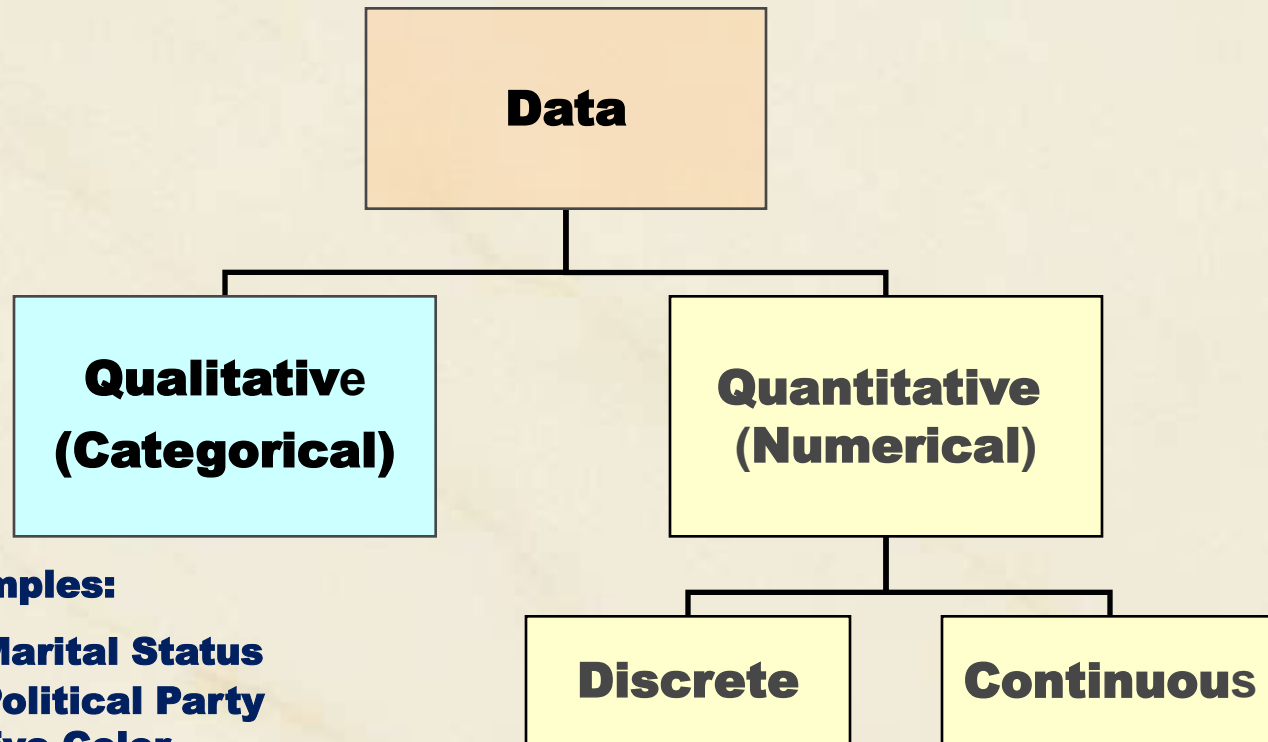
Examples:

- **Marital Status**
- **Political Party**
- **Eye Color**
(Defined categories)

Examples:

- **Number of Children**
- **Defects per hour**
(Counted items)

Examples:

- **Weight**
- **Voltage**
(Measured characteristics)

Format of variables

- You can also select format of variable

Categorical

Continuous

Binary Variable

Interval Variable

Nominal Variable

Ratio Variable

Ordinal Variable

Types of Scales

- **Nominal-** objects or people are categorized according to some criterion (gender, job category)
- **Ordinal-** Categories which are ranked according to characteristics (income- low, moderate, high)
- **Interval-** contain equal distance between units of measure- but no zero (calendar years, temperature)
- **Ratio-** has an absolute zero and consistent intervals (distance, weight)

Categorical (Nominal)

- A categorical variable is one that has two or more categories, but there is no intrinsic ordering to the categories.

eg.

Gender

Department/Division

Ordinal Variable

- An **ordinal variable** is similar to a categorical nominalvariable.

- The difference between the two is that **ordinal variable** has a clear rank/order.

  – SES (Socio Economic Status)

  – Education

- Even though we can order these from lowest to highest, the spacing between the values may not be the same across the levels of the variables.

Levels of Measurement

There are four levels of data.

Nominal level: Data that is classified into categories and cannot be arranged in any particular order.

o**EXAMPLES:** eye color, gender, religious affiliation.

Levels of Measurement

Ordinal level: involves data arranged in some order, but the differences between data values cannot be determined or are meaningless.

○ EXAMPLE: During a taste test of 4 soft drinks, Mellow Yellow was ranked number 1, Sprite number 2, Seven-up number 3, and Orange Crush number 4.

Levels of Measurement

Interval level: similar to the ordinal level, with the additional property that meaningful amounts of differences between data values can be determined. There is no natural zero point.

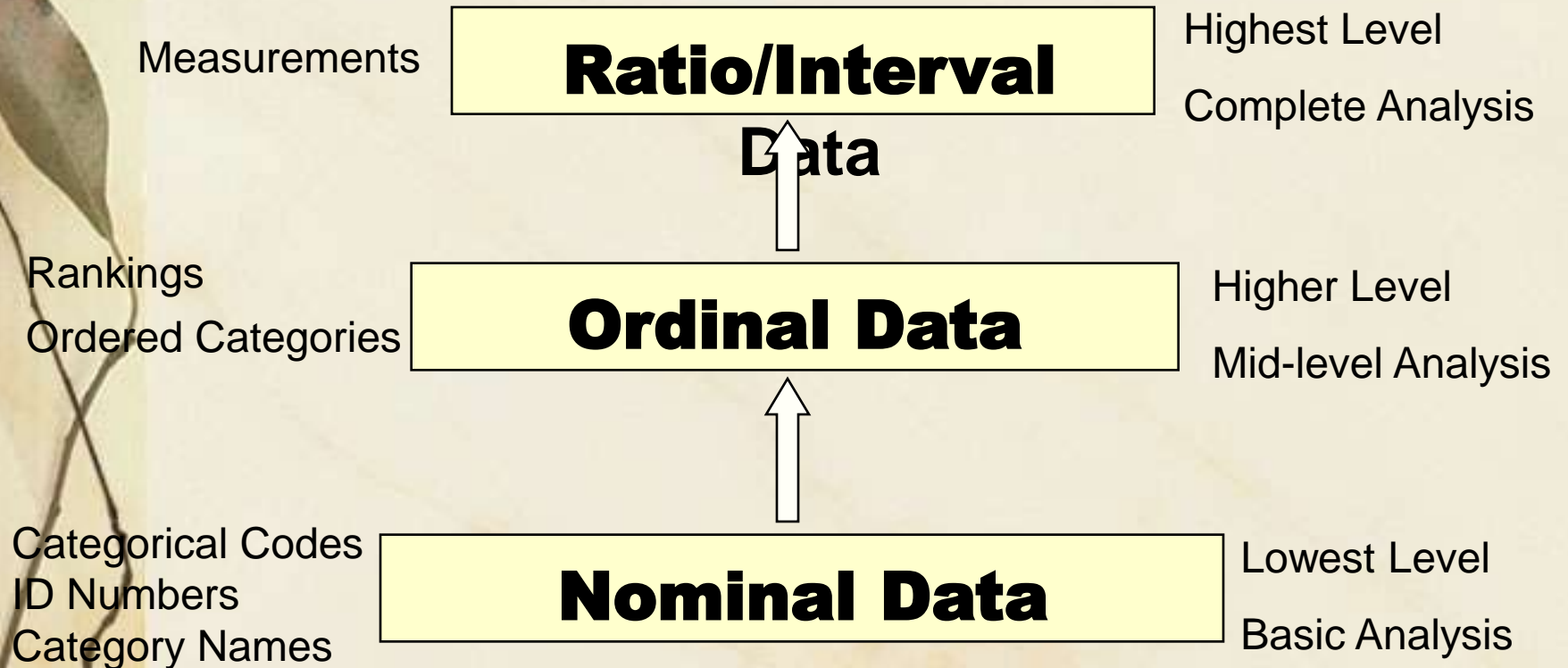○EXAMPLE: Temperature on the Fahrenheit scale.

Levels of Measurement

Ratio level: the interval level with an inherent zero starting point. Differences and ratios are meaningful for this level of measurement.

EXAMPLES:

Monthly income of surgeons, or distance traveled by manufacturer's representatives per month.

Data Measurement Levels

Measurements

Ratio/Interval Data

Highest Level

Complete Analysis

Rankings
Ordered Categories

Ordinal Data

Higher Level

Mid-level Analysis

Categorical Codes
ID Numbers
Category Names

Nominal Data

Lowest Level

Basic Analysis

The Relation Between Data Level and Statistical Methods

Data Level	Meaningful Operation	Statistical Method
Nominal; Sihalease,muslims,tamils	**Classifying & Counting**	**NP**
Ordinal; Taste test ranking of three brand of soft drink	**Classifying, Counting & Ranking**	**NP**
Interval; Calander time,Monetory Unit,Ferenhighgt,Temperature of three brand of Soft Drink	**Classifying,Counting, Ranking,Addition,substracting, Multiplication,& Division**	**P**
Ratio; Heihgt, Weight,Revenue &Expenditure,P/E ratio,Stock Turn Over .	**Classifying,Counting, Ranking,Addition, substracting,Multiplication, & Division**	**P**