

Chapter 1 Data and Statistics

Motivation: the following kinds of statements in newspaper and magazine appear very frequently,

- Sales of new homes are accruing at a rate of 70300 homes per year.
- The unemployment rate has dropped to 5.0%.
- The Dow Jones Industrial Average closed at 35000
- Census

The above numerical descriptions are very familiar to most of us since we use it in everyday life. As a matter of fact, these are part of statistics. Therefore, statistics is in our everyday life. We now give a description about statistics.

Definition of statistics:

Statistics is the art and science of collecting, analyzing, presenting, interpreting and predicting data.

Objective of this course:

Using statistics is to give the managers and decision makers a better understanding of the business and economic environment and thus enable them to make more informed and better decision.

1.1. Data:

1. Basis components of a data set:

Usually, a data set consists the following components:

Element: The entities on which data are collected.

Variable: A characteristic of interest for the element.

Observation: The set of measurements collected for a particular element.

Example 1:

We have a data set for the following 5 stocks:

Stock	Annual Sales (in million)	Earnings per share (\$)	Exchange (where to trade)
Cache Inc.	86.6	0.25	OTC
Koss Corp	36.1	0.89	OTC
Par Technology	81.2	0.32	NYSE
Scientific Tech.	17.3	0.46	OTC
Western Beef	273.7	0.78	OTC

Note:

OTC stands for “over the counter” while NYSE stands for “New York Stock Exchange”.

In the above data set,

Elements	Cache Inc. , Koss Corp, Par Technology, Scientific Tech, Western Beef
Variables	Annual Sales, Earnings per share, Exchange
Observations	(86. 6, 0. 25, OTC), (36. 1, 0. 89, OTC), (81. 2, 0. 32, NYSE), (17. 3, 0. 46, OTC), (273. 7, 0. 78, OTC)

2. Qualitative and quantitative data:

Qualitative data: labels or names used to identify an attribute of each element.

Quantitative data: indicating how much or how many

Example 1 (continue):

Qualitative data: OTC, OTC, NYSE, OTC and OTC

Quantitative data:

86. 6, 36. 1, 81. 2, 17. 3, 273. 7, 0. 25, 0. 89, 0. 32, 0. 46 and 0. 78

The variable “Exchange” is referred to as a qualitative variable.

The variables “Annual Sales” and “Earnings per share” are referred to as quantitative variables.

Note:

Quantitative data are always numeric, but qualitative data may be either numeric or nonnumeric, for example, id numbers and automobile license plate numbers are qualitative data.

Note:

Ordinary arithmetic operations are meaningful only with quantitative data and are **not** meaningful with qualitative data.

3. Cross-sectional and time series data:

Cross-sectional data: data collected at the same or approximately the same point in time.

Time series data: data collected over several time periods.

Example 2:

A recent issue of Fortune Magazine reported that the following companies had lowest sales per employee among the Fortune 500 companies.

Company	Sales per Employee	Sales Rank
Seagate Technology	\$42.2	285
SSMC	\$42.19	414
Russell	\$41.99	480
Maxxam	\$40.88	485
Dibrell Brother	\$22.56	470

- (a) How many elements are in the data set? Write down these elements.
- (b) How many variables are in the data set? Write down these variables.
- (c) How many observations are in the data set? Write down these observations.
- (d) Which of the above variables are qualitative and which are quantitative?

[Solution:]

- (a) 5 elements:

Seagate Technology, SSMC, Russell, Maxxam, and Dibrell Brother.

- (b) 2 variables: Sales per Employee and Sales Rank.

- (c) 5 observations:

(42.2, 285), (42.19, 414), (41.99, 480), (40.88, 485), (22.56, 470).

- (d) Quantitative variable: Sales per Employee

Qualitative variable: Sales Rank.