Chapter 1

Summary and Display of Univariate Data

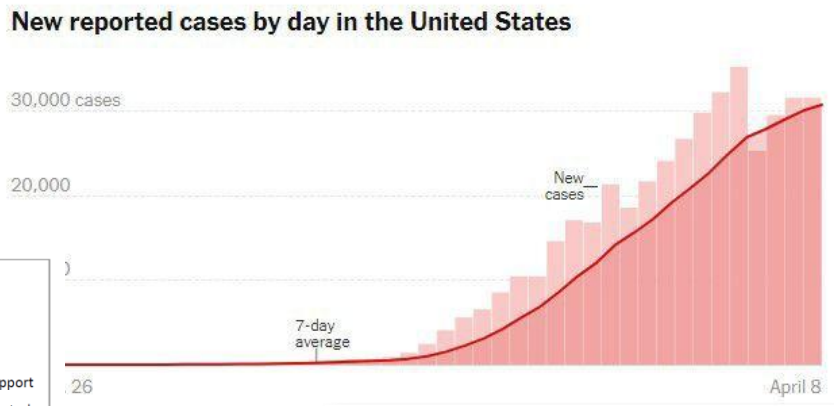
Lecture 1

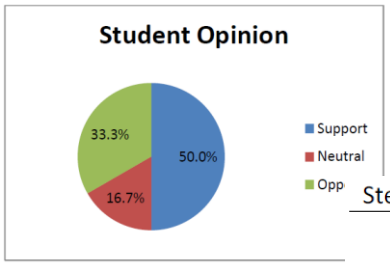
Some Key Statistical Concepts

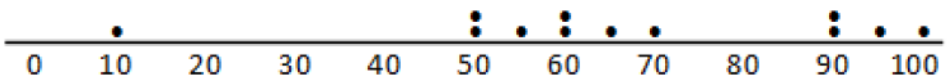
Classification of Variables

Summarizing data using tables and graphs

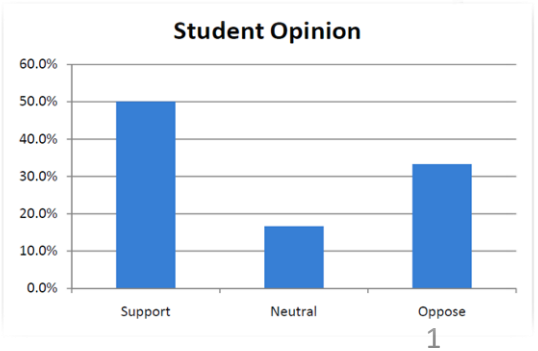
Dr. Lasantha Premarathna







Stem	Leaf
2	5
3	
4	1
5	057
6	2359
7	0255
8	0125
9	025



Chapter 1

Learning Outcomes

Demonstrate the ability to apply fundamental concepts in exploratory data analysis.

- Distinguish between different types of data.
- Interpret examples of methods for summarizing data sets, including common graphical tools (such as boxplots and histograms) and summary statistics (such as mean, median, variance and IQR).
- Assess which methods for summarizing a data set are most appropriate given data.
- Identify the features that describe a data distribution.
- Use an appropriate software tool for data summary and exploratory data analysis.

What is Statistics?

- We can mislead people w/ stats
  - ↳ garbage in garbage out
  - ↳ poor techniques / analysis
  - ↳ poor interpretation

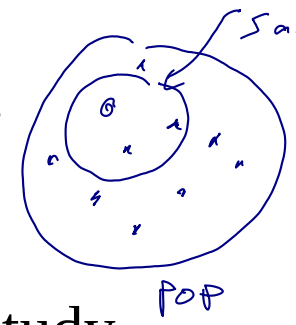
Basically similar defs

Statistics is a science involving the design of studies, data collection, summarize and analyse data, interpreting results and drawing conclusions.

Statistics is the science of learning from data, and of measuring, controlling and communicating uncertainty

Statistics is a branch of applied mathematics dealing with data collection, organization, analysis, interpretation and presentation.

Some Key Statistical Concepts...



Population and Samples

- Population: all subjects of interest in a particular study
- Sample: subset of the population

Parameter and Statistic

- Parameter : a descriptive measure of a *population*.
- Statistic : a descriptive measure of a *sample*.

Census and Sample Survey

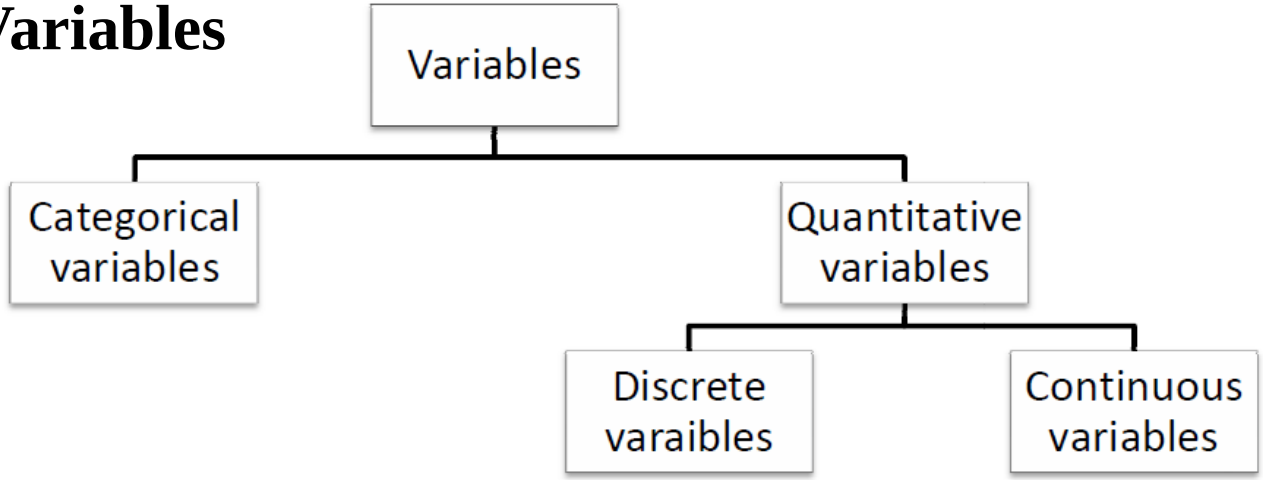
- Census : collecting data for the entire population
- Sample survey : collecting data for a sample

Descriptive vs. Inferential Statistics

- Descriptive Statistics refers to methods for summarizing the data. Summaries consist of graphs and numbers

- Inferential statistics refers to methods of making decisions or predictions about a population based on data obtained from a sample of that population.

Classification of Variables



- A variable can be classified as **categorical** if each observation belongs to one of a set of categories
- A variable is called **quantitative** if observations on it take numerical values that represent different magnitudes of the variable

Ex: Categorical(A) or quantitative (B) ?

1. Number of siblings in a family — Quantitative
2. County of residence — categorical
3. Distance (in km) of commute to school — Quantitative.
4. Blood type — Categorical

Before the next class

- Go through the Canvas Course page and read course guide and assessment regulations
- Register to iClicker Cloud, if not done already

Next Class:

- Chapter 1 : Summary and Display of Univariate Data (contd.)