www.DSinGIS.eu

Statistics

Research Methodologies and Scientific Communiation

Author(s):

Co-funded by the Erasmus+ Programme of the European Union

Statistics

Learning objectives

- Definition of statistics
- Tabular representation of the data
- Graphical representation of the data
- Statistical characteristics
- Comparison of several variables

Statistics

Definition of statistics

There are many different definitions of statistics, a few examples:

1. Statistics is the methodology for collecting, analyzing, interpreting and drawing conclusions from information. In other words, statistics is the methodology which scientists and mathematicians have developed for interpreting and drawing conclusions from collected data.

2. Statistics consists of a body of methods for collecting and analyzing data.(Agresti & Finlay, 1997)

Statistics

Definition of statistics

3. Statistics is the science of gaining information from numerical and categorical data. Statistical methods can be used to find answers to the questions like:

    - What kind and how much data need to be collected?
    - How should we organize and summarize the data?
    - How can we analyse the data and draw conclusions from it?
    - How can we assess the strength of the conclusions and evaluate theiruncertainty?

Statistics

Definition of statistics

Descriptive and inferential statistics

There are two major types of statistics. Summarization and description of data is called **descriptive statistics** and the branch of statistics concerned with using sample data to make an inference about a population of data is called **inferential statistics**.

Descriptive statistics consist of methods for organizing and summarizing information(Weiss, 1999)

Inferential statistics consist of methods for drawing and measuring the reliability of conclusions about population based on information obtained from a sample of the population.(Weiss, 1999)

Statistics

Definition of statistics

Descriptive statistics:
https://www.youtube.com/watch?v=QoQbR4lVLrs
https://www.youtube.com/watch?v=h8EYEJ32oQ8&list=PLU5aQXLWR3_yYS0ZYRA-5g5YSSYLNZ6Mc

Descriptive versus inferential statistics:
https://www.youtube.com/watch?v=OqoWtOvD8w0

Inferential statistics:
https://www.youtube.com/watch?v=sWCoj0fCtnU

(assessed June, 4th, 2019)

Statistics

Definition of statistics

Univariate analysis: this is the simplest form of analyzing data. The data has only one variable. it takes data, summarizes that data and finds patterns in the data.

Bivariate analysis: It involves the analysis of two variables, it the analysis of the relationship between the two variables.

multivariate analysis: is used to analyze the joint behavior of more than one random variable.

Statistics

Tabular representation of the data

Many application areas require a presentation of data where the nodes of a diagram are organized in a tabular way, i.e., where each node is associated to a specific row and column in the grid-like structure of a table.

Frequency table for the Ages of 50 Students enrolled in Statistics	
Age	**Frequency**
12	27
13	35
	N = 62

Statistics

Graphical representation of data

The statistical data can be represented graphically by using graphical symbols such as lines, bars, pie slices, dots etc.

Source: pixabay.com (24/6/19)

Statistics

Source: pixabay.com (24/6/19)

Source: pixabay.com (24/6/19)

Statistics

Statistical characteristics

Measure of location (describes the center of the frequency distribution)

- Mean
- Mode
- Median

Measure of dispersion (describes the variation of the frequency distribution)

- Variance
- Standard deviation
- Quartile
- Range

Statistics

Statistical characteristics

Measure of location: Describe the central tendency of the data. They include the mean, median and mode.

Measure of dispersion: Describes how spread out a set of data is. When a data set has a large value, the values in the set are widely scattered; when it is small the items in the set are tightly clustered. Very basically, this set of data has a small value:
1, 2, 2, 3, 3, 4
…and this set has a wider one:
0, 1, 20, 30, 40, 100

Statistics

Comparison of several variables

Standardization of values

https://www.youtube.com/watch?v=BN-2XOMnoCs
https://www.youtube.com/watch?v=1o-t_mVDDYQ

Covariation coefficient

https://www.youtube.com/watch?v=KDw3hC2YNFc
https://www.youtube.com/watch?v=xGbpuFNR1ME

Comparison via box-plots

https://www.youtube.com/watch?v=oBREri10ZHk
https://www.youtube.com/watch?v=CoVf1jLxgj4

Further information:

http://www.statsoft.com/Textbook/Graphical-Analytic-Techniques#icon3

(assessed June, 4th, 2019)