User Interface Design

Lecture 5: Data Analysis, Presentation and Interpretation

Dr. Obead Alhadreti

Learning Objectives

- Discuss the difference between qualitative and quantitative data and analysis.
- Make you aware of software packages that are available to help your analysis.
- Enable you to interpret and present your findings in appropriate ways.

What is Data?

- A collection of facts, such as values or measurements.
- The raw material of science.
- Contains the information waiting to be released.
- This is done through data analysis, presentation, and interpretation.

Types of Data

1. Quantitative data – expressed as numbers

2. Qualitative data – expresses the nature of elements and is represented as themes, patterns, stories

Quantitative vs. Qualitative Analysis

Explanation through numbers	Explanation through words
Objective	Subjective
Deductive reasoning	Inductive reasoning
Predefined variables and measurement	Creativity, extraneous variables
Data collection before analysis	Data collection and analysis intertwined
Cause and effect relationships	Description, meaning

Simple quantitative analysis

- Summary statistics: used to summarise a set of observations.

- Commonly used summary statistics are:
  1. measure of central tendency such as the average (mean, median or mode)
  2. measure of variability like the range and the standard deviation

Measures of central tendency

1. Mean (also called average): add up values and divide by number of data points

2. Median: middle value of data when ranked

3. Mode: figure that appears most often in the data

Example 1: Finding the Mean of a Data Set

- **Find the mean of each data set.**

Depths of Puddles (in.)						
5	8	3	5	4	2	1

- **mean:** 5 + 8 + 3 + 5 + 4 + 2 + 1 = 28

  28 ÷ 7 = 4

Add all values.

Divide the sum by the number of items.

The mean is 4 inches.

Example 2: Finding the Mean of a Data Set

- **Find the mean of each data set.**

Number of Points Scored			
96	75	84	7

Example 3: Finding the Median of a Data Set

- **Find the median of the data set.**

$100, $275, $300, $325, $350, $375, $500

First place the prices in numerical order.

$100, $275, $300, $325, $350, $375, $500

The price in the middle is the median price.

$100, $275, $300, $325, $350, $375, $500

The median price is $325

Example 4: Finding the Median of a Data Set

- **Find the median of the data set.**

12, 15, 11, 11, 7, 13

First place the prices in numerical order.

7,11,11,12,13,15

Then find the number in the middle or the average of the two numbers in the middle.

11 + 12 = 23 23 / 2 = 11.5

The median number is 11.5

Example 5: Finding the Mode of a Data Set

- **Find the mode of the data set.**

12, 15, 11, 11, 7, 13

The mode is 11

Example 6: Finding the Mode of a Data Set

- Sometimes a set of data will have more than one mode.

- For example, in the following set the numbers both the numbers 5 and 7 appear twice.

2, 9, 5, 7, 8, 6, 4, 7, 5

- The mode and this set is said to be bimodal.

Example 7: Finding the Mode of a Data Set

- Sometimes there is no mode in a set of data.

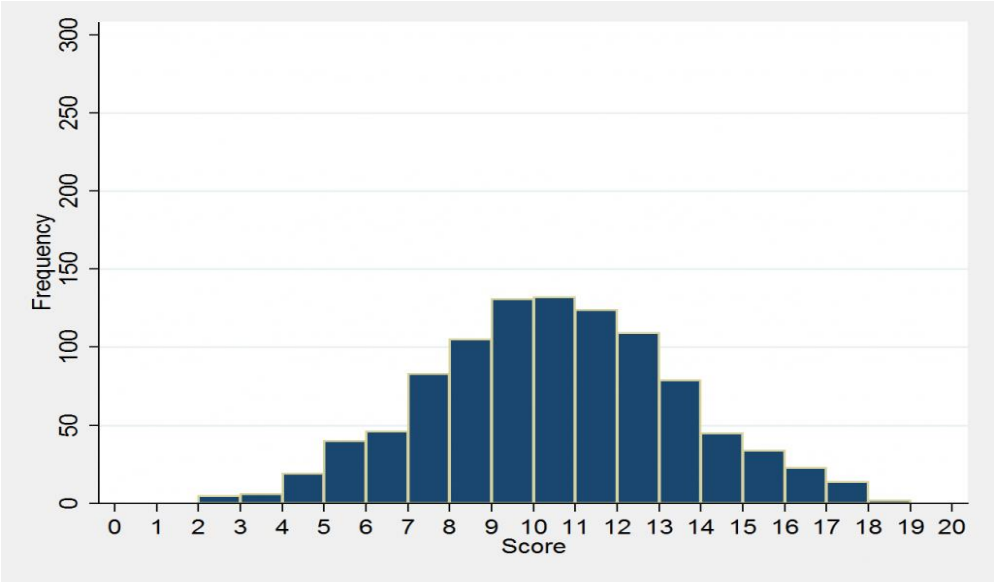
3, 8, 7, 6, 12, 11, 2, 1

- All the numbers in this set occur only once therefore there is no mode in this set.

Measures of Variability

1. Range: is the difference between the least and greatest values in the set.

2. Standard deviation: is a measure of how each value in a data set varies or deviates from the mean.



Example 8: Finding the Range of a Data Set

- Consider the following set:

40, 30, 43, 48, 26, 50, 55, 40, 34, 42, 47, and 50

- To find the range you would take the largest number, 50, and subtract the smallest number, 26.

55 – 26 = 29

The range is 29

Example 9: Finding the Standard Deviation of a Data Set

- The math test scores of five students are:

92,88,80,68 and 52.

1) Find the **mean**: (92+88+80+68+52)/5 = 76

2) Find the **deviation from the mean**:

92-76=16
88-76=12
80-76=4
68-76= -8
52-76= -24

Example 9: Finding the Standard Deviation of a Data Set

3) Square the deviation from the mean:

$$(16)^2 = 256$$

$$(12)^2 = 144$$

$$(4)^2 = 16$$

$$(-8)^2 = 64$$

$$(-24)^2 = 576$$

4) Find the sum of the squares of the deviation from the mean:

256+144+16+64+576= 1056

Example 9: Finding the Standard Deviation of a Data Set

5) Divide by the number of data items to find the variance:

1056/5 = 211.2

6) Find the square root of the variance:

$$\sqrt{211.2} = 14.53$$

- Thus the standard deviation of the test scores is 14.53

Qualitative analysis

- The primary activity of qualitative analysis is the search for patterns and explanations for those patterns.

- Categorization scheme may be emergent or pre-specified.

Figure 8.8 Building the affinity diagram of Indian ATM usage

Source: Figure 1, A. DeAngeli, U. Athavamker, A. Joshi, L. Coventry and G.I. Johnson (2004) "Introducing ATMs in India: a contextual inquiry", *Interacting with Computers* 16(1), 29–44. Reproduced with permission.

Some approaches to qualitative analysis

- Grounded Theory

  Begins with observations and no preconceived hypotheses

  Seeks to discovers patterns in the data.

- Conversation Analysis

  structure and norms around language and attached significance and meaning.

  attention to pauses, tone, stuttering etc.

Qualitative analysis process

- There are some generic strategies that are part of almost every approach to data analysis.

  - immersion in the data
  - doing preliminary and informal analysis
  - making analytic memos
  - finding codes or themes
  - connecting the codes or themes into categories
  - searching for confirming and disconfirming evidence
  - building a conceptual framework that explains the findings

Tools to support data analysis

- Spreadsheet – simple to use, basic graphs
- Statistical packages, e.g. SPSS
- Qualitative data analysis tools
  - Categorization and theme-based analysis
  - Quantitative analysis of text-based data
- Nvivo and Atlas.ti support qualitative data analysis
- CAQDAS Networking Project, based at the University of Surrey (http://caqdas.soc.surrey.ac.uk/)

Data Presentation

- Visual methods can make the point much stronger than simply describing the data.
- Appropriate use of tables and graphs can enhance the message you are delivering.
- Tables
  - Simplest way to summarize data
  - Data is presented as absolute numbers or percentages
- Charts and graphs
  - Visual representation of data
  - Usually data is presented using percentages

Tables

- Good for when exact numbers need to be presented.
- Because your STEPS report will be **the** reference for results from your STEPS survey, it will have a lot of tables!

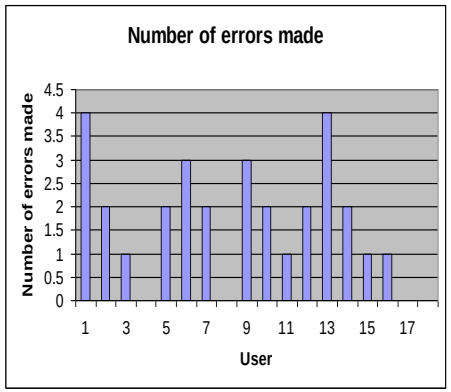
- **Best Practice:**
  - **Clear table title and column / row headings**
  - **Minimal use of grid lines**
  - **Leave enough space so columns/rows are easy to read**

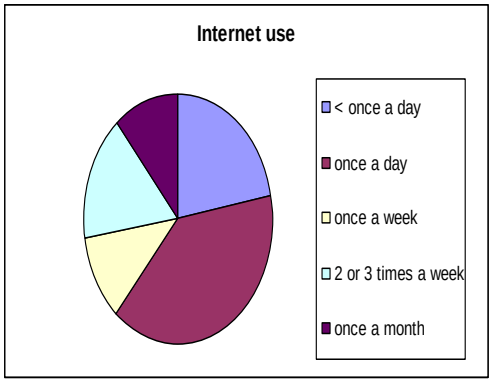
Percentage of current smokers									
Age Group (years)	Men			Women			Both Sexes		
	n	% Current smoker	95% CI	n	% Current smoker	95% CI	n	% Current smoker	95% CI
18-24	251	45.2	38.2-52.1	391	5.0	2.6-7.4	642	26.5	22.1-30.9
25-34	258	62.7	55.2-70.2	675	3.7	2.1-5.2	933	36.1	31.8-40.4
35-44	357	65.3	59.8-70.9	812	5.7	4.0-7.4	1169	33.7	29.7-37.7
45-54	483	57.5	52.1-62.9	1212	5.7	3.5-7.9	1695	29.9	26.6-33.1
55-64	523	40.1	35.1-45.1	1475	3.2	2.0-4.4	1998	20.9	18.1-23.6
18-64	**1872**	**55.5**	**52.7-58.4**	**4565**	**4.8**	**3.7-5.8**	**6437**	**30.3**	**28.4-32.2**

Charts and Graphs

- **Bar graphs** show quantities represented by horizontal or vertical bars and are useful for displaying:
  - Several categories of results at once (e.g. males vs. females)

- **Pie charts** show proportions in relation to a whole, with each wedge representing a percentage of the total and are useful for displaying:
  - Parts of a whole in percentages





Charts and Graphs: Best Practice

- Emphasize one idea at a time in a figure.
- Pay careful attention to the scaling of the graph.
- Provide a title, units and labels: the graph or table should be self-explanatory!
- If possible, mention the total sample size of the data set for which the graph or chart is made.
- Be sparing and consistent with use of colour, fonts and "enhancements".

Effective Data Presentation

- Clear
- Consistency
  - Font, Colors, Punctuation, Terminology, Line/ Paragraph Spacing
- An appropriate amount of information, less is more
- Attractive

Data Interpretation

- **Interpretation**: adding meaning to information by making connections and comparisons and by exploring causes and consequences.

- Only make claims that your data can support.