Data Analysis, Interpretation and Presentation

BTH745 – Human Factors in Computing

Aims of this chapter…

- Discuss the difference between qualitative and quantitative data and analysis

- Enable you to analyze data gathered from:
  - Questionnaires
  - Interviews.
  - Observation studies

- Make you aware of software packages that are available to help your analysis

- Identify common pitfalls in data analysis, interpretation, and presentation

- Enable you to interpret and present your findings in appropriate ways

Quantitative and qualitative

- Quantitative data – expressed as numbers

- Qualitative data – difficult to measure sensibly as numbers, e.g. count number of words to measure dissatisfaction

- Quantitative analysis – numerical methods to ascertain size, magnitude, amount

- Qualitative analysis – expresses the nature of elements and is represented as themes, patterns, stories

- Be careful how you manipulate data and numbers!

The First steps in Analyzing Data

|	Usual raw data	Example. qualitative data	Example quantitative data	Initial processing steps
Interviews	Audio/video recordings. interviewer notes.	Responses to open question, video pictures. respondent's opinions.	Age, job role, yrs experience. Responses to closed questions.	Transcription of recordings. Expansion of notes.
Questionnaires	Written responses. Online databases.	Responses to open questions. Responses in 'further comments' field.	Age, job, yrs of experience. Responses to closed questions.	Clean up data, filter into different data sets.
Observations	Observer's notes. Photographs, Audio/video recordings.	Records of behavior. Description of a task as it is undertaken.	Demographics of participants. Time on tasks. Number of people involved.	Expansion of notes. Transcription. Synchronization.

Simple quantitative analysis

- Averages
  - Mean: add up values and divide by number of data points
  - Median: middle value of data when ranked
  - Mode: figure that appears most often in the data
- Percentages: normalize (divide by maximum)
- Be careful not to mislead with numbers!
- Graphical representations give overview of data
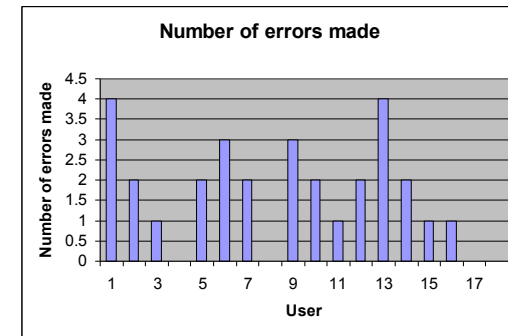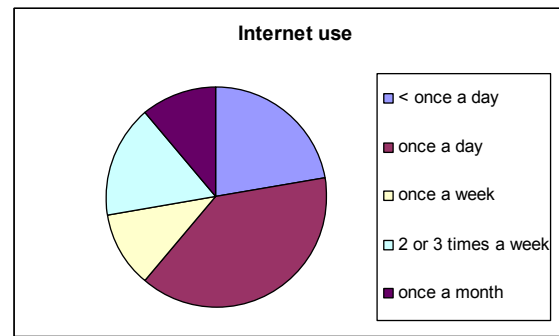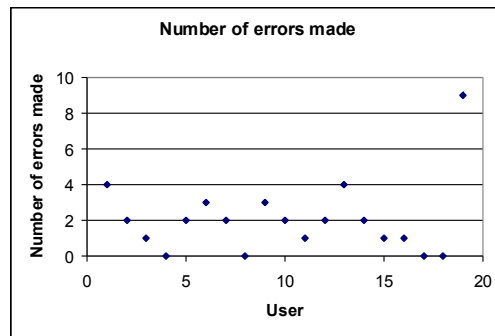
Web analytics
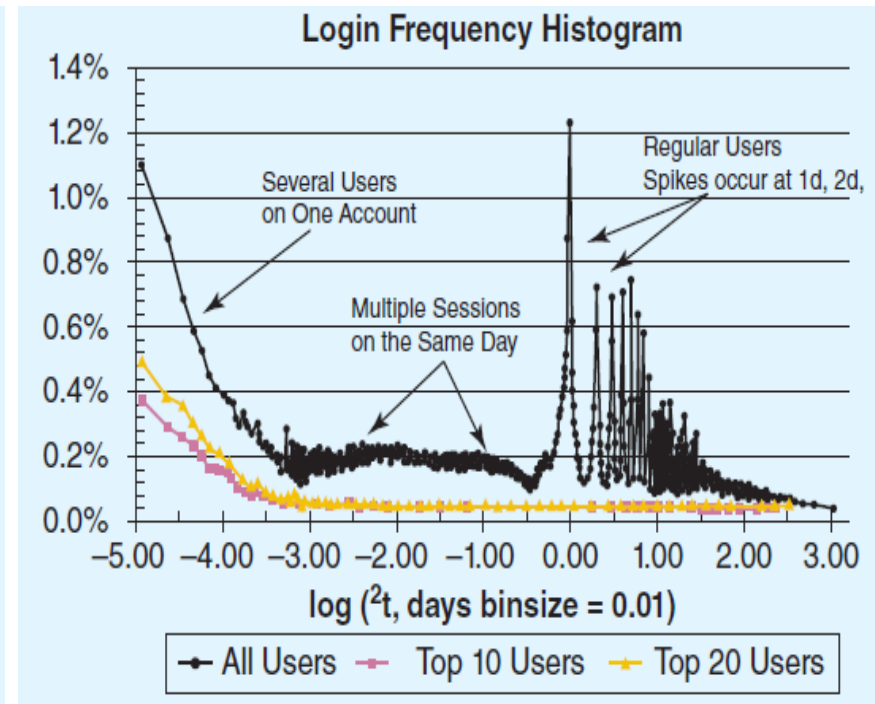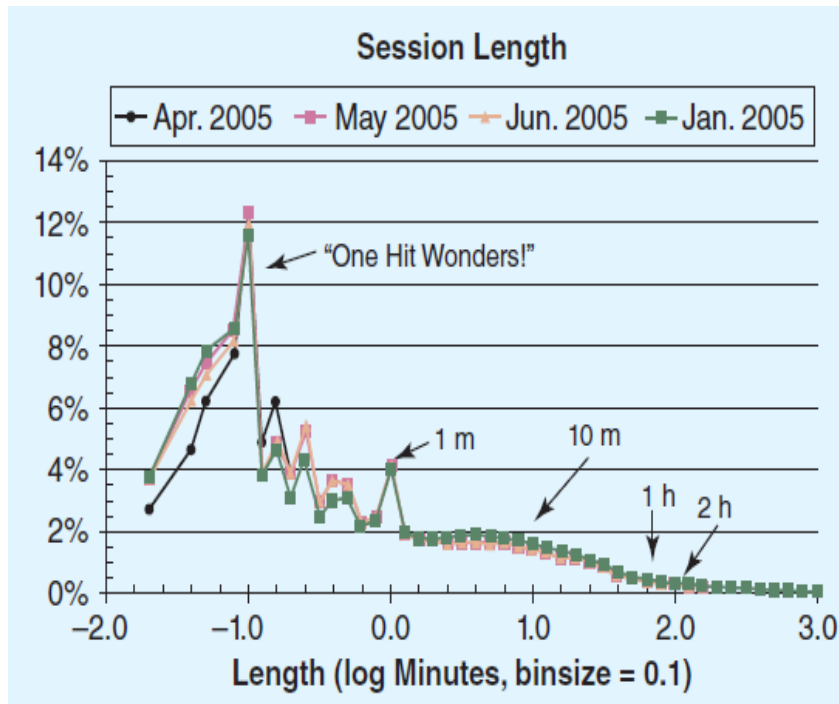
session length data of four different months from Teachers' Domain (NSDL)

Source: Khoo, M., Pagano, J., Washington, A. L., Recker, M., Palmer, B., and Donahue, R. A. (2008) Using web metrics to analyze digital libraries. *Proceedings of Joint Conference on Digital Libraries,* Pittsburgh, June 16–20. ©2008 Association for Computing Machinery, Inc. Reprinted by permission.

Simple qualitative analysis

- Categorizing data
  - Categorization scheme may be emergent or pre-specified
    - Which categories to choose
    - Appropriate granularity for the categories

Categorization Scheme

1. **Interface Problems:**
   1.1. verbalizations show evidence of dissatisfaction about an aspect of interface.
   ....
   1.10. the participant makes a suggestion for redesign of the interface.
2. **Content Problems:**
   2.1 verbalizations show evidence of dissatisfaction about aspects of the content of the electronic text.
   ....

Coded Transcription Excerpt

[I just don't concentrate very well when I'm looking at the screen, **UP1.1]**. It's nice that I'm clicking every now and then coz it just sort of organizes the thoughts. **[**But, it would be still nice to see it on a piece of paper, **UP1.10].** **[**There is so much reference to all those previously said, **UP2.1] ….** **[**so, it would be nice to have this navigation, extra links, **UP1.10]**

Simple qualitative analysis

- Looking for critical incidents

  - Helps to focus in on key events

  - Using the rest of the data as context

  - Examples:

    - After searching around the house for a while, she found a copy of the local paper for the correct week.

    - The local paper she had found did not have details of cinema that she wanted to visit

    - When trying to book the cinema tickets by phone she discovered that she needed a credit card, which she doesn't have, and so she had to give me a phone!

36

Cluster Analysis

- PCA: Principal component analysis. Reduce the number of variables to a small
  number of manageable "components".
- CA: Cluster analysis. Create categories ("clusters") based on these key components.
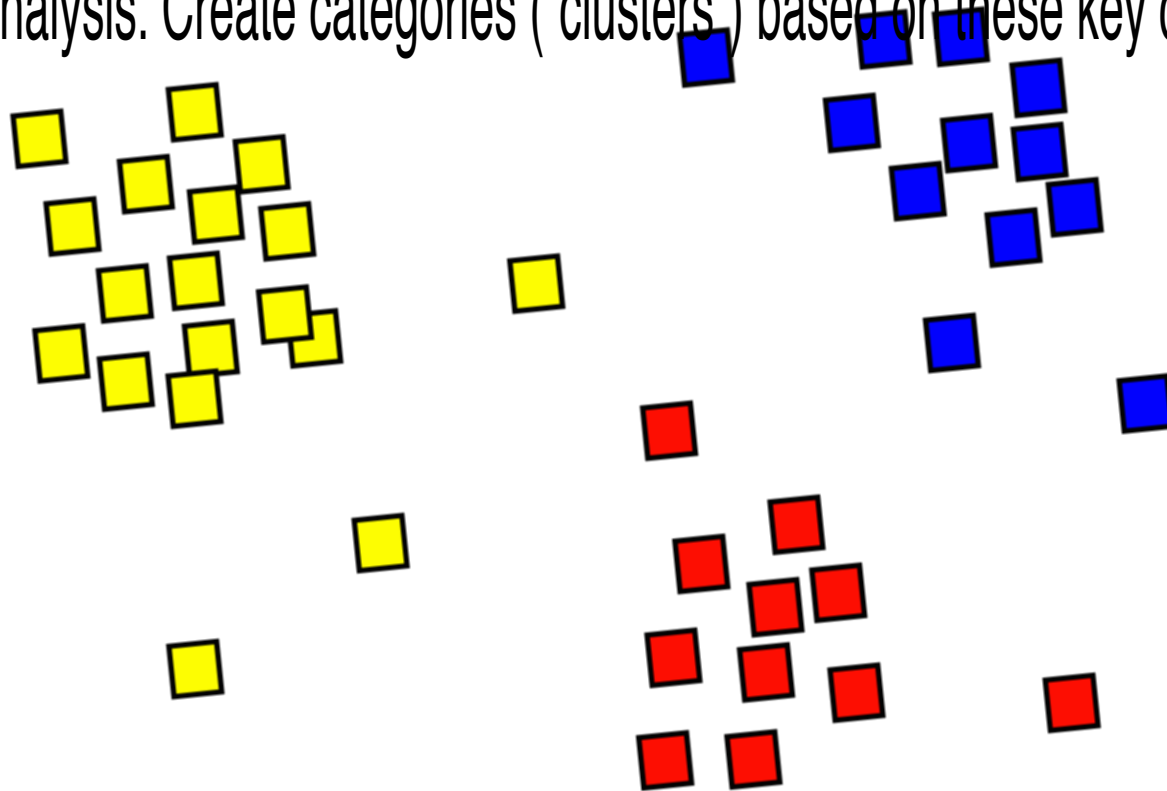
- There are numerous different algorithms that can automatically perform such analysis.

Tools to support data analysis

- Spreadsheet (Excel) – simple to use, basic graphs

- Programming environments: Excel (VBA), Python

- Statistical packages, e.g. SPSS, SAS, R (most also have scripting/programming ability)

- 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 (https://www.surrey.ac.uk/computer-assisted-qualitative-data-analysis)

Presenting the findings

- Only make claims that your data can support

- The best way to present your findings depends on the audience, the purpose, and the data gathering and analysis undertaken

- Graphical representations (as discussed above) may be appropriate for presentation.

  - Tool: Tableau

- Other techniques are:
  - Rigorous notations, e.g. UML
  - User stories, e.g. to create scenarios
  - Summarizing the findings

38

Summary

- The data analysis that can be done depends on the data gathering that was done

- Qualitative and quantitative data may be gathered from any of the three main data gathering approaches

- Percentages and averages are commonly used in Interaction Design

- Mean, median and mode are different kinds of 'average' and can have very different answers for the same set of data

- Normalization constant should be the same for different data sets if the metrics are to be compared

- Presentation of the findings should not overstate the evidence