Week 8

# Data Analysis, Interpretation and Presentation

Aims

- **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.**

Simple quantitative analysis

- **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!**

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**
- **Be careful not to mislead with numbers!**
- **Graphical representations give overview of data**
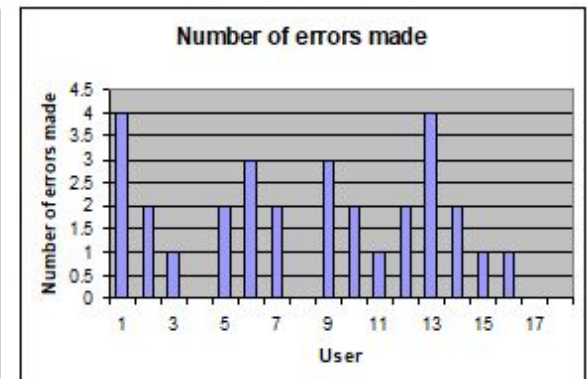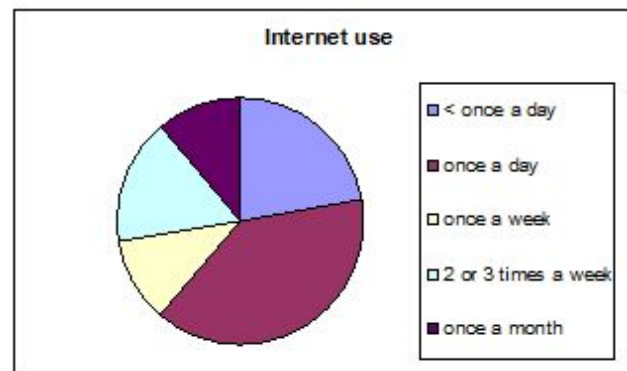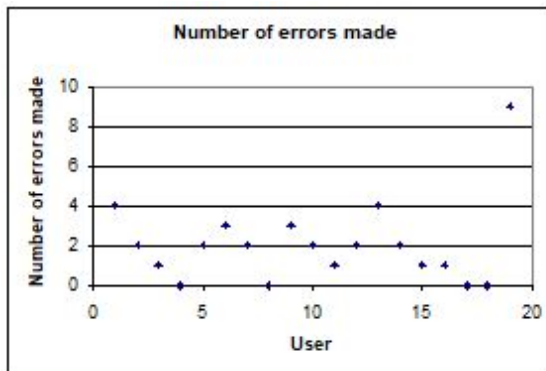
Visualizing log data
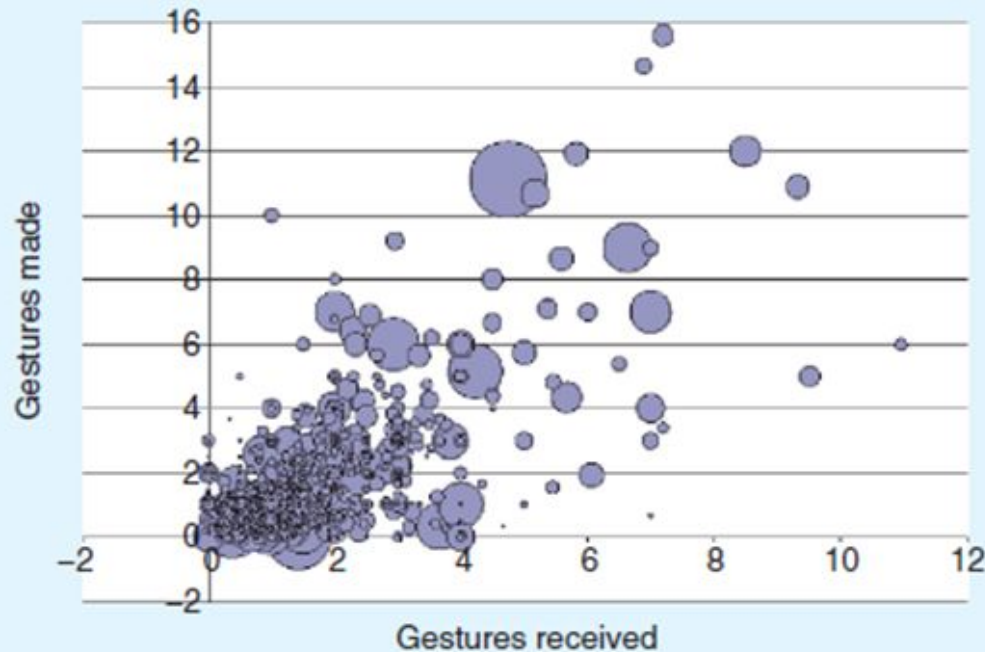
Interaction profiles of players in online game

Figure 8.6 Interaction profiles of players in the cantina

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in *Proceedings of CSCW 04*. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.
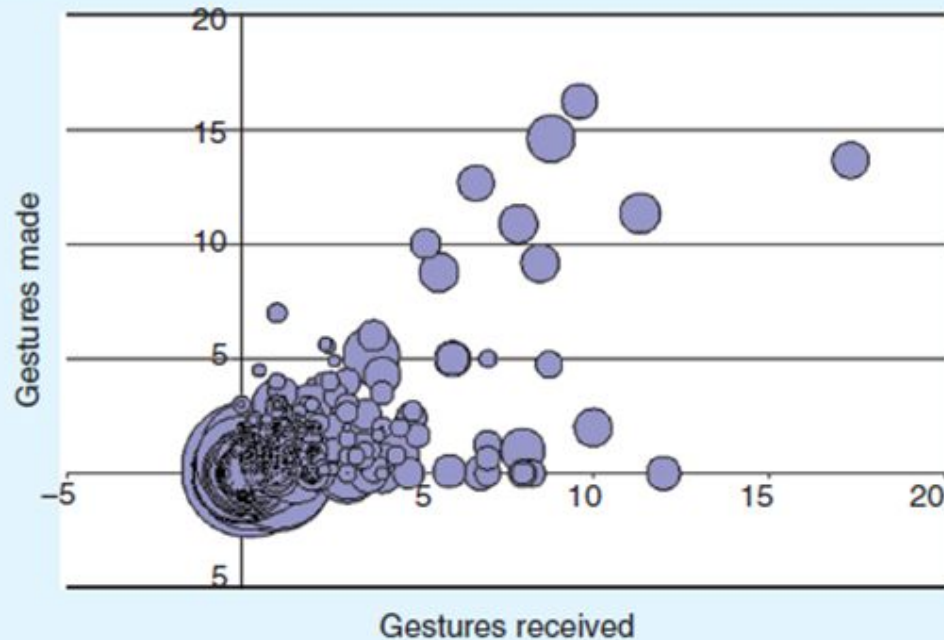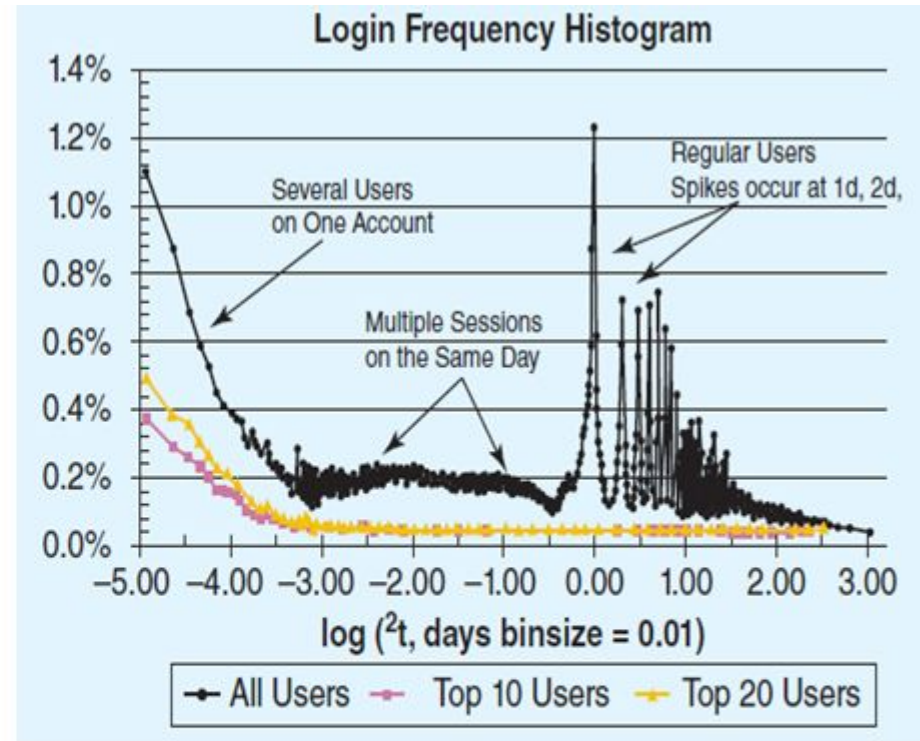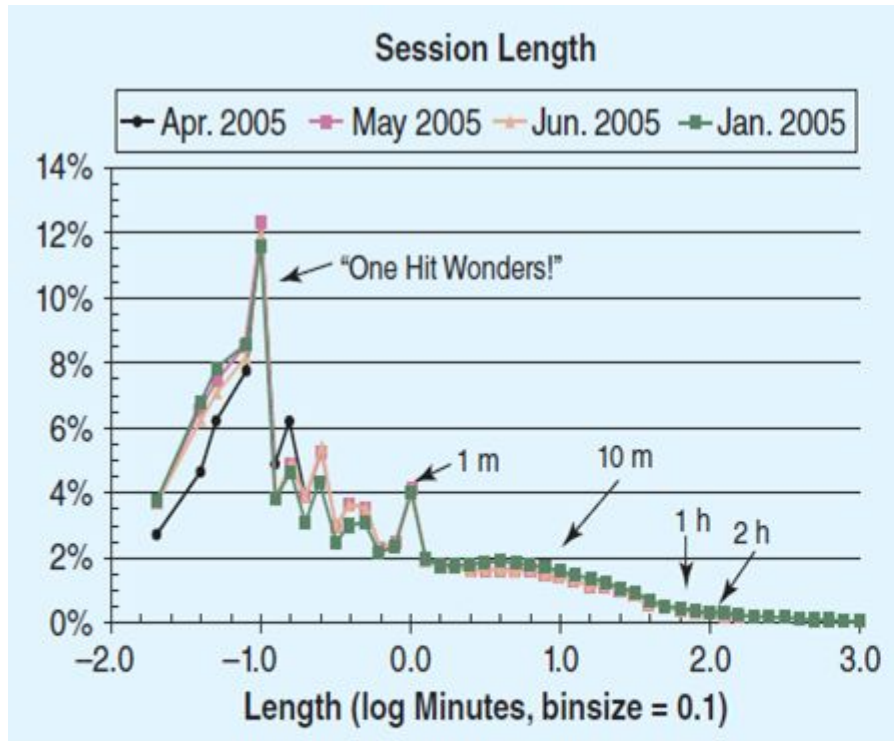
Log of web page activity

Figure 8.7 Interaction profiles of players in the starport

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in Proceedings of CSCW 04. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

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

- **Recurring patterns or themes**
  - Emergent from data, dependent on observation framework if used

- **Categorizing data**
  - Categorization scheme may be emergent or pre-specified

- **Looking for critical incidents**
  - Helps to focus in on key events

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.

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/)**

Theoretical frameworks for qualitative analysis

- **Basing data analysis around theoretical frameworks provides further insight**

- **Three such frameworks are:**
    - Grounded Theory
    - Distributed Cognition
    - Activity Theory

Grounded Theory

- **Aims to derive theory from systematic analysis of data**

- **Based on categorization approach (called here 'coding')**

- **Three levels of 'coding'**
  - Open: identify categories
  - Axial: flesh out and link to subcategories
  - Selective: form theoretical scheme

- **Researchers are encouraged to draw on own theoretical backgrounds to inform analysis**

Figure 8.13 Code book used in a grounded theory analysis of citizens' motivations to contribute to citizen science

Source: Rotman, D. et al (2014). Does motivation in citizen science change with time and culture? In Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW Companion '14). ACM, New York, NY, USA, 229–232. ©2014 Association for Computing Machinery, Inc. Reprinted by permission.
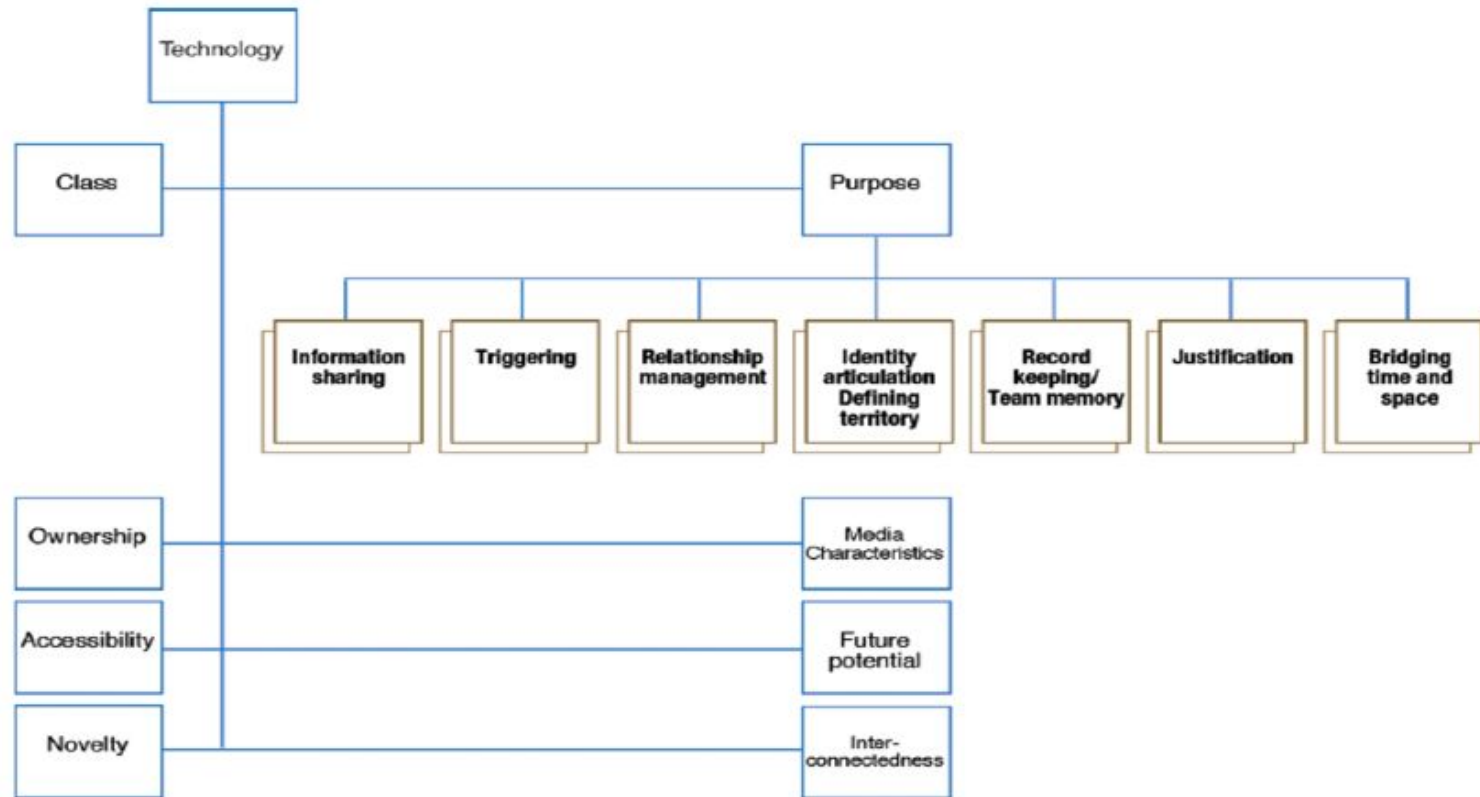
Excerpt showing axial coding

Figure 8.14 Axial coding for the technology category

Source: S. Sarker, F. Lau and S. Sahay (2001): "Using an adapted grounded theory approach for inductivetheory building about virtual team development". *The Data Base for Advances in Information Systems*, 32(1), pp. 38–56 ©2001 Association for Computing Machinery, Inc. Reprinted by permission.

Distributed Cognition

- **The people, environment & artefacts are regarded as one cognitive system**

- **Used for analyzing collaborative work**

- **Focuses on information propagation & transformation**

Activity Theory

- **Explains human behaviour in terms of our practical activity in the world**

- **Provides a framework that focuses analysis around the concept of an 'activity' and helps to identify tensions between the different elements of the system**

- **Two key models: one outlines what constitutes an 'activity'; one models the mediating role of artifacts**

Activity — Motive

Action _ Goal

Operation — Conditions
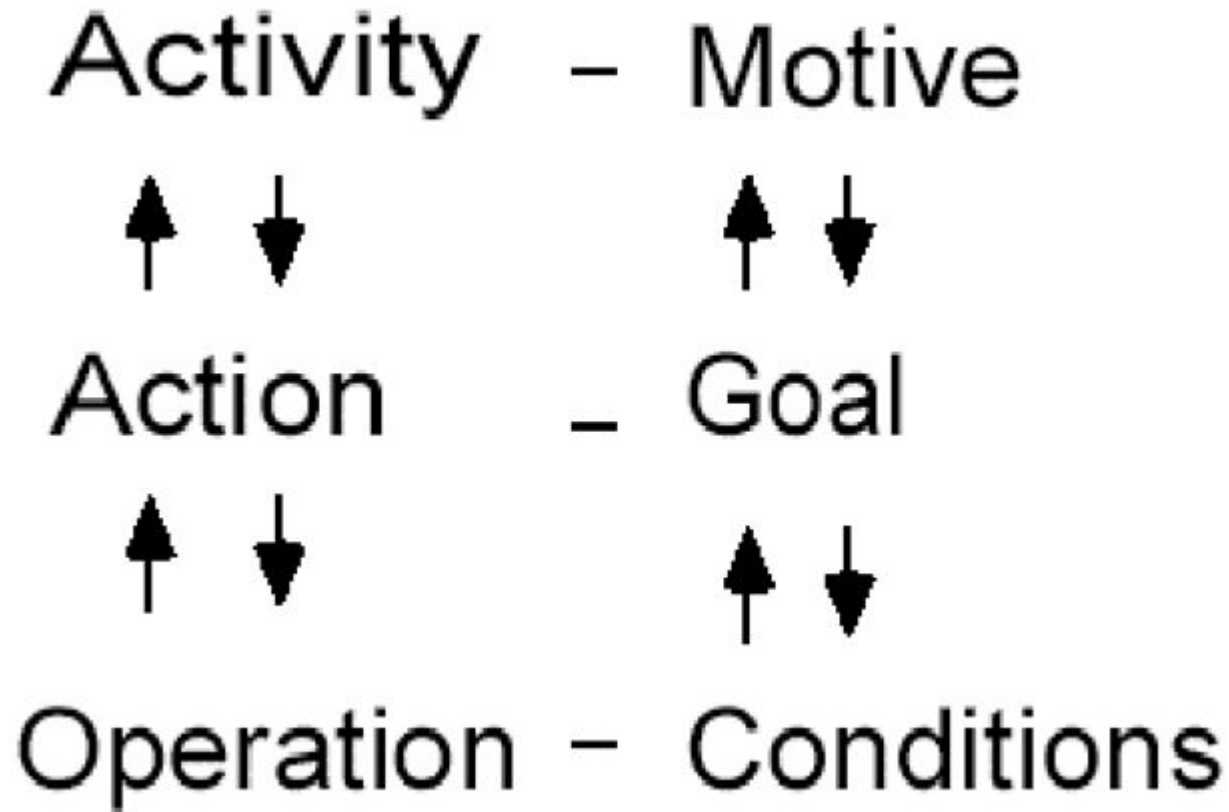
Engeström's (1999) activity system model
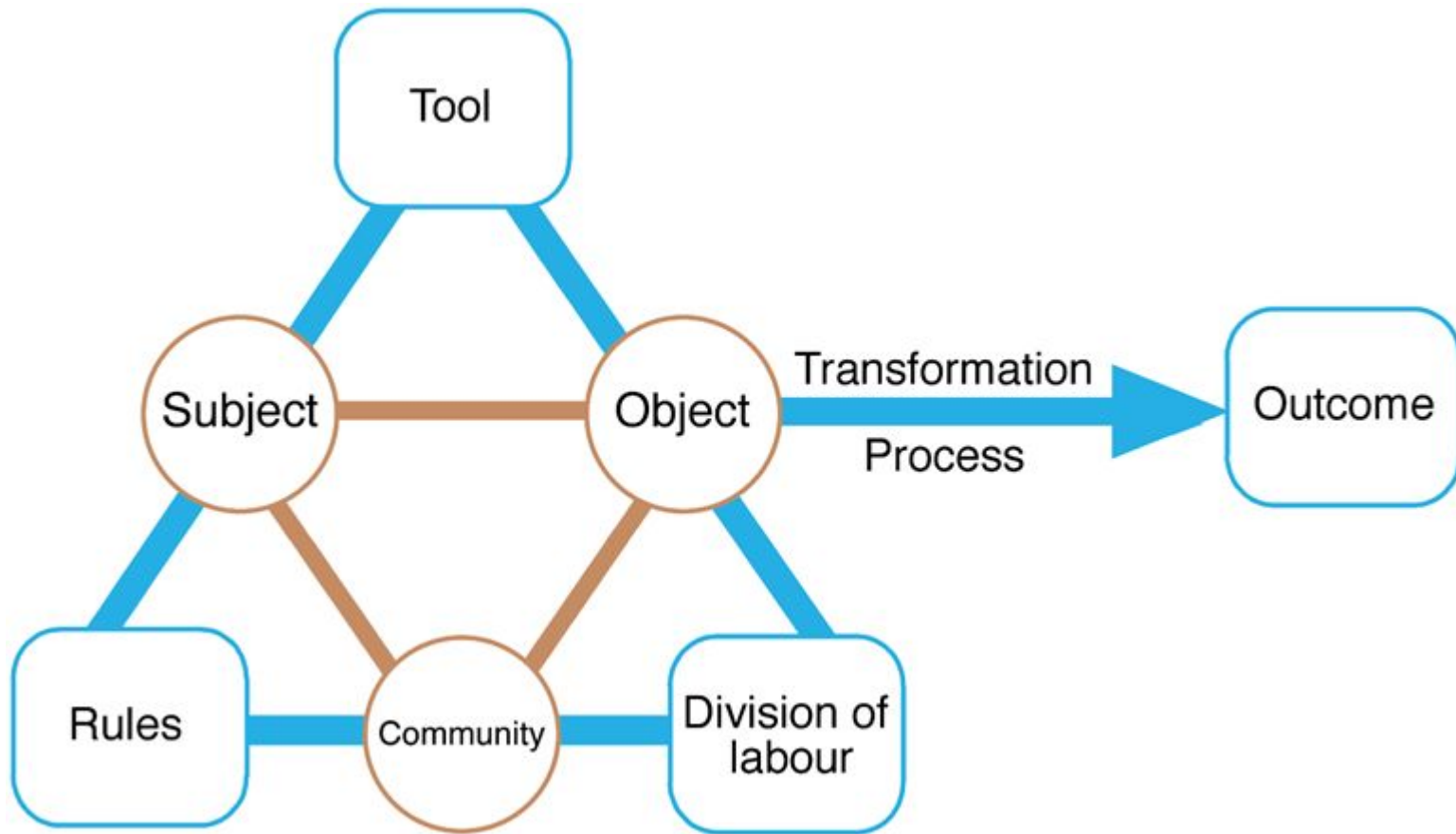
Presenting the findings

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

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

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

3. Percentages and averages are commonly used in Interaction Design

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

5. Grounded Theory, Distributed Cognition and Activity Theory are theoretical frameworks to support data analysis

6. Presentation of the findings should not overstate the evidence