

Introduction to Digital Humanities for Japanese Studies with the Japanese Visual Media Graph project

Part 2

Research funded by:

DFG Deutsche
Forschungsgemeinschaft
German Research Foundation

Zoltan Kacsuk
Stockholm University
17 January 2023

Outline of the presentation

1. Introducing Digital Humanities
2. Digital Resources for Japanese Studies
3. The Japanese Visual Media Graph Project
4. Simple Descriptive Statistics Tools
5. Visual Network Analysis

Part 2

Simple Descriptive Statistics Tools

Descriptive statistics

- Summarizing data via:
 - Summary statistics (sum, average, etc.)
 - Tables (e.g. cross tables)
 - Diagrams
 - Measures of dependence (e.g. correlation)

Descriptive statistics

- Summarizing data via:
 - Summary statistics (sum, average, etc.)
 - Tables (e.g. cross tables)
 - Diagrams
 - Measures of dependence (e.g. correlation)

The other main branch of statistics is:

Inferential statistics

- Based on probability theory
- Hypothesis testing

Software tools

- **Spreadsheet programs:** MS Excel, LibreOffice Calc, etc.
 - Easiest to use
 - High chance that one is already available on your computer
 - LibreOffice Calc is free and available for all major operating systems
- **Specialized **statistical analysis programs:**** SPSS, STATA, etc.
 - Used in fields that have a long history of using statistical methods: sociology, biology, economics, etc.
 - Graphical User Interface (GUI) and menu systems
 - Usually proprietary and not free
- **Programming languages:** R, Python, etc.
 - Free, open source, and most powerful
 - Used in data science work
 - Need to learn some programming
 - Excellent GUI coding environments: RStudio for R, Jupyter Notebooks for Python

Software tools

- **Spreadsheet programs:** MS Excel, LibreOffice Calc, etc.
 - Easiest to use
 - High chance that one is already available on your computer
 - LibreOffice Calc is free and available for all major operating systems
- **Specialized **statistical analysis programs:**** SPSS, STATA, etc.
 - Used in fields that have a long history of using statistical methods: sociology, biology, economics, etc.
 - Graphical User Interface (GUI) and menu systems
 - Usually proprietary and not free
- **Programing languages:** R, Python, etc.
 - Free, open source, and most powerful
 - Used in data science work
 - Need to learn some programming
 - Excellent GUI coding environments: RStudio for R, Jupyter Notebooks for Python

Example analysis: Applying the Census Approach to Japanese Visual Media

- https://jvmg.iuk.hdm-stuttgart.de/wp-content/uploads/sites/3/2022/12/2022_Seika_JVMG_research.pdf

Visual Network Analysis

Networks in our lives

- **Infrastructural networks** (electric grid, Internet infrastructure, road system, flights, etc.)
- **Social networks** (friends, colleagues, production networks, etc.)
 - Emails, phone calls, etc.
 - Collaboration networks
 - Erdős number, Bacon number
- **Networks of reference**
 - World Wide Web
 - Citation networks (bibliometry and scientometry)
- etc.

Software tools

- **Spreadsheet programs:** MS Excel, LibreOffice Calc, etc.
 - Only used for preparing the data to be loaded into Gephi
- **Specialized programs: Gephi**
 - The best specialized network visualization program
 - Free and open source
 - Graphical User Interface
- **Programming languages:** R, Python, etc.
 - More powerful than Gephi for analysis, but for visualization just use Gephi

Software tools

- **Spreadsheet programs:** MS Excel, LibreOffice Calc, etc.
 - Only used for preparing the data to be loaded into Gephi
- **Specialized programs: Gephi**
 - The best specialized network visualization program
 - Free and open source
 - Graphical User Interface
- **Programing languages:** R, Python, etc.
 - More powerful than Gephi for analysis, but for visualization just use Gephi

Thank you for your attention!

Get in touch at: kacsuk@hdm-stuttgart.de

Visit our project website:

<https://jvmg.iuk.hdm-stuttgart.de/>

Visit the JVMG database:

<https://mediagraph.link/>