Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Analysing Survey Data in R
FOAR705 Proof of Concept presentation

Kathryn Phillips, 43686826

Department of International Studies

June 2, 2017

Table of Contents

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Introduction
My Research
Topic
Data Analysis in
R

Scoping
Selection
Elaboration

Implementation
Creating a Tool
Chain
Elaborating on
my Chain
The Database

Results
Pros
Cons

Conclusion
Final Words
Code Repository

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

My Research Topic

During my MRes course, I plan to research the influence of Korean Pop (Kpop) within Japan.

You can find advertisements for Kpop all around Tokyo, but where is the research on this phenomenon?

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Data Analysis in R

I used R to form a code line that could analyse specific groups of responses out of the thousands I could receive from an online survey. R allows me to:

- Immediately take online responses from surveys
- Input them into a well-structured database
- Select specific groups within the database
- Create graphs of my data

With this tool chain, I am well-prepared to start my research in the months to come!

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Introduction
My Research
Topic
Data Analysis in
R

Scoping
Selection
Elaboration

Implementation
Creating a Tool
Chain
Elaborating on
my Chain
The Database

Results
Pros
Cons

Conclusion
Final Words
Code Repository

Scoping
Selection

Initially, I had several areas where using a digital tool could improve the flow and pace of my work.

- Compiling sources for literature research
- Analysis of survey data
- Transcription and Translation of Interviews
- Writing and formatting my final Thesis

Out of these four areas, I considered survey data analysis to be the most time-consuming, so I decided to work on saving time here through implementing a digital tool.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Introduction
My Research
Topic
Data Analysis in
R

Scoping
Selection
Elaboration

Implementation
Creating a Tool
Chain
Elaborating on
my Chain
The Database

Results
Pros
Cons

Conclusion
Final Words
Code Repository

Scoping
Elaboration

In order to choose the correct tool to use, I had to refine my input and output sources. These went through many iterations, but eventually ended up like this:

- Input: Raw data collected from an online survey site that is capable of producing both qualitative and quantitative data and presenting it in a spreadsheet (CSV) format.

- Output: A dataset that can receive queries for specific subsets, answer these queries and plot the results in a graphical format.

Due to my choice of input and output, I was able to pick R as my intended tool-of-choice.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Implementation
Creating a Tool Chain

I needed to prove that R could take a basic spreadsheet (CSV) input and create my desired output. To do this, I created my own small table of data, before loading it into the R environment and constructing a plot.

This tool chain was a success, so I was then able to continue onto using more complex data, in a similar style to what I intended on using in my final thesis.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Implementation
Elaborating on my Chain

Now I knew that my tool chain worked with simple data, created a more complicated chain.

Google Forms survey (real-world data collection)
$\rightarrow$ Data analysis in R
$\rightarrow$ Graph imported into a Word document

The survey was completed by my FOAR705 class. With this new data I created a tabular database that could answer all the questions I asked.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Introduction
My Research
Topic
Data Analysis in
R

Scoping
Selection
Elaboration

Implementation
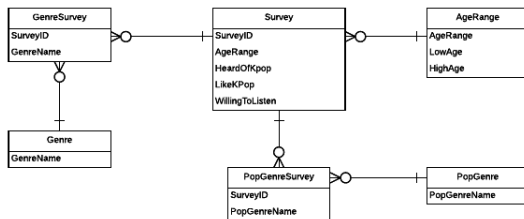Creating a Tool
Chain
Elaborating on
my Chain
The Database

Results
Pros
Cons

Conclusion
Final Words
Code Repository

Implementation
The Database

In order to avoid errors when asking my data various questions, I broke the
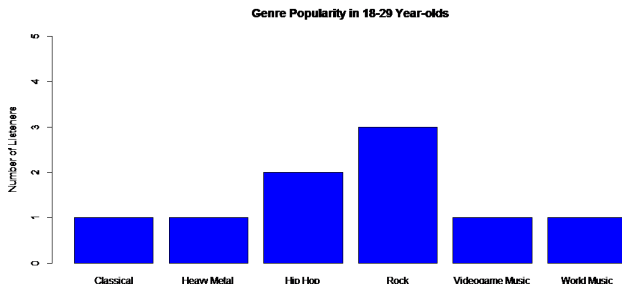database down into a number of connected tables.

These could be joined together at will when necessary for answering a question. In
order to prove my concept, I decided to examine 18-29 year olds' music tastes, and
place the data into a bar plot.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Results

Pros:

- General growth of my technical skillset
- Success in using R to create simple tool chains
- Increased knowledge of technical tools available for implementation in a range of tasks
- The structure of the database can be used for much larger spreadsheets with thousands of entries

Genre Popularity in 18-29 Year-olds

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Results
Cons:

- R support online is often outdated for the current version (code changes)
- Time is still currently an issue in terms of utilising this tool - however, I would say that if R is implemented properly, there will likely be fewer problems concerning data breaking under manipulation.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Conclusion

My implementation of R as a tool for data analysis has been successful.
I am now able to ask questions of R, and receive graphical responses with ease.
This success means that I can replicate the same process shown here with my
thesis data with minimal difficulty.

Analysing
Survey Data
in R

Kathryn
Phillips,
43686826

Conclusion
Code Repository

If you would like to have a look at my code and the collected data used in this
project, please follow the link below:

https://github.com/Kattkin/FOAR705-Coding-line