

Thesis Proposal

**Visualization methods for genealogical and
RNA-sequencing datasets**

Lindsay Rutter

Program of Study Committee:

Dianne Cook, Major Professor

Amy Toth, Major Professor

Heike Hofmann

Daniel Nettleton

James Reecy

May 16, 2016

Contents

Contents	2
1 Introduction	3
1.1 Background to data visualization	3
1.2 Problems to be addressed	3
1.3 Overview of thesis research	3
2 Visualization methods for genealogical datasets	4
2.1 Database structure	4
2.1.1 Question types	4

Introduction

1.1 Background to data visualization

Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah
blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah
blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah.

1.2 Problems to be addressed

Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah
blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah
blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah.

1.3 Overview of thesis research

Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah
blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah
blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah.

Visualization methods for genealogical datasets

2.1 Database structure

Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah
blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah
blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah. As
an example, the learning outcomes for **Topic 03** are provided in [List 1](#) below.

2.1.1 Question types

Blah blah blah Blah blah blah Blah blah blah Blah blah blah Blah blah blah
Blah blah blah Blah blah blah Blah blah blah Blah. More information about **Respondus**
can be found on their [website](#). As seen in [Figure 2.1](#).



Figure 2.1: Caption blah blah blah blah blah blah blah blah blah blah blah blah
blah blah blah blah blah blah blah

Now, we will run the [REFER TO SECTION](#) .

Table 2.1: Topic numbers and descriptions

Number	Description
01	Data
02	Descriptive Statistics for a Single Categorical Variable
03	Descriptive Statistics for a Single Quantitative Variable

List 1: Learning outcomes for Topic 03

- A. Use standardizing to determine how many standard deviations an observation is away from the mean value.

B. Use z-scores to compare observations for different quantitative variables.

C. Explain how standardizing affects the shape, center, and variability of the distribution of a quantitative variable.

Example question title: **T16.A.A.04-1.1.MC.1**

The absolute pathway to the `extdata` directory on your local computer can be determined by typing the following command into the **R** console:

```
system.file("inst/extdata/", package = "ePort")
```

```
keyHTM = system.file("inst/extdata/KeyFiles/Topic06.Questions.htm", package =
  "ePort")

refineKey(keyHTM)

keyPath = gsub("htm$", "txt", keyHTM)

dataPath = system.file("inst/extdata/DataFiles/Topic06/Topic06.A.csv", package =
  "ePort")

rewriteData(dataPath)

loPath = system.file("inst/extdata/LOFiles/Topic06.Outcomes.txt", package = "ePort")

outPath = system.file("inst/extdata/OutputFiles", package = "ePort")

makeReport(keyFile = keyPath, dataFile = dataPath, loFile = loPath, outFile =
  outPath)
```

```
merged = subsetData(mergedData, dataTable)
makeReport(outFile = outPath, unit = 2, reportType = "crossSecUnit", className =
  "Eng444", repeatLowScore = 70)
```

- One topic for one section - short version ("secTopicShort")
- One topic for one section - long version ("secTopicLong")
- One topic comparing multiple sections - short version ("crossSecTopicShort")

- One topic comparing multiple sections - long version ("crossSecTopicLong")
- One unit (group of topics) for one section ("secUnit")
- One unit (group of topics) comparing multiple sections ("crossSecUnit")