Intro to RQDA

Prabin Raj Shakya
prabinrs.github.io

Steps in Analysis of Qualitative Data

1. Preparation
2. Coding
3. Data Attributes
4. Codes abstraction
5. Code plotting and sharing
6. Theory building

1. Preparation

RQDA – based analysis

#1. Install R, RSQLite, GTK, RQDA
#2. Transform textual data into individual ASCII or .txt files
#3. Launch RQDA
#4. Create a new project
#5. Upload all .txt files into RQDA

Advantages and Disadvantages

- All data files are transportable online via emails,
- Dropbox, or using a device (e.g. USB stick)
- Ease of logistics handling when moving offices,
- involving multiple researchers
- Not all researchers are familiar with R

Let’s start with:

What is R

Free
Software for
statistical computation
and graphics

What is R Studio

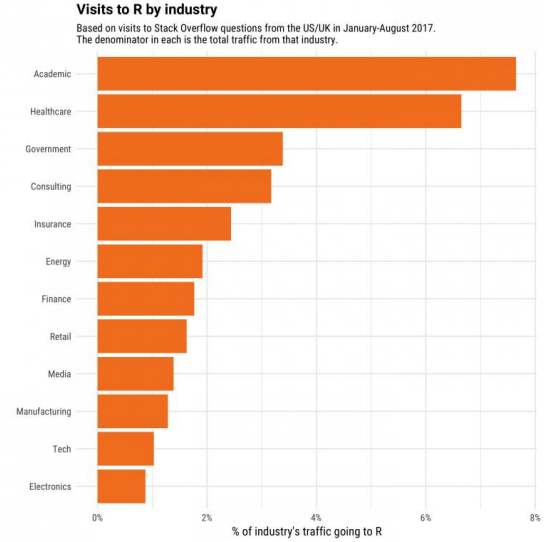
1) A company

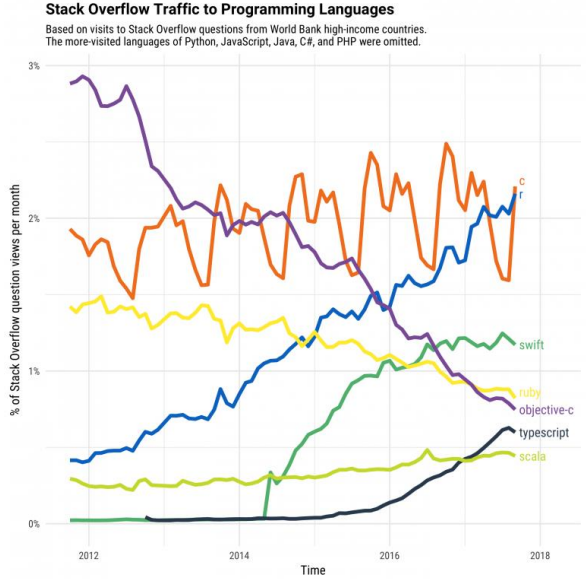
2) A software program that makes working in R easier

IDE: Integrated Development Environment

Why R

Language Rank	Types	Spectrum Ranking
1. C		100.0
2. Java		98.1
3. Python		98.0
4. C++		95.9
5. R		87.9
6. C#		86.7
7. PHP		82.8
8. JavaScript		82.2
9. Ruby		74.5
10. Go		71.9





Why R

Full programming language: customizable, extensible

Open source: contributed packages

Designed for data analysis

Why R

Support for reproducible analysis

Widely used

R Community

R Community

#rstats

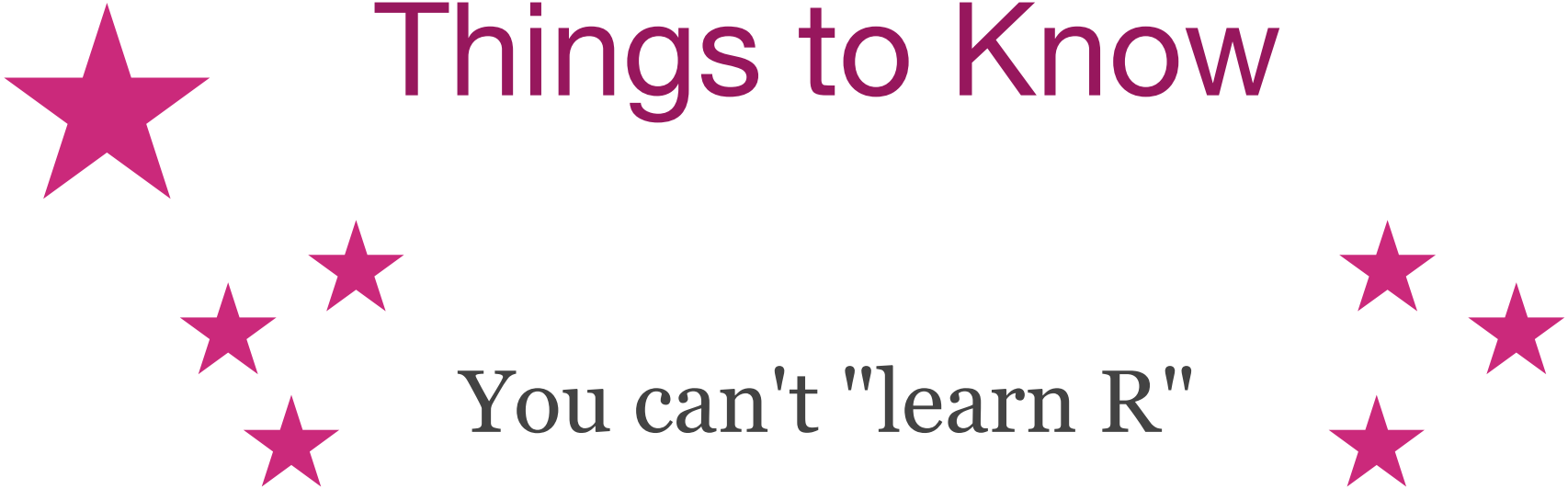
RStudio Community https://community.rstudio.com/

R Bloggers https://www.r-bloggers.com/

R User Group Nepal https://www.r-nepal.org/

Things to Know About R

Things to Know



You can't "learn R"

You learn the basics

and learn how to learn more

Things to Know

It's expected that you'll install and use additional packages

Things to Know

There are multiple ways

to do most things

Some ways are better than others

(efficient)

Things to Know

You can, and often will, have more than one dataset open in R at the same time

Things to Know

Run R in multiple ways:

Interactively: R console, RStudio, R Notebooks, Jupyter Notebooks

Batch jobs

Recently it got on cloud as well

R packages


CRAN
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
GitHub

Installing Packages

RCRAN network

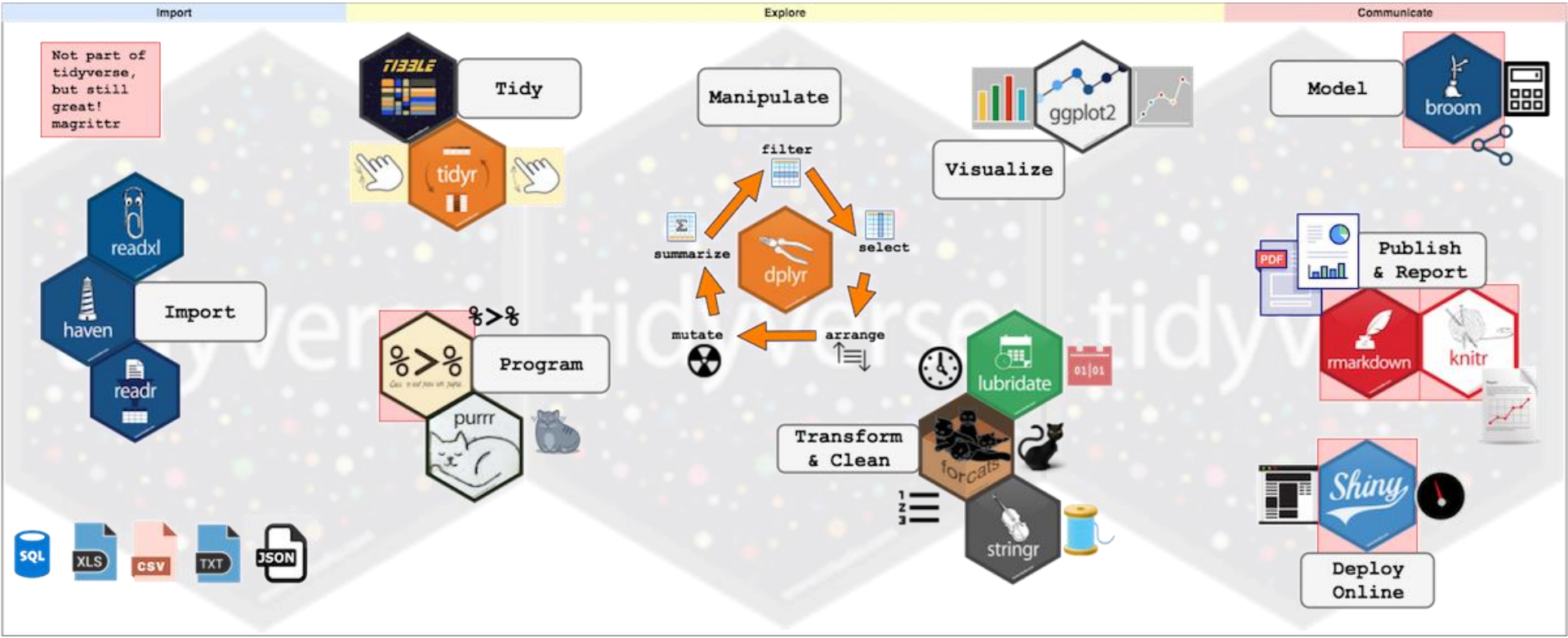
```
install.packages(“tidyverse”) #install the packages
library(tidyverse) #load the packages
```

Bioconductor

```
if (!requireNamespace("BiocManager"))
    install.packages("BiocManager")
BiocManager::install()
BiocManager::install(c("GenomicFeatures", "AnnotationDbi"))
```

Github

```
library(devtools)
devtools::install_github("strengejacke/sjPlot")
```


Import
Explore
Communicate
Not part of tidyverse, but still great! magrittr
Tidy
tibble
tidyr
readxl
haven
readr
Import
%>%
Program
purrr
SQL
XLS
csv
TXT
JSON
Manipulate
filter
select
arrange
mutate
summarize
dplyr
ggplot2
Visualize
lubridate
forcats
stringr
Transform & Clean
Model
broom
Publish & Report
PDF
rmarkdown
knitr
Shiny
Deploy Online

Install and Download

https://github.com/prabinrs/intro_RQDA

2. Coding

- One or more coder performs inductive data analysis together, by highlighting the text and clicking *“Mark”* (*“Unmark”*) to do (undo) the coding
- Other analyst(s) critiques the coding results; reconcile differences in interpretation

- Increased reliability and accuracy of data interpretation
- Ease of capturing new, unexpected insights through re-coding of data
- Possible mechanical errors for inexperienced users
- Possible loss of reflexivity or de-contextualizing data

Code and Categories

Inductive coding or Open Coding (Strauss and Corbin, 1990; Gioia et al. 2013)

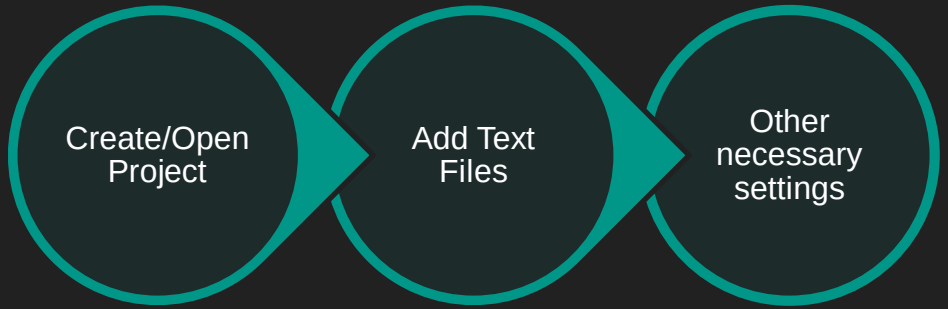
- Code a textual unit that is close to (i.e basic or lowest level) the data without being predicated on any theory.

Eg “they used my debit card to charge amounts I never authorized” : code “Monetary Loss”

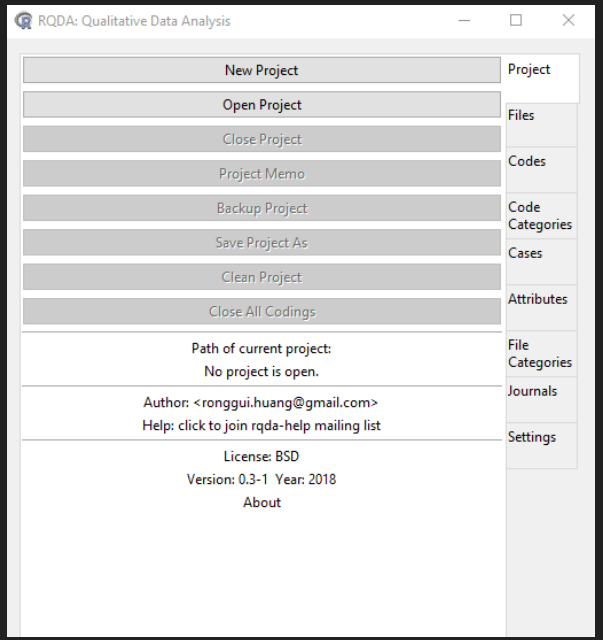
Deductive Coding: (Fereday and Muir-Cochrane, 2006; Bazeley and Jackson, 2013)

- Categories are coded based on established concept, variables or theories

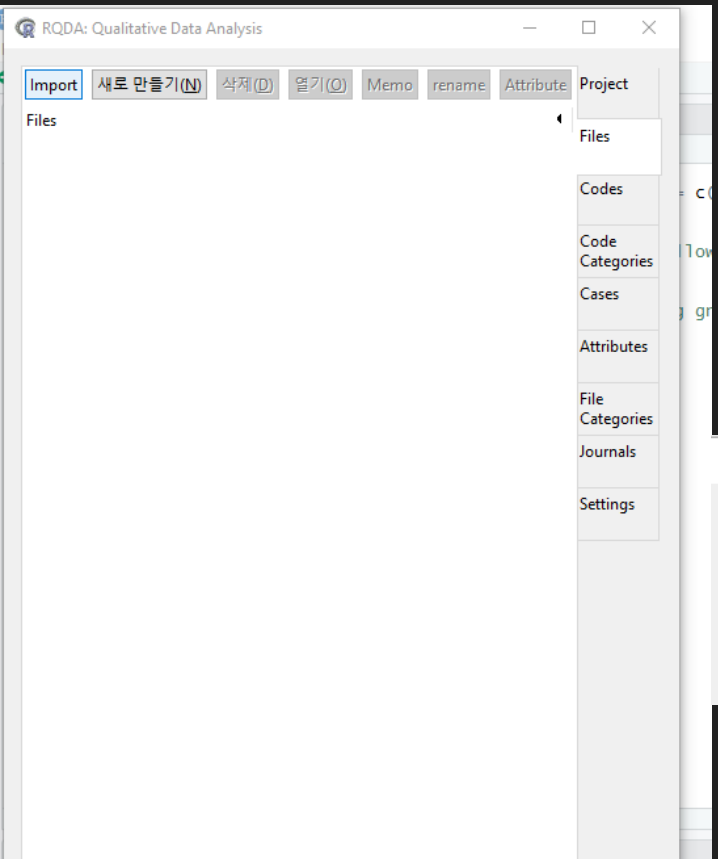
Eg “the customer service department doesn’t work well with the technical support and that’s why they took six months to solve my technical problems; Code “Market Orientatin” (based on Jaworki and Kohli, 1993)

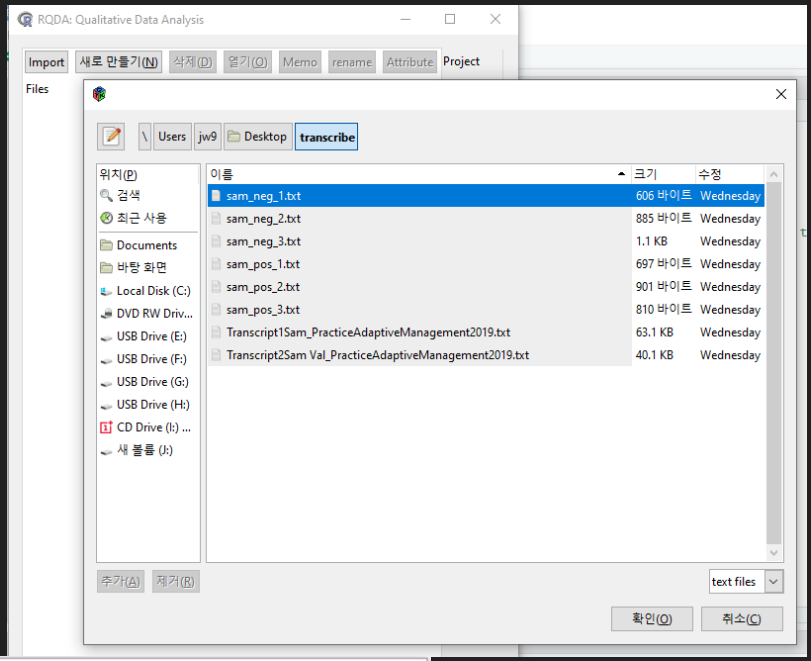

Create/Open Project
Add Text Files
Other necessary settings

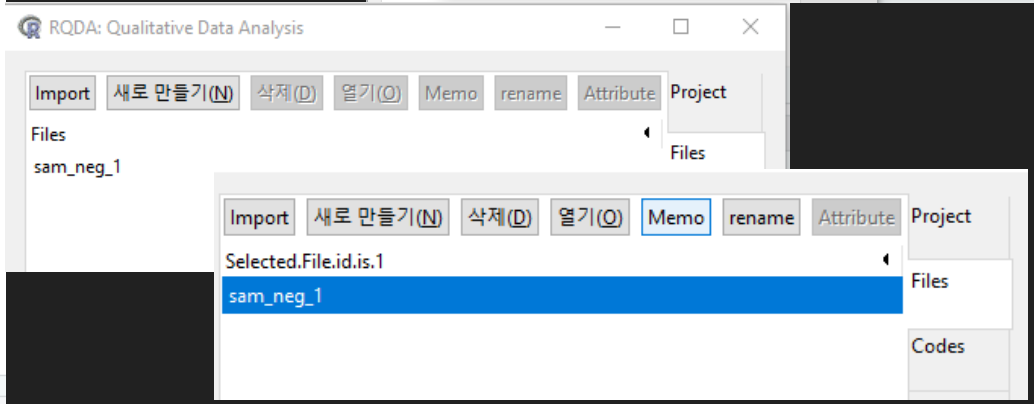
New Project


RQDA: Qualitative Data Analysis
New Project
Open Project
Close Project
Project Memo
Backup Project
Save Project As
Clean Project
Close All Codings
Path of current project:
No project is open.
Author: <ronggui.huang@gmail.com>
Help: click to join rqda-help mailing list
License: BSD
Version: 0.3-1 Year: 2018
About
Project
Files
Codes
Code Categories
Cases
Attributes
File Categories
Journals
Settings

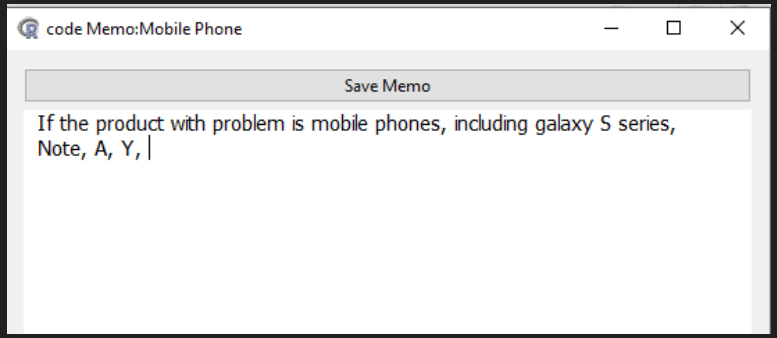
Import Files

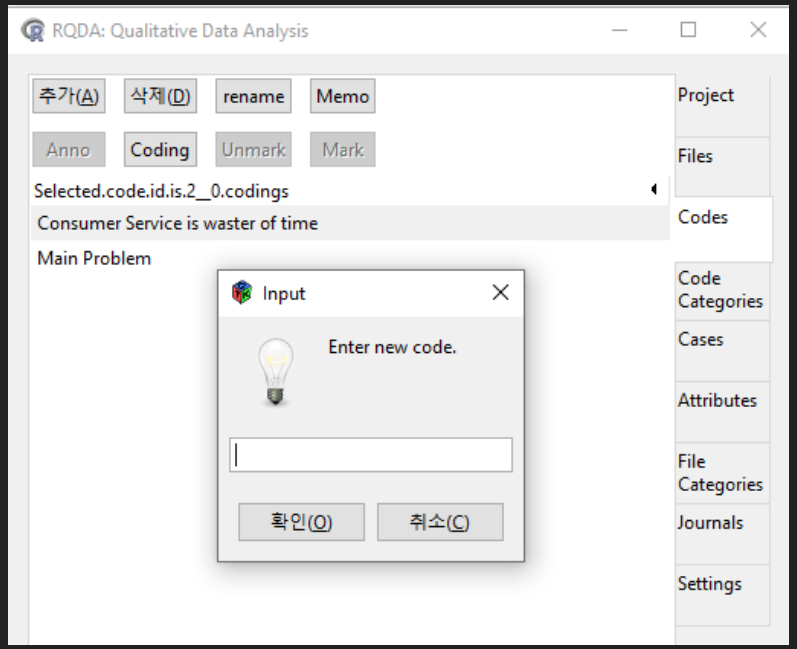






Codes


code Memo:Mobile Phone
Save Memo
If the product with problem is mobile phones, including galaxy S series, Note, A, Y,


RQDA: Qualitative Data Analysis
추가(A)
삭제(D)
rename
Memo
Anno
Coding
Unmark
Mark
Selected.code.id.is.2__0.codings
Consumer Service is waster of time
Main Problem
Input
Enter new code.
확인(O)
취소(C)
Project
Files
Codes
Code Categories
Cases
Attributes
File Categories
Journals
Settings

3. Data Attributes

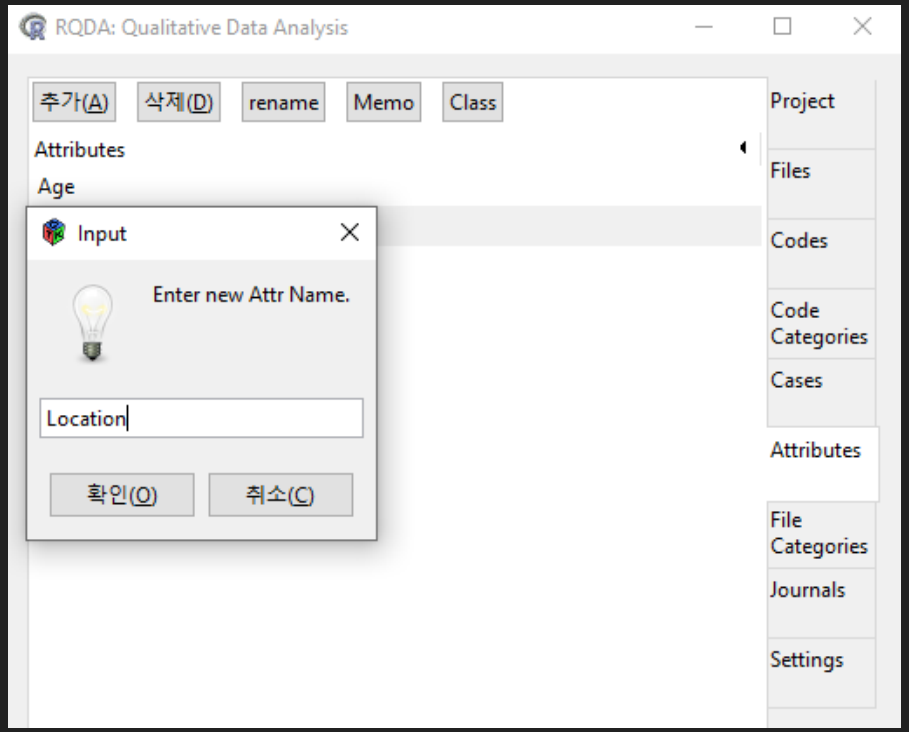
- Click “Attributes” and record key data attributes (e.g. data sources, informants’ demographics, dependent variables)
- Click *“Memo”* to write memos for each case

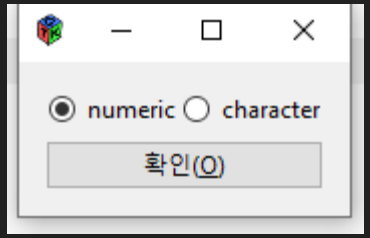
- Allow easy capture of meta information of the data and attributes of informants
- Memos can be copied-and-pasted to a Word processor to aid data analysis

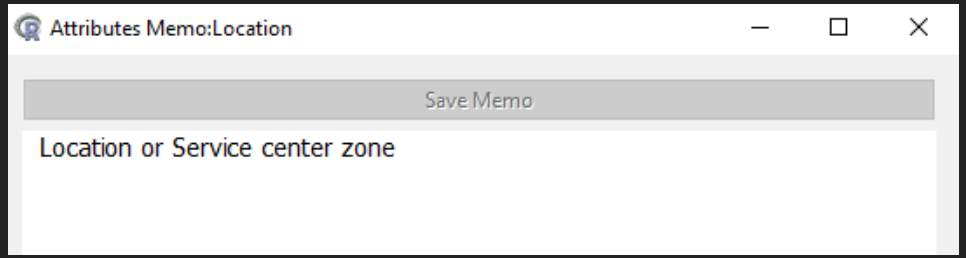
- that coders add data attributes or memos to enrich understanding
- enables users to add variables to individual files to show a subset of files and to perform statistical analysis on attributes.

Eg:
Attributes: Type of File,
Memo : Interview with p07 was conducted in presence of her colleagues and seniors.

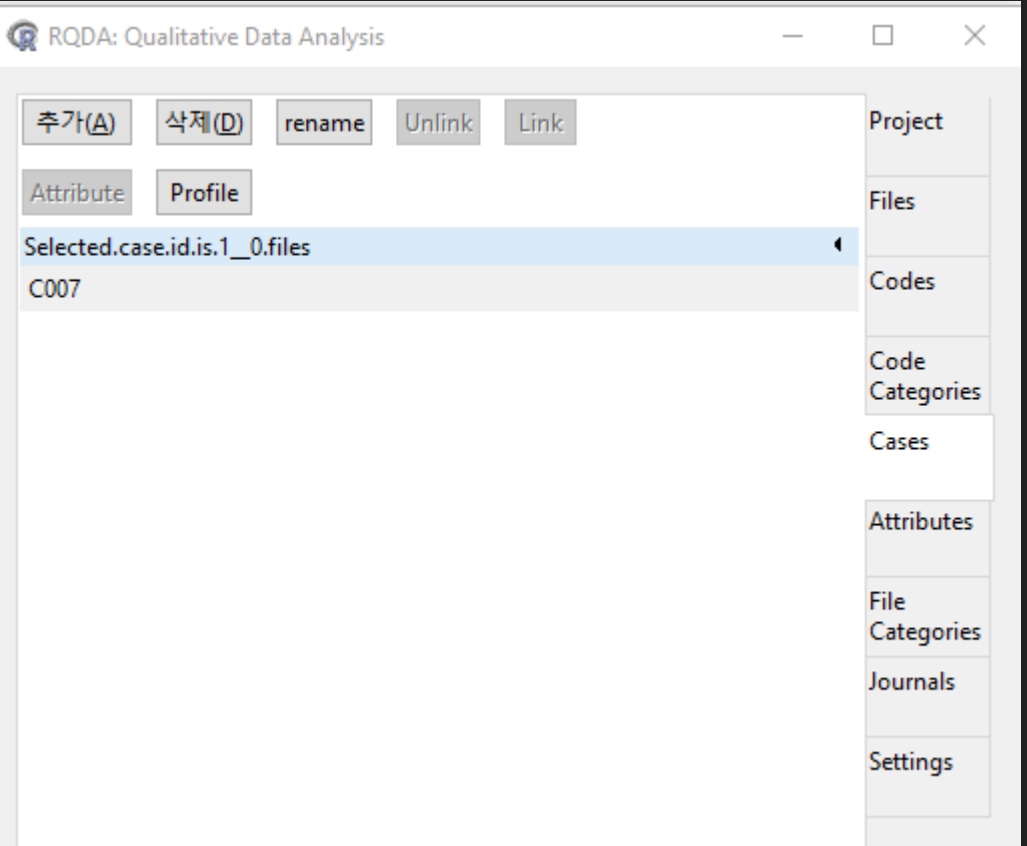
Attributes

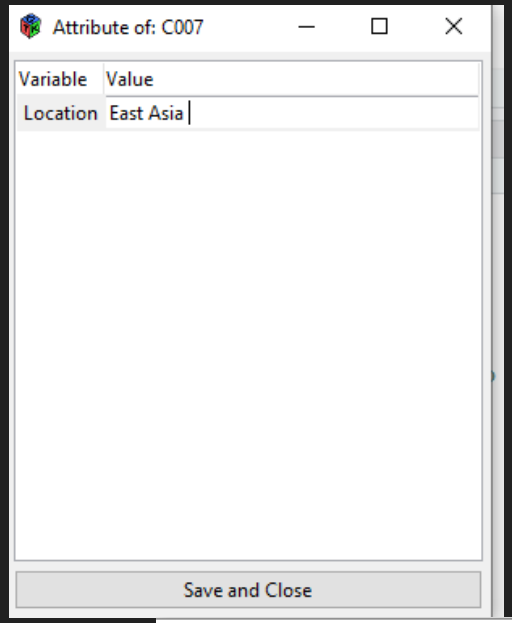

RQDA: Qualitative Data Analysis
추가(A)
삭제(D)
rename
Memo
Class
Attributes
Age
Input
Enter new Attr Name.
Location
확인(O)
취소(C)
Project
Files
Codes
Code Categories
Cases
Attributes
File Categories
Journals
Settings

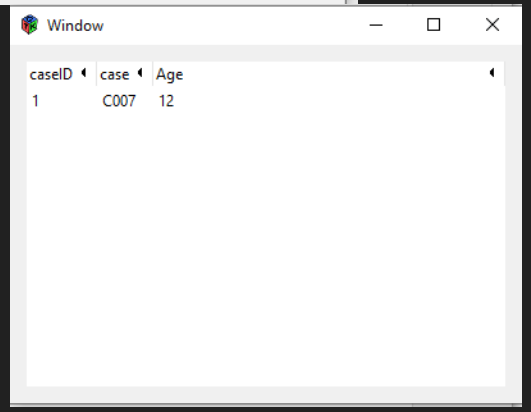

numeric
character
확인(O)


Attributes Memo:Location
Save Memo
Location or Service center zone

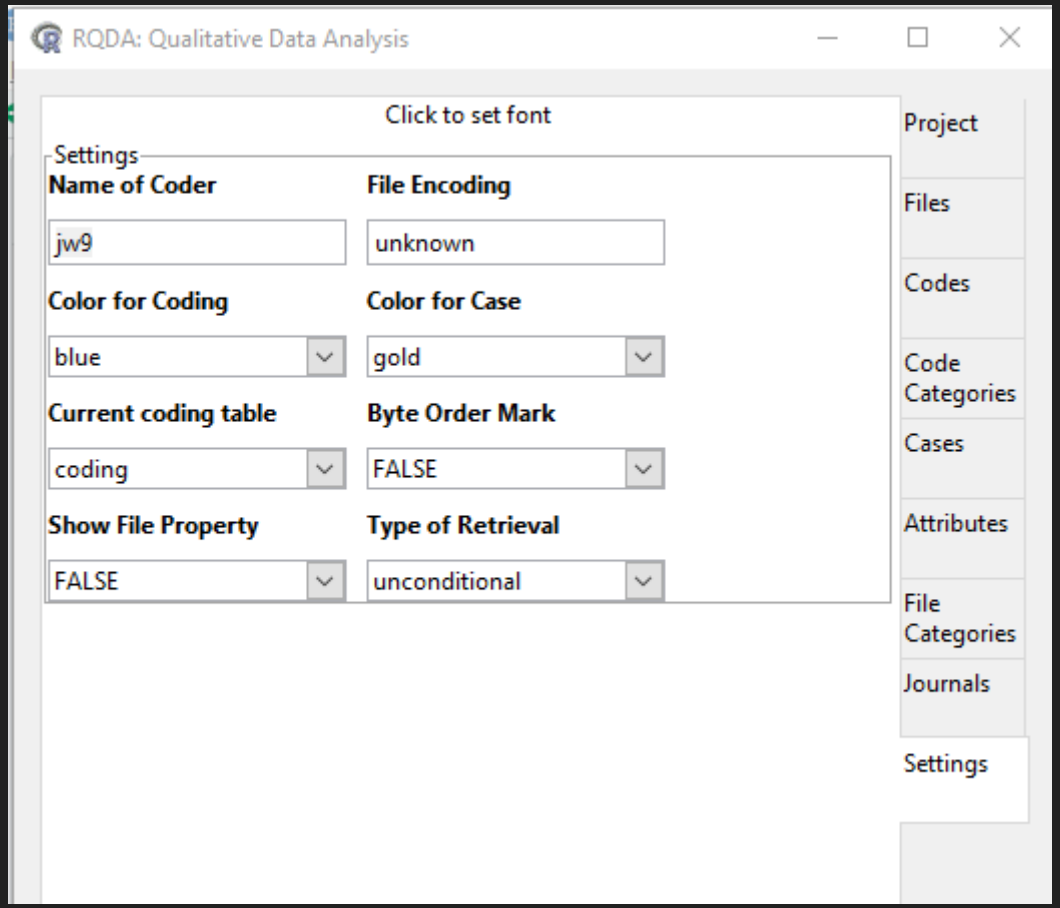
Cases


RQDA: Qualitative Data Analysis
추가(A)
삭제(D)
rename
Unlink
Link
Attribute
Profile
Selected.case.id.is.1__0.files
C007
Project
Files
Codes
Code Categories
Cases
Attributes
File Categories
Journals
Settings

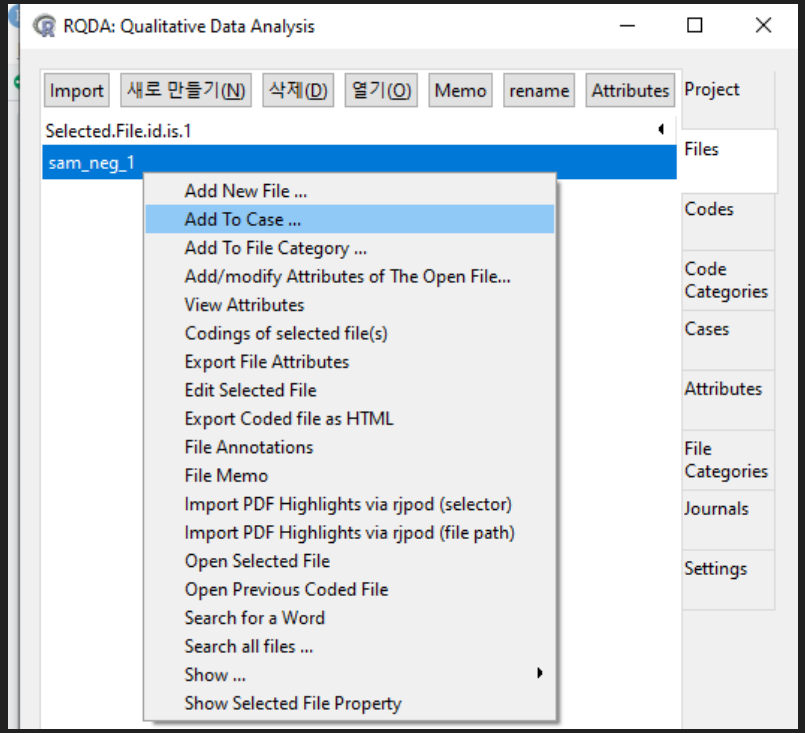

Attribute of: C007
Variable
Value
Location
East Asia
Save and Close

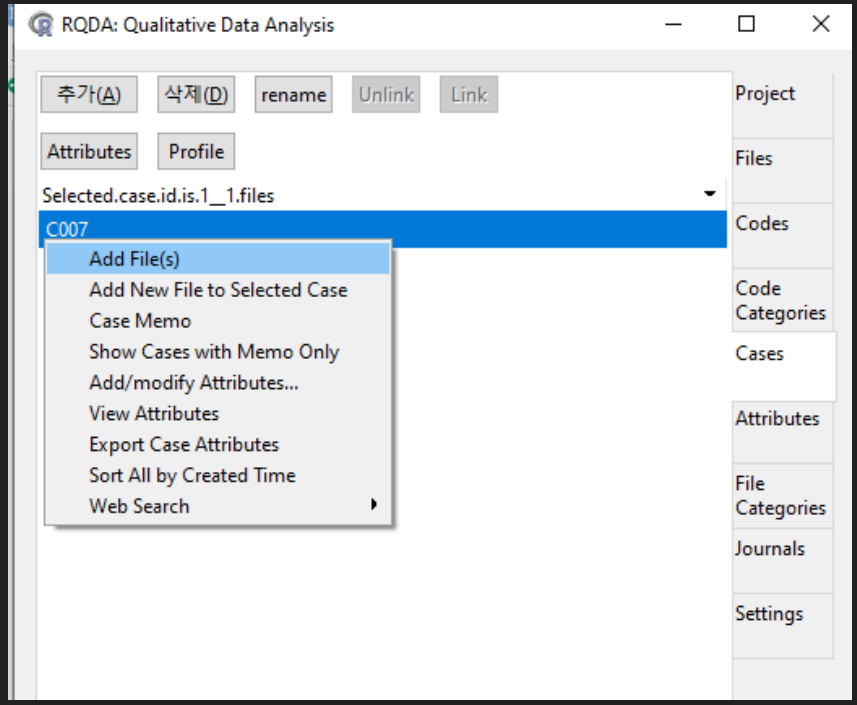

Window
caseID
case
Age
1
C007
12

Settings

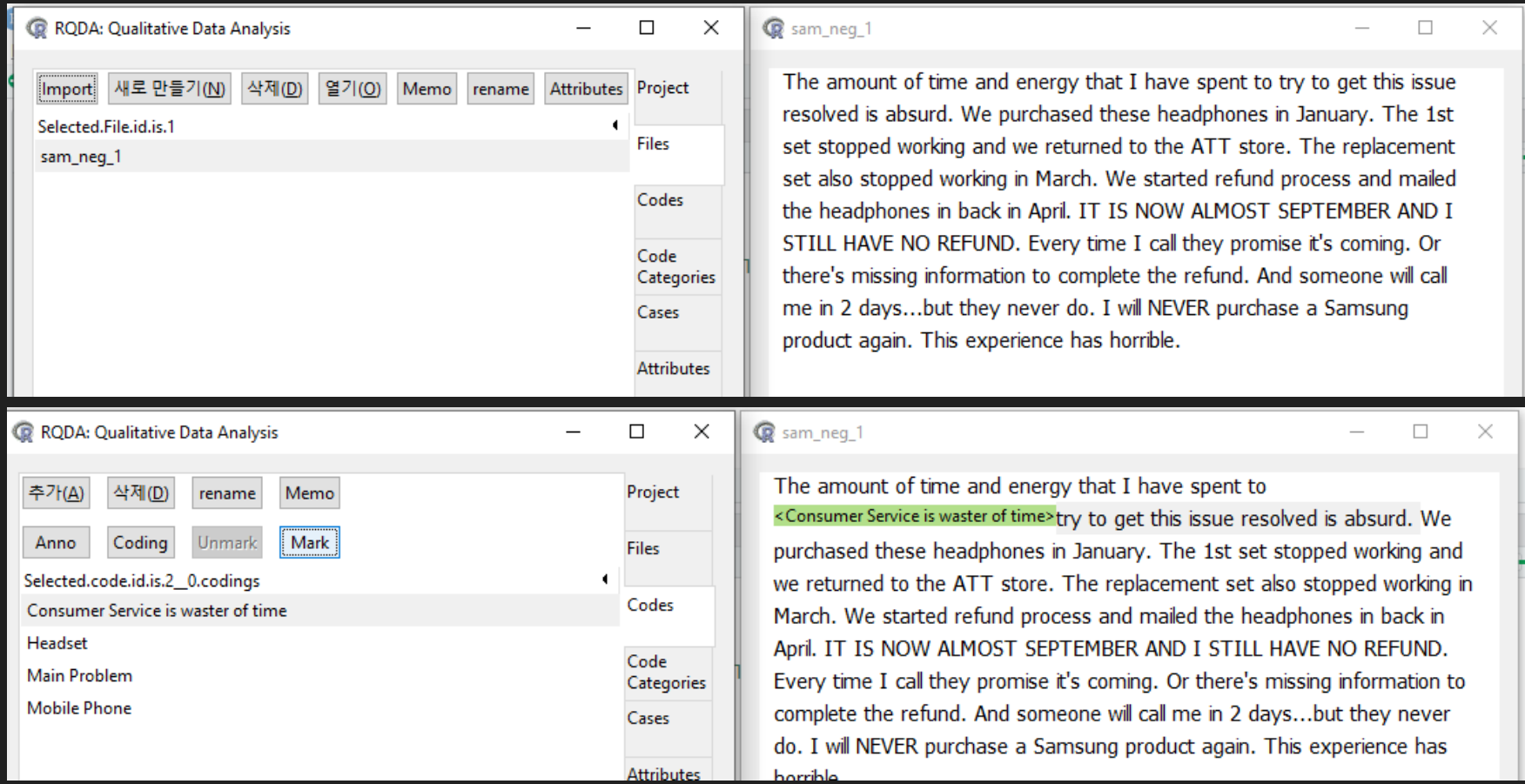

RQDA: Qualitative Data Analysis
Click to set font
Settings
Name of Coder
jw9
File Encoding
unknown
Color for Coding
blue
Color for Case
gold
Current coding table
coding
Byte Order Mark
FALSE
Show File Property
FALSE
Type of Retrieval
unconditional
Project
Files
Codes
Code Categories
Cases
Attributes
File Categories
Journals
Settings

Add File to case





Coding


RQDA: Qualitative Data Analysis
Import
새로 만들기(N)
삭제(D)
열기(O)
Memo
rename
Attributes
Project
Selected.File.id.is.1
sam_neg_1
Files
Codes
Code Categories
Cases
Attributes
sam_neg_1
The amount of time and energy that I have spent to try to get this issue resolved is absurd. We purchased these headphones in January. The 1st set stopped working and we returned to the ATT store. The replacement set also stopped working in March. We started refund process and mailed the headphones in back in April. IT IS NOW ALMOST SEPTEMBER AND I STILL HAVE NO REFUND. Every time I call they promise it's coming. Or there's missing information to complete the refund. And someone will call me in 2 days...but they never do. I will NEVER purchase a Samsung product again. This experience has horrible.
RQDA: Qualitative Data Analysis
추가(A)
삭제(D)
rename
Memo
Anno
Coding
Unmark
Mark
Project
Files
Selected.code.id.is.2__0.codings
Consumer Service is waster of time
Headset
Main Problem
Mobile Phone
Codes
Code Categories
Cases
Attributes
sam_neg_1
The amount of time and energy that I have spent to
<Consumer Service is waster of time>try to get this issue resolved is absurd. We purchased these headphones in January. The 1st set stopped working and we returned to the ATT store. The replacement set also stopped working in March. We started refund process and mailed the headphones in back in April. IT IS NOW ALMOST SEPTEMBER AND I STILL HAVE NO REFUND. Every time I call they promise it's coming. Or there's missing information to complete the refund. And someone will call me in 2 days...but they never do. I will NEVER purchase a Samsung product again. This experience has horrible

4. Codes Abstraction

- Use *“Add To”* button to assign all first order codes to second-order code categories

To create third-order code categories, copy and paste all codes and re-arrange them using color font with Excel

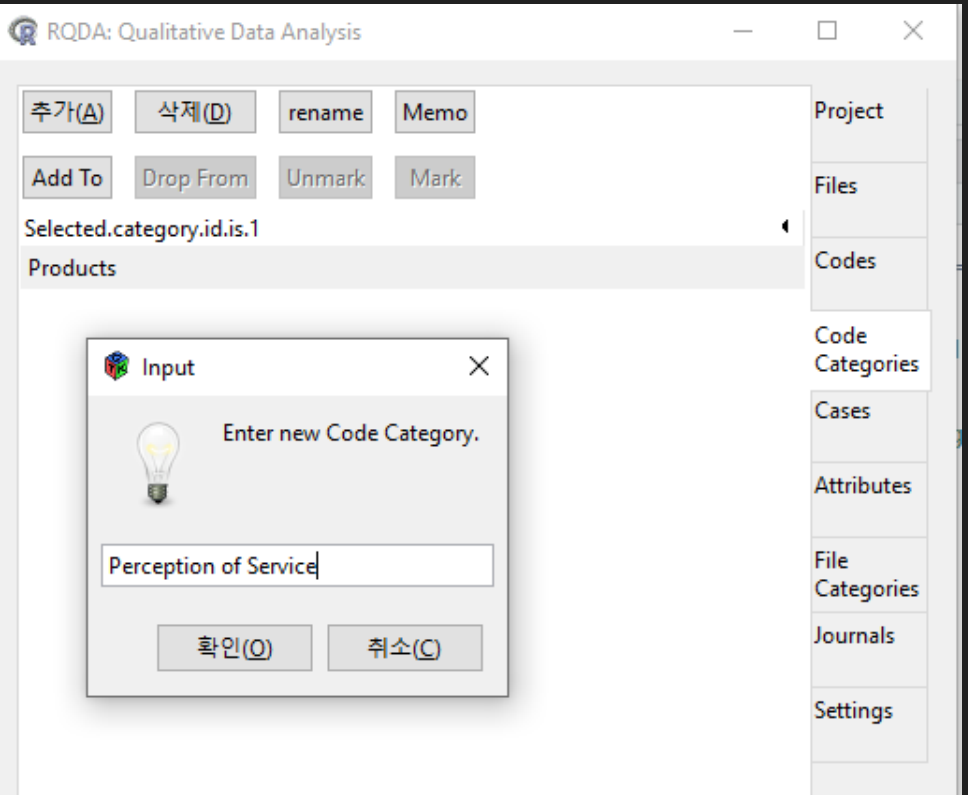
Based on Constructivist theory-building Process

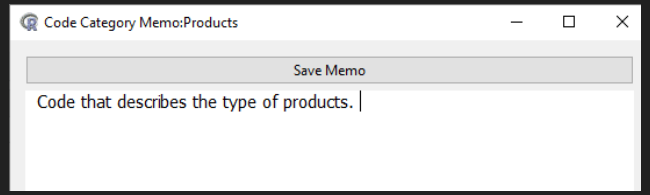
- A highly systematic way of aggregating codes to a higher level of meaning
- Challenging to aggregate too many first-order codes (thousands of codes) to second-order codes

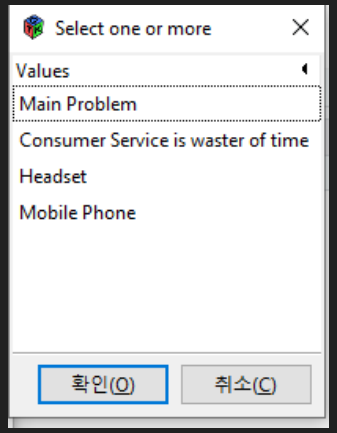
Abstraction to Second Level categories i.e after coding completion, first level codes can be aggreagated into higher-level **code categories**

Eg. “they used my debit card to charge amounts that I never authorize” and “this faulty laptop means I couldn’t work for four days” can be aggregated to a higher-level concept called **“monetary loss”**

Code Categories







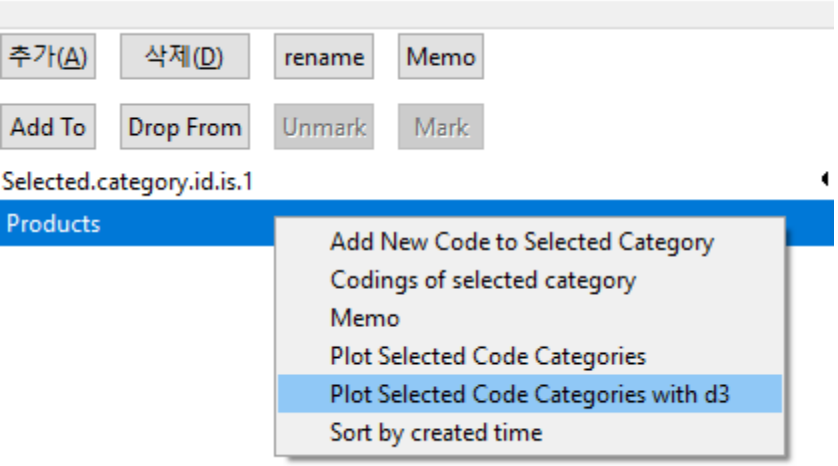
5. Codes Plotting and Sharing

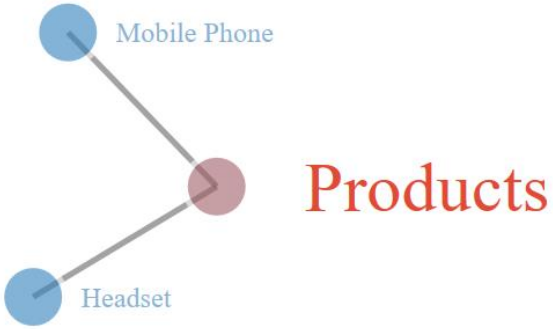
- Select all second-order using cursor, right click and press “Plot selected Code categories”.

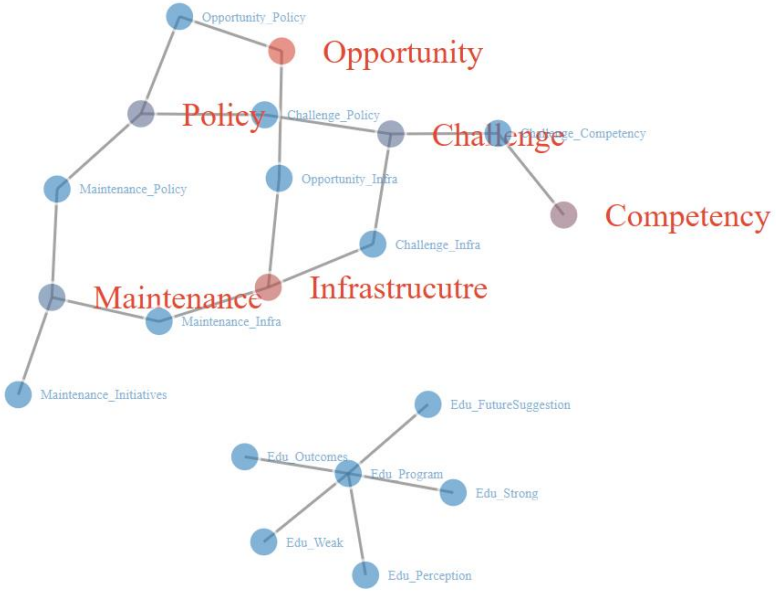
- Produce images of networks of codes
- Can re-arrange network codes using various network plot models (e.g

Allows User to visualize the coding results, enhance the data interpretation and analysis process.
Transition from an inductive to abductive theorizing process- moving back and forth between the emergent findings and literature/theories to find new concepts within the data.

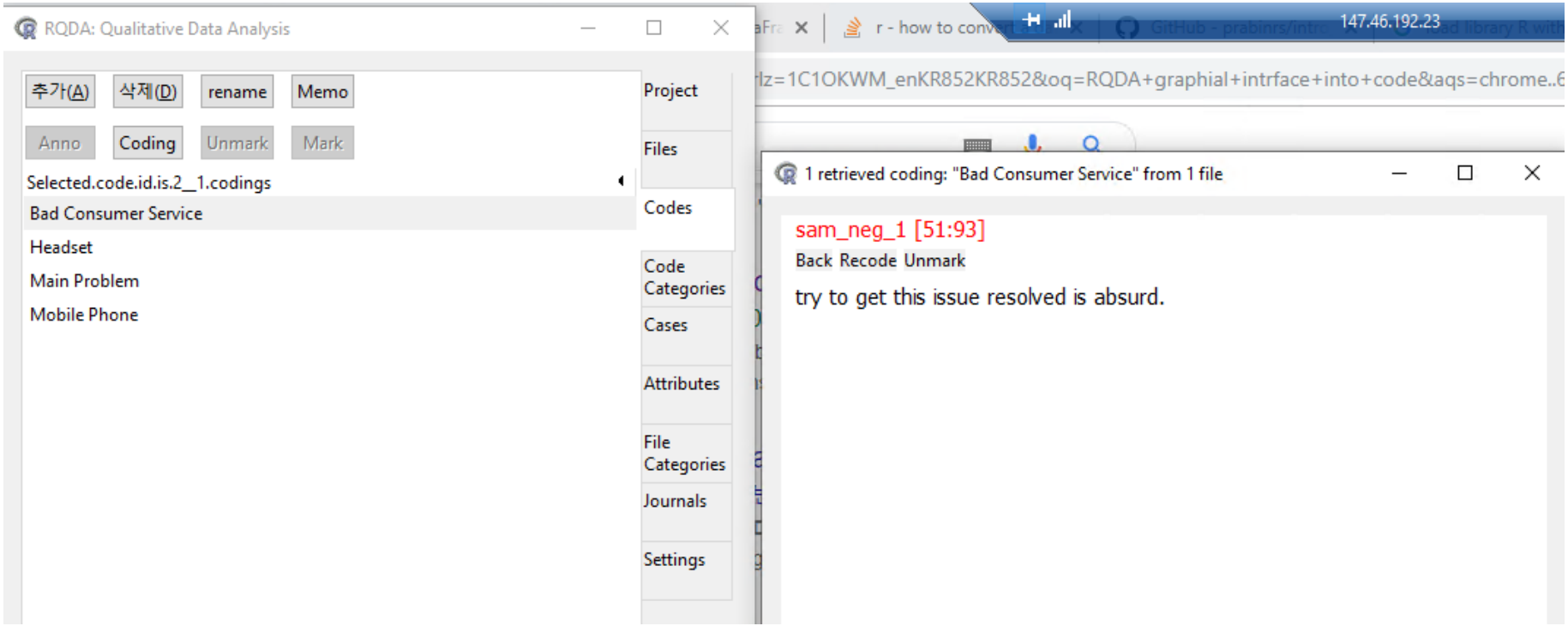
Visualize the concepts



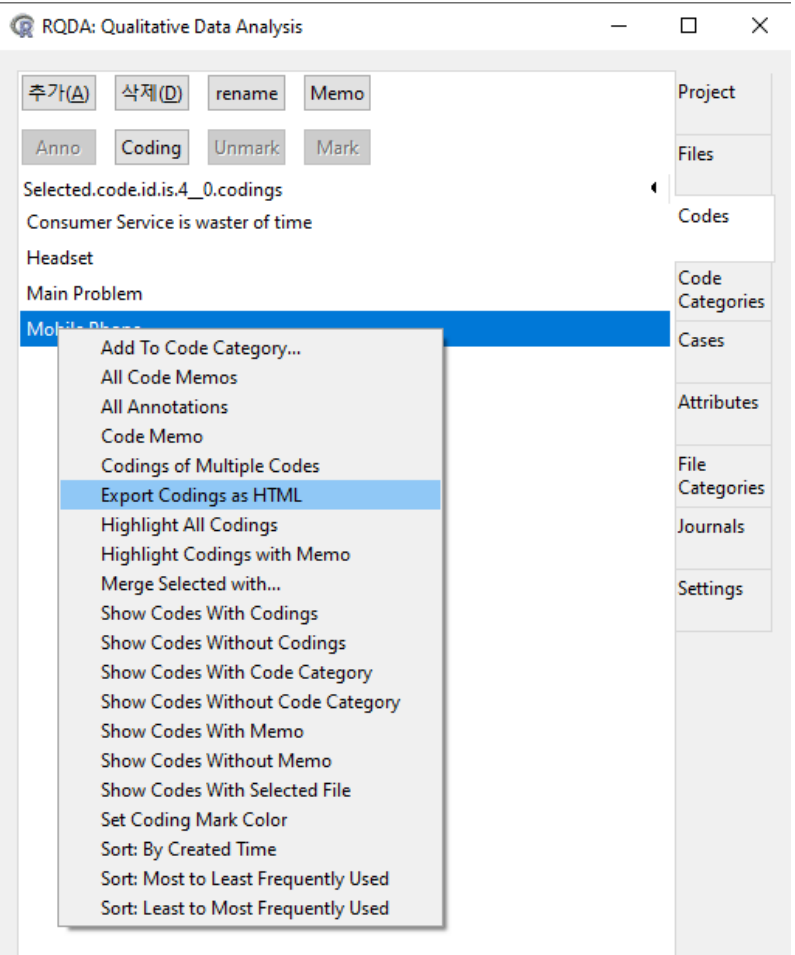


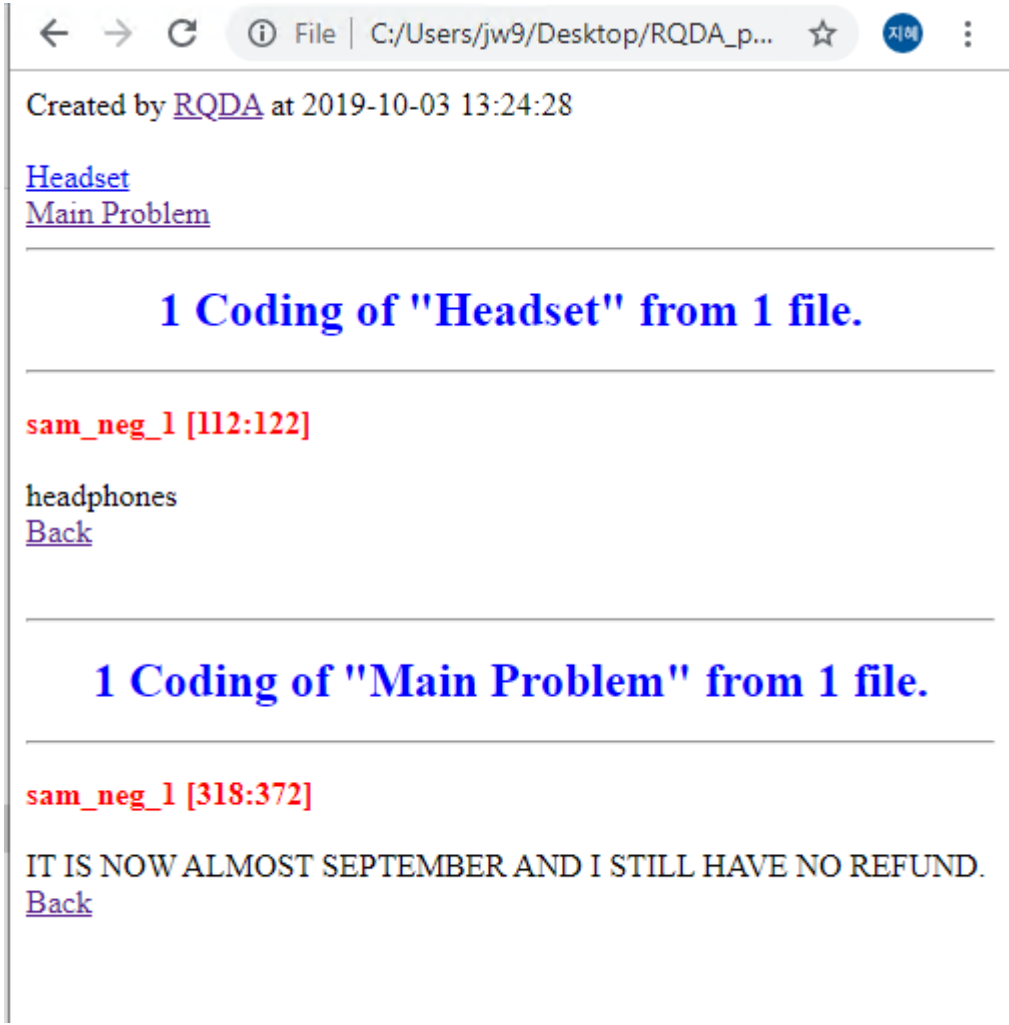


Review codes and Export



Review codes and Export





Complex Queries

- RQDA uses SQLite for the data management,
- SQL Queries can be used from R (not RQDA GUI)
- Build in Functions

```
#Load Data from RQDA

openProject("leaders.rqda")

interviews <- RQDAQuery("SELECT file FROM source")
interviews$file <- apply(interviews, 1, function(x) gsub("…", "...", x))
interviews$file <- apply(interviews, 1, function(x) gsub("'", "", x))
interviews <- Corpus(VectorSource(interviews$file))
interviews <- tm_map(interviews, stripWhitespace)
interviews <- tm_map(interviews, content_transformer(tolower))
interviews <- tm_map(interviews, removeWords, stopwords("english"))
interviews <- tm_map(interviews, removePunctuation)
interviews <- tm_map(interviews, removeNumbers)
interviews <- tm_map(interviews, removeWords, c("L01","L02","L03","L04" ))
interviews <- tm_map(interviews, removeWords, c("prabin", "raj", "shakya","jk", "like", "jieun", "know"))
interviews <- tm_map(interviews, removeWords, stopwords(kind = "en"))
interviews <- tm_map(interviews, removeWords, c("prs", "will", "can", "quite",
                                                "day", "yes", "even", "give",
                                                "done", "also", "actually", "yeah", "things", "thing",
                                                "see", "much", "just", "bit", "maybe",
                                                "got", "kind", "tow", "example"))

#wordCloud
set.seed(1985)
wordcloud(interviews, min.freq = 10, max.words =50, rot.per=0.35,
          colors = brewer.pal(8, "Blues")[-1:-5])

#Topic analysis
leader_dtm<-DocumentTermMatrix(interviews)
leader_dtm<-removeSparseTerms(leader_dtm,0.99)
leaderTerms<-LDA(leader_dtm, k=5)
terms(leaderTerms,10)
```

Complex Queries

help(RQDA)
help(RQDATables")
help("RQDAQuery")

```
# Few useful inbuild functions
filesByCodes() # files contain codes with frequency
getCaseIds(fid) # Retrieve the number of cases and the case name
getCaseNames(caseID) # Return the names of the IDs
getCaseIds(fid) #Return the case Name or IDs
summaryCodings() # return coding frequencies
getCodingTable() # get details on the codes
getCodingsByOne() # serach for specific code in the data.
codingBySearch() # finding speciific text in file and apply specific code.
```

```
filesByCodes() # files contain codes with frequency
fid  filename codedBy.Bad Consumer Service codedBy.Headset codedBy.Main Problem
  1 sam_neg_1                            1               1                     1
```

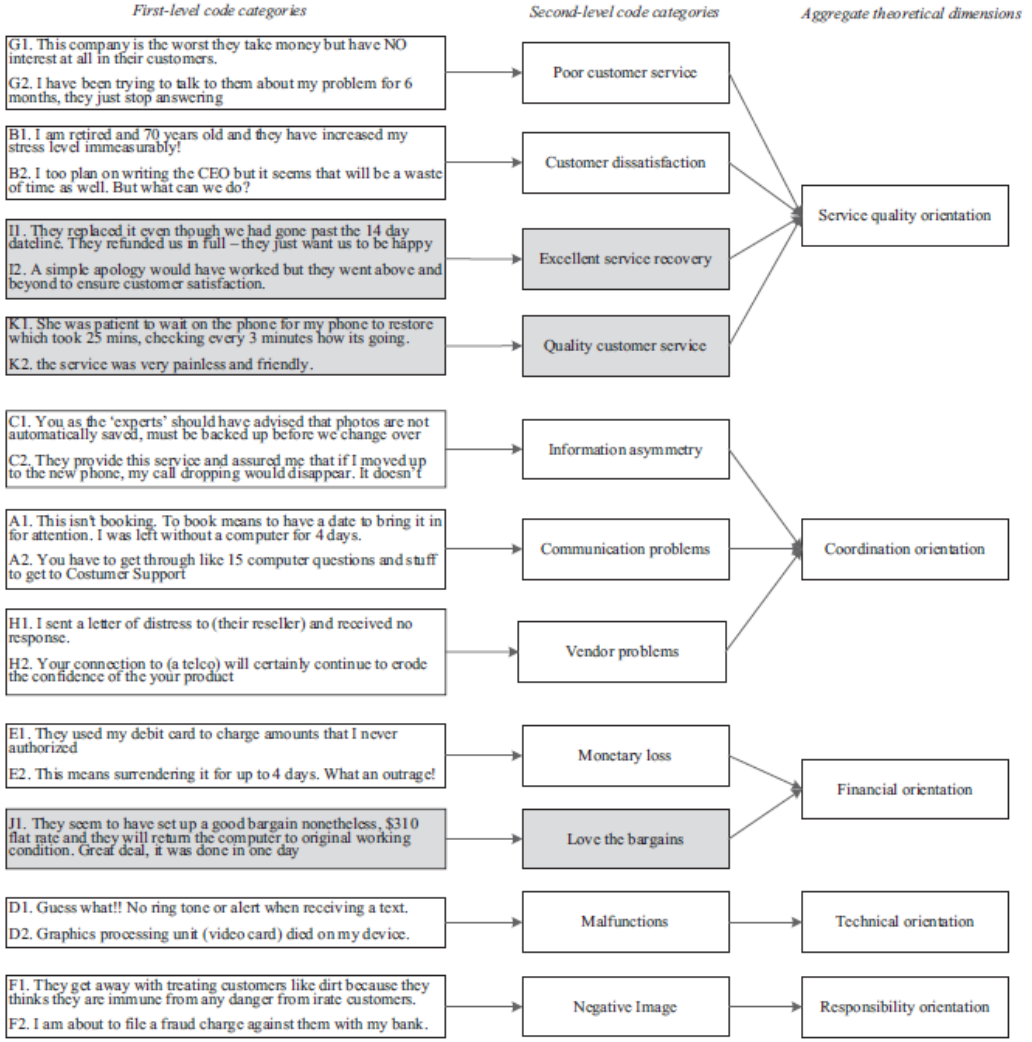
Codes are available in gitub, please open Rproj file and play around.

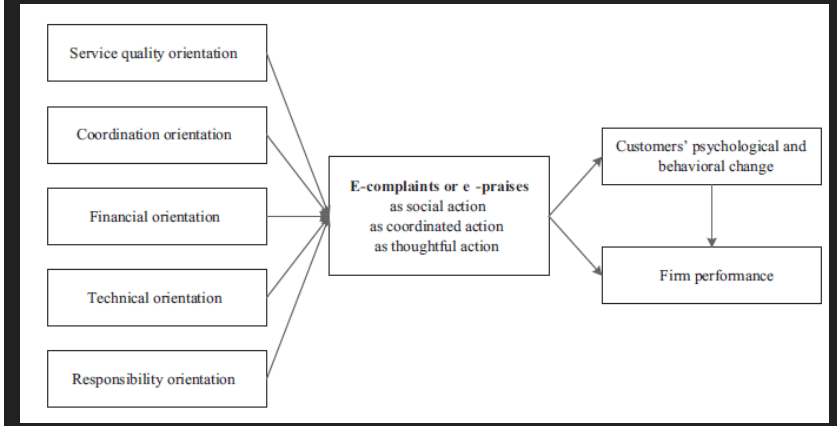
6. Theory Building

- Identify the relationships among themes / concepts / variables that emerged to create a process mode.

- This is manual process that can only be done using human interpreter.

iteratively re-categorized all first- and second-level code categories in a “data structure” to refine codes to the best possible higher-level categories that describe and explain the empirical reality and the literature. Should give clear ideas on who’s and why’s


First-level code categories
Second-level code categories
Aggregate theoretical dimensions
G1. This company is the worst they take money but have NO interest at all in their customers.
G2. I have been trying to talk to them about my problem for 6 months, they just stop answering
Poor customer service
B1. I am retired and 70 years old and they have increased my stress level immeasurably!
B2. I too plan on writing the CEO but it seems that will be a waste of time as well. But what can we do?
Customer dissatisfaction
I1. They replaced it even though we had gone past the 14 day dateline. They refunded us in full – they just want us to be happy
I2. A simple apology would have worked but they went above and beyond to ensure customer satisfaction.
Excellent service recovery
Service quality orientation
K1. She was patient to wait on the phone for my phone to restore which took 25 mins, checking every 3 minutes how its going.
K2. the service was very painless and friendly.
Quality customer service
C1. You as the 'experts' should have advised that photos are not automatically saved, must be backed up before we change over
C2. They provide this service and assured me that if I moved up to the new phone, my call dropping would disappear. It doesn't
Information asymmetry
A1. This isn't booking. To book means to have a date to bring it in for attention. I was left without a computer for 4 days.
A2. You have to get through like 15 computer questions and stuff to get to Costumer Support
Communication problems
Coordination orientation
H1. I sent a letter of distress to (their reseller) and received no response.
H2. Your connection to (a telco) will certainly continue to erode the confidence of the your product
Vendor problems
E1. They used my debit card to charge amounts that I never authorized
E2. This means surrendering it for up to 4 days. What an outrage!
Monetary loss
J1. They seem to have set up a good bargain nonetheless, $310 flat rate and they will return the computer to original working condition. Great deal, it was done in one day
Love the bargains
Financial orientation
D1. Guess what!! No ring tone or alert when receiving a text.
D2. Graphics processing unit (video card) died on my device.
Malfunctions
Technical orientation
F1. They get away with treating customers like dirt because they thinks they are immune from any danger from irate customers.
F2. I am about to file a fraud charge against them with my bank.
Negative Image
Responsibility orientation


Service quality orientation
Coordination orientation
Financial orientation
Technical orientation
Responsibility orientation
E-complaints or e -praises
as social action
as coordinated action
as thoughtful action
Customers' psychological and behavioral change
Firm performance