Exploring Qualtrics Surveys with R

Ambreen Chaudhri

1/13/2023

R Markdown Installation

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

To install R Markdown, you will need to run the following install packages lines and load library lines:

```
# install packages
#install.packages('rmarkdown')

# load packages
library(rmarkdown)

# load additional packages
library(qualtRics)
library(dplyr)
library(stringr)
library(vtable)
#library(stargazer)
```

Including Code Chunks and Results

You can embed an R code chunk in the space below. Note that this line will show you both the R command you used and the results.

```
# Clear Workspace
rm(list=ls())

# set working directory
setwd("~/MKTG_531")

# RUN THIS STEP FIRST TIME ONLY, TO GENERATE ".Renviron" FILE
#qualtrics_api_credentials(api_key = "INSERT YOUR API KEY HERE",
#                          base_url = "ca1.qualtrics.com",
#                          install = TRUE,
#                          overwrite = TRUE)

readRenviron("~/.Renviron")
outpath = "~/MKTG_531"

# Show all surveys and select one project by name
surveys   <- all_surveys()
surveys
```

```
## # A tibble: 86 x 6
##    id                 name            ownerId lastModified creationDate isActive
##    <chr>              <chr>           <chr>   <chr>        <chr>        <lgl>
##  1 SV_0H63qPlpEPK7WPs Post-Survey: D~ UR_4OW~ 2022-08-10T~ 2022-07-21T~ TRUE
##  2 SV_2gHUqeygiXUF74F Faculty Subscr~ UR_0fk~ 2018-10-03T~ 2017-07-11T~ TRUE
##  3 SV_37AJQNTlXAdBUaN Academic Props~ UR_cIS~ 2021-05-11T~ 2020-06-15T~ TRUE
##  4 SV_51MlEKTaJKqujrg Academic Props~ UR_cIS~ 2022-05-18T~ 2021-11-30T~ TRUE
##  5 SV_6A4GI1rCxDTISF0 Querying Large~ UR_4OW~ 2022-11-29T~ 2022-11-29T~ TRUE
##  6 SV_6iepepDgfuyzn5b Emotion and Ul~ UR_3Uf~ 2018-03-27T~ 2018-03-05T~ FALSE
##  7 SV_7O3AFgjw9hgFPz8 Demographics S~ UR_1Sb~ 2023-01-06T~ 2022-07-06T~ TRUE
##  8 SV_9RGpeLaha178FAV Trust Game X R~ UR_3Xk~ 2018-03-28T~ 2018-03-13T~ TRUE
##  9 SV_b3MA3u5En0VMmRE Academic Props~ UR_cIS~ 2022-05-18T~ 2022-05-18T~ TRUE
## 10 SV_b3NY495VdBuAwdv Reaction to Mo~ UR_3Xk~ 2018-04-11T~ 2018-03-06T~ TRUE
## # ... with 76 more rows
```

```
myproject <- filter(surveys, name=="Demographics Survey - eLab - 2022 pool refresh")
myproject
```

```
## # A tibble: 1 x 6
##   id                 name            ownerId lastModified creationDate isActive
##   <chr>              <chr>           <chr>   <chr>        <chr>        <lgl>
## 1 SV_7O3AFgjw9hgFPz8 Demographics Su~ UR_1Sb~ 2023-01-06T~ 2022-07-06T~ TRUE
```

```
myproject$id[1]
```

```
## [1] "SV_7O3AFgjw9hgFPz8"
```

```
#Fetch the survey for that project as .RDS file
mysurvey <- fetch_survey(surveyID = myproject$id[1],
                         save_dir = outpath,
                         force_request = TRUE,
                         verbose = TRUE)
```

```
##   |                                                                      |
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   StartDate = col_datetime(format = ""),
##   EndDate = col_datetime(format = ""),
##   Progress = col_double(),
##   `Duration (in seconds)` = col_double(),
##   Finished = col_logical(),
##   RecordedDate = col_datetime(format = ""),
##   RecipientLastName = col_logical(),
##   RecipientFirstName = col_logical(),
##   RecipientEmail = col_logical(),
##   ExternalReference = col_logical(),
##   LocationLatitude = col_double(),
##   LocationLongitude = col_double(),
##   `Q2 Birth year` = col_double(),
##   `Q9 Zip code` = col_double(),
##   MTurkCode = col_double()
## )
## i Use `spec()` for the full column specifications.
```

```
## Warning: 38 parsing failures.
## row col           expected   actual
##  14  -- value in level set < 1 year
##  16  -- value in level set < 1 year
##  42  -- value in level set < 1 year
##  49  -- value in level set < 1 year
##  71  -- value in level set < 1 year
## ... ... .................. ........
## See problems(...) for more details.
```

```
nrow(mysurvey)
```

```
## [1] 6705
```

```
# save the approximate day/time of when you downloaded the survey
Sys.time()
```

```
## [1] "2023-01-12 14:24:49 CST"
```

```
download_time <- format(Sys.time(), "%a %b %d %X %Y")
download_message <- paste("I downloaded the", myproject$name, "survey data on:", download_time, sep=" ")
download_message
```

```
## [1] "I downloaded the Demographics Survey - eLab - 2022 pool refresh survey data on: Thu Jan 12 02:24
```

```
# save download time to log file
logfile_path <- paste(outpath, "/logfile.text", sep="")
writeLines(download_message, logfile_path)


# Read RDS file into dataframe "mysurvey", and save copy as CSV
rdspath <- paste(outpath, "/", myproject$id[1], ".rds", sep="")
mysurvey <- readRDS(file = rdspath)
csvpath <- paste(outpath, "/", myproject$id[1], ".csv", sep="")
ret <- write.csv(x=mysurvey, file=csvpath)


# Fix problem of spaces in column names!
names(mysurvey) <- str_replace_all(names(mysurvey),c(" "=".", ","=""))
colnames(mysurvey)
```

```
##  [1] "StartDate"             "EndDate"               "Status"
##  [4] "IPAddress"             "Progress"              "Duration.(in.seconds)"
##  [7] "Finished"              "RecordedDate"          "ResponseId"
## [10] "RecipientLastName"     "RecipientFirstName"    "RecipientEmail"
## [13] "ExternalReference"     "LocationLatitude"      "LocationLongitude"
## [16] "DistributionChannel"   "UserLanguage"          "Consent.text.6"
## [19] "Q1.Gender"             "Q2.Birth.year"         "Q3.Hispanic"
## [22] "Q4.Race_1"             "Q4.Race_2"             "Q4.Race_3"
## [25] "Q4.Race_4"             "Q4.Race_5"             "Q4.Race_6"
## [28] "Q5.marital"            "Q6.Eng.first"          "Q7.Eng.skill"
## [31] "Q8.Home.lang"          "Q8.Home.lang_5_TEXT"   "Q9.Zip.code"
## [34] "Q10.Citizenship"       "Q11.Years.in.US"       "Q12.School"
## [37] "Q13.Edu.level"         "Q14.Parent.edu"        "Q15.Employment"
## [40] "Q16.Household.inc"     "Q17.Size.household"    "Q18.Under.18.house"
## [43] "Q19.Pol.party"         "Q20.Pol.orientation"   "Q21.Religious"
## [46] "Q22.Religious.group"   "Q23.Pet"               "Q24.Surroundings"
```

```
## [49] "MID"                    "MTurkCode"
```

```
### FILTERING
# Eliminate test responses
mysample <- filter(mysurvey, Status=="IP Address")
nrow(mysample)
```

```
## [1] 6468
```

```
# Eliminate outliers for response time
quant_duration <- quantile(mysample$`Duration.(in.seconds)`, probs=seq(0,1,.05), na.rm=FALSE, names=FALS
quant_duration
```

```
##   [1]      2.00     96.00    112.00    122.00    131.00    141.00    150.00
##   [8]    158.00    170.00    180.00    191.00    204.00    220.00    237.00
## [15]    257.00    286.00    325.00    384.00    480.00    690.65 377912.00
```

```
mysample <- filter(mysample, `Duration.(in.seconds)` > quant_duration[2])
mysample <- filter(mysample, `Duration.(in.seconds)` < quant_duration[20])
nrow(mysample)
```

```
## [1] 5816
```

```
# Proportion of respondents who are female
females <- filter(mysample, Q1.Gender == "Female")
pct_female <- nrow(females) / nrow(mysample)
pct_female
```

```
## [1] 0.5428129
```

```
# Create a table
partyXgender <- table(mysample$Q19.Pol.party, mysample$Q1.Gender,
                      exclude = c(NA,"I do not identify with either of these.",
                                  "I do not identify with any of these"))
partyXgender
```

```
##
##                                Male Female
##   Democratic                   1194   1571
##   Republican                    742    811
##   An independent or third party  535    577
```

Embedding the results of a R code within text

R markdown also allows you to embed the results of some R code within your text. For instance:

The proportion of females respondents was 54.3 percent in our sample.

Creating a PDF or HTML document.

Finally, when you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document