HomeworkTwo

Kazmiera Breest

8/10/2020

# COVID DATA

## Describing the Data

This dataset contains survey responses from a questionaire aimed at gathering insight on how folks in the UK experience the crisis caused by COVID-19. More specifically, Italian citizens took this optional survey which explores their self reported behaviors during the pandemic. The survey collected responses to 62 different questions, which resulted in a data set that has 80 variables, or columns, and 3,460 individual observations, or rows, with each row indicating a unique survey response. Outside of the 62 questions, the remaining 18 rows are made up of identifying information such as the start and end date of the survey response, contacts information, language, and names.

setwd("C:/Users/Kassie/Documents/DACSS 601/R")  
library(tidyverse)

## -- Attaching packages ---------------------------------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.3 v dplyr 1.0.1  
## v tidyr 1.1.1 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

## -- Conflicts ------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

covidData <- read\_tsv("CovidResponses.tsv")

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## Progress = col\_double(),  
## Duration..in.seconds. = col\_double(),  
## Finished = col\_logical(),  
## RecipientLastName = col\_logical(),  
## RecipientEmail = col\_logical(),  
## ExternalReference = col\_logical(),  
## LocationLatitude = col\_double(),  
## LocationLongitude = col\_double(),  
## SelfReported\_Behavio\_1 = col\_double(),  
## SelfReported\_Behavio\_2 = col\_double(),  
## SelfReported\_Behavio\_3 = col\_double(),  
## SelfReported\_Behavio\_4 = col\_double(),  
## SelfReported\_Behavio\_5 = col\_double(),  
## SOB\_1 = col\_double(),  
## SOB\_2 = col\_double(),  
## SOB\_3 = col\_double(),  
## SOB\_4 = col\_double(),  
## Geldstrafe\_1\_1 = col\_double(),  
## Geldstrafe\_2\_1 = col\_double(),  
## age = col\_double()  
## # ... with 1 more columns  
## )

## See spec(...) for full column specifications.

head(covidData)

## # A tibble: 6 x 80  
## StartDate EndDate Status Progress Duration..in.se~ Finished RecordedDate  
## <chr> <chr> <chr> <dbl> <dbl> <lgl> <chr>   
## 1 3/17/20 ~ 3/17/2~ Surve~ 100 281 TRUE 3/17/20 17:~  
## 2 3/17/20 ~ 3/17/2~ Surve~ 100 442 TRUE 3/17/20 20:~  
## 3 3/18/20 ~ 3/18/2~ IP Ad~ 100 4 TRUE 3/18/20 7:33  
## 4 3/18/20 ~ 3/18/2~ IP Ad~ 100 1013 TRUE 3/18/20 7:58  
## 5 3/18/20 ~ 3/18/2~ IP Ad~ 100 79 TRUE 3/18/20 8:00  
## 6 3/18/20 ~ 3/18/2~ IP Ad~ 100 3327 TRUE 3/18/20 8:00  
## # ... with 73 more variables: ResponseId <chr>, RecipientLastName <lgl>,  
## # RecipientEmail <lgl>, ExternalReference <lgl>, LocationLatitude <dbl>,  
## # LocationLongitude <dbl>, DistributionChannel <chr>, UserLanguage <chr>,  
## # Q1 <chr>, SelfReported\_Behavio\_1 <dbl>, SelfReported\_Behavio\_2 <dbl>,  
## # SelfReported\_Behavio\_3 <dbl>, SelfReported\_Behavio\_4 <dbl>,  
## # SelfReported\_Behavio\_5 <dbl>, Reflection <chr>, T2 <chr>, social <chr>,  
## # handshake <chr>, stores <chr>, curfew <chr>, SOB\_1 <dbl>, SOB\_2 <dbl>,  
## # SOB\_3 <dbl>, SOB\_4 <dbl>, financialpunishment <chr>, Geldstrafe\_1\_1 <dbl>,  
## # Geldstrafe\_2\_1 <dbl>, perceivedreaction <chr>, Q36 <chr>, Q37 <chr>,  
## # Q23 <chr>, perceivedeffectivnes <chr>, anxiety\_1 <chr>, anxiety\_2 <chr>,  
## # anxiety\_3 <chr>, anxiety\_4 <chr>, anxiety\_5 <chr>, Q24 <chr>, Q25 <chr>,  
## # Q25\_13\_TEXT <chr>, Q26 <chr>, Q26\_11\_TEXT <chr>, age <dbl>, gender <chr>,  
## # gender\_3\_TEXT <chr>, zipcode <chr>, health <chr>, health\_conditions <chr>,  
## # personality\_1 <chr>, personality\_2 <chr>, personality\_3 <chr>,  
## # personality\_4 <chr>, personality\_5 <chr>, personality\_6 <chr>,  
## # personality\_7 <chr>, personality\_8 <chr>, personality\_9 <chr>,  
## # personality\_10 <chr>, Q27\_5 <chr>, Q35 <chr>, Q62\_5 <chr>, Q62\_6 <chr>,  
## # Q62\_7 <chr>, Q62\_8 <chr>, Q62\_9 <chr>, Q62\_11 <chr>, Q62\_12 <chr>,  
## # Q62\_13 <chr>, Q62\_14 <chr>, Q62\_15 <chr>, rid <chr>, treatment <chr>,  
## # finish <dbl>

colnames(covidData)

## [1] "StartDate" "EndDate" "Status"   
## [4] "Progress" "Duration..in.seconds." "Finished"   
## [7] "RecordedDate" "ResponseId" "RecipientLastName"   
## [10] "RecipientEmail" "ExternalReference" "LocationLatitude"   
## [13] "LocationLongitude" "DistributionChannel" "UserLanguage"   
## [16] "Q1" "SelfReported\_Behavio\_1" "SelfReported\_Behavio\_2"  
## [19] "SelfReported\_Behavio\_3" "SelfReported\_Behavio\_4" "SelfReported\_Behavio\_5"  
## [22] "Reflection" "T2" "social"   
## [25] "handshake" "stores" "curfew"   
## [28] "SOB\_1" "SOB\_2" "SOB\_3"   
## [31] "SOB\_4" "financialpunishment" "Geldstrafe\_1\_1"   
## [34] "Geldstrafe\_2\_1" "perceivedreaction" "Q36"   
## [37] "Q37" "Q23" "perceivedeffectivnes"   
## [40] "anxiety\_1" "anxiety\_2" "anxiety\_3"   
## [43] "anxiety\_4" "anxiety\_5" "Q24"   
## [46] "Q25" "Q25\_13\_TEXT" "Q26"   
## [49] "Q26\_11\_TEXT" "age" "gender"   
## [52] "gender\_3\_TEXT" "zipcode" "health"   
## [55] "health\_conditions" "personality\_1" "personality\_2"   
## [58] "personality\_3" "personality\_4" "personality\_5"   
## [61] "personality\_6" "personality\_7" "personality\_8"   
## [64] "personality\_9" "personality\_10" "Q27\_5"   
## [67] "Q35" "Q62\_5" "Q62\_6"   
## [70] "Q62\_7" "Q62\_8" "Q62\_9"   
## [73] "Q62\_11" "Q62\_12" "Q62\_13"   
## [76] "Q62\_14" "Q62\_15" "rid"   
## [79] "treatment" "finish"

dim(covidData)

## [1] 3460 80