WorkingWithDataAssignment

João Mota

Table of contents

1 Limits

There are multiple limitations to my analysis to be noted. Firstly, the data collected is limited to cyber attacks that were detected, there is variety of attacks that have gone unnoticed and therefore the data has a systematic tendency to underestimate the real level of breach attacks.(Department For Digital 2020) Secondly the missing data generated by is biased towards smaller and less digital institution as they don't have the infrastructure to detect and make the proper assessment of the attack.

/*
A few hours of trial and errors can save you a few minutes of reading the proper documentation :) https://quarto.org/docs/output-formats/pdf-basics.html
Go to terminal tab down there and type quarto install tool tinytex
NOTE TO SELF!!!! using quarto is the same as playing restart Rstudio simulator 2022 because nothing is properly recached and they have a worse garbage collecter than assembly so if you still get the same error after changing the just restart rstudio and remember to never ever ever change the initial format or add anything close to it because it will break the pdf and start generating html also please be smart and read https://quarto.org/docs/reference/formats/pdf.html for the formating
*/

```
library(haven)
library(tidyverse)
```

```
-- Attaching packages --------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.3.6      v purrr   0.3.4
```

```
v tibble  3.1.8      v dplyr   1.0.10
v tidyr   1.2.1      v stringr 1.4.1
v readr   2.1.3      v forcats 0.5.2
-- Conflicts --------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)
library(geometry)
#install.packages("Rmpfr")
library(formatR)
#install.packages("devtools")
devtools::install_github("cardiomoon/autoReg")
```

```
Skipping install of 'autoReg' from a github remote, the SHA1 (6bc789c9) has not changed since las
  Use `force = TRUE` to force installation
```

```
## Very important documentation for the 2018 data set //it is a
## surprise toll that will help us later
technicalAnnex2018 = "https://doc.ukdataservice.ac.uk/doc/8406/mrdoc/pdf/8406_cyber_security_b

## this is the loading the first year of this level of survey data set
## after burning my entire brain, replacing it with the backup one and
## also burning that one I discovered that it is just these lines that
## aren't being formatted in pdf because they are absolutely huge but
## at least it works for the other ones #FicaADica I assume it was
## thanks to formatR ?? I won't bother to redo every single bloody step
## again, enough pain and stack for the day
dataCyberSecuritySurvey2018 = read_spss("C:/AppliedDataScienceAndStatistics/Applied-Data-Scien

## adding the variable year because none of the data sets have any
## proper way to distinguish between the years of each survey
dataCyberSecuritySurvey2018$year = "2018"
```

2 Now we do the same for the other years before we merge them

```
## loading the second year of this level of survey data set
dataCyberSecuritySurvey2019 = read_spss("C:/AppliedDataScienceAndStatistics/Applied-Data-Scien

## adding the variable year because none of the data sets have any
## proper way to distinguish between the years of each survey
dataCyberSecuritySurvey2019$year = "2019"
```

```r
## loading the third year of this level of survey data set
dataCyberSecuritySurvey2020 = read_spss("C:/AppliedDataScienceAndStatistics/Applied-Data-Scien

## adding the variable year because none of the data sets have any
## proper way to distinguish between the years of each survey
dataCyberSecuritySurvey2020$year = "2020"
```

```r
## loading the forth year of this level of survey data set
dataCyberSecuritySurvey2021 = read_spss("C:/AppliedDataScienceAndStatistics/Applied-Data-Scien

## adding the variable year because none of the data sets have any
## proper way to distinguish between the years of each survey
dataCyberSecuritySurvey2021$year = "2021"
```

```r
## loading the fifth and final year of this level of survey data set
dataCyberSecuritySurvey2022 = read_spss("C:/AppliedDataScienceAndStatistics/Applied-Data-Scien

## adding the variable year because none of the data sets have any
## proper way to distinguish between the years of each survey
dataCyberSecuritySurvey2022$year = "2022"
```

```r
## Now that we have all data loaded lets start by tidying up data set
## by data set start from 2018


## for some sweet sweet documentation about the questions starting from
## page 26 TODO comment in case of fire or debugging
browseURL(technicalAnnex2018)
```

```r
## This entire code snippet is tidying up the type of organisation for
## the 2018 survey renaming the bloody variables to a more java like
## name

dataCyberSecuritySurvey2018TidyName = rename(dataCyberSecuritySurvey2018,
    instituitionTypes = "samptype")

## if instituitionTypes is 1 it is a business if it is 2 it is a
## charity and in the future 3 is for schools and education


## daily reminder that there is a boolean type but it is called logical
## Numeric -\tSet of all real numbers Integer -\tSet of all integers, Z
## Logical - -\tTRUE and FALSE Complex -\tSet of complex numbers
## Character -\t"a", "b", "c", …, "ç", "#", "~", …., "1", "2", …etc
```

```r
## it is a string so lets make it a proper numeric code

dataCyberSecuritySurvey2018TidyName$instituitionTypes = as.integer(dataCyberSecuritySurvey2018


## typex is 1-2 for businesses and 3 for charities so redundant and can
## be removed

dataCyberSecuritySurvey2018TidyName = dataCyberSecuritySurvey2018TidyName %>%
    select(-typex)


## dataCyberSecuritySurvey2018TidyName never forget if R can't show all
## displayed text from a computation it breaks both the rendering and
## ##the refreshing of the rendered code for some reason ¯\_( )_/¯
## future edit anything and everything breaks for no reason at all,
## just kill it and reopen refer to the first NOTE TO SELF for more
## information
```

```r
## see questioner documentation start from page 27

technicalAnnex2019 = "https://assets.publishing.service.gov.uk/government/uploads/system/uploa
## TODO comment in case of fire or debugging
browseURL(technicalAnnex2019)
```

```r
## see questioner documentation start from page 31

technicalAnnex2020 = "https://assets.publishing.service.gov.uk/government/uploads/system/uploa

## TODO comment in case of fire or debugging
browseURL(technicalAnnex2020)
```

```r
## see questioner documentation start from page 28

technicalAnnex2021 = "https://assets.publishing.service.gov.uk/government/uploads/system/uploa

## TODO comment in case of fire or debugging
browseURL(technicalAnnex2021)
```

```r
## see questioner documentation start from page 36

technicalAnnex2022 = "https://assets.publishing.service.gov.uk/government/uploads/system/uploa

## TODO comment in case of fire or debugging
browseURL(technicalAnnex2022)
```

```r
## trying not to get arrested for DDoSing the uk goverment by making a
## request to all the pdfs after rendering the page for the nth because
## I can't code nor debug (challenge impossible) bonus points if I get
## an exeter ip banned because of it
```

```r
## time to recycle the code for the 2018 survey that gets a 'neat' code
## of the institution types

## This entire code snippet is tidying up the type of organisation for
## the 2019 survey renaming the bloody variables to a more java like
## name

dataCyberSecuritySurvey2019TidyName = rename(dataCyberSecuritySurvey2019,
    instituitionTypes = "samptype")

dataCyberSecuritySurvey2019TidyName$instituitionTypes = as.integer(dataCyberSecuritySurvey2019

str(dataCyberSecuritySurvey2019TidyName$instituitionTypes)
```

```
int [1:2080] 1 1 1 1 1 1 1 1 1 ...
```

```r
## typex is redundant be we already have an indentifies for each type
## of institution and can be removed same for questtype since this
## questioner has more redundancy than amazon and google data centers
## combined

dataCyberSecuritySurvey2019TidyName = dataCyberSecuritySurvey2019TidyName %>%
    select(-one_of("typex", "questtype"))
```

```r
## I continue to save the planet by recycling as much as I can, mostly
## recycled code from the previous snippet today though this time we do
## have the concept of education institutions as our code just annoy me
## after I thought they should be converted to boolean like a getter in
## java

dataCyberSecuritySurvey2020TidyName = rename(dataCyberSecuritySurvey2020,
    instituitionTypes = "samptype")

dataCyberSecuritySurvey2020TidyName$instituitionTypes = as.integer(dataCyberSecuritySurvey2020

str(dataCyberSecuritySurvey2020TidyName$instituitionTypes)
```

```
int [1:1900] 1 1 1 1 1 1 1 1 1 ...
```

```r
## typex is redundant be we already have an indentifies for each type
## of institution and can be removed same for questtype since this
## questioner has more redundancy than amazon and google data centers
## combined

dataCyberSecuritySurvey2020TidyName = dataCyberSecuritySurvey2020TidyName %>%
    select(-one_of("typex", "questtype"))
```

```r
## saving the planet one recycled snippet of code at a time


dataCyberSecuritySurvey2021TidyName = rename(dataCyberSecuritySurvey2021,
    instituitionTypes = "samptype")

dataCyberSecuritySurvey2021TidyName$instituitionTypes = as.integer(dataCyberSecuritySurvey2021

str(dataCyberSecuritySurvey2021TidyName$instituitionTypes)
```

```
int [1:1900] 1 1 1 1 1 1 1 1 1 1 ...
```

```r
## typex is redundant be we already have an indentifies for each type
## of institution and can be removed same for questtype since this
## questioner has more redundancy than amazon and google data centers
## combined

dataCyberSecuritySurvey2021TidyName = dataCyberSecuritySurvey2021TidyName %>%
    select(-one_of("typex", "questtype"))
```

```r
## this comment was already dealt by the garbage collector unlike the
## previous ones

dataCyberSecuritySurvey2022TidyName = rename(dataCyberSecuritySurvey2022,
    instituitionTypes = "samptype")

dataCyberSecuritySurvey2022TidyName$instituitionTypes = as.integer(dataCyberSecuritySurvey2022

str(dataCyberSecuritySurvey2022TidyName$instituitionTypes)
```

```
int [1:2157] 1 1 1 1 1 1 1 1 1 1 ...
```

```r
## questtype is redundant be we already have an indentifies for each
## type of institution and can be removed

dataCyberSecuritySurvey2022TidyName = dataCyberSecuritySurvey2022TidyName %>%
    select(-questtype)
```

```r
## now that we have started the data wrangling we will categorize all
## institutions by size remember that for some wicked reason they use
## -97 for missing values for anything without a proper missing value
## code for each question I will start by simply nulling every single
## -97 so we can see how much is missing and then possibly make a table
## with custom missing values for each like I did in C (remember to
## start from -1000 to -1999 like standard ACLs)


############################################## 2018
############################################## ##############################################
numberOfCycles = length(dataCyberSecuritySurvey2018TidyName$sizea)

dataCyberSecuritySurvey2018TidyNameSize = dataCyberSecuritySurvey2018TidyName

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2018TidyNameSize$sizea[i] == -97) {
        dataCyberSecuritySurvey2018TidyNameSize$sizea[i] = NA
    }
    if (dataCyberSecuritySurvey2018TidyNameSize$sizeb[i] == -97) {
        dataCyberSecuritySurvey2018TidyNameSize$sizeb[i] = NA
    }

}

############################################## 2019
############################################## ##############################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyName$sizea)

dataCyberSecuritySurvey2019TidyNameSize = dataCyberSecuritySurvey2019TidyName

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2019TidyNameSize$sizea[i] == -97) {
        dataCyberSecuritySurvey2019TidyNameSize$sizea[i] = NA
    }
    if (dataCyberSecuritySurvey2019TidyNameSize$sizeb[i] == -97) {
        dataCyberSecuritySurvey2019TidyNameSize$sizeb[i] = NA
    }

}

############################################## 2020
############################################## ##############################################
```

```
numberOfCycles = length(dataCyberSecuritySurvey2020TidyName$sizea)

dataCyberSecuritySurvey2020TidyNameSize = dataCyberSecuritySurvey2020TidyName

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2020TidyNameSize$sizea[i] == -97) {
        dataCyberSecuritySurvey2020TidyNameSize$sizea[i] = NA
    }
    if (dataCyberSecuritySurvey2020TidyNameSize$sizeb[i] == -97) {
        dataCyberSecuritySurvey2020TidyNameSize$sizeb[i] = NA
    }
}

############################################## 2021
############################################## ################################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyName$sizea)

dataCyberSecuritySurvey2021TidyNameSize = dataCyberSecuritySurvey2021TidyName

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2021TidyNameSize$sizea[i] == -97) {
        dataCyberSecuritySurvey2021TidyNameSize$sizea[i] = NA
    }
    if (dataCyberSecuritySurvey2021TidyNameSize$sizeb[i] == -97) {
        dataCyberSecuritySurvey2021TidyNameSize$sizeb[i] = NA
    }
}

############################################## 2022
############################################## ################################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyName$sizea)

dataCyberSecuritySurvey2022TidyNameSize = dataCyberSecuritySurvey2022TidyName

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2022TidyNameSize$sizea[i] == -97) {
        dataCyberSecuritySurvey2022TidyNameSize$sizea[i] = NA
    }
    if (dataCyberSecuritySurvey2022TidyNameSize$sizeb[i] == -97) {
```

```
        dataCyberSecuritySurvey2022TidyNameSize$sizeb[i] = NA
    }
}
```

```
## we don't need neither the combined regions (since those are for
## business analyzes and we don't do those) same for sector_comb1 and
## 2.

## region_comb? throw it in the trash. sector_comb1? throw it in the
## trash. sector_comb2? throw it in the trash.

################################################ 2018
################################################ ################################################

dataCyberSecuritySurvey2018TidyNameSize = dataCyberSecuritySurvey2018TidyNameSize %>%
    select(-one_of("region_comb", "sector_comb1", "sector_comb2"))

################################################ 2019
################################################ ################################################

dataCyberSecuritySurvey2019TidyNameSize = dataCyberSecuritySurvey2019TidyNameSize %>%
    select(-one_of("region_comb", "sector_comb2"))

################################################ 2020
################################################ ################################################

dataCyberSecuritySurvey2020TidyNameSize = dataCyberSecuritySurvey2020TidyNameSize %>%
    select(-one_of("region_comb", "sector_comb2"))

################################################ 2021
################################################ ################################################

dataCyberSecuritySurvey2021TidyNameSize = dataCyberSecuritySurvey2021TidyNameSize %>%
    select(-one_of("region_comb", "sector_comb2"))

################################################ 2022
################################################ ################################################

dataCyberSecuritySurvey2022TidyNameSize = dataCyberSecuritySurvey2022TidyNameSize %>%
    select(-one_of("region_comb", "sector_comb2"))
```

```
## removing social media questions that are irrelevant because they are
## absolutely terrible metrics to understand the digitalization of an
## institution Note to self: if I have time get all of these type of
## functions in try catch because them breaking up with the select
## error is no good and it makes me cry every time I have to manually
```

```r
## run a part of the snippet and see which is one the bad one
## https://r-lang.com/r-trycatch-function/ ## #FicaADica

################################################ 2018
################################################ ################################################

dataCyberSecuritySurvey2018TidyNameSize = dataCyberSecuritySurvey2018TidyNameSize %>%
    select(-(online1:online11))

################################################ 2019
################################################ ################################################

dataCyberSecuritySurvey2019TidyNameSize = dataCyberSecuritySurvey2019TidyNameSize %>%
    select(-(online1:online11))

################################################ 2020
################################################ ################################################

dataCyberSecuritySurvey2020TidyNameSize = dataCyberSecuritySurvey2020TidyNameSize %>%
    select(-(online1:online11))

################################################ 2021
################################################ ################################################

dataCyberSecuritySurvey2021TidyNameSize = dataCyberSecuritySurvey2021TidyNameSize %>%
    select(-(online1:online11))

################################################ 2022
################################################ ################################################

dataCyberSecuritySurvey2022TidyNameSize = dataCyberSecuritySurvey2022TidyNameSize %>%
    select(-(online1:online14))
```

```r
## removing the question about the mobile usage because it also is a
## terrible indicator of a company digitalization

################################################ 2018
################################################ ################################################

dataCyberSecuritySurvey2018TidyNameSize = dataCyberSecuritySurvey2018TidyNameSize %>%
    select(-mobile)

################################################ 2019
################################################ ################################################

dataCyberSecuritySurvey2019TidyNameSize = dataCyberSecuritySurvey2019TidyNameSize %>%
    select(-mobile)
```

```
############################################### 2020
############################################### ################################################

dataCyberSecuritySurvey2020TidyNameSize = dataCyberSecuritySurvey2020TidyNameSize %>%
    select(-mobile)

############################################### 2021
############################################### ################################################

dataCyberSecuritySurvey2021TidyNameSize = dataCyberSecuritySurvey2021TidyNameSize %>%
    select(-mobile)

############################################### 2022
############################################### ################################################

dataCyberSecuritySurvey2022TidyNameSize = dataCyberSecuritySurvey2022TidyNameSize %>%
    select(-mobile)
```

```
## question about the attitude and outsourcing of cyber security have
## been removed the the surveys starting from 2020 so it doesn't make
## sense to keep them in the 2018 and 2019 data set

## I will start doing some proper garbage collection and this time I am
## not just taking myself out I will only ever have the original data
## and the most recent modified one


############################################### 2018
############################################### ################################################

dataCyberSecuritySurvey2018TidyNameSize = dataCyberSecuritySurvey2018TidyNameSize %>%
    select(-(outsource:attitude4))

############################################### 2019
############################################### ################################################

dataCyberSecuritySurvey2019TidyNameSize = dataCyberSecuritySurvey2019TidyNameSize %>%
    select(-(outsource:attitude4))
```

```
## since we want to have access to some proper data we will tidy the
## questions about how big of a priority is cyber security and how
## often are the higher ups updated about it this could really use some
## try catches because for the some weird reason -97 == NA does not
## return true or false, this is like javascript levels of bad

## also excepting this very first one the removals will be at the end
```

```r
## so they are computed as if they were a transaction because try and
## catch is a lie to sell more lines of codes

## thanks to a blessing for our lord not finding the object only gives
## a warning and we ignore those as long as it still lets run the rest
## of the code

rm(dataCyberSecuritySurvey2018TidyName)
rm(dataCyberSecuritySurvey2019TidyName)
rm(dataCyberSecuritySurvey2020TidyName)
rm(dataCyberSecuritySurvey2021TidyName)
rm(dataCyberSecuritySurvey2022TidyName)

############################################## 2018
############################################## ##############################################
numberOfCycles = length(dataCyberSecuritySurvey2018TidyNameSize$priority)

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSize

## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$priority[i] == -97) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$priority[i] = NA
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$update[i] == -97) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$update[i] = NA
    }

}

############################################## 2019
############################################## ##############################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyNameSize$priority)

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSize


## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$priority[i] == -97) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$priority[i] = NA
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$update[i] == -97) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$update[i] = NA
```

```
    }

}

############################################## 2020
############################################## ##############################################

numberOfCycles = length(dataCyberSecuritySurvey2020TidyNameSize$priority)

dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSize


## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$priority[i] == -97) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$priority[i] = NA
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$update[i] == -97) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$update[i] = NA
    }
}

############################################## 2021
############################################## ##############################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyNameSize$priority)

dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSize


## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$priority[i] == -97) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$priority[i] = NA
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$update[i] == -97) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$update[i] = NA
    }
}

############################################## 2022
############################################## ##############################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyNameSize$priority)
```

```r
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSize


## apparently we have to be careful because an already inserted NA on
## the variable breaks the
for (i in 1:numberOfCycles) {
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$priority[i] == -97) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$priority[i] = NA
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$update[i] == -97) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$update[i] = NA
    }
}

## garbage man? Well, of course I know him. He is me.

rm(dataCyberSecuritySurvey2018TidyNameSize)
rm(dataCyberSecuritySurvey2019TidyNameSize)
rm(dataCyberSecuritySurvey2020TidyNameSize)
rm(dataCyberSecuritySurvey2021TidyNameSize)
rm(dataCyberSecuritySurvey2022TidyNameSize)
```

```r
## questions about reason of investment in cybersecuirty were removed
## form the pre-pilot survey in 2020

############################################## 2018
############################################## ################################################

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(reason1:reason27))

############################################## 2019
############################################## ################################################

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(reason1:reason28))

## the rest were already deleted
```

```r
## removing the cyber security insurance claims because they don't give
## us relevant data to what we are analyzing in the data set pro tip:
## having insurance does not make you more or less likely to be
## targeted nor does it change the costs of the attack

############################################## 2018
############################################## ################################################
```

```r
dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(insurex:noinsure19))

############################################## 2019
############################################## ##############################################

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(insurex:noinsure19))

############################################## 2020
############################################## ##############################################

dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(insurex:claim))

############################################## 2021
############################################## ##############################################

dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(insurex:claim))

############################################## 2022
############################################## ##############################################

dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(insurex:claim))
```

```r
## we are removing the questions about asking for info, advice,
## guidance about cyber security or government schemes

############################################## 2018
############################################## ##############################################

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(info1:trainwho7))

############################################## 2019
############################################## ##############################################

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(info1:trainwho7))

############################################## 2020
############################################## ##############################################

dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(info1:scheme5))
```

```
############################################## 2021
############################################## ##############################################

dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(info1:scheme5))

############################################## 2022
############################################## ##############################################

## here we have more that we need to remove because they introduced
## questions similar to the ones that we are removing but with the
## government as a source for reason to make implementations which we
## are also not interested in

dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(info1:govtact44))
```

3 Now is when the policies and procedures are evaluated

```
## manage 1 - Board members/trustees with responsibility for cyber
## security manage 2 - outsourcing cyber security manage 3 - formal
## policy or policies in place covering cyber security risks manage 4 -
## Business Continuity Plan manage 5 - Staff members whose job role
## includes information security or governance //it stopped being used
## after the 2020 survey manage 6 - don't know/missing data manage 7 -
## absolutely nothings, good luck have fun (rip bozzo) manage 8 -
## written list of what is critical to protect (only exists in the
## survey of 2022)

## altura de me desemerdar que esta aqui esta mesmo grossa não faz
## frio, nem orvalho, está a chover para caralho converting the final
## value to a collection so I can append all the values //facepalm

##################################################### 2018
##################################################### ######################################

numberOfCycles = length(dataCyberSecuritySurvey2018TidyNameSizeCyber$manage1)

dataCyberSecuritySurvey2018TidyNameSizeCyber$management = c(0)

## brb = c() brb = append(brb,7)

## TODO TODO TODO TODO should i make it binary since i cannot make it
## into a collection inside a dataframe
```

```r
for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage1[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = 1
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage2[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage3[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            100
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage4[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1000
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage5[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10000
    }
    ## not needed since by default they are all = NA, just here to
    ## remind me
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage6[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = NA
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$manage7[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1e+06
    }


}


######################################################## 2019
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyNameSizeCyber$manage1)

dataCyberSecuritySurvey2019TidyNameSizeCyber$management = 0

for (i in 1:numberOfCycles) {
```

```
    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage1[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage2[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage3[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            100
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage4[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1000
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage5[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10000
    }

    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage6[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = NA
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$manage7[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1e+06
    }


}


######################################################## 2020
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2020TidyNameSizeCyber$manage1)

dataCyberSecuritySurvey2020TidyNameSizeCyber$management = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs
```

```
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage1[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage2[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage3[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            100
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage4[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1000
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage5[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10000
    }

    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage6[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = NA
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$manage7[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1e+06
    }


}

######################################################## 2021
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyNameSizeCyber$manage1)

dataCyberSecuritySurvey2021TidyNameSizeCyber$management = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage1[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1
```

```
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage2[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage3[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            100
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage4[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage5[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            10000
    }

    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage6[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = NA
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$manage7[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1e+06
    }


}

######################################################## 2022
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyNameSizeCyber$manage1)

dataCyberSecuritySurvey2022TidyNameSizeCyber$management = 0

## This part of the data set has an huge amount of missing values and
## now I will be force to do try and catch because the nevermind
## sharrooq saved me by mentioning the is.na so I will be cheesing with
## it instead, truly a blessing, studyroom at 21 just hits harders


## can't make this a switch case because they are all different
## variables ffs

for (i in 1:numberOfCycles) {
```

```
`?`(replace_na)
## using this to make the values -10001 and I will remove them
## after the computation dear god I am being biased again because I
## wasn't recording the missing values in the answers this is such
## a mess, how will I be able to apply proper missing data modules
## with this

dataCyberSecuritySurvey2022TidyNameSizeCyber$manage1[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage2[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage3[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage4[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage5[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage6[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage7[i] = replace_na(dataCyberSecuritySurv
    -10001)
dataCyberSecuritySurvey2022TidyNameSizeCyber$manage8[i] = replace_na(dataCyberSecuritySurv
    -10001)

if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage1[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
        1
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage2[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
        10
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage3[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
        100
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage4[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
        1000
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage5[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
        10000
}

if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage6[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = NA
```

```
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$manage7[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$management[i] = dataCyberSecuritySurvey20
            1e+06
    }



}

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(manage1:manage7))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(manage1:manage7))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(manage1:manage7))
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(manage1:manage7))
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(manage1:manage8))
```

```
## now we are removing the reasons why they don't have the appropriate
## measures because we are more interested in the questions about
## security after these ones also it was deleted after 2019

## that is a catch 22
dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(nopol1:nopol22))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(nopol1:nopol22))
```

```
## sadly all the questions about measures done in the last 12 moths
## have changes quite a bit during the years which makes it impossible
## to have a good year to year analysis when we aren't comparing the
## same thing

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(ident1:ident8))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(ident1:ident8))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(ident1:ident11))
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(ident1:ident11))
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(ident1:ident7))
```

```r
## this only exists in 2022 so it makes no sense to look at
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(comply1:audit))
```

```r
## This is where the fun begins with some proper policies simping for
## incremental backups


## rule 1 - applying software updates rule 2 - up to date maleware
## protection rule 3 - well configured firewalls rule 4 - proper
## permission configuration rule 5 - monitoring user activity rule 6 -
## encrypting personal data // only used in 2018 rule 7 - security
## controls on company devices rule 8 - only allowing access from
## company devices rule 9 - segregated guest wireless / so basically a
## DMZ rule 10 - don't know rule 11 - none (YOLO) rule 12 - strong
## passwords //only used in 2018 rule 13 - backup data to the cloud
## (diskette robots in data center go brrrrrrr) rule 14 - backup the
## data to another place that isn't the cloud rule 15 - storing and
## moving data/files securely //wasn't used in 2018 rule 16 - 2 factor
## authentication // only used in 2019 rule 17 - policy for strong
## passwords rule 18 - VPN (virtual private network) //only in 2022
## rule 19 - phishing procedure // only in 2022 rule 20 -
## authentication when accessing the network // only in 2022


## TODO TODO TODO TODO unfuck this mess as well because I just don't
## know anymore good news is that I have a solution, bad news is that
## it is not a perfect solution at least I won't be able to cause a
## stack overflow because with a precision of 53 bits, and represents
## to that precision a range of absolute values from about 2e-308 to
## 2e+308


######################################################## 2018
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2018TidyNameSizeCyber$rules1)

dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
```

```
        1
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules1[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        10
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules3[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        100
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules4[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1000
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules5[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        10000
}
## yeah we really didn't need this one since it was only present in
## 2018 my bad
## if(dataCyberSecuritySurvey2018TidyNameSizeCyber$rules6[i] == 1){
## dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] =
## NA }

if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules7[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+06
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules8[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+07
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules9[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+08
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules10[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = NA
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules11[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+10
}
if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules12[i] == 1) {
    dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+11
}
```

```
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules13[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1e+12
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$rules14[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1e+13
    }


}


########################################################## 2019
########################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules1)

dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules = 0

## it was at this moment I realised, I fucked up, like really hard
## because I have to change a crap ton of things to collection because
## I was only receiving the last value from them like an idiot brb =
## c(1,2,3) brb = append(brb,5)

for (i in 1:numberOfCycles) {

    ## handing missing data that will be
    ## dataCyberSecuritySurvey2019TidyNameSizeCyber$rules6[i] =
    ## #replace_na(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules6[i],
    ## -10004)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$rules12[i] = replace_na(dataCyberSecuritySurv
        -10004)

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules3[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
```

```
        100
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules4[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1000
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules5[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        10000
}
## yeah we really didn't need this one since it was only present in
## 2018 my bad
## if(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules6[i] == 1){
## dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] =
## NA }

if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules7[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+06
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules8[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+07
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules9[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+08
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules10[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = NA
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules11[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+10
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules12[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+11
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules13[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+12
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules14[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+13
}
```

```
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$rules15[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1e+14
    }
    ## we don't care about something that was only analyzed during one
    ## year if(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules16[i]
    ## == 1){
    ## dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] =
    ## 16 }



}


########################################################## 2020
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2020TidyNameSizeCyber$rules1)

dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules = 0



for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs
    dataCyberSecuritySurvey2020TidyNameSizeCyber$rules12[i] = replace_na(dataCyberSecuritySurv
        -10004)

    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules3[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            100
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules4[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1000
    }
```

```
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules5[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        10000
}
## not needed since the code was dropped after 2018, just here to
## remind me i can't believe I can't read oh great heavens
## if(dataCyberSecuritySurvey2018TidyNameSizeCyber$manage6[i] ==
## 1){ dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] =
## NA }

if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules7[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+06
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules8[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+07
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules9[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+08
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules10[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = NA
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules11[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+10
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules12[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+11
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules13[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+12
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules14[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+12
}
if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules15[i] == 1) {
    dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+13
}
## we don't care about something that was only analyzed during one
## year if(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules16[i]
```

```
    ## == 1){
    ## dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] =
    ## 16 } I guess we can have a look at the 17th one since we still
    ## look at 2020:2022
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$rules17[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1e+15
    }

}

######################################################## 2021
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyNameSizeCyber$rules1)

dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs
    dataCyberSecuritySurvey2021TidyNameSizeCyber$rules12[i] = replace_na(dataCyberSecuritySurv
        -10004)

    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules3[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            100
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules4[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules5[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10000
    }
    ## not needed since the code was dropped after 2018, just here to
    ## remind me
```

```
## if(dataCyberSecuritySurvey2018TidyNameSizeCyber$manage6[i] ==
## 1){ dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] =
## NA }

if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules7[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+06
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules8[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+07
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules9[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+08
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules10[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = NA
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules11[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+10
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules12[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+11
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules13[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+12
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules14[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+13
}
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules15[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+14
}
## we don't care about something that was only analyzed during one
## year if(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules16[i]
## == 1){
## dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] =
## 16 } I guess we can have a look at the 17th one since we still
## look at 2020:2022
if (dataCyberSecuritySurvey2021TidyNameSizeCyber$rules17[i] == 1) {
    dataCyberSecuritySurvey2021TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
```

```
            1e+16
    }

}


######################################################## 2022
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyNameSizeCyber$rules1)

dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules = NA

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs
    dataCyberSecuritySurvey2022TidyNameSizeCyber$rules8[i] = replace_na(dataCyberSecuritySurve
        -10004)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$rules9[i] = replace_na(dataCyberSecuritySurve
        -10004)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$rules12[i] = replace_na(dataCyberSecuritySurv
        -10004)

    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules1[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules3[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            100
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules4[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            1000
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules5[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
            10000
    }
    ## not needed since the code was dropped after 2018, just here to
    ## remind me
    ## if(dataCyberSecuritySurvey2018TidyNameSizeCyber$manage6[i] ==
```

```
## 1){ dataCyberSecuritySurvey2018TidyNameSizeCyber$management[i] =
## NA }

if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules7[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+06
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules8[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+07
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules9[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+08
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules10[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+09
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules11[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+10
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules12[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+11
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules13[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+12
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules14[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+13
}
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules15[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
        1e+14
}
## we don't care about something that was only analyzed during one
## year if(dataCyberSecuritySurvey2019TidyNameSizeCyber$rules16[i]
## == 1){
## dataCyberSecuritySurvey2019TidyNameSizeCyber$internalRules[i] =
## 16 } I guess we can have a look at the 17th one since we still
## look at 2020:2022
if (dataCyberSecuritySurvey2022TidyNameSizeCyber$rules17[i] == 1) {
    dataCyberSecuritySurvey2022TidyNameSizeCyber$internalRules[i] = dataCyberSecuritySurve
```

```
            1e+16
    }
    ## rules 18,19 and 20 are only for this year so there is no point
    ## of having a look at them here

}


## now we can remove all those rules columns that we are no longer
## using

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(rules1:rules14))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(rules1:rules16))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(rules1:rules17))
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(rules1:rules17))
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(rules1:rules20))
```

```
## we do a little policing but sadly not the one QoS type on cisco
## servers


## policy 1 - what can be stored in the removable devices policy 2 -
## remote working policy 3 - what staff are permitted to do on your
## organisations IT devices policy 4 - use of personally-owned devices
## for business activities policy 5 - Use of new digital technologies
## such as cloud computing (seriously what the hell is this question
## smh) policy 6 - data classification policy 7 - a Document Management
## System policy 8 - don't know (estudasses) policy 9 - none of these
## (YOLO) policies 10,11 and 12 were only made in 2022 so not work
## comparing


## TODO TODO TODO TODO unfuck this mess as well because I just don't
## know anymore good news is that I have a solution, bad news is that
## it is not a perfect solution at least I won't be able to cause a
## stack overflow because with a precision of 53 bits, and represents
## to that precision a range of absolute values from about 2e-308 to
## 2e+308


########################################################## 2018
########################################################## ########################################
```

```r
numberOfCycles = length(dataCyberSecuritySurvey2018TidyNameSizeCyber$policy1)

dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies = 0

for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy1[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy2[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy3[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy4[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy5[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy6[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy7[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy8[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2018TidyNameSizeCyber$policy9[i] = replace_na(dataCyberSecuritySurv
        -10005)

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy1[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            1
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy2[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            10
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy3[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            100
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy4[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            1000
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy5[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            10000
```

```
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy6[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            1e+05
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy7[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            1e+06
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy8[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = NA
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$policy9[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2018
            1e+08
    }

}


########################################################## 2019
########################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyNameSizeCyber$policy1)

dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies = 0

for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy1[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy2[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy3[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy4[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy5[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy6[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy7[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy8[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2019TidyNameSizeCyber$policy9[i] = replace_na(dataCyberSecuritySurv
        -10005)
```

```
    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy1[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            1
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy2[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            10
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy3[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            100
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy4[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            1000
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy5[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            10000
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy6[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            1e+05
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy7[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            1e+06
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy8[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = NA
    }
    if (dataCyberSecuritySurvey2019TidyNameSizeCyber$policy9[i] == 1) {
        dataCyberSecuritySurvey2019TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2019
            1e+08
    }

}

######################################################## 2020
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2020TidyNameSizeCyber$policy1)

dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies = 0
```

```
for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy1[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy2[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy3[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy4[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy5[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy6[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy7[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy8[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2020TidyNameSizeCyber$policy9[i] = replace_na(dataCyberSecuritySurv
        -10005)

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy1[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            1
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy2[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            10
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy3[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            100
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy4[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            1000
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy5[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            10000
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy6[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            1e+05
```

```
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy7[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            1e+06
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy8[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = NA
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$policy9[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2020
            1e+08
    }

}


######################################################## 2021
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyNameSizeCyber$policy1)

dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies = 0

for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy1[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy2[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy3[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy4[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy5[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy6[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy7[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy8[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2021TidyNameSizeCyber$policy9[i] = replace_na(dataCyberSecuritySurv
        -10005)

    ## can't make this a switch case because they are all different
    ## variables ffs
```

```
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy1[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            1
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy2[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            10
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy3[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            100
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy4[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            1000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy5[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            10000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy6[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            1e+05
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy7[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            1e+06
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy8[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = NA
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$policy9[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2021
            1e+08
    }

}



##################################################### 2022
##################################################### ########################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyNameSizeCyber$policy1)

dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies = 0
```

```
for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy1[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy2[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy3[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy4[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy5[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy6[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy7[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy8[i] = replace_na(dataCyberSecuritySurv
        -10005)
    dataCyberSecuritySurvey2022TidyNameSizeCyber$policy9[i] = replace_na(dataCyberSecuritySurv
        -10005)

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy1[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            1
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy2[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            10
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy3[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            100
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy4[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            1000
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy5[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            10000
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy6[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            1e+05
```

```
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy7[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            1e+06
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy8[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = NA
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$policy9[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$Policies[i] = dataCyberSecuritySurvey2022
            1e+08
    }

}

## another day of garbage collection of unused columns

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(policy1:policy9))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(policy1:policy9))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(policy1:policy9))
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(policy1:policy9))
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(policy1:policy12))
```

```
## taking care of the columns that are only in the 2018 survey

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(doc1:doc6))

## removing the question about if they know about the 10 steps for
## cyber security Spoiler alert: knowing about it doesnt mean you apply
## it and you can learn about it from other sources either way
## https://www.ncsc.gov.uk/collection/10-steps/risk-management

## same for the next question about the cyber essential scheme

## nevermind they removed all the rest of the questions until business
## standard on 2019 and 2020 (Q43)

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(tensteps:implemb))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(tensteps:implemb))
```

```r
## TODO: think if I should keep the review of cyber security
## documentation colum removing it since i don't have it on 2018 and I
## don't think I will be using it

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-review)
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-review)
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-review)
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-review)

## in 2022 they asked some proper questions about cyber security
## training and cyber security strategy that will be removed for lack
## of comparrison with the other years

dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(trained:corprisk))


## Removing the question about cyber security conserns in the suppliers
## because 1 - most institutions evaluated here won't be in a scale
## where that is an important question 2 - if you were a big
## institution you would just have taken of most of the inside managed
## and now would worry about the suppliers on that level you will just
## get multiple suppliers in case your main supplier fails removing the
## SPOF(single point of failure) that way like what would they do
## anyway such a poorly written question, just hire me to write next
## year survey instead

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-supply)

dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-supply)

dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(supplyrisk1:supplyrisk2))

dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(supplyrisk1:supplyrisk2))

dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(supplyrisk1:supplyrisk2))

## questions about supplier standards were deleted after 2019
```

```r
dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(adhere1:cloud))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(adhere1:cloud))

## only asked in 2022 so not relevant for comparrison

dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(barrier1:barrier8))
```

```r
## type of attacks that targetted the institution type 01 - ramsonware
## infection type 02 - spyware, malware or other type of infection type
## 03 - DDOS (distributed denial of service) type 04 - hacking online
## bank accounts type 05 - phising - impersonating your organisation
## type 06 - phising - fraudulent emails or website targetting staff
## type 07 - unauthorized access by internal staff members type 08 -
## unauthorized access by outsiders type 09 - other type of cyber
## attacks type 10 - don't know type 11 - (don't care) none of these
## type 12 - refused to answer type 13 - unauthorized access by
## students (to be merged with type 7) // only used starting from 2022
## type 15 and 16 were only collect in 2022 so not to be compared and
## type 14 just does not exist for some reason

######################################################## 2018
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2018TidyNameSizeCyber$type1)

dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type1[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type2[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type3[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            100
    }
```

```
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type4[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1000
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type5[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10000
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type6[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+05
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type7[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type8[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+07
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type9[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+08
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type10[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = NA
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type11[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+10
    }
    if (dataCyberSecuritySurvey2018TidyNameSizeCyber$type12[i] == 1) {
        dataCyberSecuritySurvey2018TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+11
    }

}

######################################################## 2019
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2019TidyNameSizeCyber$type1)

dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes = 0

for (i in 1:numberOfCycles) {
```

```
## can't make this a switch case because they are all different
## variables ffs

if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type1[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type2[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        10
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type3[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        100
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type4[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1000
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type5[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        10000
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type6[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1e+05
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type7[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1e+06
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type8[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1e+07
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type9[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1e+08
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type10[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = NA
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type11[i] == 1) {
    dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
        1e+10
}
if (dataCyberSecuritySurvey2019TidyNameSizeCyber$type12[i] == 1) {
```

```
        dataCyberSecuritySurvey2019TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+11
    }

}


######################################################## 2020
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2020TidyNameSizeCyber$type1)

dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type1[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type2[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type3[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            100
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type4[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1000
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type5[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10000
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type6[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+05
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type7[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }
```

```
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type8[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+07
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type9[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+08
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type10[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = NA
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type11[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+10
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type12[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+11
    }
    if (dataCyberSecuritySurvey2020TidyNameSizeCyber$type13[i] == 1) {
        dataCyberSecuritySurvey2020TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }

}


######################################################## 2021
######################################################## ########################################

numberOfCycles = length(dataCyberSecuritySurvey2021TidyNameSizeCyber$type1)

dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes = 0

for (i in 1:numberOfCycles) {

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type1[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type2[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10
    }
```

```
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type3[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            100
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type4[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type5[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10000
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type6[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+05
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type7[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type8[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+07
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type9[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+08
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type10[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = NA
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type11[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+10
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type12[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+11
    }
    if (dataCyberSecuritySurvey2021TidyNameSizeCyber$type13[i] == 1) {
        dataCyberSecuritySurvey2021TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }

}

########################################################## 2022
```

```
##################################################### ########################################

numberOfCycles = length(dataCyberSecuritySurvey2022TidyNameSizeCyber$type1)

dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes = 0

for (i in 1:numberOfCycles) {

    dataCyberSecuritySurvey2022TidyNameSizeCyber$type13[i] = replace_na(dataCyberSecuritySurve
        -10001)

    ## can't make this a switch case because they are all different
    ## variables ffs

    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type1[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type2[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type3[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            100
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type4[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1000
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type5[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            10000
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type6[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+05
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type7[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type8[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+07
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type9[i] == 1) {
```

```
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+08
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type10[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = NA
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type11[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+10
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type12[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+11
    }
    if (dataCyberSecuritySurvey2022TidyNameSizeCyber$type13[i] == 1) {
        dataCyberSecuritySurvey2022TidyNameSizeCyber$attackTypes[i] = dataCyberSecuritySurvey2
            1e+06
    }

}


dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(type1:type12))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(type1:type12))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-(type1:type13))
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-(type1:type13))
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
    select(-(type1:type16))


## i think I will also remove the frequency of the attack since I won't
## be using them for anything and also removing the //TODO review this
## choice

dataCyberSecuritySurvey2018TidyNameSizeCyber = dataCyberSecuritySurvey2018TidyNameSizeCyber %>
    select(-(freq:numbb))
dataCyberSecuritySurvey2019TidyNameSizeCyber = dataCyberSecuritySurvey2019TidyNameSizeCyber %>
    select(-(freq:numbb))
dataCyberSecuritySurvey2020TidyNameSizeCyber = dataCyberSecuritySurvey2020TidyNameSizeCyber %>
    select(-freq)
dataCyberSecuritySurvey2021TidyNameSizeCyber = dataCyberSecuritySurvey2021TidyNameSizeCyber %>
    select(-freq)
dataCyberSecuritySurvey2022TidyNameSizeCyber = dataCyberSecuritySurvey2022TidyNameSizeCyber %>
```

```
    select(-freq)

## now to register both outcome and impact
```

```
## for the frequency of attacks in the last 12 months I am not sure if
## I am interested in that data TODO I am temporarly removing them if I
## can add thhem back if needed (more interested in the outcome of the
## attacks)

## outcomes from the attacks outcome 01 - Software or systems were
## corrupted or damaged outcome 02 - Personal data was altered,
## destroyed or taken outcome 03 - Permanent loss of files (other than
## personal data) outcome 04 - Temporary loss of access to files or
## networks outcome 05 - Lost or stolen assets, trade secrets or
## intellectual property outcome 06 - Money was stolen outcome 07 -
## (DDOS) website or online services were taken down or made slower
## outcome 08 - Lost access to any third-party services you rely on
## outcome 09 - Don't know (NA) outcome 10 - none of these outcome
## 11,12 and 13 are only present in 2022 so we won't use them to make
## comparrisons

########################################################## 2018
########################################################## ########################################
```

::: {.cell}

```{.r .cell-code}
1 + 1
```

```
[1] 2
```

:::

```
1 + 1
```

```
[1] 2
```

```
1 + 1
```

```
[1] 2
```

```
1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
  1 + 1
```

```
[1] 2
```

```
1 + 1
```

```
[1] 2
```

4 Quarto

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

5 Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
1 + 1
```

```
[1] 2
```

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).

Department For Digital, Culture. 2020. "Cyber Security Breaches Survey, 2020." UK Data Service. https://doi.org/10.5255/UKDA-SN-8638-1.