

data__clean

Yiling Yang

11/30/2019

```
library(tidyverse)
library(ggplot2)
library(broom)
library(plotly)
library(highcharter)
library(rworldmap)
library(gganimate)
library(transformr)
library(corrplot)
library(leaps)
library(kableExtra)
```

```
maindata <- read_csv("~/Desktop/master 2.csv")
```

```
maindata <- maindata %>%
  janitor::clean_names() %>%
  select(-("hdi_for_year")) %>%
  mutate(suicideRate = (suicides_no/population)*100,
         country = as.factor(country),
         age=str_remove(age, "years"),
         gdp_for_year = as.numeric(gsub(",","",gdp_for_year)),
         age_group=as.factor(age),
         suicides_no = as.integer(suicides_no),
         population = as.integer(population),
         sex = as.factor(sex),
         generation = as.factor(generation)) %>%
  filter(year!=2016)
```

```
maindata$year <- as.integer(maindata$year)
```

```
maindata <- maindata %>% group_by(year,country,sex,age,population,gdp_per_capita) %>% summarise(suicideRate = (suicides_no/population)*100)
```

```
write_csv(maindata, "maindata.csv",row.names = FALSE)
```