Datathon Testing doc

Team 34 - Ananya Trivedi, Omkaar Kamath, Amy Lai

11/20/2021

```
setwd("./") # Run this line to setup environment
print(getwd())
```

```
## [1] "/Users/omkaar/Downloads/datathon"
```

Data Retreival

```
suicide <- read.csv("datathon_data/suicide.csv", header=TRUE,
   sep=",")
life.expectancy <- read.csv("datathon_data/life.expectancy.csv", header=TRUE,
   sep=",")
co2.emissions <- read.csv("datathon_data/co2.emissions.csv", header=TRUE,
   sep=",")
obesity <- read.csv("datathon_data/obesity.csv", header=TRUE,
   sep=",")
head(suicide)
```

```
##   country year    sex         age suicides_no population suicides.100k.pop
## 1 Albania 1987   male 15-24 years          21     312900              6.71
## 2 Albania 1987   male 35-54 years          16     308000              5.19
## 3 Albania 1987 female 15-24 years          14     289700              4.83
## 4 Albania 1987   male   75+ years           1      21800              4.59
## 5 Albania 1987   male 25-34 years           9     274300              3.28
## 6 Albania 1987 female   75+ years           1      35600              2.81
##   country.year HDI.for.year gdp_for_year.... gdp_per_capita....      generation
## 1  Albania1987           NA    2,156,624,900                796    Generation X
## 2  Albania1987           NA    2,156,624,900                796          Silent
## 3  Albania1987           NA    2,156,624,900                796    Generation X
## 4  Albania1987           NA    2,156,624,900                796 G.I. Generation
## 5  Albania1987           NA    2,156,624,900                796         Boomers
## 6  Albania1987           NA    2,156,624,900                796 G.I. Generation
```

Data Cleaning

Data Processing

Eliminate variates that have weak to no correlation

Data Merging

Data Exploration

Growth of attributes over time

Multi-variate analysis with lines of different colors for countries