Assignment 2 Report

**Statistical and Mathematical methods for Data Science**

 | FAST National university of computers and emerging science

EDA on dataset Who suicides

Rafia Shaikh

2018

# Introduction

I have taken dataset from kaggle website, “WHO Suicide Statistics “Basic historical (1979-2016)

Dataset is in CSV (comma separated file) format, having six columns (country, year, age, sex, suicides\_no and population) and dataset contains about 43776 rows

Using these datasets perform data analysis that mostly who commit suicides and predict either the ratio of population of suicides over years increases or decreases

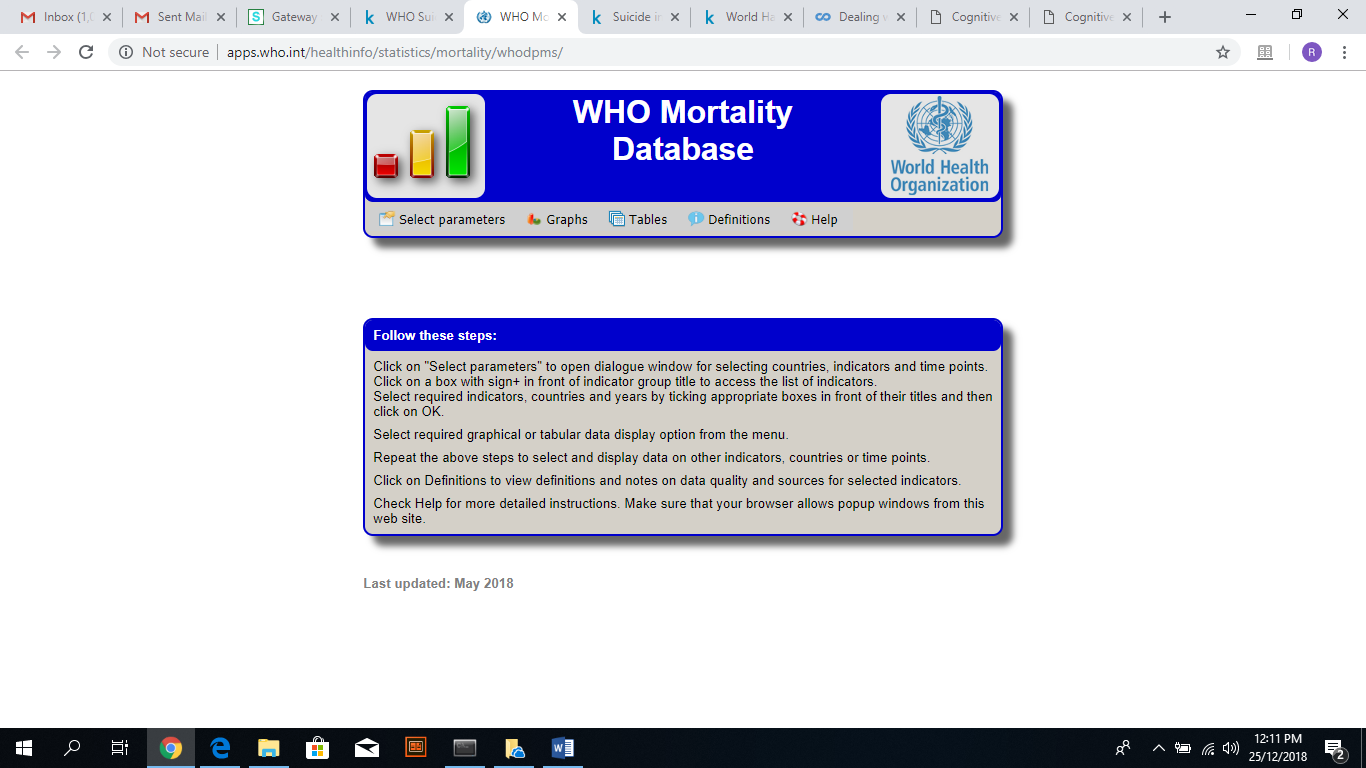
# Objective

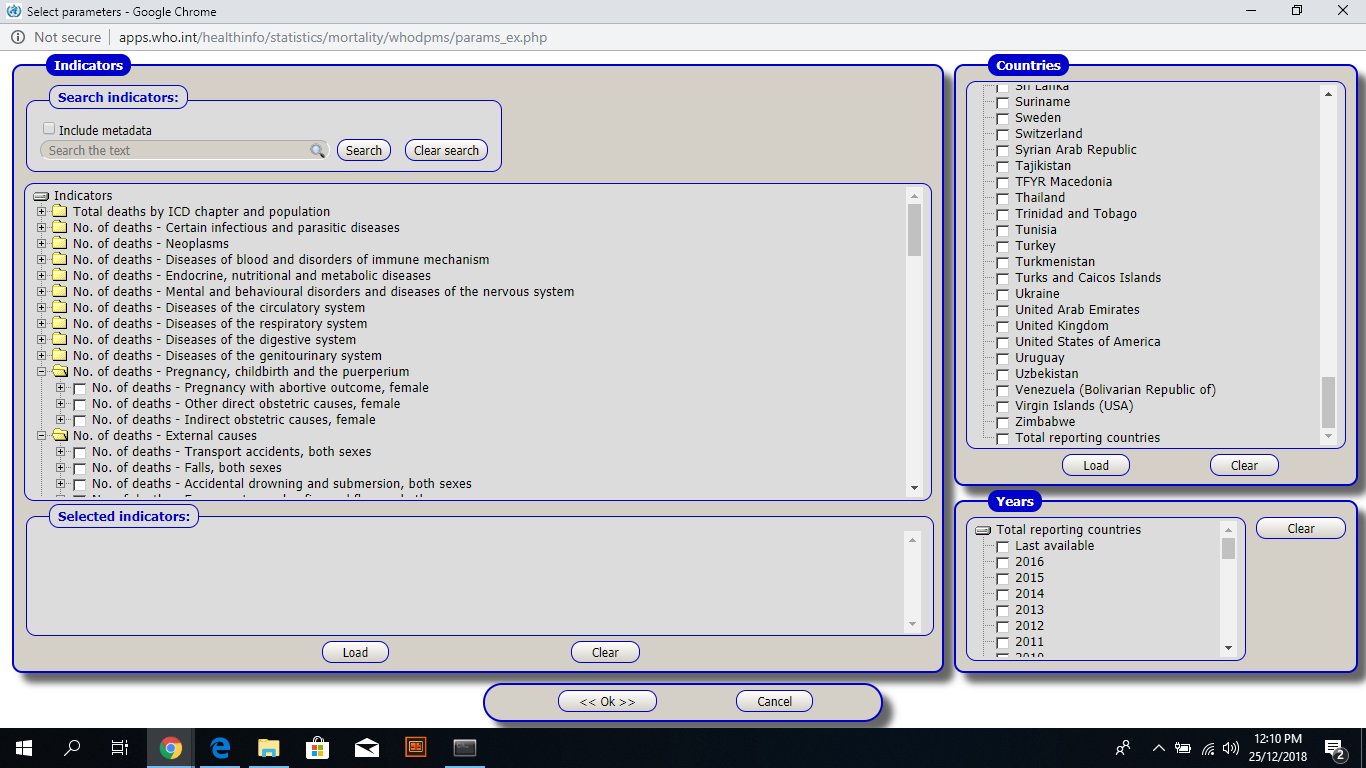
perform EDA on dataset

My question: is suicides getting better or worse over years?

Who suicides mostly male or female?

# Data set





# Tool

Python 3

IDE – Anaconda3

# Project working and code:

Project code and working is available in zip folder

# Results and conclusions:

After visualizing data set, it is concluded that most ratio of suicides is male, and female ratio is rarer as compare to male

And it is concluding that over years’ suicide rate decrease

