



# DATA WIZARDS

PEER REVIEW/PROJECT  
PRESENTATION



# PROBLEM STATEMENT

Suicide is a complex issue and therefore suicide prevention efforts require coordination and collaboration among multiple sectors of society, including the health sector and other sectors such as education, labour, agriculture, business, justice, law, defense, politics, and the media. These efforts must be comprehensive and integrated as no single approach alone can make an impact on an issue as complex as suicide.

- \* This Project is aimed at accessing Mental health generally by analyzing global suicide rate data
- \* Predicting sentiment and Suicide intentions in tweets from Twitter.com

AIM:

TO FORECAST AND PREDICT  
SUICIDE RATES AROUND  
THE WORLD AND  
EXTRAPOLATE DATA FROM  
THE SAME.

# CONTENTS OF THIS PRESENTATION:

- 1) INTRODUCTION TO THE TOPIC
- 2) QUESTIONS POSED BY PEERS WHO REVIEWED US
- 3) EDA COMPONENT
- 4) MODELS USED
- 5) PAPER REPORT

TEAM THAT REVIEWED US

# WORK COMPLETED + MODELS USED



So far we have completed the Exploratory Data Analysis (EDA) for our project and are on track to complete our first model soon. This model shall be the first version of many, to help us achieve accuracy and precision in our predictions, forecasts and inferences.

We have already created an SLR Model.

Our final goal is to finish a project that is super accurate at extracting suicidal tendencies from tweets + figure out high stress situations that cause these feelings in the first place.

## QUESTIONS POSED BY REVIEWING TEAM:

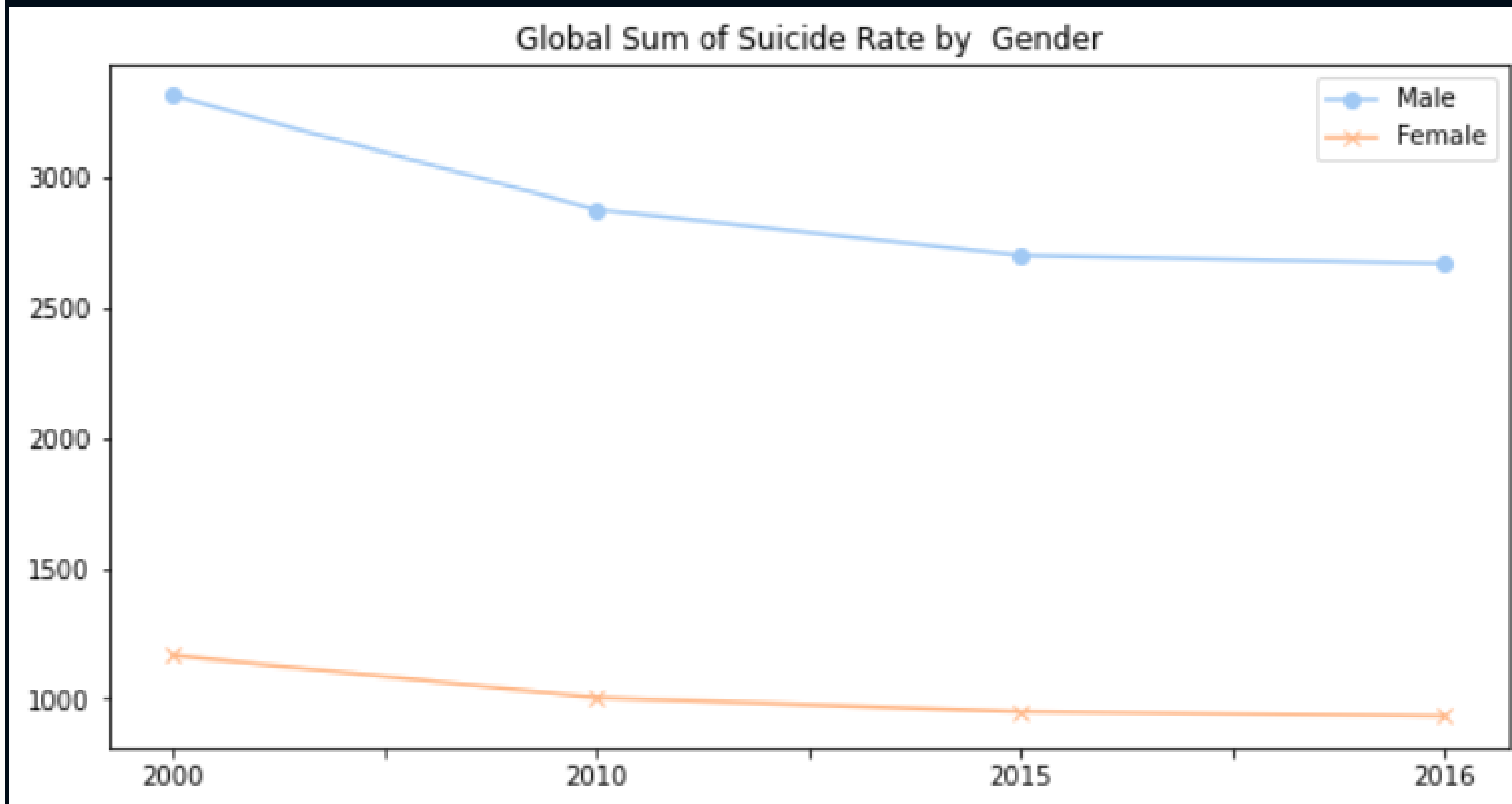
- 1) WHAT IS THE EDA PERFORMED
- 2) HOW ACCURATE IS THE DATA AND HOW HAVE MISSING VALUES, BEEN TAKEN INTO ACCOUNT?
- 3) WHY ONLY ONE MODEL? ARE THERE PLANS TO IMPLEMENT ANY MORE?

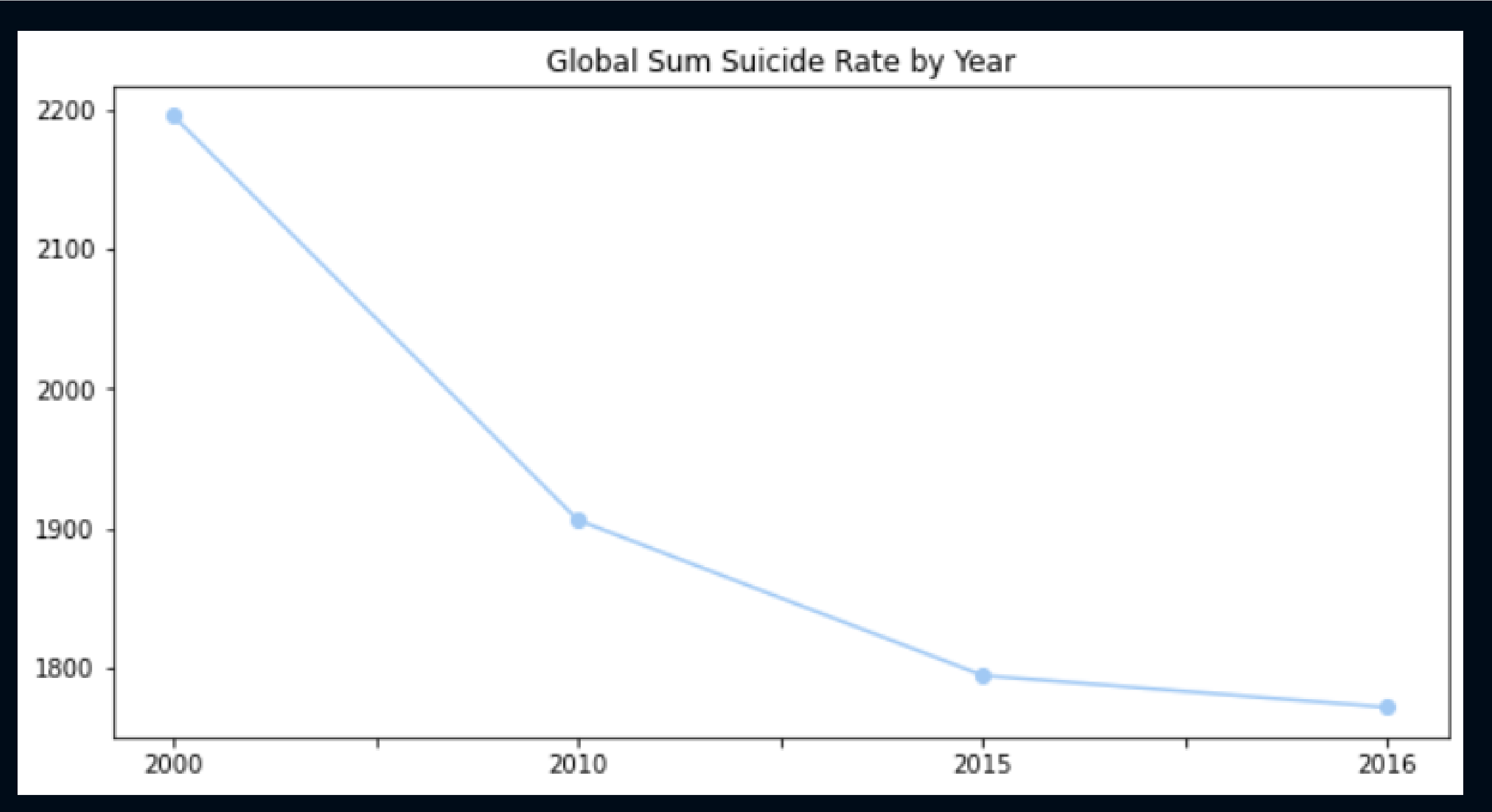
HAS THE TEAM PERFORMED EDA?

YEP. OUR EDA IS QUITE EXTENSIVE

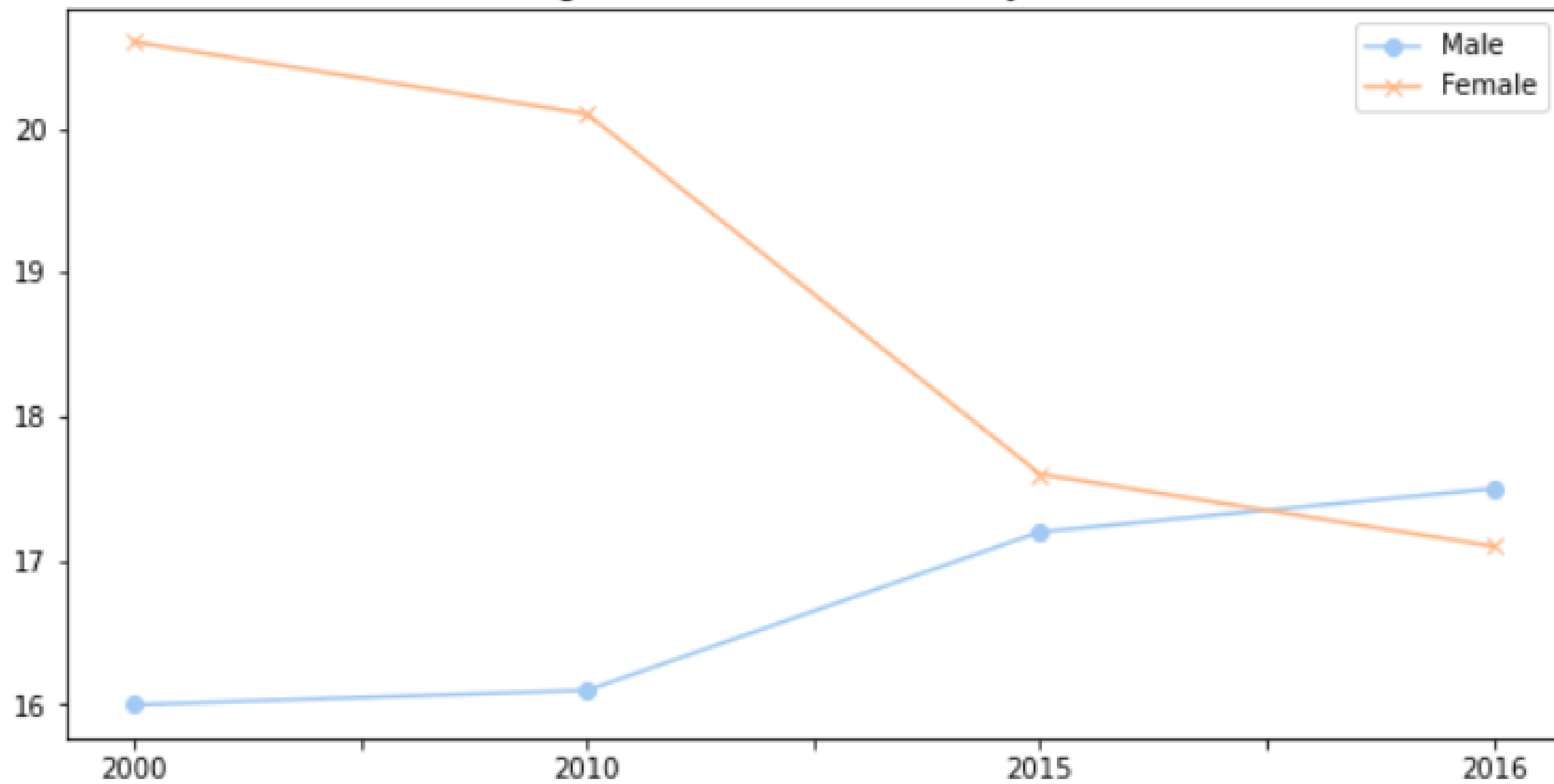


<matplotlib.legend.Legend at 0x164449c1670>

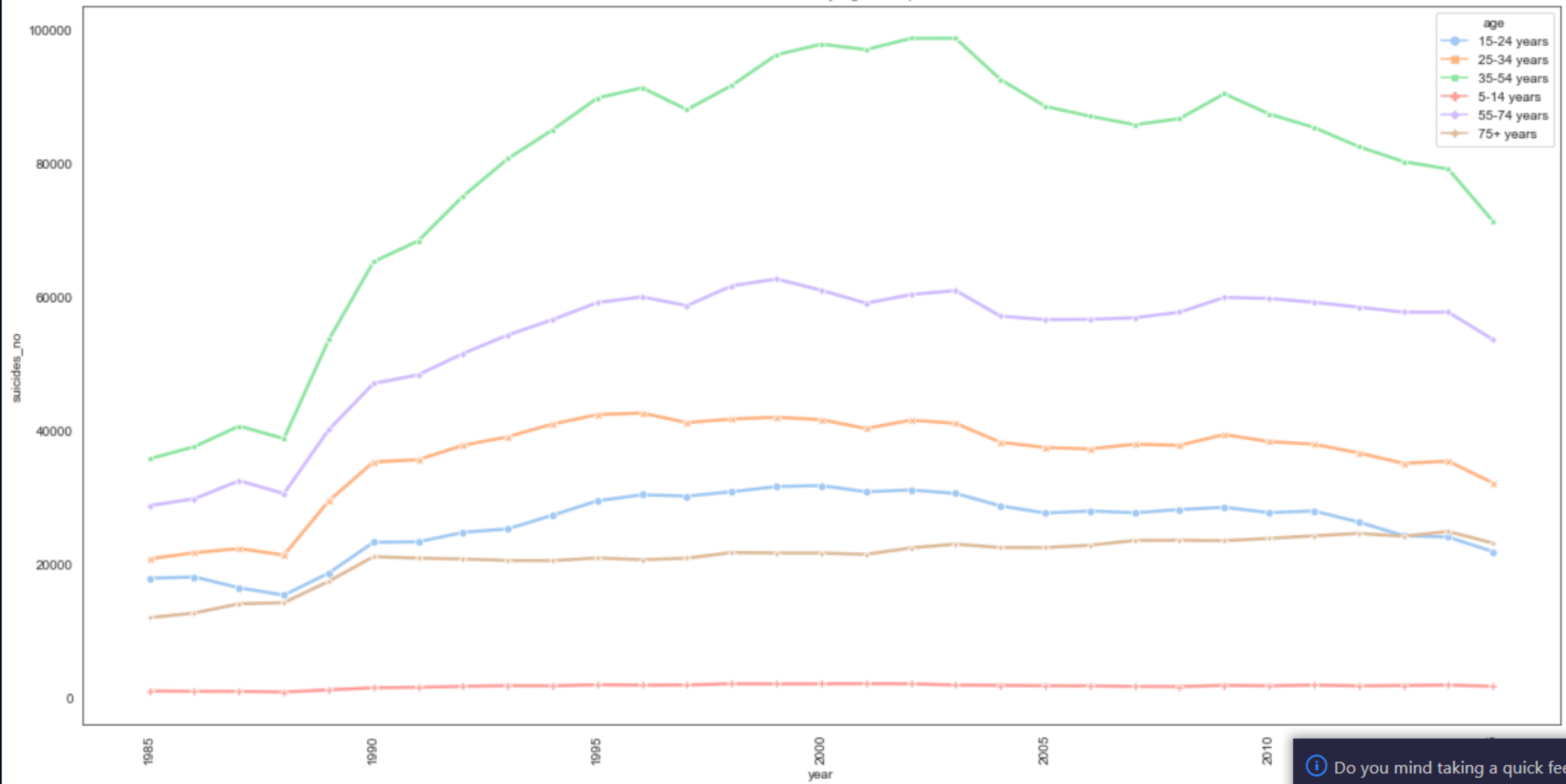




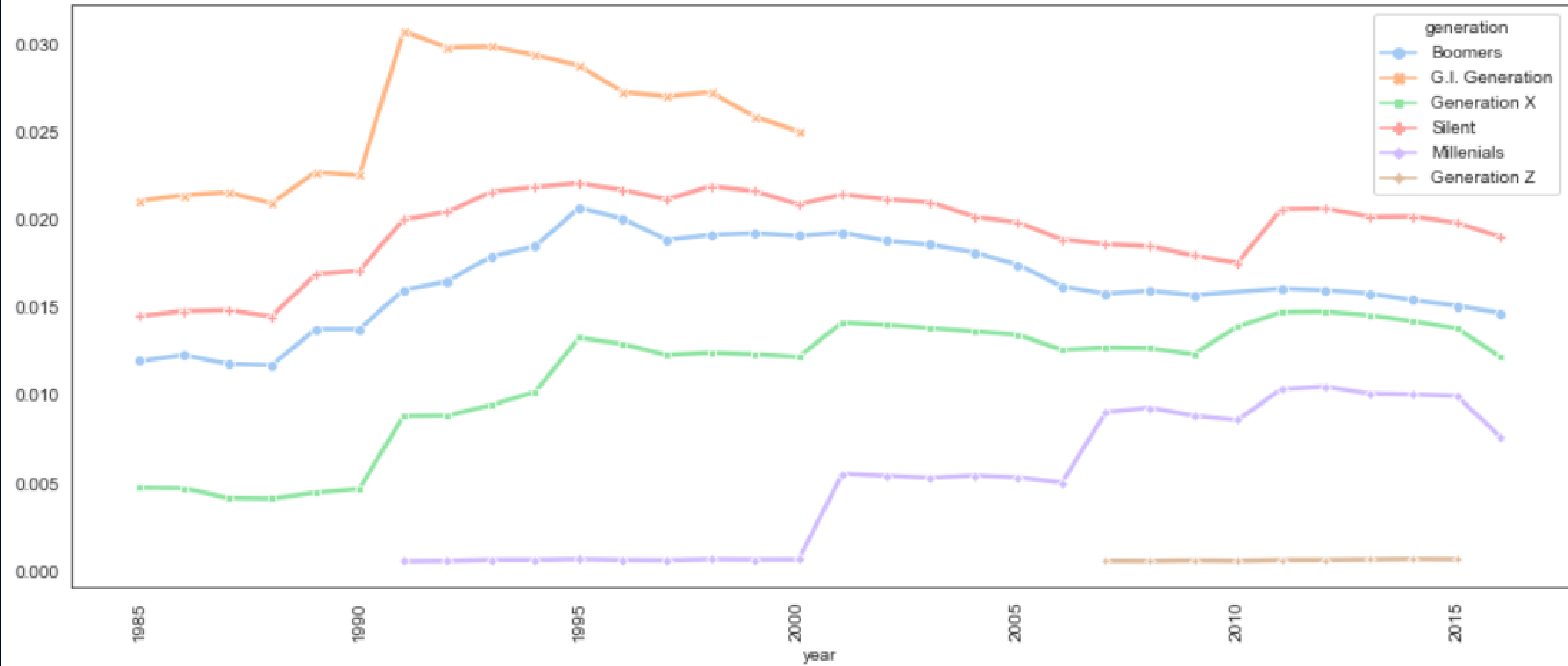
Nigeria Sum of Suicide Rate by Year



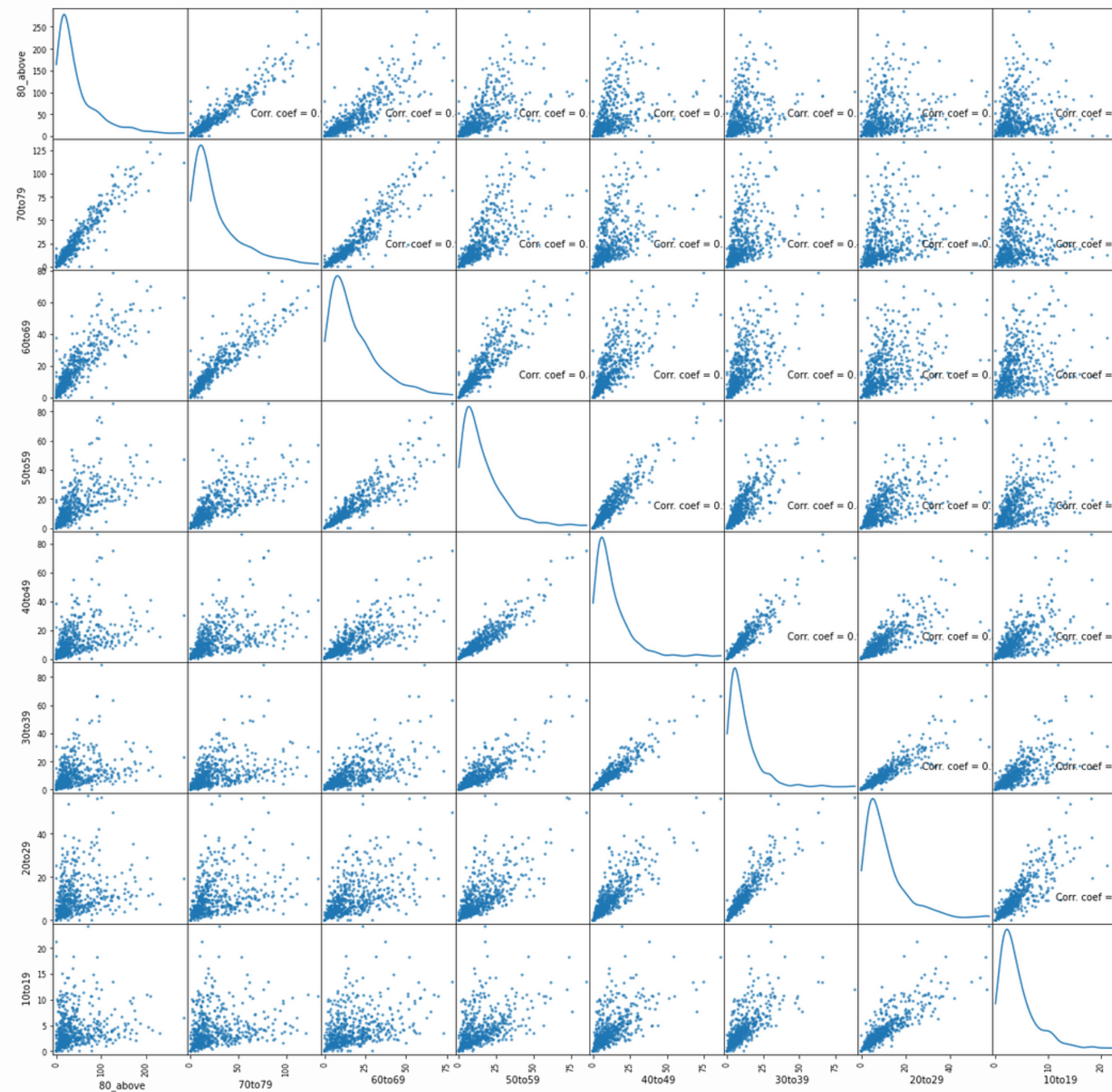
Suicide Rate by Age Group



Suicide Rate by Generation



Scatter and Density Plot

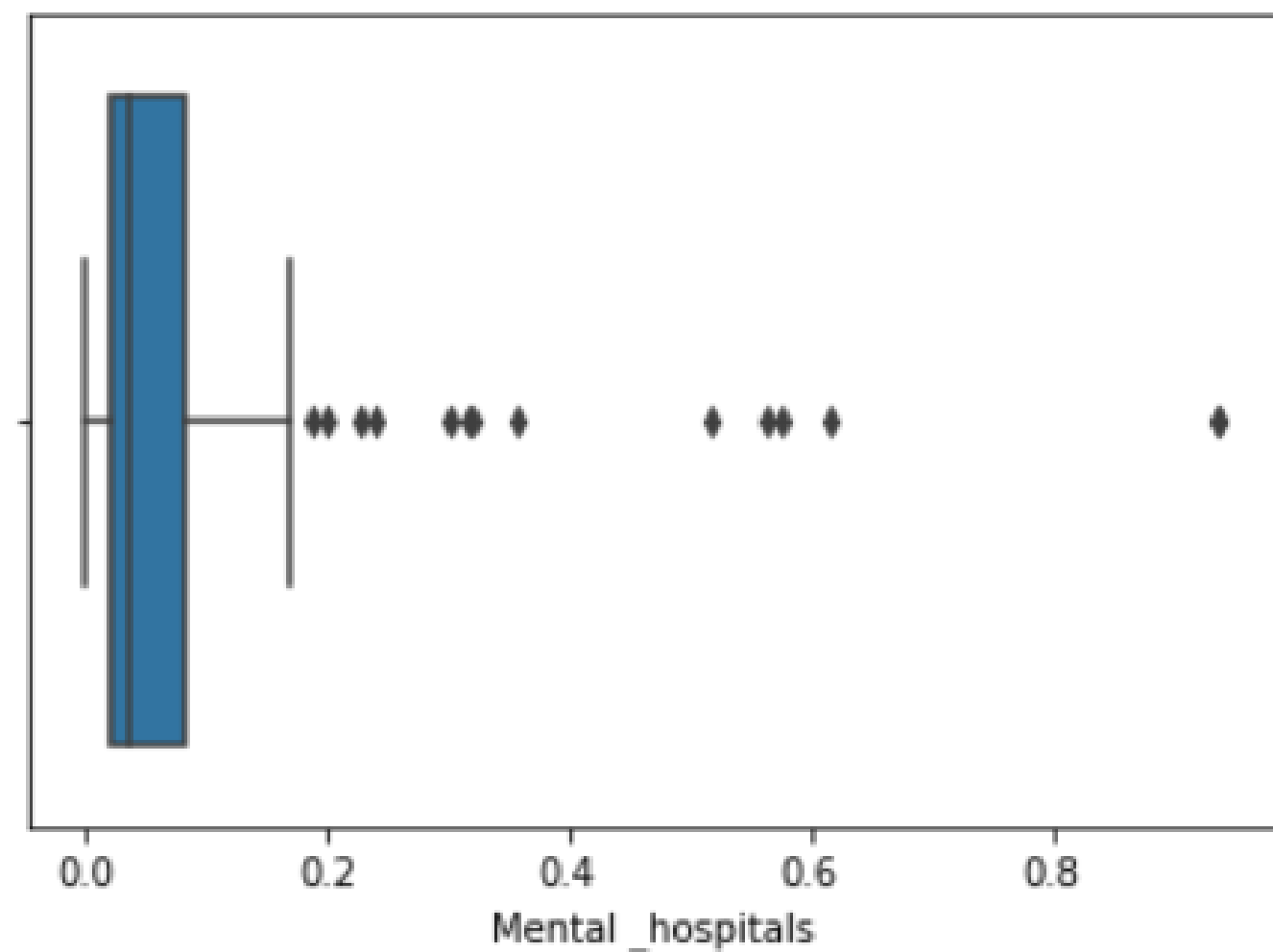


H A S   T H E   T E A M   P R E P R O C E S S E D   T H E   D A T A ?

Y E P

```
median = facilities['Mental_hospitals'].median()
facilities["Mental_hospitals"] = np.where(facilities["Mental_hospitals"] > 1, median, facilities['Mental_hospitals'])
```

```
sns.boxplot(x=facilities["Mental_hospitals"]);
```





HAS THE TEAM USED ONE OR MORE  
MODELS?

YES.

WE HAVE USED SIMPLE LINEAR  
REGRESSION (SLR) TO OBTAIN SOME  
INSIGHT

# Simple Linear Regression

df\_age

89]

	Country	Sex	2016	2015	2010	2000	Sum
0	Afghanistan	Both sexes	6.4	6.6	7.4	8.1	28.5
1	Afghanistan	Male	10.6	10.9	12.5	14.3	48.3
2	Afghanistan	Female	2.1	2.1	2.1	1.7	8.0
3	Albania	Both sexes	5.6	5.3	7.7	5.8	24.4
4	Albania	Male	7.0	6.7	9.5	8.2	31.4
...	...	...	...	...	...	...	...
544	Zambia	Male	17.5	17.4	17.9	21.9	74.7
545	Zambia	Female	6.2	6.1	6.2	7.5	26.0
546	Zimbabwe	Both sexes	19.1	18.9	20.6	21.7	80.3
547	Zimbabwe	Male	29.1	28.7	32.3	35.5	125.6
548	Zimbabwe	Female	11.1	11.1	11.1	9.8	43.1

549 rows × 7 columns

```
df_reg = df_age.sort_values("Sum", ascending = False)
df_reg.reset_index(drop = True, inplace = True)
df_reg.head()
```

```
#Error checks
MAE= mean_absolute_error(y_test,pred)
MSE= mean_squared_error(y_test,pred)
RMSE= np.sqrt(mean_squared_error(y_test,pred))
r2= r2_score(y_test,pred)
```

50]

```
print( "MAE:", MAE)
print( "MSE:", MSE)
print( "RMSE:", RMSE)
print( "r2:", r2)
```

51]

```
• MAE: 5.302658864579461
  MSE: 52.41057017998232
  RMSE: 7.239514498913744
  r2: 0.15999888695836462
```

```
coeff= pd.DataFrame(model_lr.coef_,x.columns,columns=['Coefficient'])
coeff
```

## PAPER REPORT:

AFTER EXTENSIVE MODELLING AND BRIDGING THE GAP BETWEEN ACCURACY AND PRECISION, IN ORDER TO PREDICT AND FORECAST SUICIDES ACROSS THE WORLD IN VARYING LIVING CONDITIONS. THE SIMPLE LINEAR MODEL WE EMPLOYED, FORECASTED WITH CERTAINTY AND HAD A RELATIVELY LOW RMSE VALUE, WHICH IS REQUIRED FOR BETTER PREDICTIONS. OUR INFERENCES STATE CLEARLY THAT BEING MALE IS A MAJOR FACTOR IN DETERMINING THE PREDICTION OF DEATHS BY SUICIDE SINCE IT INCREASES THE RISK OF SUICIDE BY 8.08 POINTS. AS THE YEARS PROGRESS, SUICIDE RATES ARE COMPARATIVELY RISING, AND THIS IS SOMETHING THAT IS ALARMING. BEING BORN IN GEN Z, IT IS IMPERATIVE WE TAKE STEPS TO ENSURE LONGER LIFESPANS FOR HUMANS AND REDUCE THE MENTAL STRESS AND PROBLEMS THAT COME WITH OUR DAY-TO-DAY LIVES.

NOVELTY: USED SLR TO PREDICT OUTCOMES

IDEATION:  
SUICIDE PREVENTION

THANK YOU!

VIBHAV DEEPAK

KSHITIJ SAHA

SRI CHARAN