

Project Report

On

Suicide Rate Overview From 1985-2006 By using Data Analytics

**Submitted as Major Project of the degree of Master of Computer Application**

Siliguri Institute of Technology

Department of Computer Application

Guide: Submitted By:

Mahendra Dutta Abhishek Basu

Senior Project Developer mca(6th sem)

Ardent Technology Kolkata Roll: 33601017013

Mail-id: Infobasu96@gmail.com

ACKNOWLEDGEMENT

**T**he satisfaction that accompanies that the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all the efforts with success. We are grateful to our project guide **Mr. Mahendra Dutta** for the guidance, inspiration and constructive suggestions that are helpful to us in the preparation of this project.

I also thank my colleagues who have helped in successful completion of the project.

Abhishek Basu

Roll: 3361017013

INDEX

1. Introduction and Objective of the Project
2. Project Category(RDMS/OOPS/NETWORKING/DATA ANALYTICS
3. TOOLS/PLATFORM,HARDWARE AND SOFTWARE REQUIREMENTS
4. Problem Definition, Requirement Specification
5. Scope of the solution
6. Numbers Of modules and their description to provide an estimation of the students’ effort on the project, and along with process logic.
7. Implementation of Methodology
8. Data Analytics Prediction
9. Future Scope of the Project
10. Bibliography

**Introduction and Objective of the Project**

Suicide is one of the major problem of the world that world is facing now. About 800000 people commit suicide worldwide every ware. Death by suicide is extremely complex issue that causes pain to hundreds of thousands of people every year around the world. The objective of this project is to contribute to an informed, open debate about ways to prevent suicide. If you are dealing with suicidal death then our project give you full overview of worldwide suicide rate from year 1985 to 2016 in a organized manner. Not only full data in an pictorial or table form we also predict the data.

Suicide is one of the most complex issue, and while it is not possible to pin down its’ causes. There are risk factors that have been identified, mainly through correlations. Mental health, specifically depression, is widely recognized as the most important risk factor.

In our project we give you a vivid description of data about suicide

rate of various country from 1985 to 2016 in a analyzed form and also predict the data about future as well as also calculate the accuracy of prediction. This project is based on Data Analytics feature where various tools and modules are used to analyze the data. Here we provide the specific information from raw data which we collect from internet. Based on that data we provide a vivid representation of suicide rate of the world based on their countries and other parameters.

**Project Category:**

This project is based on the **Data Analytics** Method. Here We first take some data in the form of csv file(Coma separated values).Then by using data analytics method with the help of python language we analysis the data,wich contain various information about the country wise yearly suicide rate from 1985-2016.From that comma separated values we provide a visualization of data in a bar graph and pie chart manner. That any non-technical person can easily understand about the data.

Not only provide only visualization of the data in graphical manner but also provide suicide rate prediction also with the help of machine learning method. Here we use various types of python modules which contain some predefined built-in method. Like NUMPY,Scikit-Learn,Matplotlib,random,flask ,pandas,seaborn etc.

**Platforms/Hardware/Software Requirements**:

1. **Hardware Requirements:-**

* Computer:HP core I3 7th generation
* Processor: Intel core i3 7th generation processor
* Ram:4GB or Higher Is needed
* Cache Memory::512 MB
* Hard Disk:1 TB or 500 GB is applicable
* Input Devices: Mouse, Keyboard etc.

1. **Software Requirements:-**

* Anaconda distribution for development of the

Project.

* Spyder ide (integrated development Environment).
* Various types of modules like Numpy,Pandas,Scikit-learn,

Matplotlib, Tensar Flow,etc.

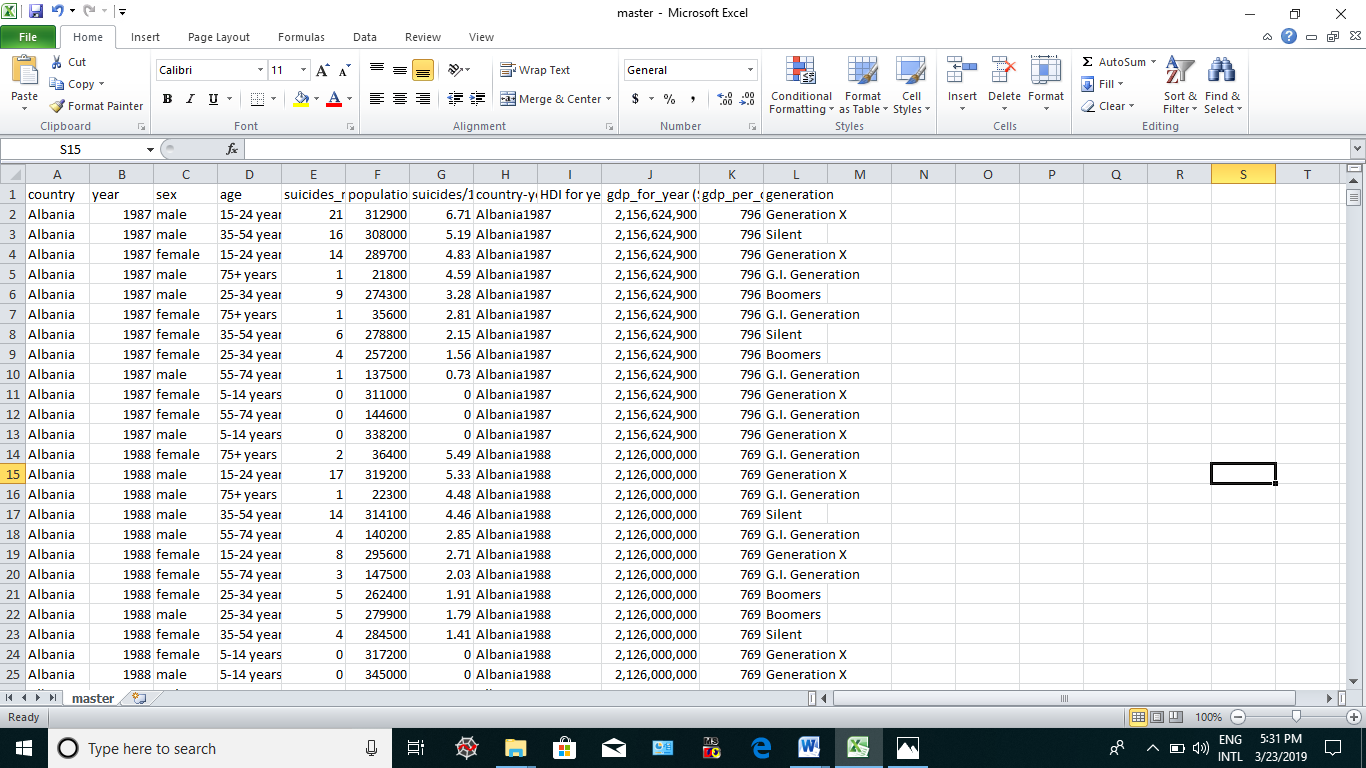
* Operating System: Windows 10 Operating System.

**Problem Requirements and Specification:-**

**S**uicide is one of the major problem that world is facing from some decades large number of people commit suicide worldwide ever year, of these 17% are resides in India. The male to female suicide ratio has been about 2:1 in India . On an average total number of suicide rate in india is per day 300.Suicide is a act of deliberately killing oneself or, more specifically ,an act deliberately initiated and perform by the person concerned in the full knowledge, or expectation, of its fatal outcome. Suicide prevention is one of the term for collective efforts of local citizen organizations, health professionals and related professionals to reduce the incidence of suicide. Other than direct interventions to stop an impending suicide. In this project we analyze the suicide rates of various countries based on their age ,sex from the year of 1985-2006. Where we give the pictorial representation of the data in an organized manner. By using machine learning technique we also provide the predication of the data.Find out the suicide rate of future year and give an overall view of suicide rate of various countries. Which we think help government to take the right decision to prevent the suicide attempt.

**Database Creation:**

A dataset is a collection of data. Mostly a data set refers to the contents of a single database table, or a single statistical data matrix, where every table column represents a particular variable, and each table row corresponds to a given members of the data set. The objective of project is to find the relations between the dataset to predict the future dataset.



**Technology Used:-**

**Machine Learning:**

Machine learning is a study of computer Science that provides computers the

ability to learn without being explicitly programmed. Machine learning is used

to used to study algorithms that learn from and make predictions on data.

Machine learning is related to computational statistics, which also focuses on

Prediction making. With in the field of data analytics, machine learning is a

Method used to devise complex, models and algorithms that lend themselves

to prediction;in commercial use,this is known as predictive analytics.Machine

learning focuses on the development of computer programs that can access

data and use it learn for themselves.The learning process begins with observa

tions or data,examples,direct experience,or instruction,in order to find patter

ns in data make better decisions in the future based on the examples provides

The objective is to allow the computers learn automatically without human as

-Istence and adjust actions accordingly.

Modules:-

Data analytics:

Data analysis is a process of inspecting,cleansing transforming and modeling

data with the goal of discovering useful information,suggesting conclusions,

and supporting decision making.Data mining is a data analysis technique.

Data analytics is performed on the suicide database so that the data can be

Cleaned and data that is not required can be deleted. It is used to model the

Complex suicide data to simpler form so that it can be used as input for predict

-tions and visualizations. Data visualizations is one of the way of presenting

Data in pictorial or graphical format. It enables decision makers to see analytic

Presented visually to understand difficult concept or to identify new patterns.

Visualization help to make charts and graphs for more details, thereby changing

what data you see and how it’s processed. This helps us to understand which

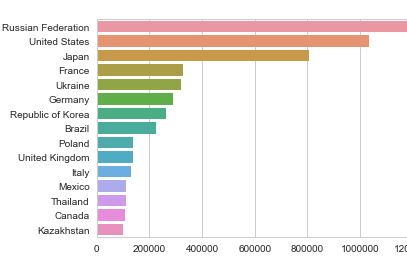
features of data have strong relations between them.

Linear Regression:-

Linear Regression is a linear approach used for modeling the relationship between a scalar dependent variable y and one or more independent variable X.

Linear regression has many practical uses.If the goal is prediction,or forecasting,or error reduction, linear regression can be used to fit predictive model to an observed data set of y and x values. After developing this model to an observed.

With simple linear regression as shown in the equation **y=mx+c**.**here y is the output variable,x is the input variable and c and m are coefficients we need to estimate.c called the intercept because it determines wheather line intercepts the y axis.The M term is called slopes.**

** Fig1.1 Suicide rate data Analytics from 1985-2006 countrywise.**

**Fig 1.2 Suicide rate of male and female year wise bar graph**

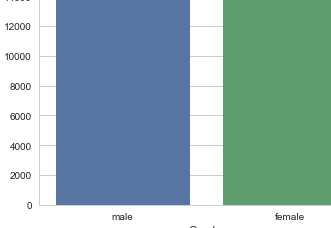
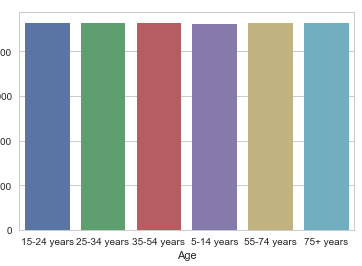
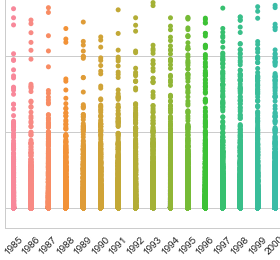
****

Fig 1.3 Bar graph representation age wise suicide rate from 1985 to 2016:-



Fig 1.4 suicide rate of various countries based on the year 1985-2016

***Number Python Modules used in this projects and their vivid description:-***