A

PROJECT REPORT

ON

"Empirical Exposition Model to Analyze Suicide Causes in India"

SUBMITTED

TO

SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE

FOR THE PARTIAL FULFILLMENT

**MASTER IN COMPUTER APPLICATION
(MCA-I,SEM-II)**

SUBMITTED BY

MR. MIHIR DIWAKAR LIMJE

UNDER THE GUIDANCE OF

DR. POONAM SAWANT

**SINHGAD INSTITUTE OF MANAGEMENT AND COMPUTER APPLICATION,
PUNE
A.Y. 2021-22**

CERTIFICATE

This is to certify that, the project entitled **"Empirical Exposition Model to Analyze Suicide Causes in India"** being submitted for the partial fulfillment of Master of Computer Application by her/him to Sinhgad Institute of Management and Computer Application affiliated to Savitribai Phule Pune University, Pune is the result of the original work completed by Mihir Diwakar Limje under the guidance of Dr. Poonam Sawant.

To the best of our knowledge and belief, this work has not been previously submitted by the award of any degree or diploma of Savitribai Phule Pune University or any other University.

PLACE:

DATE:

Dr. Poonam Sawant   **Dr. Ashwini Brahme**   **Dr. Vilas Nandavadekar**

Internal Guide   **Project Co-Ordinator**   **Director SIMCA-MCA**

External Examiner

DECLARATION

I hereby declare that the project entitled **"Empirical Exposition Model to Analyze Suicide Causes in India"** submitted for the partial fulfillment of **Master of Computer Application (MCA)** by me to **Savitribai Phule Pune University, Pune** is the outcome of my original research work carried out under the guidance of **Dr. Poonam Sawant** and has not previously formed the basis for the award of any Degree or Diploma or other similar title of this or any other University or Examining Body.

Place: Pune

Date:

Mihir Diwakar Limje

(Student)

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my project. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank Dr. Poonam Sawant, for providing me an opportunity to do the project and giving us all support and guidance which made me complete the project duly. I am extremely thankful to her for providing such a nice support and guidance.

I take this opportunity to thank my project coordinator Dr. Ashwini Brahme and our Director Dr. Vilas Nandawadaker, for encouragement and guidance throughout the progress of this report.

Mihir Diwakar Limje

PREFACE

Suicide is the third leading cause of death among young adults worldwide. There is a growing recognition that prevention strategies need to be tailored to the region-specific demographics of a country and to be implemented in a culturally-sensitive manner. This review explores the historical, epidemiological and demographic factors of suicide in India and examines the strategies aimed at the prevention of suicide. There has been an increase in the rates of suicide in India over the years, although trends of both increases and decline in suicide rates have been present. Distinct from global demographic risk factors, In India, marital status is not necessarily protective and the female: male ratio in the rate of suicide is higher. The motives and modes of suicide are also distinct from western countries. Preventive strategies implemented at a community level and identifying vulnerable individuals maybe more effective than global strategies.

Although suicide is a deeply personal and an individual act, suicidal behavior is determined by a number of individual and social factors. Ever since Esquire wrote that "All those who committed suicide are insane" and Durkheim proposed that suicide was an outcome of social / societal situations, the debate of individual vulnerability vs social stressors in the causation of suicide has divided our thoughts on suicide. Suicide is best understood as a multidimensional, multifactorial malaise. Suicide is perceived as a social problem in our country and hence, mental disorder is given equal conceptual status with family conflicts, social maladjustment etc. According to the official data, the reason for suicide is not known for about 43% of suicides while illness and family problems contribute to about 44% of suicides.

Divorce, dowry, love affairs, cancellation or the inability to get married (according to the system of arranged marriages in India), illegitimate pregnancy, extra-marital affairs and such conflicts relating to the issue of marriage, play a crucial role, particularly in the suicide of women in India.

CONTENTS

Keywords

- Suicide
- India
- Risk factors
- Depression
- Python
- Data science
- PowerBI
- Reduce
- adopted
- Causes
- States
- Education profile
- Social profile

List of Figures

List of Table

Chapter 1
Introduction

1.1   Introduction

Each suicide is a personal tragedy that prematurely takes the life of an individual and has a continuing ripple effect, dramatically affecting the lives of families, friends and communities. Every year, more than 1,00,000 people commit suicide in our country. There are various causes of suicides like professional/career problems, sense of isolation, abuse, violence, family problems, mental disorders, addiction to alcohol, financial loss, chronic pain etc. NCRB collects data on suicides from police recorded suicide cases. Most public attention in India has focused on suicide in farmers. The age- and sex-specific death totals, rates and risks and the mode of suicide in India's markedly diverse socio-demographic populations are not well understood. The WHO states that about one-third of its Member States provide high-quality suicide data. It will be important to examine the accuracy and reliability of suicide data from India and China in particular since the results from these mega-nations will strongly affect conclusions about global suicide rates and factors that influence them.

1.2   Motivation

India's National Crime Records Bureau (NCRB) has published yearly reports on suicides since 1967 in the States and Delhi and the Union Territories. Their reports showed an increased suicide rate per 100,000 of the total population from 6.3 in 1978 to 8.9 in 1990. The reported rate then fluctuated, rising between 2006 and 2011, when it reached 11.25. Calculated age-standardized suicide rates across the years 2001-2013 using NCRB data. Male rates ranged around 14 per 100,000, while female rates decreased from 9 to 7 per 100,000 over the 13 years. The NCRB reported the total number of suicides in India in 2015 as 133,623, the rate being 10.6 per 100,000 with a male to female ratio of 2.25.

A number of suicide studies conducted in India between the 1970s and the early 2000s have been listed. Most were from defined regions (e.g., West Bengal) with several relating to villages or rural areas, while others focused on cities and data from tertiary care hospitals. Several case-control studies allowed the analysis of possible causative factors in those areas. Verbal autopsy studies in various rural locations reported high suicide rates. Reported an annual suicide rate of 92.1 per 100,000 in a catchment area of Tamil Nadu in 2000-2002, the male to female ratio being 1.5:1. The authors commented that inefficient civil registration systems,

incomplete reporting of deaths, variable standards in certifying death and the legal and social consequences of suicide were the major obstacles to investigating suicide, but in their area, the obstacles were overcome from 1985 onwards by comprehensive data collection involving community health workers who lived in the villages.

In India, unnatural deaths are reported to police, who then investigate and compile a First Information Report (FIR); this states the apparent cause of death, based on the collection of evidence and (where available) autopsy reports. FIRs are provided to the NCRB. Thus, NCRB data are dependent on what the community reports. There is an incomplete medically certified cause of death system that covers only a small proportion of deaths in India and has variable coverage across the States. It is generally agreed that the NCRB under-reports the frequency of suicide.

1.3   Justification of the Problem

A remarkable observation to be made about suicide studies in India is that suicide rates and patterns differ a lot between different States and territories. Possible explanations for these differences may point to factors relevant to the causation of suicide. Referring to 2001-2013 data, southern States had the highest and northern the lowest rates for both genders, while rates in central and western States were mainly mid-level. For example, in Tamil Nadu, the male suicide rate in 2013 was 29.9 per 100,000 and female rate 14.6; Maharashtra's corresponding rates were 19.7 and 8.5; Uttar Pradesh's were 2.9 and 2.1 and Bihar's corresponding rates were 1.1 and 0.8.

The suicide rates in 2016 reported for different States by the GBD Study were proportionally higher (to a variable extent) than those derived from NCRB data. More economically developed States had 3 to 4 times higher rates than less economically developed States, but part of this might be attributed to their better registration of suicide cases. Populations with higher agricultural employment, States with higher levels of male unemployment and States with higher literacy rates had higher risks of suicide. A substantially lower suicide rate than Hindus in northeastern States , where the former were in the majority, whereas elsewhere in India (with Hindus in the majority), rates were lower in Hindu, and rates in Sikh and Muslim were lower still. Rates were lower among scheduled caste and scheduled tribe populations, and 'other backward classes', especially in regions dominated by these groups.

1.4 Significance of the Study

Despite the under-reporting, the NCRB offers significant insights for the planning of suicide interventions. However, the demonstrable heterogeneity of NCRB data recording demands a need to understand the reporting practices at State level. Around the world, there is a variation between jurisdictions in documented numbers and proportions of persons recorded as having died as a result of 'Event of Undetermined Intent' (EUI; introduced as a diagnostic category in the 8th edition of the International Classification of Diseases), or coded as having died from an ill-defined or unknown cause of death[10]. Variable recording of so-called 'undetermined deaths' between States would affect recording of suicide rates. Clearly, the analysis of factors related to causation needs to include attention to factors considered at a State level, and to factors leading to non-inclusion of completed suicide cases among data provided to the NCRB.

1.5 Scope of the Problem

Suicide is the ultimate outcome of a complicated correlation of genetic, biological, physiological, sociological, and environmental factors. The number of suicides in the whole world is estimated at nearly 9000000 people, from which 170,000 are from India. This accounts for nearly 18% of suicide worldwide. There is a valuable reason to understand and study Indian suicide data. It is because of what outcomes and results we get as compared to data from other countries. These findings display the relationship for the causation of suicide and thus helps to find prevention strategies. Properly examining such data which includes differences in rates& patterns in gender, age group, population over some time may provide the key to understanding the suicidal purpose and therefore plan intervention aimed towards forestalling suicide.

1.6 Statement of the Problem

Suicide is an important issue in the Indian context. More than one lakh (one hundred thousand) lives are lost every year to suicide in our country. In the last two decades, the suicide rate has increased from 7.9 to 10.3 per 100,000. There is a wide variation in the suicide rates within the country. The southern states of Kerala, Karnataka, Andhra Pradesh and Tamil Nadu have a suicide rate of > 15 while in the Northern States of Punjab, Uttar Pradesh, Bihar and Jammu and Kashmir, the suicide rate is < 3. This variable pattern has been stable for the last twenty years. Higher literacy, a better reporting system, lower external aggression, higher

socioeconomic status and higher expectations are the possible explanations for the higher suicide rates in the southern states.

The majority of suicides (37.8%) in India are by those below the age of 30 years. The fact that 71% of suicides in India are by persons below the age of 44 years imposes a huge social, emotional and economic burden on our society. The near-equal suicide rates of young men and women and the consistently narrow male: female ratio of 1.4: 1 denotes that more Indian women die by suicide than their Western counterparts. Poisoning (36.6%), hanging (32.1%) and self-immolation (7.9%) were the common methods used to commit suicide. Two large epidemiological verbal autopsy studies in rural Tamil Nadu reveal that the annual suicide rate is six to nine times the official rate. If these figures are extrapolated, it suggests that there are at least half a million suicides in India every year. It is estimated that one in 60 persons in our country are affected by suicide. It includes both, those who have attempted suicide and those who have been affected by the suicide of a close family or friend. Thus, suicide is a major public and mental health problem, which demands urgent action.

1.7 Objectives of the Study

The researcher has set the following main objectives of the study.

1. To study the suicidal rates in India

2. Analyzing suicides by categorizing them under various categories and considering various constraints and parameters.

1.8 Hypothesis

H1: There is a significance difference between the suicide rate analysis using gender and age of people.

H2: There is a significance difference between the latest year and previous year suicide count

1.9 Research Methodology

In this statistical analysis, we used secondary data from the worlds largest data science community Kaggle. In which we took the data set of suicides happening in India in the past decade from 2001 to 2012. Analysis is performed by considering various parameters like States-wise suicides, considering the factor of Age and gender in increasing and decreasing the number

of suicides in India. Though reports for the years 201114 were available, detailed data in a comparable format were not available for disaggregated analysis as for 200112.

We applied a three-stage approach in which we analyzed the complete data set by breaking them down into different prospects, namely 1.Input, 2.Processing, and 3.output.As input we put together our complete data set and further for the next process we divided the suicides prospects into multiple formats, then for processing these formats we worked on defining the problem statement, data cleaning, data analysis and validation.. Further, we designed a visual representation to properly understand, compare useful insights and trends. In the end, we complied and integrate all of these data representations, comparisons, and trends to gain the desired output.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 237519 entries, 0 to 237518
Data columns (total 7 columns):
 #   Column      Non-Null Count    Dtype
---  ------      --------------    -----
 0   State       237519 non-null   object
 1   Year        237519 non-null   int64
 2   Type_code   237519 non-null   object
 3   Type        237519 non-null   object
 4   Gender      237519 non-null   object
 5   Age_group   237519 non-null   object
 6   Total       237519 non-null   int64
dtypes: int64(2), object(5)
memory usage: 12.7+ MB
```

Fig. 1.1 : suicide dataset

Data Source: NCRB

Rate of suicides in the past decade from the 2001- 2012 year:

According to the survey conducted by WHO, there are nearly 900,000 people worldwide committing suicides every year about 200,000 in China, 170,000 in Inia, and 140000 in high-income countries. Rates which is being recorded for the suicides were about 14.9 and 15.4

suicides per 100,000 population according to the survey conducted in the year 2001 and 2010 respectively World Suicide Prevention Day theme Suicide Prevention across the Globe: Strengthening Protective Factors and Instilling Hope is celebrated on September 12, 201. The NCRB (National Crime Records Bureau) reported 104720 suicides in the year 2001, increasing to 130653 in the year 2010, and so on.

1.9.1 Data Collection

Researcher has collected massive amount of relevant multi-structured data from various social media sites, product review sites, newspapers, online forms and government website as shown in table 1.1.

Data Sources		
Sr. No.	**Name of Site**	**Nature of Data**
1	https://www.kaggle.com/	Unstructured
2	https://ncrb.gov.in/en	
3	https://www.ncbi.nlm.nih.gov/	

Table 1.1: Sources of Data Collection

1.9.2 Research Work Stages

The research experiment is carried out in 3 different stages Literature Review, System Design and System Execution as shown in Figure 1.2 and elaborated further. The output of previous step is taken as the input for the next step. The secondary data collected in earlier in the previous section is used to get the exact result at every stage.

Stage-1: Literature Review

The researcher has carried out a literature review in stage 1 to identify the research gap between existing work and the present study.

- To understand the theoretical concept of the topic.

- To study the frameworks shared by researchers.

- To study the Methodology used by researchers.

- To identify the research gap between current research work and future research work

Researcher has identified, gathered and carried out detailed literature survey. Qualitative and quantitative information is obtained from various literature sources like national and international research journals, conference, thesis and books. Literature review is carried out in various sections like, State wise suicides, age and gender wise suicides, and year wish suicide rate. Each section is concluded with some conclusion at the end. Research gap identified which is included in chapter Number 2.
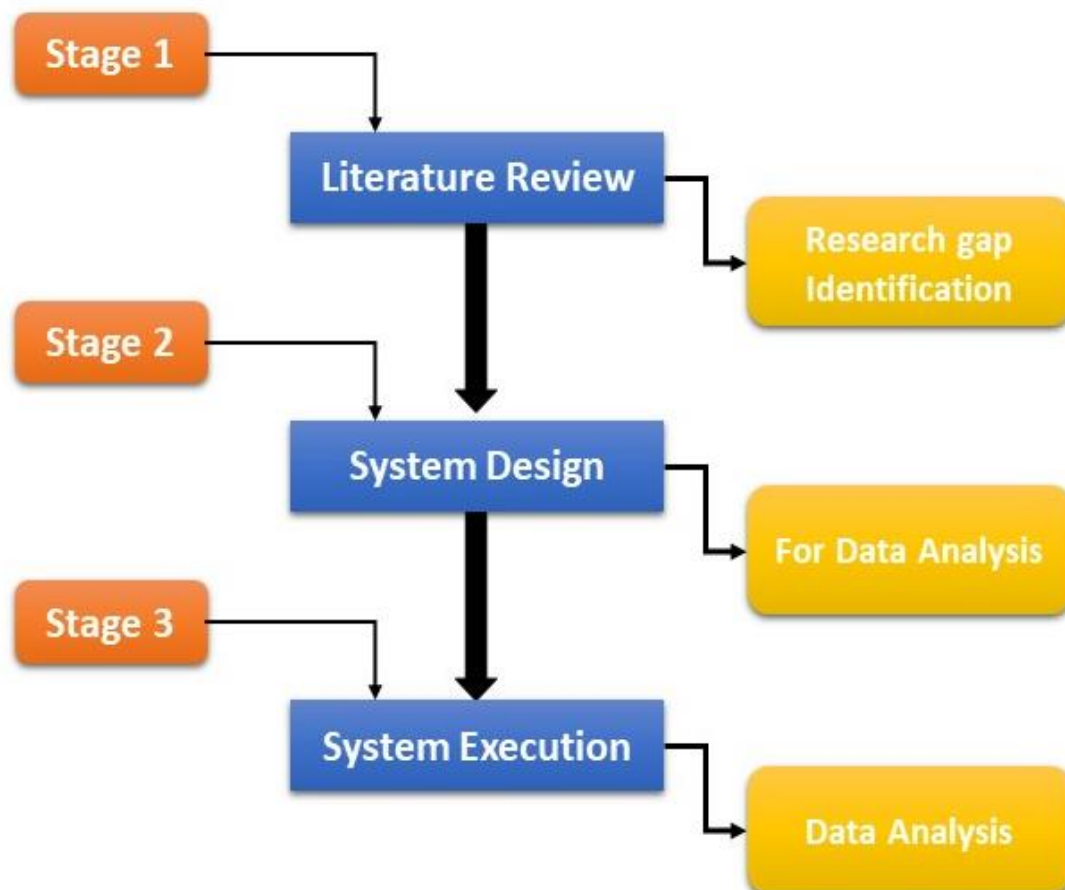
Figure 1.2: Different Stages of Research Work

Stage-2: System Design

To design the system we have carried out following steps elaborated.

1. Collected data from the official web sites of NCRB and other social networking sites as mentioned.
2. Data is pre-processed by removing noise and converted in CSV format.
3. Researcher has selected google colab and python framework. For data visualization we have selected PowerBI.
4. Researcher has proposed data analysis to analyse suicide rate in India.
5. Researcher has developed methodology to analyse states and year wise suicide rate using data set.

Figure 1.3: Steps of Data analysis

Stage-3: System Execution

In this stage we have executed system with large data sets. Collected data is stored in dataset and cleaned using google colab. Then we have performing some analysis on dataset. We have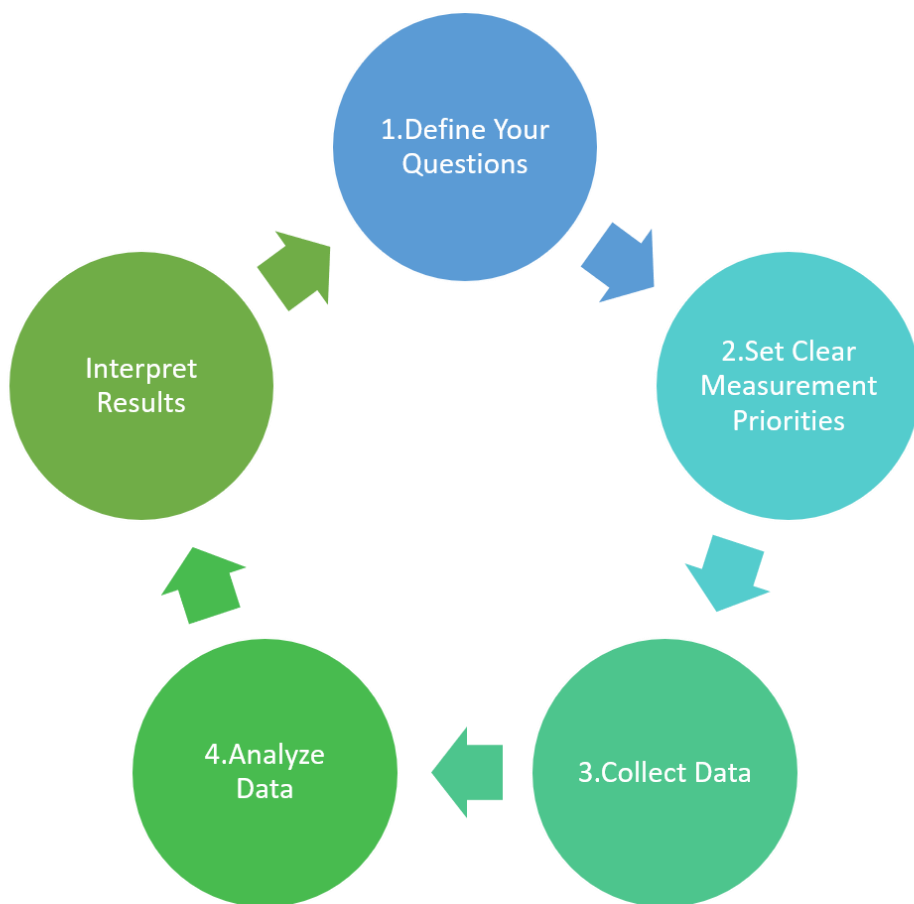 analysis the suicide data like causes, means adopted, education status, social profile, gender of the people who attempted suicides.

1.10   Limitations of the Study

Following are the limitations of the study.

- The present study is limited to predict suicide rate.
- The study is confined to analyze only multi structure data not images, videos etc.
- Only analysis are used to find the suicide rate.

1.11   Thesis Outline

The thesis entitled "**Empirical Exposition Model to Analyze Suicide Causes in India**" comprises of five chapters described as follows.

Chapter 1: Introduction

In this chapter we have given the overall introduction of topic and research methodology. This model of interpreting suicidality has great relevance in preventative approaches, since it gives the opportunity of intercepting suicidal trajectories at several different stages. However, this may not be the case for many situations, and the hypothesis of a continuum can be true only in a limited number of cases, probably embedded with a specific psychopathological scenario (e.g. depression) and with frequency that should not permit generalizations. This paper reviews the available evidence about the existence and validity of this construct, and discusses its practical implications.

Chapter 2: Review of Literature

This chapter elaborates the research already carried out in the related area and the future studies are indicated with the review of literature. The chapter provides detail review of research topic with related fields such as Suicide is the final outcome of complex interactions of biological, genetic, psychological, sociological and environmental factors. It is an increasingly important public health issue and the research gap between existing research and present research is identified in detail is mentioned in this chapter.

Chapter 3: Data Analytic Framework

This chapter gives theoretical aspects of the study area. It discusses about all methods and techniques used to carry out the research like data loading, data cleaning and data analysis. Related frameworks and architectures also described in detail in this chapter.

Chapter 4: Empirical Exposition Model

In this chapter we have included our main research execution. System design and analysis carried out using google colab python framework and PowerBI. In this chapter some experiments are carried out to select appropriate technologies and data analysis techniques.

Chapter 5: Result, Suggestion and Conclusion

This is concluding chapter which gives a synthesis of the whole study. Findings of the study and some constructive suggestions based on the study are discussed. Chapter also focuses the constraints, problems and prospects of developed system.

1.12 Chapter Summary

This chapter gives the broad introduction about the research carried out. In this chapter we discussed motivation of the research, Origin of the problem, significance and methodology used. Controlling the suicide rate is challenging problems faced by government of India. The present research provides organizations with greater opportunities by exposing suicide causes patterns from dataset. This information can be useful to control and reduce the rate of suicide. In our statistical analysis, we found various outcomes like, the rate at which suicides in India increasing is alarming as it shows considerable growth from the year 2001 to 2019. To come up with the suicide data set we find the analysis of the suicide rate in India.

1.13 References

1. Vijayakumar L. Suicide in India in Suicide in Asia. In: Yip PS, editor. Hong Kong Univ Press; 2008. pp. 121–31. [Google Scholar]
2. Accidental Deaths and suicides in India. National Crime Records Bureau. Ministry of home affairs. Government of India. 2007 [Google Scholar]

3. Rao VA. Attempted suicide. Indian J Psychiatry. 1965;7:253–64. [Google Scholar]

4. Nandi DN, Banarjee G, Boral GC. Suicide in West Bengal - A century apart. Indian J Psychiatry. 1978;20:155–60. [Google Scholar]

5. Hegde RS. Suicide in rural community. Indian J Psychiatry. 1980;22:368–70. [PMC free article] [PubMed] [Google Scholar]

6. Shukla GD, Verma BL, Mishra DN. Suicide in Jhansi city. Indian J Psychiatry. 1990;32:44–51. [PMC free article] [PubMed] [Google Scholar]

7. Banerjee G, Nandi DN, Nandi S, Sarkar S, Boral GC, Ghosh A. The vulnerability of Indian women to suicide. A field study. Indian J Psychiatry. 1990;32:305–8. [PMC free article] [PubMed] [Google Scholar]

8. Lal N, Sethi BB. Demographic and socio demographic variables in attempted suicide by poisoning. Indian J Psychiatry. 1975;17:100–7. [Google Scholar]

9. Badrinarayana A. Suicidal attempt in Gulbharga. Indian J Psychiatry. 1977;19:69–70. [Google Scholar]

10. Venkoba Rao A. Marriage, parenthood, sex and suicidal behavior. Indian J Psychiatry. 1974;16:92–4. [Google Scholar]

11. Nandi DN, Mukherjee SP, Banerjee G, Ghosh A, Boral CG, Chowdhury A, Bose J. Is suicide preventable by restricting the availability of lethal agents. A rural survey of West Bengal? Indian J Psychiatry. 1979;21:251–5. [Google Scholar]

12. Chowdhary AN, Brahma A, Banerjee S, Biswas MK. Pattern of domestic violence among non fatal deliberate self harm attempters: A study from primary care of West Bengal. Indian J Psychiatry. 2009;51:96–100. [PMC free article] [PubMed] [Google Scholar]

13. Singh AN. Discussion and evaluation of present suicide prevention activities and suggestions for future prevention activities. Indian J Psychiatry. 1972;14:369–74. [Google Scholar]

14. Rao VA. Toward suicide prevention. Indian J Psychiatry. 1999;41:280–8. [PMC free article] [PubMed] [Google Scholar]

15. Jena S, Sidhartha T. Non-fatal suicidal behavior in adolescents. Indian J Psychiatry. 2004;46:310–8. [PMC free article] [PubMed] [Google Scholar]

16. Vijayakumar L. Suicide and its prevention: The urgent need in India. Indian J Psychiatry. 2007;49:81–4. [PMC free article] [PubMed] [Google Scholar]

17. Gupta SC, Anand R, Trivedi JK. Development of a suicidal intent questionnaire. Indian J Psychiatry. 1983;25:57–62. [PMC free article] [PubMed] [Google Scholar]

2.1 Introduction

Suicide is the final outcome of complex interactions of biological, genetic, psychological, sociological and environmental factors. It is an increasingly important public health issue: from 1990 to 2010 the number of global suicides increased by 32%. It is particularly important among young adults 15 to 49 years of age among whom it accounts for 4.8% of all female deaths and 5.7% of all male deaths. Eighty-four percent of global suicides occur in low and middle-income countries (LMICs); India and China alone account for 49% of global suicides. There is substernal variability both in the prevalence of suicide and in the factors that influence the occurrence of suicide between geographic regions, cultures, and over time, so country-specific analyses are needed to develop targeted suicide prevention efforts. The World Health Organization (WHO) estimates that of the nearly 900,000 people who die from suicide globally every year, 170,000 are from India However, India's National Crime Records Bureau (NCRB) – which report official suicide rates based on police reports – estimated only 135,000 suicides in 2011. One possible reason for under-reporting of suicide by the NCRB is that suicide remains a crime in India. Detailed independent verbal autopsy investigations of all unnatural deaths in several rural areas of the country report suicide rates that are up to five-fold higher than the official national average.

2.2   Methods of Suicide

Hanging was the most frequently reported method of suicide in most of the studies, accounting for 10 to 72% of all suicides. The second most frequently reported method was self-poisoning, which accounted for 16 to 49% of all suicides. The proportion of all suicides attributed to drowning ranged from 3 to 39% and the proportion attributed to burning or self-immolation ranged from 6 to 57%. Other reported methods of suicide include jumping on heights (0.5 to 2% of all suicides), being run over by a train (6 to 13% of all suicides) and using a firearm (3% of all suicides). Some studies report gender-based differences in method preference. A community-based surveillance study by Prasad and colleagues reported in 2006 found that significantly more women chose drowning and burning as modes of suicide than

men, while significantly more men chose hanging. A similar pattern of gender-based method preference was reported by Abraham and colleagues among persons 55 years of age and older. Other studies report higher rates of suicides by hanging in males than females, a predominance of males in suicide decedents who use other violent methods (e.g., jumping in front of a train), and a predominance of females among suicides by self-immolation.

```
┌─────────────────────────────┐
│        Data Collect         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│         Data Load           │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│         Data Store          │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Data Process         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Data Analysis        │
└─────────────────────────────┘
```
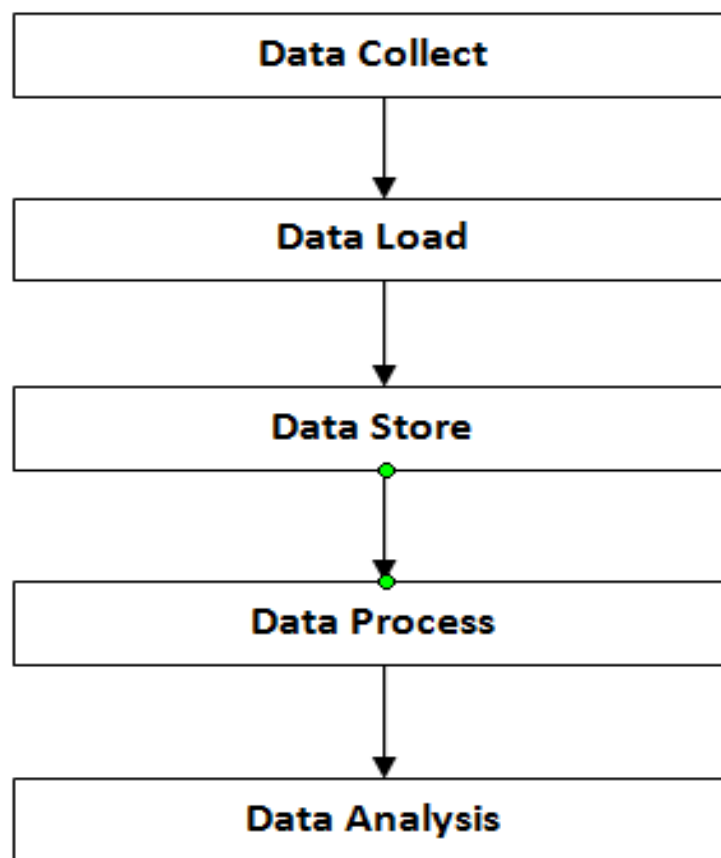
Figure 2.1: Framework for suicide rate analysis

10 Steps in Data Analysis

1. Define the question

2. Define the ideal data set

3. Obtain data

4. Clean the data

5. Exploratory Data Analysis

6. Statistical Prediction/modeling

7. Interpret results

8. Challenge results

9. Communicate results

10. Build a Data Product

2.3 RESEARCH GAP ANALYSIS

There have been several studies reporting the incidence of suicide in India. Over the years the studies have reported incidence rates ranging from 2.36 to 42 per 100,000 populations. The majority of these have been hospital based studies along with a few community based samples.

In one of the some article on attempted suicide published by the IJP, Venkoba Rao reported an incidence rate of 43 / 100,000 in Madurai. He also reported that 1 in 12 cases of suicide attempts were fatal. Studied incidence rates in Bengal using data available in the public domain across a hundred year period (1872-1972) and reported that the incidence of suicide had increased significantly from 2.36/100,00 in 1872 to 15.96 in 1972. The study also revealed that there was preponderance of male suicides,
the vulnerable age group being those between the ages of 18 to 30 and the most common method employed was poisoning. Suicide in a rural community in northern Karnataka reported an incidence rate of 9.3/100,000. The study also reported a male (67%) preponderance. The study also revealed that rural suicide patterns did not vary from urban.

In contrast to these reports in their study on the incidence of suicides in Jhansi city reported more suicides among women (34 / 100,000) than men (24 / 100,000). Several other gender related differences were also reported, women were significantly younger (24.6 years) compared to the men (28.9 years), self immolation was the most frequent method of suicide by women while for men it was being run over by a train. Domestic strife and mental illness were identified as the most common causative factors. The study reported an incidence rate of 29/100,000.

These findings were supported by who studied the vulnerability of Indian women. They found that the incidence of suicide was 43/100,000 in Bengal and that women (79.3%) outnumbered men. 75% of the victims were below 25 years of age and the commonest cause for suicide in women was quarrel with husband, while in men it was with parents. Ingestion of insecticide was the most common method of committing suicide.

2.4 Chapter Summary

In our statistical analysis, we found various outcomes lie, the rate at which suicides in India increasing is alarming as it shows considerable growth from the year 2001 to 2019. States having the highest number of suicides include Maharashtra, Tamil Nadu, and West Bengal which are almost highly populated states. From our findings, we reached the fact that men commit suicides almost twice as compared to women which can be evident by the fact that there are 2602974 women and 4868760 men committed suicides in the past decade. Also, in our analysis, we found that the people under the age group of 5-19 and 30-44 having cases 1534037 and 1471599 respectively are contributing the most to the number of people committing suicides in India. I consider the parameter of education level, then we can infer that people having the educational qualifications of Postgraduation and graduation are more likely to commit suicides as compared to others.

2.5 REFERENCES

1. Bagadia VN, Ghadiali HN, Shah LP. Unemployment and attempted suicide. Indian J Psychiatry. 1976;18:131–9. [Google Scholar]

2. Srivastava MK, Sahoo RN, Ghotekar LH, Dutta S, Danabalan M, Dutta TK, Das AK. Risk factors associated with attempted suicide. Indian J Psychiatry. 2004;46:33–8. [PMC free article] [PubMed] [Google Scholar]

3. Narang RL, Mishra BP, Mohan, Nitesh Attempted suicide in Ludhiana. Indian J Psychiatry. 2000;42:83–7. [PMC free article] [PubMed] [Google Scholar]

4. Bagadia VN, Abhyankar RR, Shroff P, Mehta P, Doshi J, Chawla P, et al. Suicidal behavior: A clinical study. Indian J Psychiatry. 1979;21:370–5. [Google Scholar]

5. Gupta SC, Singh H. Psychiatric illness in suicide attempters. Indian J Psychiatry. 1981;23:69–74. [PMC free article] [PubMed] [Google Scholar]

6. Mahla VP, Bhargava SC, Dogra R, Shome S. The psychology of self immolation in India. Indian J Psychiatry. 1992;34:108–13. [PMC free article] [PubMed] [Google Scholar]

7. Jain V, Singh H, Gupta SC, Kumar S. A study of hopelessness, suicidal intent and depression in cases of attempted suicide. Indian J Psychiatry. 1999;41:122–30. [PMC free article] [PubMed] [Google Scholar]

8. Khan FA, Anand B, Devi GM, Murthy K. Psychological autopsy of suicide: A cross sectional study. Indian J Psychiatry. 2005;47:73–8. [PMC free article] [PubMed] [Google Scholar]

9. Badrinarayana A. Study of suicidal risk factors in depressive illness. Indian J Psychiatry. 1980;22:81–3. [PMC free article] [PubMed] [Google Scholar]

10. Srivastava S, Kulsreshtha N. Expression of suicidal intent in depressives. Indian J Psychiatry. 2000;42:184–7. [PMC free article] [PubMed] [Google Scholar]

11. Anand R, Trivedi JK, Gupta SC. Suicidal communication in psychiatric patients. Indian J Psychiatry. 1983;25:121–8. [PMC free article] [PubMed] [Google Scholar]

12. Ponnudurai R, Jeyakar J, Saraswathy M. Attempted suicides in Madras. Indian J Psychiatry. 1986;28:59–62. [PMC free article] [PubMed] [Google Scholar]

13. Suresh Kumar PN. An analysis of suicide attempters versus completers in Kerala. Indian J Psychiatry. 2004;46:144–9. [PMC free article] [PubMed] [Google Scholar]

14. Devi PS, Rao VA. The premenstrual phase and suicidal attempts. Indian J Psychiatry. 1972;14:375–9. [Google Scholar]

15. Palaniappan V, Ramachandran V, Somasundaram O. Suicidal ideation and biogenic amines in depression. Indian J Psychiatry. 1983;25:286–92. [PMC free article] [PubMed] [Google Scholar]

16. Rao VA, Parvathi Devi S. Psychobiology of suicide behavior. Indian J Psychiatry. 1987;29:299–305. [PMC free article] [PubMed] [Google Scholar]

17. Rao VA. Attempted suicide and suicide among students in Madurai. Indian J Psychiatry. 1972;14:389–97. [Google Scholar]

18. Kumar S, Chandrasekharan R. As study of psychosocial and clinical factors associated with adolescent suicide attempts. Indian J Psychiatry. 2000;42:237–42. [PMC free article] [PubMed] [Google Scholar]

19. Sharma R, Vijay L, Chaturvedi S. Suicidal behavior amongst adolescent students in South Delhi. Indian J Psychiatry. 2008;50:30–3. [PMC free article] [PubMed] [Google Scholar]

20. Rao VA, Madhavan T. Depression and suicide behavior in the aged. Indian J Psychiatry. 1983;25:251–9. [PMC free article] [PubMed] [Google Scholar]

21. Rao VA, Mahendran N, Reddy GC, Prabhakar KT, Swaminathan ER, Andal BC, et al. One hundred female burns cases: A study in suicidology. Indian J Psychiatry. 1989;31:43–50. [PMC free article] [PubMed] [Google Scholar]

22. Jacob R, SureshKumar M, Rajkumar R, Palaniappun V. A study to assess depression, its correlates and suicidal behavior in epilepsy. Indian J Psychiatry. 2002;44:161–4. [PMC free article] [PubMed] [Google Scholar]

23. . Latha KS, Bhat SM. Suicidal behavior among terminally ill cancer patients. Indian J Psychiatry. 2005;45:79–83. [PMC free article] [PubMed] [Google Scholar]

24. Satyavati K. Attempted suicide in Psychiatric patients. Indian J Psyhiatry. 1971;13:37–48. [Google Scholar]

3.1   Introduction

In a world of data space where organizations deal with petabytes and exabytes of data, the era of Big Data emerged, the essence of its storage also grew. It was a great challenge and concern for industries for the storage of data until 2010. All those fancy Sci-fi movies you love to watch around can turn into reality by Data Science. Data Science is kinda blended with various tools, algorithms, and machine learning principles. Most simply, it involves obtaining meaningful information or insights from structured or unstructured data through a process of analyzing, programming and business skills. It is a field containing many elements like mathematics, statistics, computer science, etc. It's not an easy thing to do but not impossible too. You need to start from data, it's visualization, programming, formulation, development, and deployment of your model. In the future, there will be great hype for data scientist jobs. In this chapter we have given detailed introduction about big data technologies we used for Prediction and Analysis.

3.2   Python Programming Language

Python is a high-level, general-purpose and a very popular programming language. Python programming language (latest Python 3) is being used in web development, Machine Learning applications, along with all cutting edge technology in Software Industry. The essential things relevant to our research are listed below.

3.2.1   Python Eco-System

The Python ecosystem of libraries, frameworks, and tools is enormous and growing. Python is used for web scraping, data analysis, web development, internet of things development (IoT), machine learning, DevOps, general scientific computing, and many other computing and scripting uses. Python's standard library (plain old vanilla Python) is large, powerful, and utilitarian. For everything the standard library doesn't cover (or do well at), there are thousands of third-party modules and libraries. Awesome Python is a great place to get a sense for what these libraries/modules can be used for. Many of these libraries/frameworks are mature and have been battle-tested for 5-10+ years. This makes Python an attractive choice for startups and established enterprise companies alike.

The Python consists of following components :-

1. **Pandas:** Pandas (which stands for Python Data Analysis) is an open source data analysis library.
2. **Numpy:** NumPy's multidimensional array can perform very large calculations much more easily and efficiently than using the Python standard data types.
3. **Matplotlib:** Matplotlib is a Python 2D plotting library included within the SciPy ecosystem.
4. **Seaborn:** Seaborn is one of an amazing library for visualization of the graphical statistical plotting in Python.
5. **Django:** Django is an open source web application development framework.

3.2.2   Core Components of Python

There are many components of python that make it a high-level programming language. The most important components are as follows:

1) **Expression:** An Expression is a phrase of code that Python assesses to produce a value. The simplest expressions are literals and identifiers. You build different expressions by getting sub expressions together with the operators and delimiters.

2) **Statements**: It is a logical unit of code that can be executed by the Python interpreter. At the point when we type the statement at the command prompt, it will execute the code in the statement and give the result, as long as the code is error-free. In general, the programming code saved in the document contains scripts; these scripts are a combination of statements to perform units of work.

3) **Comments**: As programs get bigger and more complicated, they get harder to read. Formal languages are thick, and it is frequently hard to take a look at a piece of code and sort out what it is doing, or why.

4) **Functions**: Functions are a convenient strategy to separate your code into useful blocks, permitting us to arrange our code, make it more readable, reuse it, and save some time. Also, functions are a key way to characterize interfaces so programmers can share their code.

5) **Blocks and Indentation**: One of the most distinctive features of Python is its utilization of indentation to mark blocks of code.

3.3.3 Python Pandas

Pandas (which stands for Python Data Analysis) is an open source data analysis library. Pandas is very fast and <u>used by researchers</u> who have large amounts of data that they need to clean, analyze, and organize. Pandas provides flexible and efficient data structures like **Dataframe** and **Series** that make working with tabular data, excel files, and csv is far easier in Python. Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the **NumPy** library. Pandas is fast and it has high performance & productivity for users.

Series: <u>Pandas Series</u> is a one-dimensional labeled array capable of holding data of any type (integer, string, float, python objects, etc.).Pandas Series is nothing but a column in an excel sheet. In the real world, a Pandas Series will be created by loading the datasets from existing storage, storage can be SQL Database, CSV file, an Excel file.

DataFrame: <u>Pandas DataFrame</u> is a two-dimensional size-mutable, potentially heterogeneous tabular data structure with labeled axes (rows and columns). A Data frame is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Pandas DataFrame consists of three principal components, the data, rows, and columns.

Why Pandas is used for Data Science?

Pandas are generally used for data science but have you wondered why? This is because pandas are used in conjunction with other libraries that are used for data science. It is built on the top of the **NumPy** library which means that a lot of structures of NumPy are used or replicated in Pandas. The data produced by Pandas are often used as input for plotting functions of **Matplotlib**, statistical analysis in **SciPy**, and machine learning algorithms in **Scikit-learn**. Pandas program can be run from any text editor but it is recommended to use Jupyter Notebook for this as Jupyter given the ability to execute code in a particular cell rather than executing the entire file. Jupyter also provides an easy way to visualize pandas data frames and plots.

3.2.2.1 Python Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Array in Numpy is a

table of elements (usually numbers), all of the same type, indexed by a tuple of positive integers. In Numpy, number of dimensions of the array is called rank of the array. A tuple of integers giving the size of the array along each dimension is known as shape of the array. An array class in Numpy is called as **ndarray**. Elements in Numpy arrays are accessed by using square brackets and can be initialized by using nested Python Lists.

3.2.3   Matplotlib

Matplotlib is an amazing visualization library in Python for 2D plots of arrays. Matplotlib is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader SciPy stack. It was introduced by John Hunter in the year 2002.

One of the greatest benefits of visualization is that it allows us visual access to huge amounts of data in easily digestible visuals. Matplotlib consists of several plots like line, bar, scatter, histogram etc. Matplotlib comes with a wide variety of plots. Plots helps to understand trends, patterns, and to make correlations. They're typically instruments for reasoning about quantitative information.

Pyplot:

Pyplot is a Matplotlib module which provides a MATLAB-like interface. Matplotlib is designed to be as usable as MATLAB, with the ability to use Python and the advantage of being free and open-source. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc. The various plots we can utilize using Pyplot are **Line Plot**, **Histogram**, **Scatter**, **3D Plot**, **Image**, **Contour**, and **Polar**.

Syntax :

> *matplotlib.pyplot.plot(*args,   scalex=True,   scaley=True,   data=None, **kwargs)*

The plot function marks the x-coordinates and y-coordinates in a linear graph with specified scales.

Parameters: This function accepts parameters that enables us to set axes scales and format the graphs. These parameters are mentioned below :-

- **plot(x, y):** plot x and y using default line style and color.

- **plot.axis([xmin, xmax, ymin, ymax])**: scales the x-axis and y-axis from minimum to maximum values
- **plot.(x, y, color='green', marker='o', linestyle='dashed', linewidth=2, markersize=12):** x and y co-ordinates are marked using circular markers of size 12 and green color line with — style of width 2
- **plot.xlabel('X-axis')**: names x-axis
- **plot.ylabel('Y-axis')**: names y-axis
- **plot(x, y, label = 'Sample line ')** plotted Sample Line will be displayed as a legend.

3.2.4 Seaborn

Seaborn is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and color palettes to make statistical plots more attractive. It is built on the top of <u>matplotlib</u> library and also closely integrated to the data structures from <u>pandas</u>.

Seaborn aims to make visualization the central part of exploring and understanding data. It provides dataset-oriented APIs, so that we can switch between different visual representations for same variables for better understanding of dataset.

Different categories of plot in Seaborn:

Plots are basically used for visualizing the relationship between variables. Those variables can be either be completely numerical or a category like a group, class or division. Seaborn divides plot into the below categories :-

- **Relational plots:** This plot is used to understand the relation between two variables.
- **Categorical plots:** This plot deals with categorical variables and how they can be visualized.
- **Distribution plots:** This plot is used for examining univariate and bivariate distribution.
- **Regression plots:** The regression plots in seaborn are primarily intended to add a visual guide that helps to emphasize patterns in a dataset during exploratory data analyses.
- **Matrix plots:** A matrix plot is an array of scatterplots.
- **Multi-plot grids:** It is an useful approach is to draw multiple instances of the same plot on different subsets of the dataset.

Some basic plots using seaborn:

- Dist plot **:** Seaborn dist plot is used to plot a histogram, with some other variations like kdeplot and rugplot.

- Line plot **:** The line plot is one of the most basic plot in seaborn library. This plot is mainly used to visualize the data in form of some time series, i.e. in continuous manner.

- Lmplot **:** The lmplot is another most basic plot. It shows a line representing a linear regression model along with data points on the 2D-space and x and y can be set as the horizontal and vertical labels respectively.

3.3 Chapter Summary

For customer behavior analysis using reviews and comments automatic rating prediction is essential. The large scale data analytics framework elaborates theoretical concepts of various data science technologies for data analysis and prediction to discover of hidden behavioral patterns. The frame work focuses on the different steps required for developing predictive analytics and how python libraries like pandas, numpy, matplotlib and seaborn are useful to achieve this. This chapter also describes the importance of machine learning algorithms in behavior prediction.

3.4 References

https://www.geeksforgeeks.org

https://www.w3schools.com/datascience/

Matplotlib Tutorial – javatpoint

Empirical Exposition Model

4.1 Introduction

According to the survey conducted by NDTV on 2013, It was declared that for every hour there are 15 suicides getting registered across India. In fact, more than 371 suicides every day, the reason behind all this suicides are instability in thinking, lack of motivation. Among all the suicide cases, 5.5% of all suicide victims are students, on the other side farmers. Medical care claims that "person who attempts suicide shall be presumed to be suffering from mental illness at that time". In India,addressing this issue, a Mental Health care Bill got introduced in 2013 towards decriminalizing suicides. As per the records collected form 2001-2019, mostly students and farmers are vulnerable to suicides. Students committed suicide, due to failure in exams. The other side, the farmers. It is commonly know that the agriculture is the backbone of India. But, many are unaware of pledge of the Indian farmer. Maharashtra a state in India constitutes for half of the farmer suicides in India. The government has been battling farmer suicides over 15 years. The problems being faced by the farmers are of grave importance. Vidarbha is a village in Maharashtra, where dams are constructed for irrigation. But, the outlet have not been started. This is the fundamental problem for the former as they are dependent on rain for their crops to grow.

4.2 Types

There are five types available in the dataset:-

4.2.1 Causes

Type of Causes	Count
Causes Not known	3332
Other Causes (Please Specity)	3065
Family Problems	3013
Other Prolonged Illness	2680

Insanity/Mental Illness	2572
Love Affairs	1697
Poverty	1568
Death of Dear Person	1511
Fall in Social Reputation	1496
Suspected/Illicit Relation	1420
Drug Abuse/Addiction	1406
Unemployment	1404
Property Dispute	1393
Failure in Examination	1379
Bankruptcy or Sudden change in Economic	1373
Cancer	1233
Professional/Career Problem	1160
Cancellation/Non-Settlement of Marriage	1116
Paralysis	1105
Not having Children(Barrenness/Impotency	1039
Illness (Aids/STD)	1011
Dowry Dispute	871
Divorce	847
Physical Abuse (Rape/Incest Etc.)	731
Ideological Causes/Hero Worshipping	471
Illegitimate Pregnancy	420
Bankruptcy or Sudden change in Economic Status	126
Not having Children (Barrenness/Impotency	92

Table 4.1 : types of causes

Using the Table 4.1 we analyze and create a plot bar to show the maximum and minimum number of causes type that show in Figure 4.1. Which is used to identify the causes of suicide for what reason.

From Figure 4.1 and Table 4.1 can interpret that Family Problems, illness, love affairs and poverty of count in committing suicide is more, whereas Not having Children and bankruptcy committing suicide is less.
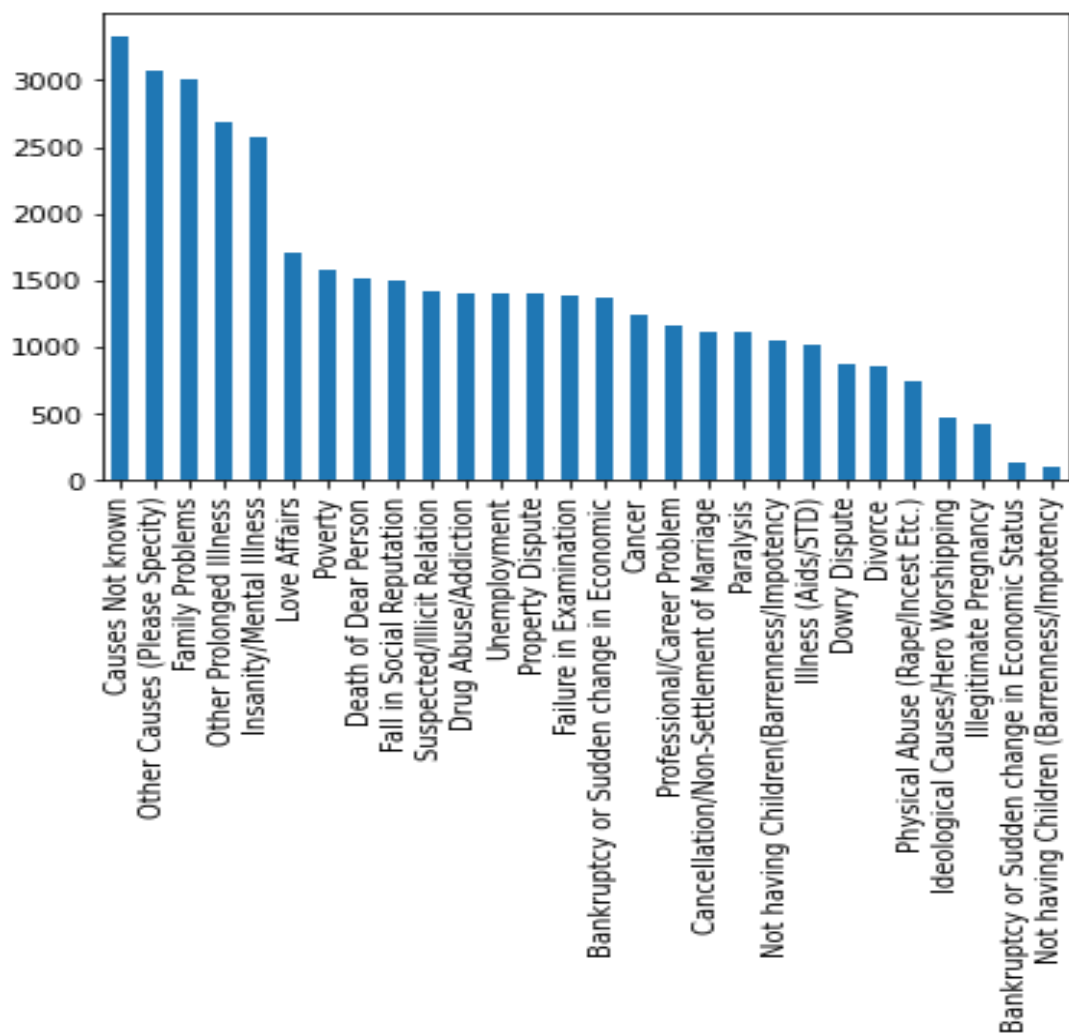

Fig 4.1 : Plot of type of Causes

4.2.2 Means adopted

Means Adopted	Count
By Hanging	3514
By Consuming Other Poison	2952
By Drowning	2609
By Consuming Insecticides	2608
By Fire/Self Immolation	2582
By Other means (please specify)	2415

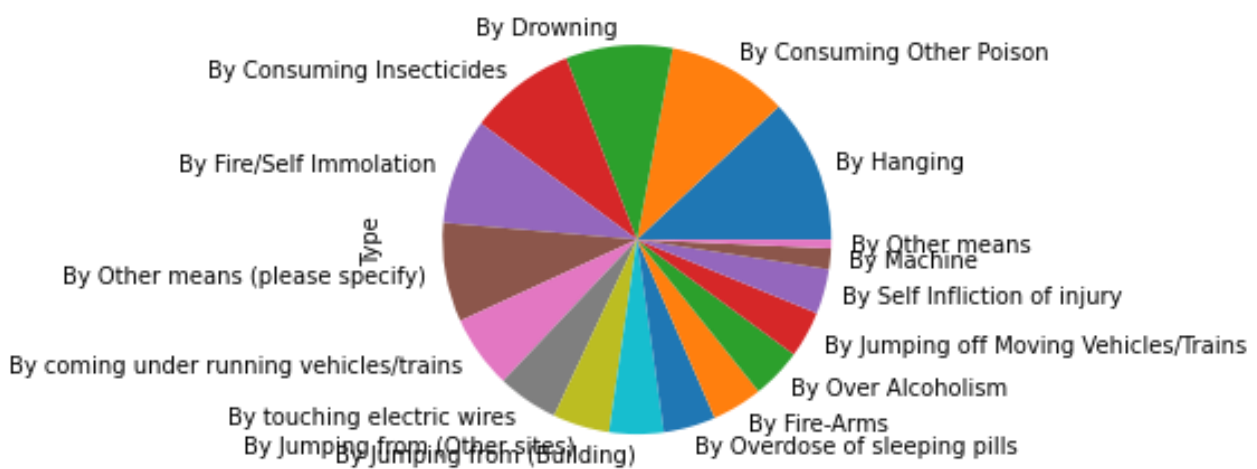|	
By coming under running vehicles/trains	1782
By touching electric wires	1483
By Jumping from (Other sites)	1384
By Jumping from (Building)	1325
By Overdose of sleeping pills	1272
By Fire-Arms	1236
By Over Alcoholism	1228
By Jumping off Moving Vehicles/Trains	1137
By Self Infliction of injury	1093
By Machine	504
By Other means	213

Table 4.2 : types of mean adopted

Fig 4.2 : Plot of types of mean adopted

From Figure 4.2 and Table 4.2 can interpret that by hanging and consuming poison of count in committing suicide is more, whereas by machine and self-infliction of injury committing suicide is less.

4.2.3   Professional Profile

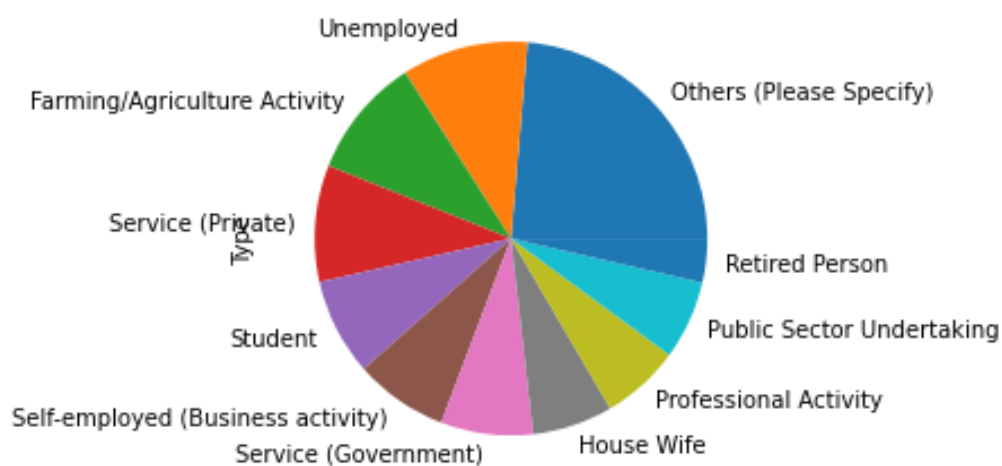| Professional Profile	Count
Others (Please Specify)	5509
Unemployed	2424
Farming/Agriculture Activity	2310
Service (Private)	2230
Student	1854
Self-employed (Business activity)	1793
Service (Government)	1771
House Wife	1543
Professional Activity	1513
Public Sector Undertaking	1506
Retired Person	838

Table 4.3: Type of Professional Profile

Fig 4.3 : Plot of types of professional Profile

From Figure 4.3 and Table 4.3 can interpret that unemployed and farming person of count in committing suicide is more, whereas retried person committing suicide is less.

4.2.4   Education Status

Education Status	Count
Middle	811
Primary	803
Matriculate/Secondary	803
No Education	773
Hr. Secondary/Intermediate/Pre-Universit	733
Graduate	643
Diploma	529
Post Graduate and Above	507

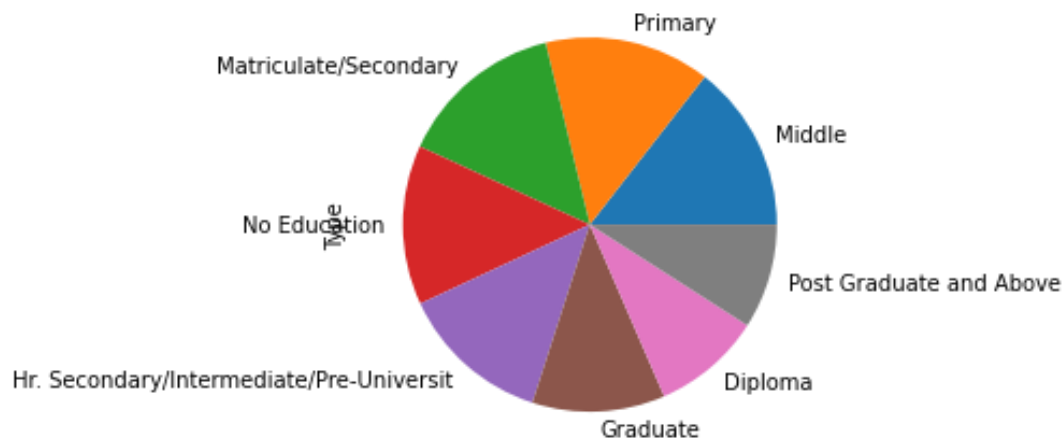**Table 4.4: Type of Education Status**

Fig 4.4 : Plot of types of education status

From Figure 4.3 and Table 4.3 can interpret that the Primary and middle educated person of count in committing suicide is more, whereas post graduated and above educated person committing suicide is less.

4.2.5 Social Status

Social Status	Count
Married	824
Never Married	813
Widowed/Widower	626
Separated	566
Divorcee	520

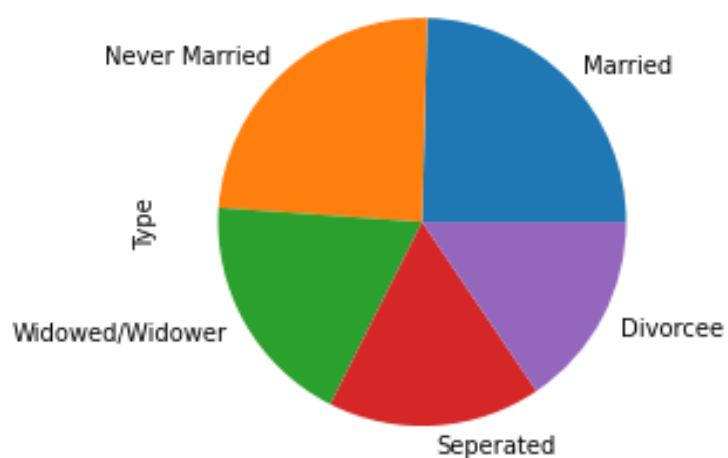**Table 4.5: Type of Social Status**

Fig 4.5 : Plot of Types of social status

From Figure 4.5 and Table 4.5 can interpret that the count of married and unmarried person in committing suicide is more, there has not lot of difference between them.
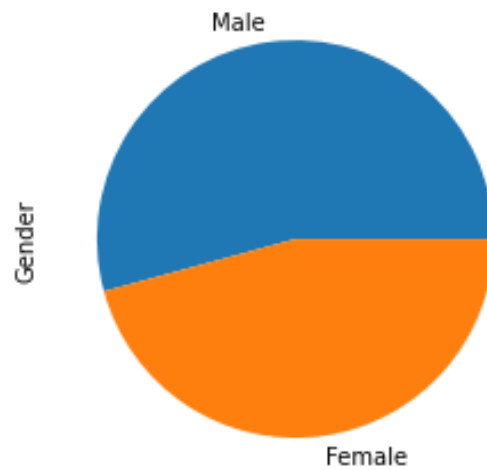
Fig. 4.6 : Male female Ratio.

From Figure 4.6, we can understand that, percentage of suicide is 55% in male, whereas in female, it is 45%. It indicates that, Male gender have more depression levels compared to female gender.
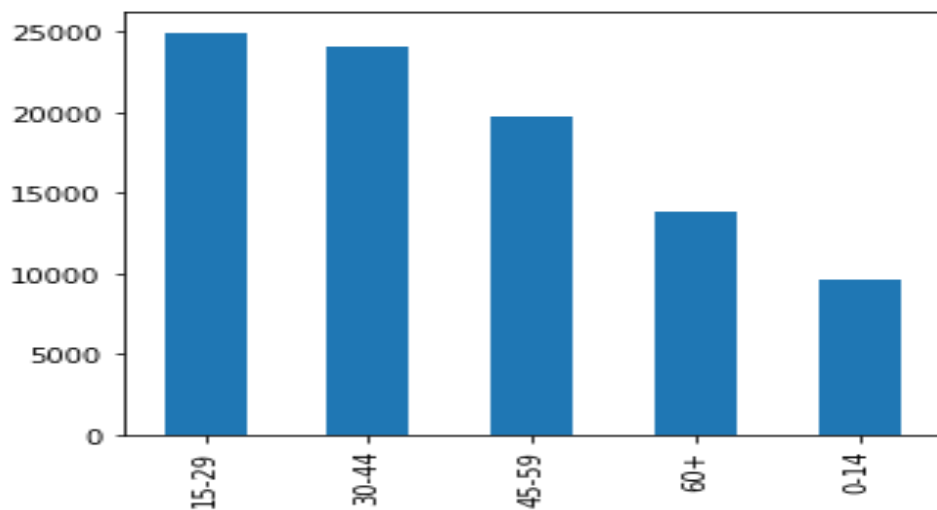

Fig. 4.7 : Age Group with Count.

From Figure 4.7 one can interpret that at 15-29 age group of count in committing suicide is more, whereas at the age group of 0-14 committing suicide is less.

5.1   Result

Understanding the representations of neuron with in death-related and suicide-related concepts. And measures the performance of classifier in determining the state of participant. In Figure 4.2, year wise total number of suicides based on different age group is plotted for better understanding. In Figure 5.2 the year wise plot (2001-2019) and causes for committing suicides are plotted. In which most of the suicides occur due to family problems and it is surprising to notice from 2006 it is gradually increasing and by 2012 it has reached to 35000 cases as per the dataset considered. In Figure 5.1, state wise total number of suicide cases registered are plotted. In which, Maharashtra, Tamil Nadu , Andhra Pradesh and West Bengal are top among the list.
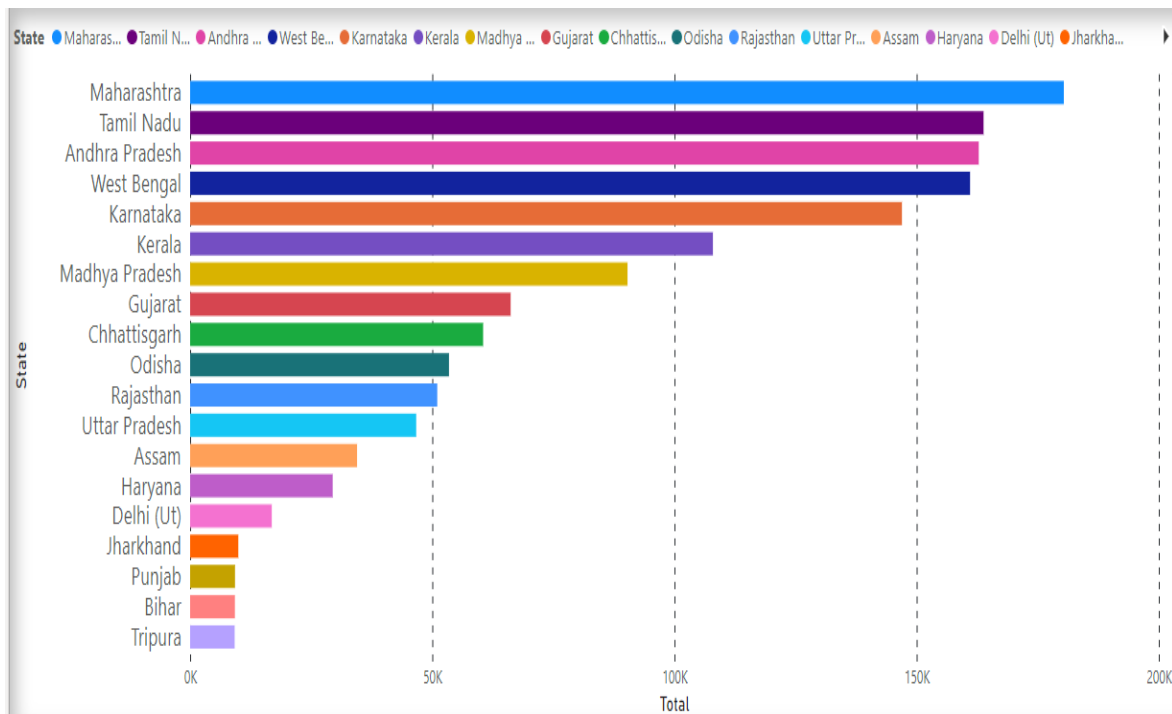
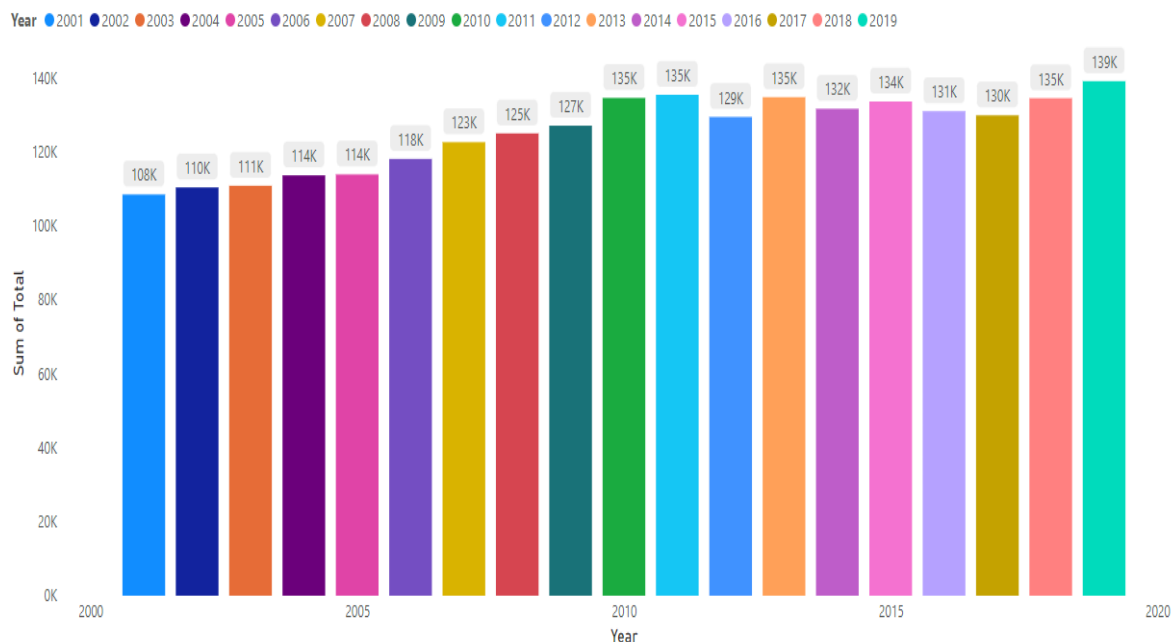**Fig 5.1 : State Wise Total Number of Suicides Year Wise**

Fig 5.2 : Total Number of Suicides Year Wise (2001-2019)

Result 1 : Family Problem, illness and love affair are the major causes for suicide in between 2001yr to 2019yr.

Result 2 : By hanging, consuming poison and drowning are the means adopted for the attempted suicide.

Result 3 : In the view of professional profiles there are unemployed , farmers and private job of people attempted maximum number of suicide.

Result 4 : The married and unmarried number of people ratio approximately same, their haven't major difference between them for suicide attempted.

Result 5 : The suicide ration between male and female are 55% of male where as the female ratio is 45%.

Result 6 : According to the age, the number of people between 15 to 29 are attempted suicide more and furthermore in between 30 to 44 age of people attempted suicide is second highest.

Result 7 : Maharashtra, west Bengal, Tamil Nadu and Andhra Pradesh this stats are highest ratio of the suicide whereas the Manipur , Nagaland and Daman & Diu are less suicide rate from India.

Result 8 : When we talk about years, the number of suicides continuously increasing from 2001 to 2019. The total number of death in this year are 2384218.

5.2   Suggestion

It needs to be a priority agenda not only for the health department (for accessible and quality support during a crisis), but also for the education department (to prevent suicide amongst school and college students), and the women and child development department (to reduce deaths by suicide amongst married women, which exceed 20,000 annually). Additionally, we need to change how we think and talk about suicide as a country, because that reflects in our systems and our society For instance, even though suicide in India today is decriminalized, the National Crime Records Bureau reports deaths by suicide as part of crime data, under the category of 'accidental death and suicide'. Instead, deaths by suicide should be under the health department, as is the case in so many countries worldwide. ,suicide prevention measures also need to be community-based rather than individual-focused. Most people who are vulnerable consciously or unconsciously share verbal or written hints in times of extreme distress. Thus, training and risk assessment programs can reduce attempts to take their own lives.

5.3   Conclusion

In our research, we found that there is a need to do much research on ways of controlling suicide by continuously improving the conceptualization of suicides thoughts and behaviors and by improvising etiological understanding of a individual problem. The students with age group 15-24 are committing suicides as their performance in academics is not that good. In addressing this problem parents needs to improve the level of skill in accepting a challenge and make them learn how to grow along with the problem. The other side, the farmers at the age group 29-45 are also committing suicide due to their commitments made with the family. The facilities provided by the government of India towards irrigation needs to be outlet successfully. In future, with this understanding on suicides, we would like to work on applying machine learning algorithms in evaluating the impact factor of each risk factor and correlation exists among them. This will help in providing the recommendation both to the farmer and the students at different levels of their life.

Bibliography

Research Papers

1. Vijayakumar L. Suicide in India in Suicide in Asia. In: Yip PS, editor. Hong Kong Univ Press; 2008. pp. 121–31. [Google Scholar]

2. Accidental Deaths and suicides in India. National Crime Records Bureau. Ministry of home affairs. Government of India. 2007 [Google Scholar]

3. Rao VA. Attempted suicide. Indian J Psychiatry. 1965;7:253–64. [Google Scholar]

4. Nandi DN, Banarjee G, Boral GC. Suicide in West Bengal - A century apart. Indian J Psychiatry. 1978;20:155–60. [Google Scholar]

5. Hegde RS. Suicide in rural community. Indian J Psychiatry. 1980;22:368–70. [PMC free article] [PubMed] [Google Scholar]

6. Shukla GD, Verma BL, Mishra DN. Suicide in Jhansi city. Indian J Psychiatry. 1990;32:44–51. [PMC free article] [PubMed] [Google Scholar]

7. Banerjee G, Nandi DN, Nandi S, Sarkar S, Boral GC, Ghosh A. The vulnerability of Indian women to suicide. A field study. Indian J Psychiatry. 1990;32:305–8. [PMC free article] [PubMed] [Google Scholar]

8. Lal N, Sethi BB. Demographic and socio demographic variables in attempted suicide by poisoning. Indian J Psychiatry. 1975;17:100–7. [Google Scholar]

9. Badrinarayana A. Suicidal attempt in Gulbharga. Indian J Psychiatry. 1977;19:69–70. [Google Scholar]

10. Venkoba Rao A. Marriage, parenthood, sex and suicidal behavior. Indian J Psychiatry. 1974;16:92–4. [Google Scholar]

11. Nandi DN, Mukherjee SP, Banerjee G, Ghosh A, Boral CG, Chowdhury A, Bose J. Is suicide preventable by restricting the availability of lethal agents. A rural survey of West Bengal? Indian J Psychiatry. 1979;21:251–5. [Google Scholar]

12. Bagadia VN, Ghadiali HN, Shah LP. Unemployment and attempted suicide. Indian J Psychiatry. 1976;18:131–9. [Google Scholar]

13. Srivastava MK, Sahoo RN, Ghotekar LH, Dutta S, Danabalan M, Dutta TK, Das AK. Risk factors associated with attempted suicide. Indian J Psychiatry. 2004;46:33–8. [PMC free article] [PubMed] [Google Scholar]

14. Narang RL, Mishra BP, Mohan, Nitesh Attempted suicide in Ludhiana. Indian J Psychiatry. 2000;42:83–7. [PMC free article] [PubMed] [Google Scholar]

15. Bagadia VN, Abhyankar RR, Shroff P, Mehta P, Doshi J, Chawla P, et al. Suicidal behavior: A clinical study. Indian J Psychiatry. 1979;21:370–5. [Google Scholar]

16. Gupta SC, Singh H. Psychiatric illness in suicide attempters. Indian J Psychiatry. 1981;23:69–74. [PMC free article] [PubMed] [Google Scholar]

17. Mahla VP, Bhargava SC, Dogra R, Shome S. The psychology of self immolation in India. Indian J Psychiatry. 1992;34:108–13. [PMC free article] [PubMed] [Google Scholar]

18. Jain V, Singh H, Gupta SC, Kumar S. A study of hopelessness, suicidal intent and depression in cases of attempted suicide. Indian J Psychiatry. 1999;41:122–30. [PMC free article] [PubMed] [Google Scholar]

19. Khan FA, Anand B, Devi GM, Murthy K. Psychological autopsy of suicide: A cross sectional study. Indian J Psychiatry. 2005;47:73–8. [PMC free article] [PubMed] [Google Scholar]

20. Badrinarayana A. Study of suicidal risk factors in depressive illness. Indian J Psychiatry. 1980;22:81–3. [PMC free article] [PubMed] [Google Scholar]

21. Srivastava S, Kulsreshtha N. Expression of suicidal intent in depressives. Indian J Psychiatry. 2000;42:184–7. [PMC free article] [PubMed] [Google Scholar]

22. Anand R, Trivedi JK, Gupta SC. Suicidal communication in psychiatric patients. Indian J Psychiatry. 1983;25:121–8. [PMC free article] [PubMed] [Google Scholar]

23. Ponnudurai R, Jeyakar J, Saraswathy M. Attempted suicides in Madras. Indian J Psychiatry. 1986;28:59–62. [PMC free article] [PubMed] [Google Scholar]

24. Suresh Kumar PN. An analysis of suicide attempters versus completers in Kerala. Indian J Psychiatry. 2004;46:144–9. [PMC free article] [PubMed] [Google Scholar]

25. Devi PS, Rao VA. The premenstrual phase and suicidal attempts. Indian J Psychiatry. 1972;14:375–9. [Google Scholar]

26. Palaniappan V, Ramachandran V, Somasundaram O. Suicidal ideation and biogenic amines in depression. Indian J Psychiatry. 1983;25:286–92. [PMC free article] [PubMed] [Google Scholar]

27. Rao VA, Parvathi Devi S. Psychobiology of suicide behavior. Indian J Psychiatry. 1987;29:299–305. [PMC free article] [PubMed] [Google Scholar]

28. Rao VA. Attempted suicide and suicide among students in Madurai. Indian J Psychiatry. 1972;14:389–97. [Google Scholar]

29. Kumar S, Chandrasekharan R. As study of psychosocial and clinical factors associated with adolescent suicide attempts. Indian J Psychiatry. 2000;42:237–42. [PMC free article] [PubMed] [Google Scholar]

30. Sharma R, Vijay L, Chaturvedi S. Suicidal behavior amongst adolescent students in South Delhi. Indian J Psychiatry. 2008;50:30–3. [PMC free article] [PubMed] [Google Scholar]

31. Rao VA, Madhavan T. Depression and suicide behavior in the aged. Indian J Psychiatry. 1983;25:251–9. [PMC free article] [PubMed] [Google Scholar]

32. Rao VA, Mahendran N, Reddy GC, Prabhakar KT, Swaminathan ER, Andal BC, et al. One hundred female burns cases: A study in suicidology. Indian J Psychiatry. 1989;31:43–50. [PMC free article] [PubMed] [Google Scholar]

33. Jacob R, SureshKumar M, Rajkumar R, Palaniappun V. A study to assess depression, its correlates and suicidal behavior in epilepsy. Indian J Psychiatry. 2002;44:161–4. [PMC free article] [PubMed] [Google Scholar]

34. Latha KS, Bhat SM. Suicidal behavior among terminally ill cancer patients. Indian J Psychiatry. 2005;45:79–83. [PMC free article] [PubMed] [Google Scholar]

35. Satyavati K. Attempted suicide in Psychiatric patients. Indian J Psychiatry. 1971;13:37–48. [Google Scholar]

36. Gupta SC, Singh H, Trivedi JK. Evaluation of suicidal risk in depressives and schizophrenics: A two year follow up study. Indian J Psychiatry. 1992;34:298–310. [PMC free article] [PubMed] [Google Scholar]

37. Srivastava AS, Kumar R. Suicidal ideation and attempts in patients with major depression: Socio demographic and clinical variables. Indian J Psychiatry. 2005;47:225–8. [PMC free article] [PubMed] [Google Scholar]

38. Goel Capt DS. Attempted suicide in the armed forces. Indian J Psychiatry. 1975;17:163–73. [Google Scholar]

39. Chakraborthy PK. The significance of attempted suicide in armed forces. Indian J Psychiatry. 2002;44:277–82. [PMC free article] [PubMed] [Google Scholar]

40. Chavan BS, Singh GP, Kaur J. Psychological autopsy of 101 suicide cases from northwest region of India. Indian J Psychiatry. 2008;50:34–8. [PMC free article] [PubMed] [Google Scholar]

41. Behere PB, Behere AP. Farmers' suicide in Vidarbha region in Maharashtra State: A myth or reality? Indian J Psychiatry. 2008;50:124–7. [PMC free article] [PubMed] [Google Scholar]

42. Sethi BB, Gupta SC, Singh H. Psychosocial factors and personality characteristics in cases of attempted suicide. Indian J Psychiatry. 1978;20:25–30. [Google Scholar]

43. Sarkar P, Sattar FA, Gode N, Basanar DR. Failed suicide and deliberate self harm: A need for specific nomenclature. Indian J Psychiatry. 2006;48:78–83. [PMC free article] [PubMed] [Google Scholar]

44. Das PP, Grover S, Avasthi A, Chakrabarthi S, Malhotra S, Kumar S. Intentional self harm seen in psychiatric referrals in a tertiary care hospital. Indian J Psychiatry. 2008;50:187–91. [PMC free article] [PubMed] [Google Scholar]

45. Chowdhary AN, Banerjee S, Brahma A, Biswas MK. Pesticide poisoning in non fatal deliberate self harm: A public health issue. Indian J Psychiatry. 2007;49:117–20. [PMC free article] [PubMed] [Google Scholar]

Sample Data Set

Content :-

Time Period: 2001 – 2019

Granularity: Yearly
Location: States of India

Parameters:

a) Suicide causes
b) Education status
c) By means adopted
d) Professional profile
e) Social status

Execution

path="/content/drive/MyDrive/csv files/Suicides in India 2001-2012.csv"

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df=pd.read_csv(path)

df.head()

Number of Rows : 237519
Number of Columns : 07

print("No. of row in dataframe is   :",df.shape[0])

print("No. of column in dataframe is :",df.shape[1])

No. of row in dataframe is   : 237519

No. of column in dataframe is : 7

Information about Data se

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 237519 entries, 0 to 237518
Data columns (total 7 columns):

```
 #  Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0  State      237519 non-null  object
 1  Year       237519 non-null  int64
 2  Type_code  237519 non-null  object
 3  Type       237519 non-null  object
 4  Gender     237519 non-null  object
 5  Age_group  237519 non-null  object
 6  Total      237519 non-null  int64
dtypes: int64(2), object(5)
memory usage: 12.7+ MB
```

Check values whether is null or not

df.isna().sum()

```
State      0
Year       0
Type_code  0
Type       0
Gender     0
Age_group  0
Total      0
dtype: int64
```

copy of dataset

df1=df.copy()

Check State column and deleting unwanted data

df=df1.copy()

df["State"].value_counts().count()

```
38
df.drop(df.index[df["State"]=="Total (All India)"], axis=0, inplace=True)
df.drop(df.index[df["State"]=="Total (States)"], axis=0, inplace=True)
df.drop(df.index[df["State"]=="Total (Uts)"],axis=0, inplace=True)
df["State"].value_counts().count()
35
```

Years start from 2001 to 2012

print(df['Year'].min())

print("to")

print(df['Year'].max())

```
2001
to
2012
```

```python
df["Year"].unique()
```
```
array([2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011,
       2012])
```
```python
df["Total"].value_counts()
```
```
0       135473
1        16038
2         9939
3         6697
4         5110
         ...
1774         1
1332         1
732          1
1548         1
2658         1
Name: Total, Length: 1764, dtype: int64
```

Removing zero vaues from dataset

```python
df.drop(df[df["Total"]==0].index, inplace=True)
```

```python
df.head()
```

```python
df['Type'].unique()
```
```
array(['Love Affairs', 'Other Causes (Please Specify)',
       'Other Prolonged Illness', 'Failure in Examination',
       'Causes Not known', 'Family Problems', 'Insanity/Mental Illness',
       'Death of Dear Person', 'No Education', 'Middle', 'Primary',
       'Matriculate/Secondary', 'Graduate',
       'Hr. Secondary/Intermediate/Pre-Universit', 'By Hanging',
       'By Fire/Self Immolation', 'By Consuming Other Poison',
       'By Drowning', 'Student', 'Others (Please Specify)', 'House Wife',
       'Service (Private)', 'Public Sector Undertaking',
       'Service (Government)', 'Farming/Agriculture Activity',
       'Retired Person', 'Married', 'Never Married', 'Unemployment',
       'Post Graduate and Above', 'Self-employed (Business activity)',
       'Fall in Social Reputation', 'Suspected/Illicit Relation',
       'Cancellation/Non-Settlement of Marriage',
       'Not having Children(Barrenness/Impotency', 'Poverty',
       'By Consuming Insecticides', 'By Other means (please specify)',
       'Unemployed', 'Professional/Career Problem', 'Paralysis',
       'Bankruptcy or Sudden change in Economic', 'Divorce', 'Cancer',
       'Property Dispute', 'By touching electric wires',
       'Professional Activity', 'Divorcee', 'Widowed/Widower',
       'Seperated', 'Diploma', 'By Fire-Arms',
       'By Overdose of sleeping pills', 'Illness (Aids/STD)',
       'Physical Abuse (Rape/Incest Etc.)', 'Drug Abuse/Addiction',
       'Ideological Causes/Hero Worshipping', 'Dowry Dispute',
       'Illegitimate Pregnancy',
```

'By coming under running vehicles/trains',
'By Jumping off Moving Vehicles/Trains',
'By Jumping from (Other sites)', 'By Over Alcoholism',
'By Jumping from (Building)', 'By Self Infliction of injury',
'By Machine', 'Bankruptcy or Sudden change in Economic Status',
'Not having Children (Barrenness/Impotency', 'By Other means'],
dtype=object)

After deleting unwanted data :
Rows is : 101110
Columns is : 07

df.shape

df.to_csv("df")

State wise suicide in INDIA

df["State"].unique()

array(['A & N Islands', 'Andhra Pradesh', 'Arunachal Pradesh', 'Assam',
    'Bihar', 'Chandigarh', 'Chhattisgarh', 'D & N Haveli',
    'Daman & Diu', 'Delhi (Ut)', 'Goa', 'Gujarat', 'Haryana',
    'Himachal Pradesh', 'Jammu & Kashmir', 'Jharkhand', 'Karnataka',
    'Kerala', 'Lakshadweep', 'Madhya Pradesh', 'Maharashtra',
    'Manipur', 'Meghalaya', 'Mizoram', 'Nagaland', 'Odisha',
    'Puducherry', 'Punjab', 'Rajasthan', 'Sikkim', 'Tamil Nadu',
    'Tripura', 'Uttar Pradesh', 'Uttarakhand', 'West Bengal'],
    dtype=object)
df.info()
<class 'pandas.core.frame.DataFrame'>
Int64Index: 101110 entries, 13 to 237518
Data columns (total 7 columns):
 # Column   Non-Null Count   Dtype
--- ------   --------------   -----
 0 State   101110 non-null object
 1 Year   101110 non-null int64
 2 Type_code 101110 non-null object
 3 Type   101110 non-null object
 4 Gender   101110 non-null object
 5 Age_group 101110 non-null object
 6 Total   101110 non-null int64
dtypes: int64(2), object(5)
memory usage: 6.2+ MB

Type_Code

 ☐ Causes
 ☐ Means_adopted

☐ Professional_Profile
☐ Education_Status
☐ Social_Status

```python
df["Type_code"].value_counts()
```

```
Causes                39531
Means_adopted         29337
Professional_Profile  23291
Education_Status       5602
Social_Status          3349
Name: Type_code, dtype: int64
```

1. Causes

```python
causes= df[df["Type_code"]=="Causes"]
causes.head()
```

```python
CState=causes.groupby("State")["Total"].sum()
```

```python
CState.sort_values(ascending=False).plot.bar()
```

Total number of Suicides between 2001 to 2012 by Causes : 1440974

```python
causesDeath=causes['Total'].sum()
causesDeath
```

1440974

Type of Causes

```python
causes["Type"].value_counts()
```

Cause	Count
Causes Not known	3332
Other Causes (Please Specity)	3065
Family Problems	3013
Other Prolonged Illness	2680
Insanity/Mental Illness	2572
Love Affairs	1697
Poverty	1568
Death of Dear Person	1511
Fall in Social Reputation	1496
Suspected/Illicit Relation	1420
Drug Abuse/Addiction	1406
Unemployment	1404
Property Dispute	1393
Failure in Examination	1379
Bankruptcy or Sudden change in Economic	1373
Cancer	1233
Professional/Career Problem	1160
Cancellation/Non-Settlement of Marriage	1116

```
Paralysis                                           1105
Not having Children(Barrenness/Impotency            1039
Illness (Aids/STD)                                  1011
Dowry Dispute                                        871
Divorce                                             847
Physical Abuse (Rape/Incest Etc.)                   731
Ideological Causes/Hero Worshipping                 471
Illegitimate Pregnancy                              420
Bankruptcy or Sudden change in Economic Status      126
Not having Children (Barrenness/Impotency           92
Name: Type, dtype: int64
```

causes["Type"].value_counts().plot.bar()

Yearwise suicide cases

```
c1=causes.groupby("Year")["Total"].sum()
c1.to_csv('c1')
c1=pd.read_csv('c1')
c1.columns=['Year','Total']
```

2. Means_adopted

```
Means_adopted= df[df["Type_code"]=="Means_adopted"]
Means_adopted.head()
MState=Means_adopted.groupby("State")["Total"].sum()
MState.sort_values(ascending=False).plot.bar()
Total number of Suicides between 2001 to 2012 by Means_adopted : 1455931
adoptedDeath=Means_adopted["Total"].sum()
adoptedDeath
1455931
```
Means_adopted["Type"].value_counts()
```
By Hanging                               3514
By Consuming Other Poison                2952
By Drowning                              2609
By Consuming Insecticides                2608
By Fire/Self Immolation                  2582
By Other means (please specify)          2415
By coming under running vehicles/trains  1782
By touching electric wires               1483
By Jumping from (Other sites)            1384
By Jumping from (Building)               1325
By Overdose of sleeping pills            1272
By Fire-Arms                             1236
By Over Alcoholism                       1228
By Jumping off Moving Vehicles/Trains    1137
By Self Infliction of injury             1093
```

By Machine   504
By Other means   213
Name: Type, dtype: int64

Type of Means_adopted

Means_adopted["Type"].value_counts().plot.pie()

Yearwise suicide cases

```
m1=Means_adopted.groupby("Year")["Total"].sum()
m1.to_csv('m1')
m1=pd.read_csv('m1')
m1.columns=['Year','Total']
```

Thank You…