Devi Ahilya Vishwavidhyalaya, Indore

School of Statistics and School of Data Science and Data Forecasting

Session: 2021- 2025

A project Report submitted for the partial fulfilment of the Degree of Bachelor of Science (Hons)

Statistical Analysis of Suicides in India

Submitted to:

Dr. Rashmi Awad,

Assistant Professor

School of Statistics

DAVV, Indore

Submitted by: -

Shraddha Singh

B.Sc. (Hons) Applied Statistics & Analytics

Forward

I feel immense pleasure to forward the project entitled, **"Statistical Analysis of Suicides in India"** submitted by **Shraddha Singh** of the School of Statistics, DAVV, Indore enrolled in B.Sc. (Hons) Applied Statistics and Analytics (V Semester).

Date: -

Dr. V. B. Gupta
Head of Department
School Of Statistics
DAVV, Indore.

Certificate

The research work embodied in the project work entitled "**Statistical Analysis of Suicides in India**" has been carried out under the supervision of Dr. Rashmi Awad. The work reported herein is original.

Shraddha Singh
B.Sc. (Hons.) Applied Statistics
And Analytics

Dr. Rashmi Awad
Asst. Professor
School of Statistics
DAVV, Indore

Declaration

I hereby declare that this project entitled **"Statistical Analysis of Suicides in India**" has been prepared by me, under the supervision of Dr. Rashmi Awad, Asst. Professor, School of Statistics, DAVV, Indore for the partial fulfilment of the B.Sc. (Hons.) degree in Applied Statistics and Analytics.

Shraddha Singh

Date:

Place: Indore

Acknowledgement

I would like to thank my Project Supervisor and Asst. Professor Dr. Rashmi Awad for her continuous guidance and encouragement throughout the project.

Throughout the project phase, she was very kind, patient, accommodating and very helpful. Her valuable feedback on my work kept the project going in the right direction and at the right pace.

In addition, I would like to thank my family and friends for their constant support and patience throughout the phases of this research.

Table of Content

Abstract

Background: According to the reports published by the National Crime Records Bureau (NCRB) India, every year, the number of suicides is continuously increasing. The number of suicidal deaths increased by 26.29% from 2017 to 2021.

Objectives: To briefly overview the recent trends in the number of suicides and study risk factors for suicides, in India. We explore the association of suicidal deaths and some attributes such as gender (sex), and age group of suicide victims.

Method: Correspondence analysis (CA) and subset CA are used to study the association of attributes.

Results: In India, family problems and health problems are major reasons for suicides. The association between the risk factors for suicides, gender, and age group of suicide victims is studied and explored using biplots.

Conclusions: The leading risk factors for suicides are family problems, illness, drug addiction, failure in examination, etc. As per the association of attributes studied, the government of India must launch 'Anti-suicide campaigns' at all levels regularly and the campaign should consider gender, and age groups while structuring the campaign.

Keywords: Suicides, Risk factors for suicides, Odds ratio, CA, Biplots

Introduction

Suicide (Latin suicidium, from sui caedere, 'to kill oneself') is a leading cause of death among teenagers and adults under 35 years of age, ranked among the top 13 causes of death for individuals of all ages worldwide by the World Health Organization (WHO) and the National Safety Council rates it sixth in the United States. It is believed that the most dramatic increase in suicide mortality will be observed in third-world countries because of socioeconomic and behavioural factors.

Suicide has been around for as long as human society, ranking among the top 13 causes of death in all ages worldwide, and continues to challenge our collective wisdom.

A total of 1,64,033 suicides were reported in the country during 2021 showing an increase of 7.2% in comparison to 2020 and the rate of suicides has increased by 6.2% during 2021 over 2020. Most suicides were reported in Maharashtra (22,207) followed by 18,925 suicides in Tami Nadu, 14,965 suicides in Madhya Pradesh, 13,500 suicides in West Bengal, and 13,056 suicides in Karnataka accounting for 13.5%, 11.5%, 9.1%, 8.2% and 8.0% of total suicides respectively. These 5 States together accounted for 50.4% of the total suicides reported in the country. The remaining 49.6% of suicides were reported in the remaining 23 States and 8 UTs. Uttar Pradesh, the most populous State (16.9% share of the country's population) has reported a comparatively lower percentage share of suicidal deaths, accounting for only 3.6% of the total suicides reported in the country.

There is no single reason why someone may try to take their own life, but certain factors can increase the risk, such as illness, family problems, financial loss, and harmful use of alcohol, act cumulatively to increase a person's vulnerability to suicidal behaviour, etc.

If we consider the youngsters, "failure in examination" was cited as the cause of suicide in 1,673 cases, with 991 male victims and 682 female victims. Notably, there were no transgender suicides attributed to this cause in 2021.

Knowledge about suicidal behaviour has increased greatly in recent decades. Research at different levels has shown the importance of the interplay between biological, psychological, social, environmental, and cultural factors in determining suicidal behaviours. At the same time, literature has helped to identify many risks and protective factors for suicide, both in the general population and in vulnerable groups

Objective

The aim of this project includes:

- **To study the recent trend of suicidal deaths in India**.

- **To explore the association between risk factors of suicides and gender across age categories for youngsters and adults**
  This was done to understand the factors that lead our youth to commit suicide.

- **To study the association between various attributes such as social status, Economic Status, Educational Status, and gender as to which portion of the population is more likely to commit suicide**.

Research Problem

The youth and adults are the building blocks of any Country. When studied it was found that nearly 34.6% of the people in the age group 18-30 years commit suicides, which is a huge problem. Also, The NCRB data shows that a total of 13,089 students died by suicide in 2021.

Hence, this raises the question of **what leads them to commit suicide.**

To answer this, this project was conducted to study the association between the risk factors of suicides and gender across the ages of the victims.

Research Methodology

- ## About the data

  To conduct this project, secondary data was used. Data on suicides were collected from Accidental Deaths & Suicides in India Report 2021: NCRB.

- ## Software used.

  R Studio was used for data visualization and conducting analysis. The graphs were plotted using ggplot2, tidyverse. While Factoshiny was used for conducting the Correspondence analysis.

- ## Method

  The present study includes the association of various attributes such as gender, age category of victims, social status, economic status, educational status, professional status, risk factors for suicides, means of suicides, etc. for the number of suicides in India during 2021. At the primary level, an odds ratio is calculated to measure an association between gender and risk factors for suicides. Further, we use CA to study the association between various attributes related to suicides. We summarize the conclusions and findings which may be helpful for policymakers, researchers, NGOs, etc. to develop preventive measures of suicides. The chi-square test for independence is commonly used to determine whether there is a significant association between the categorical variables (CVs) expressed in a contingency table. Usually, if the CVs are found to be dependent through this test, the interest would be, how the levels of CVs are related with each other. To portray the interrelations between the CVs, CA is used. CA is a multivariate statistical technique to visualize graphically the association between CVs in a contingency table. Particularly Simple CA (SCA) and Multiple CA (MCA) are useful for exploring the relations between two and more than two CVs respectively. The association between the categories of the variables are visualized on a map, called a biplot, allowing interpretations of their similarities and differences. The biplot shows the best two-dimensional approximations of the distances between row and column profiles of matrix data. The distance between any row points or column points gives a measure of their similarity (or dissimilarity). A biplot consists of lines and dots. Lines are used to reflect the variables of the dataset, and dots are used to show the observations. In a biplot, the length of a line approximates the variance of the variable. The longer the line, the higher the variance. The angle between the lines, or to be more precise, the cosine of the angle between the lines, approximates the correlation between the variables they

represent. The closer the angle to 900 or 2700, the smaller the correlation, while an angle of 00 or 1800 degrees reflects a correlation of 1 or −1, respectively. In common practice, CA includes all the categories of the CVs under consideration in the analysis since this gives the most comprehensive and global view of their interrelationships. Sometimes it is likely that after a global view of the data, it would be of interest to focus attention on a reduced set of response categories of the data. Another reason for restricting the categories to be analysed for a subset is when there are many variables and thus many biplot points in the graphical display. In such a situation, it is better to study a subset of variables of interest and their categories. This approach is known as subset CA.
The subset CA based on separate SVDs is applied to study the association of key risk factors for suicides with other attributes related to the number of suicides in India.

Statistical Analysis

The data about the number of accidental and suicidal deaths in India has been recorded and analysed by NCRB, India every year since 1967. The data is maintained according to various attributes such as gender, age group, social, economic, educational status, profession, etc. Further, the data are also classified according to the risk factors for suicides and the means of suicides. The suicide death rate for a year represents the number of suicides per 100,000 of the population during the reference year. It is to be noted that the data provided excludes farmer causes.

Suicide death rates have been increasing every year, which calls for the need to take preventive measures. The below graph shows the trend in suicide deaths for the past five years.

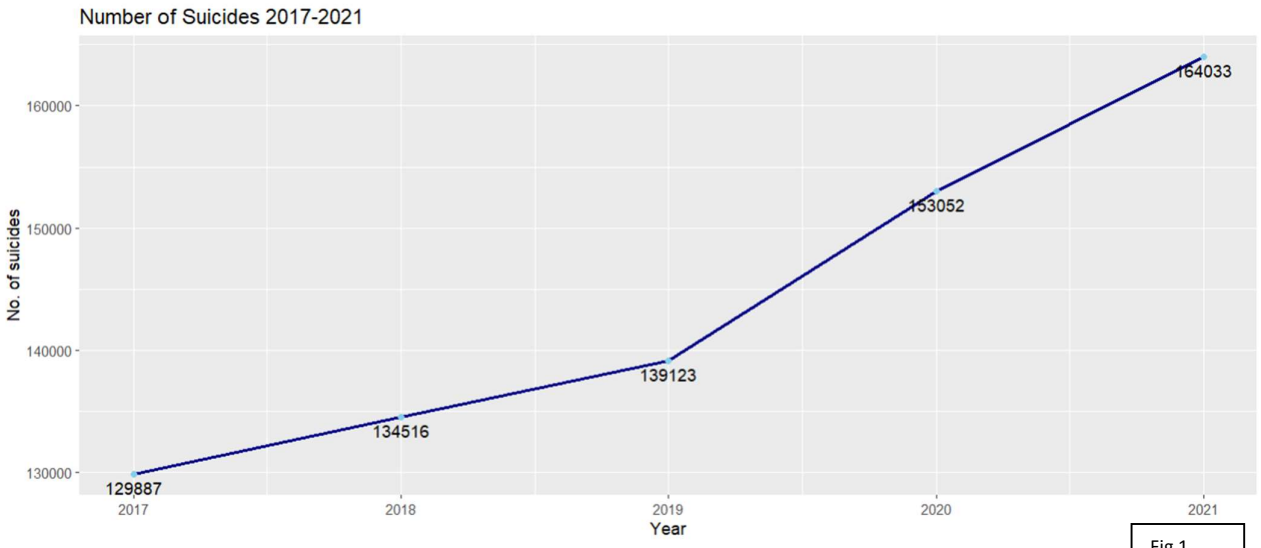
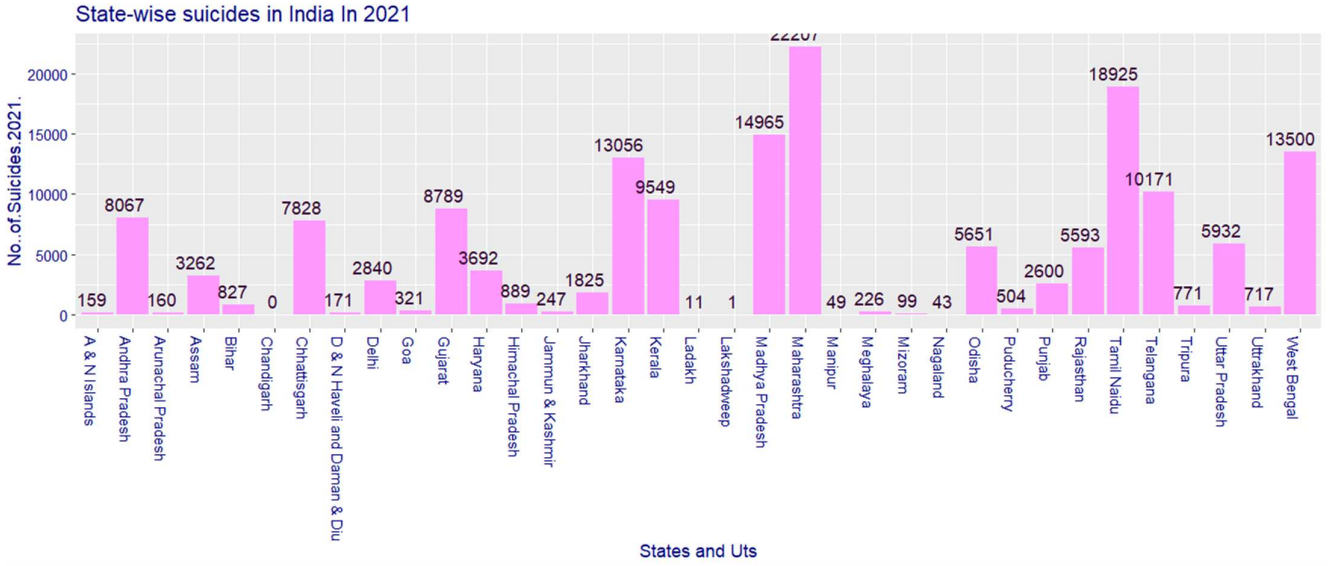


Fig 1

There is an increase in no. of suicidal deaths of 26.29% from 2017 to 2021. There is a wide variation of suicidal deaths in the States. Maharashtra is on top where 22207 cases were reported comprising 13.5% of the total cases reported, followed by 18925 suicidal deaths in Tamil Naidu, 14965 in Madhya Pradesh, 13500 in West Bengal and the list goes on.



Out of the total suicidal deaths in the year 2021, 118979 suicides were committed by males (i.e.,72.53%), 45026 were by females and 28 were by transgender. The percentage share of male suicidal deaths is continuously increasing, as in 2017, it was 68.5%. The below graph shows the gender-wise total suicides from 2017 to 2021.

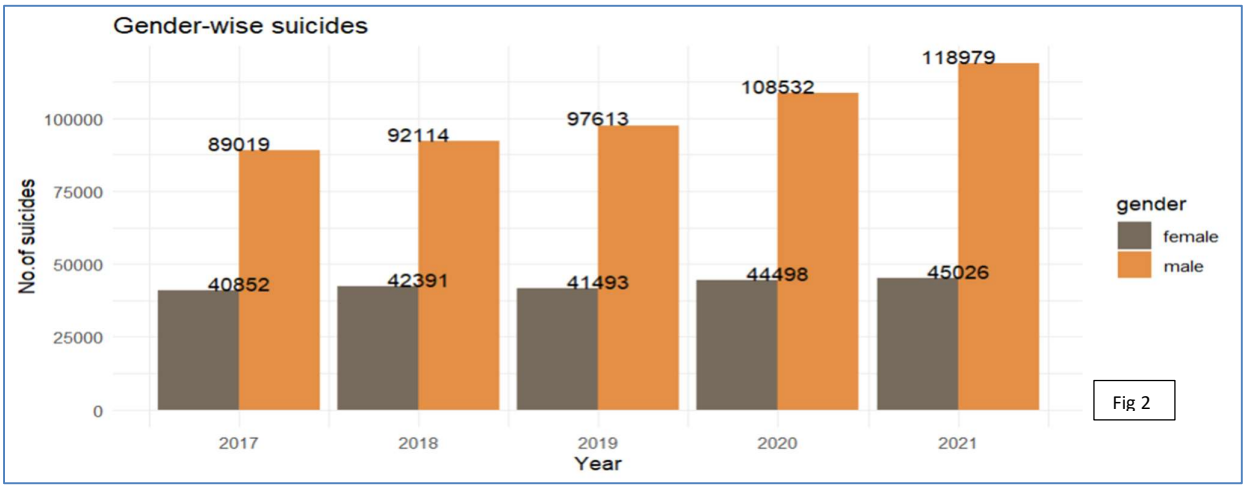


Fig 2

Most of the suicidal deaths were reported among youngsters and adults. Fig 3 shows the percentage share of the different age groups. It is seen that the per cent share of male suicidal deaths for the age group 18 – 30 and 30 – 45 is too high as compared to the females. Nearly 34.46% of the suicides were committed by people between the age group 18 – 30, this act by the younger age group results in a huge social, emotional, and economic burden on the society.

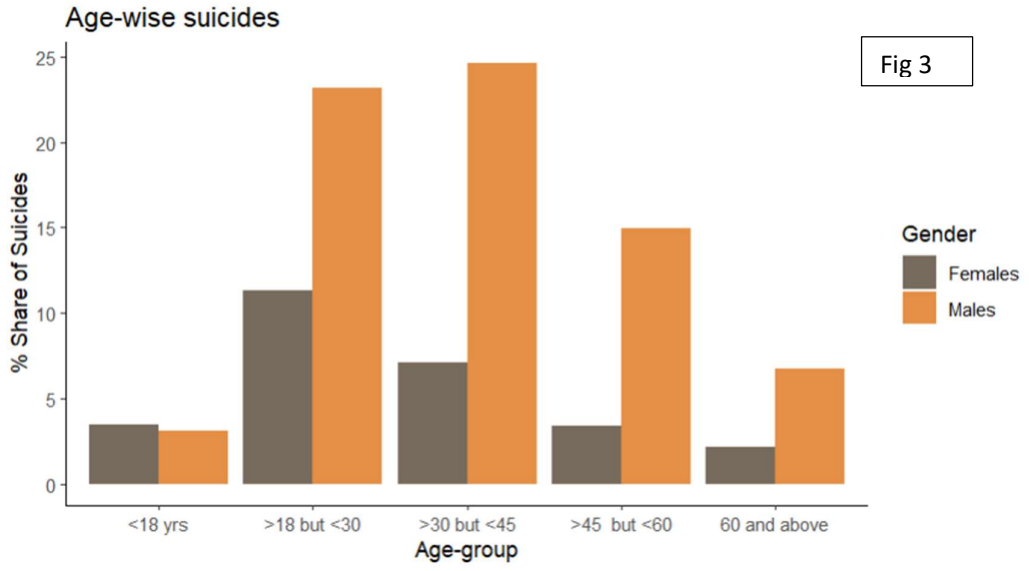


Fig 3

If we talk about the social status and gender of the victims, married males have the highest percentage share of the total deaths (49.4%), followed by married females 17.5%, and unmarried males (16.65%). This can be due to various reasons such as extramarital affairs, family pressure, dowry-related issues and so on. Figure 4 shows the percentage share of the victims according to their social status.

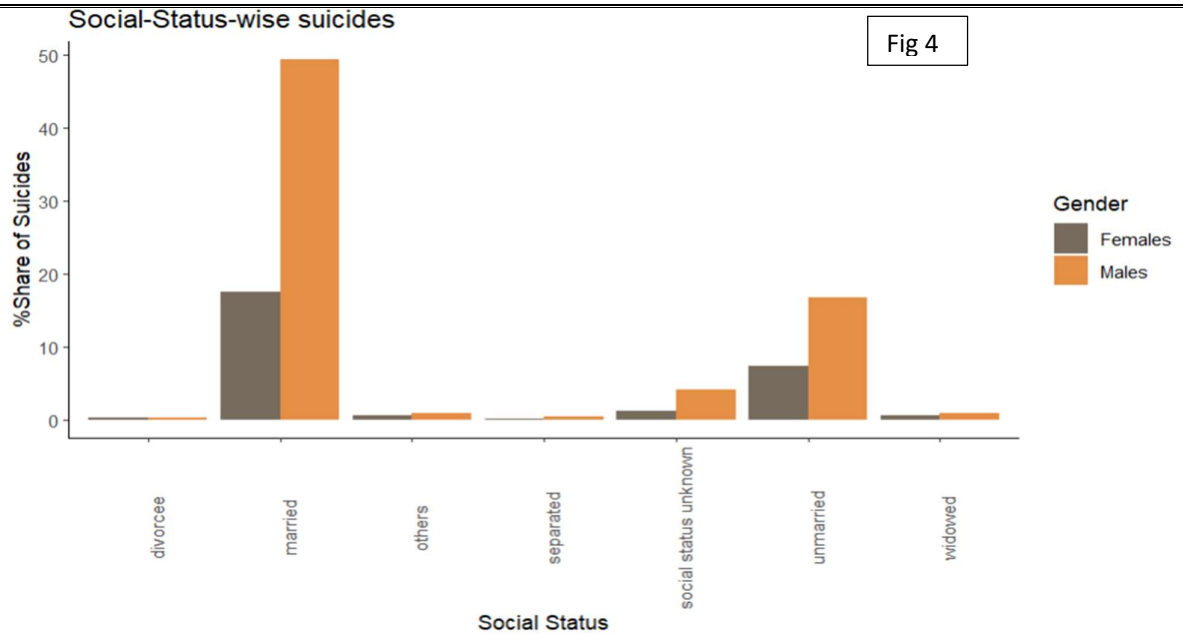


Fig 4

If educational status is considered, professionals are less likely to commit suicide while people who have received education up to matriculate are more likely to commit suicide. Figure 5 shows the education status and gender-wise percentage share.

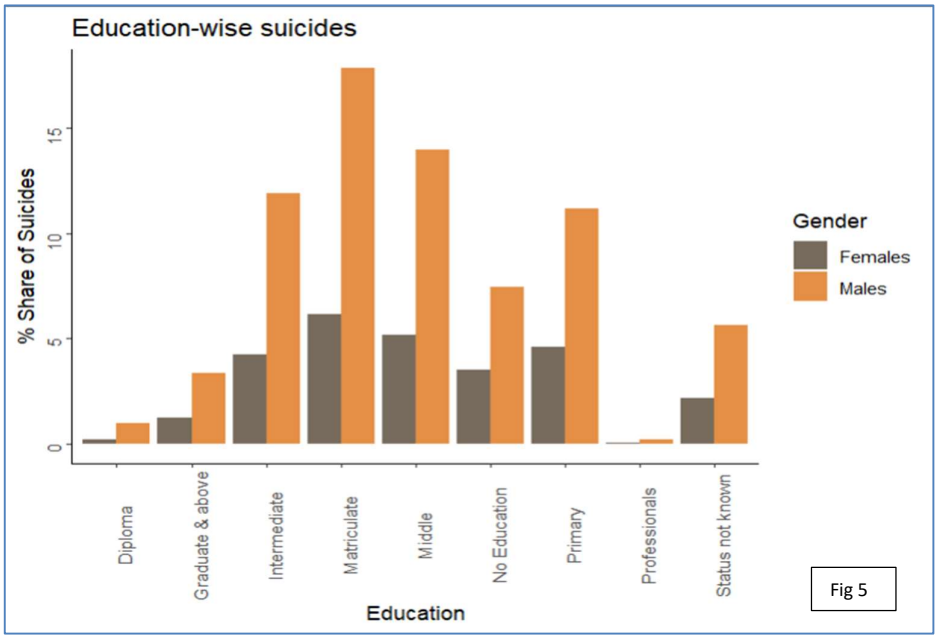


Fig 5

If the Economic status is considered, around 64.14 % of the total victims belonged to a group earning less than 1 lakh annually. This shows the impact of poverty on suicides. Figure 6 represents the percentage share of the victims according to their gender and economic status.

The parameters such as age, sex, social, economic, educational, and professional status play a significant role towards suicidal behaviour. However, the risk factor for suicides needs to be studied for the prevention of such cases. A brief study of the risk factors for suicides for age category and gender is presented.

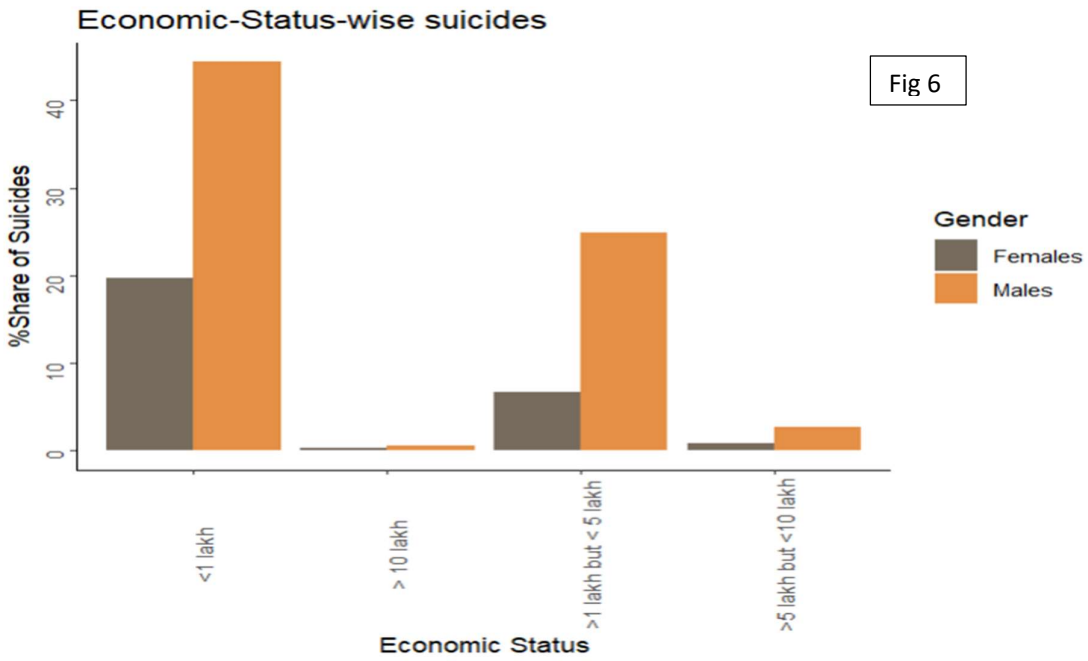
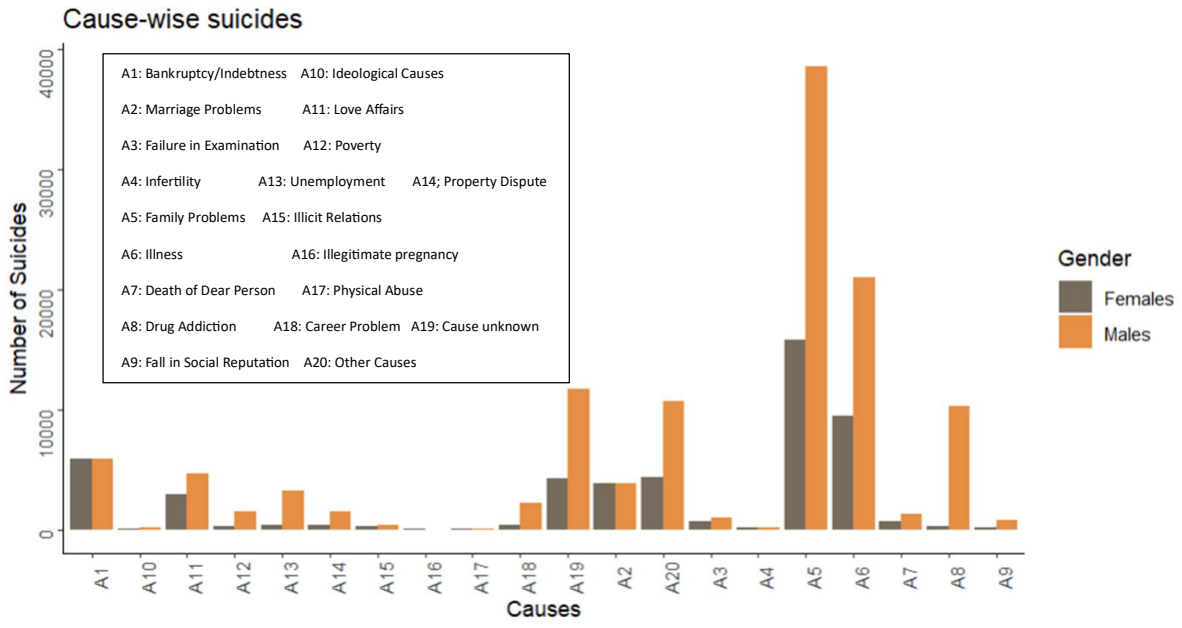


Fig 6

Study of the Risk Factors of Suicide

There are various risk factors which ultimately result in suicide. Suicidal behaviours can be observed when there is a situation or event that the person finds overwhelming, such as ageing, the death of a loved one, drug or alcohol use, emotional trauma, serious physical illness, unemployment, money problems, etc. So, it is imperative to design interventions that can address distress among various demographic groups, and not aggravate the problem by focusing on health and family problems alone.



Family problems apart from other marriage-related issues and illness are the major key risk factors for suicides. Apart from these two causes, other causes that contribute to suicides are marriage-related problems, love affairs, bankruptcy, and others.

If family problems are handled with care and precautions, most suicides can be prevented. If patients are given proper therapy/consultation, suicides can be prevented to some extent.

To further explore the risk factors for suicides, we consider the key risk factors only. We study the risk factors for suicides by gender through odds ratio (Tab 1). Table 1 gives the odds ratios for respective risk factors.

Sno.	Causes	Males	Females	Total	Odds favoring males	Odds favoring females
1	Bankruptcy	5879	482	6361	12.197	0.082
2	Marriage related problems	3833	4069	7902	0.94	1.06
3	Family Problems	38634	15769	54403	2.45	0.41
4	Illness	21014	9426	30440	2.23	0.45
5	Drug Abuse	10284	275	10559	37.396	0.028
6	Love affairs	4667	2894	7561	1.61	0.62
7	Poverty	1535	244	1779	6.29	0.16
8	Unemployment	3203	337	3540	9.50	0.10
9	Physical abuse	25	53	78	0.47	2.12

Table 1: Odds ratio for causes of suicides for males and females.

From the above table, males are 37 times more likely to commit suicide due to drug abuse than females. They are 12 times more likely to commit suicide due to bankruptcy than females. Unemployment is also a subsidiary cause of male suicides. Males are 9 times more likely to commit suicide due to unemployment than females. Thus, bankruptcy, unemployment, and drug abuse are the common risk factors for suicides among males.

To check the pattern of association of these attributes concerning the number of suicidal

We will first perform a Chi-square test for deaths in a specific age group.

- **Hypothesis**
  Null Hypothesis: There is no association between risk factors and gender across age categories of victims.
  Alternative Hypothesis: There is an association between the risk factors and gender across age categories of victims.

The result obtained:

```
The chi-square of independence between the two variables is equal to 17945.43 (p-value = 0).
```

Since the p-value is less than 0.05, we accept an alternative hypothesis, therefore, there is an association between the attributes concerning the number of suicidal deaths.

- **Exploration of the association between risk factors for suicide and gender across Age-Category of suicide victims**.

Our main objective is to see what are the risk factors for suicides in the youth and adults, as they are the ones who build the economy of any country.

We perform Correspondence Analysis for the categorical variables (CVs), gender and risk factors(A) across age- category (C) of suicide victims. The Categories of the CVs are summarized in Table 2.

For the year 2021, the distribution of the number of suicidal deaths regarding 3-attributes is presented in a 3-way contingency table (Table 4). The biplot points far from the origin and close to each other are considered for interpretations, since the more the vector length, the better the discrimination ability.

If we consider all the risk factors, there will be too many biplot points which are complex to interpret; hence, we will consider only 10 key risk factors that have a percentage share in the total cases (A1, A2, A3, A5, A6, A8, A11, A12, A13, A15). We will drop out the risk factors with a relatively small percentage share. We will perform a subset based on separate SVDs to the cross-classified data with the key risk factors.

Name of the CV	Categories of CVs
Gender	Male(M), Females(F)
Risk Factors for Suicides(A)	A1: Bankruptcy/Indebtness A2: Marriage Problems A3: Failure in Examination A4: Infertility A5: Family Problems A6: Illness A7: Death of Dear Person A8: Drug Addiction A9: Fall in Social Reputation A10: Ideological Causes A11: Love Affair A12: Poverty A13: Unemployment A14: Property Dispute A15: Illicit Relations A16: Illegitimate pregnancy A17: Physical Abuse A18: Career Problem A19: Cause unknown A20: Other Causes
Age Category(C)	C1: Below 18yrs, C2: 18-30yrs C3: 30-45yrs

Table 2: Categories of CVs for risk factors.

The column-wise results are given in Fig 7. From the biplot (fig 8) it is observed that Males Below 18 years commit suicide due to *Failure in Examination* while females in the same age group commit suicide due to *Failure in Examination* and *Love Affairs*. The other conclusions about the association of risk factors for suicides with gender and age group from biplot are summarized in Table 3 sorted on the attributes' discrimination ability and strength of association of pair of attributes.

Age group	**Gender**	**Pair of Biplot points**	**Interpretation of pair of biplots.**
Below 18	Males Females	(M.C1, A3) (F.C1, A3) (F.C1, A11)	Males below 18 tend to commit suicide due to Failure in Examination while Females of the same age group commit suicide due to Failure in Examination and Love Affairs.
18-30	Males Females	(M.C2, A8) (M.C2, A12) (M.C2, A13) (F.C2, A2)	Males of this age group commit suicide due to Drug addiction, Poverty and Unemployment. While females of this age group commit suicide due to Marriage-related issues.
30-45	Males Females	(M.C3, A8) (M.C3, A13) (M.C3, A1) (F.C3, A2) (F.C3, A15)	Males of this age group commit suicide due to Drug addiction, bankruptcy, and unemployment. While females commit suicide due to Marriage-related issues and illicit relations.

Table 3: Association between risk factors for suicides and gender across age categories.

Columns

	Iner*1000	Dim.1	ctr	cos2	Dim.2	ctr	cos2	Dim.3	ctr	cos2
M.C1	36.130	0.724	18.537	0.572	0.560	21.186	0.342	0.249	18.933	0.067
M.C2	8.725	-0.010	0.026	0.003	0.099	5.326	0.356	-0.125	38.465	0.567
M.C3	47.246	-0.361	39.898	0.942	0.061	2.202	0.027	0.055	8.062	0.022
F.C1	34.714	0.765	23.062	0.741	0.430	13.933	0.234	0.045	0.699	0.003
F.C2	41.234	0.354	17.969	0.486	-0.349	33.435	0.473	-0.064	5.164	0.016
F.C3	19.570	0.075	0.508	0.029	-0.374	23.918	0.713	0.192	28.677	0.189

Fig 7

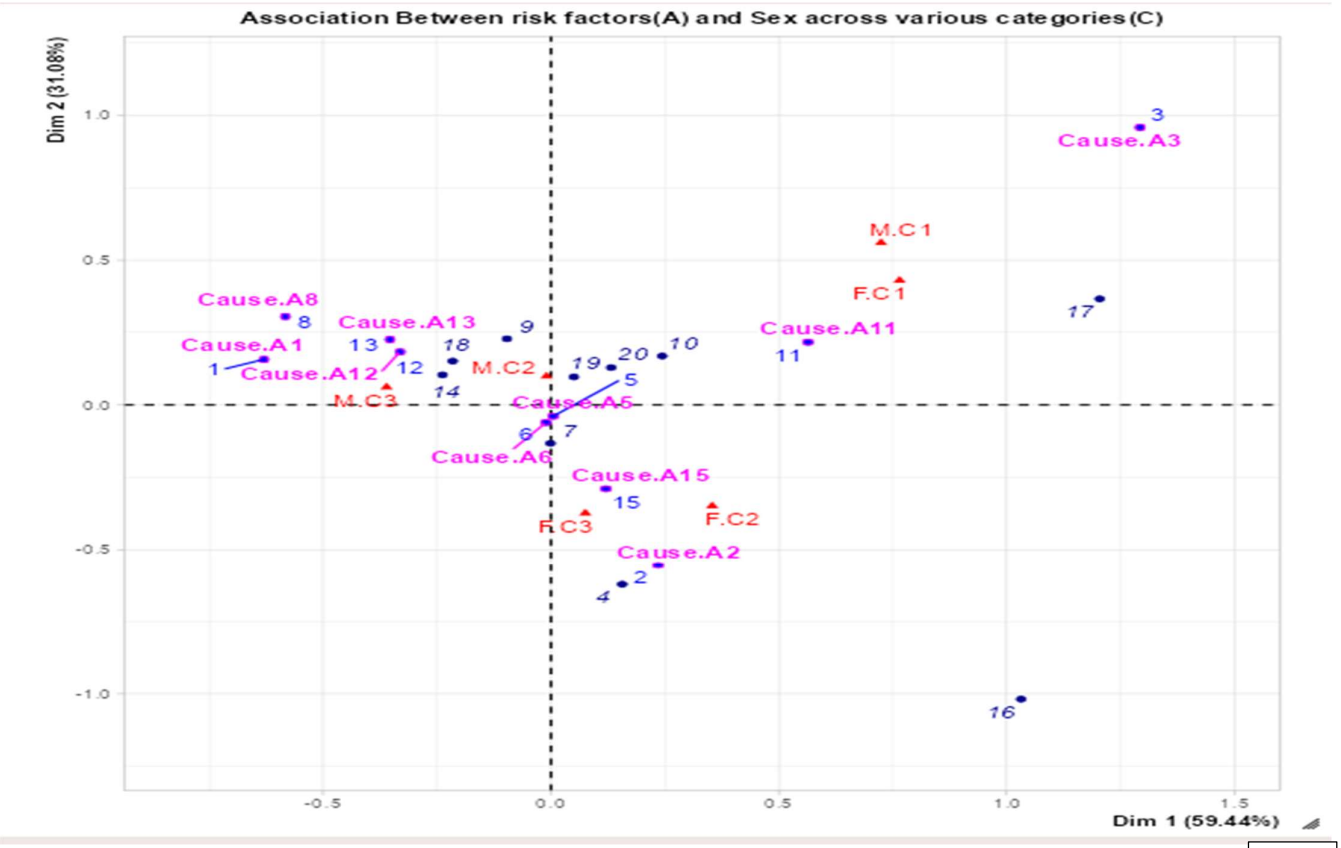


Fig 8

S. No.	Cause of Suicide	Males			Females			Total
		C1	C2	C3	C1	C2	C3	
1	A1	16	1100	2517	10	125	219	3987
2	A2	68	1772	1581	126	2631	1153	7331
3	A3	467	449	56	397	265	16	1650
4	A4	0	43	95	13	102	81	334
5	A5	1630	12345	13764	1603	6786	4515	40643
6	A6	596	4973	5955	812	2537	2436	17309
7	A7	38	352	374	42	180	193	1179
8	A8	83	2774	4138	33	78	87	7193
9	A9	37	214	230	23	52	40	596
10	A10	10	43	33	9	22	8	125
11	A11	585	3167	817	910	1652	314	7445
12	A12	43	372	590	26	81	74	1186
13	A13	55	1316	1264	28	157	99	2919
14	A14	38	425	528	31	103	113	1238
15	A15	13	170	161	22	143	96	605
16	A16	0	0	0	7	31	12	50
17	A17	6	5	8	26	19	7	71
18	A18	58	786	809	47	197	103	2000
19	A19	672	4138	4068	664	1812	1069	12423
20	A20	660	3497	3427	826	1615	1003	11028

Table 4: Contingency table

Results related to the risk factors for suicides.

- Males below 18 years commit suicide due to 'Failure in Examination', while females of the same age group commit suicide due to 'Love Affairs' and 'Failure in Examination'.
- Males between age group 18-30 commit suicide due to 'Drug Addiction', 'Poverty' and 'Unemployment', while females of this age group commit suicide due to 'Marriage related issues'.
- Males between the age group 30-45 commit suicide due to Bankruptcy, Unemployment, and drug addiction while females of this age group commit suicide due to marriage-related issues and Illicit relations.

Conclusions

Overall, the suicide death rate in India is observed to be continuously increasing. The percentage of male suicides is observed to be more as compared to the percentage of females.

The leading risk factors for suicides are family problems, illness, drug addiction, failure in examination etc. We observe that the proportion of female victims is comparatively higher under heads of marriage-related issues while for males it is higher under the head of drug abuse, unemployment etc. Family problems and illness are the major issues for committing suicide in India.

When we talk about youth, they tend to commit suicide due to failure in examination and love affairs, which shows that the youths lack proper guidance and receive negligence when they are under depression or immense pressure. While the adults commit suicide due to marriage-related issues, unemployment, and poverty.

Bibliography

- ADSI Report 2021: NCRB
- ADSI Report 2020: NCRB
- Biplot Interpretation: *https://bookdown.org/danieljcarter/socepi/biplots-and-interpretation.html*
- Correspondence Analysis using R Programming.: *http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/120-correspondence-analysis-theory-and-practice/#distance-or-similarity-between-row-profile*