PROJECT NAME:

Suicides In India Visualizations Using Tableau

Suicides In India Visualizations Using Tableau

Category: Data Analytics

Skills Required:

Exploratory Data Analysis, MySQL, Databases, Tableau

Team leader :   PANGA SAI PRUDHVI

Team members :   1) PALIVELA NAGA PUJITHA

2) NODAGALA AVINESHWARA SAI

3) MOSUGANTI JEDIDIAH

Project Description:

The project is majorly distributed into 4 parts :

1. Downloading the data set and loading it in Tableau.
2. Plotting the data using different types of options available.
3. Altering the visualizations to make them meaningful using the features available in Tableau.
4. Exploring features like forecasting without writing codes the way you need to do in ML.

Technical Architecture:
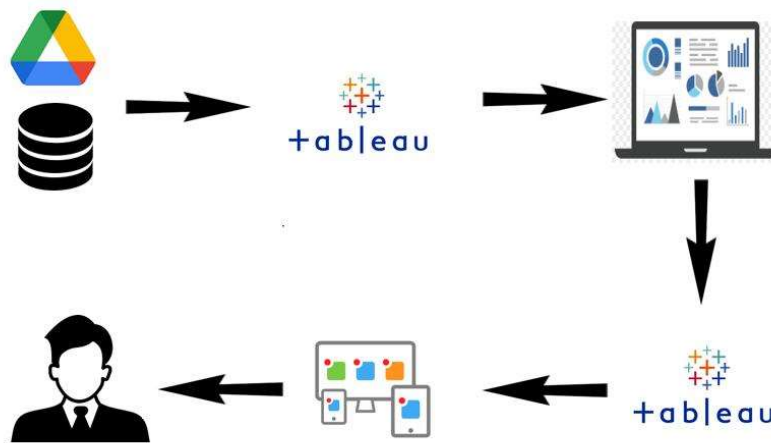
Pre-Requisites

For Completing this project these are some of the prerequisites needed

- A system with a minimum 4GB RAM and 128GB Hard Disk
- Good Internet Connection
- Google Drive / Any of the Database Server with Management Studio
- MySQL:
- **SQL Server Management Studio:**
- **Tableau Desktop:**
- Tableau Public Account:
  https://public.tableau.com/app/discover
- Html, CSS or Bootstrap

Prior-Knowledge

To Complete this project, one must understand the below concepts and able to work with the tools

- **Data Visualization:**
-
- **Univariate, Bi- Variate and Multi-Variate Analysis**
- **Chart Types:**
- **Tableau:**
- **Business Intelligence:**

Project Objectives

By the end of this project, you will:

- Able to Connect Tableau with different data sources
- Know fundamental concepts and techniques used for Data Visualization.
- Gain a broad understanding about data and different types of charts.
- Have knowledge of developing Visualizations, Dashboards and Story.
- Able to Integrate the developed dashboard and story with the web application

Project Flow

 o Users can create multiple analysis graphs/charts.
 o Using the analysed chart creation of Dashboard is done.
 o Saving and Visualizing the final dashboard in the personal Tableau public profile.
 o To accomplish this, we have to complete all the activities and tasks listed below
 o Working with the Dataset
 o Understand the Dataset
 o Loading the Dataset
 o Visualizations of Suicides in India (2001-2012)
 o Problem Statement 1: Why Are Suicides Happening?
 o Problem Statement 2: Suicides Per Year (Forecasting)
 o Problem Statement 3: Classification on basis of age group
 o Problem Statement 4: State-wise suicide counts for different reasons

Data Collection

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate

outcomes and generate insights from the data.

Working With Dataset

Understand The Data

The data was sourced from a Github Repository. You can see the data set used for the project by clicking here.

Let's understand the data we're working with and give a brief overview of what each feature represents or should represent

5. State: Name of the Indian state where the data belongs to.

6. Year: Year ranges from 2001-2012, the complete field contains data for that year respectively.

7. Type_code: Major classification of "Why Did People Commit Suicide".

8. Type: Subclassification of type code.

9. Gender: Either the person committing suicide was a male/female.

10.   Age_group: What age group does the person belong to who committed suicide.

11.   Total: Count of people who committed suicide and has the common above 6 mentioned parameters.

Loading The Dataset

Before you can build a view and analyze your data, you must first connect Tableau to your data. Tableau supports connecting to a wide variety of data, stored in a variety of places.

The data might be stored on your computer in a spreadsheet or a text file, or in a big data, relational, or cube (multidimensional) database on a server in your enterprise.

In our case, we will be using a spreadsheet or text file for making our analysis. Watch the video for understanding the connection of the dataset in Tableau.

Data Visualization

Data visualization is the process of creating graphical representations of data in order to help people understand and explore information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.
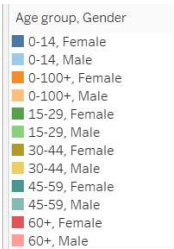
Visualizations Of Suicides In India (2001-2012)

Using the Suicides in India 2001-2012 dataset, we plan to create a dashboard showing the facts about suicides which we're not able to understand by just the numbers. These visualisations help us to get a better understanding of the data in a single look, as well as, is easily understandable to a layman.

Age group
- 0-14
- 0-100+
- 15-29
- 30-44
- 45-59
- 60+

ars | ⊞ Dashboard 2 | horizontal bars | packed bubbles | dual combination(story) | symbol maps | stacked bars(story) | gant views | dual lines | area chart | ◫ Story 1 | side by side circles

Age group, Gender
- 0-14, Female
- 0-14, Male
- 0-100+, Female
- 0-100+, Male
- 15-29, Female
- 15-29, Male
- 30-44, Female
- 30-44, Male
- 45-59, Female
- 45-59, Male
- 60+, Female
- 60+, Male

Age group:	0-100+
Gender:	Male
Type:	Married
Count of Suicides_in_India_2001-2012.csv:	456
Total:	19,95,021

shboard 2 | horizontal bars | **packed bubbles** | dual combinati | ual lines | area chart | ◫ Story 1 | side by side circles | box and w

Reasons Behind Suicides Happening In India (2001 - 2012)

For visualizing the above, we will require the following data

- Type_Code
- Total
- Gender

-

-
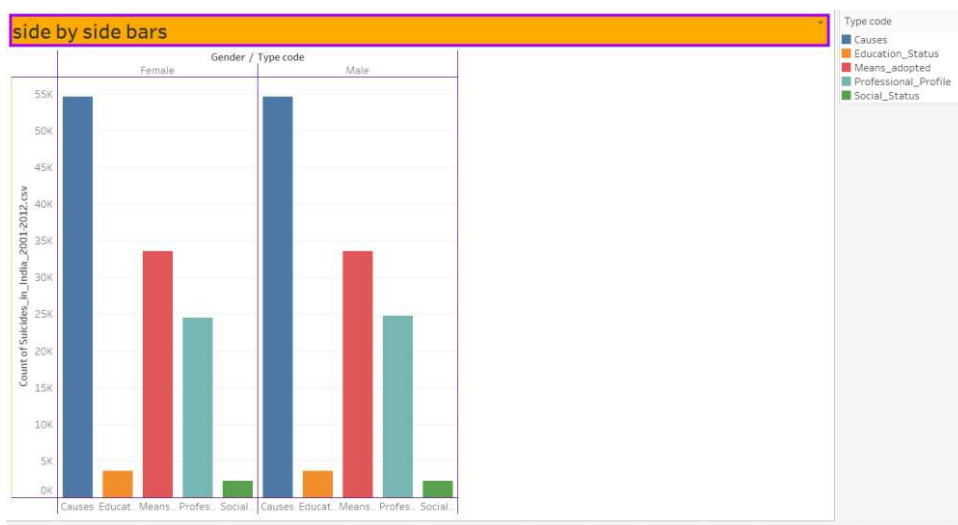
From this data, we will be plotting a "Side by Side Bars" to see the comparison b/w males and females as well as the plot the total no of suicides (all years combined) vs reason of suicide.

Refer to the video below to create the visualization for the above statement.

Age Group Wise Classification

In this milestone, we will analyze age-wise suicide analysis
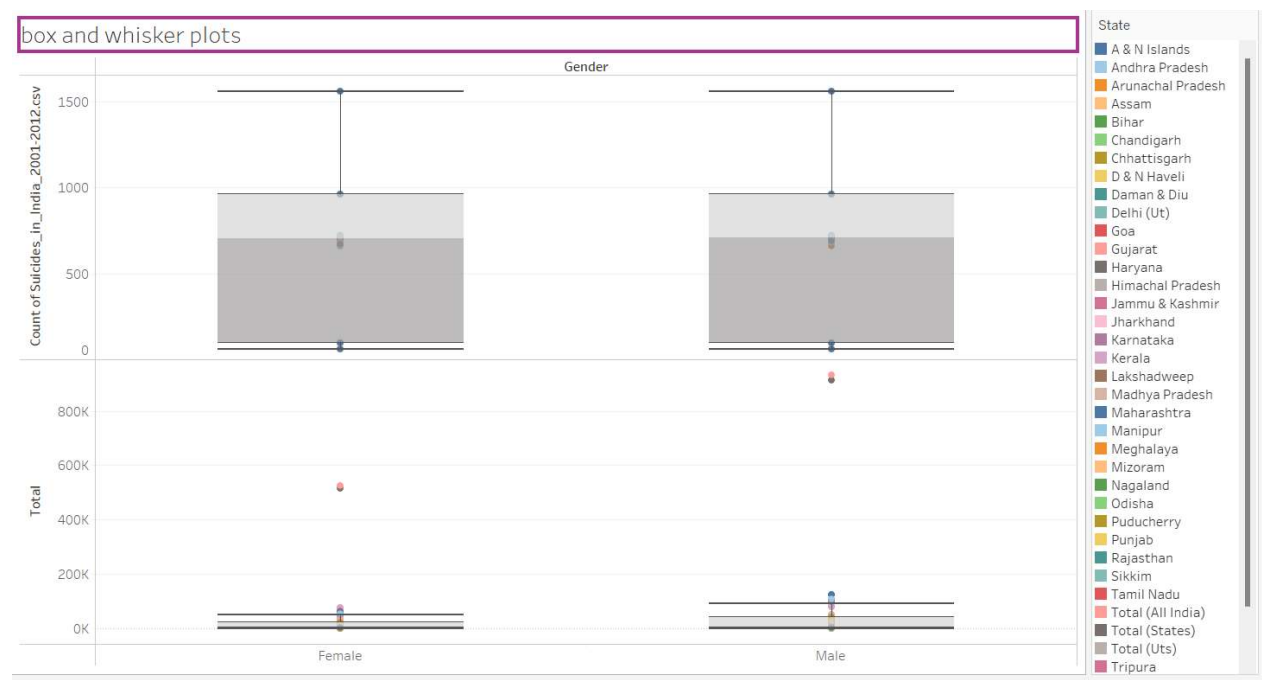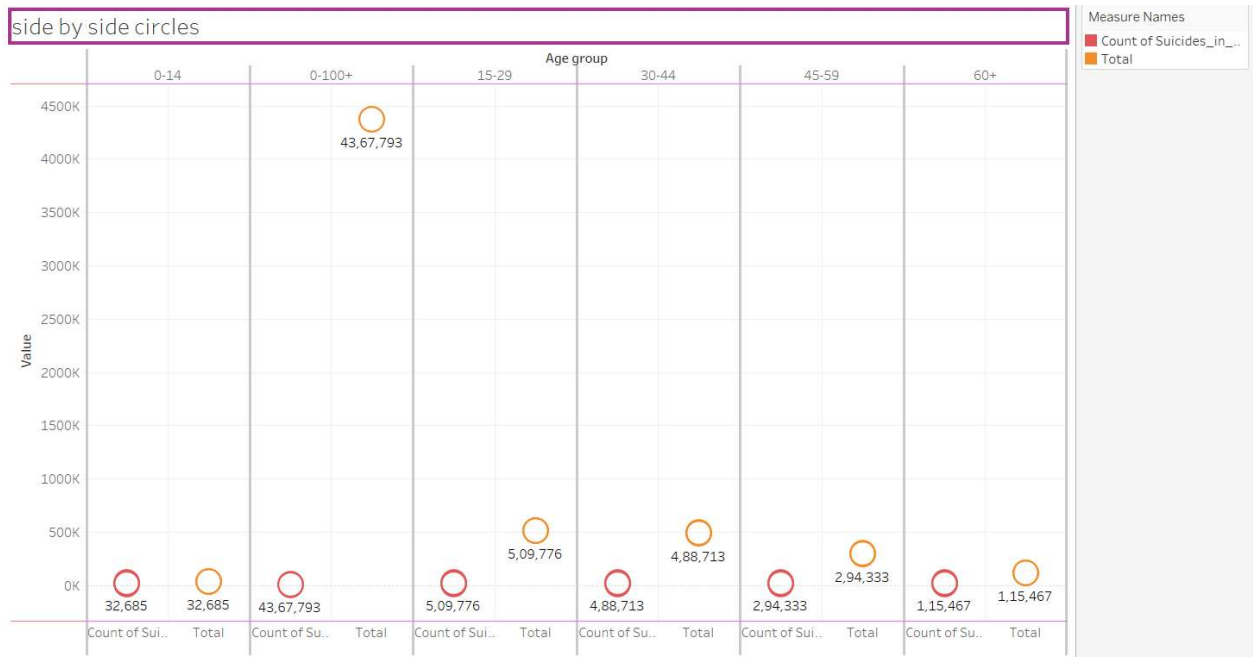
Age Wise Analysis Of Suicide Data

We will be plotting a simple bar chart to show the number of suicide cases according to age groups. One thing that we need to notice here is, majority of our data did not have the correct age group values, so they were given a trivial value [0-100+], so we had to remove that column due to which the data got a little biased, but it was still good enough for our visualisation. Problems like these are very common and a data visualization is considered good only if they are plotted to keep these problems in mind and to avoid such things. The visualisation answers the question, which age groups need more mentoring and taken care of so as to decrease the majority of cases.

State Vs No. Of Suicides Based On Reason

In this milestone, we will analyze the state-wise suicide rate for multiple reasons

Suicides In Each State Over The Course Of 12 Years

As in our data, there are many states mentioned, so to see which part of India is having more suicides, we decided to plot a "**Maps**" type plot. Here the colour gradient form is used and a heatmap kind of plot is generated, where the number of suicides directly affects the colour of the region, to which the data belong to. Here, we also created a dropdown menu to choose the reasons one by one in order to see, which state is suffering more due to that particular reason.

These visualisations can help govt. of India in understanding what problems are more common in which region in order to prepare campaigns and help people by stopping them to commit suicide.
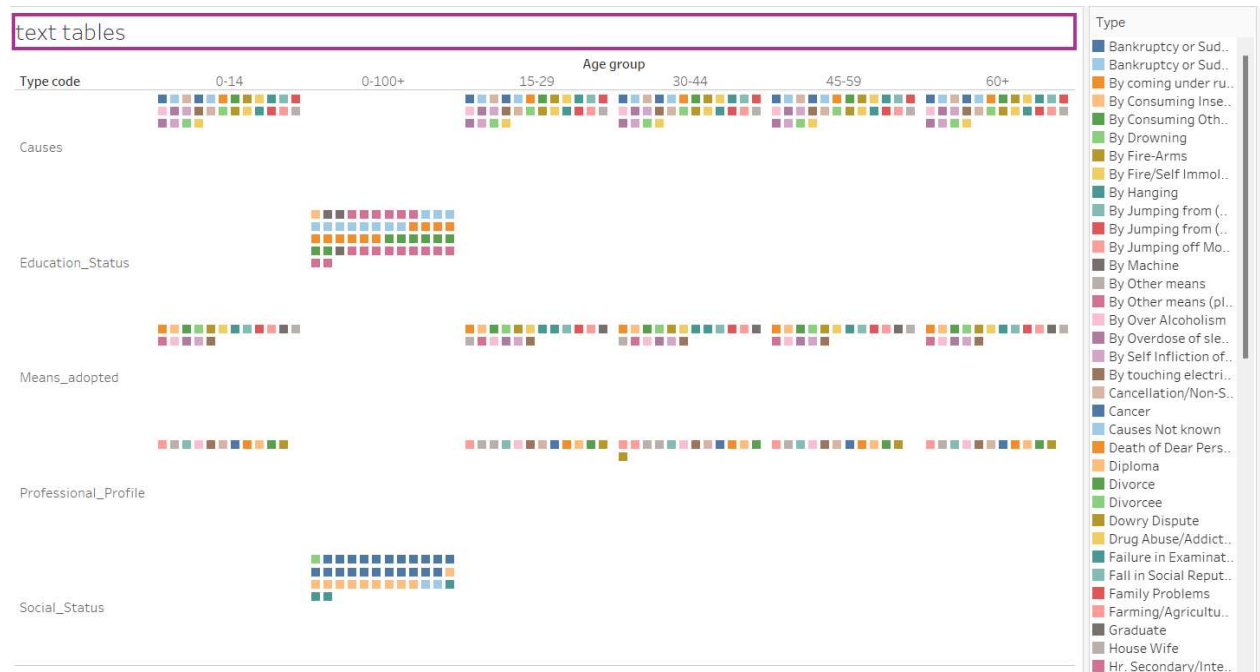
Dashboard

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

SUICIDES ANALYSIS IN INDIA DASHBOARD-1

side by side bars

pie chart
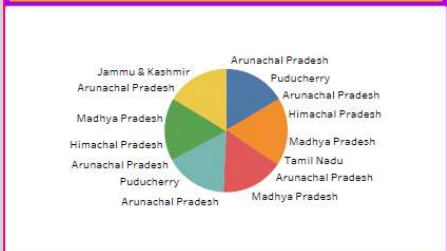
Age group
- 0-14
- 0-100+
- 15-29
- 30-44
- 45-59
- 60+

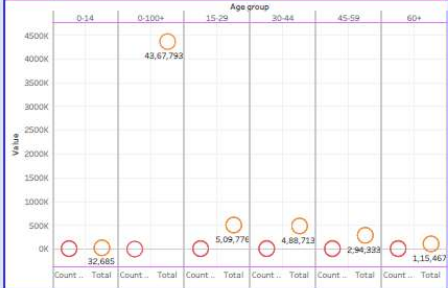next

horizontal bars
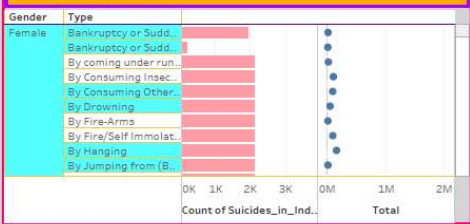
packed bubbles
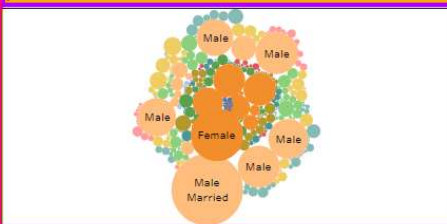
Age group, Gender
- 0-14, Female
- 0-14, Male
- 0-100+, Female
- 0-100+, Male
- 15-29, Female
- 15-29, Male
- 30-44, Female
- 30-44, Male
- 45-59, Female
- 45-59, Male
- 60+, Female
- 60+, Male

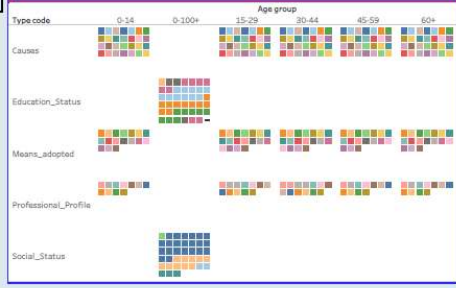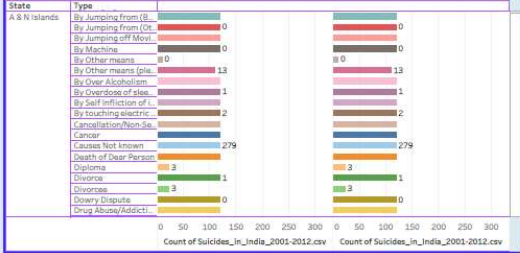**SUICIDES ANALYSIS IN INDIA DASHBOARD-2**

side by side circles

Measure Names
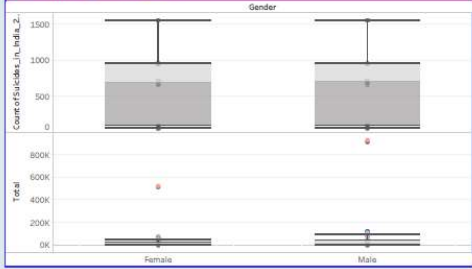- Count of Suicides_in_India_200..
- Total

text tables

Type
- Bankruptcy ..
- Bankruptcy ..
- By coming u..
- By Consumi..
- By Consumi..
- By Drowning
- By Fire-Arms
- By Fire/Self..
- By Hanging
- By Jumping ..
- By Jumping ..
- By Machine
- By Other me..
- By Other me..
- By Over Alco..
- By Overdose..
- By Self Inflic..
- By touching ..
- Cancer
- Causes Not ..
- Death of De..
- Diploma
- Divorce
- Divorcee

- Dowry Disp..
- Drug Abuse/..
- Failure in Ex..
- Fall in Social..
- Family Probl..
- Farming/Ag..
- Graduate
- House Wife
- Hr. Seconda..
- Ideological ..
- Illegitimate ..
- Illness (Aids..
- Insanity/Me..
- Love Affairs
- Married
- Matriculate/..
- Middle
- Never Marri..
- No Education
- Not having ..
- Other Cause..
- Other Prolo..
- Others (Plea..
- Paralysis
- Physical Ab..

- Post Gradu..
- Poverty
- Primary
- Professiona..
- Professiona..
- Property Di..
- Public Sect..
- Retired Per..
- Self-emplo..
- Seperated
- Service (Go..
- Service (Pri..
- Student
- Suspected/..
- Unemploye..
- Unemploym..
- Widowed/V..

bullet graphs

box and whisker plots

State
- A & N Islands
- Andhra Prad..
- Arunachal P..
- Assam
- Bihar
- Chandigarh
- Chhattisgarh
- D & N Haveli
- Daman & Diu
- Delhi (Ut)
- Goa
- Gujarat
- Haryana
- Himachal Pr..
- Jammu & Ka..
- Jharkhand
- Karnataka
- Kerala
- Lakshadweep
- Madhya Pra..
- Maharashtra
- Manipur
- Meghalaya
- Mizoram
- Nagaland
- Odisha
- Puducherry
- Punjab
- Rajasthan
- Sikkim
- Tamil Nadu
- Total (All In..
- Total (States)
- Total (Uts)
- Tripura
- Uttar Prade..
- Uttarakhand
- West Bengal

Story

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

dual combination(story)

The trends of count of Suicides_in_India_2001-2012.csv and Total for Year broken down by Gender. Color shows details about count of Suicides_in_India_2001-2012.csv and Total. For pane Sum of Total: The marks are labeled by Gender.

Publishing And Web Integration

Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others.

Publishing Dashboard and Reports To Tableau Public

Step 1: Go to Dashboard/story, click on the share button on the top ribbon

Step 2: Once you click on connect it will ask you for the tableau public username and password

Once you login into your tableau public using the credentials, the particular visualization will be published into the tableau public
Note: While publishing the visualization to the public, the respective sheet will get published when you click on the share option.
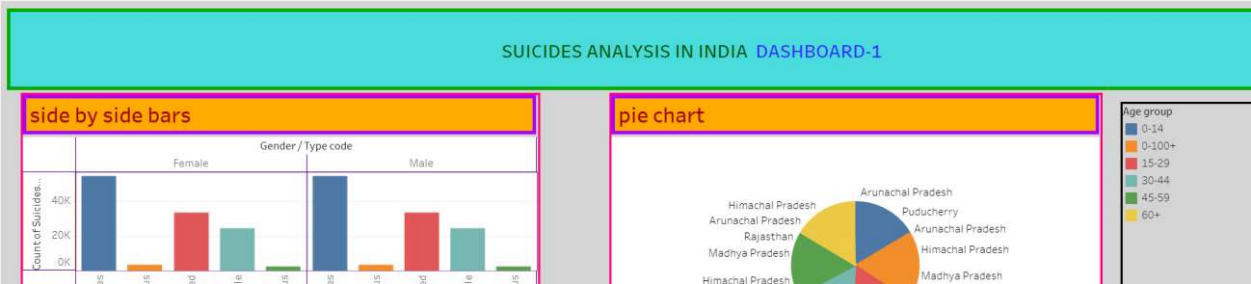
Integrating In Web With Embedded Code

Story

A data story is a way of presenting data and analysis in a narrative format,with the goal of making the information more engaging and easier to understand.