

CAUSE OF DEATH

- Importing the dataset by using PANDAS library.
- Read the dataset and assign data to a variable.
- Check the column type either it is float or object.
- Two columns in dataset are present in as object.
- And rest of all are in as integer.
- Now, let's check for null values present or not in dataset.
- Good, no null values are there in dataset.
- It is a very good indication that there are no null values.
- Now, its time to encode those both columns that are in as object.
- Let's import Label Encoder.
- Encode the selected columns.
- Now, let's do "describe method" to check a lot of information about dataset.
- "Describe Method" gives about mean, standard deviation, quantiles.
- Now, import libraries for visualization.
- First, we plot distribution plot to see how data is distributed.
- By distribution plot, it seems like that most of data is right skewed.

- Now, again plot boxplot to see the outliers present in dataset.
- A lot of outliers are present in dataset.
- Now, again distribution plot individually for each column.
- Now, it is more-clear that data is right skewed.
- If I try to remove outliers then whole dataset eliminates.