

# SPPU-TE-COMP-CONTENT – KSKA Git

Q1) List out the different types of plot patterns of data.

ANS. Plots help us to visibly visualize data in the order to reveal the patterns, relationships

## 1. Distribution of Data.

### (A) Histogram

- Divides the data into bins and shows the frequency of data points within each bin.

### (B) KDE Plot

- Visualizes the p.d.f. (Probability density function) on the continuous data.

### (C) Box Plot.

- Shows the spread of data, highlighting Medians and outliers.

## 2. Relationship between Variables.

### (A) Scatter Plots

- Visualizes relationships between two continuous variables.

### (B) Line Plots

- Shows the trends over time or a sequence order.

### (C) Bubble Plot:-

- Adds a third variable through size of the database to a scatter plot.

## 3. Time Series Analysis.

### (A) Time Series Analysis.

- Displays data trends over time

### (B) Area Plot.

# SPPU-TE-COMP-CONTENT – KSKA Git

Highlights cumulative trends over time.

## 4. Specialized Plots

### (A) Pie Chart

Visualize the properties in Categorical Data.

### (B) Swarm Plot

Visualizes the data points while avoiding overlay categorical variables.

Q2) Explain when you will use Distribution Plot and when you will use categorical plots.

ANS.

Use of Distribution Plot.

No

Use of Categorical Plot.

(1) To Analyze how numerical data is spread or distributed.

(2) To Analyze relationship frequencies or comparisons involving the categorical data.

(2) For Eg:-

(2) For Eg:-

- ① Histogram
- ② Boxplot.

① Boxplot

② countplot.

(3) Identifying the shape of Data distribution. (Normal, skewed, etc.)

(3) Comparing numerical data (average) across categories.

Q3) Write the conclusion from the following plot. (Consider the titanic dataset.)

ANS.

The swarm plot, based on titanic Dataset in-depth visualization of survival.

(1) Survival Rates by Gender:-

# SPPU-TE-COMP-CONTENT – KSKA Git

(A) Female Passenger:-

- A significant proportion of Gender survived the disaster.

(B) Male Passenger:-

- Males have a disproportionality higher number non-survivors (blue dots)

(2) Impact of Age on Survival.

(A) children

- Young Boys shows higher survival rate compared to older designing with the "Women and Children First."

(B) Adults and Elderly.

- Among males, survival data significantly drop few older groups where dots are mostly Blue.

Q4. > Which parameters is used to add another categorical variable to the violin plot

ANS. Violin plots are powerful tools for visualizing distribution of dataset, especially when you to compare the distribution of multiple category.

o Adding another parameter categorical variable

i) The hue parameter in a violin plot issued to break down or differentiate the data within each categorical variable.

ii) It allows us to split the data based on another categorical features, by making it easy to compare the distributions across different sub-categories within each main category.

# SPPU-TE-COMP-CONTENT – KSKA Git

## SYNTAX:-

```
sns.violinplot(x='categorical', y='continuous',  
hue='categorical_variable', data='your data.')
```

## CODE:-

```
import seaborn as sns  
import matplotlib.pyplot as plt
```

```
tips = sns.load_dataset('tips')
```

```
sns.violinplot(x='day', y='total_bill', hue='sex')
```

## OUTPUT :-

- The Resulting plot will show violins for each of the week, with one half of the violin representing male customers and the other half representing the female customers.
- Each half shows the distributing total values for the corresponding gender.