|  |  |
| --- | --- |
| 10 | **Data Visualization III** Download the Iris flower dataset or any other dataset into a Data Frame.  Scan the dataset and give the inference as: 1. List down the features and their types (e.g., numeric, nominal) available in the dataset. 2. Create a histogram for each feature in the dataset to illustrate the feature distributions. 3. Create a box plot for each feature in the dataset. 4. Compare distributions and identify outliers. |

|  |  |
| --- | --- |
| 1 | **import** numpy **as** np  **import** matplotlib.pyplot **as** plt  **import** pandas **as** pd  **import** seaborn **as** sns  **import** warningswarnings.filterwarnings('ignore')  df **=** pd.read\_csv('iris.csv')  df.head() |
| 2 | df.isnull().sum() |
| 3 | df['PetalLengthCm']**=**df['PetalLengthCm'].fillna(np.mean(df['PetalLengthCm'])) |
| 4 | df.isnull().sum() |
| 5 | 1. **List down the features and their types (e.g., numeric, nominal) available in the dataset**   df.info() |
| 6 | **Hence the dataset contains 4 numerical columns and 1 object column**  np.unique(df["Species"]) |
| 7 | df.describe() |
| 8 | 1. **Create a histogram for each feature in the dataset to illustrate the feature distributions.**   fig, axes **=** plt.subplots(2, 2, figsize**=**(12, 6), constrained\_layout **= True**)  **for** i **in** range(4):  x, y **=** i **//** 2, i **%** 2 axes[x, y].hist(df[df.columns[i **+** 1]])  axes[x, y].set\_title(f"Distribution of {df.columns[i **+** 1][:**-**2]}") |
| 9 | 1. **Create a boxplot for each feature in the dataset.**   \_to\_plot **=** [df[x] **for** x **in** df.columns[1:**-**1]]  fig, axes **=** plt.subplots(1, figsize**=**(12,8))  bp **=** axes.boxplot(data\_to\_plot) |
| 10 | **4. Compare distributions and identify outliers.**   1. **If we observe closely for the box 2, interquartile distance is roughly around 0.75 hence the values lyingbeyond this range of (third quartile + interquartile distance) i.e. roughly around 4.05 will be consideredas outliers. Similarly outliers with other boxplots can be found.** |