**3. Source Code**

**import numpy as np**

**import pandas as pd**

**import matplotlib.pyplot as plt**

**Import seaborn as sns**

**from google.colab import files**

**Data\_files = files.upload()**

**dealer\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Dealers")**

**Customers\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Customers")**

**Claims\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Claims")**

**Parts\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Parts")**

**Transactions\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Transactions")**

**Vendor\_df = pd.read\_excel("Dataset.xlsx", sheet\_name = "Vendors")**

**Parts\_df.head()**

**Claims\_df.head()**

**Customers\_df.head()**

**dealer\_df.head()**

**x = Claims\_df["Dealer\_ID"].value\_counts()**

**x.head(10)**

**Claims\_df["Part\_ID"].value\_counts().head(10)**

**Claims\_df.head()**

**Parts\_df.head()**

**Transactions\_df.head()**

**Vendor\_df.head()**

**mergedclaim\_transaction = pd.merge(Claims\_df,Transactions\_df, how = 'left',on ='claim\_id')**

**mergedclaim\_transaction.head()**

**sns.barplot(x="Part\_ID",y = 'transaction\_amount',hue ='Repair\_or\_Replace',data = mergedclaim\_transaction)**

**mergedclaim\_transaction["Vendor\_Location"] = mergedclaim\_transaction['Dealer\_ID'].map(dealer\_df['City'])**

**mergedclaim\_transaction.head()**

**sns.barplot(y='transaction\_amount',x = 'Vendor\_Location',hue = 'Repair\_or\_Replace',data = mergedclaim\_transaction)**

**mergedclaim\_transaction.groupby("City)['Part\_ID'].value\_counts().head()**

**mergedclaim\_transaction = pd.merge(mergedclaim\_transaction,Parts\_df,how ='left',on ='Part\_ID')**

**mergedclaim\_transaction['Age'] = mergedclaim\_transaction['claim\_date'] - mergedclaim\_transaction['Manf\_Date']**

**mergedclaim\_transaction**

**import datetime**

**mergedclaim\_transaction["Age"] = mergedclaim\_transaction["Age"].apply(lambda x: x.days)**

**sns.violinplot(y='Age',x = 'Part\_ID',hue = 'Repair\_or\_Replace',data = mergedclaim\_transaction)**

**sns.jointplot(y='Age',x = 'Part\_ID',hue = 'Repair\_or\_Replace',data = mergedclaim\_transaction)**

**mergedclaim\_transaction.columns**

**'Vendor\_Location\_Naoida','Vendor\_Location\_Panji']]**

**y=mergedclaim\_transaction['Repair\_or\_Replace\_rpr']**

**from sklearn.tree import DecisionTreeClassifier**

**classifier = DecisionTreeClassifier()**

**cross\_val\_score(classifier,x1,y1,cv=10,scoring='accuracy').mean()**

**from sklearn.ensemble import RandomForestClassifier**

**rfc = RandomForestClassifier(n\_estimators = 100)**

**cross\_val\_score(rfc,x1,y1,cv=10,scoring='accuracy').mean()**

**from sklearn.tree import DecisionTreeClassifier**

**classifier = DecisionTreeClassifier()**

**cross\_val\_score(classifier,x2,y2,cv=10,scoring='accuracy').mean()**