

```

Import pandas as pd

Import numpy as np

# Load data

df = pd.read_csv('credit_card_data.csv')

# Check for missing values

Missing = df.isnull().sum()

# Option 1: Drop rows with missing values

df = df.dropna()

# Option 2: Fill missing values (example using median)

# df.fillna(df.median(), inplace=True)

Print("Missing values handled.")

duplicates = df.duplicated().sum()

df = df.drop_duplicates()

Print(f"{duplicates} duplicate records removed.")

Import seaborn as sns

Import matplotlib.pyplot as plt

# Visualizing using boxplot (for 'Amount' feature as example)

Sns.boxplot(x=df['Amount'])

Plt.title('Outlier Detection in Amount')

Plt.show()

# Option: Remove outliers using IQR

Q1 = df['Amount'].quantile(0.25)

Q3 = df['Amount'].quantile(0.75)

IQR = Q3 - Q1

Df = df[~((df['Amount'] < (Q1 - 1.5 * IQR)) | (df['Amount'] > (Q3 + 1.5 * IQR)))]

Print("Outliers treated.")

# Convert object types to appropriate formats

df['Time'] = pd.to_numeric(df['Time'], errors='coerce')

```

```
df['Amount'] = pd.to_numeric(df['Amount'], errors='coerce')
```

Example: Convert date columns to datetime

```
# df['TransactionDate'] = pd.to_datetime(df['TransactionDate'])
```

```
Print("Data types converted and consistency checked.")
```

```
From sklearn.preprocessing import LabelEncoder, OneHotEncoder
```

```
# Label Encoding example
```

```
# df['Category'] = LabelEncoder().fit_transform(df['Category'])
```

```
# One-Hot Encoding example
```

```
# df = pd.get_dummies(df, columns=['Category'], drop_first=True)
```

```
Print("Categorical variables encoded.")
```

```
From sklearn.preprocessing import StandardScaler
```

```
Scaler = StandardScaler()
```

```
Df[['Amount', 'Time']] = scaler.fit_transform(df[['Amount', 'Time']])
```

```
Print("Features normalized/standardized.")
```