```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
from sklearn.datasets import load_iris
%matplotlib inline

df = pd.read_csv("loan_data_set.csv")

df.head()

df.shape

df.info()

df.describe()

df.isna().sum()

df["ApplicantIncome"].plot(kind="hist")
plt.show()

df["ApplicantIncome"].fillna(df["ApplicantIncome"].mean(), inplace=True)

df["CoapplicantIncome"].plot(kind="hist")
plt.show()

df["CoapplicantIncome"].fillna(df["CoapplicantIncome"].mean(), inplace=True)
```

```
df["LoanAmount"].plot(kind="hist")
plt.show()

df["LoanAmount"].fillna(df["LoanAmount"].mean(), inplace=True)

df["Credit_History"].plot(kind="hist")
plt.show()

df["Credit_History"].fillna(np.random.randint(0,2), inplace=True)

grouped_df = df[["ApplicantIncome", "CoapplicantIncome", "LoanAmount",
"Credit_History"]].groupby(df["Loan_Status"])

mean = grouped_df.mean()
mean

median = grouped_df.median()
median

min = grouped_df.min()
min

max = grouped_df.max()
max

std = grouped_df.std()
std

iris = load_iris()
iris.keys()
```

```
iris_df = pd.DataFrame(iris.data, columns = iris.feature_names)
iris_df["label"] = iris.target

iris.target_names

iris_df.shape

iris_df.head()

iris_df.info()

iris_df.describe()

setosa = iris_df[iris_df["label"] == 0].drop("label", axis=1)

setosa.describe()

plt.figure(figsize=(10,7))
box = setosa.boxplot()
medians = setosa.median()
for i in range(len(medians)):
    box.annotate(medians[i], (i+1, medians[i]), ha="center", va="center", color="red", size=10)
plt.show()

versicolor = iris_df[iris_df["label"] == 1].drop("label", axis=1)

versicolor.describe()

plt.figure(figsize=(10,7))
box = versicolor.boxplot()
medians = versicolor.median()
```

```
for i in range(len(medians)):
    box.annotate(medians[i], (i+1, medians[i]), ha="center", va="center", color="red", size=10)
plt.show()

virginica = iris_df[iris_df["label"] == 2].drop("label", axis=1)

virginica.describe()

plt.figure(figsize=(10,7))
box = virginica.boxplot()
medians = virginica.median()
for i in range(len(medians)):
    box.annotate(medians[i], (i+1, medians[i]), ha="center", va="center", color="red", size=10)
plt.show()
```