In-class exercise on Feb 14: Statistical analysis with iris data

import pandas as pd

import seaborn as sns

iris = sns.load\_dataset('iris')

1. Basic Exploration:

Display the first 5 rows of the dataset.

Display the summary statistics of the dataset using describe().

1. Mean, Median, and Mode:

Calculate the mean of the 'sepal\_length'.

Find the median of 'sepal\_width'.

Determine the mode of 'species'.

1. Variance and Standard Deviation:

Compute the variance of 'petal\_length'.

Calculate the standard deviation of 'petal\_width'.

1. Skewness and Kurtosis:

Find the skewness of the entire dataset.

Determine the kurtosis of the 'sepal\_length'.

1. Quantiles:

Calculate the 25th, 50th, and 75th percentiles of 'petal\_length'.

In-class exercise on Feb 26: Statistical analysis with iris data (2)

1. A brief overview of the Iris dataset with its features:

sepal length, sepal width, petal length, and petal width.

1. Select pairs of variables to examine relationships, for example, sepal length and sepal width, and petal length and petal width.
2. Calculate the covariance matrix for the Iris dataset and the selected pairs.

Discuss the results and what the covariance values indicate about the variable relationships.

1. Compute the correlation matrix.
2. Visualization: create scatter plots for their selected variable pairs using libraries like matplotlib or seaborn; include the correlation coefficients on the scatter plots for better interpretation.