This is an academic project on 'Data Mining and Machine Learning'. The project details is given below:

Use telescope_data.csv for this assignment. The dataset has 10 real attributes, and the last one is simply the class label, which is categorical, and which you will ignore for this assignment. Assume that attributes are numbered starting from 1.

Write Python Code in a Notebook to answer the following questions:

1. Compute the multivariate mean vector.
2. Compute the sample covariance matrix as inner products between the columns of the centered data matrix.
3. Compute the sample covariance matrix as outer product between the centered data points.
4. Compute the correlation between Attributes 1 and 2 by computing the cosine of the angle between the centered attribute vectors. Plot the scatter plot between these two attributes.
5. Assuming that Attribute 1 is normally distributed, plot its probability density function.
6. Plot Empirical CDFs of Attributes,2,5, and7.
7. Which attribute has the largest variance, and which attribute has the smallest variance? Print these values.
8. Which pair of attributes has the largest covariance, and which pair of attributes has the smallest covariance? Print these values.
9. Plot attribute 6 vs attribute 8. Mark each point as a circle.