Exercise 4

The dataset below represents casino customer behavior with two features per customer.

Use principal component analysis to project the data from 2 dimensions to 1 dimension

Show complete solutions

Casino Customers Dataet

Customer	Avg Bet per Visit (USD)	Visits per Month
A	1	2
B	2	3
C	3	4
D	4	5

Tasks

1. Plot the dataset on a 2 dimensional plane. **(5 pts)**
2. Compute the mean of each feature. **(5 pts)**
3. Center the dataset by subtracting the mean from each sample. **(5 pts)**
4. Write down the centered data matrix. **(5 pts)**
5. Compute the sample covariance matrix of the centered data. **(5 pts)**
6. Compute the eigenvalues of the covariance matrix. (25 pts)
7. Compute the corresponding eigenvectors. (25 pts)
8. Normalize the eigenvectors. **(5 pts)**
9. Check if the eigenvectors and eigenvalues are correct using $Av = \lambda v$ **(5 pts)**
10. Identify the **first principal component**. **(5 pts)**
    - Explain why this component is chosen over the others.
11. Project the centered data onto the first principal component. **(5 pts)**
12. Write down the resulting 1-dimensional representation of each sample. **(5 pts)**