

Statistics Project

To analyze customer behavior using statistical and mathematical tools in Python, to identify factors affecting customer retention and churn.

Get the dataset from :

https://drive.google.com/file/d/1VX-LFyvIpQf7T_6IMnco0DKn6-5PtA1q/view?usp=drive_link

Dataset Description:

Assume a dataset with the following columns:

1. CustomerID
2. Gender
3. Region
4. PurchaseAmount
5. ProductCategory
6. Churn (Yes/No)
7. CampaignGroup (A/B)

Questions

1. What is the average, median, and mode of PurchaseAmount?
2. Are there any outliers in the PurchaseAmount data?
3. Is there any skewness or kurtosis in the PurchaseAmount distribution?
4. Is there a significant difference in spending between male and female customers?
5. Is there a relationship between ProductCategory and customer churn?
6. Does PurchaseAmount vary significantly across different regions?
7. Which email campaign (A or B) performed better in terms of average PurchaseAmount?
8. Can we assume PurchaseAmount follows a normal distribution?
9. What insights can we gain by applying the Central Limit Theorem?
10. What is the 95% confidence interval for the average PurchaseAmount?