**Worksheet Questions for Churn Modelling Dataset**

**Dataset Understanding**

1. How many rows and columns does the dataset have?
2. What are the data types of each column?
3. Are there any missing values in the dataset? If yes, in which columns?
4. Which columns are identifiers (not useful for prediction)?
5. What is the target variable in this dataset?

**Univariate Analysis**

1. What is the distribution of the target variable **Exited**?
2. What is the most frequent **Geography** among customers?
3. What is the distribution of **Gender** in the dataset?
4. Plot the distribution of **Age**. What do you observe?
5. What is the distribution of **CreditScore** values?
6. How many customers have a **HasCrCard** value of 1 and 0?
7. What is the distribution of **IsActiveMember**?
8. Plot the distribution of **Tenure**. What do you observe?
9. What is the distribution of **EstimatedSalary**?
10. What is the range of **Balance**?

**Bivariate Analysis**

1. What is the churn rate among males vs females?
2. Does churn differ across different **Geography** values?
3. Compare the **Age** of customers who churned vs who stayed.
4. Compare the **Balance** of customers who churned vs who stayed.
5. Compare the **CreditScore** of churned vs retained customers.
6. Does **Tenure** have any visible relationship with churn?
7. Compare churn rates between customers with and without a credit card.
8. Compare churn rates between active and inactive members.
9. Does the number of products (**NumOfProducts**) affect churn?

**Multivariate Analysis**

1. Create a pairplot of **CreditScore, Age, Balance, EstimatedSalary** with respect to churn. What do you observe?
2. Which numerical variables seem most related to churn visually?
3. Do older customers with high balance churn more compared to younger customers?
4. Do customers with low credit score and high balance churn more?
5. Does the combination of **Geography and Gender** show any pattern in churn?
6. Which group of customers (by **Age + Geography + IsActiveMember**) appears most likely to churn?