Customer Segmentation Project - Interview Q&A

1. What was the objective of your Customer Segmentation project?

The objective was to segment customers based on their purchasing behavior using RFM (Recency, Frequency, and Monetary) analysis and apply K-Means clustering to group them into actionable segments. This allowed the marketing team to run personalized campaigns, improve retention, and increase ROI.

2. Why did you use RFM analysis in this project?

RFM analysis helps quantify customer value. Recency tells us how recently a customer purchased, Frequency shows how often they buy, and Monetary indicates how much they spend. These three metrics give a clear behavioral snapshot that's highly effective for segmentation.

3. How did you choose the number of clusters for K-Means?

I used the Elbow Method, which involves plotting the within-cluster sum of squares (WCSS) for different values of `k`. I selected the `k` value where the curve began to flatten, indicating the optimal number of clusters.

4. What challenges did you face in this project, and how did you handle them?

One challenge was dealing with duplicate and cancelled transactions in the dataset. I handled it by filtering out returns and zero-quantity records. Another challenge was ensuring the number of clusters didn't exceed the number of customers - I added dynamic checks in the code to avoid that error.

5. How did this project add business value?

The project helped identify VIP customers, at-risk customers, and low-spending occasional buyers. This allowed the business to run more targeted marketing campaigns, focus retention efforts where needed, and increase customer lifetime value.

6. Why did you choose K-Means over Hierarchical Clustering or DBSCAN?

I chose K-Means because it's efficient for large datasets and works well when the number of clusters is known or can be estimated. Since I could apply the Elbow Method to determine optimal `k`, it was suitable for this business case.

7. How did you handle scaling and normalization?

Since RFM features are on different scales, I used `StandardScaler` to normalize them. This ensures that no single feature dominates the clustering algorithm, which is sensitive to feature magnitude.

8. How did you validate or evaluate the clustering results?

Customer Segmentation Project - Interview Q&A

I used visual validation by applying PCA to reduce the data to two dimensions and plotting the clusters. I also profiled each cluster based on RFM averages to ensure they made business sense.

9. Can you explain the logic behind your cluster labels?

Yes. After clustering, I calculated the average Recency, Frequency, and Monetary values for each group. Based on these, I labeled segments like VIPs, At-Risk, etc., in a language the marketing team could act on.

10. How would you update this project if new customer data arrived daily?

I would automate the RFM feature generation and retrain the clustering model periodically. Alternatively, I could build a pipeline that appends new data, recalculates RFM, and reassigns cluster labels using a pre-trained KMeans model.

11. What did you learn from this project?

I learned how to bridge technical models with business outcomes. It taught me to convert raw purchase data into customer intelligence and communicate results clearly.

12. How would you use this segmentation in a real company setting?

I would integrate the cluster labels into the company's CRM. Campaigns can then be personalized - VIPs get loyalty programs, at-risk users receive win-back discounts, and new buyers get onboarding emails.

13. If your manager said the segments were not useful, what would you do?

I would revisit the cluster profiles and validate if they align with key business goals. I might also interview stakeholders and adjust features or try alternative clustering methods.

14. What would you improve in your current approach?

I would explore more advanced clustering techniques like Gaussian Mixture Models or use more features such as product category, time of day, or seasonality.

15. How does this project demonstrate your business analytics skills?

It shows that I can turn raw transactional data into strategic customer insights using data-driven techniques, and present it in a business-friendly way.