



Customer Retention & Cohort Analysis

Exploratory Data Science Project

Introduction & Objectives



- UNDERSTAND
CUSTOMER
RETENTION TRENDS



- BUILD COHORTS
AND MEASURE
RETENTION



- SEGMENT
CUSTOMERS USING
RFM ANALYSIS



- PROVIDE DATA-
DRIVEN BUSINESS
RECOMMENDATIONS

Dataset & Preparation

- Fields: InvoiceNo, InvoiceDate, CustomerID, Quantity, UnitPrice, Country

- Data cleaning: remove missing IDs, convert datatypes

- Feature engineering: TotalPrice, InvoiceMonth, CohortMonth, CohortIndex

Cohort Analysis – Methodology



- CUSTOMERS GROUPED BY FIRST PURCHASE MONTH (COHORTMONTH)



- COHORTINDEX = MONTHS SINCE FIRST PURCHASE



- RETENTION RATE = ACTIVE CUSTOMERS ÷ INITIAL COHORT SIZE

Cohort Analysis – Key Findings



- Retention drops sharply after month 1 (20–30%)

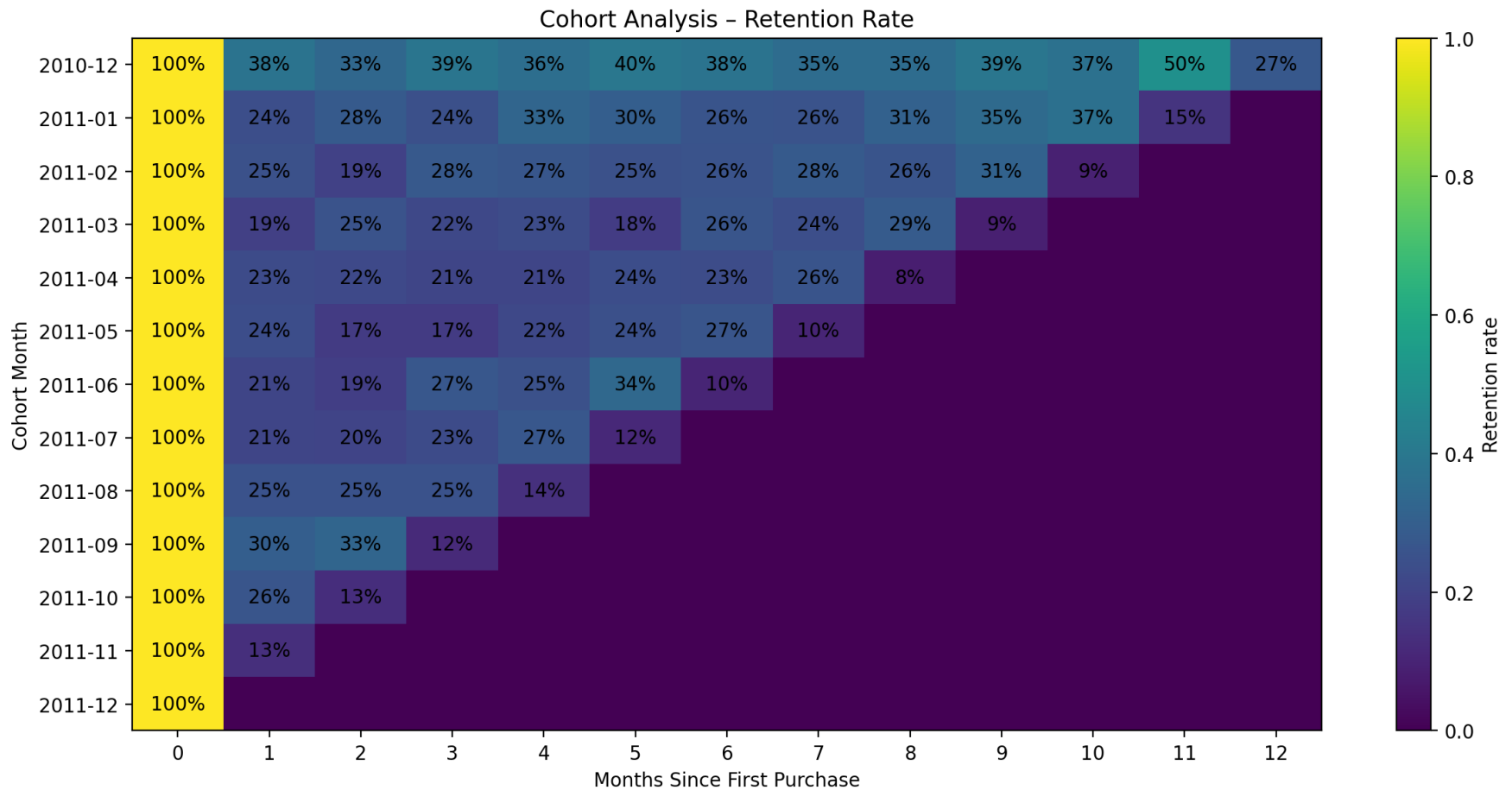


- Few customers remain active after 6 months

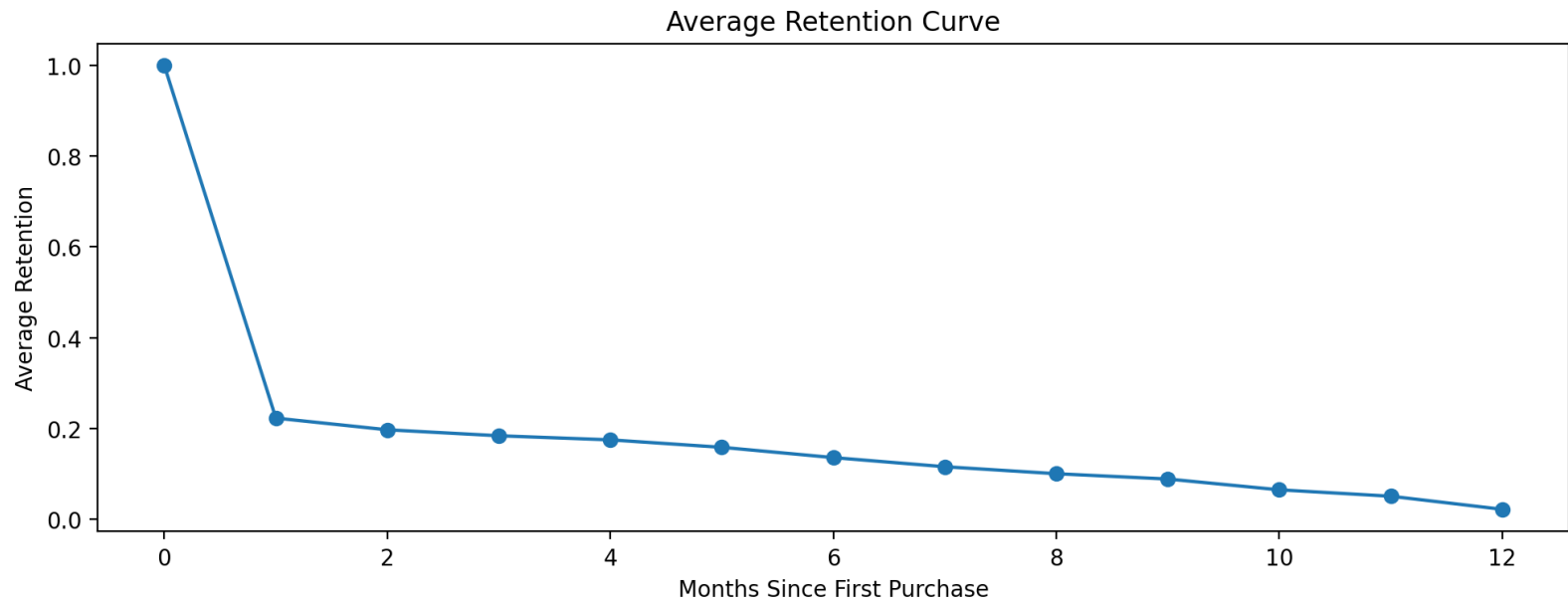


- Early cohorts showed stronger retention than later cohorts

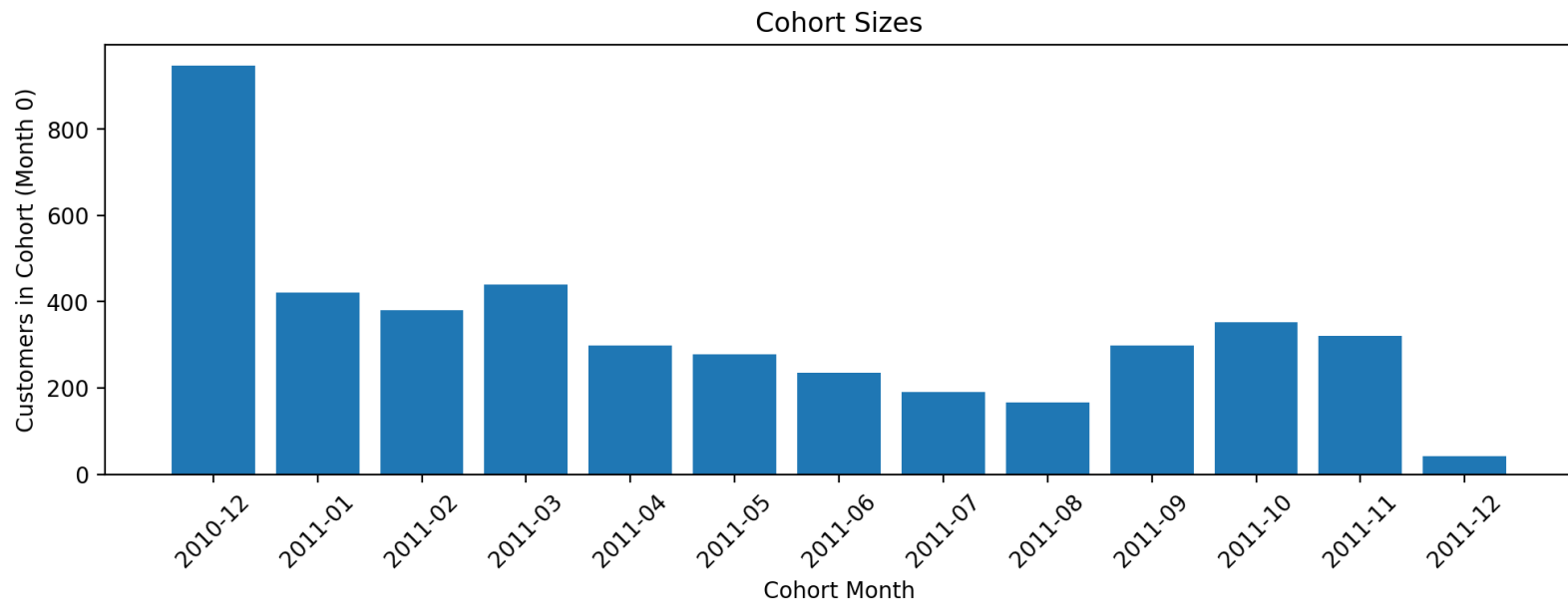
Retention Heatmap



Average Retention Curve



Cohort Sizes



RFM Segmentation – Methodology

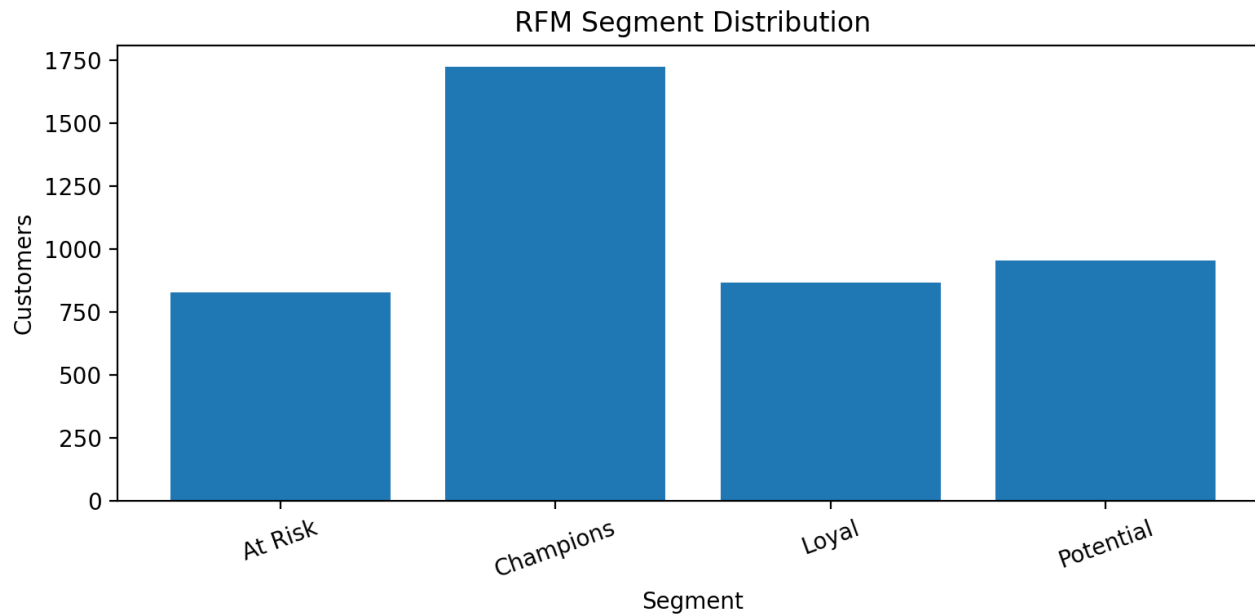
- Recency: days since last purchase

- Frequency: total transactions

- Monetary: total spend

- Customers scored 1–4 per metric → Combined RFM Score

RFM Segment Distribution



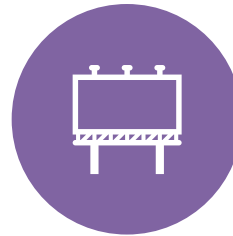
Business Insights



- RETENTION DROPS QUICKLY → STRENGTHEN ONBOARDING



- LOYAL CUSTOMERS DRIVE REVENUE → REWARD THEM



- MANY POTENTIAL CUSTOMERS → TARGET WITH ENGAGEMENT CAMPAIGNS



- AT RISK CUSTOMERS NEED REACTIVATION STRATEGIES

Recommendations



- Personalized onboarding campaigns



- Loyalty and rewards programs



- Win-back offers for at-risk customers



- Data-driven marketing & segmentation strategies

Conclusion

Cohort and RFM analysis revealed clear retention challenges.

By implementing targeted strategies, businesses can improve retention,

increase customer lifetime value, and achieve sustainable growth.