AUA American University of Armenia

Report II

Data Visualisation: IESM 343

Revenue Trends Across Age and Gender: An Interactive Analysis of Cardholder Transactions

Professor: Gurgen Hovakimyan
E-mail: ghovakimyan@aua.am

Team:	**E-Mail:**
Arman Manukyan	arman_manukyan@edu.aua.am
Mariam Manukyan	mariam_manukyan@edu.aua.am

Introduction

Purpose and Objectives of the Project

The primary objective of this project is to analyze consumer behavior using transactional sales data from cardholders. We aim to identify patterns and trends based on gender, age, purchase time, and product types. The insights are visualized in an interactive dashboard to support informed business decisions such as targeted promotions, customer segmentation, and time-based campaign planning.

Dataset and Preprocessing

The dataset used for this project includes **370 cardholder transactions** across **41 unique customers**. Each record contains information on gender, age, product purchased, date and time of purchase, revenue, and quantity. Using R, we performed several data preparation steps:

- Converted raw dates and times into standard formats using lubridate
- Extracted derived features: **Hour**, **Weekday**, and **Month**
- Grouped customers into **age bands** (e.g., 0–18, 19–25, etc.)
- Removed or handled missing and anomalous entries
- Computed summary statistics such as **mean**, **standard deviation**, **skewness**, and **kurtosis**

This prepared dataset allowed us to conduct a multi-dimensional analysis across demographics, time, and revenue.

Understanding this data provides valuable insight into **when, who, and how much** consumers spend, which is crucial for revenue optimization and customer experience strategies.

	Gender	Birthdate	Age	Product	Date	Time	Quantity	Revenue	Datetime	Hour	Weekday	Month
1	Женский	1998-01-28	27	16 հեքիաթ շարքից՝ Սուտասանը	2023-08-14	21:58:17	1	288	2023-08-14T17:58:17Z	21	Monday	2023-08
2	Женский	2005-07-27	20	Լատիներենի դասագիրք	2023-12-03	22:02:14	1	1406.5	2023-12-03T18:02:14Z	22	Sunday	2023-12
3	Мужской	1990-01-04	35	Կոշկավոր կատուն - համայնապատկեր	2023-03-26	21:59:52	1	1447.5	2023-03-26T17:59:52Z	21	Sunday	2023-03
4	Женский	1992-08-11	33	Մայրենի 3-1	2023-03-09	22:05:12	1	1450	2023-03-09T18:05:12Z	22	Thursday	2023-03
5	Женский	1985-03-05	40	Հանճարեղ մեխանիկը	2023-11-24	22:44:53	1	1462.5	2023-11-24T18:44:53Z	22	Friday	2023-11
6	Женский	2005-07-27	20	Երկիր անապատ	2023-04-05	22:08:49	1	1611.55	2023-04-05T18:08:49Z	22	Wednesday	2023-04
7	Женский	1990-08-20	35	Ընթերցանության վարժարան 1-2-րդ դասարաններ	2023-04-17	21:59:05	1	1640.5	2023-04-17T17:59:05Z	21	Monday	2023-04
8	Женский	1981-03-03	44	Մաթեմատիկա 1 Առաջին մաս	2023-08-31	22:11:35	1	1657.4	2023-08-31T18:11:35Z	22	Thursday	2023-08
9	Женский	1990-02-13	35	Երկարագուլպա Պիպին	2023-02-19	22:09:32	1	1680.32	2023-02-19T18:09:32Z	22	Sunday	2023-02
10	Женский	1984-01-21	41	Վիշապ. Տուրբոկնոջ հարձակումը	2023-05-29	17:11:29	1	1714.37	2023-05-29T13:11:29Z	17	Monday	2023-05

Showing 1 to 10 of 370 entries Previous 1 2 3 4 5 … 37 Next

Figure 1. Dataset final_structured_cardholders1

Justification for Visualization Choices

- **Bar plots** were chosen to compare average revenue across gender and age groups.
- **Density plot** was used to understand the skewness in revenue distribution.
- **Line charts** show revenue trends over months and gender-wise comparisons.
- **Boxplots** visualize variability in revenue by age group.
- **Histograms** provide intuitive overviews of time-of-day and weekday purchase patterns.
- **Interactive bar chart (ggplotly)** ranks top products using revenue and unit price.

Findings & Insights

Revenue Distribution and Descriptive Statistics

We first explored how revenue values are distributed. The results showed a **highly right-skewed distribution** (Skewness = 5.29), with a **long tail of high-revenue transactions**, as shown in the **Revenue Density Plot** . The **kurtosis of 42.24** confirms the presence of extreme outliers.

Statistical Summary:

- **Mean Revenue**: 4,977.81
- **Median**: 4,132.50
- **Max Revenue**: 42,098\
- **Coefficient of Variation**: 0.76

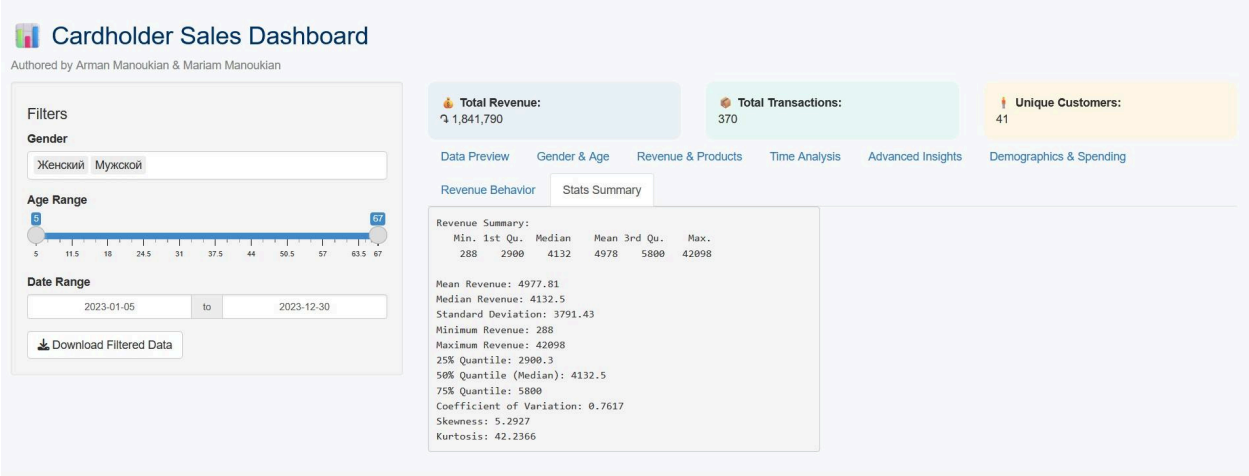

Figure 2. Statistical Summary

The boxplot by age group (Figure 3) supports this by highlighting large revenue variability in age ranges 31–50, which contain several high-value purchases.

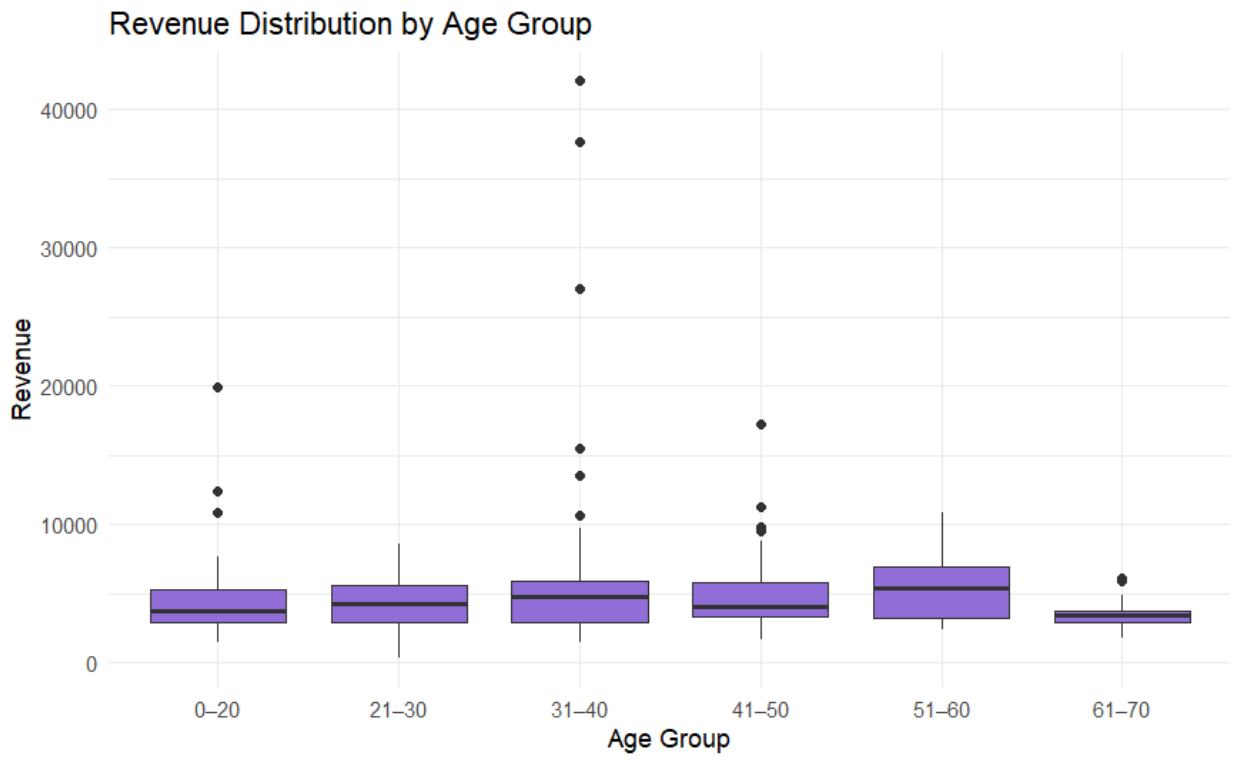

Figure 3. Revenue Distribution by Age group

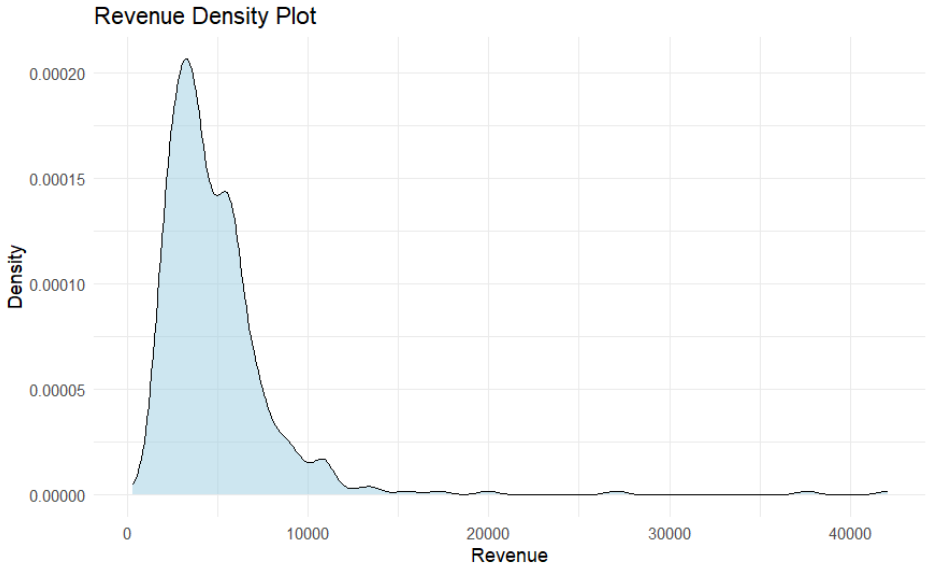

Figure 4. Revenue Density plot

Demographic Insights: Gender and Age Behavior

A large portion of transactions were made by **female customers (≈84%)** (Figure 5), which initially suggests higher engagement. However, deeper analysis revealed that **male customers** had a **higher average revenue per transaction** (Figure 6), indicating fewer but higher-value purchases.

The **age-based revenue analysis** (Figure 7) shows that customers aged **36–45** spend the most on average, followed by **56–65**, making these prime segments for premium marketing. In contrast, the **65+ segment** spends significantly less.

The **age histogram** (Figure 8) showed a bimodal distribution, with activity peaks around **20–30** and **40–50 years** — possibly representing younger adults and mid-career individuals.

Figure 5. Gender Distribution

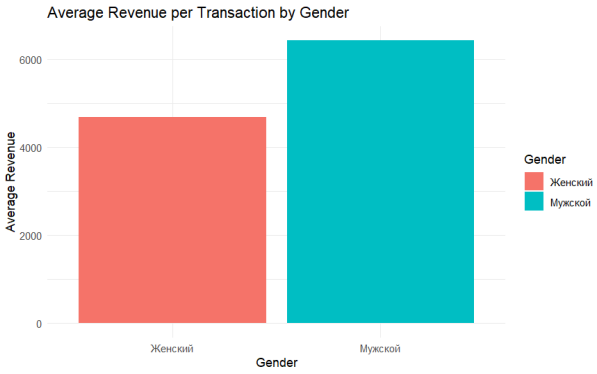

Figure 6. Average revenue per transaction

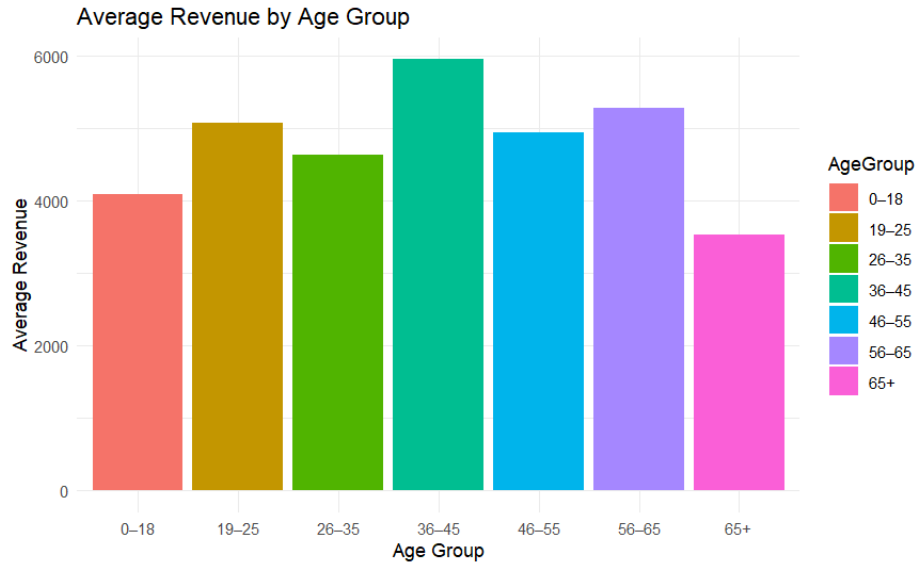

Figure 7. Average revenue by Age group

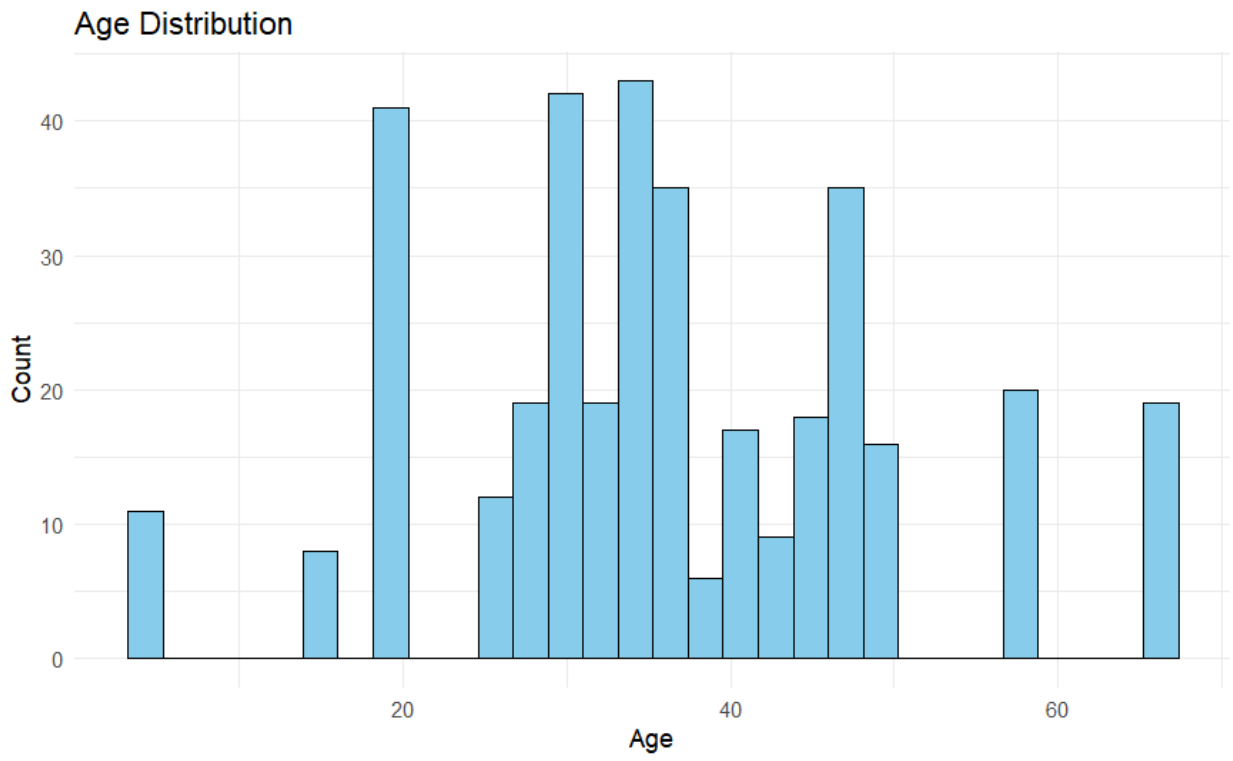

Figure 8. Age Histogram

Time-Based Behavior: Hour and Weekday Patterns

Using temporal attributes, we uncovered clear trends in **when** customers are most active:

- **Hourly Purchases** peaked at **21:00–23:00** (Figure 9), with 22:00 being the busiest time.
- **Weekend Shopping** dominated, especially **Sundays and Saturdays** (Figure 10), suggesting leisure-time browsing or payday effects.

These patterns provide actionable insights for scheduling **discount campaigns** or **advertising windows**.

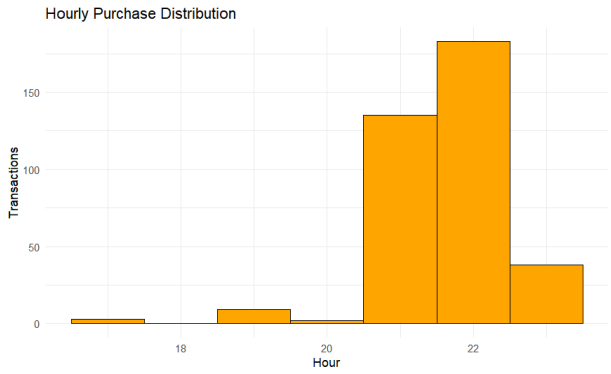

Figure 9. Hourly Purchases

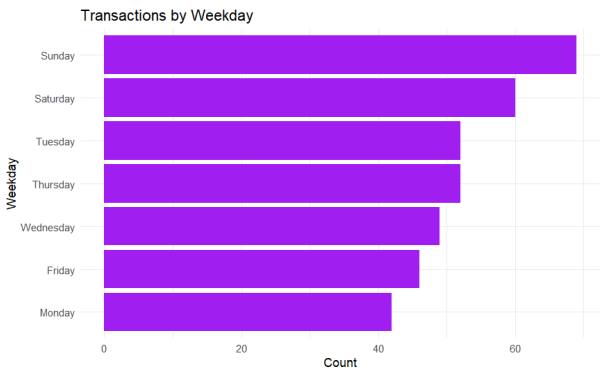

Figure 10. Age Distribution

Monthly Trends and Gender Comparison

Figure 11 shows total revenue fluctuating moderately throughout the year, with a **significant spike in December**, likely driven by holiday shopping. This seasonal effect was stronger among **female customers** (Figure 12), which suggests that gender-focused seasonal promotions could yield strong results.

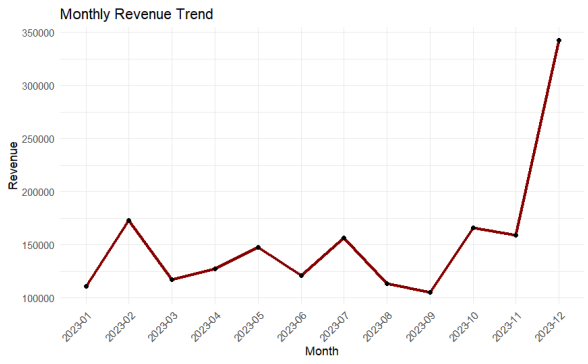

Figure 11. Monthly Revenue Trend

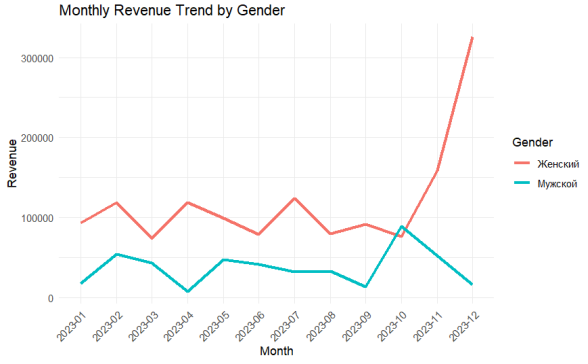

Figure 12. Monthly Revenue Trend by Gender

Product-Level Insights

In Figure 13, we visualized the **Top 15 products by revenue**, also coloring the bars by average price. This dual encoding helped identify both **best sellers** and **premium-priced items**. Some products are bought frequently at moderate prices, while others contribute large revenues through **fewer but high-value purchases** — a critical insight for pricing and inventory decisions.

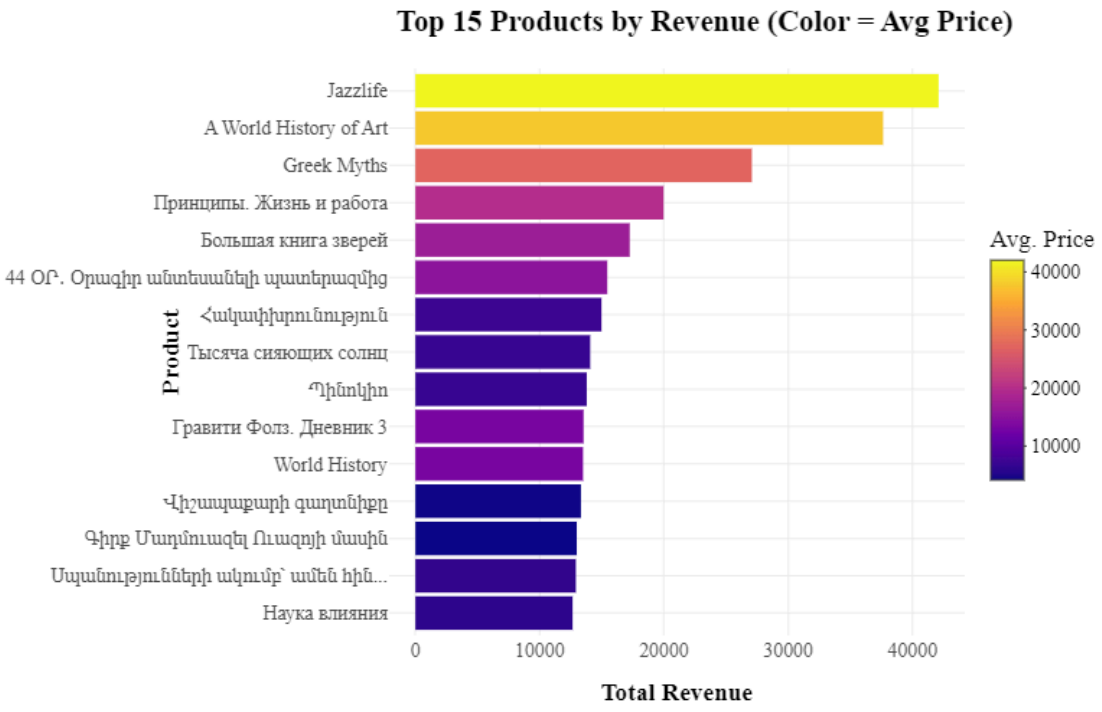

Figure 13. Top 15 products by revenue

Summary and Business Implications

This analysis highlights several actionable insights:

- **Target high-revenue customers (male, age 36–55)** for upselling.\
- **Promote during evening hours and weekends** when purchase activity peaks.
- **Use seasonal spikes**, such as December, for time-sensitive marketing.

- Identify and promote **top-performing products** based on unit price and sales volume.
- Consider loyalty or referral campaigns for **female users**, who shop frequently but at lower per-transaction values.

Tools and Visualization Techniques

We developed all visualizations in **R**, using packages such as:

- ggplot2, dplyr, plotly, viridis, lubridate, moments and etc.
- Visuals included: bar plots, line graphs, histograms, boxplots, density plots, and interactive charts

Each visualization was selected to clearly support the insight being discussed and was styled for accessibility, contrast, and minimalism.

Conclusion

Through visual analytics, we revealed how demographic and temporal factors affect spending behaviour. Key takeaways include:

- **Males aged 36–55** are fewer in number but offer high revenue potential\
- **Evening and weekend times** represent optimal moments for marketing
- **Holiday seasons**, especially December, drive revenue spikes
- **Outlier purchases** have significant impact and should be studied further
- **Top products** offer strategic opportunities for upselling or bundling

We recommend launching **personalized campaigns targeting high-spending segments**, aligning promotions with time-based behavior, and leveraging top-performing products in loyalty programs.

References

- Chang, W. (2018). *R Graphics Cookbook: Practical Recipes for Visualizing Data*. O'Reilly Media.
- Wickham, H. (2021). *Mastering Shiny*. Retrieved from https://mastering-shiny.org
- Mane Davtyan (2024). *Enhancing Book Recommendations in Local Bookstores: A Case Study of Zangak Bookstore*. American University of Armenia. https://cse.aua.am/files/2024/05/Enhancing-Book-Recommendations-in-Local-Bookstores-A-Case-Study-of-Zangak-Bookstore.pdf