Analysis and Review of Trends in Revenue for a Financial Services Company in 2022

Naida Dzigal, June 2023
dzigal_naida@student.ceu.edu
Central European University
Vienna, Austria

Background 2
Data Analysis Questions 2
Data Processing 2
Revenue Analysis 3
Discussion and Key Findings 3
Conclusion 4

Abstract

This work provides an overview of a capstone project, conducted from April-June 2023 as part of the CEU Masters in Business Analytics. The aim of the work was to demonstrate advantages of using R for data analytics, without compromising on the company's operational and analysis needs. In general, R offers better capabilities in handling large datasets, conducting statistical analysis, and ensuring the reproducibility of results. By adopting R, the company aimed to streamline its data analysis process, uncover hidden insights, and make more informed business decisions and recommendations to its clients.

The project involved creating a prototype data analysis pipeline that took .txt documents as input and generated a comprehensive report on key performance metrics, insights, and trends of a client's revenue streams, with a focus on the year 2022. The pipeline included data cleaning, transformation, and visualization steps, leveraging advanced data analytics tools based in R. This transition to a more data-savvy ecosystem aimed to help the company stay ahead of the competition in its industry.

Background

In today's data-driven business landscape, companies are constantly seeking ways to leverage advanced data analytics tools to gain insights and make informed decisions. This paper summarizes a project conducted by the student at a mid-level consultancy company that aimed to map out and analyze commission revenue streams for a client company over a five-year period (2018-2022). The project focused on transitioning one type of analysis from Excel to R, a powerful statistical programming language, to overcome the limitations of Excel in handling large datasets and enhance the company's data analysis process.

It is worth mentioning that the visualizations showcased in the following sections serve the purpose of illustrating how the fundamental analysis capabilities in Excel can be replicated and improved upon using R. For confidentiality reasons, the data used in this report has been fully anonymized. During the anonymization process, two data files were inadvertently corrupted. Rather than rectifying this issue in the technical report, it was decided to retain the data in its current state (see Figure 6 depicting January 2020 and August 2021 show large deviations from the monthly revenue trend lines). Additionally, a key objective of this project was also to showcase the complete automation of the revenue analysis process.

Data Analysis Questions

The project sought to answer several important analysis questions about the client's revenue, including:}

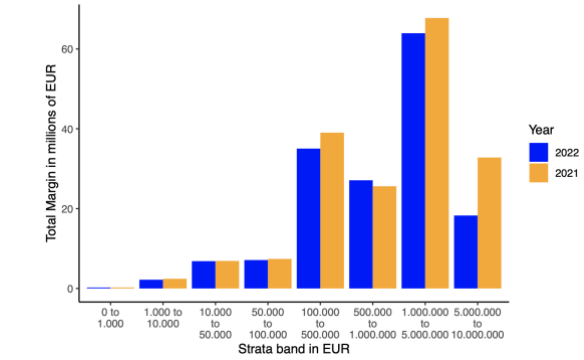
1. What is the total transaction volume?
2. What is the total transaction volume and margins per creditor?
3. Are there any differences in transaction volumes and margins between 2021 and 2022?
4. How is the revenue distributed by creditor group?
5. What is the total revenue per consistency unit?
6. Who are the top customers in each consistency unit?
7. Are there any significant differences in monthly revenue, commission rates, and revenue by margin level?
8. What is the average income for the client per transaction in euros?

Data Processing

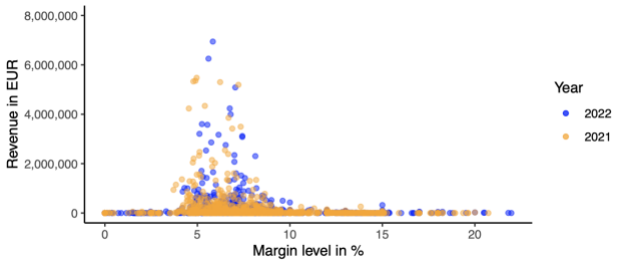
To ensure the data was in the right format for analysis, the project involved data loading, cleaning, and preprocessing steps. The data, received in .txt format, consisted of multiple files representing each month for 5 years (12 files per year). Key columns included date, month, year, merchant currency, creditor, amount, margin, and quantity. Additional files containing merchant names, creditor codes, and currency conversion rates were also used. R packages such as data.table, stringr, tidyr, lubridate, and dplyr were employed to clean, format, and also convert the data to a common currency (Euro).

Revenue Analysis

The revenue analysis phase focused on grouping the data by various attributes such as merchant, strata band, and consistency unit. This grouping allowed for an examination of transaction volumes, margins, revenue trends, and customer profiles. The analysis revealed that a significant portion of revenue was generated by a few large merchants, with the composition of the merchant portfolio remaining relatively stable over time. The analysis also identified trends in revenue, commission rates, and average transaction values, providing valuable insights for strategic decision-making.
See the below figures for summary findings on: the total income margin per customer group (Figure 1); Total revenue in EUR per merchant per margin level in percent comparing 2022 (blue) and 2021 (orange) (Figure 2); Total Revenue per Consistency Unit in 2022 (Figure 3); Total Revenue per Consistency Unit in 2021 (Figure 4); Treeplot showing the (anonymised) largest 20 customers in 2022 (Figure 5); Monthly revenue for the first 8 months of 2020, 2021 and 2022 (Figure 6); Linear model of average euros charged by Company from merchants from 2018-2022 (Figure 7).



CU Merchant	# of Creditors 2022	EUR Margin 2022	% Total
0001	909 856	35 120 481	21.85
0002	1 484	0	0.00
0003	205 927	10 076 768	6.27
0004	4 920	0	0.00
0005	222	3	0.00
0006	1 415	8 672	0.01
0007	283	19	0.00
0008	506	0	0.00
0009	2 274	0	0.00
0010	528 771	114 833 708	71.44
0011	17 941	705 875	0.44



CU Merchant	# of Creditors 2021	EUR Margin 2021	% Total
0001	979 797	39 914 180	21.91
0002	1 608	0	0.00
0003	176 004	8 298 859	4.56
0004	5 321	0	0.00
0005	218	0	0.00
0006	1 085	3 586	0.00
0007	288	25	0.00
0008	488	0	0.00
0009	2 731	0	0.00
0010	602 916	133 832 181	73.46
0011	4 379	137 856	0.08

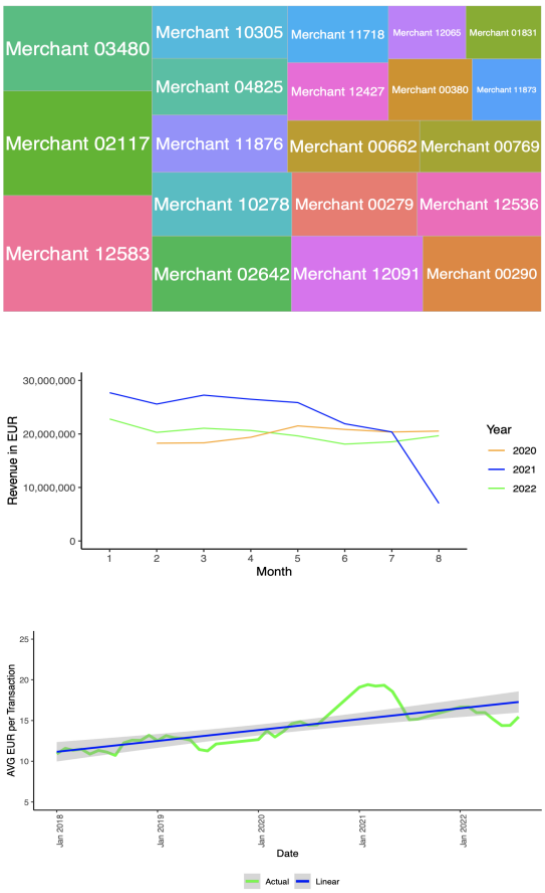
Top left (Figure 1): Total revenue per strata band (customer group); bottom-left (Figure 2): Total revenue in euros per merchant per margin level in percent comparing 2022 (blue) and 2021 (orange); top right (Figure 3): Total Revenue per CU in 2022; bottom right (Figure 4): Total Revenue per CU in 2021

Discussion and Key Findings

The analysis conducted aimed to address several important questions pertaining to the client's business performance. Here is a summary of the key questions explored and their corresponding findings:

1. Total transaction volume and margins per creditor: The data was grouped and summarized to reveal that the total transaction volume in 2022 amounted to 160 million EUR, a decrease from 182 million EUR in 2021.
2. Differences in transaction volumes and margins between 2021 and 2022: The revenue distribution across strata bands exhibited a consistent trend for both years, except for the 5-10 million euro band, which experienced a significant decline in revenue. (Figure 1)
3. Revenue distribution by creditor groups: The analysis showed that the 1-5 million EUR group generated the highest revenue, contributing 30\% of the total. (Figure 1)
4. Difference in Total Revenue per Consistency Unit between 2021 and 2022: The analysis identified variations in total revenue between the years, with Consistency Unit 0010 being the most profitable (Figures 3 and 4).

5. Top 20 customers in each Consistency Unit: The top customers for each Consistency Unit were identified, including Merchant 12583 and Merchant 02117 for CU 0010, Merchant 03480 for CU 0003, and Merchant 02642 for CU 0001 (Figure 5).
6. Monthly revenue differences: With the exception of the COVID-19 lockdown period, where changes in consumption behavior impacted revenue, the overall revenue trends remained consistent across 2020, 2021, and 2022 (Figure 6).
7. Client's most typical commission rate: Contracts with commission rates ranging from 6.99 to 7.99 EUR per transaction generated the highest revenue and were the most commonly charged commission rates. This corresponds to a margin level between 4 and 10 percent (Figure 2).
8. Lowest commission rates and associated customers: Analysis revealed that transactions with the lowest commission rates (0-3.99 EUR) were predominantly associated with businesses based in one particular country, likely due to specific regulations governing charges by online stores in that country.
9. Revenue by margin level: Transactions charging 4-10 EUR exhibited the highest revenue, with no significant differences observed between 2021 and 2022 (Figure 2).
10. Average amount of euros charged per transaction: The average EUR per transaction displayed a steady growth trend, increasing from approximately 11.5 EUR in 2018 to nearly 15 EUR by the end of 2022 (Figure 7).



Top (Figure 5 in text): Treeplot showing the (anonymized) largest 20 customers in 2022. Note that the size of the area of each rectangle corresponds to the total margin generated per customer.
Middle (Figure 6): Monthly revenue for the first 8 months of 2020, 2021 and 2022. Note that during the anonymization process, the .txt files for January 2020 and August 2021 were corrupted.
Bottom (Figure 7): Linear model (blue) of average euro charged by the client company (green) from merchants from 2018-2022.

These findings provide valuable insights into various aspects of the client's business, including transaction volumes, revenue distribution, customer behavior, and commission rates. They can assist the client in making informed decisions, optimizing their operations, and identifying opportunities for growth and improvement such as contacting customers whose revenue income has dropped, renegotiating commission contracts, etc.

Conclusion

In conclusion, the project successfully demonstrated the benefits of transitioning from Excel to R for data analytics in a mid-level consultancy company. This not only improved the efficiency and accuracy of data analysis but also enabled the company to delve deeper into the data, uncovering valuable insights that were previously inaccessible. This enabled the company to handle large datasets, conduct advanced statistical analysis, ensure reproducibility, and uncover valuable insights in the client's revenue streams. The prototype data analysis pipeline developed in this project can serve as a foundation for future data analysis endeavors, empowering the company to provide more informed business recommendations and maintain a competitive edge in the industry.