Final Report

This report can help the store with its employee promotion decision, customer behavior analysis, and performance indicators for the near future. The integration of both Python and SQL is used to generate statistical summaries and data visualizations which help us make the final decision in each case.

The store needs to promote one current employee to the manager position. The best fit would be someone who has the overall most experience. Therefore, by analyzing metrics such as how many sales each sales person makes and how much revenue each of them brings we can potentially find the candidate. This process starts with joining the customers, invoices, and employees tables together. From there we group the data by employee and sum up the total transactions for each one of them. The final table output shows that Jane Peacock should be promoted. The result is shown in Appendix 1.

After the promotion, we are interested in knowing how often customers visit the store. We take the data from the invoices table first, group them by customers, sort the invoices by dates and calculate the interval between each visit as frequency. The result of the histogram indicates that the majority of customers will return and purchase the item from our store around every three months or eight months. The result is shown on Appendix 2.

One extra metric that we think will be beneficial to track for the store is the popularity of different music genres. In this case, we can quantity this by measuring the total sales amount for each music genre. We first join invoice_items, tracks, genres, and albums table together. From there we group the new data by genre and sum up each employee's total sales amount. The result shows that rock is extremely popular being the best selling genre. The second to fourth places are

Latin, metal, alternative & punk. Appendix 3.1 shows the result for the most popular genre for top 50 best-selling albums in the store.

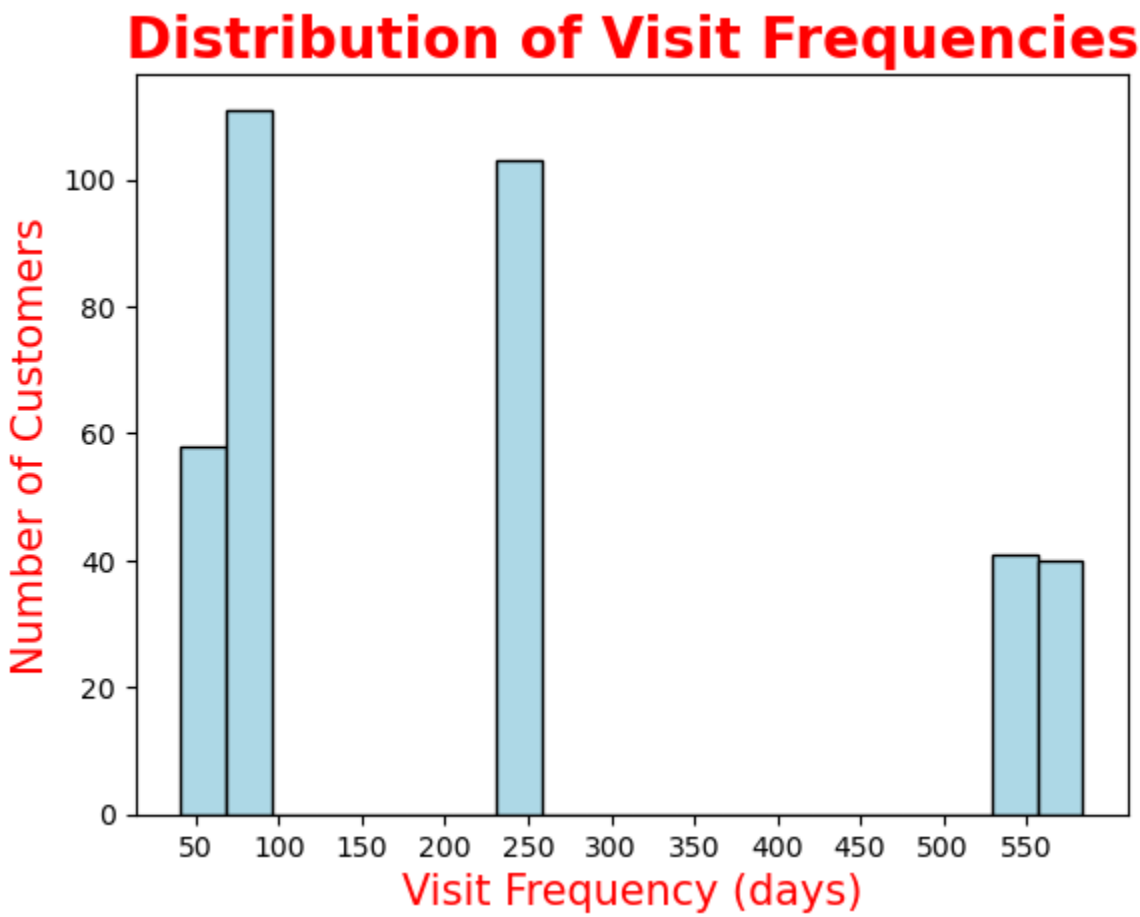
Lastly, we need to solve the specific SQL problem on the purchase amount and personal information of those specific customers (ranking 2, 3, 5, 8, or 12). This is done by calculating the total spending for each customer, then ranking them based on their spending and retrieving details of customers who hold specific spending ranks for further analysis or decision-making purposes. Richard Cunningham, Luis Rojas, Julia Barnett, Terhi Hamalainen are ranked from second to fifth in purchasing while Luís Gonçalves, François Tremblay, Bjørn Hansen, Dan Miller, Heather Leacock, and Wyatt Girard tied for twelfth highest on the list. Their name, address, phone, email, total amount, and spending ranks are shown in Appendix 4.

Appendix

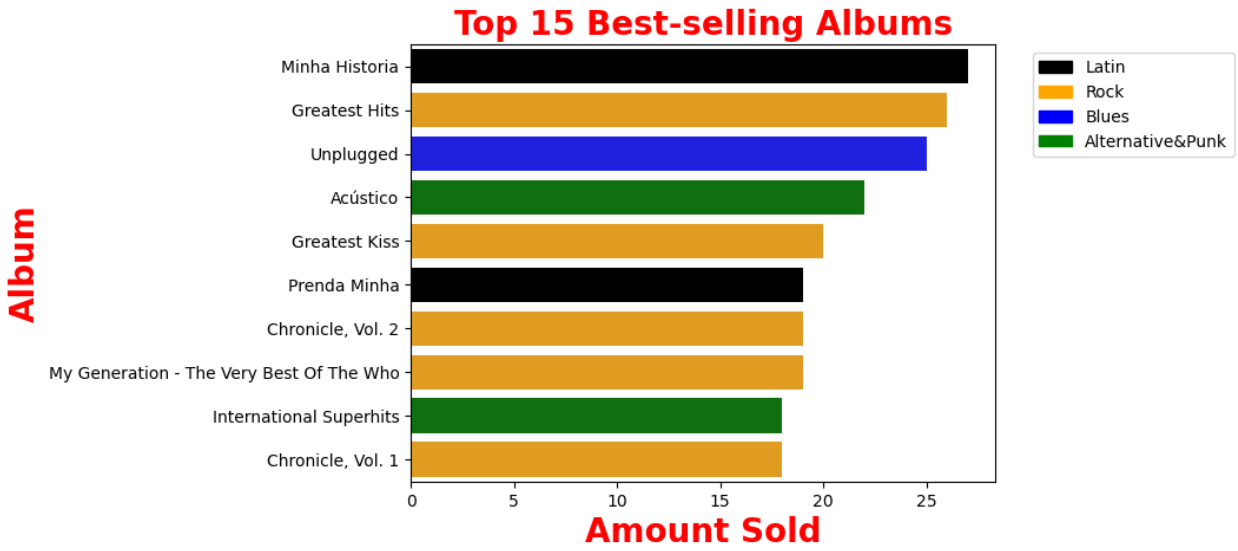
Appendix 1

	EmployeeId	FirstName	LastName	Total_Transaction	Total
0	3	Jane	Peacock	146	833.04
1	4	Margaret	Park	140	775.40
2	5	Steve	Johnson	126	720.16

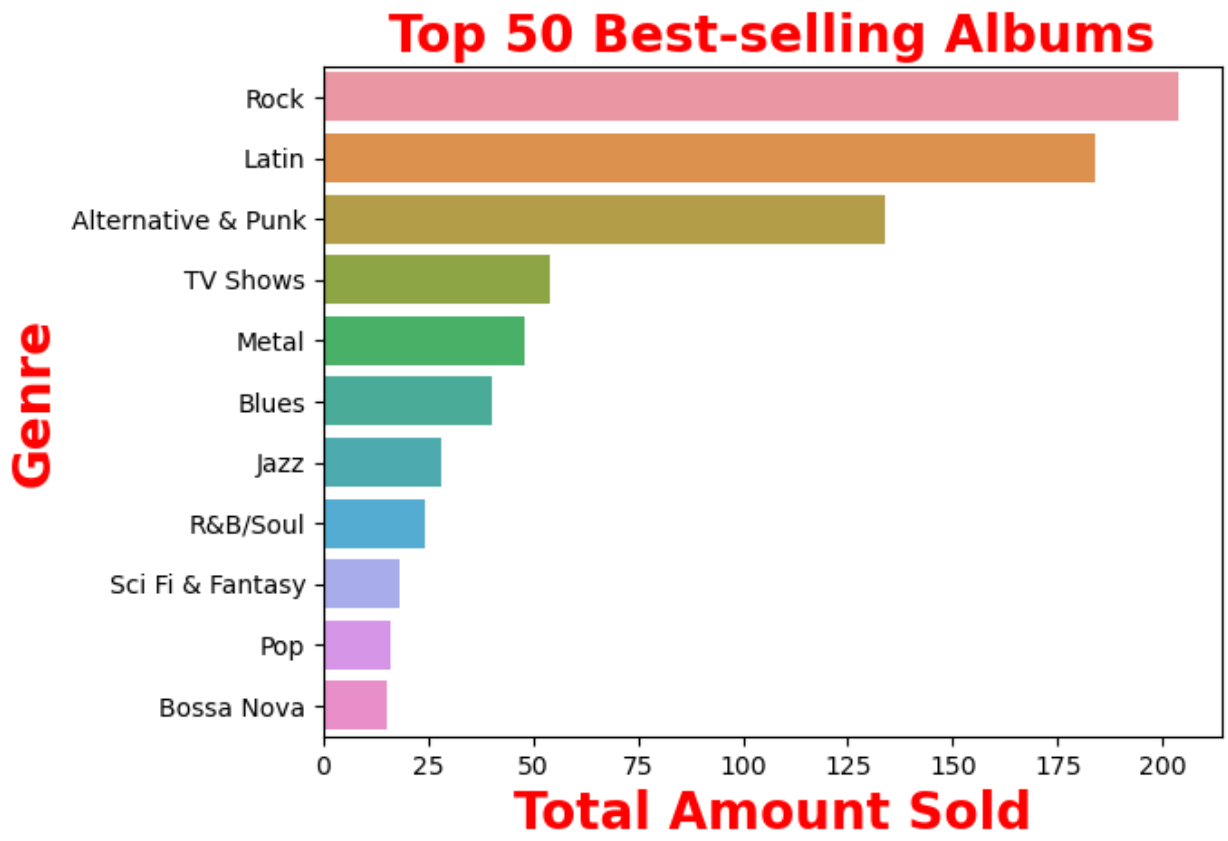
Appendix 2 Visit frequency distribution



Appendix 3.1 Top 15 best selling albums



Appendix 3.2 Most popular genre for top 50 best-selling albums



Appendix 4 Customer Spending Rank

	FirstName	LastName	Address	Phone	Email	Total_amount	Spending_rank
0	Richard	Cunningham	2211 W Berry Street	+1 (817) 924-7272	ricunningham@hotmail.com	47.62	2
1	Luis	Rojas	Calle Lira, 198	+56 (0)2 635 4444	luisrojas@yahoo.cl	46.62	3
2	Julia	Barnett	302 S 700 E	+1 (801) 531-7272	jubarnett@gmail.com	43.62	5
3	Terhi	Hämäläinen	Porthaninkatu 9	+358 09 870 2000	terhi.hamalainen@apple.fi	41.62	8
4	Luís	Gonçalves	Av. Brigadeiro Faria Lima, 2170	+55 (12) 3923-5555	luisg@embraer.com.br	39.62	12
5	François	Tremblay	1498 rue Bélanger	+1 (514) 721-4711	ftremblay@gmail.com	39.62	12
6	Bjørn	Hansen	Ullevålsveien 14	+47 22 44 22 22	bjorn.hansen@yahoo.no	39.62	12
7	Dan	Miller	541 Del Medio Avenue	+1 (650) 644-3358	dmiller@comcast.com	39.62	12
8	Heather	Leacock	120 S Orange Ave	+1 (407) 999-7788	hleacock@gmail.com	39.62	12
9	Wyatt	Girard	9, Place Louis Barthou	+33 05 56 96 96 96	wyatt.girard@yahoo.fr	39.62	12