

CANADA AIRLINE LOYALTY ACTIVITY ANALYSIS

NeuroTECH Academy

Catherine Habib

Group: 29 Dec 2024 | Online

Data Analysis Final Project Report: Methodology, Analysis, and Conclusions

A Comprehensive Examination and Findings

Introduction

This analysis report provides a detailed examination of the methodology employed, the core findings derived through analysis, and the conclusions drawn from the study. The report aims to ensure clarity, rigor, and traceability in all aspects of the research process.

Methodology

Research Design

The research utilized a **descriptive analysis** approach, highlighting qualities (**dimensions**) and quantities (**measures**) in the data to ensure robust and multifaceted insights. This approach was chosen to highlight, describe, and focus on the summary and presentation of the data.

The goal is to present the data in an orderly fashion and answer the question, “What happened?”.

Data Collection

Used the provided dataset: **Airline Loyalty Data**.

File Name	Rows	Columns	Format	Last Modified	MD5 Hash
Airline Loyalty Data Dictionary.csv	1,349	577	XLS Worksheet	2/5/2024 1:32 AM	34D1A261
Calendar.csv	115,117	7,234	XLS Worksheet	2/5/2024 1:32 AM	D3456A4D
Customer Flight Activity.csv	11,073,106	2,428,249	XLS Worksheet	2/5/2024 1:32 AM	15F22328
Customer Loyalty History.csv	1,733,749	284,015	XLS Worksheet	2/5/2024 1:32 AM	048AA4E0
Flight Log.csv	47,331	47,048	XLS Worksheet	2/5/2024 1:32 AM	034AAB0 MA 5CE45051215

Analysis

EDA

Used SQL queries to get distinct values and understand the data.

Query1.sql - LA...therine Habib (61)*

```

1 -- 7 years of enrollments 1 Distinct country: Canada with 11 provinces and 29 Cities
2 SELECT
3 DISTINCT country, province, city
4 FROM [Customer Loyalty History]
5 ORDER BY 1, 2, 3;
6
7 -- Distinct gender:
8 SELECT
9 DISTINCT gender
10 FROM [Customer Loyalty History];
11
12
13 SELECT
    DATENAME(YEAR, MIN(data)) MAX(data))

```

Results Messages

	country	province	city
1	Canada	Alberta	Calgary
2	Canada	Alberta	Edmonton
3	Canada	Alberta	Peace River
4	Canada	British Columbia	Dawson Creek
5	Canada	British Columbia	Kelowna
6	Canada	British Columbia	Vancouver
7	Canada	British Columbia	Victoria
8	Canada	British Columbia	West Vancouver
9	Canada	British Columbia	Whistler
10	Canada	Manitoba	Winnipeg
11	Canada	New Brunswick	Fredericton
12	Canada	New Brunswick	Moncton
13	Canada	Newfoundland	St. John's
14	Canada	Nova Scotia	Halifax
15	Canada	Ontario	Kingston
16	Canada	Ontario	London
17	Canada	Ontario	Ottawa
18	Canada	Ontario	Toronto

Query executed successfully.

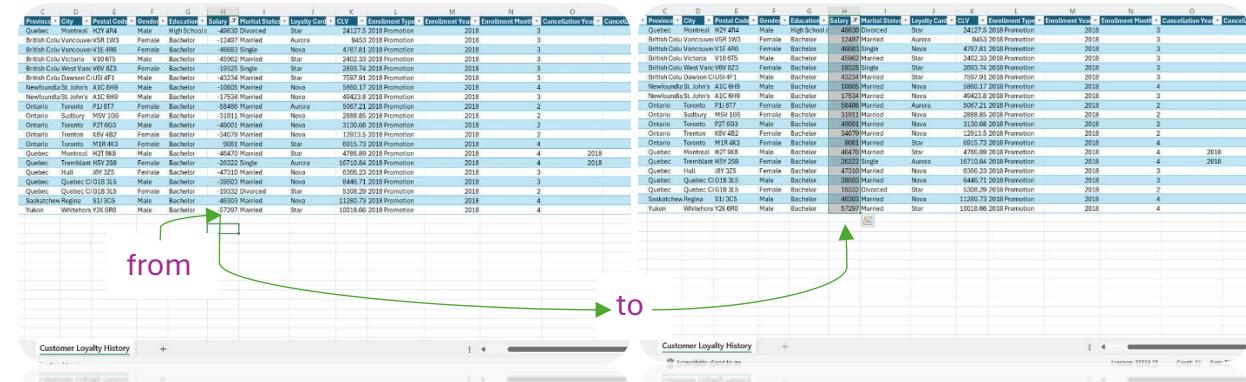
Example: Distinct data derived from table (Customer Loyalty History)



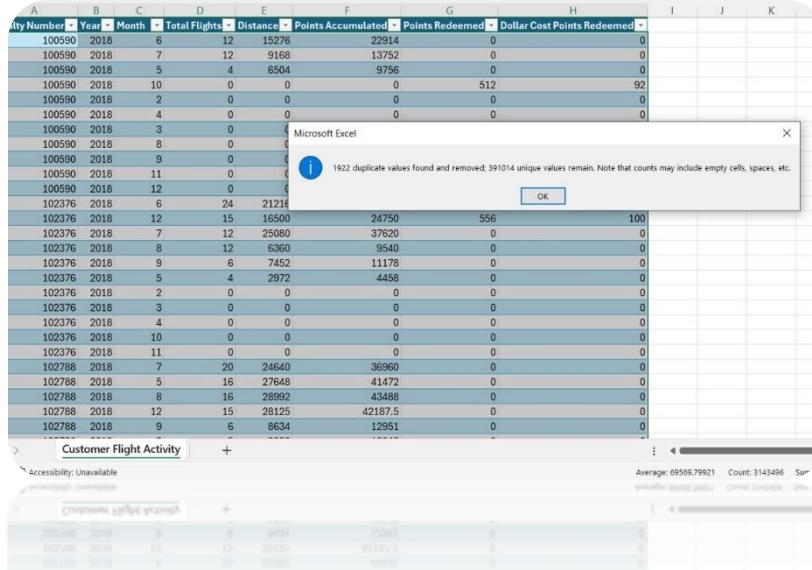
Data Cleaning and Preparation

Used Excel to clean, remove duplicates and prepare the data for further analysis and visualization.

Formatted data values in the **salary** field in table **Customer Loyalty History** from negative values to actual salary values.

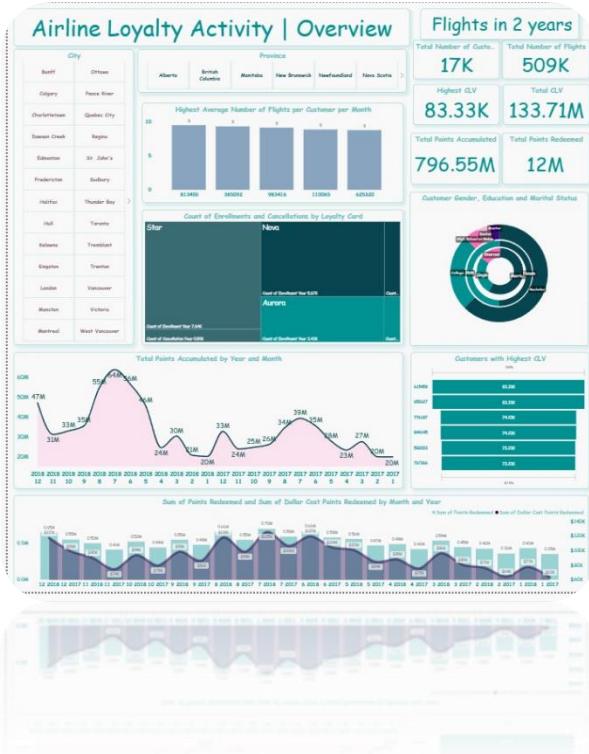


Removed 1,922 duplicates from table **Customer Flight Activity**.



Key Analytical Insights

Used Power BI to further clean the data, perform advanced data analysis and generate meaningful visualizations (charts identifying trends, patterns, and correlations within the dataset).



Questions that can be answered from this research:

1. What is the total number of customers?
2. What is the total number of flights?
3. What is the highest CLV?
4. Which customers have the highest CLV?
5. What is the average number of flights per customer?
6. Which customers have the highest average of flights per month?
7. Which customers have the highest average of flights per month per city?
8. Which customers have the highest average of flights per month per province?
9. What is the count and percentage of the customers per gender?
10. What is the count and percentage of the customers per education?
11. What is the count and percentage of the customers per marital status?
12. What is the total number of the accumulated points?
13. Which month has the highest accumulated points?
14. Which month has the least accumulated points?
15. Which year has the highest accumulated points?
16. What is the total number of redeemed points?
17. What is the value in dollars for the redeemed points per month?
18. What is the number of enrollments per loyalty card?
19. What is the number of cancelations per loyalty card?

Conclusions

Key Findings

- The airline of this dataset provides services to customers coming from only **1** Country which is **Canada** from **11** different Provinces and **29** different Cities in **Canada**.
- The dataset contains approximately **17K** customers enrolled in the airline services (exactly **16,737** customers).
- Most customers are **female** gender.
50.25% are Female.
49.75% are Male.
- There are **5** different degrees: The Majority have a **Bachelor's** degree.
62.59% have a **Bachelor's** degree.
25.32% have a **College** degree.
4.67% have a **High School** degree or Below.
4.39% have a **Doctor** degree.
3.04% have a **Master's** degree.
- There are **3** categories: The Majority are **Married**
58.16% are Married.
26.79% are Single.
15.04% are Divorced.
- There are **2** types of enrollments: **Standard** and **2018Promotion**.
- There are **3** types of Loyalty Cards: The Majority have a **Star Card**
7,637 are enrolled in a **Star Card**, **901** of which were cancelled.
5,671 are enrolled in a **Nova Card**, **717** of which were cancelled.
3,429 are enrolled in an **Aurora Card**, **449** of which were cancelled.
All enrollments were between **2012** and **2018**
&
All cancellations were between **2013** and **2018**.

- The Total of all CLVs measures approximately **133.71M**.
- Customers with the highest CLVs: **615459** and **652627**.
- All flights in this dataset were only in **2** years **2017** and **2018**.
- The dataset contains approximately **509K** recorded flights (exactly **508,808** flights).
- The total number of flights per customer ranges from **0** to **28** flights per month.
- The average number of flights per customer ranges from **0** to **9** flights per month.
- Most travelling customers: **813455** and **385092**.
- The Highest Numbers of Points Accumulated were in July of each of both years **2017** and **2018** with approximately **39M** and **64M**, respectively.
2018 has the higher numbers of accumulated points.
- The total of the points accumulated in both years is approximately **796.55M**.
- The total of the points redeemed in both years is approximately **12M**.
- The highest value of dollar cost points redeemed in **July 2018** with value **125,280\$**.