# Analytics Startup Plan

**Synopsis: *This document provides a high-level walkthrough of the activities required to guide completion of the analysis.***

|  |  |
| --- | --- |
| **Project** | *Air CC Customer Satisfaction Analysis* |
| **Requestor** | *Air CC Company* |
| **Date of Request** | *2022-07-06* |
| **Target Quarter for Delivery** | *2022-08-22* |
| **Epic Link(s)** | *Not applicable* |
| **Business Impact** | *Business can acquire insights from this analysis for improving customer experience and satisfaction to seek long-term growth.* |

## 1.0 Business Opportunity Brief

|  |  |
| --- | --- |
|  | Clearly articulated business statement of the Ask, opportunity, or problem you are trying to solve for. An important step is to understand the nature of the business, system or process and the desired problems to be addressed. This will be communicated back to All stakeholders for alignment. |

Air CC is a young Canadian airline, mainly undertaking intra-Canadian routes and a few important international routes. As competition among aerospace companies becomes more intense, Air CC is facing huge competitive adjustments, with operational issues such as customer loss, underutilization of resources, and marketing strategies.

Customer churn has a huge negative impact on profit growth. The longer the customer stays with the airline, the more profit the airline will engage. When the customer has high satisfaction, the same airline will continue to be chosen the next time. But to get new customers, it requires marketing techniques and a lot of money to attract customers. And new customers can be recommended by old customers. In this context, in order to retain customers, the most relevant factors for satisfaction need to be identified based on the analysis of historical customer satisfaction survey data.

Due to the general environment of the epidemic, there is a significant decline in revenue in 2020 and 2021 compared to previous years, and the airline industry should see a growth after the end of the epidemic. The company hopes to seize this opportunity to retain and attract customers and increase its overall revenue.

**The specific ask:**

*Clearly articulate the specific task you will be conducting to help achieve the opportunity*

Identifying key elements causing customers’ dissatisfaction help company concentrate sources on improving services or hardware facilities to improve customer experience and increase company revenue.

## 1.1 Supporting Insights

|  |  |
| --- | --- |
|  | Define any supporting insights, trends and research findings. Where relevant, list key competitors in the market. What are their key messages, products & services? What is their share of market, nationally and regionally? |

Air CC's major competitors include Air Canada, Westjet, Air Transat, Emirates, Air France and others. According to statista data, in 2020, air Canada held 43.7% of the Canadian domestic market share, WestJet held 36.5% and other air companies held only less than 20% totally (statista, 2022). The entire aircraft industry has been hit hard by the impact of the epidemic， during which people travel much less, and flights for leisure travel and visiting relatives and friends will increase significantly after COVID-19 (Bouwer, Saxon, & Wittkamp, 2021). Remote working patterns and flexible work arrangements during the pandemic are likely to remain in place for a long time, making it difficult for business travel to return to pre-pandemic proportions (Bouwer, Saxon, & Wittkamp, 2021).

## 1.2 Project Gains

|  |  |
| --- | --- |
|  | *Describe any revenue gains, quality improvements, cost and time savings (as applicable). What will you do differently and why would our customers care. What are the implications if we do nothing? This section is particularly key for prioritization against company goals and KPI’s.* |

By implementing this project, the key factors that determine customer satisfaction are identified, and the company can strategically adjust the status quo based on these factors, such as optimizing the online ticketing system, training flight attendants, improving in-flight meals, and other measures to enhance the customer experience. As a result, customer loyalty will be improved and thus long-term growth is achieved.

Without this project, the company may not target to improve some aspects that are truly important enough, which will result in a waste of resources and time.

## *Note: Completion of the following sections is possible only after a careful assessment and triage of the Ask. This is required to determine scope, resource, time, priority and data availability.*

## 2.0 Analytics Objective

|  |  |
| --- | --- |
|  | List the key questions, assumptions and define the hypotheses. Often the deliverable may not just be an analysis output, however a recommended operating model or blueprint for a pilot etc.  Note: Asking the right questions and truly understanding the problem will lead to the right data, right mathematics, and right techniques to be employed. |

Which flight factors such as seat comfort, food and beverage, entertainment, cleanness, flight delays, etc. do have the greatest impact on the customer's overall satisfaction?

Assumptions: The cleanness level, seat comfort level and on-board service could be more important in determining customer satisfaction.

Which flight factors do have the greatest impact on the disloyal customer’s overall satisfaction?

Assumptions: The cleanness level, seat comfort level, Arrival delay time and on-board service could be more important in determining disloyal customer satisfaction.

## 2.1 Other related questions and Assumptions:

|  |  |
| --- | --- |
|  | *List any assumptions that may affect the analysis* |

What factors do greatly influence satisfaction of people travelling for personal purpose?

Assumptions: The cleanness level, seat comfort level, inflight entertainment and so on.

## 2.2 Success measures/metrics

|  |  |
| --- | --- |
|  | *What does success look like? Define the key performance indicators (success definition/indicators, drivers and key metrics) against which the objectives will be analyzed. These should be drawn from the interlock meeting with key stakeholders and will inform the approach and methodology for the analysis.* |
|  | Customer satisfaction rate increases 10% next year.  Customer satisfaction rate of disloyal customers increases 5% next year.  The number of loyal customers increases.  The overall revenue increases. |
|  |  |

## 2.3 Methodology and Approach

|  |  |
| --- | --- |
|  | *Now that you have a good understanding of the Ask and deliverable, detail the recommended approach/methodology.* |

**Type of Analysis:** decision tree, logistic regression, neural network

I will use decision tree analysis to find out which variables are highly correlated with customer satisfaction. Regression analysis and neural network will also be used to verify the conclusion.

**Methodology**: Descriptive analysis will be used to do descriptive statistics on the data and observe potential problems, then use EDA to examine, clean and transform data. Multiple models will be used to analyze which factors are more prominent in affecting overall customer satisfaction and specific customer group satisfaction, then compare the accuracy of these models to determine the best model and its results.

**Output:**The results will include an analysis of important factors affecting customer satisfaction and recommendations for improving customer satisfaction based on the conclusions of the analysis. Factors and recommendations that affect the satisfaction of specific customer segments will also be provided.

## 3.0 Population, Variable Selection, considerations

**Audience/population selection**:

Historical passenger aged from 7 to 85, including loyal customers and disloyal customers

**Observation window:** Not applicable

**Inclusions:**

All 23 current variables and records are included before doing exploratory data analysis.

**Exclusions:**

There may be some exclusions of high correlated response variables when doing dimension reduction.

**Data Sources:**

Kaggle, https://www.kaggle.com/datasets/sjleshrac/airlines-customer-satisfaction

**Audience Level:**

The operations team，Diretors

**Variable Selection:**

Satisfaction, Gender, Customer Type, Age, Type of Travel, Class, Flight Distance, Seat comfort, Departure/Arrival time convenient, Food and drink, Gate location, Inflight wifi service, Inflight entertainment, Online support, Ease of Online booking, On-board service, Leg room service, Baggage handling, Checkin service, Cleanliness, Online boarding, Departure Delay in Minutes, Arrival Delay in Minutes

**Derived Variables:**

Indicator\_Missing (If value is missed, then “1”. If not, then “0”.)

Others may be derived in future analytics.

**Assumptions and data limitations:**

Assumptions:

These customers are randomly selected.

Data limitations:

* The source company is not specific. I cannot access additional background information of this dataset.
* Most of variables are survey data, which could involve bias.
* There are missing values existing in one variable, which needs to be imputed.
* This data was updated 2 years ago in Kaggle, but without specific time frame. If this data set is too old, the results may not reflex the reality of recent time.

## 4.0 Dependencies and Risks

|  |  |
| --- | --- |
|  | Identification of key factors that may influence the outcome of the project and likelihood of it happening: |

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Likelihood (based on historical data)** | **Delay (based on historical data)** | **Impact** |
| Survey answers from underage passengers may be completed by their guardians instead. | *Median* |  | The records of come young passengers may not reflect their real feelings and evaluations, which may mislead the relevant analysis results. |
| This survey is interviewed and recorded by humans, and the results may contain human error and multiple bias. | *Low* |  | Artificial data errors may affect the accuracy of analysis conclusions. |
|  |  |  |  |

## 5.0 Deliverable Timelines

|  |  |
| --- | --- |
|  | List key dates and timelines as a work-back schedule. Activate line items based on complexity and line-of-sight required. Will set the stakeholder expectations for the process. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item** | **Major Events / Milestones** | **Description** | **Days** | **Date** |
| 1. | Kick-off / Formal Request | Instruction of Capstone | 1 | July 5 |
| 2. | Data Finalization | Data feasibility discussion and finalization of data | 5 | July 10 |
| 3. | Analysis Plan | Statement of analysis objects, methodology clarification, timelines and so on | 7 | July 15 |
| 4. | EDA | Exploration of data structure, data quality and variable properties Preparations like cleansing, transforming for modeling | 7 | July 22 |
| 5. | Modeling | Decision Tree  Regression  Neural Network | 10 | August 3 |
| 6. | Governance | Risk management | 2 | August 5 |
| 7. | Documentation | Documentation in the form of a report with project introduction, process and results. | 7 | August 12 |
| 8. | Presentation | PowerPoint presentation for stakeholders explaining project gains | 10 | August 15 |
| 9. | Portfolio | Enhancement of Portfolio by including this project analysis | 2 | August 19 |

# References

Bouwer, J., Saxon, S., & Wittkamp, N. (2021, April 2). *Back to the future? Airline sector poised for change post-COVID-19*. Retrieved from McKinsey & Company: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/back-to-the-future-airline-sector-poised-for-change-post-covid-19

statista. (2022, July 12). *Major airlines' domestic market share in Canada in 2020\**. Retrieved from statista: https://www-statista-com.centennial.idm.oclc.org/statistics/545642/air-carrier-canada-domestic-market-share/