# Analytics Startup Plan

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

|  |  |
| --- | --- |
| **Project** | Airbnb |
| **Requestor** | Centennial College |
| **Date of Request** | 2022-07-06 |
| **Target Quarter for Delivery** | 2002-08-22 |
| **Epic Link(s)** |  |
| **Business Impact** | Improving revenue and attracting more house host |

## 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. |

………

**The specific ask:**

Rental property owners may get a feel of whether a property's location or neighborhood is worth investing in by analyzing the accurate data provided by Airbnb. Precise analysis can help them choose the optimal rental strategy. Airbnb's revenue is rising as more people use the service to rent out vacation homes.

## 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? |

Airbnb is a United States-based online community marketplace and hospitality agency. Local hosts can use Airbnb as a platform to accommodate guests for short-term stays and engage in tourism-related activities. Various search options, such as types of lodging, dates, price, and locations, will be available to guests. They can also narrow their search to certain types of homes, vacation homes, and bed and breakfasts. The hosts provide information such as the rental fee, location, number of guests, restrictions, home types, and other facilities.

Airbnb's most recent data shows it has more than six million listings in more than 100,000 cities and towns and more than 220 countries. New York City is the third-largest Airbnb market in the world and one of the oldest.

Airbnb's competitors include Booking.com, Expedia, TripAdvisor, and Home Exchange. The Competition among current businesses indicates the number of rivals that present Airbnb with strong competition. Extreme competition demonstrates that Airbnb is exposed to intense pressure from its competitors, which can limit each other's growth potential. In such industries, profit margins are low due to firms' aggressive targeting and pricing methods.

## 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.* |

For Airbnb: Every time a reservation is made, Airbnb receives commissions from both hosts and guests. Airbnb charges the guest 6-12 percent of the booking fee for each reservation. Additionally, Airbnb charges the host 3% of each successful transaction. Due to Airbnb's revenue model, it is essential to growing the number of bookings to ensure the company's profitability. One way would be for hosts to offer lodgings and experiences at reasonable prices and to provide prospective guests with efficient recommendations of popular listings. Airbnb can increase its revenue by helping hosts quote fair pricing and understanding market trends and guests' expectations.

For hosts: Several variables might affect the pricing of any listing, making it difficult for hosts to come up with a reasonable price. An understanding of guest willingness to book in light of various features of hosts and houses has been gathered from historical data collected from Airbnb listings. Hosts can gain a competitive advantage by focusing on aspects like location, amenities, and history of pleasant experiences to set pricing at mutually desirable levels.

## 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. |

1. What variables influence listing price?

Location, room type, review rating, instant bookable, amenities

1. How can hosts optimize price?

Increase properly quality, amenities, and experience reviews

## 2.1 Other related questions and Assumptions:

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

According to New York's tourism laws, an Airbnb listing for a room in a home occupied by a permanent resident is lawful. Renting a whole apartment to an Airbnb guest for fewer than 30 days is against the law. Is long-term profitable for the entire home rental?

## 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.* |
|  | Airbnb increase 2% number of reservations each quarter.  More homes and rooms list on Airbnb. |
|  |  |

## 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:** descriptive analysis, predictive analysis: decision tree, regression

**Methodology:** Regression is used to train the model and to forecast the listing price, and the model is initially constructed using all variables as predictors. Various evaluation metrics such as Variance Explained Score, Mean Absolute Error, R2-score, RMSE (Root Mean Squared Error) can be used in the project.

**Output:** The output will be the most critical variable affecting list price. And the output will give the rental homeowner a reference about how the home will be priced.

## 3.0 Population, Variable Selection, considerations

|  |  |
| --- | --- |
|  | Capture learning about the data available today location, structure, and reliability; this would include data in operational systems including dealer sourced, data warehouse and any CRM or email marketing systems available today. |

**Audience/population selection:** Not applicable

**Observation window:** June 05, 2021 – June 05, 2022

**Inclusions:** 16 Variables

**Exclusions:** 55 Variables

**Data Sources:** Inside Airbnb <http://insideairbnb.com/get-the-data/>

**Audience Level:** Hosts and Airbnb operating team

**Variable Selection:** host\_is\_superhost, host\_identity\_verified, neighbourhood\_cleansed,

neighbourhood\_group\_cleansed, latitude, longitude, room\_type, bathrooms\_text, bedrooms,

beds, amenities, price, availability\_365, number\_of\_reviews, review\_scores\_rating,instant\_bookable

**Derived Variables:** Not applicable

**Assumptions and data limitations:**

## 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** |
| The majority-minority classification problem | High |  | Class imbalance in the large dataset can significantly distort the performance of classifiers, resulting in a prediction bias for the majority class. For example, most hosts in the listing are not super hosts. |

## 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** | **Scope** | **Days** | **Date** |
| 1. | Kick-off / Formal Request | * Giving the information and requirements of the capstone |  | 1 | July 05 |
| 2. | Analysis Plan | * Statement of the research objectives and hypothesis * Identifies the dataset to be used. Inclusion and exclusion criteria. |  | 7 | July 15 |
| 3. | Data Exploration & Analysis | * Utilizing summary statistics and graphical representations to perform early work on data in order to detect trends, identify errors, test hypotheses, and verify hypotheses. |  | 7 | July 22 |
| 4. | Modeling | * Applying decision tree, regression, and Neural Network |  | 10 | *August 03* |
| 5. | Governance | * Data management |  | *2* | *August 05* |
| 6. | Documentation | * Forming the project into a report |  | 7 | *August 12* |
| 7. | Presentation | * PowerPoint presentation |  | *10* | *August 22-24* |
| 8. | Portfolio | * Publishing the project on the portfolio website |  | 2 | *August 26* |