Project Plan

Analysing AirBnB Data in Manly NSW

Georgia Platt

Jackson Scown

Gia Huy Lieu

Table of Contents

[1.0 Introduction 3](#_Toc46748287)

[1.1 Problem Background 3](#_Toc46748288)

[1.2 Scope 3](#_Toc46748289)

[1.3 Document contents 3](#_Toc46748290)

[2.0 Work Breakdown Structure 4](#_Toc46748291)

[3.0 Activity Definition & Estimation 5](#_Toc46748292)

[4.0 Gantt Chart 6](#_Toc46748293)

# Introduction

## Background

*The Sydney AirBnB Data Analysis project aims to analyse a detailed dataset containing information regarding AirBnB listings in Sydney Australia. The project involves various tasks, including retrieving and reporting listing information for accommodation in Manly between 2012 and 2023, creating a price distribution chart, conducting keyword-based searches, analysing customer comments related to cleanliness, and analysing booking trends using Time Series Analysis. The goal is to provide valuable insights into the AirBnB market in Sydney and enhance decision-making for hosts, guests, and stakeholders.*

## Scope

1. Data Collection and Preparation:

Obtain the Sydney AirBnB dataset, clean and preprocess the data for analysis. This involves handling missing values, data validation, and data formatting. Need to clean “name” column of dataset.

1. Data Retrieval and Reporting:

Develop a system to retrieve and display information about Manly listings between 2012 and 2015. This includes listing details such as property name, room type, price, number of reviews, etc.

1. Price Distribution Chart:

Generate a visual representation (chart) of the distribution of property prices between 2012 and 2023. This chart will help understand the range and spread of prices.

1. Keyword-based Search:

Implement a search functionality to retrieve records containing specific user-entered keywords (e.g., "pool," "pet"). This feature enables users to find listings that match their preferences.

1. Cleanliness Analysis:

Analyse customer reviews to identify factors related to cleanliness. Develop a method to extract relevant reviews and quantify the frequency of cleanliness-related keywords.

1. Time Series Analysis:

Perform an innovative analysis, which explores trends in booking patterns over time. Analyse with the assumption that the last review was completed the day the booking came to an end.

## Document contents

1. Introduction:

Problem Statement:

Describe the objective of the project, which is to analyse a comprehensive dataset of AirBnB listings in Sydney to provide valuable insights for hosts, guests, and stakeholders.

Project Scope:

Define the scope of the project, including tasks such as data retrieval for Manly listings, price distribution chart creation, keyword-based searches, cleanliness analysis, and time series analysis.

Document Overview:

Provide an outline of the sections covered in the document.

1. Work Breakdown Structure (WBS):

WBS Overview:

Present a hierarchical breakdown of the project tasks, starting from high-level components (e.g., Data Collection) and drilling down into sub-tasks (e.g., Data Cleaning, Data Preprocessing).

WBS Diagram:

Include a visual representation of the WBS with task relationships and dependencies.

1. Activity Definition and Estimation:

Task Definitions:

Detail the tasks identified in the WBS, including a brief description of each task's purpose.

Task Duration Estimation:

Provide estimated durations for each task based on team expertise, historical data, and potential complexities.

1. Gantt Chart:

Project Timeline:

Present the timeline of the project with tasks plotted over time.

Milestones:

Highlight key milestones, such as Data Collection Completion, Analysis Phases, and final reporting.

Dependencies:

Depict task dependencies to show the order in which tasks need to be executed.

# Work Breakdown Structure

*This section should include the work breakdown structure for the whole project. The elements from the WBS should be used to generate your activity definition and those activities should then be scheduled in the Gantt Chart. Remember to consider ALL project activities – anything you do or will need to do should be included in the WBS*

*WBS’s are usually presented as some kind of hierarchical diagram/chart etc. The details what is involved each work unit should be provided in section 3:* ***Activity Definition***

*You do NOT need to do a WBS Dictionary for this project – the activity definition (whilst slightly different) will suffice. The WBS is focussed on SCOPE. The Activity definition is focussed on TIME.*

# Activity Definition & Estimation

*From your WBS, define the activities required for your project. You will revise this document and add more detail for part B as you discover more about the project.*

*Each activity should be clearly identified by a number and should match up to your Gantt chart. You should provide some estimations for the time you think each activity will take. This should make it easy to prepare your Gantt chart.*

# Gantt Chart

*This section should contain your Gantt chart. The items in the Gantt chart should match the activity definition from section 3. You should also submit your Gantt chart file separately.*