High Level Design (HLD)

Travel Data Analysis

Domain: Business Intelligence

Revision Number:1.0

Last Date Of Revision: 03.12.2023

Mohammad Tippu Sohail

Mohammad Salman

Document Version Control

Date issued	Version	Description	Author
Nov 25, 2022	1.1	First Draft	Mohammad Tippu & Mohammad Salman

Contents

Abstract……2

Introduction 3

What is High-Level Design Document? 3

Scope 3

General Description 3

Definitions 3

Product Description 4

Airbnb is an online platform that connects travelers with unique lodging options provided by individual hosts, offering a personalized and authentic travel experience. 4

Problem Statement 4

Since the launch of Airbnb in 2008, both travelers and hosts have found this platform more interesting, enhancing travel experiences with a more personal touch. 4

Proposed solution 5

Using all the standard techniques used in the life cycle of a Data Analysis project starting from Data Exploration, Data Cleaning , Data transforming , data Visualizations and also building a Dashboard with greater visuals using Tableau ..... 5 Tools used 5

Design Details 6

Process Flow 7

KPIs (Key Performance Indicators) : 9

Conclusion..... 16

Deployment 15

Abstract

Since 2008, Airbnb has played a pivotal role in transforming travel experiences, enabling both guests and hosts to engage in a more personalized and unique exploration of the world. As the platform continues to shape the way people experience travel, this dataset serves as a valuable resource for unraveling patterns and uncovering valuable insights within the context of one of Netherland's vibrant cities. This dataset specifically from 2017 provides a comprehensive overview of Airbnb listing activity in Amsterdam capturing essential information for understanding hosts, geographical availability, and key metrics relevant for analysis and it also opens avenues for exploring trends in travel preferences, host behavior, and the overall impact of Airbnb on the hospitality landscape in Amsterdam. The dataset aims to identify top earners among hosts and explore potential relationships between monthly earnings and pricing strategies. With respect to or Regarding neighborhoods, the focus shifts to uncovering locations with the highest booking numbers and understanding the relationship between pricing and specific areas. It also explores reviews to know how good they are compared to prices. Lastly, looking into how prices relate to the amenities offered and the location of the place. Overall, this dataset helps us see how Airbnb works in Amsterdam and gives us insights into what hosts are doing, where people prefer to stay, and how prices play a role in all of it.

Introduction

What is High-Level Design Document?

The goal of this HLD or a high-level design document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of design aspects and define them in detail
- Describe all user interfaces being implemented Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and architecture of the project
- List and describe the non-functional attributes such as security, reliability, maintainability, portability, reusability, application compatibility. resource utilization, serviceability

Scope

The HLD documentation presents the structure of the system, such as database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly technical terms which should be understandable to the administrators of the system.

General Description

Product Description

Airbnb is an online platform that connects travelers with unique lodging options provided by individual hosts, offering a personalized and authentic travel experience.

Problem Statement

Since the launch of Airbnb in 2008, both travelers and hosts have found this platform more interesting, enhancing travel experiences with a more personal touch.

Objectives for Data Analysis:

1. Host Analysis:

- Identifying Top Earners
- Monthly Earnings and Prices Relationship

2. Neighborhood Analysis:

- Identifying specific locations (neighborhoods) in Amsterdam that receives the maximum number of bookings
- Price and Location Relationship

3. Reviews Analysis:

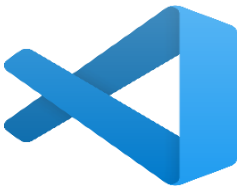
- Quality and Price Relationship
- Price vs Amenities
- Price Relation with Location

Proposed solution

Using all the standard techniques used in the life cycle of a Data Analysis project starting from Data Exploration, Data Cleaning , Data transforming , data Visualizations and also building a Dashboard with greater visuals using Tableau

Tools used

Python programming language and frameworks such as NumPy, Pandas and few other libraries were used to build the whole model.



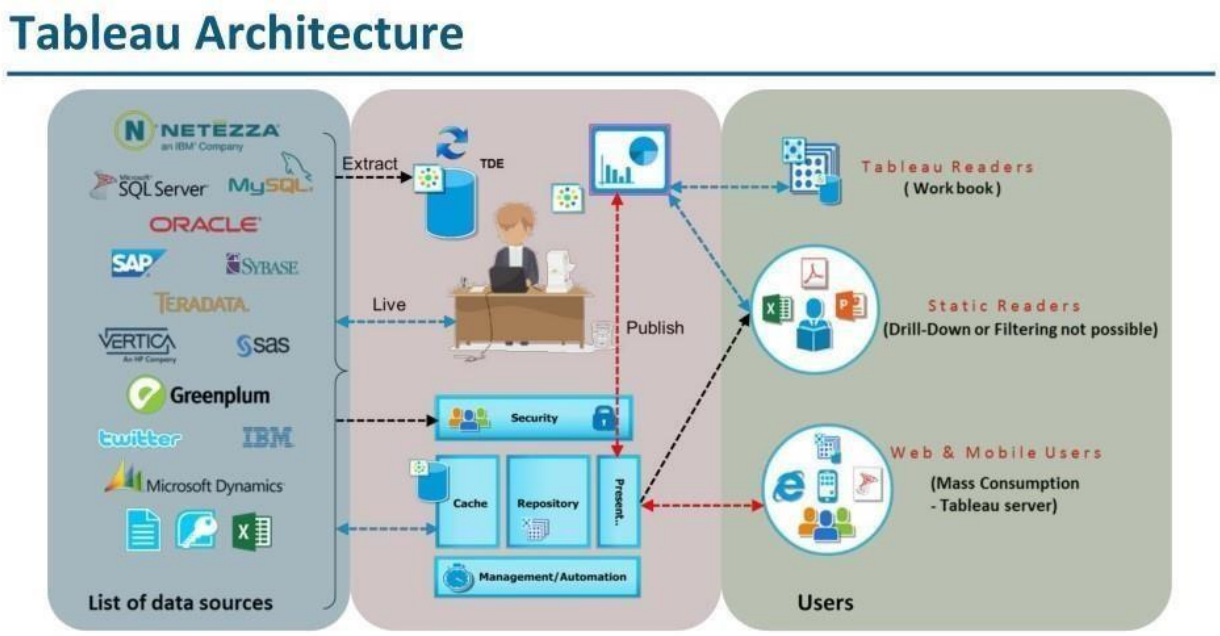
4. Price Analysis:

- For visualization tasks, matplotlib, seaborn and plotly were used
- GitHub is used as version control system

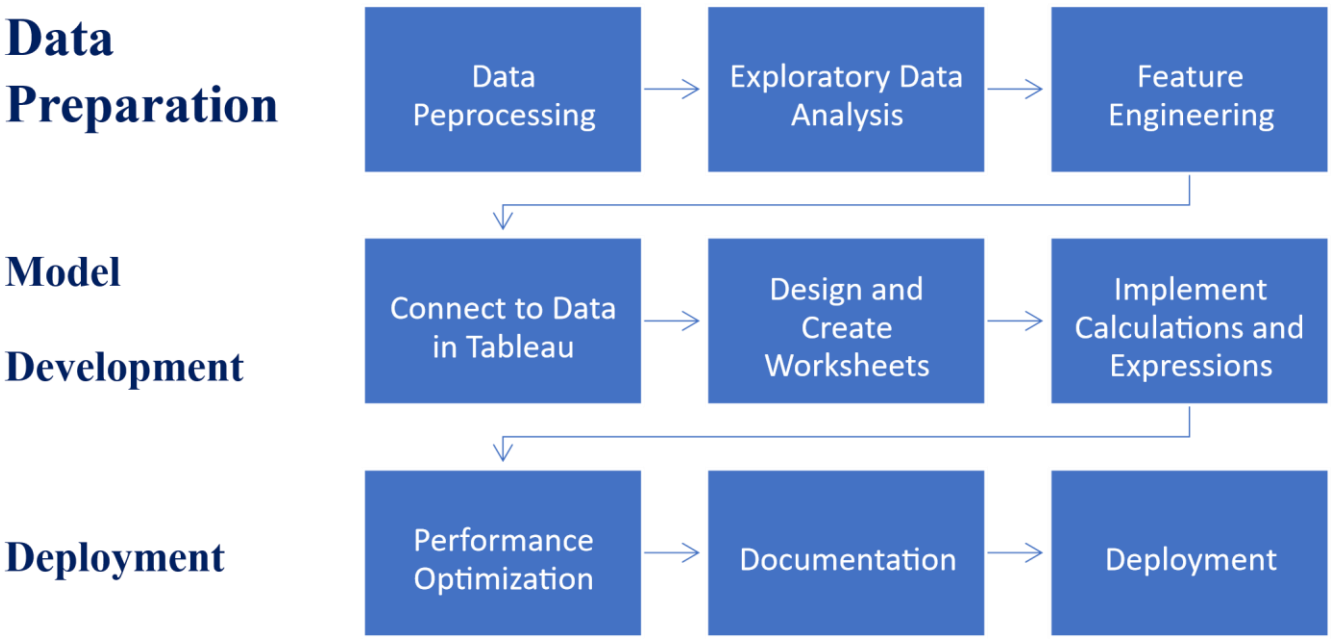
- NumPy and Pandas were used to clean and interpret data

- Another powerful data visualization and business intelligence tool used is Tableau.

Design Details



Process Flow



Optimization

Your data strategy drives insights

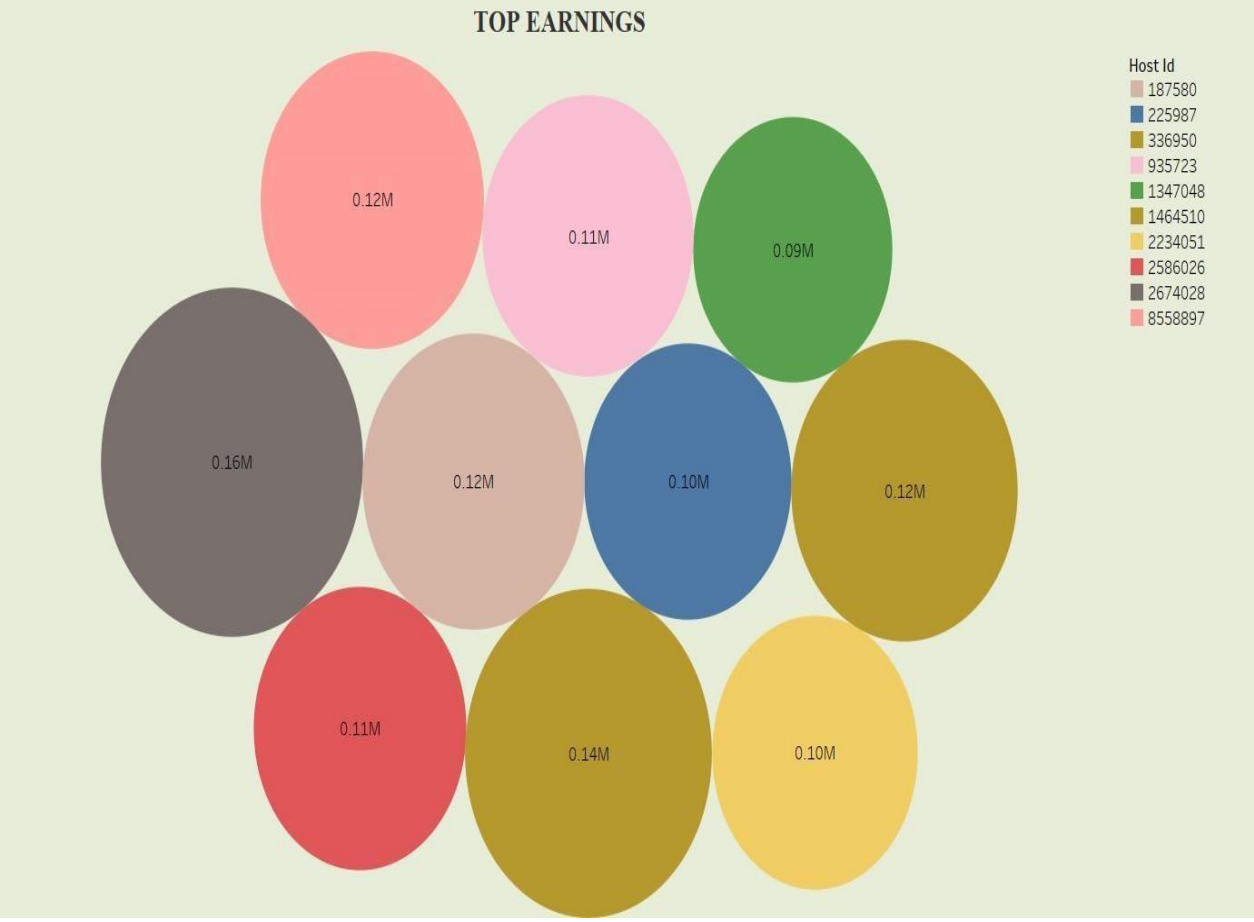
- Cleaning Data by imputing missing values or removing missing values.
- Changing and converting the column data types.
- Creating measures to create values for simple card views. Optimize space to ensure every chart or visuals is fitting and telling the same story.

Reduce the marks (data points) in your view

- Practice guided analytics. There's no need to fit everything you plan to show in a single page. Compile related visuals and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
- Creating charts that explore relationship between different variables in the dataset

KPIs (Key Performance Indicators) :

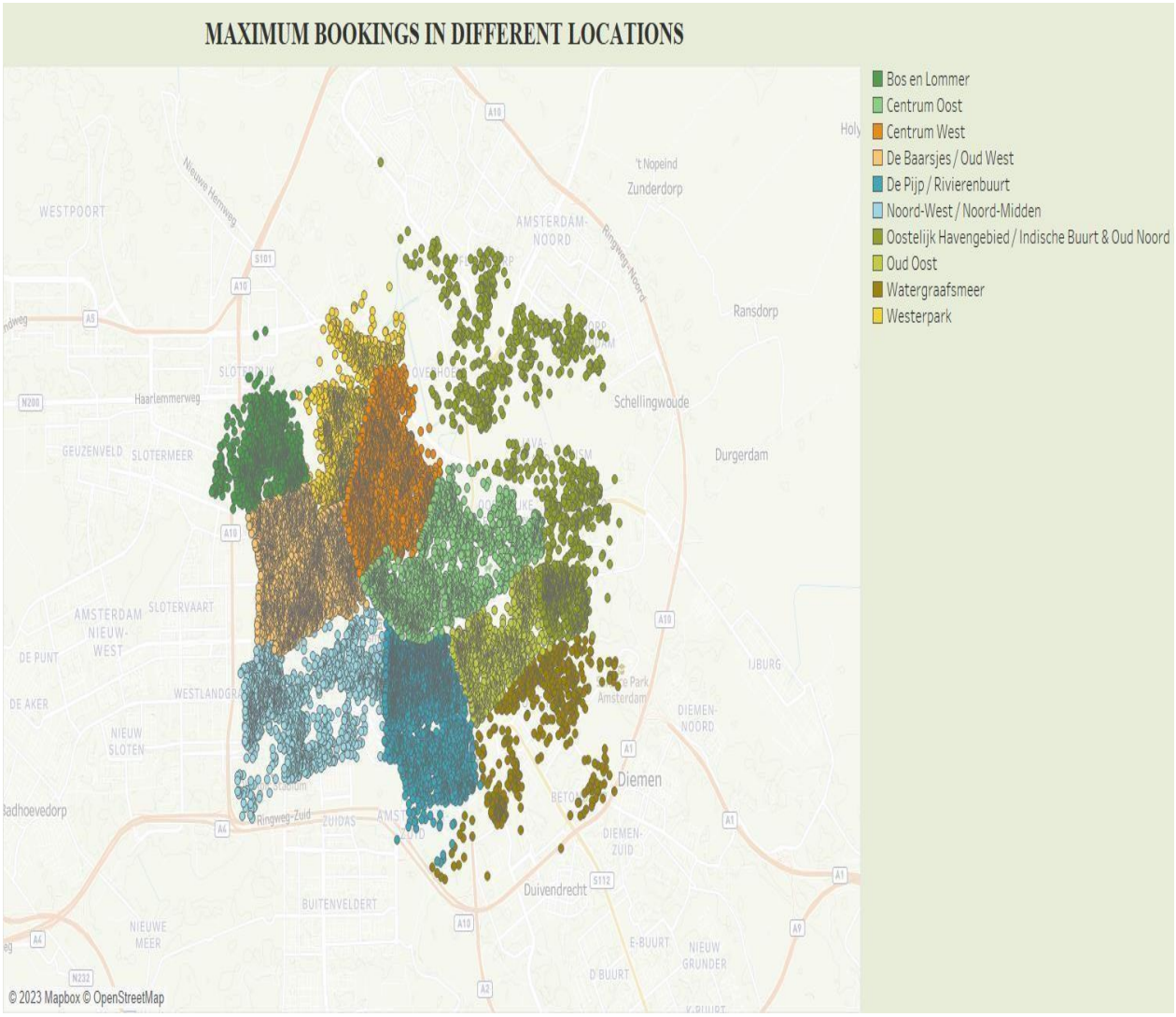
TOP EARNERS AMONG THE HOSTS:



Summary:

- Sum of Earned. Color shows details about Host Id. Size shows sum of Earned.
- The marks are labeled by sum of Earned.
- The view is filtered on Host Id, which keeps 10 of 15,943 members.

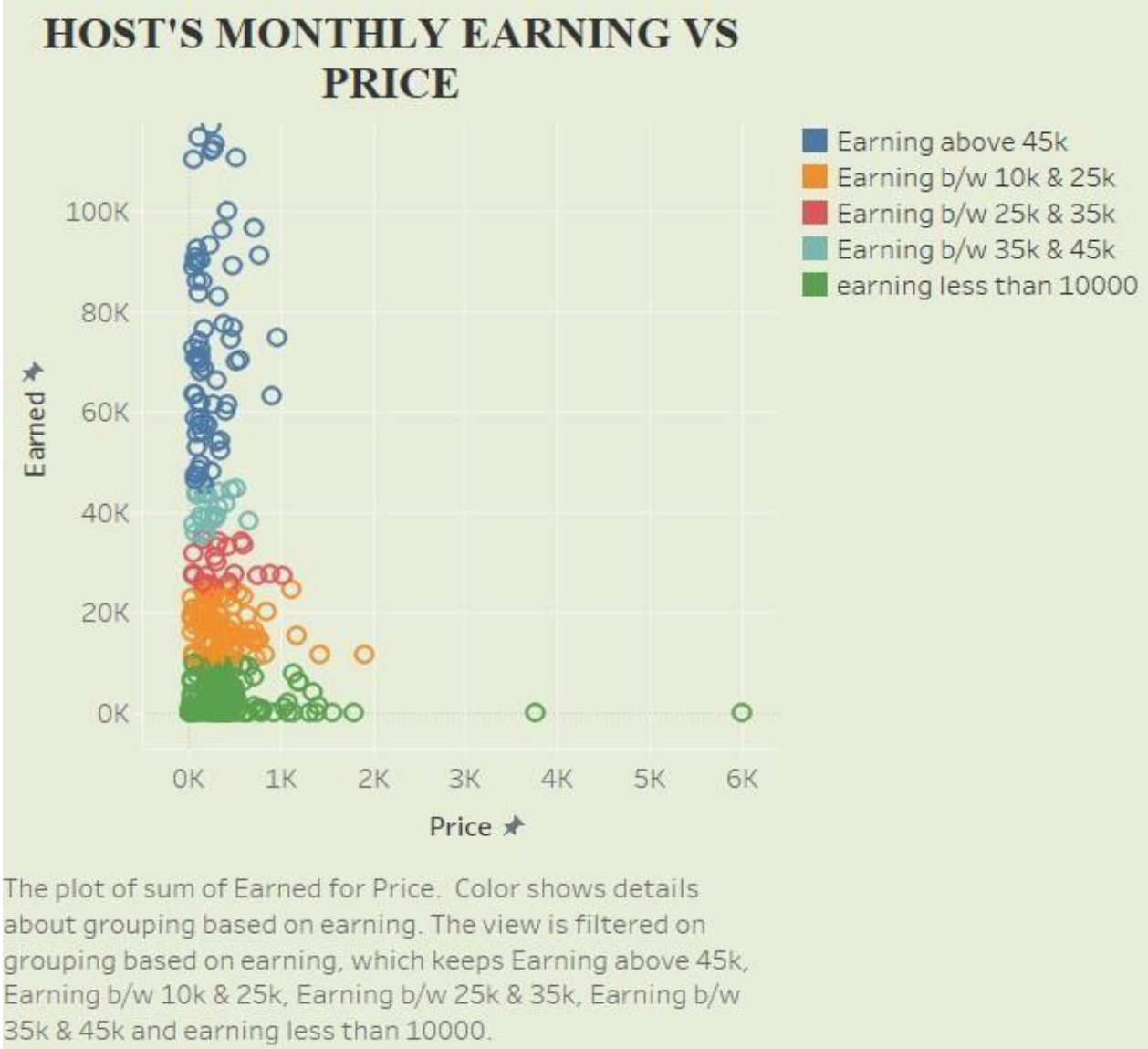
MAXIMUM BOOKINGS AMONG DIFFERENT LOCATIONS :



Summary:

- Map based on Longitude and Latitude. Color shows details about Neighborhood (group).
- The view is filtered on Neighborhood (group), which has multiple members selected.

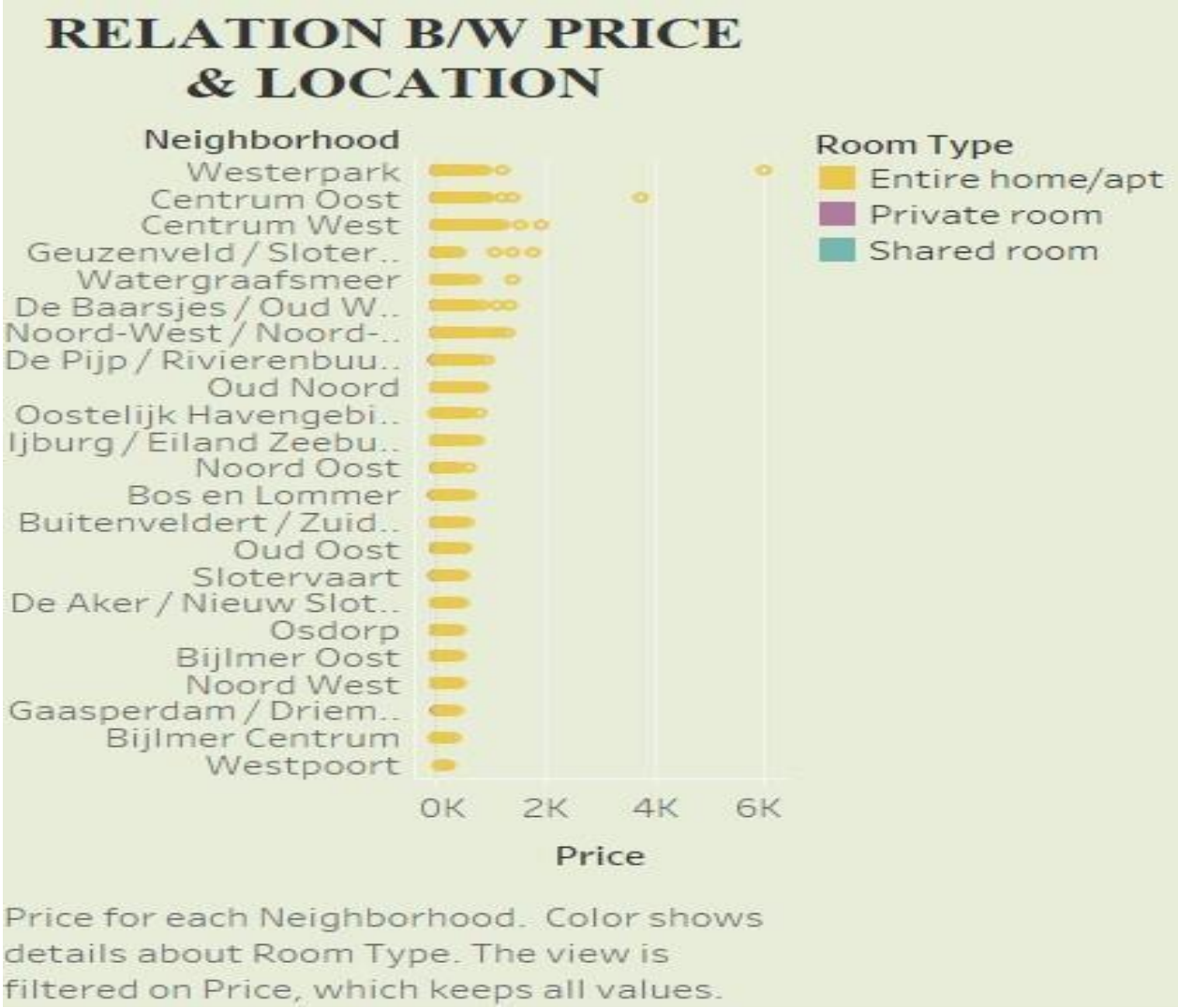
HOSTS MONTHLY EARNING VS PRICE:



Summary:

- The plot of sum of Earned for Price. Color shows details about grouping based on earning.
- The view is filtered on grouping based on earning, which keeps Earning above 45k, Earning b/w 10k & 25k, Earning b/w 25k & 35k, Earning b/w 35k & 45k and earning less than 10000

RELATION BETWEEN PRICE AND LOCATION :



Summary:

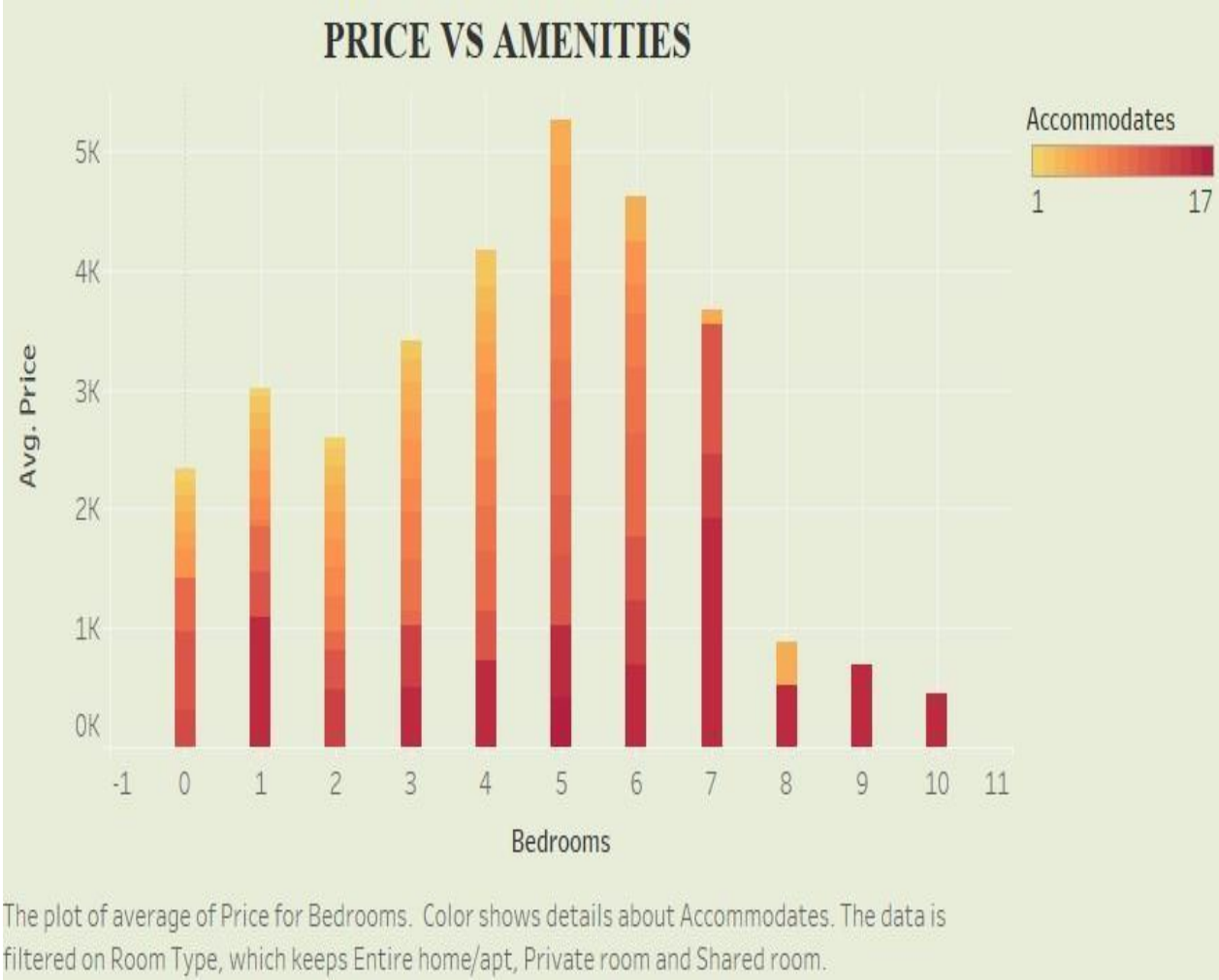
- Price for each Neighborhood.
- Color shows details about Room Type.
- The view is filtered on Price, which keeps all values.

QUALITY VS PRICE:

Summary:

- The plot of median of Price for Overall Satisfaction.
- Shape shows details about sum of Reviews.
- The marks are labeled by median of Price.

PRICE VS AMENITIES:



The plot of average of Price for Bedrooms. Color shows details about Accommodates. The data is filtered on Room Type, which keeps Entire home/apt, Private room and Shared room.

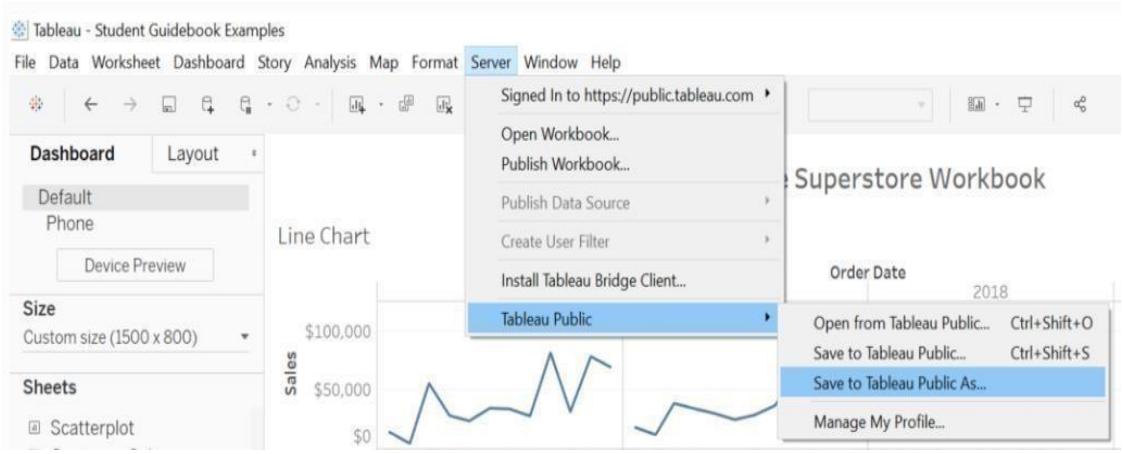
Summary:

- The plot of average of Price for Bedrooms.
- Color shows details about Accommodates.
- The data is filtered on Room Type, which keeps Entire home/apt, Private room and Shared room.

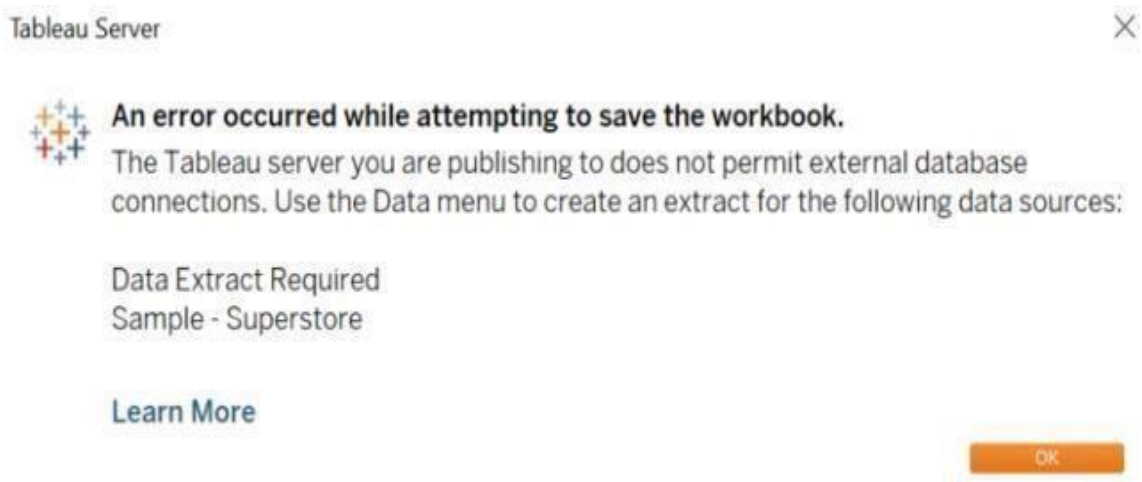
Deployment :

Once you've completed your dashboard, follow these steps:- **Server, Tableau Public, Save to Tableau Public As**

You may be prompted to log into your Tableau Public profile first if this is your first time publishing.

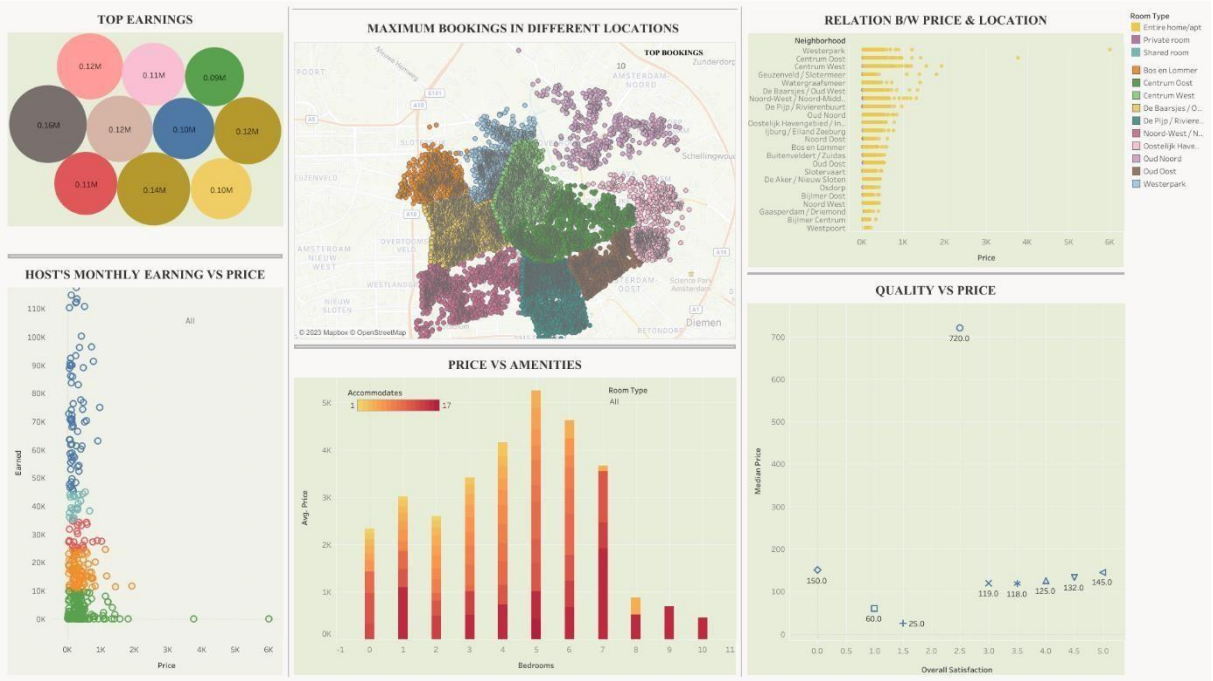


Next, fill out the title you want your viz to have and click "save".



This message means that your connection to the Sample-Superstore data set is a live connection. Tableau Public cannot host live connections, so you'll need to convert your connection to an extract (like a frozen screenshot of your data).

Here in the below screenshot, we can see that out workbook has been published to tableau public.



Conclusion

The results will help stakeholders in the AirBNB ecosystem, including hosts and travelers, make informed decisions for a more personalized and enjoyable experience.