Assessment 2

|  |  |
| --- | --- |
| Student Name/ID Number |  |
| **Unit Number and Title** | **14: Business Intelligence** |
| Academic Year | 2022 |
| Unit Tutor |  |
| **Assignment Title** | * Show your data set and describe basic information and statistical information related to the data set. * List and describe different type of charts for the variable in the data ser. * Apply data preprocessing steps (clean/scrub/filter, so on) for the data set (if have). Why do we need the steps? * Using tableau desktop to transform your dataset to a tableau dashboard. |
| **Issue Date** |  |
| Submission Date |  |
| IV Name & Date |  |

|  |
| --- |
| **Submission Format** |
| Part I: Project submission. This should be a zip / rar folder of your project, including all necessary files to run your project.  Part II: all students must present their writing report for the second assignment. This report should be written in a concise, formal business style using single spacing and font size 12. You are required to make use of headings, paragraphs and subsections as appropriate, and all work must be supported with research and referenced using the Harvard referencing system. Please also provide a bibliography using the Harvard referencing system.  Part III: an oral presentation and public speaking are not required for all students, but it is an important aspect for student who want to get high score for the 2nd assessment. Moreover, your presentation must explain the reasons why you select the tools and techniques to transform data to useful information in your topic. |

Table of Contents

[Table of Figures 1](#_Toc97970133)

[Introduction 2](#_Toc97970134)

[Task 1 Data set and basic information and statistical information related to the data set. 2](#_Toc97970135)

[1.1. Data set 2](#_Toc97970136)

[1.2. Pre-process steps on dataset(s) 2](#_Toc97970137)

[Task 2: List and describe different type of charts for the variable in the data ser. 5](#_Toc97970138)

[Task 3: Dashboard and purpose of dashboard 8](#_Toc97970139)

[3.1. Each dashboard contains charts (Purpose of chart) 8](#_Toc97970140)

[3.2. Interaction on dashboards (Result / Conclusion / Suggestion) 9](#_Toc97970141)

[Task 4: Data preprocessing steps and its important 10](#_Toc97970142)

[4.1. The important of BI to the business 10](#_Toc97970143)

[4.2. The important of preprocessing steps 11](#_Toc97970144)

[CONCLUSION 12](#_Toc97970145)

[REFERENCES 13](#_Toc97970146)

Table of Figures

[Figure 1 Dataset CSV 3](#_Toc97970147)

[Figure 2 Tableau UI 4](#_Toc97970148)

[Figure 3 Dataset after imported 4](#_Toc97970149)

[Figure 4 Table in tableau 5](#_Toc97970150)

[Figure 5 Column graph Price/Date 5](#_Toc97970151)

[Figure 6 Dual line graphs 6](#_Toc97970152)

[Figure 7 Maps graph 7](#_Toc97970153)

[Figure 8 Comparison dashboard 8](#_Toc97970154)

[Figure 9 Dashboard 9](#_Toc97970155)

**ASSIGNMENT 2’S ANSWER**

# Introduction

This report is about the Business Intelligent helps people in their work like Data analyze, Data collect. There are 4 parts will be presented in this report, those are Dataset, Pre-process steps on dataset, List and different type of charts and dashboard. Last one is the Data processing steps to analyze data in dataset and give the summary about data above. The dataset in this report is about the price house in North Bend WA from 2014-2015.

# Task 1 Data set and basic information and statistical information related to the data set.

* 1. **Data set**

About the dataset, this is the data about the price of house around the world, special in North Bend WA. This dataset has 21 columns and approximately 21000 rows data of each column. About the number of rooms in each house like bedrooms, bathrooms, floors, view, and features like waterfront, condition, grade, lat, long. The data also have the zipcode to know exactly where the data has been gotten and most of those data is from North Bend WA.

This dataset named "house\_price.csv" contains whether a user purchased the product or not. The dataset has some fields are id, date, price, bedrooms, bathrooms, sqft\_living, sqft\_lot, floors, waterfront, view, etc.

* 1. **Pre-process steps on dataset(s)**

**Step 1: Prepare the dataset for work**

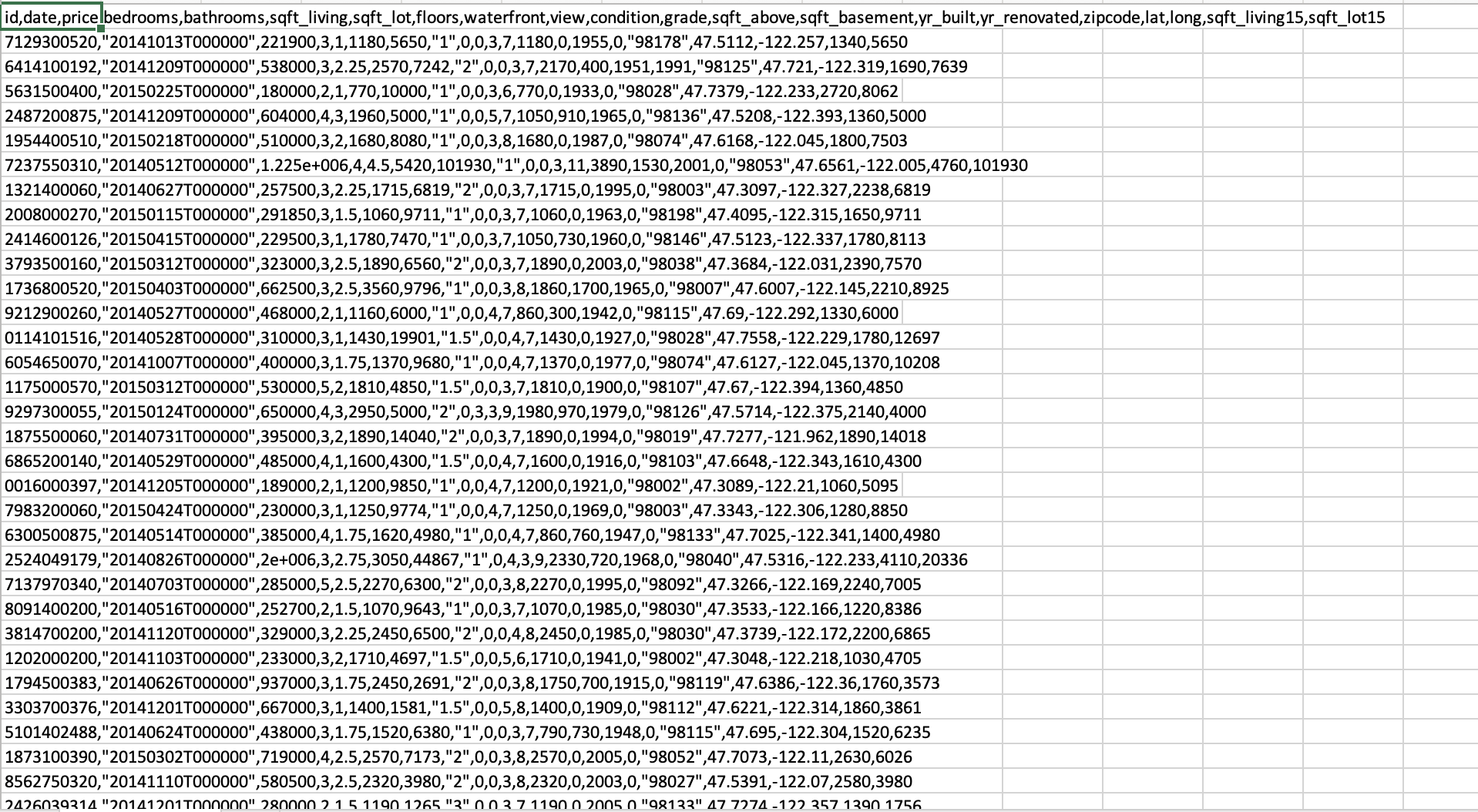


Figure Dataset CSV

**Step 2: Import the dataset to Tableau**

Tableau's interface has three primary components, as shown in Figure. Part 1 is Connect, which is about uploading data from existing files. Tableau allows you to import data from a variety of sources, including Excel, Text, Json, and more. Part 2 is Tableau's primary interface, where you'll find product and data information, as well as my project's visualizations. Discover is the third section, which includes tutorials on how to develop a data analytics solution as well as how and where to source outside. When you create a new Workbook in Tableau, the UI will look like this:

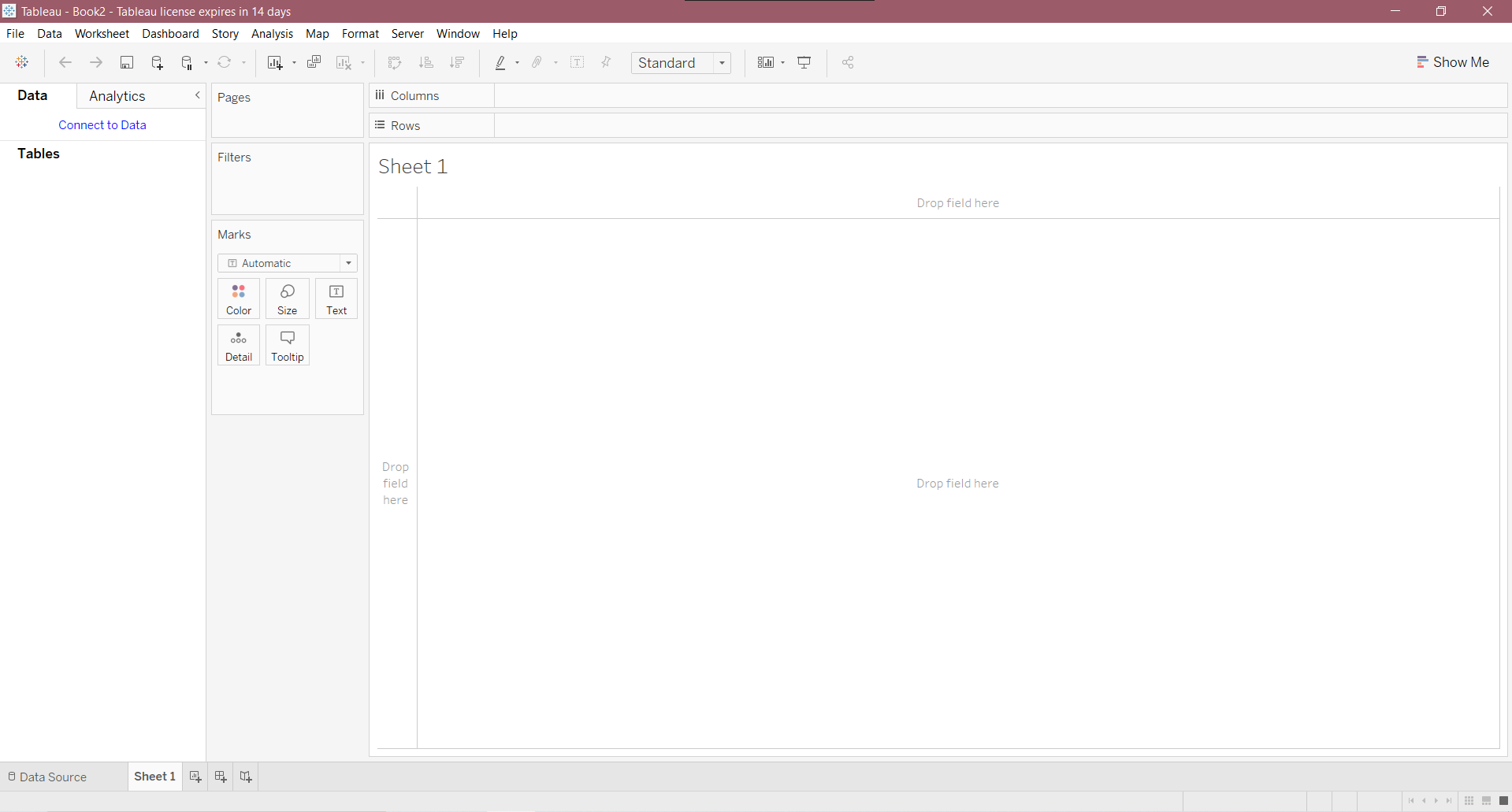


Figure Tableau UI

To analyze the dataset and provide a summary for this project, create a graph (line, table, etc.).

For example, we can use the table graph to readily observe the prices of all houses and the features of each house. With data from Excel, the user can open Tableau and select Microsoft Excel from the drop-down menu to a file. In my situation, the dataset was kept in a csv file, therefore I imported it using a TXT file. Next, select the type of data you wish to import into the dataset. The data from the csv file will be imported into the dataset and displayed on the screen.

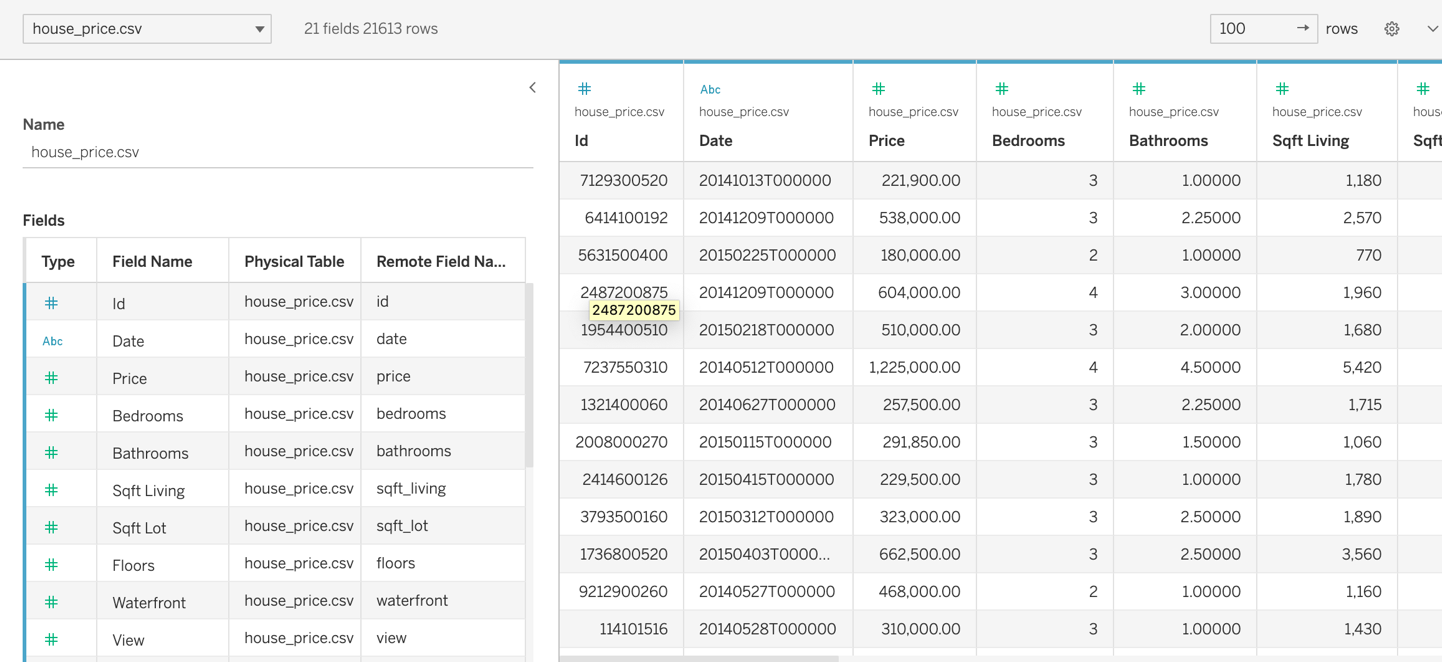


Figure Dataset after imported

**Task 2: List and describe different type of charts for the variable in the data ser.**

First, I create a new workbook and select the table graph to construct a new data graph for analysis after importing the dataset into Tableau. Bathrooms, Bedrooms, Views, Floor, Waterfront, and Price are the six data fields.

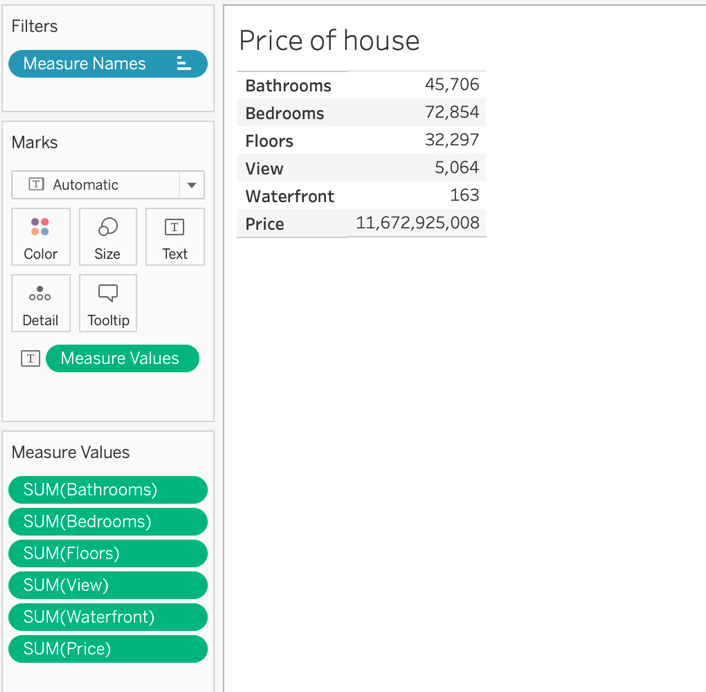


Figure Table in tableau

We can see the summary price of all of our houses in this table by looking at the Price row, as well as the qualities that contribute to that price. We can also use the division to calculate the average price of each one using that information. Another example is a line graph that shows how the price of a house changes over time, whether it increases or decreases.

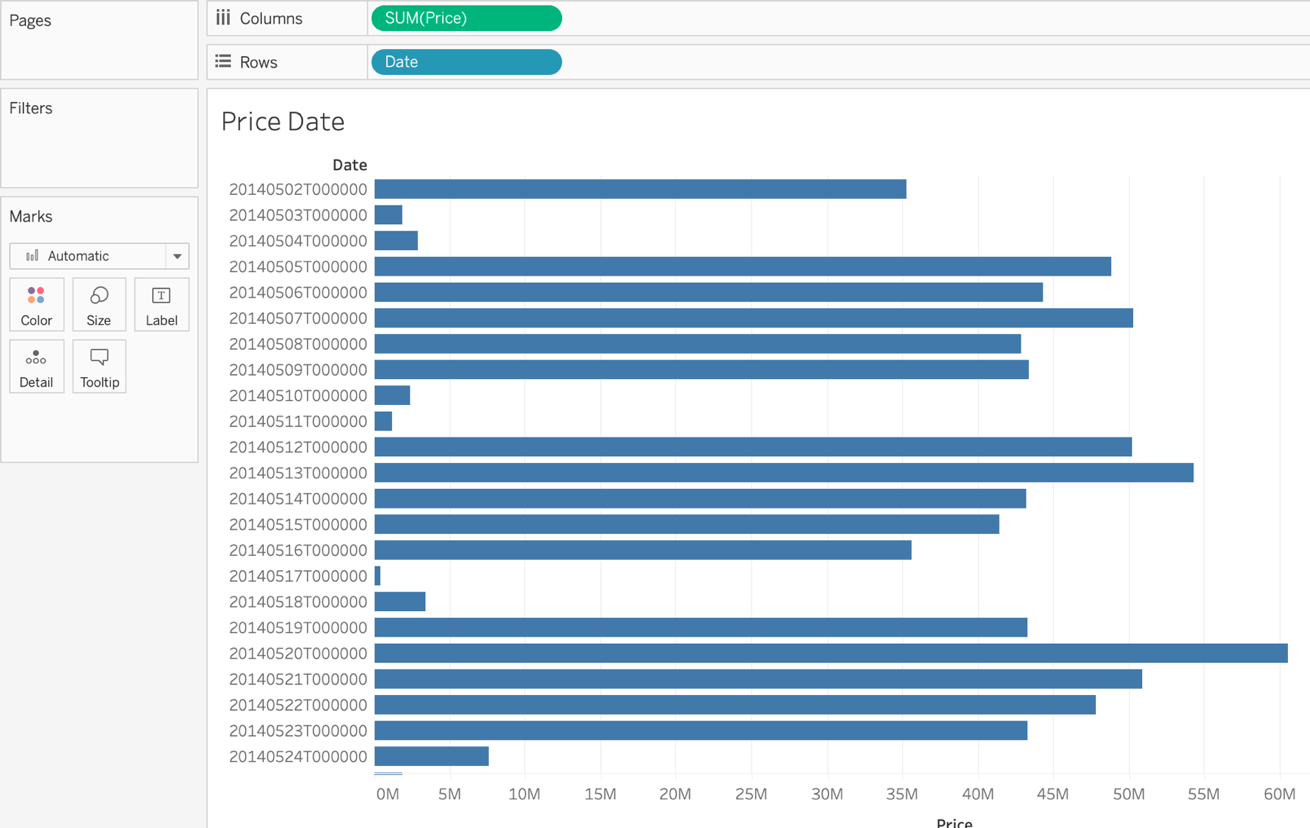


Figure Column graph Price/Date

This one informs us that the cost of our project varies depending on the month of the year. In this line graph, we can see that the maximum price of our house was around 60 million dollars on May 21, 2014, and the lowest was around 500 thousand dollars.

Next, to examine how many total bedrooms and baths were added over time, I decided to utilize dual lines, so you can see two lines of each above on the graph for easy analysis.

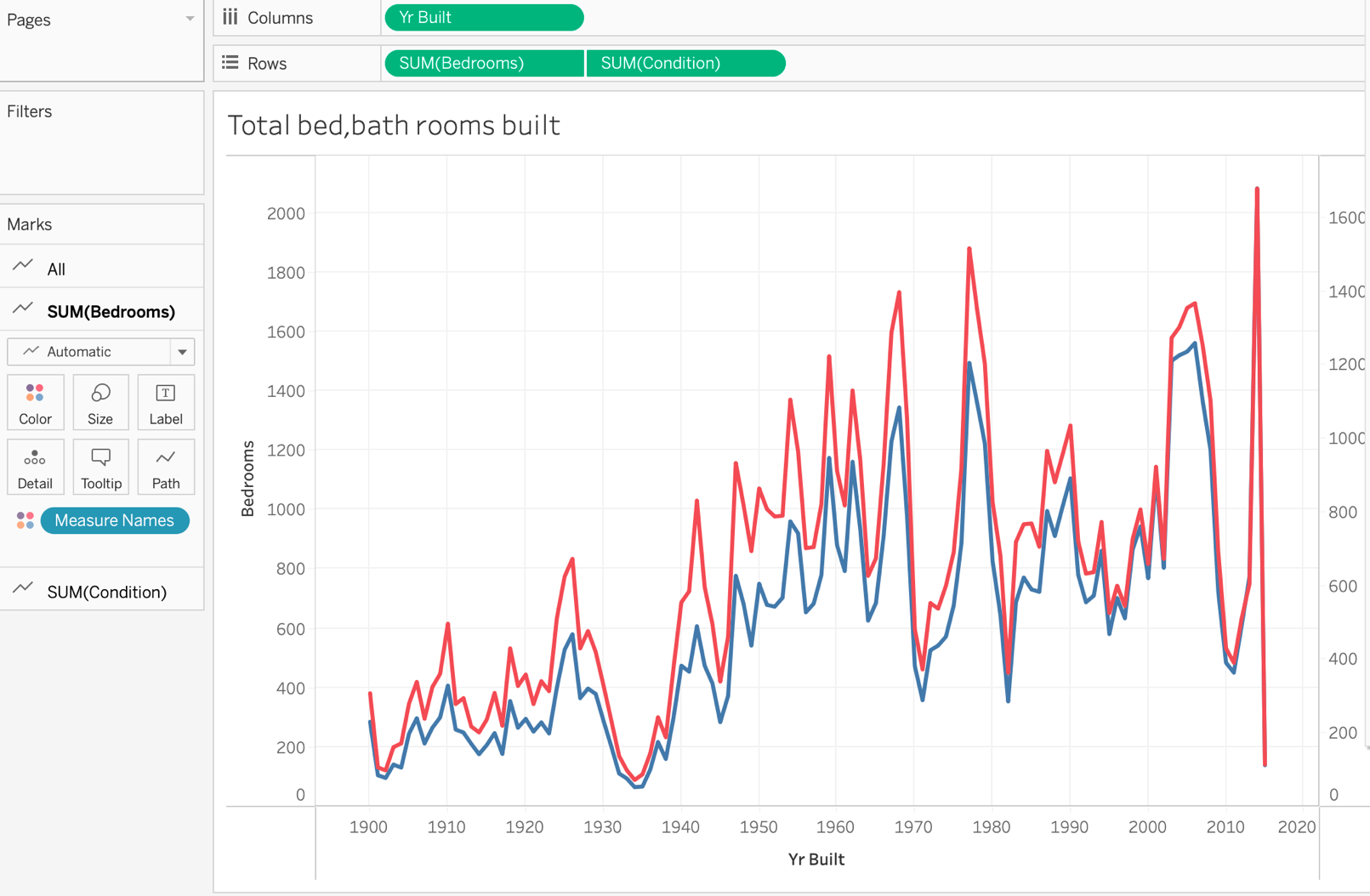


Figure Dual line graphs

We may also use a map with two fields, Zip code and Bathrooms, to show which countries or areas have a strong market for house sales all over the world. Depending on the Zip code, a map of the area will emerge, with a circle indicating the number of bathrooms that have been sold.

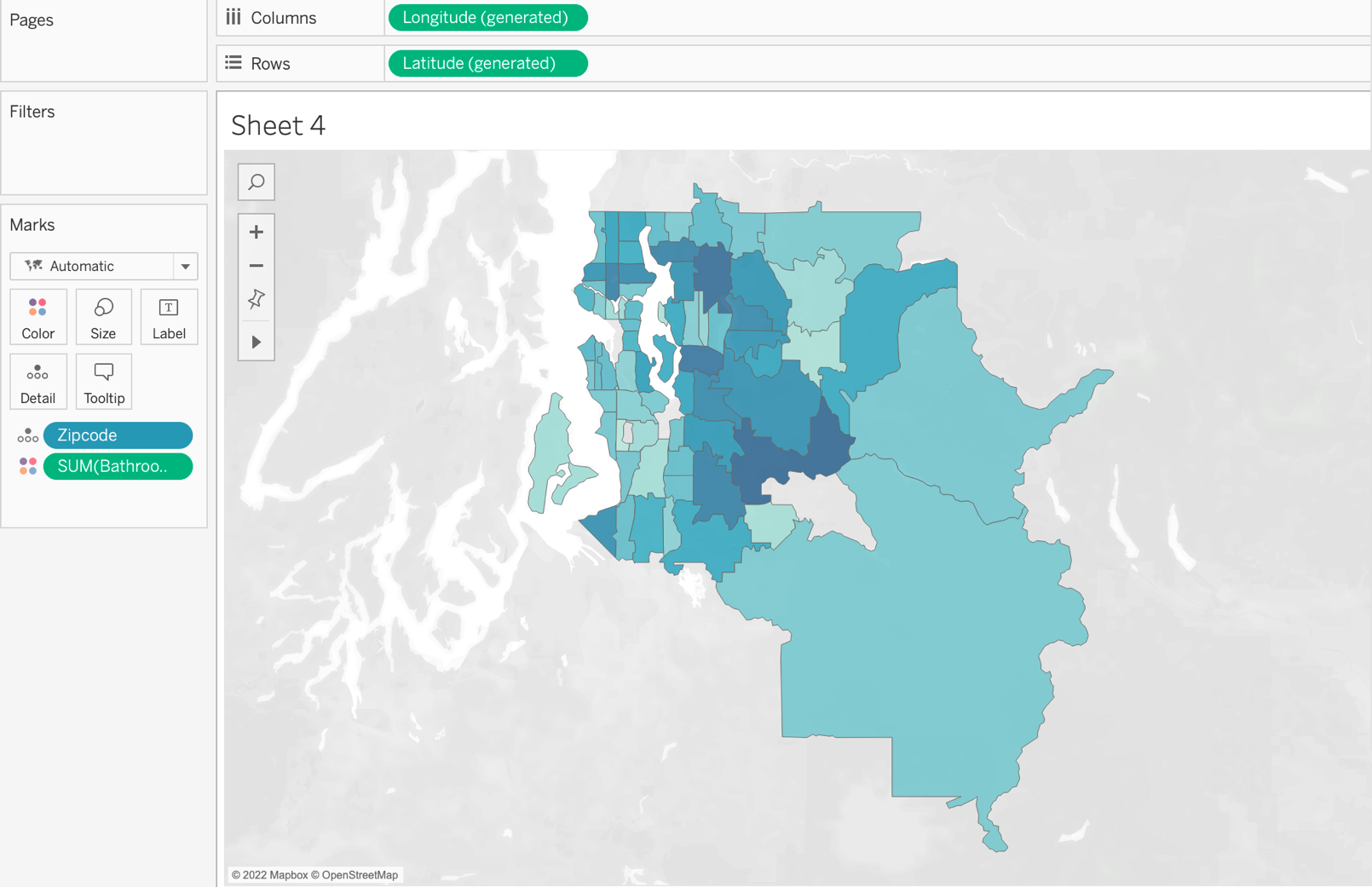


Figure Maps graph

Organizations can utilize business intelligence dashboards to make complex data comprehensible and approachable for non-technical users. Business users can design and see their own dashboards, with content created by IT as a starting point. Non-technical individuals can engage with data via self-service BI. Chipotle utilizes dashboards to improve their analytical procedures and create a consistent view of its restaurant locations. Data visualizations based on dashboards can assist business users in identifying trends. Positive trends can be detected, bad trends can be isolated, and predicted insights can be provided. When Charles Schwab enabled thousands of bank branches to construct their own performance dashboards, they were able to track consumer satisfaction with their products without having to sift through pages of spreadsheets.

**Task 3: Dashboard and purpose of dashboard**

The design of the suggested BI Tool that Courts would use to analyze the Price and Date of their assets every year is depicted in the image above. This would be used to help with a better understanding of how Houses are carried out in comparison to how much money is made when they sell their houses.

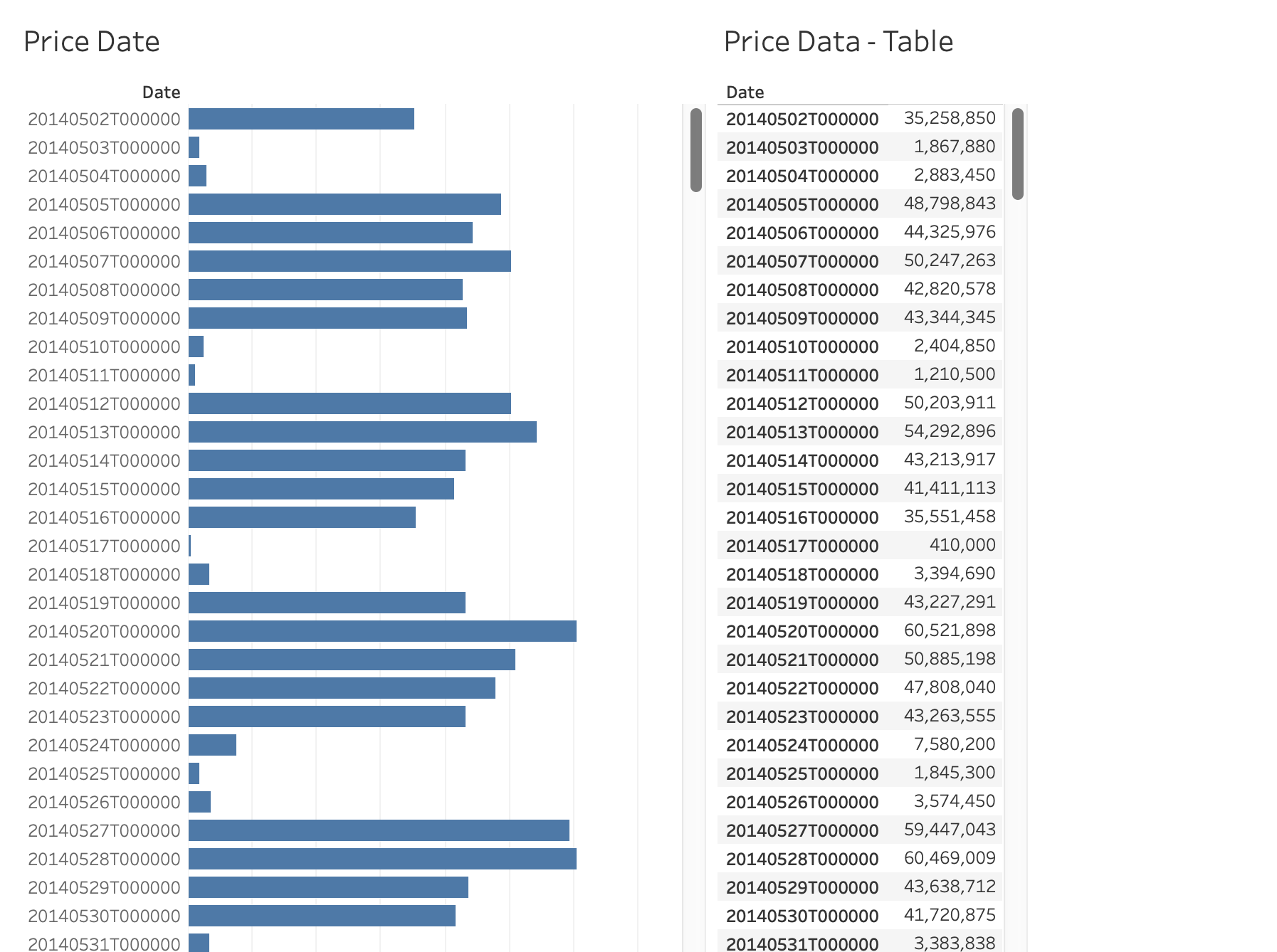
****

Figure Comparison dashboard

**3.1.** **Each dashboard contains charts (Purpose of chart)**

House Price - Table: To view the entire number of rooms/features as well as their overall cost.

Bathrooms / Maps - Maps: Understand the area that has a strong market for sold houses due to previous sales.

Total number of bed/bath rooms constructed – dual lines: This one creates two lines to make it easier to control which room is the most significant to people.

Price Data - Column Graph: See the price of a house each year and the profit from selling a house each year.

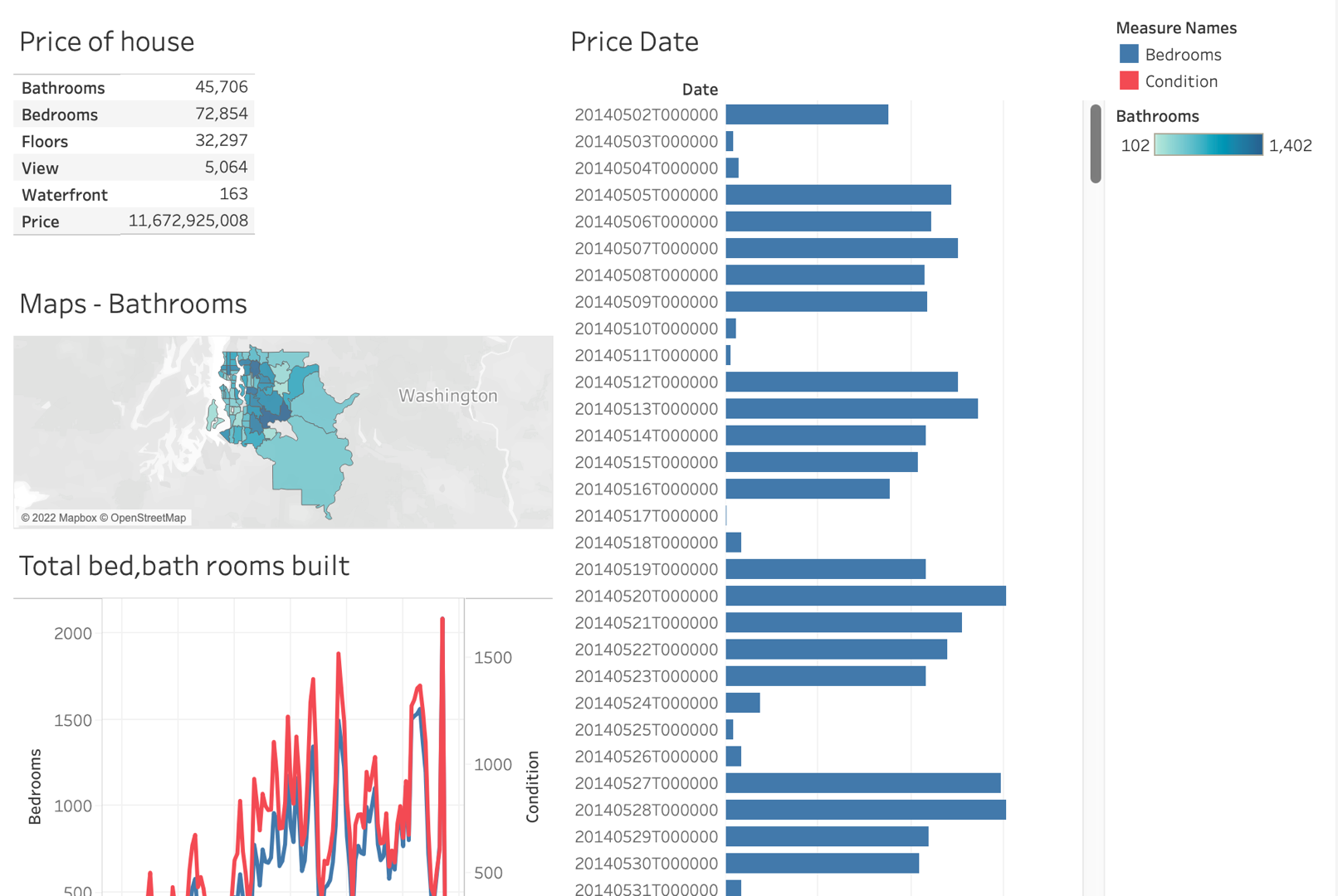


Figure Dashboard

**3.2. Interaction on dashboards (Result / Conclusion / Suggestion)**

The images above were taken to show how the slicer and period were used with specific filters to provide data with a variety of dates, times, and branches.

The Dash Boards for the Business Intelligence Tool are depicted in the images above. This tool includes the four analyses described previously: Sales vs Cost, Most Purchased House's rooms, Number of Rooms, and House price of each year. These assessments will now be merged to create a single business intelligence tool that will assist courts by visually illustrating numerous parameters that are important to them.

This tool would also have two filters for Transaction Date and Branch, in addition to a variety of analysis. These filters would allow users to choose specific criteria on the tool and receive an output that reflected their selections. As an example, take a look at the image below.

The standards that was followed for the design of this tool was based from Gestalt Principles of Designs. Some of the standards that were followed are Similarity, Proximity and Common Fate.

* **Similarity Principle** – This deals with certain elements that is categorised in to groups that have similar qualities. Due to human nature it was natural instinct that we would group elements that are visually alike. Based on the dashboard above, it shows to similarities with all 4 graphs. Two graphs on the left-hand side are strictly vertical blue lines where the two graphs on the right-hand side is blue with orange. majority of the charts.
* **Proximity Principle** – This deals with items that are placed closely together which would appear to have a lot more relatable information than when they are spaced apart. Due to proximity, these graphs have been placed together to allow users to recognise it as one main entity. Based on the design of the dash board it illustrates how all the analyses have been arranged and sorted. It shows that all the filters are located in one general area which is at the top right-hand corner, whereas the other area has the respected graphs that would be affected by these same filters.
* **Common Fate** – This deal with elements that are moving together which would be viewed as a single grouped entity once they are moving together. Based on the design on the dash board, it was specifically designed in this manner to place all graphs together to work in one, however four of these graphs were placed together to give a better understanding of what is transpiring throughout Courts.

# Task 4: Data preprocessing steps and its important

* 1. **The important of BI to the business**

Applying the data preprocessing steps to the project, I have divided it into 2 main phases: Data quality assessment, data cleaning

**Data quality assessment**

After carefully reviewing the collected dataset data for its suitability for the project and the consistency it provides, it is well suited for analysis and evaluation for future plans. The datasets with a simple structure that are easy to understand are clearly divided to make the collection easy to check and avoid data mismatches, some mixed data needs to be homogenized.

However, it is impossible to avoid some risks from the built-in data structure, the return of 0 or missing data will cause the analysis to go astray.

However, the aggregated data is quite complete and there are no missing or gaps in the dataset, which saves me from having to process the data and clean up. However, it is also necessary to double check the data carefully to give accurate analysis.

**Data cleaning**

The process of adding missing data and correcting or removing it from the dataset. The date of the house\_price dataset is the data that needs to be cleaned the most in this process because the date data is added in the format of the date and time combined with the time however the time is not needed in this project because the data will be analyzed and evaluated on a daily basis over a period of 2 years (2014-2015)

* 1. **The important of preprocessing steps**

The main reasons to invest in a solid BI strategy and system are:

* **Gain New Customer Insights**: One of the primary reasons companies are investing their time, money, and efforts into Business Intelligence is because it gives them a greater ability to observe and analyse current customer buying trends. Once you utilize BI to understand what your consumers are buying and the buying motive, you can use this information to create products and product improvements to meet their expectations and needs and, as a result, improve your organization’s bottom-line.
* **Improved Visibility**: Business Intelligent organizations have better control over their processes and standard operating procedures, as the visibility of these functions is improved by a BI system. The days of skimming through hundreds of pages of annual reports to assess performance are long gone. Business Intelligence illuminates all areas of your organization helps you to readily identify areas for improvement and allow you to be prepared instead of reactive.
* **Actionable Information**: An effective Business Intelligence system serves as a means to identify key organizational patterns and trends. A BI system also allows you to understand the implications of various organizational processes and changes, allowing you to make informed decisions and act accordingly.
* **Efficiency Improvements**: BI Systems help improve organizational efficiency which consequently increases productivity and can potentially increase revenue. Business Intelligence systems allow businesses to share vital information across departments with ease, saving time on reporting, data extraction, and data interpretation. Making the sharing of information easier and more efficient permits organizations to eliminate redundant roles and duties, allowing the employees to focus on their work instead of focusing on processing data.
* **Sales Insight**: Sales and marketing teams alike want to keep track of their customers, and most utilize Customer Relationship Management (CRM) application to do so. CRMs are designed to handle all interactions with customers. Because they house all customer communications and interactions, there is a wealth of data and information that can be interpreted and used to strategic initiatives. BI systems help organizations with everything from identifying new customers, tracking and retaining existing ones, and providing post-sale services.
* **Real-Time Data**: When executives and decision-makers have to wait for reports to be compiled by various departments, the data is prone to human error and is at risk of being outdate before it’s even submitted for review. BI systems provide users with access to data in real-time through various means including spread sheets, visual dashboards, and scheduled emails. Large amounts can be assimilated, interpreted, and distributed quickly and accurately when leveraging Business Intelligence tools.
* **Competitive Advantage**: In addition to all of these great benefits, Business Intelligence can help you gain insight into what your competitors are doing, allowing your organization to make educated decisions and plan for future endeavours.

# CONCLUSION

Business Intelligent is an important part which cannot be missed in every company. This report has mentioned all about the BI in business and the pre-processing to complete the work with dataset collection.

With the dataset, I have designed 4 charts for easily analyse data and also with the dashboard, charts is summarized into it and show out for the users. However, many parts cannot be completed or done because of missing data or incorrect data.

# REFERENCES

Jake, F (2021) Business Intelligence (BI) [online] Available at: <https://www.investopedia.com/terms/b/business-intelligence-bi.asp>

Tableau () Real-World Applications of Business Intelligence (BI) [online] Available at: <https://www.tableau.com/learn/articles/business-intelligence-examples>

Jayanthi, R (2009) BUSINESS INTELLIGENCE: CONCEPTS, COMPONENTS, TECHNIQUES AND BENEFITS. Pg. 61-62

Jayanthi, R (2009) BUSINESS INTELLIGENCE: CONCEPTS, COMPONENTS, TECHNIQUES AND BENEFITS. Pg. 68-69

Omnisci () Business Intelligence [online] Available at: <https://www.omnisci.com/technical-glossary/business-intelligence>

Elitsa, K (2020) How to Create a Matplotlib Bar Chart in Python? [online] Available at: <https://365datascience.com/tutorials/python-tutorials/bar-chart-python-matplotlib/>

Analytics (2019) Why Business Intelligence Is Important [online] Available at: <https://analytiks.co/why-business-intelligence-is-important/>