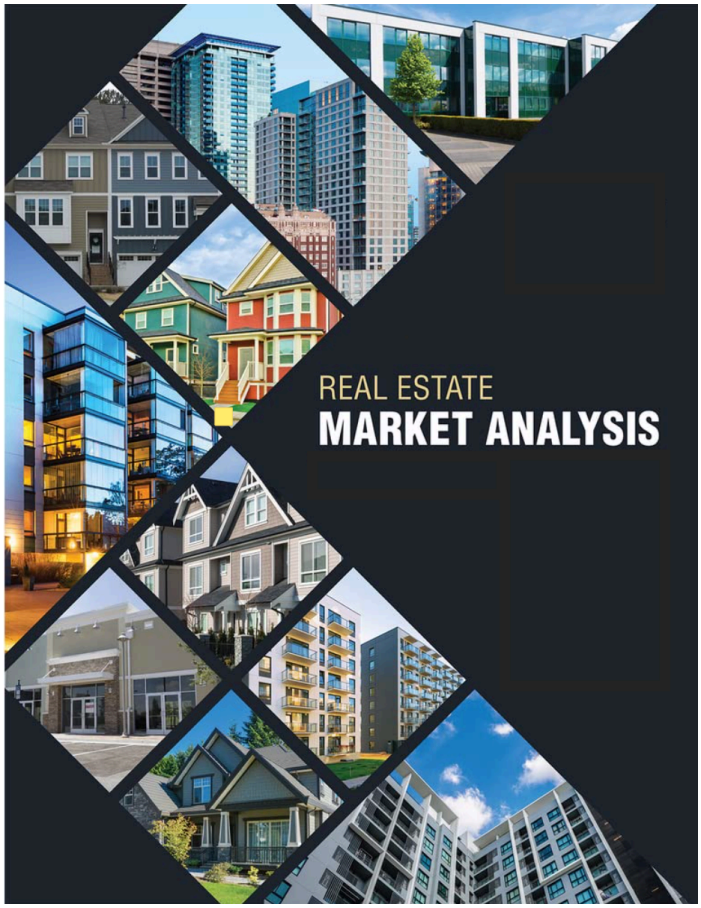


NARRATIVE VISUALISATION PROJECT (MELBOURNE REAL ESTATE MARKET ANALYSIS)

Name: Parul

Student id: 29507960

Tutorial Number: 01-P1

Tutor Name: Xiaojiao Du

Table of Contents

Introduction

For living, Melbourne is considered as the number one city according to the global liveability surveys because of its thriving art and welcoming culture. Melbourne is beloved by young and independent people for the buzz in the city, bustling cosmopolitan atmosphere and everything at doorsteps. Melbourne offers a unique network of inner-city laneways with one of the most extensive tram systems, suburbs on boundaries with waterfront appeal. This makes the Real Estate Business in Melbourne to be one of the most busy and successful Industry.

The Real Estate business is growing, the customers are no doubt struggling to find and choose a perfect house according to their budget and needs. Housing prices are fluctuating over past few years. There are numerous agents in the market trying to provide best deals with a percentage of brokerage. Some well-known brokers have set foots well in the market. Buying property is a very big decision to make. In order to make a smart decision, one needs to have a better understanding of the market and housing prices in different areas.

This visualisation talks about the Housing prices in Melbourne in past few years. This visualisation shows the average price of different areas, how the price fluctuates over time, how the sales trend changed for different types of properties. This also begets the question of price distribution for different properties in the nearby areas. And how can the crime incidents report rate in that area be related to the property price of that area.
The intended audience of this visualisation includes Real estate brokers and the citizens of Australia, planning to invest in some Real Estate in Malborne.

Design

Following is the description the design of visualisation based on 5 design sheets:

Design Sheet 1:

This sheet is a brainstorm of ideas, that could be implemented on the dataset for visualisation such as,

- Choropleth map – to show the average price of properties by council area
- Bubble map – size of the bubble showing the crime incident reported rate for each council area.
- Scatter plot– to show the reported crime incidence by council area
- Nightingale rose map – to show sales size each month every year
- Box plot – for price distribution of properties within a suburb
- Bubble plot – for price by quarter year, with size of bubble by sales.

Design Sheet 2:

This sheet is developed on first combined and refined idea of layout. This sheet includes a combined view of choropleth map and bubble map of Local Government Area of Melbourne with colour saturation based on the average price of each Council Area and size of the bubble showing the rate of crime incidents reported, with some filters available on the number of bedrooms in a property and its distance from CBD. On clicking the Council Area in choropleth map, there will be a focused view contains a box plot of distribution of price in the suburbs for a selected Council Area. The popup will include the median, maximum and min price for that suburb.

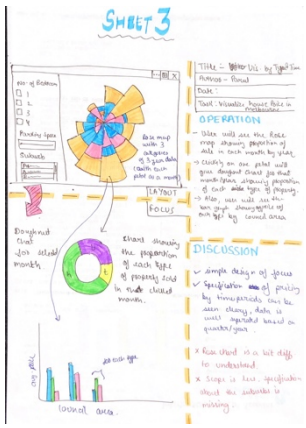
Advantage: can compare suburbs and council areas, scope of the data is covered by covering all suburbs.

Disadvantage: difficult to implement, comparison in based on all type of properties.

Design Sheet 3:

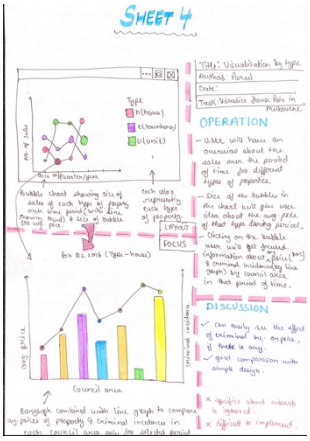
This sheet is developed on the second category. This visualisation starts with showing a nightingale rose map for proportion of sales every year each month. User will choose an area for a selected timeframe. The focused view for the given timeframe will show to plots. First, the doughnut map with 3 sections for each type of data. The second plot will be multiclass bar plot by Council Area, with 3 bars for 3 types of property for each council area. These two graphs will be displayed in one row.

Advantage: simple design of focus, data is well separated based on month and year.



Disadvantage: specification about the suburb is missing, rose chart is a bit difficult to understand.

Design Sheet 4:

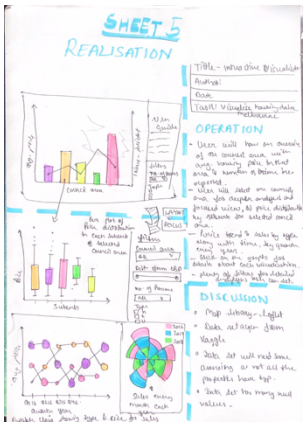


This sheet includes another idea using R. The idea was to allow the user to analyse based on the type of property and Quarter of year. The layout will show the scatter plot with quarter for every year on the x-axis and average price of houses in that period of time with different point for each type in different. The size of the point corresponds to the number of sales. On selecting a particular type or time period, user will get a focused view of price of the selected property type for each council area using bar graph and the line graph for the rate of crime incidents reported in each council area in the selected year.

Advantage: can easily compare the effect of crime rate on price in that area, simple design to understand.

Disadvantage: user cannot compare diff type of property in same area, specification about the suburb is ignored.

Design Sheet 5:



Sheet 5 is the final Design sheet with combination of sheets 1,2 and 3. This includes a user interactive bar chart combiner with line chart, a bubble chart, box plot and a polar bar chart. Among the data, the user can apply multiple filters at a time, for detailed analysis such as select council area, number of rooms in a property, type of property, distance from CBD, price range. Bar graph shows the average price and line shows the crime incidents by council area. Box plot shows the price distribution of suburbs for selected Council area, bubble chart shows the average price trend and sales with time. The last chart shows how busy the market was every month each year.

Further modifications were made to the final design, by adding a leaflet map to get localised view of properties, in clusters or without clusters.

Implementation

The dataset used to create visualisation is the wrangled dataset of Melbourne housing data (MELBOURNE_HOUSE_PRICES_LESS.csv) and crime incident report (criminal_incident.csv). The project is implemented using R shiny due to prior knowledge of the language considering the completion of project within the given timeframe and familiarity with the libraries and their implementations. The list of libraries used is given bellow:

- library(ggplot2) – for plotting of polar bar chart
- library(dplyr) – for filtering options
- library(stringr) – to perform operation on string
- library(plotly) – for plotting of bar, bubble and box plot
- library(shiny) – To create shiny app
- library(leaflet) – for map plotting
- library(shinythemes) – select the theme for shiny
- library(RColorBrewer) – for colour palette
- library(shinydashboard) – to create UI

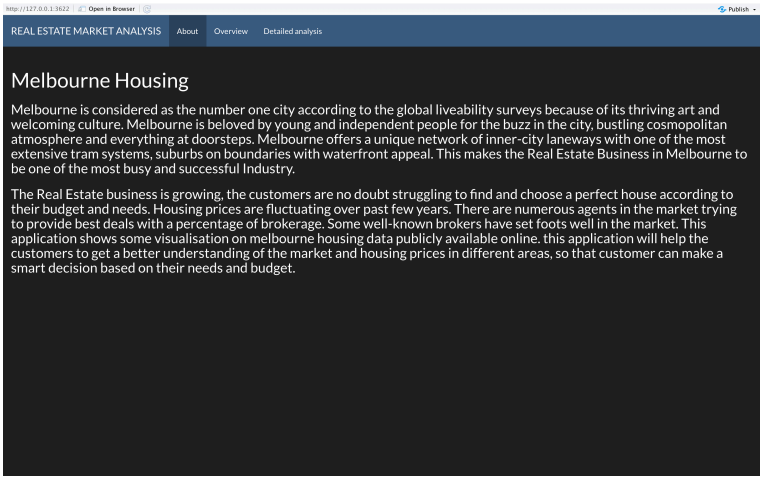
The final visualisation is created using a dashboard in R shiny. Shiny provides an easy way to build an interactive visualisation on large data efficiently. The Visualisation consist of three tabs for introduction (Message), an overview and a detailed analysis. The map is plotted using leaflet package to create interactive maps. The Nightingale rose chart was plotted using ggplot2 library as the plot using plotly was difficult to understand by the user. Other chart like bubble chart or combined bar and line chart are plotted using plotly, it provides the best way of plotting an interactive bubble chart. Plotly is a more sophisticated data visualization library that is better suited for creating elaborate plots more efficiently.

User Guide

The dashboard contains highly interactive visualisations, to provide user all the information in the dataset easily and perform some detailed analysis for better understanding of the Real estate market in Australia. The visualisation contains 3 tabs as follow:

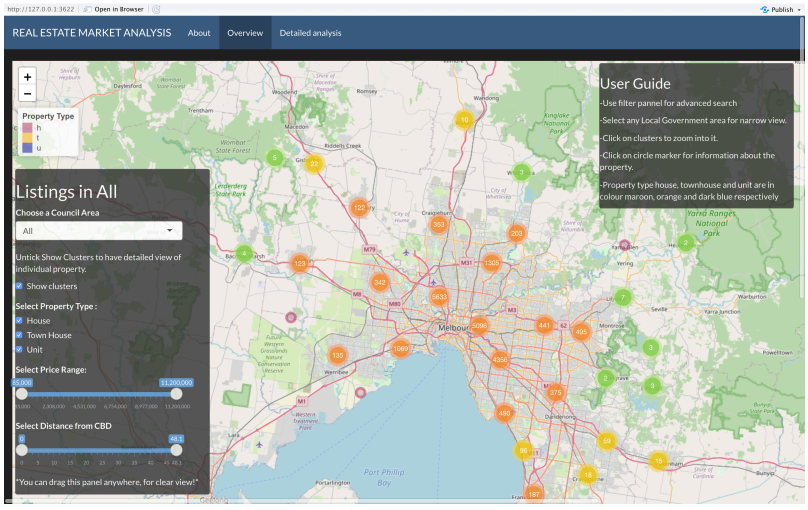
TAB 1: MESSAGE (about the topic)

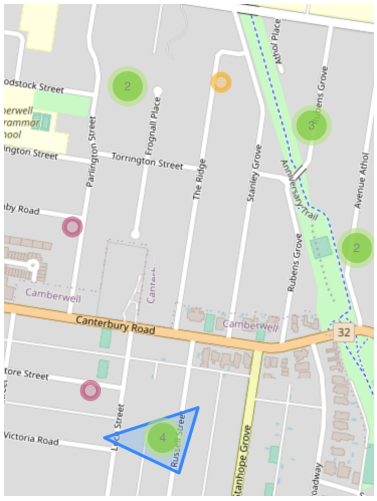
This tab contains a message for the user to give some basic idea about the topic of the visualisation.



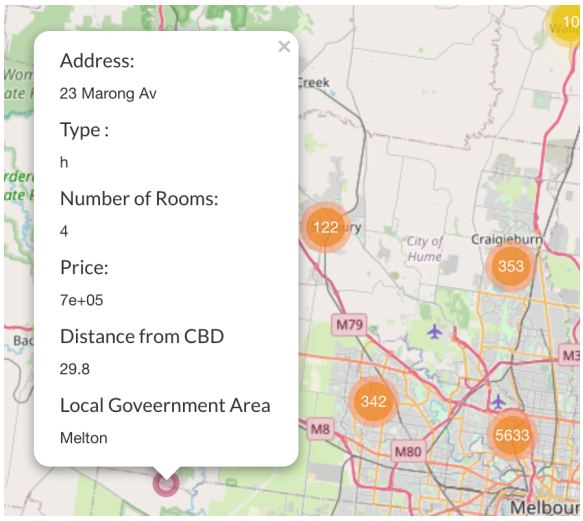
TAB 2: OVERVIEW

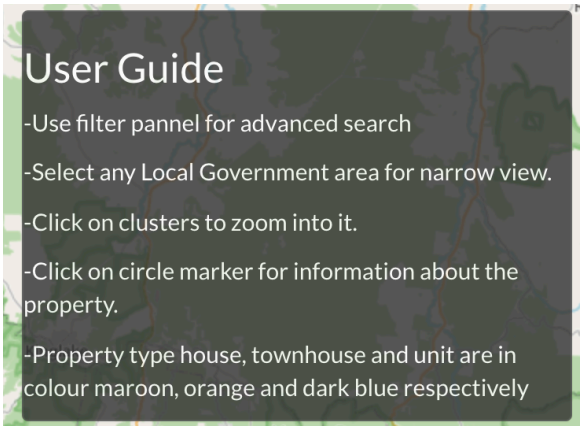
On clicking overview the below tab will be displayed. This tab contains a layout of properties on the map. The clustered leaflet map shows a location of the properties using clusters. User can click the clusters to zoom in and visualise individual property. Click on the cluster will further split the cluster into sub clusters or the individual circle marker of property locations. The page can be scrolled down or click and drag in any direction for clear view.



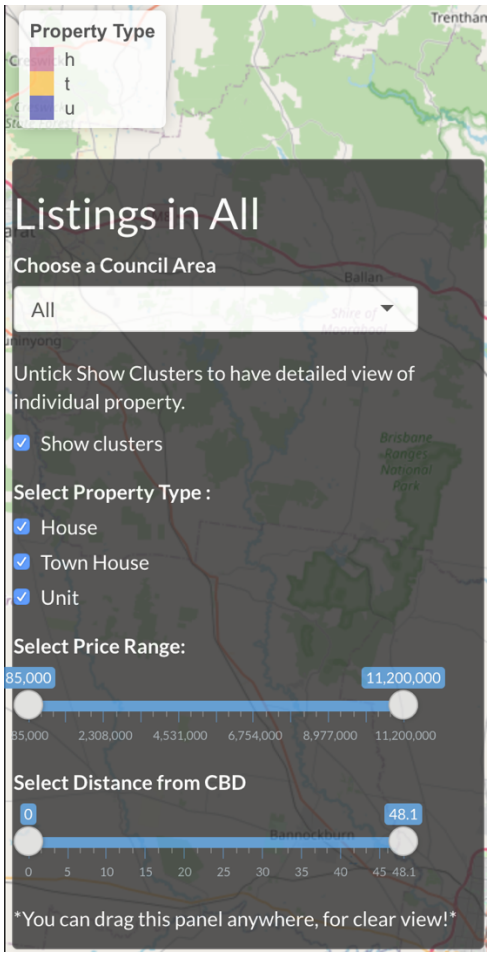


The circle markers have three different colours for each type of property, these markers can be clicked for more details about the property such as address, type of property, price, number of rooms, distance from CBD and Local government area as shown below.





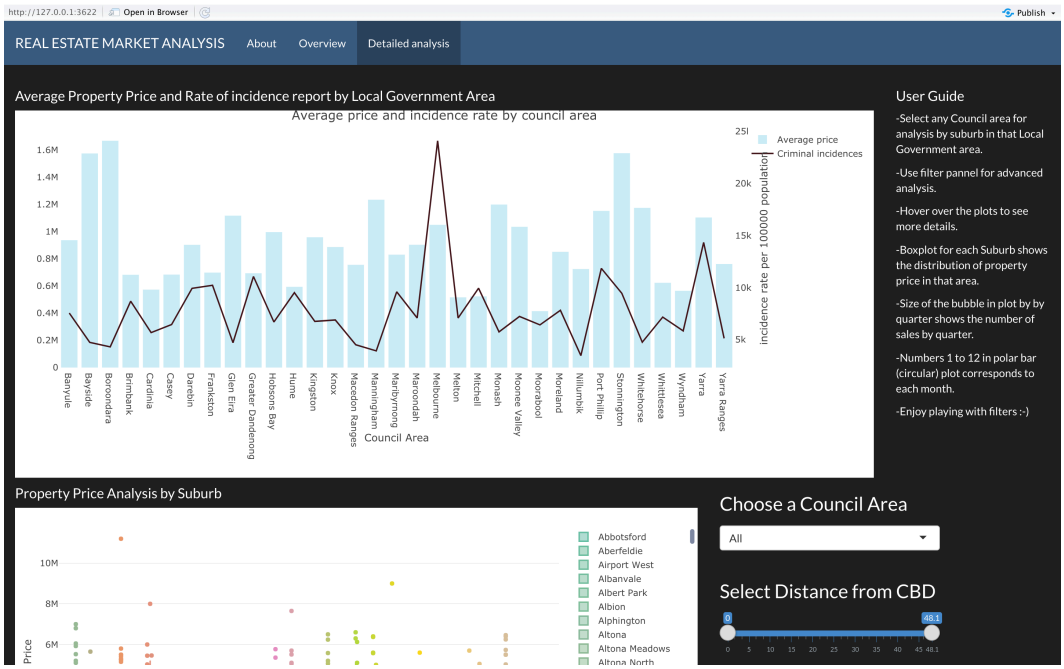
The right image above shows a user guide panel for the user to help control the dashboard and visualise by making changes and playing with given options. The image below shows some more options of filters available for the users to analyse data based on different categories. Also, there are 3 colours maroon, orange and dark blue for 3 types of property house, townhouse and unit respectively as seen in image below.



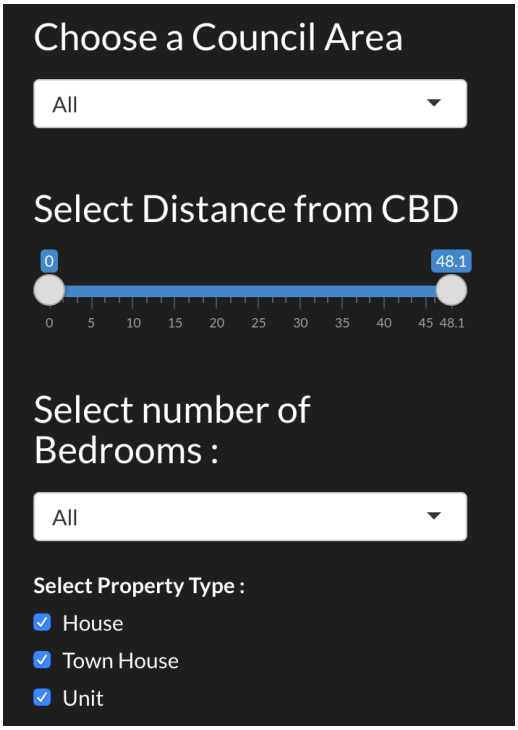
- Select input to select Council Area: the data is for 33 council areas in Melbourne. Choose any one council area for clicking on the name in the list. The selected area will show in the box.
- Checkbox for cluster: untick the cluster box if want an individual circle marker for each property with no clusters.
- Checkbox for selecting property types: there are three types house, townhouse and unit of property presented. Tick and untick the options to choose one or more than one type of property. Just untick the type you are not interested in.
- Slider for adjusting Price range: the price of the property varies between 85000 to 11,200,00. Drag and adjust the slider on the both sides to set the minimum (left) and maximum (right) price range.
- Slider to adjust distance from CBD: there are properties in the CBD as well as approx. 48.1 km away from CBD. Drag and adjust the slider on the both sides to set the minimum (left) and maximum (right) distance from CBD.
- Drag the whole filter panel and User Guide panel if needed.

TAB 3: DETAILE ANALYSIS

On clicking the "Detailed Analysis" the application will display the third tab with all the components. This may take a few seconds to load. This tab contains all the detailed analysis about the Real estate housing market. The pages can be scrolled down to see the rest of the visualisations.

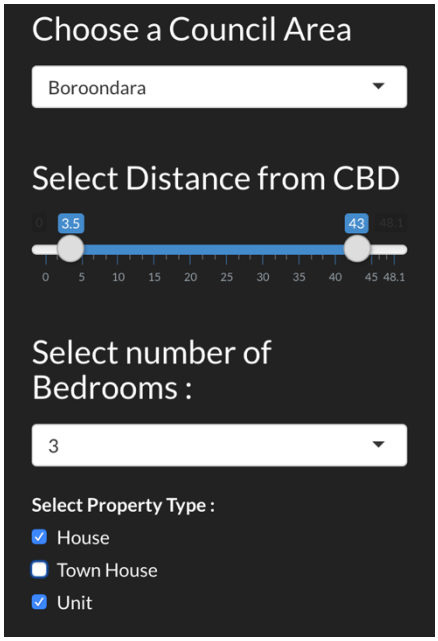


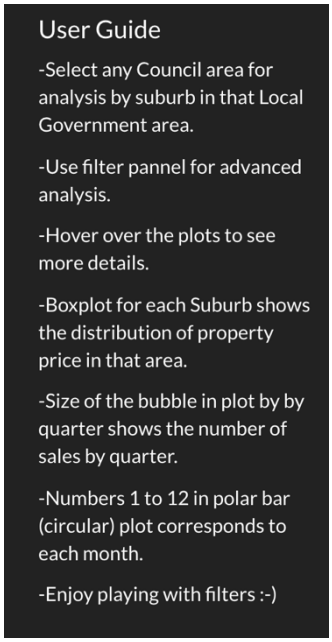
Filter panel and User Guide:



- Select input to select Council Area: the data is for 33 council areas in Melbourne. Choose any one council area for clicking on the name in the list. The selected area will show in the box.
- Slider to adjust distance from CBD: there are properties in the CBD as well as approx. 48.1 km away from CBD. Drag and adjust the slider on the both sides to set the minimum (left) and maximum (right) distance from CBD
- Select input to select Council Area: there are properties with 1 to 12 number of rooms. Choose any one option in the list for clicking on the number. The selected number will show in the box.
- Checkbox for selecting property types: there are three types house, townhouse and unit of property presented. Tick and untick the options to choose one or more than one type of property. Just untick the type you are not interested in.

Below is the image of filter panel with different filters applied.

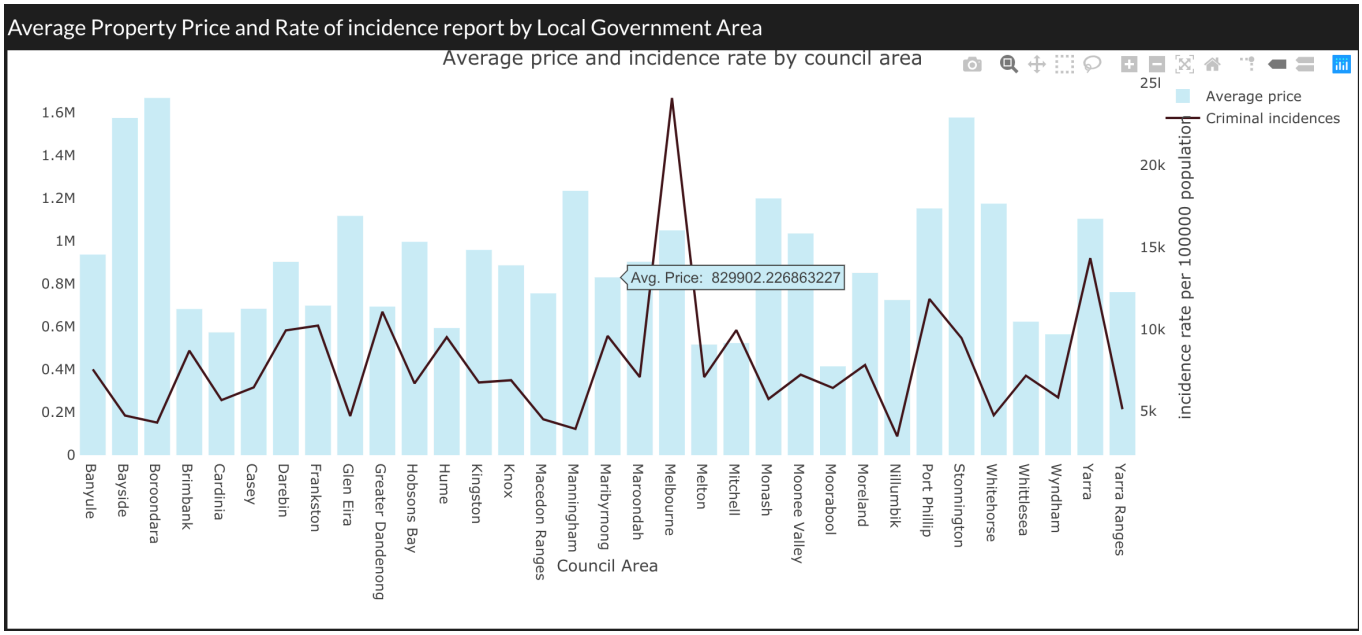




The right image above shows a user guide panel for the user to help control the dashboard and visualise data by making different changes to filter panel according o their choice and playing with given options to do some detailed analysis to reach any conclusion.

There will be four different graphs each showing different information. Without selecting any filters, the information displayed in these charts will be for all the Council areas or each type of property. The four graphs are as follow:

Chart 1: bar chart (for average price) combined with line chart (crime incident report) by council area.

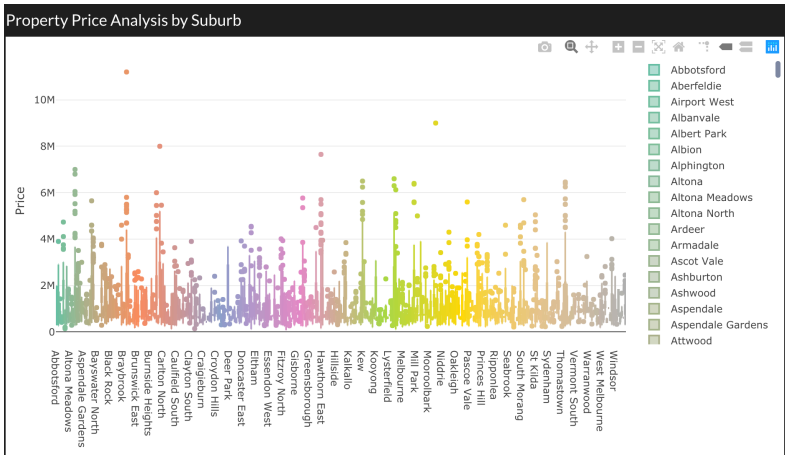


Hover over the bars or the line to get detail about the value of average price or the crime incidents reported.

Chart 2: Box plot for price distribution in each suburb

This graph shows the distribution of price of property among a suburb through data quartiles and averages, and how much the price varies within that Suburb area. One box plot shows the median, upper bound maximum, lower bound and minimum price range for one suburb. Hover over the candles for details about the price range in that suburb.

The left image shows the plot without any filters., for all the Council areas. The Right image shows the plot for the suburbs in Boroondara Council Area only, with filter applied on Council area.



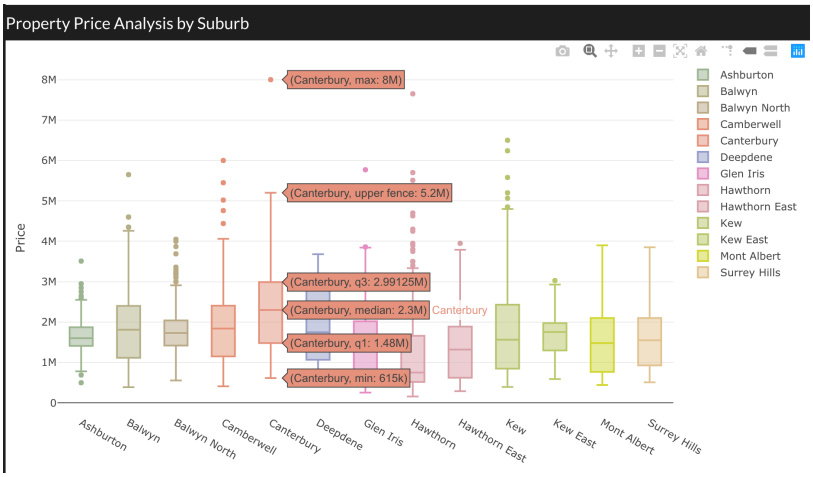
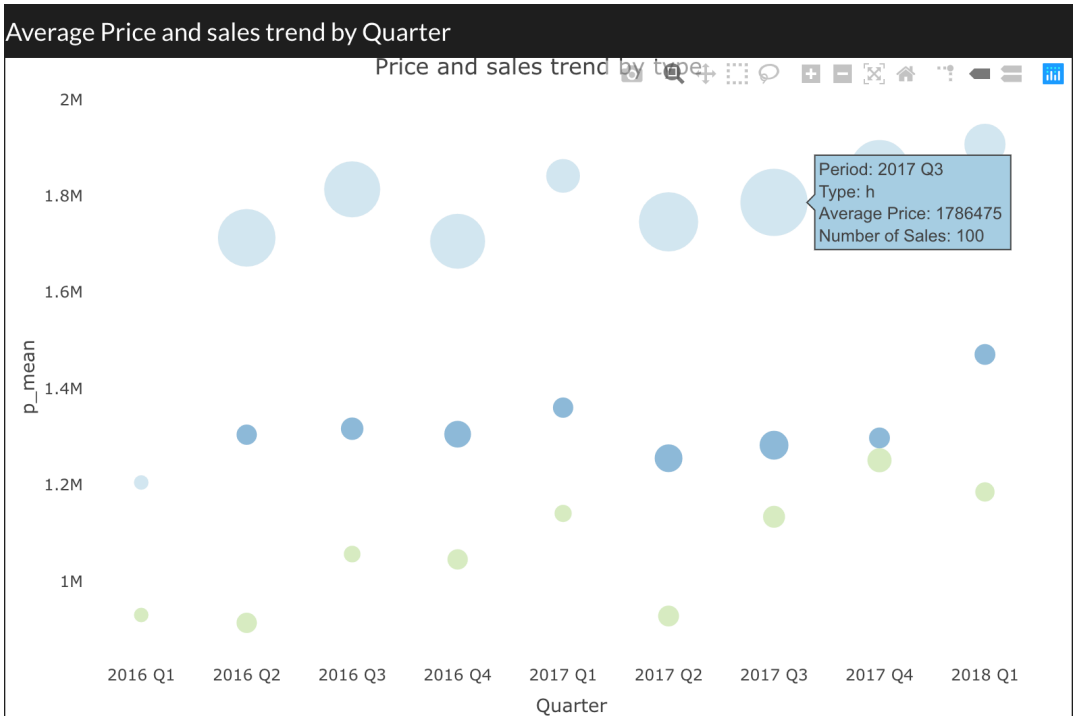


Chart 3: Bubble Plot for Average price and sales trend by Quarter each year.

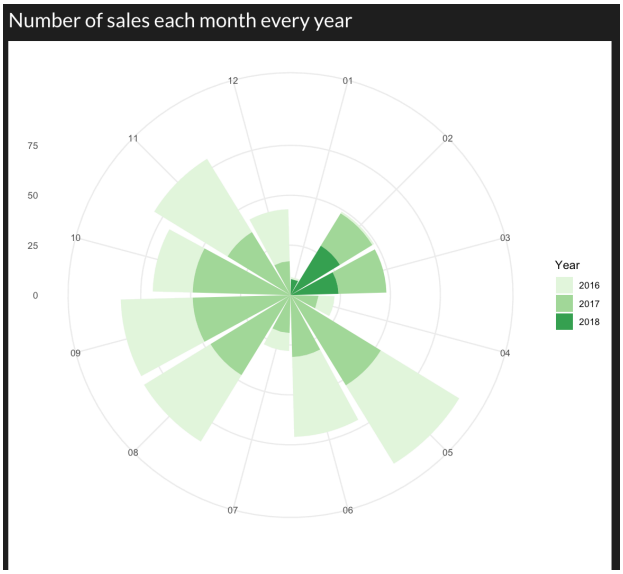
The third plot is a bubble plot for Average price and sales trend by Quarter each year. The position of the bubble shows the average price rate during that time frame, the size of the bubble corresponds to the number of sales took place in that time period and different colour of the bubble is for different property type. The bigger the bubble, the more properties were sold. The higher the bubble, the higher the average price rate was.



The above plot is with filters applied above in filter panel. Hover over the bubble to see details about the bubble.

Chart 3: Polar Bar chart for number of sales every month each year.

This graph shows the sales every month each year (2016, 2017, 2018). The numbers for each sector from 1 to 12 corresponds to month from January to December. The area of the sector is proportional to the sales. The graph below is with filters applied above in filter panel.



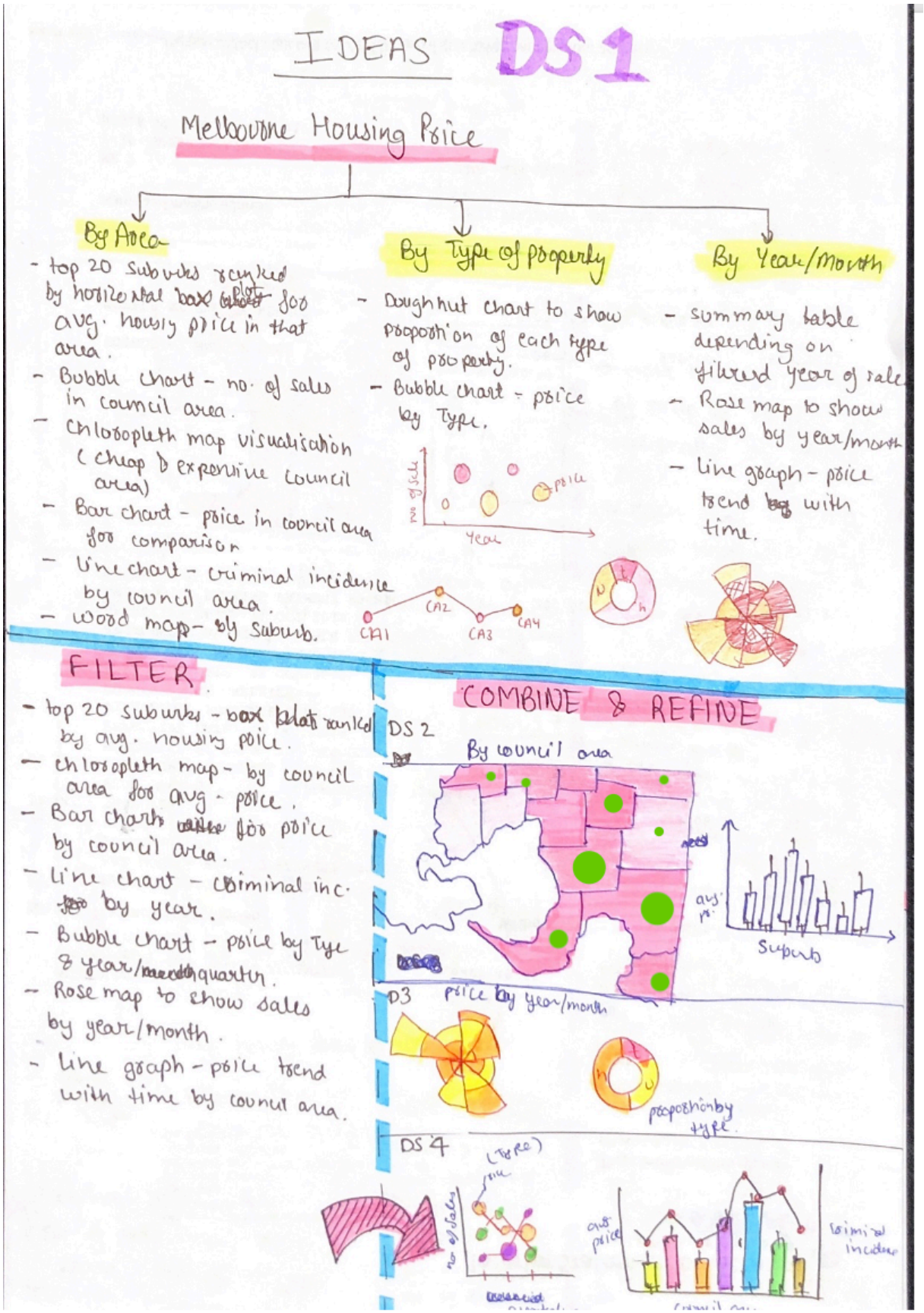
Conclusion

After the detailed analysis of the market using the visualisation made, we have a plenty of knowledge about the housing market to make decisions. Housing has a long-term role to play in investment portfolios, which needs a clear understanding. The data has enabled us to find some really interesting and important insights about Melbourne housing market. Suburb is an important predictive to find if the property is expensive. Boroondara is the most expensive Local Government Area, within which Canterbury is the most expensive suburb.
Also, units are cheaper compared to houses and townhouses in almost every Council Area, while houses are the most expensive. There are more sales of houses than the rest of the two types. Also, the prices seem to fall by the start of year 2017 till quarter-3, year 2017. By the end of the year, market is hot, with higher number of sales.

Reflecting on what I have learnt from the project:

- the dataset had many null values for price which were dropped during exploration process, some changes were made to the process and modify the data.
- Data pre-processing plays a vital role in exploration visualisation. This process can have a huge impact on the results.
- R is better for static visualisation. Doing project on D3 could have allowed me to be more interactive and creative and cover the scope of the data to fullest.
- The proposal consists of two data sets, but no correlation could be obtained between the two.

Appendix


IDEAS
DS 1
Melbourne Housing Price
By Area
- top 20 Suburbs ranked by horizontal box plot for avg. housing price in that area.
- Bubble chart - no. of sales in council area.
- Chloropleth map visualisation (cheap & expensive council area)
- Bar chart - price in council area for comparison
- line chart - criminal incidence by council area.
- word map - by Suburb.
By Type of property
- Doughnut chart to show proportion of each type of property.
- Bubble chart - price by Type.
price
Year
CA1
CA2
CA3
CA4
By Year/Month
- summary table depending on filtered year of sale
- Rose map to show sales by year/month
- line graph - price trend with time.
FILTER
- top 20 suburbs - box plot ranked by avg. housing price.
- chloropleth map - by council area for avg. price.
- Bar charts for price by council area.
- line chart - criminal inc. by year.
- Bubble chart - price by Tye & year/quarter.
- Rose map to show sales by year/month.
- line graph - price trend with time by council area.
COMBINE & REFINE
DS 2
By council area
Suburb
DS 3
price by year/month
DS 4
proportion by type

SHEET 2

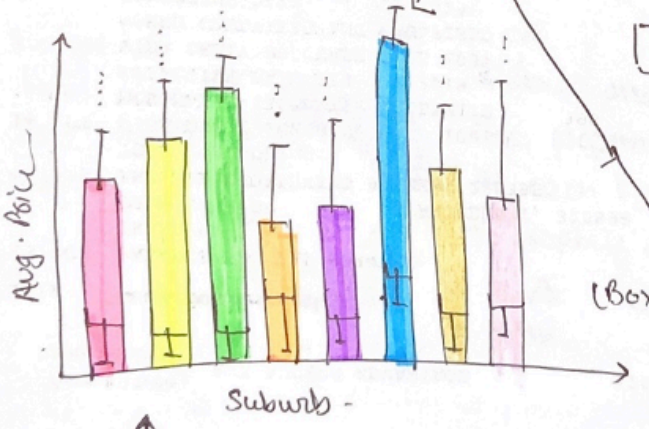
Title - Interactive Vis by Area
Author - Parul
Date - 03/06/2010
Task : Visualize house price by area

LAYOUT

average house price for each local area by chloropleth map.

criteria options (how far from CBD & no. of Bedrooms in property)

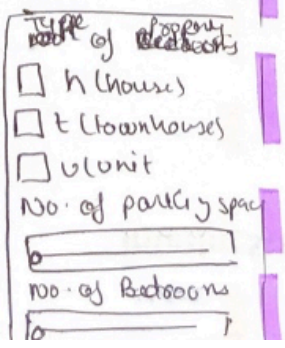
FOCUS



(Box Plot)

showing avg. housing price of each suburb in the selected council area.

criteria filter option (no. of Bedrooms, parking space, type of property)



OPERATION

→ User has the overview of map of melbourne seperated by council area.

→ each area is filled with colour, indicating the avg. house price in that area. (higher the price, darker color). Expensive areas will have darker color.

→ ~~Suburbs~~ & avg price & council area in tooltip.

→ Click on council area to zoom-in for results by suburb in that area.

→ down the map bar chart with suburbs showing price range. by.

DISCUSSION

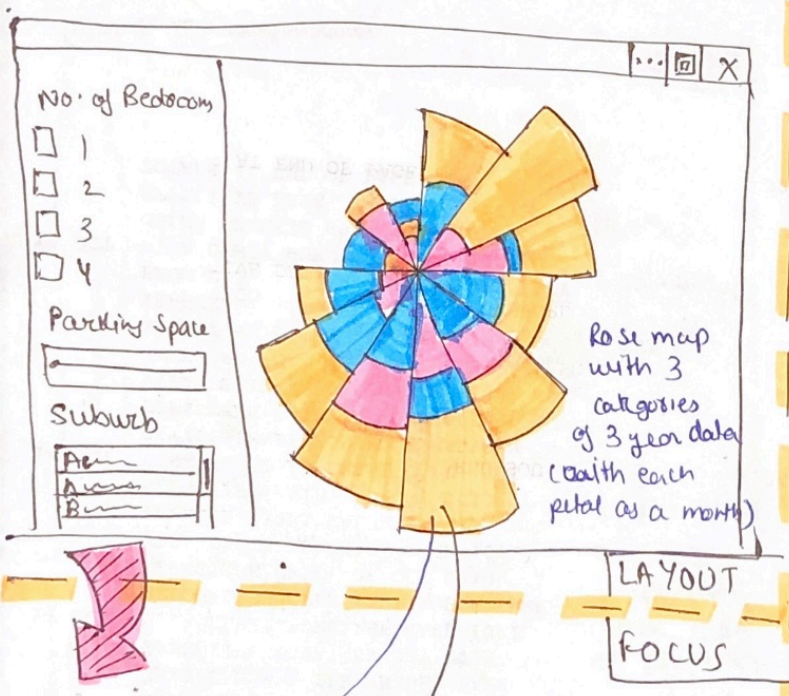
✓ Scope of the data is correct by taking all the suburbs.

✓ can compare suburbs as well as council areas.

X Comparison is based on avg. price, cafe all types of properties (not similar properties)

X Dataset is huge for D_3

SHEET 3



Doughnut chart for selected month.

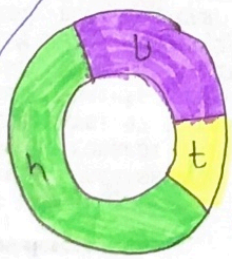


Chart showing the proportion of each type of property sold in that chosen month.

Title :- Vis. by Type & Time

Author - Parul

Date :

Task : Visualize house price in Melbourne

OPERATION

- User will see the Rose map showing proportion of sale in each month by year
- → clicking on one petal will give doughnut chart for that month/year showing proportion of each type of property.
- → Also, user will see the bar graph showing average of each type by council area

DISCUSSION

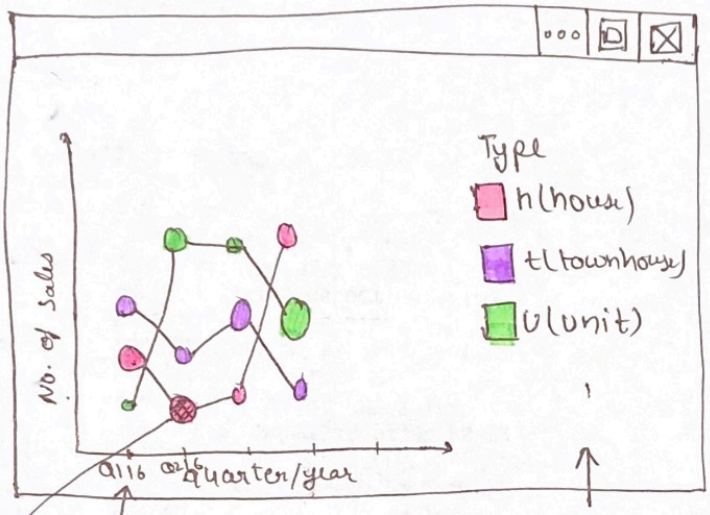
✓ simple design of focus

✓ Specification of pricing by time periods can be seen clearly, data is well separated based on quarter/year.

× Rose chart is a bit diffi. to understand.

× Scope is less, specification about the suburbs is missing.

SHEET 4



Bubble chart showing size of sales of each type of property over time period (with line showing trend) & size of bubble for avg. price.

each color representing each type of property.

for Q2 2016 (Type - house)

Title: Visualisation by type

Author: Parul

Date:

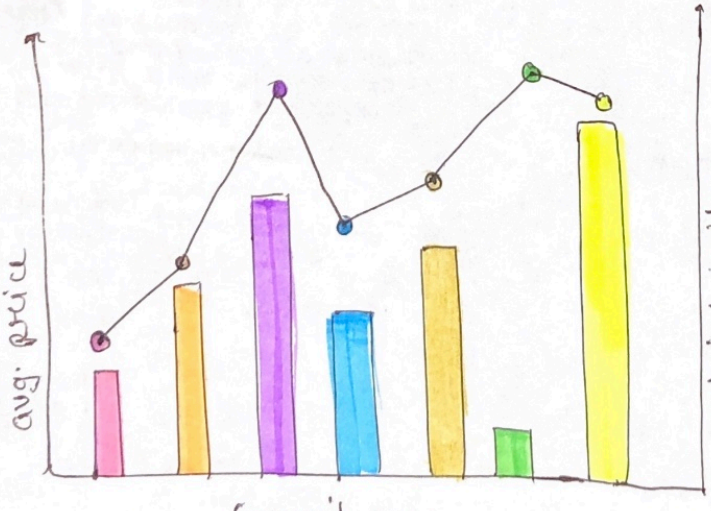
Task: Visualize house Price in Melbourne

OPERATION

- User will have an overview about the sales over the period of time for different types of properties.
- Size of the bubbles in the chart will give user idea about the avg. price of that type during period.
- Clicking on the bubble user will get focused information about avg. price ($ bar) & criminal incidences (by line graph) by council area in that period of time.

LAYOUT

FOCUS



Bargraph combined with line graph to compare avg. prices of property & criminal incidences in each council area only for selected period

DISCUSSION

✓ can easily see the effect of criminal inc. on price, if there is any.

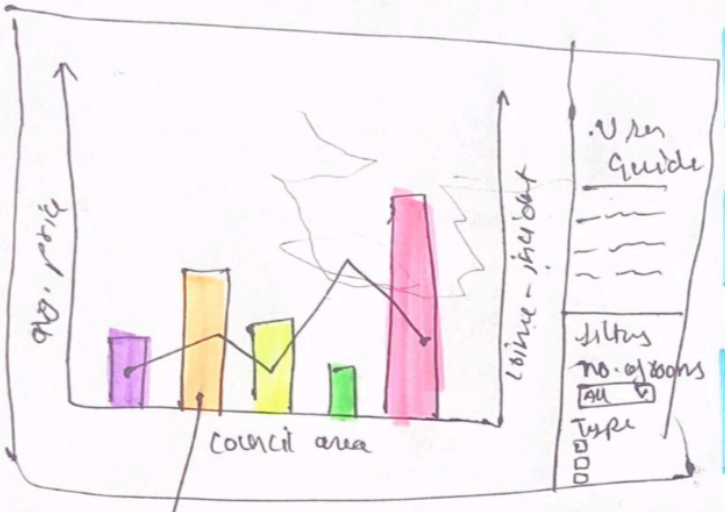
✓ good comparison with simple design.

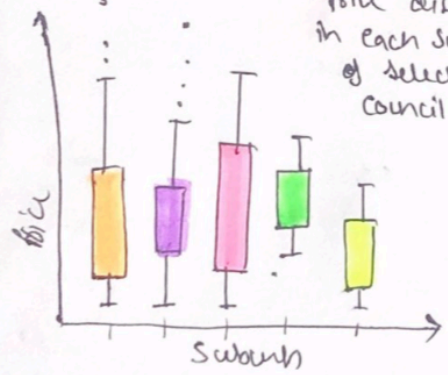
× specifics about suburb is ignored.

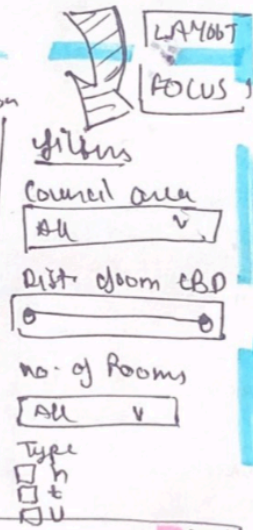
× difficult to implement.

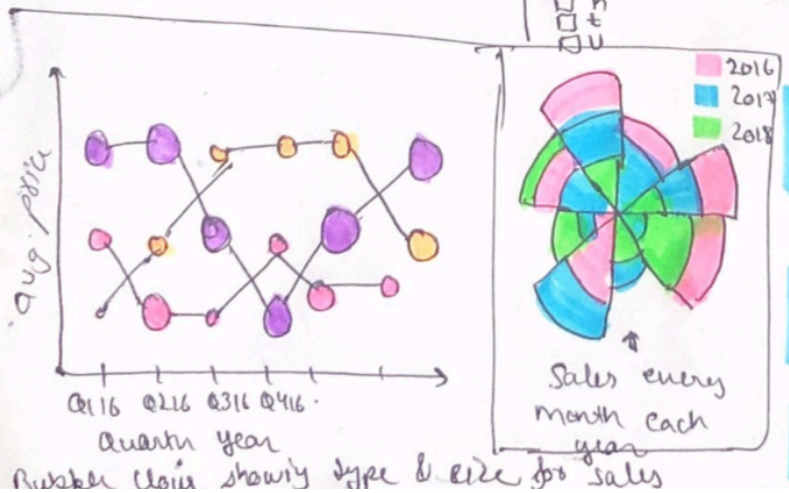
SHEET 5

REALISATION









Bubble chart showing type & size for sales

Title - Interactive Visualisation

Author:

Date

Task: Visualise housing data Melbourne

OPERATION

- User will have an overview of all council area with avg. housing price in that area & number of crime incidents reported.
- User will select one council area for deeper analysis and focused views, as price distribution by suburb for selected council area.
- Price trend & sales by type along with time, by quarter every year.
- Click on the graphs for details about each visualisation.
- plenty of filters for detailed analysis user can set.

DISCUSSION

- Map library - leaflet
- Data set given from Kaggle
- Data set will need some cleansing as not all the properties have typ.
- Data set has many null values.