FIT5147 Data Exploration and Visualisation

Yehezkiel Efraim Darmadi
34078215
ydar0001@student.monash.edu

Table of Contents

Data Loading, Checking and Cleaning

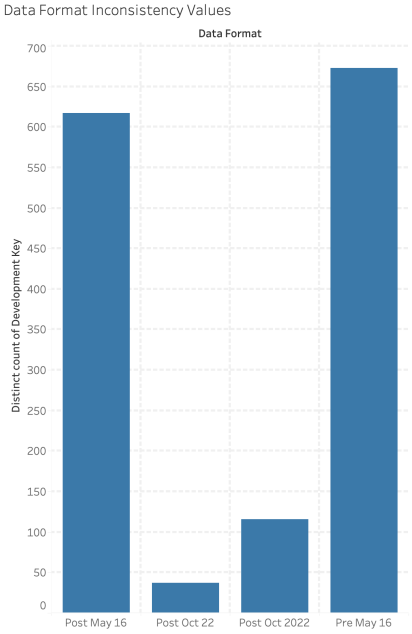

Figure 1. The Inconsistency in Data Format Column and the Number of Inconsistent Values

Figure 1 highlights inconsistencies in the Data Format column, where 'Post Oct 22' and 'Post Oct 2022' should represent the same value. To standardize the values, I created a calculated field using an if statement to convert all 'Post Oct 2022' entries to 'Post Oct 22'. I chose 'Post Oct 22' because it aligns with the format used by the other entries, even though it appears less frequently than 'Post Oct 2022'.

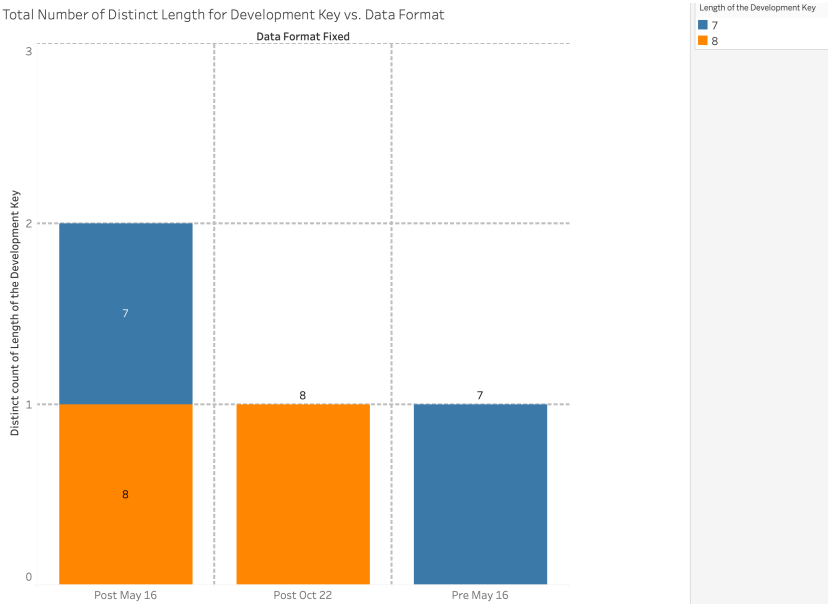

Figure 2. The Inconsistency in Development Key's Length per Data Format

The Development Key's length is inconsistent across the Data Format values, as shown in Figure 2, which reveals two different lengths. I used a calculated field to standardize the length to 8, aligning with the format indicated by the last format the dataset uses.

Street Address Inconsistency

Property Id	Street Address	Town Planning Application	Data Format Fixed	
100108	366-422 Abbotsford Street NORTH MELBOURNE VIC 3051	TPM-2021-5/B	Post Oct 22	Abc
	366-422 Abbotsford Street, North Melbourne, 3051	TPM-2021-5	Post May 16	Abc

Figure 3. The Inconsistency in Street Address Column

Figure 3 highlights inconsistencies in the Street Address column, where values are in capital letters and include 'VIC' and commas. I used a calculated field to convert the street addresses to lowercase and remove the 'VIC' and commas.

Data Exploration and Presentation

Question A

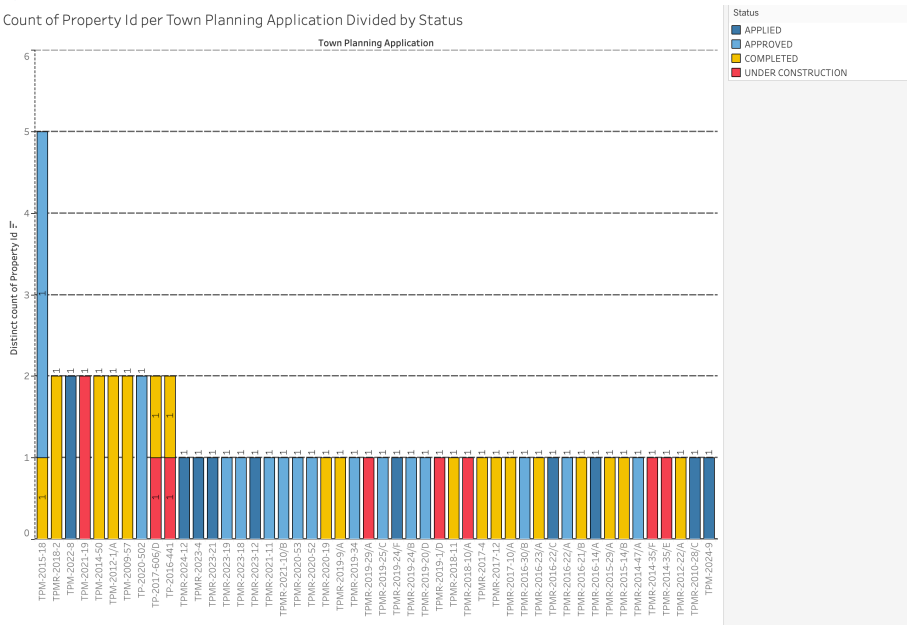

Figure 4. Count of Property Id for each Town Application Planning coloured by Status

The PDF detailing the data specifies that each Town Planning Application can be associated with multiple Property IDs (City of Melbourne, 2016). Figure 4 illustrates that some Town Planning Applications indeed have multiple Property IDs, indicating that the Town Planning Application is not a unique development identifier. The figure also shows that, on average, each Town Planning Application is intended for a single Property ID or

building. Additionally, it demonstrates that not all buildings or Property IDs associated with a single application will be completed at the same time, as evidenced by the example of Town Planning Application 'TPM-2015-18'.

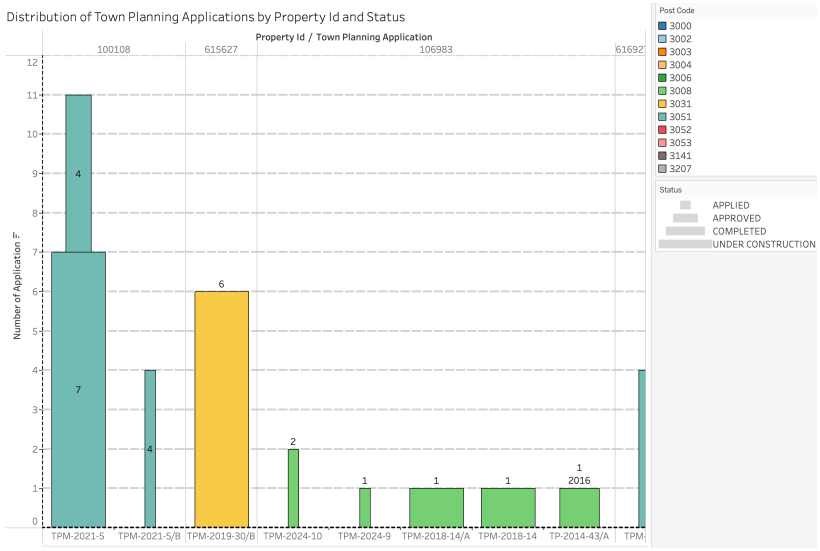

Figure 5. Distribution of Town Planning Application by Property Id Coloured by Post Code and Shaped by Status

Figure 5 reveals that the combination of Town Planning Application, Property ID, and Status is insufficient to uniquely identify each row. The figure also indicates that many Property IDs are clustered within the same neighbourhoods (same postcode). Additionally, it shows that the distribution of Status within a single Property ID varies, suggesting that there might be a prioritisation in the development of buildings.

Number of Distinct Value of Development Keys

Count of DAM-S22024PE1.csv	Count of Fixed Development Key
1,443	1,443

Figure 6. Number of Distinct Values of Development Key Column

Both Figures 4 and 5 demonstrate that the combination of Town Planning Application, Property ID, and Status is not sufficient as a unique identifier for the dataset. This leaves the Development Key as the only categorical variable that could serve as a unique identifier. Figure 6 confirms this by showing that the total number of rows in the dataset matches the total number of distinct Fixed Development Keys, indicating that each row can be uniquely identified by the Development Key. Therefore, the Fixed Development Key is the unique identifier for the dataset.

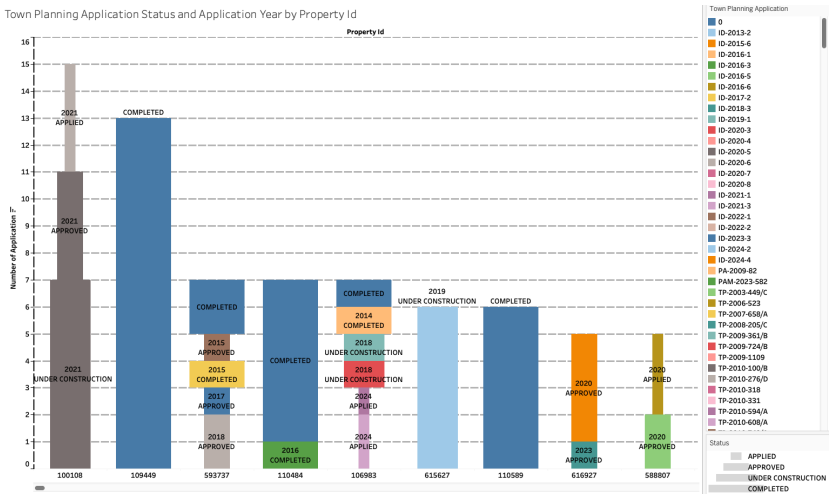

Figure 7. Town Planning Application Status by Property Id

Figure 7 shows that some Property IDs have multiple application submissions in different years, which may indicate renovations or the addition of new towers. The figure also shows the presence of properties with recent applications marked as 'Approved' or 'Under Construction' suggests ongoing or upcoming development activities, indicating growth in those areas.

Question B

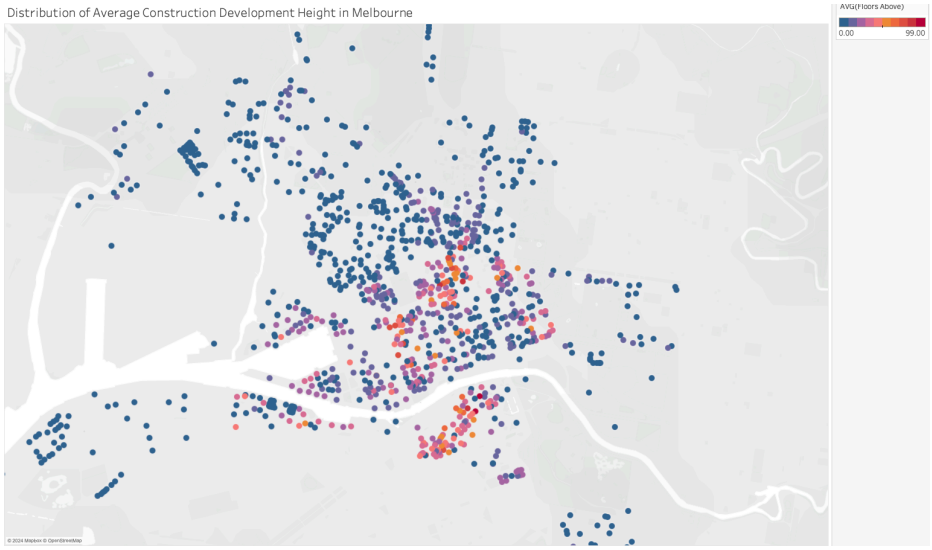

Figure 8. Distribution of Average Construction Development Height in Melbourne

Figure 8 illustrates the average distribution of construction development in Melbourne. Blue dots indicate areas with lower construction heights, while red dots represent areas with higher construction heights. The figure reveals that taller buildings are concentrated in the heart of Melbourne's city centre and the Southbank area.

In contrast, the blue dots, which are scattered farther from the city centre, likely correspond to the development of residential neighbourhoods or small commercial areas. This analysis of construction patterns can be valuable for various stakeholders. For students, it may inform migration planning, while investors could use it to identify potential growth areas for future investments.

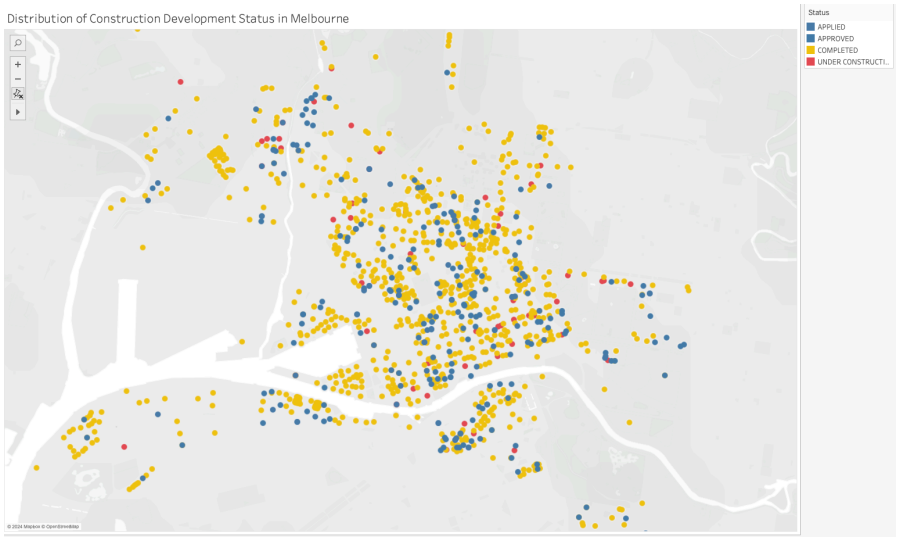

Figure 9. Distribution of Construction Development Status in Melbourne

Figure 9 illustrates the distribution of construction development statuses in Melbourne. Yellow dots indicate completed constructions, blue dots represent buildings that are in the early stages (either just applied or approved but not yet built), and red dots show buildings currently under construction. The figure suggests that the majority of constructions in the dataset have already been completed, highlighting significant development activity in Melbourne in recent years.

Notably, the concentrated clusters of red dots on the map suggest that certain areas are experiencing ongoing development, potentially marking these zones as key areas in the near future. Meanwhile, the scattered distribution of blue dots across the map indicates areas with potential for long-term growth.

Question C

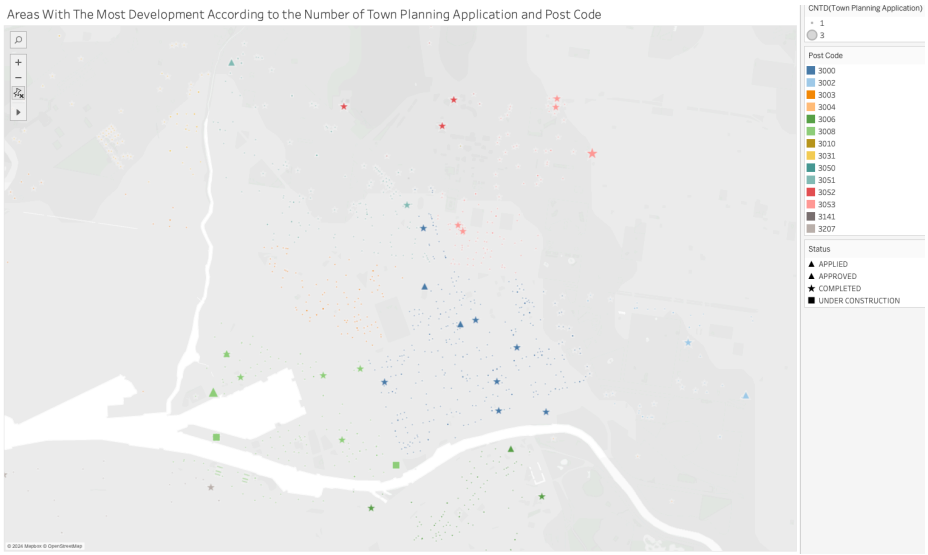

Figure 10. Areas With The Most Development According to the Number of Town Planning Application, Post Code, and Status

In Question A, it is explained that each Property ID can have multiple Town Planning Applications, and each Property ID belongs to a single postcode. If a Property ID has multiple applications, it suggests that the property may have undergone renovations, rebuilding, or the construction of multiple towers.

Figure 10 illustrates the distribution of the number of Town Planning Applications for properties across different postcodes and statuses. The figure highlights that the areas with postcodes 3008 (Docklands) and 3000 (CBD) are the central hubs of development in Melbourne. Additionally, it indicates that postcode 3053 (Carlton) is an attractive area for migration, as most of the construction there has already been completed.

References

City of Melbourne. (2016). Development Activity Monitor – May 2016: Metadata information. City of Melbourne. https://data.melbourne.vic.gov.au/api/datasets/1.0/development-activity-monitor/attachments/dam_info_dam_may_2016_meta_data_info_pdf/ (Accessed August 12, 2024).

City of Melbourne. (n.d.). Development activity monitor [Data set]. Data Melbourne. https://data.melbourne.vic.gov.au/explore/dataset/development-activity-monitor/map/