

## OPEN INNOVATIONS IN DATA SCIENCE (OIDS) COURSEWORK REPORT

Do not write your name on your work unless your lecturer has explicitly told you to do so.

Student ID number	Title of degree studying	Level/Year
2200918	Bachelor of Science (Honours) Data Science and Analytics	6

Short unit name:	M32366 - OIDS			Due date: 14 March 24	Deadline: 14 Mar
Full unit name:	OPEN INNOVATIONS IN DATA SCIENCE				
Unit lecturer name:	Mr. Jackson Maximillian Yap			Group: <i>(if applicable)</i>	
Additional items e.g. CD/disk/USB:	Yes		No <input checked="" type="checkbox"/>	Details:	

All additional items should be clearly labelled with ID number and unit name and securely attached to your work.

Candidates are reminded that the following are defined as Assessment Offences and will be dealt with in accordance with the University's Code of Student Discipline:

- a) Any attempt to complete an assessment by means considered to be unfair.
- b) Plagiarism, which the University defines as the incorporation by a student in work for assessment of material which is not their own, in the sense that all or a substantial part of the work has been copied without any adequate attempt at attribution or has been incorporated as if it were the student's own work when in fact it is wholly or substantially the work of another person or persons.

Please note: Group coursework will be filed under the first Student ID number at the top of the list. Ensure you know all group member's ID numbers.

NB: Coursework not collected will be disposed of six months after the hand-in date.

---

### FOR OFFICIAL USE ONLY

Date received/Office stamp
----------------------------

Administration Office

Provisional mark % / Comments
-------------------------------

Academic Staff Member

### Team Members:

UP2200910, UP2200923, UP2200918, UP2200933, UP2200919, UP2200934

# Table of Contents

<b>1. Evolution of your mini project (planning, development, MVP) (192 words)</b>	<b>3</b>
<b>2. Analysis of the data provided using data visualization and predictive modelling (528 words)</b>	<b>4</b>
<b>3. Lessons learnt (183 words)</b>	<b>7</b>
<b>4. In your opinion, have you met the learning outcomes of this module? Please elaborate. (97 words)</b>	<b>8</b>
<b>5. References</b>	<b>9</b>

## **1. Evolution of your mini project (planning, development, MVP) (192 words)**

In the planning phase, we laid the groundwork by defining our objectives and identifying the dataset from the Housing Development Board of Singapore. Our goal was to analyse housing transaction data to aid clients in understanding property values. This phase involved selecting tools such as Python for data manipulation, Tableau for visualization, and external APIs for enriching our dataset with geospatial information.

The development stage saw us diving into data cleaning, analysis, and integration. We harnessed Python's Pandas library to preprocess the data, Tableau to start crafting visual insights, and explored APIs like OneMap for additional geospatial context. This period was marked by iterative testing and refinement, as we sought to identify meaningful patterns and trends.

Transitioning to the MVP, we synthesized our findings into a functional dashboard that provided predictive insights on housing prices, alongside visual representations of demographic trends and geospatial influences. The MVP served as a tangible proof of concept, showcasing our ability to translate complex data into actionable intelligence for our clients, within the constraints of our defined scope and resources. This process underscored the importance of agile methodology, collaboration, and focused problem-solving in achieving our project goals.

## 2. Analysis of the data provided using data visualization and predictive modelling (539 words)

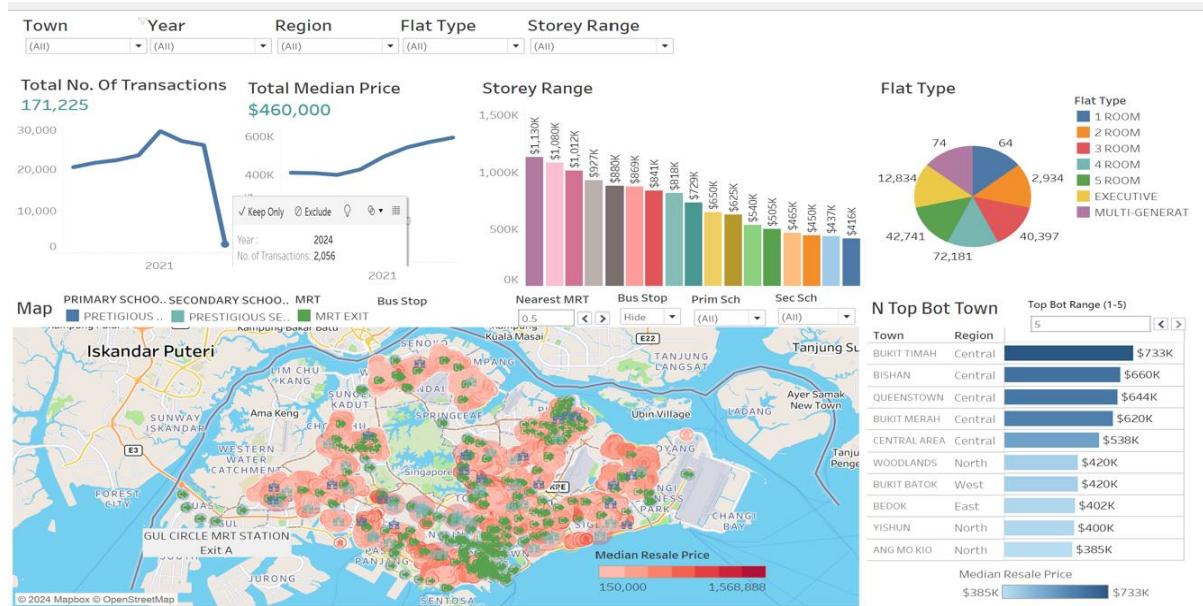


Figure 1: Dashboard Overview

Between 2017 and 2024, there were **171,225 property transactions**. However, the data for **2024** shows a significant drop to just **2,056** transactions. Furthermore, out of the entire dataset, transactions for **Multi-Generation and 1-Room flats** are particularly low, with **only 74 and 64 units** respectively. Given their scarcity, transactions from **2024** will be **omitted** from our analysis to **maintain data consistency**.

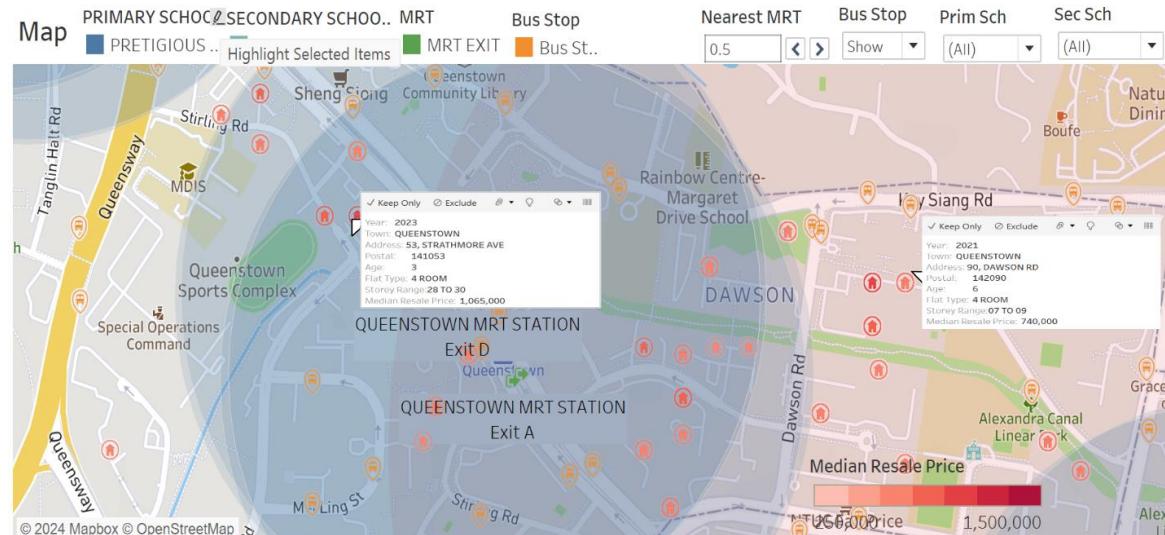


Figure 2: Amenity Proximity Index Map for Property Valuation

This interactive map visualization provides a detailed **overview of properties**, highlighted with their **median resale prices**. Key urban amenities are marked, including the **nearest MRT stations, bus stops, schools, supermarkets, sports complexes, and libraries**. Each property is indicated by a point on the map, which, when selected, displays its specific **details** such as **transaction year, address, flat type, and storey range**.

The gradient bar at the bottom illustrates the spectrum of median resale prices, emphasizing the correlation between property value and proximity to amenities. Properties closer to essential amenities like MRT stations and schools often command higher prices due to the convenience and desirability factors they bring. For instance, properties within walking distance to Queenstown MRT station and surrounded by multiple schools tend to show a higher median resale price.

The visualization suggests that resale prices vary significantly with storey range — higher floors may attract premium prices, potentially due to better views or reduced noise pollution. The integration of amenities, highlighted by convenient filters for transport, education, and shopping options, underscores their significance in property valuation.

Such a comprehensive geospatial analysis enables potential buyers and investors to assess the value-added aspects of each property, making this visualization a powerful tool for making informed real estate decisions.

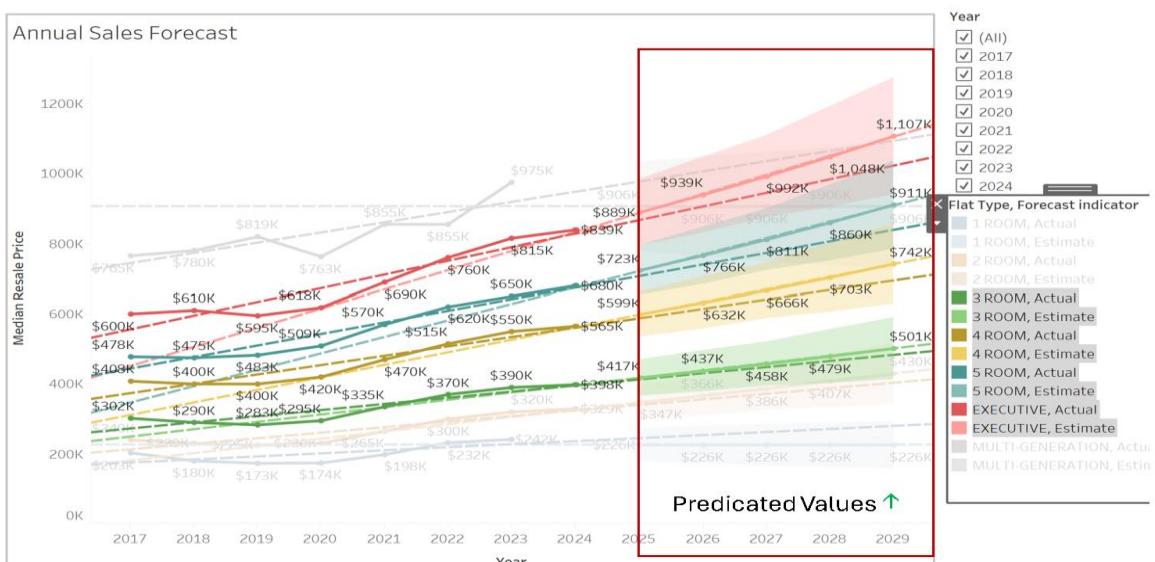


Figure 3 – Upward Sales Trends for 2-Room to 5-Room & Executive Flat: A 5-Year Forecast

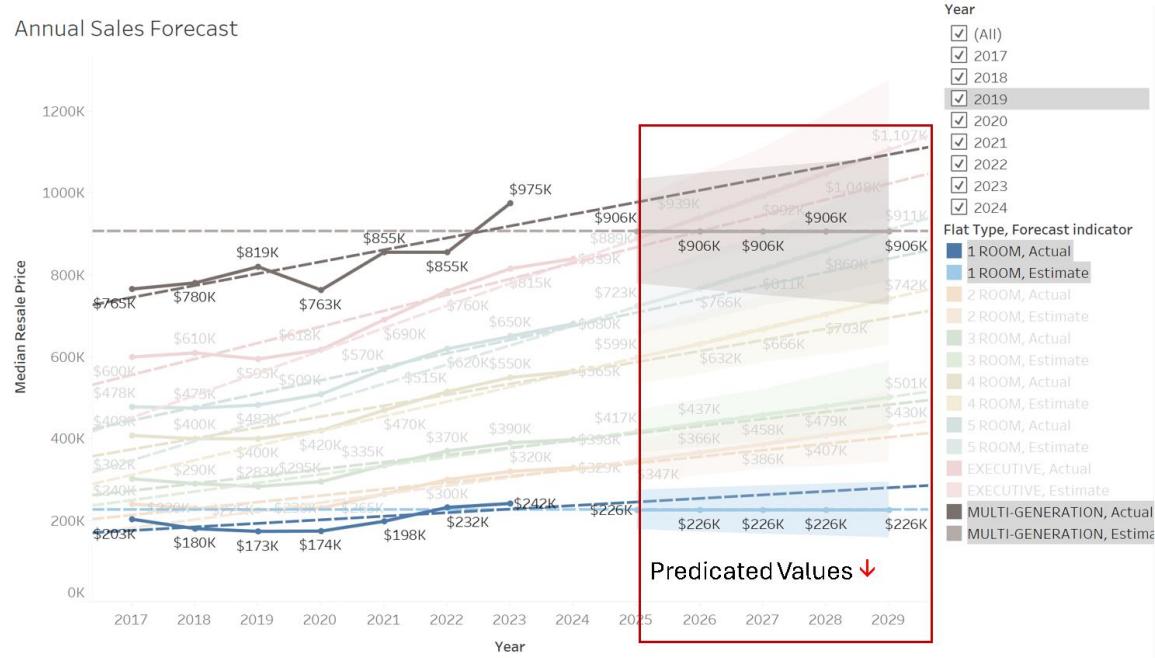


Figure 4 – Downward Sales Trends for 1-Room and Multi-Generation: A 5-Year Forecast

- Forecasting and Strategic Recommendations:**

The upward trajectory for 2 to 5-Room and Executive flats signals a bullish outlook, with a minimum anticipated profit margin of \$100,000 from 2023 to 2029. This forecast advocates for a 'hold and sell' strategy, capitalizing on the property's appreciating value. Conversely, the declining trend for 1-Room and Multi-Generation flats poses a challenge. While these types may benefit from proximity to top-tier amenities, the forecast suggests that without such locational advantages, owners should consider selling before further market contraction.

The comparative analysis of top and bottom towns indicates town disparities in investment returns. Top-performing towns with robust infrastructure developments should be focal points for buyers aiming for long-term growth, while lower-ranked towns may offer immediate rental yield opportunities given their potentially lower entry costs and upcoming urban revitalization projects.

- Data-Driven Decision Making:**

The intersection of transaction volume, median prices, and geographic analysis underpins our advisory. For well-located properties, awaiting market maturity is advised, while less favourable locations warrant a more immediate divestment to maximize returns. Our dashboard's geospatial intelligence provides clients with a lucid visualization of these trends,

ensuring that choices are not only grounded in solid data but also reflective of each property's unique position within the market tapestry.

In conclusion, our predictive analytics offer a nuanced strategy that guides clients through the complexities of the real estate market, ensuring their decisions are both data-informed and strategically sound for the forthcoming years.

### 3. Lessons learnt (210 words)

- **New Technology:**

This project introduced us to working with map data and advanced analysis, making us better at organizing data with Python's Pandas and showing data with Tableau. We also learned how to use Tableau's own tools to predict housing prices, making predictions easier and well-integrated with our visual presentations. Additionally, getting data through APIs, especially map-related data from OneMap, was a key skill we picked up. This allowed us to make our analysis more detailed with location insights. This experience taught us more about handling different types of data and showed us how important map information can be in solving real problems.

- **Working as a Team:**

The project highlighted the essence of teamwork in analytical endeavours. By pooling our diverse skills to integrate external data, we gained varied perspectives, enriching our analyses. This collaborative approach underscored the importance of effective communication and adaptability, revealing the strength of collective problem-solving in revealing comprehensive insights.

- **Personal Development:**

The project was a key milestone in my personal growth, deepening my understanding of the factors affecting housing prices. It honed my critical thinking and analytical skills, encouraging a multidimensional approach to problem-solving. The experience broadened my appreciation for the role of external variables in market dynamics, a valuable insight applicable across various sectors.

**4. In your opinion, have you met the learning outcomes of this module?  
Please elaborate. (97 words)**

Throughout this module, we've engaged deeply with complex datasets, uncovering insights amid incomplete and at times conflicting information. Our critical analysis, especially in handling sparser data for certain flat types, required a discerning approach to ensure robust results. We've applied data science principles inventively, developing a predictive model that integrates traditional housing data with geospatial variables.

This novel contribution not only enhanced our solution's feasibility but also its practical relevance. Moreover, working largely autonomously, we've adapted our skills to a new domain, showcasing independence and a commitment to data-driven problem-solving, thus affirmatively meeting the module's learning outcomes.

## 5. References

1. *HDB Resale Dashboard – Tableau.* (n.d.).  
[https://public.tableau.com/app/profile/kj.ng/viz/HDBResale\\_22Feb/OverviewDash3?  
publish=yes](https://public.tableau.com/app/profile/kj.ng/viz/HDBResale_22Feb/OverviewDash3?publish=yes)
2. *Home / HDB Resale Prices sg.* (n.d.). HDB Resale Prices. <https://hdb-resale-prices.sg/>
3. *Mini Project 1 Powerpoint Slide – Google Drive.* (n.d.).  
[https://drive.google.com/drive/folders/1IhR\\_TRR3LxYsIxqs76KfaYCnYhwvebfG?us  
p=sharing](https://drive.google.com/drive/folders/1IhR_TRR3LxYsIxqs76KfaYCnYhwvebfG?usp=sharing)
4. Marc Reid. (2022, February 8). *Tableau BUFFER functions - a complete TUTORIAL on using buffers for spatial analysis* [Video]. YouTube.  
<https://www.youtube.com/watch?v=pqb6RMGFVl0>
5. Matthew Hefferon. (2018, September 8). *Tableau Tutorial - Top N & Bottom n* [Video]. YouTube. <https://www.youtube.com/watch?v=C72JHMjf-h0>