QUANTITATIVE STRATEGY CASE INTERVIEW

Introduction

The dataset for this case interview would be the HDB resale property transactions, using open data from data.gov.sg.

Please answer **all** questions in a deck of PowerPoint slides, which will be presented during your interview. Please also ensure that any code generated is documented and sent to us, along with the deck of slides.

Assume that you are an analyst at HDB, and that the series of questions below come from senior management. During your presentation, assume that you are presenting to HDB senior management. The presentation will last for 1 hour, including time for Q&A.

To aid your presentation, your Deputy Director has advised you that a presentation to senior management should contain the following:

- A clear narrative with recommendations that leads up to a conclusion;
- Each slide should have a banner statement that contains the insights, and any visualisation or table should be directly relevant to that insight; and
- A clean layout without visual elements that distracts from the narrative

All scenarios, tasks and other information contained in this assessment are purely hypothetical and does not reflect reality.

Preparation

The data on HDB resale property transactions comes in 4 separate datasets (link: https://data.gov.sg/dataset/resale-flat-prices) differentiated by the date of the transaction:

- 1990 – 1999
- 2000 – Feb 2012
- Mar 2012 – Dec 2014
- Jan 2015 onwards

Prior to attempting the questions below, please merge the datasets together. You may merge in additional datasets should there be a need to do so.

Questions are organised by difficulty, indicated by the number of asterisks.

Section 1: Data Visualisation

HDB senior management is thinking of creating dashboards to equip potential buyers of HDB resale flats with the necessary information to make an informed decision. Your task is to utilise suitable tools (e.g. QlikSense or Tableau) to create a series of dashboards.

Question A*: Show an overview of the number of property transactions, median price across the years. Provide both a view at the national level, as well as by HDB towns. Your dashboard should also provide functionality to filter based on Flat Type (e.g. look at only 5 Room flats).

Question B:** Some buyers would want to get the largest flat possible within a given budget. Create a dashboard to allow potential buyers to input their budget, and then suggest towns where such flats exist based on historical transactions.

Question C*:** There are buyers who would like to optimise given the proximity of a flat to important locations in the neighbourhood, such as the nearest MRT station. Create a dashboard that allows buyers to: i) input their budget; and ii) optimise flat selection given distance to important locations around the neighbourhood.

Further Resources:

You might want to use geocoding APIs to obtain the geographic coordinates of a flat's location given its address (e.g. "Blk 175 Yishun Ave 7 760175" -> (1.437854, 103.832516)) so that it can be visualised on a map. An example of a geocoding API is OneMap (link: https://docs.onemap.sg/#introduction), but bear in mind that there is a limit of 250 API calls per minute.

Section 2: Data Modeling

HDB wants to know if a resale flat transaction fits market expectations. Your task is to create statistical models to answer the following questions.

Question A*: Predict a resale flat price's transaction price in 2014. Use the following characteristics: flat type, flat age and town. Propose and implement a minimum of three models, select the best model, and explain the reasons for your choice.

Question B:** A flat was sold in Nov 2017 with the following characteristics:

- Flat type: 4 ROOM
- Town: Yishun
- Flat Model: New Generation
- Storey Range: 10 to 12
- Floor Area (sqm): 91
- Lease Commence Date: 1984
- Resale Price: 550,800

Was this a reasonable price for the transaction? How confident are you in your assessment?

Question C*:** Someone mistakenly deleted the column containing data on Flat Type in the database. While backups exist, these data are critical to HDB's daily operations, and time would be needed to restore these data from the backup. Senior management would like you to create a model to predict flat type given a transaction's other characteristics. Explain the reasons for choosing this model.

Section 3: Policy Analysis

HDB is considering and reviewing a slew of policies, and wants to adopt an evidence-based approach to evaluate these policies.

Question A*: Yishun has received a negative reputation as "Crazy Town", and property prices might have been impacted. Are Yishun flats the cheapest in the country?

Question B*: Some members of public have been saying that flat sizes have gotten smaller over the years. Is there any truth in this statement?

Question C:** The Downtown Line Stage 2 connects the Bukit Panjang heartland to the city. Have prices increased for resale flats in the towns served by this Line? You might want to use a difference-in-differences model for this task.

Question D*:** There have been comments online that people are buying flats in towns further from the city so that the cost savings can be used for a car. Are resale prices in HDB estates in areas further away from the city (i.e. Sengkang and Punggol) impacted by Certificate of Entitlement (COE) prices for cars?

END
