2021 ADB Final Project Specifications
Show Me the House Prices

Background

Property Prices Index, provided by the Numbeo database, is one of the indices that measures the historical housing prices and its changes. Two of the commonly used indices are Price-to-Income Ratio and Price-to-Rent Ratio, which are both indicators of the affordability of the housing prices in one area. According to Numbeo, in 2020, Taiwan ranked 12 in terms of Price-to-Income Ratio, with Hong Kong being ahead among all developed countries. Report shows that it costs about 16 years of disposable income to buy a house in Taipei City.

Many factors may contribute to this high housing price phenomena, such as the opaqueness of the real estate market. To increase transparency in real estate market and to reduce information asymmetry and price gouging, government in Taiwan proposed a policy called Actual Price Registration (實價登錄). Under the policy, all real estate transactions must be registered and the details must be revealed. These records are then released to the public by the government on a quarterly basis.

In this project, your goal is to utilize this data to build a system that allows users to query for past transactions. The details of specification are provided below.

Requirements

The only requirement is that you need to adopt spatial databases and graph databases in your system. Explanations and examples are listed below.

1. **Spatial Database**

At the first stage, you need to collect the past real estate transaction records. After gathering the data, you can utilize a spatial database to store the data and process users' queries. You may want to provide a graphical interface and visualize the query results on the map.

Below are some of the possible queries you may face.

- Find a list of houses with two rooms, two halls, and one berth in Daan District, Taipei.
- Find a list of houses under the 15 million NTD budget.
- Rent a house within a 20-minutes driving distance of a specific working place.

2. Graph Database

With the use of graph databases such as **Neo4j**, you can reveal the underlying relationship among heterogeneous data. Common applications such as relationship visualization or recommendation can be realized. Some of the examples are illustrated below.

- Relationship Visualization: As shown in Figure 1, assume that each house is built by a single construction company, and assume that each company focuses on the construction cases in specific regions. Thus, we can visualize the relationship among houses, construction companies, and regions.

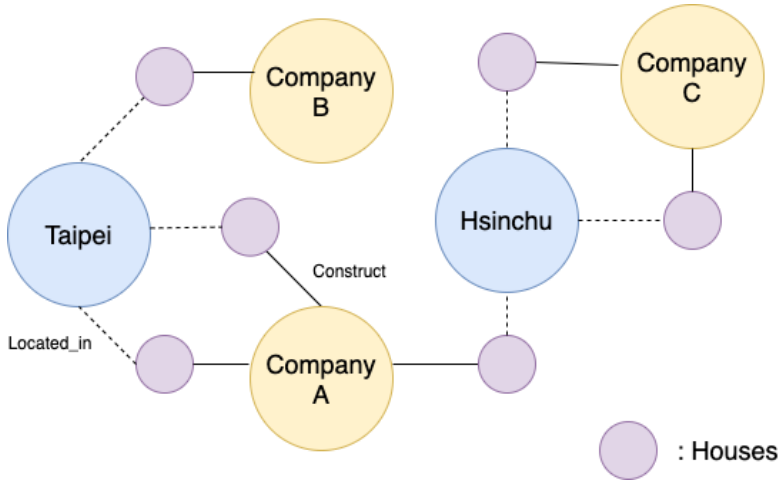

Figure 1: The relation graph between the houses, the construction companies, and the regions. Houses, companies, and regions are represented as nodes, and the “located in” and “construct” relationships are represented as edges.

- House Recommendation: A construction company may build houses with its own style. Buyers who already own their houses may want to buy other houses with similar style. Assume that buyers who own the houses from the same construction company share similar style preferences. We can recommend the buyers to buy other houses based on these common preferences. For example, as shown in Figure 2, we can recommend houses constructed by Company C to Buyer A.

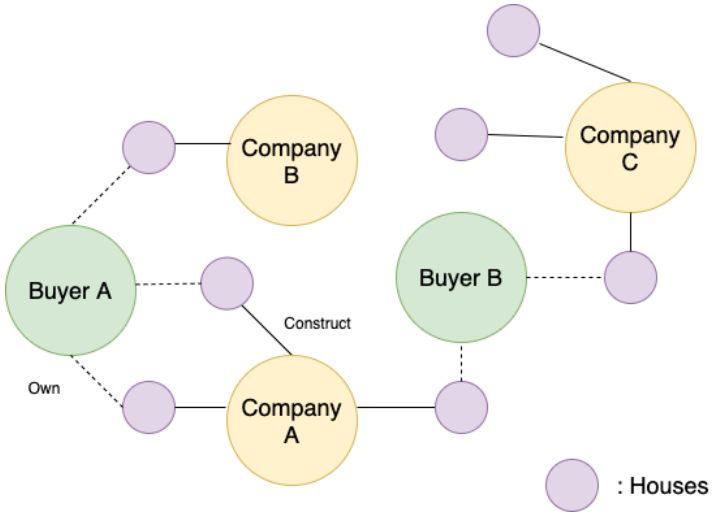

Figure 2: The relationship graph between the houses, the construction companies, and the buyers. Houses, companies, and buyers are represented as nodes, and the “own” and “construct” relationships are represented as edges.

As a reminder, we only provide basic instructions in this spec. You can apply other databases or add all sorts of fancy features to your system.

Grouping

In this project, each group can have 4 to 5 people with at least one domestic student. After you find your teammates, register your group in this google spreadsheet.

Grading Policy

1. **Poster and System Demo (50%)**

You need to prepare a poster and a live demo to showcase your system. The professor and TAs will ask questions about the implementation details or ask you to perform specific queries.

2. **Report (50%)**

Explain the concepts and techniques you use, and which parts are related to the Advanced Database course. The content of the report can include but not limited to system architecture, data model and the database design, and demonstrations. Additionally, you need to provide **detailed contributions of each member** in your group.

References and Datasets

We list some references and datasets below. You can search for other useful datasets or even crawl the data by yourself.

內政部不動產成交案件實際資訊資料供應系統

內政部:::不動產交易實價查詢服務網

建案建商查詢 (Searching the construction cases and construction companies)