# Investigating the proportion of different housing types in the Toronto rental market to infer which housing type will have more advantages or in other words, creats more benefit in the future*

Yichen Shi

23 January 2024

**Abstract**

This study examines the intricacies of the rental market in Toronto, specifically studying the distribution of different types of properties. This study employs an extensive dataset of rental listings to discern notable patterns and trends in the occurrence of various property kinds, including Condominiums, Apartments, and other classifications. The dataset, obtained from a trustworthy source, has extensive information for each listing, encompassing the unique identifier (X_id), operator registration number, address, unit, postal code, property type, ward number, and ward name.In this article, I will analyze the housing types that are likely to garner greater interest in the future rental market in Toronto. The aim is to ascertain which type is more conducive for investment.

## Contents



*Code and data are available at https://https://github.com/EthanShi1314/Rental-housing-Toronto.git

Table 1: ID of Rental Housing and the Property Type

| X_id | property_type |
|---|---|
| 679263 | Apartment |
| 679264 | Condominium |
| 679265 | Townhouse/ Row House |
| 679266 | Condominium |
| 679267 | Condominium |
| 679268 | Condominium |
| 679269 | Duplex/Triplex/Fourplex |
| 679270 | Duplex/Triplex/Fourplex |
| 679271 | Apartment |
| 679272 | Apartment |

# 1 Introduction

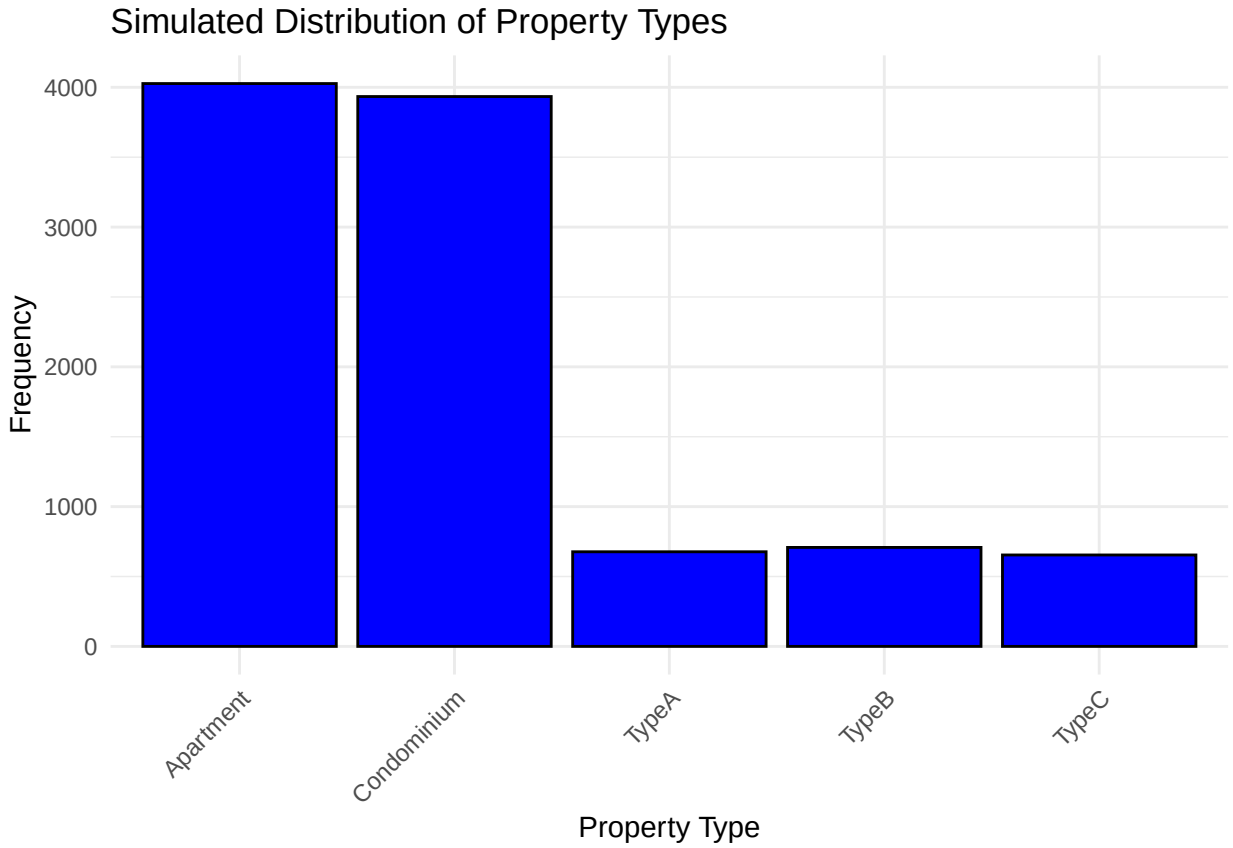
This comprehensive study aims to investigate the distribution of property types within the dynamic Toronto rental market, which is a dynamic and ever-changing landscape. Investors, urban planners, and prospective renters are just a few of the many groups who have a vested interest in keeping up with the latest rental property trends and subtleties in an otherwise thriving real estate market. An broad array of rental listings, each item richly detailed with information spanning unique IDs, property attributes, and geographical markers, provide the substantial data set that this study utilises. A comprehensive study and simulation of the dataset, which has been cleaned and standardized to the highest standards to guarantee accuracy, forms the backbone of the research. In order to provide a complete image of the rental market's composition, the main goal is to extract the substance of the dataset in order to show the proportional representation of different property kinds. This will be done by focusing on major categories like apartments and condominiums. A strategic vantage point that will enhance comprehension and drive educated decision-making in Toronto's booming rental market is promised by this study, which intends to uncover patterns, determine trends, and estimate future trajectories by diving deep into this rich information.

# 2 Data

## 2.1 The housing rental market in Toronto, here's a detailed descrption of the dataset

The dataset chosen for this study provides a comprehensive look at the rental property market in Toronto, with a wide range of variables that give a detailed view of the sector. The dataset consists of 7,922 observations and is organised into eight variables, each providing valuable information about the listed properties. X_id: This variable acts as a unique identifier for each listing, ensuring that every property can be easily recognized and referenced. It plays a crucial role in maintaining data integrity and enabling individual analysis. Property_Type: One of the key factors to consider is the property type, which classifies each listing into categories like Condominium, Apartment, Duplex/Triplex/Fourplex, Single/Semi-detached House, Townhouse/ Row House. This variable plays a crucial role in analysing the market's composition and gaining insights into the popularity and prevalence of various rental accommodations. Above are two important variables I used for the paper.

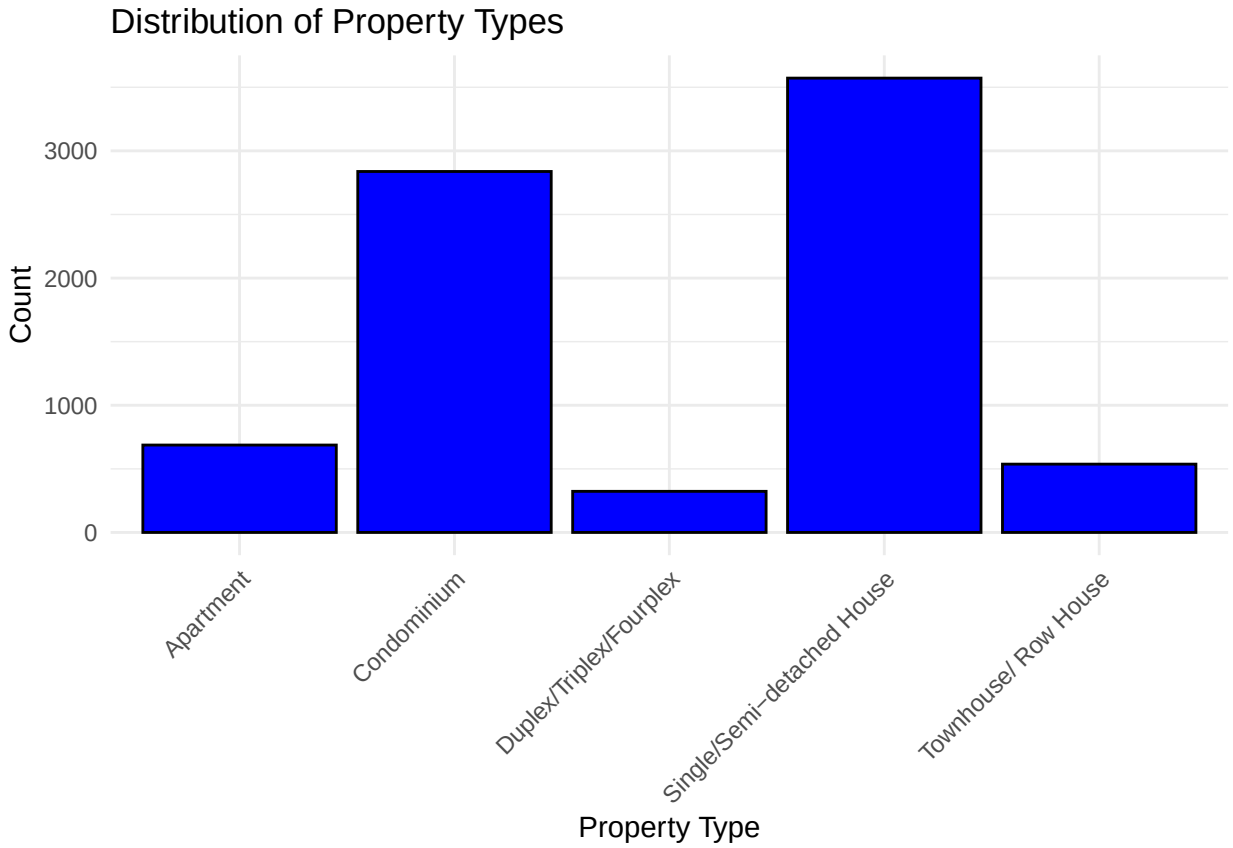
In this step, I selected two variables in the file: ID and room type. I made them into a table to reflect the room types corresponding to each listing. After the table was generated, in order to check whether the code was correct, I intercepted the first six rows and the last six rows. As expected, the ID and room type are correctly matched.




Through the data on the website, I got the proportion of rental housing in 2023, and I found that the proportion of condos and apartments reached 80.87%. The rest are other properties. So I chose to simulate the data and used the set.seed code to ensure this was a fixed simulation. I tried simulating 10000 times and got this bar chart. As shown in the figure, condos and apartments account for the majority, and the remaining three room types are typeA, TypeB, and TypeC, which account for about 20%.

# 3 Acquire

After completing the data simulation, I need to analyze it through the actual data on the official website. First I select the IDs of the listings and their room types. After selecting them, I will make a distribution map of each room type in all ranges. Through these distribution maps, we can determine which room types are popular.




We can clearly see that this is somewhat different from what I've gathered before. First of all, we can see that in the simulation, I chose the data in December 2023. Among them, apartments and condos account for about 80%. However, in the actual data, single/semi detached houses seem to account for the majority. Therefore, I believe that the factors that affect the market are not just room types, but sometimes weather, environment, and geography have a lot to do with it. For example, when international students return to Toronto to attend classes, the demand for condos and apartments will definitely be higher than usual.

# 4 Discussion

The dataset contains a wide range of 7,922 properties in Toronto's rental market, which are carefully categorised into different types. The data reveals a diverse landscape with Single/Semi-detached Houses being the most common type of listing, making up around 44.93% of the total. This is followed by Condominiums at 35.61%, indicating their strong presence in the market. Apartments, although not as common, have a significant presence at 8.63%. Townhouses/Row Houses and Duplex/Triplex/Fourplex types make up 6.75% and 4.08% respectively, providing a range of housing options.

In contrast to the simulated data, which suggested that Condominiums and Apartments would make up a significant 80.87% of the market, the actual data tells a different story. It's clear that Single/Semi-detached Houses are much more prevalent than expected, despite Condominiums having a significant market share. This difference in data emphasises the subtle and potentially evolving preferences in the housing market, which may indicate a growing demand for larger and more autonomous living spaces. These changes were not accurately accounted for in the original simulation assumptions.

The discrepancy between the actual data and the simulation results regarding the number of Apartments could be attributed to various factors such as market dynamics or demographic preferences that may be more inclined towards other types of properties. It's interesting to observe the significant number of Townhouses/Row Houses and Duplex/Triplex/Fourplex types, indicating a variety of housing preferences. This could be

influenced by factors such as family size, budget limitations, or the aspiration for communal living spaces.

This analysis highlights the significance of basing simulation models on reliable and up-to-date data. The noticeable disparities between the projected expectations and the factual data highlight the ever-changing nature of the real estate market, which is shaped by various factors such as economic trends, urban development policies, and evolving lifestyle preferences. For stakeholders, whether they are policymakers, investors, or potential renters, this divergence serves as a reminder of the importance of constantly updating models and strategies based on real-time data. It is crucial to make informed, relevant, and responsive decisions that align with the current market conditions.

# 5 References