

1. The Problem area: What is your area of interest? Within it, what challenges or opportunities could your project address?

My area of interest is the impact of gentrification on housing affordability for low-income individuals and renters in Vancouver. In recent years, average rent in the city has risen much faster than average income, intensifying the pressure on those already vulnerable to displacement. As the housing crisis continues to disproportionately affect low-income families and individuals, numerous housing development plans have been proposed and enacted throughout the Lower Mainland. However, a critical question remains: do we truly understand what the long-term impacts of these housing projects will be, particularly on affordability? The Broadway Plan, approved in 2022, is a major redevelopment strategy that aims to reshape Vancouver's Broadway corridor over the next 30 years through various means such as increased density and the construction of high-rise rental buildings. While it is framed as a response to the city's housing needs, the plan has sparked concern among housing advocates and community members who worry that its emphasis on high-rise development will accelerate gentrification. Critics argue that these towers are often shaped more by developer interests and profitability models than by a holistic understanding of community needs. In many cases, they fear the demolition of older, more affordable rental units will lead to their replacement by expensive apartments out of reach for existing residents. This project aims to use machine learning to analyze patterns in housing affordability before and after the construction of new housing developments, particularly high-rises, in Vancouver. By identifying correlations between development features and affordability outcomes, the model could then be used to predict how future projects—such as those under the Broadway Plan—might influence affordability for renters in the surrounding neighborhoods.

2. The User: Who experiences these problems? How would they benefit from the outcomes of your project?

The people most affected by Vancouver's ongoing affordability crisis are low to moderate income renters, including students, service workers, immigrants, seniors, and long-time residents who are increasingly at risk of displacement due to rising rents and redevelopment pressures. These individuals are the ones who most directly experience the consequences of gentrification, often losing access to stable housing in the communities they've long called home. In addition, housing advocates, neighborhood associations, and city planners are all key stakeholders who must navigate the tension between development and social equity. This project would ideally culminate in a detailed, data-driven report that presents the results of applying a predictive

algorithm to assess the potential impacts of new housing developments proposed under the Broadway Plan. By modeling how similar developments have affected affordability in the past, the report could provide valuable insights into which areas or projects are likely to exacerbate displacement and which may genuinely improve housing access. This information could help housing advocates push for more inclusive policies, support concerned residents in making their voices heard, and offer city planners a rigorous analytical tool to evaluate the potential consequences of their development strategies.

3. The Big Idea: How can machine learning bring solutions to these areas? Research how other people have approached the problem previously. Refer to the "Intro to Capstone" slides on synapse for an overview of different machine learning approaches.

While there is a great deal of public discussion and criticism surrounding housing development plans, much of it remains speculative or based on anecdotal evidence. This project seeks to move beyond that by using machine learning to identify and analyze patterns in historical housing data, offering a more rigorous, data-based assessment of how different types of development projects—particularly high-rise construction—may affect affordability in surrounding neighborhoods. With the growing availability of detailed open data on development projects, rental rates, construction permits, and neighborhood demographics, there is a real opportunity to use machine learning to generate actionable insights. Similar approaches have already been explored in a number of professional projects and research papers. For example, the paper “House Price Prediction Model Using Machine Learning” (JETIR, 2024) employed various machine learning algorithms (linear regression, decision trees, random forests, etc.) to model home prices based on housing characteristics and neighborhood factors. The Urban Institute’s 2023 report “Land Use Reforms and Housing Costs” used machine learning to track the impact of zoning reforms on affordability and housing supply. Another study, “An Optimal House Price Prediction Algorithm: XGBoost” (arXiv:2402.04082, 2024), demonstrated that neighborhood-level data plays a crucial role in accurately predicting housing market trends. This project builds on these methodologies, applying them to Vancouver’s context and focusing specifically on affordability for renters, rather than just changes in property values.

4. The Impact: What societal or business value do you anticipate your project to add? If possible, try to quantify the scale of the problem (in dollars, in CO2, in time spent, ...).

This project offers significant societal value by providing a rigorous, data-based analysis of housing affordability trends in local Vancouver communities. Gentrification, rising rents, and

diminishing housing availability are urgent and ongoing crises affecting tens of thousands of low to medium income residents in the city. Over 50% of households in Vancouver are renters, and a growing proportion of them spend more than 30%—and often over 50%—of their income on housing, placing them in a state of housing need. As redevelopment accelerates, particularly under large-scale plans like the Broadway Plan, this project aims to bring analytical clarity to a moment of high public concern. We are in a critical period where residents, community organizations, and housing advocates are actively engaging in public consultation processes, voicing concerns, and pushing for more equitable development. The stakes of these decisions are not abstract: they involve people's homes, livelihoods, and ability to remain in neighborhoods where they have roots, jobs, and social networks. By contributing a data science perspective to the dialogue of dissent and resistance, this project hopes to strengthen the social justice movement for fair and inclusive housing. It also has practical value for urban planners and policymakers seeking to understand the likely outcomes of their development strategies—helping them avoid unintended harm and promote more equitable urban growth.

5. The Data: Identify several possible datasets in this subject area and describe them at a high level. Include references. If you struggle to find more than one or two datasets, this might mean a Data Science approach to the problem will be challenging. Check in with your Educator.

Several high-quality datasets are available to support this project's analysis of housing affordability and development trends in Vancouver. The [Canada Mortgage and Housing Corporation](#) (CMHC) provides detailed data on rental market statistics, including average rents, vacancy rates, unit types, and new housing construction across different regions of Vancouver. The [Metro Vancouver Housing Data Book](#) offers comprehensive charts and tables on housing supply, demand, and affordability indicators, making it useful for analyzing regional trends over time. [BC Assessment](#) provides property-level data on assessed values, property types, and construction dates, which can help track how new developments impact local housing markets. Additionally, the [City of Vancouver Open Data Portal](#) includes datasets on zoning, development permits, building footprints, and neighborhood profiles—key inputs for modeling the relationship between new construction and affordability. Another option is to scrape rental housing websites like Zillow, PadMapper, or Craigslist to obtain more granular and timely rent data across different neighborhoods. These sources together provide a strong foundation for conducting a data-driven analysis of gentrification and housing affordability.

6. The Alternative: In a few sentences, summarise a problem in an alternative subject area that also interests you.

If major obstacles arise in analyzing trends tied to new housing developments—such as a lack of detailed, time-specific data connecting projects to affordability outcomes—an alternative

approach would be to use clustering algorithms to identify neighborhoods at risk of gentrification based on changes in rent, income, and development patterns. This could still support housing justice efforts by highlighting areas vulnerable to displacement. If I need to shift subject areas entirely, another idea is to use machine learning to analyze adoption trends for rescue cats. I volunteer at a major cat rescue organization with access to detailed data, which could be anonymized and used to identify which cats are at risk of not being adopted. The project could then suggest targeted advertising strategies or adopter profiles to help increase successful placements.