

Applied Data Science Capstone

Brandon Wiebe

April 2021

1 Introduction

The city of Vancouver, located on the west coast of Canada, is consistently ranked as one of the nicest cities in North America. Surrounded by water, with scenic views of mountains to the north and a mild and pleasant climate, it's not very surprising that it makes for a desirable place to live. Vancouver has an incredibly diverse population, and has hosted many international conferences and events, including the 2010 Winter Olympics. Unfortunately, it is also ranked as one of the most expensive cities to live in on the globe, and is the most expensive city to live in in Canada. Given the desirable nature of the Vancouver area, as well as the potentially prohibitive cost, many people look to the suburbs and cities surrounding Vancouver as a viable alternative place to live.

Imagine that there is specific neighborhood in Vancouver that an individual would like to move to, however, the cost of living is simply too high. Which of the surrounding suburbs and cities should that person move to if they want to live in an area that is most similar to their desired neighborhood? Or, imagine that someone who lives in Vancouver is looking to move outside of the city to save on living expenses, but they would ideally like to move to a similar neighborhood to the one they've been living in. Where should their realtor recommend they look?

In this report we will use machine learning techniques to first characterize the different neighborhoods in the city of Vancouver, as well as the more affordable communities outside of the city limits. We will then compare the neighborhoods in Vancouver to those other communities, in order to determine which of them are similar to those neighborhoods. This information can then be used by those individuals who are looking for the most authentic Vancouver experience without the price tag.