

parkerdeshazo / Phase-2-Project

Public

☆ 1 star

🍴 1 fork

★ Starred

👁 Watch ▾

<> Code

🕒 Issues

🔗 Pull requests

🎬 Actions

📁 Projects

📖 Wiki

🛡 Security

📈 Insights

🔗 main ▾

...

SBeyzaU

Update README.md ...

6 minutes ago

🕒 39

View code

☰ README.md

✎

Phase-2-Project

Authors: Vincent Amadio, Parker DeShazo, Beyza Ulasti

Overview

We identified a business problem, cleaned the data, set parameters, dropped outliers, and used OLS models to infer house prices in King County, Washington.

Business Problem

For this project, our stakeholder is a young and growing family looking to reloacte to King County, Washington. They are looking for a single family home with enough space for the family to grow. This includes looking at the more than two bedrooms, a structually safe home, and to live within a top rated school district for their kids education.

https://github.com/parkerdeshazo/Phase-2-Project

1/4

Data Understanding

The data set used in this analysis is open-source data available directly from the county website in King County, Washington (<https://kingcounty.gov/services/data.aspx>). This data set covers various aspects of real estate transactions including date of sale, square footage of house and lot, bedroom, bathrooms, and condition, etc. The data set consists of over 30,000 data points. Additionally, we have gathered school district data from Niche.com (<https://www.niche.com/places-to-live/search/zip-codes-with-the-best-public-schools/c/king-county-wa/>). For budget purposes, we have also used data from level.fyi (<https://www.levels.fyi/t/data-scientist/locations/greater-seattle-area>).

Data Limitation: More data on renovations would have improved our insight. (Cost of renovations, any added sqft, what was renovated)

Bottom Line

1. School
 2. Safety
 3. Space
- We have identified a strong school district, safety of the house, and extra space as the most important features to our stakeholders. From there, we added these features to our model for price consideration. We have found that grade, condition, bedrooms, bathroom, and area (based on school district zip code) explain roughly 10% of the variation in price. 10% of an average priced house in King County is about \$100,000.

Models

First, we gathered school district data for King County. We grouped the zip codes we have identified as belonging to the strongest school districts in the county, into four areas. The areas we used for the models are: Redmond, Sammamish, North King County, and Newcastle. From there we made a different model for each area, keeping all the other features constant, so we could compare the areas. Here's what went into our models:

1. Area
2. Number of bathrooms
3. Number of bedrooms
4. Grade of the house
5. Condition of the house

Findings & Recommendations

- When compared to the average price of a house in King County, houses in Redmond and Sammamish cost more. Among the top four areas we have identified, Newcastle is the most economical.
- Adding an extra bathroom to the house will cost less than the other features we have modeled while increasing the space.
- Our model found minimal variation in grade, condition, and the number of bedrooms between Redmond, Sammamish, North King County, and Newcastle.

Next Steps

1. House Appreciation
2. Renovation costs
3. Rental Properties

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Contributors 2



SBeyzaU



parkerdeshazo

Languages

● Jupyter Notebook 100.0%