

King County Housing Price Project

Regression Modeling and Analysis: Data Science

Orlando Vilar, Sep 2022

Introduction

Introduction

Overview & Business Understanding

- House as an asset is deeply ingrained in our culture;
- Which features impact the most housing prices?
- How can we assess housing prices using property's characteristics and location within King County - WA;
- **Relevance:** overall implication to investors decision-making, realtors, contractors, and the general public.

Data & Analysis

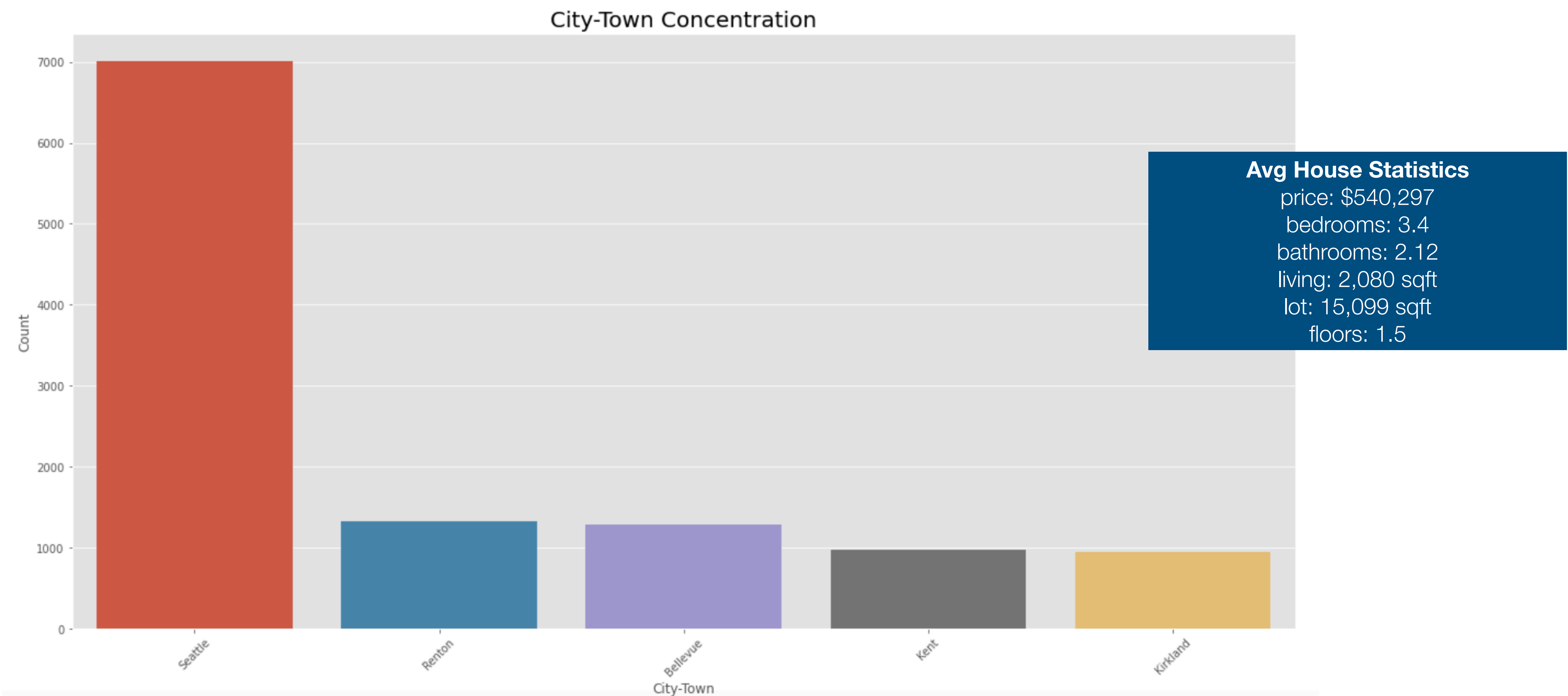
Data Understanding

Sources of Information

- Databases:
 - King County House Sales;
- Variables:
 - Date, view, living/lot area and surroundings, year built, year renovated, zip code, latitude and longitude ;
- Additional feature:
 - Matching of latitude and longitude to provide city/town location.

City-Town Property Concentration

Most representative localities



Closing Remarks

Strategies

Three key insights

- **Most expensive:** Mercer Island, newer/renovated, have an excellent view and condition. Mansions with fewer rooms, 2.5 floors and 3 bathrooms;
- **Least expensive:** Tacoma, is older with an average grade, and in fair condition;
- Living area (in sqft) is related with a price increase;
- The model explains 82% of price variation;
 - Price prediction is more assertive within the Seattle area.

Limitations

Further analysis

- Adding demographic data about each specific area;
- Adding polynomial features and testing non-linear behaviors;
- Data seems not random normally drawn.

Thank you!

<https://github.com/ovilar>