Predicting House Prices in King County, WA

Jesse Markowitz
Dillon Medd
Ronak Parekh

Summary

Three key factors that influence price:

- **Square feet** of living space
- **Distance** from Seattle (city center)
- **Grade** given to housing unit *(based on King County grading system)*

Outline

- Business Problem
- Data & Methods
- Results
- Conclusions
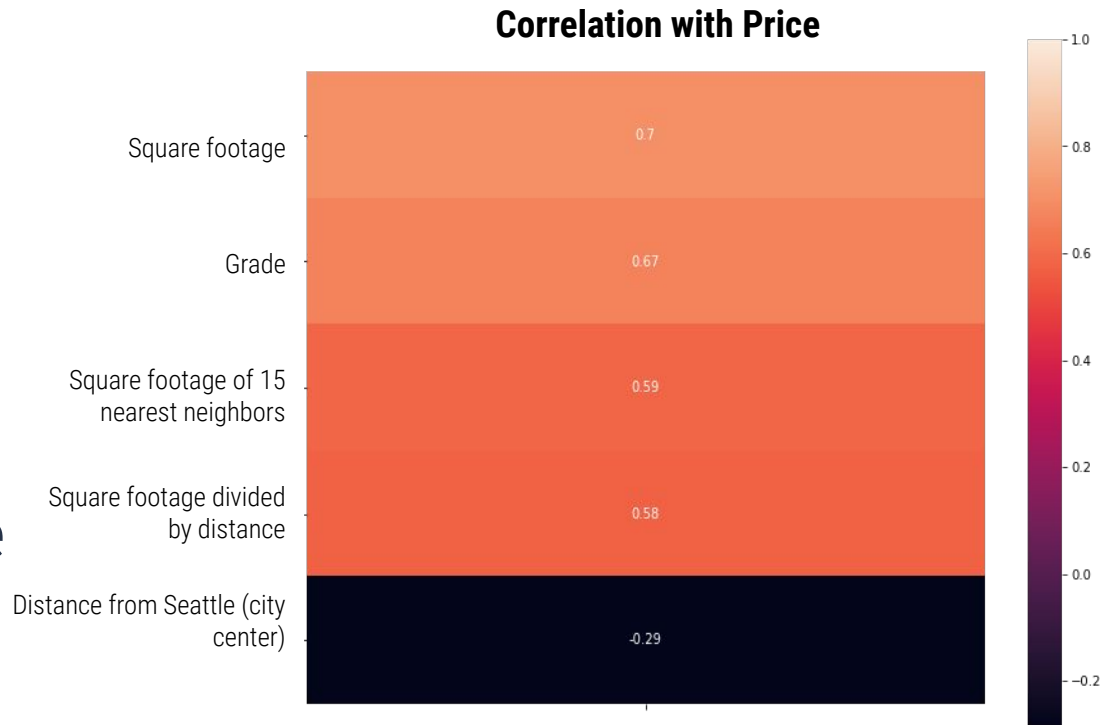- Next Steps

Business Problem

- **Purpose of Analysis:** Predict prices for houses in King County, WA for home buyers

- Priorities:
  - ▷ Location
  - ▷ Size
  - ▷ Quality

Data & Methods

- **Data source:** King County housing sale data
  - ▷ Prices and features for houses sold from May 2014–May 2015
- **Methods:** Multiple linear regression
  - ▷ Iterative modeling to identify key features
  - ▷ Feature engineering to capture patterns (distance)
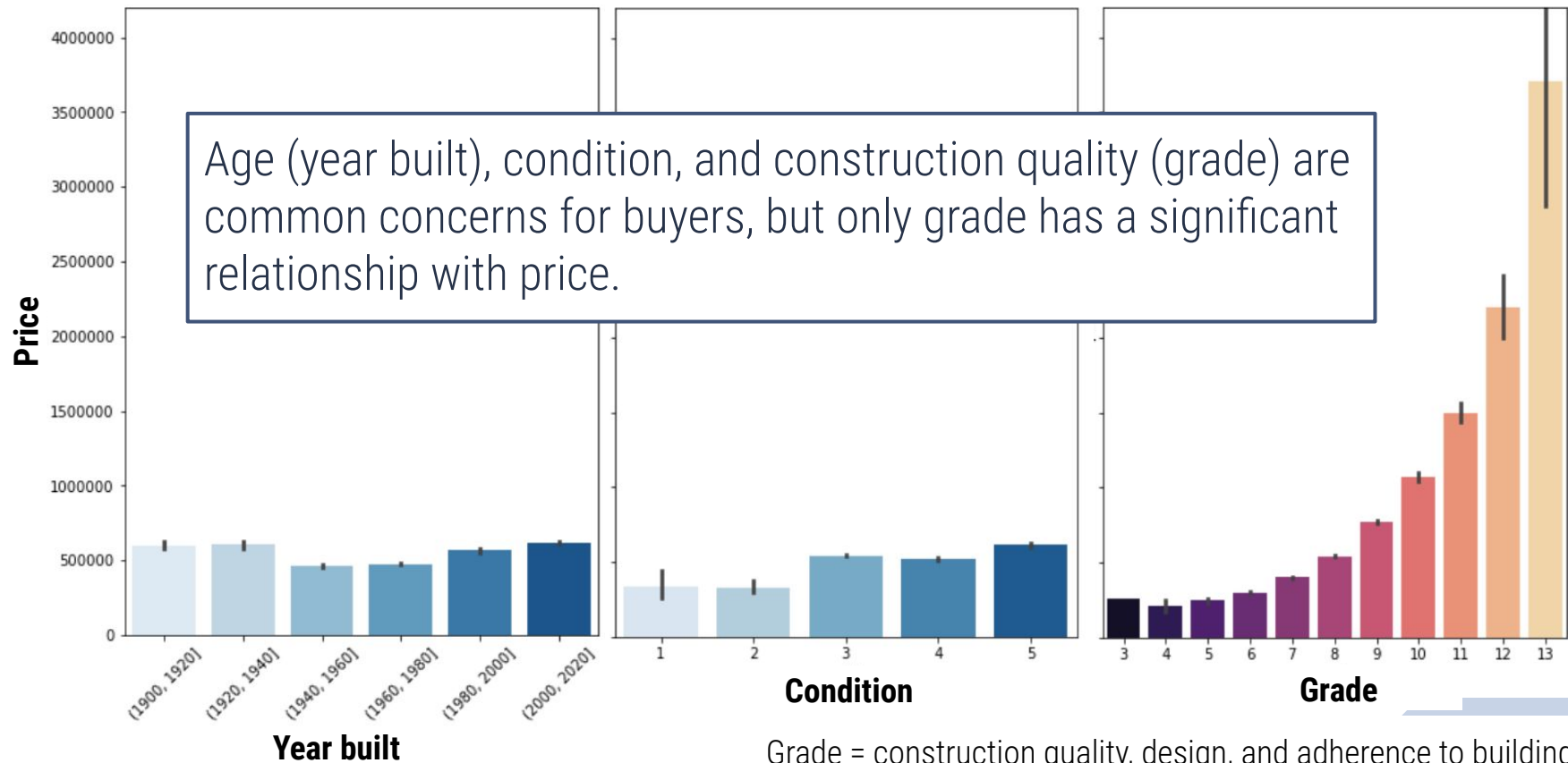
Question 1: Which factors correlate with price?

- Square footage

- Grade (1–13)

- Square footage of 15 closest homes

- Square footage divided by distance
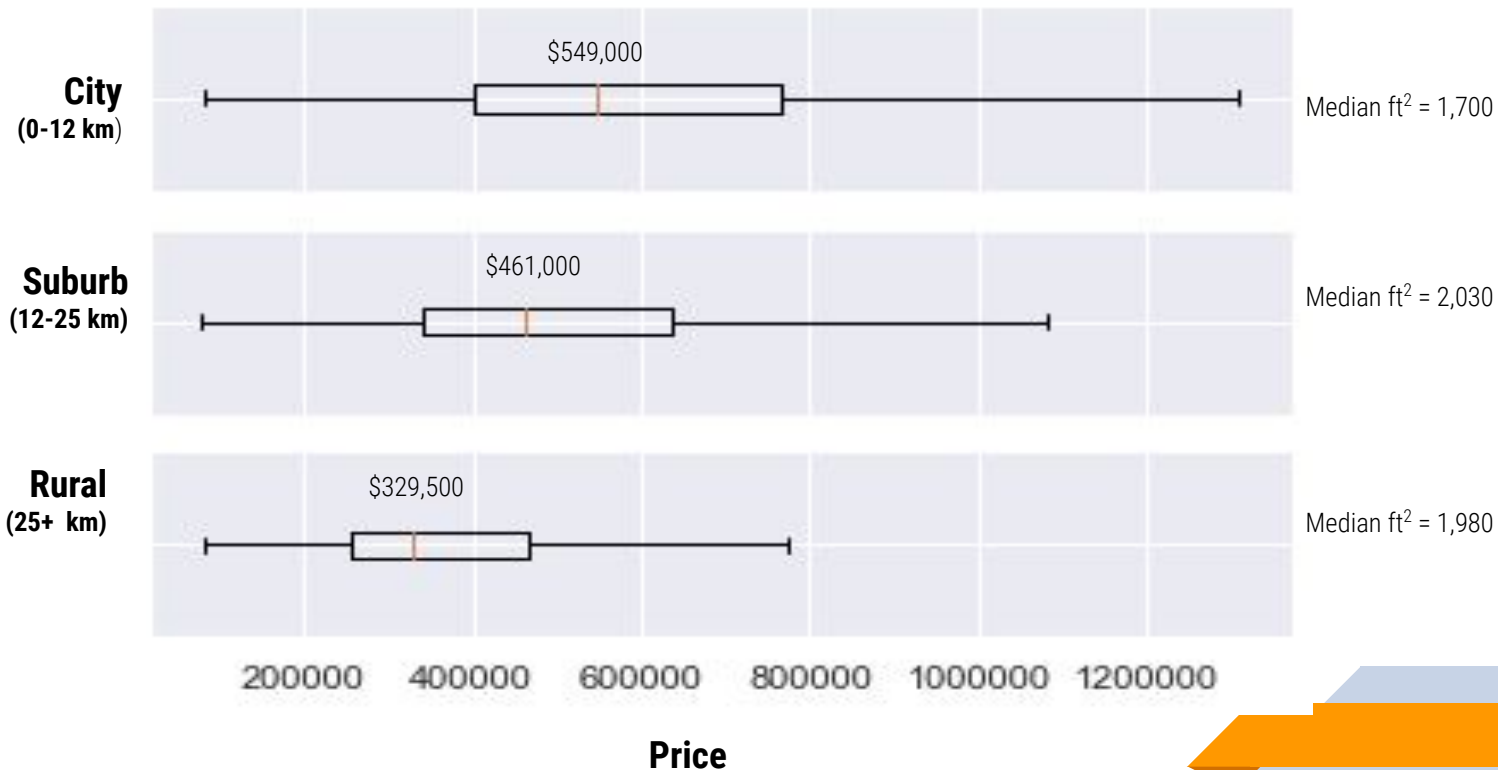
- Distance from Seattle city center

Correlation with Price

Question 1: Which factors correlate with price?

Age (year built), condition, and construction quality (grade) are common concerns for buyers, but only grade has a significant relationship with price.

Price

4000000

3500000

3000000

2500000

2000000

1500000

1000000

500000

0

(1900, 1920] (1920, 1940] (1940, 1960] (1960, 1980] (1980, 2000] (2000, 2020]

Year built

1 2 3 4 5

Condition

3 4 5 6 7 8 9 10 11 12 13

Grade

Grade = construction quality, design, and adherence to building codes

https://info.kingcounty.gov/assessor/esales/Glossary.aspx?type=r#b

Median price decreases with distance from city center, despite increase in median size

City (0-12 km) — $549,000 — Median ft^2 = 1,700

Suburb (12-25 km) — $461,000 — Median ft^2 = 2,030

Rural (25+ km) — $329,500 — Median ft^2 = 1,980

Price axis: 200000, 400000, 600000, 800000, 1000000, 1200000

Price

Question 2: How does location affect price and square footage?

Price per square foot decreases when distance from city center increases

<u>Final linear regression model:</u>

Price is affected most by square footage, grade, square footage of neighboring houses, and distance from Seattle city center

- Accounts for 72% of variance in housing prices *(r^2 = 0.723)*

- Consistent predictions for both training and unseen data

Conclusions

- Price per square foot is largely predicted by distance from Seattle city center
  - ▷ Higher in the city, lower farther out

- Buyers must balance optimal size and location for their budget

- Grade (structural quality) matters more than age or condition

Next Steps

1. Compare recent sale price to original purchase price to identify trends, good areas for investment

2. For given budget, identify for buyers affordable house size (ft^2) by location

3. More detailed mapping to identify homes in urban, suburban, and rural areas.

Thank You!

Jesse Markowitz

Email:

jess.markowitz@gmail.com

GitHub:

@jmarkowi

Ronak Parekh

Email:

ronakp2001@live.com

GitHub:

@Rparekh123

Dillon Medd

Email:

dillonmedd1@gmail.com

GitHub:

@dmedd98