

KING COUNTY DATA SET

DS EDA PROJECT | ANALYSIS FOR CLIENT NICOLE JOHNSON

CLIENT

- NICOLE JOHNSON
- BUYER
- PREFERENCES:

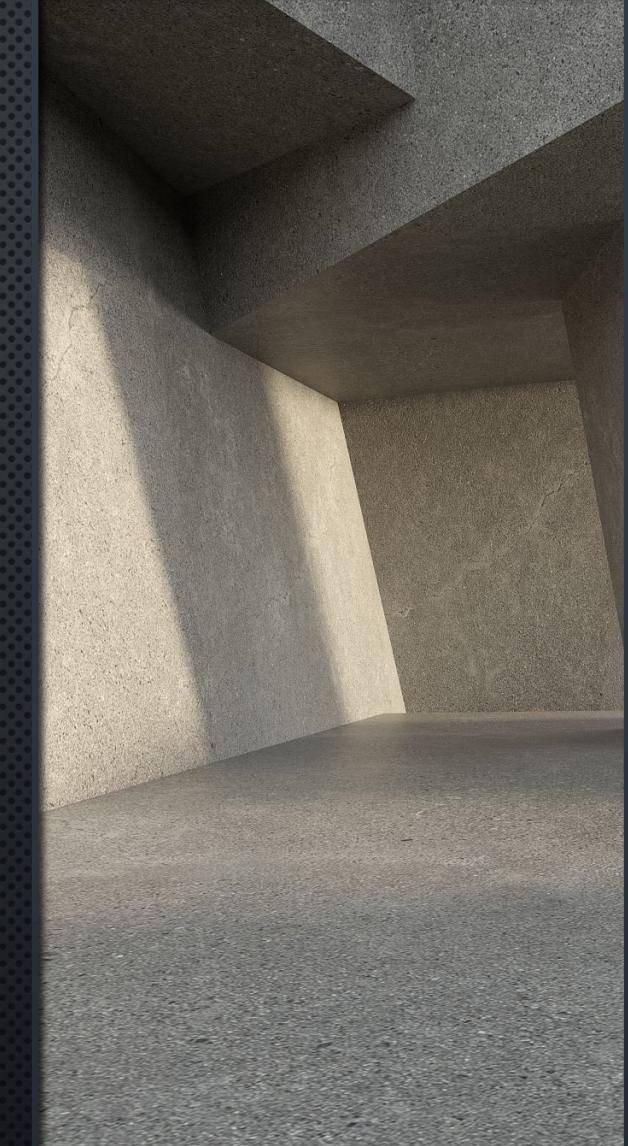
LIVELY, CENTRAL NEIGHBORHOOD,
MIDDLE PRICE RANGE, RIGHT TIMING
(WITHIN A YEAR)





THE ORIGINAL DATASET

- THE **DATASET CONTAINS RECORDS OF HOUSE SALES** IN KING COUNTY, INCLUDING DETAILED PROPERTY FEATURES AND SALES INFORMATION OF 21,597 SALES
- **KEY FEATURES:** EACH RECORD INCLUDES VARIABLES SUCH AS HOUSE ID, SALE DATE, SELLING PRICE, NUMBER OF BEDROOMS AND BATHROOMS, ...
- **GEOGRAPHICAL DATA:** THE DATASET INCLUDES GEOGRAPHICAL COORDINATES (LATITUDE AND LONGITUDE) AND ZIP CODES, ALLOWING FOR LOCATION-BASED ANALYSIS OF PROPERTY VALUES.
- **TEMPORAL ASPECT:** SALES DATES ARE INCLUDED, ENABLING ANALYSIS OF TRENDS OVER TIME WITHIN THE DATASET.
- **EXTENDED FEATURES:** ADDITIONAL FEATURES LIKE SQUARE FOOTAGE OF ABOVE AND BASEMENT AREAS, AS WELL AS LIVING AND LOT AREA FOR NEIGHBORING PROPERTIES (15 NEAREST), PROVIDE COMPREHENSIVE INSIGHTS INTO PROPERTY CHARACTERISTICS AND THEIR SURROUNDING CONTEXT.

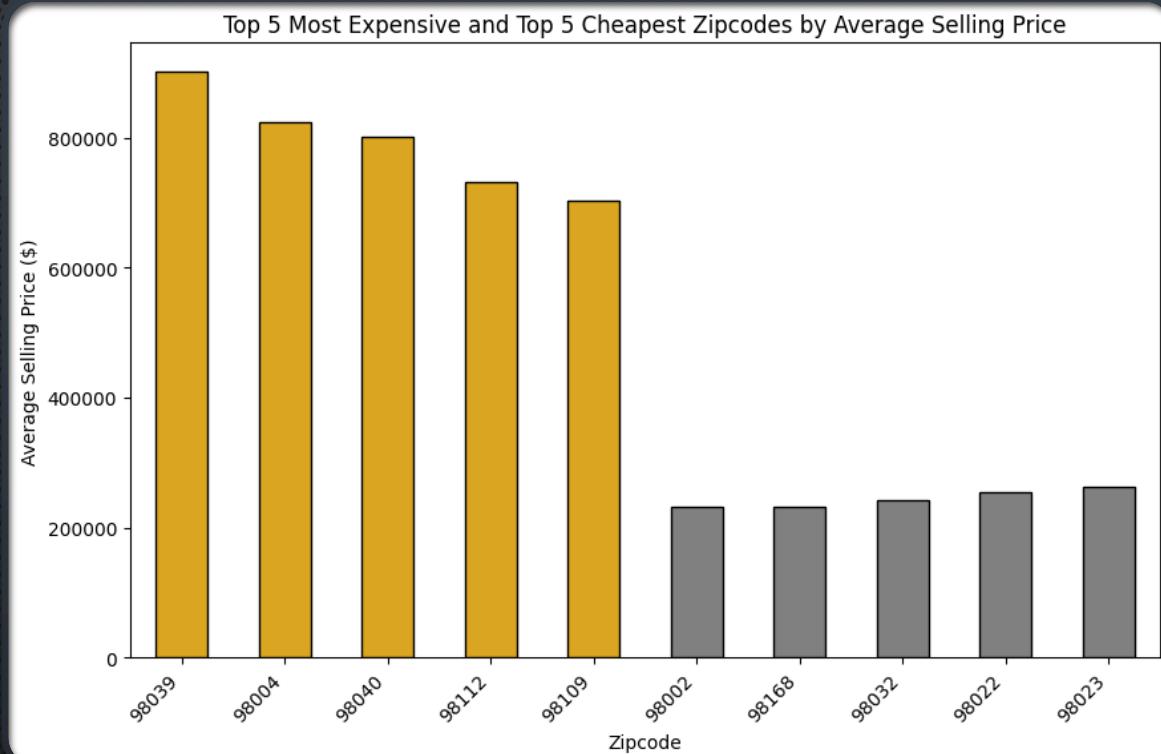


DATA CLEANING

- MISSING VALUES ON RENOVATION, BASEMENT, WATERFRONT, VIEW WAS TREATED AS IF THE CORRESPINDING PROPERTIES DID NOT HAVE THESE FEATURES
- DUPLICATES HAVE BEEN REMOVED
- DATA TYPES HAVE BEEN CORRECTED (E.G. PRICE DATE WAS AN OBJECT)
- OUTLIERS HAVE BEEN REMOVED

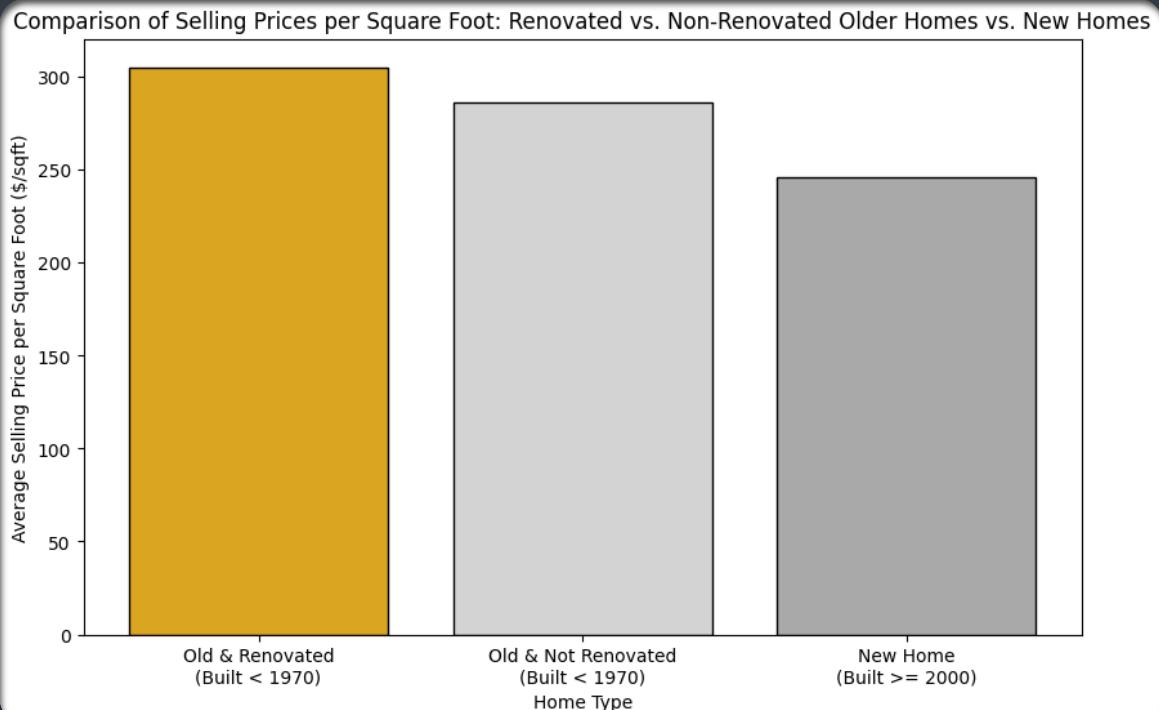
KEY INSIGHT
#1

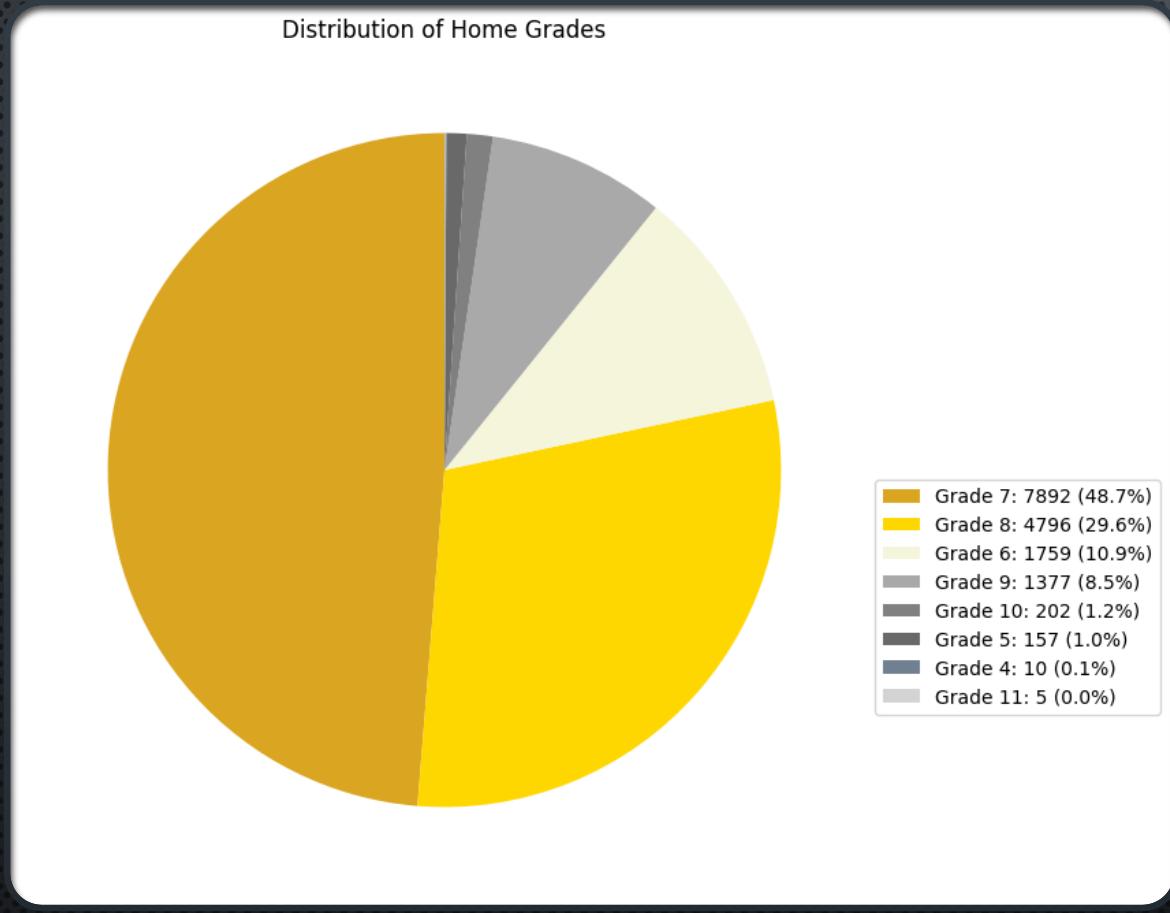
LOCATION
DETERMINES
PROPERTY
VALUES



KEY INSIGHT #2

OLD IS MORE
EXPENSIVE
THAN NEW





KEY INSIGHT
#3

ABOVE-
AVERAGE
GRADES



CENTRAL & LIVELY NEIGHBOORHOOD ASSUMPTIONS

- **HIGH POPULATION DENSITY:** HIGHER POPULATION DENSITY OFTEN CORRELATES WITH CENTRAL AND LIVELY AREAS.
→ AVERAGE LIVING SPACE OF NEARBY HOMES / AVERAGE LOT SIZE OF NEARBY HOMES
- **URBAN ZIP CODES:** CENTRAL LOCATIONS ARE LIKELY TO BE IN MORE URBANIZED AREAS.
→ IDENTIFICATION OF THE TOP 20 ZIP CODES WITH THE MOST LISTINGS IN THE DATASET AS POTENTIAL URBAN AREAS.



MIDDLE PRICE RANGE ASSUMPTION

- THE MIDDLE PRICE RANGE FOR THE HOUSES IN THE DATASET IS BETWEEN **\$300,000 AND \$556,825**. THIS RANGE REPRESENTS THE MIDDLE 50% OF HOUSE PRICES

TOP 3 PROPERTIES THAT MEET THE CRITERIA

1. Property ID: 1294300038

- **Selling Price:** \$450,000
- **ZIP Code:** 98116
- **Density Ratio:** 1.980
- **Square Footage (Living):** 1,810 sq ft
- **Lot Size:** 914 sq ft
- **Bedrooms:** 3
- **Bathrooms:** 2.5

2. Property ID: 7660100238

- **Selling Price:** \$329,950
- **ZIP Code:** 98144
- **Density Ratio:** 1.578
- **Square Footage (Living):** 1,300 sq ft
- **Lot Size:** 812 sq ft
- **Bedrooms:** 3
- **Bathrooms:** 2.5

3. Property ID: 7904700134

- **Selling Price:** \$390,000
- **ZIP Code:** 98116
- **Density Ratio:** 1.497
- **Square Footage (Living):** 1,370 sq ft
- **Lot Size:** 913 sq ft
- **Bedrooms:** 3
- **Bathrooms:** 3.25