King County Housing Sales

"Data Analysis & Prediction Model"

Filiz Camuz
July 2019

Methodology

OSEMN

Business Question

## How to split land into lot and living area price efficiently?

- Builders behavior over the years
- Number of houses that has big lots
- Lot & living area vs. price
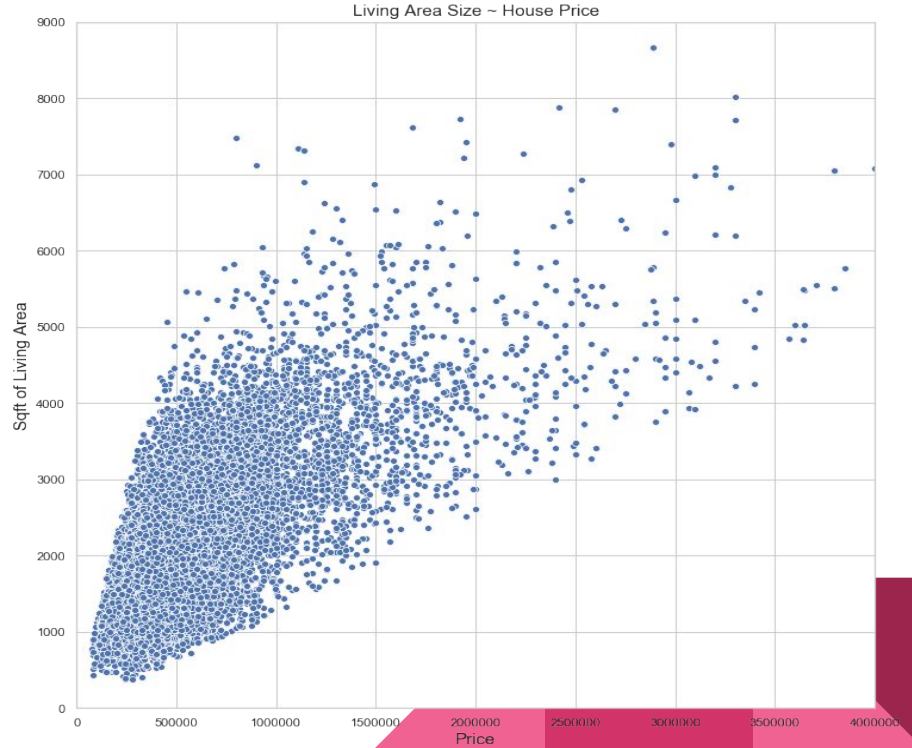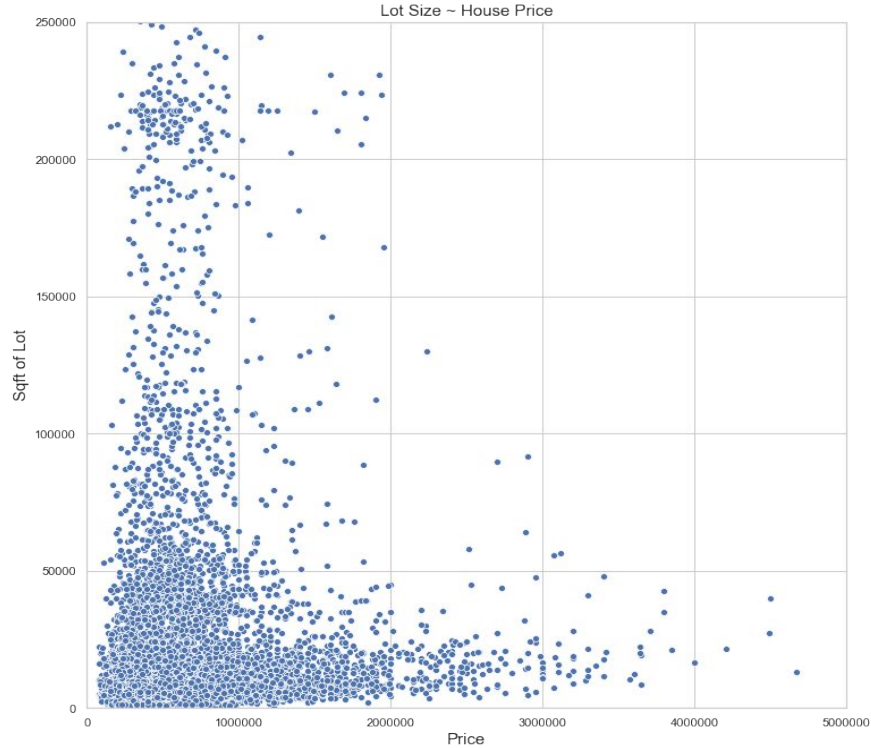- Local distributions by price, living area and lot size?

Do builders make the lot size bigger comparing to living area as the years pass?

Percentage of Lot to the Living Area
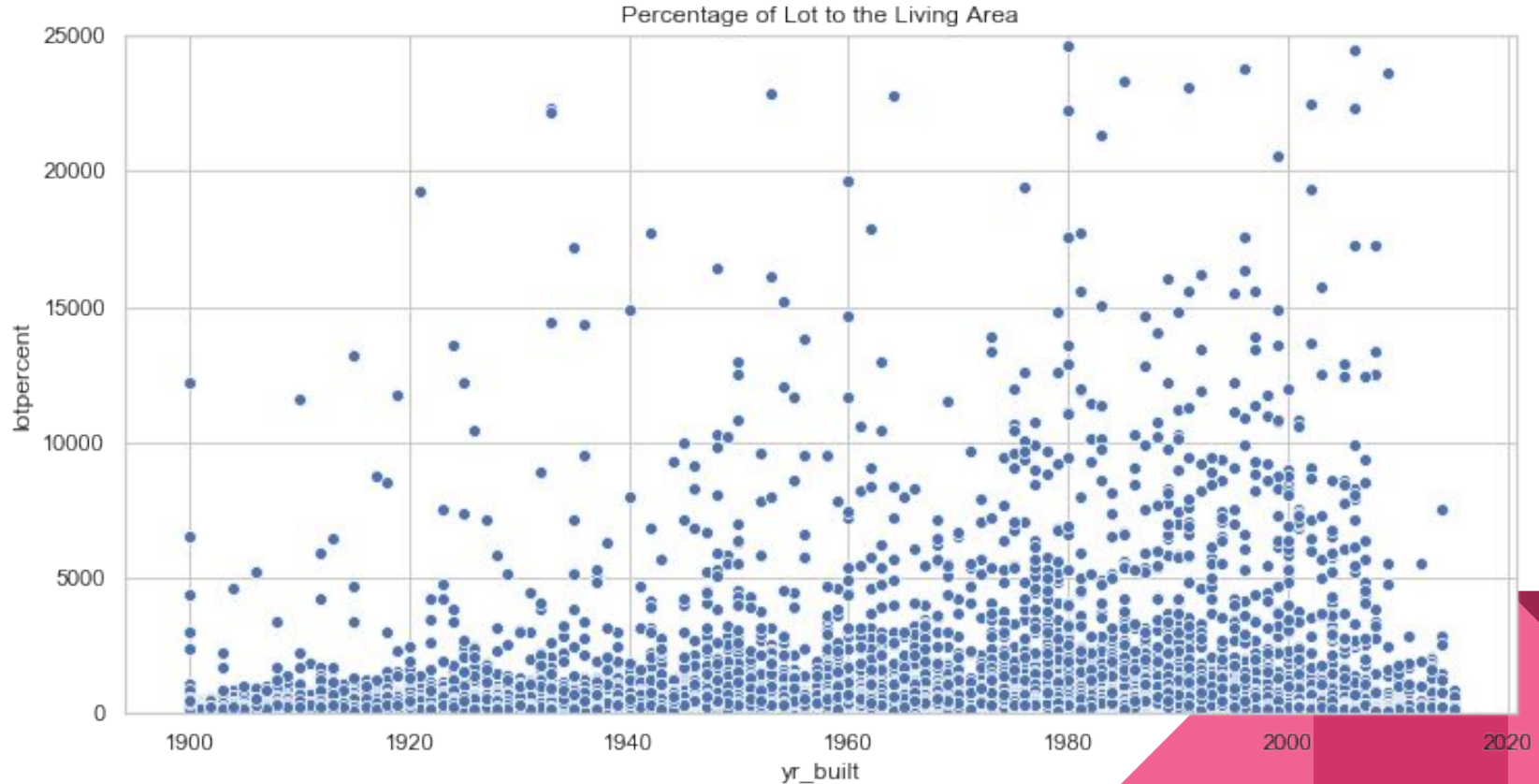
Were more houses built that has large lot size (>500% of the living area) over the years?

Total Number of the Houses Built vs Lot Size > 500 % of the Living Area
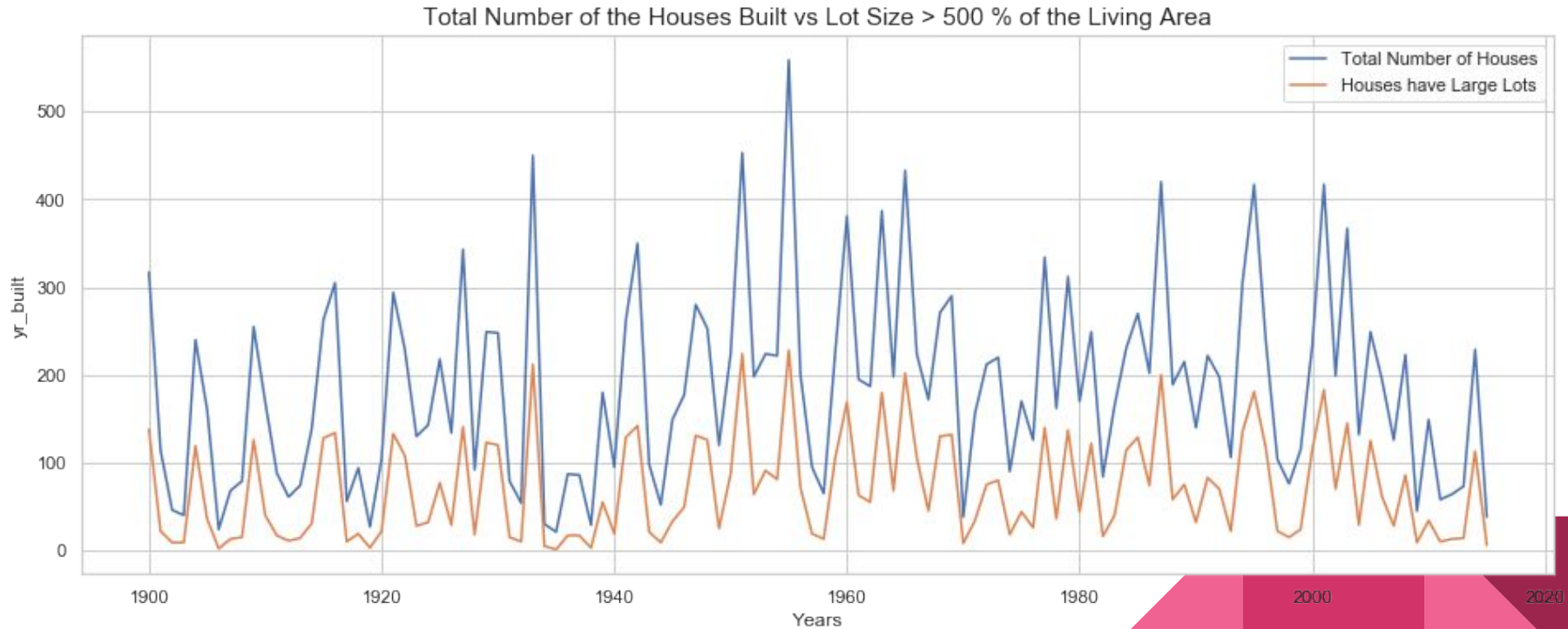
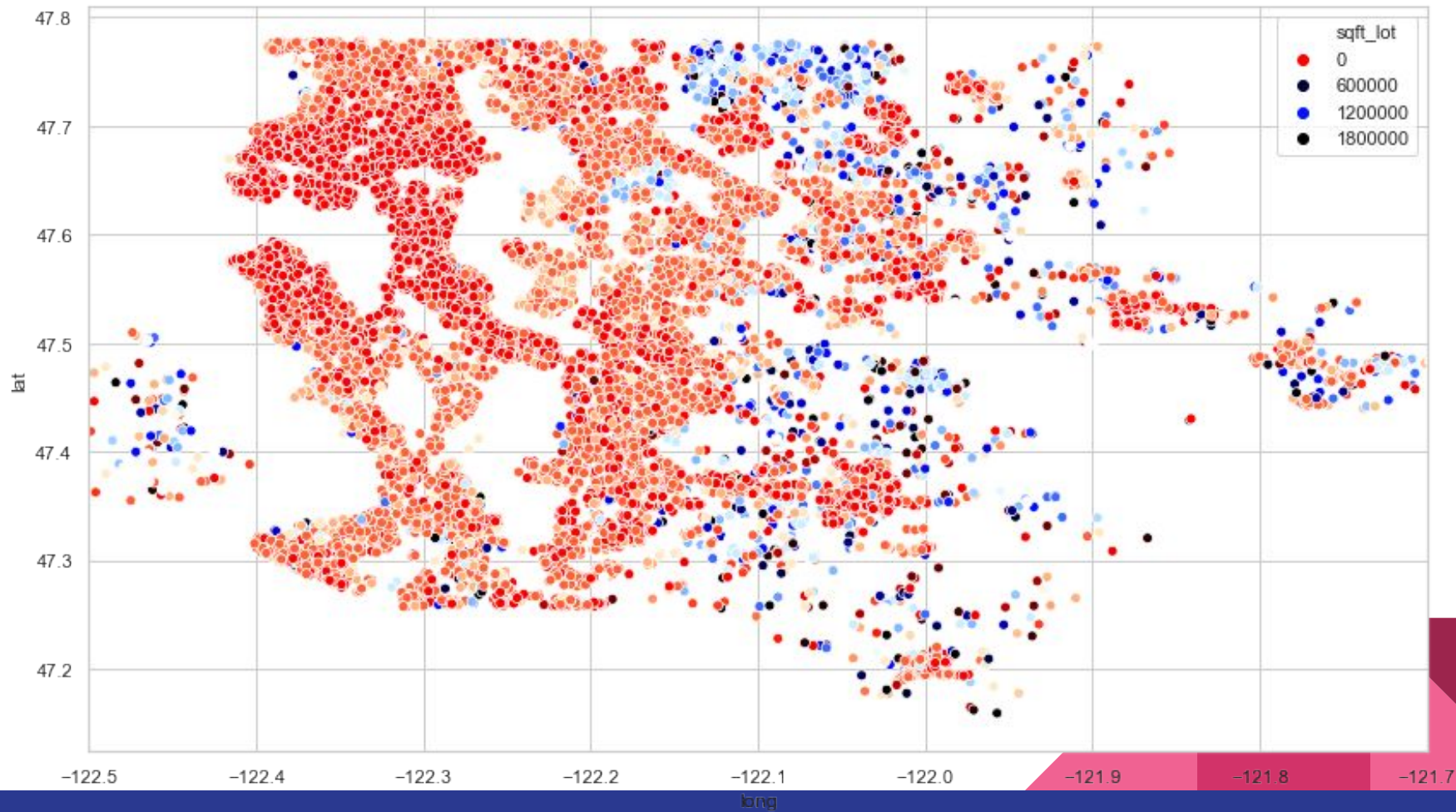# How does the lot & living size effect the price?

How houses are distributed on the map ?

- **By lot size**
- **By living area**
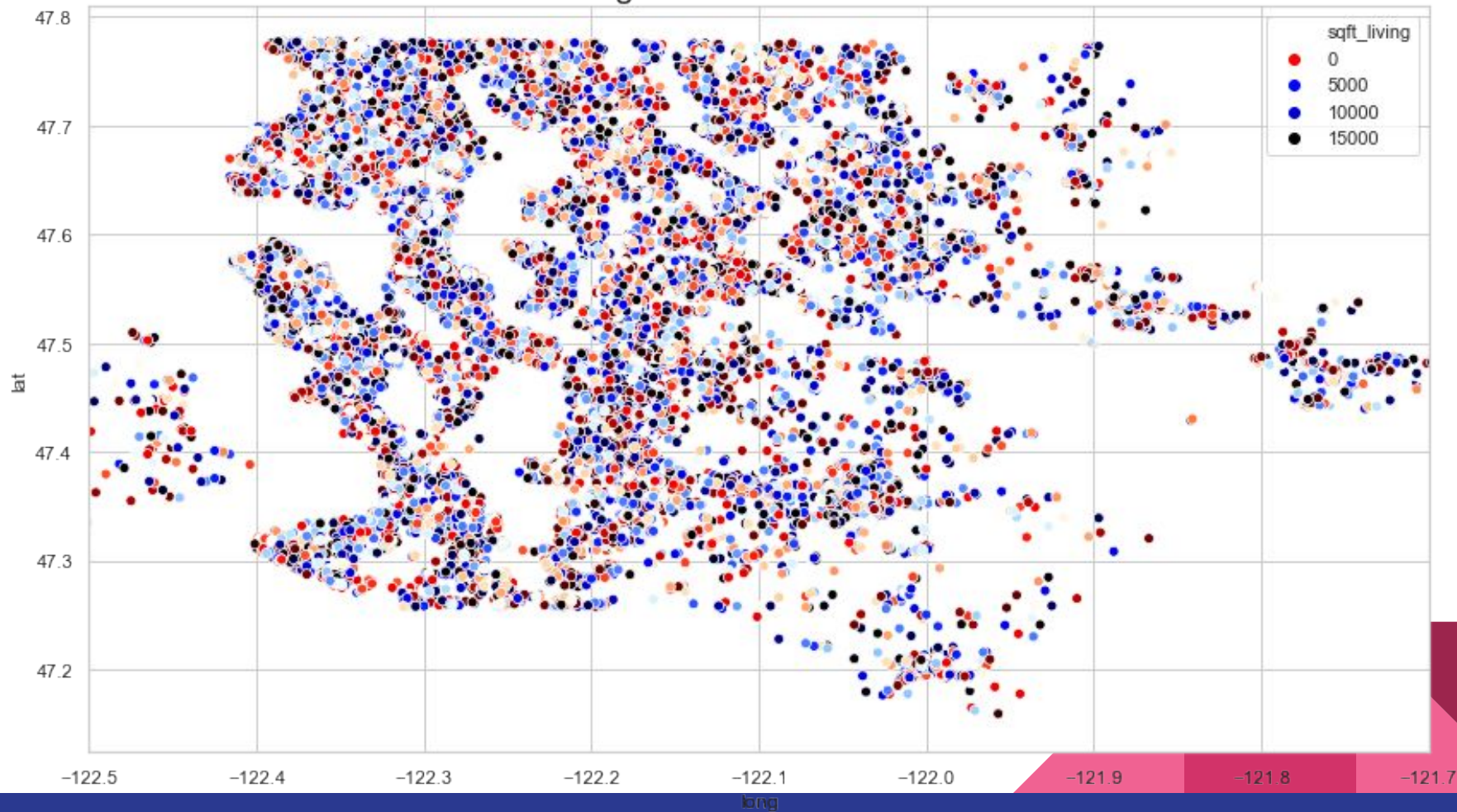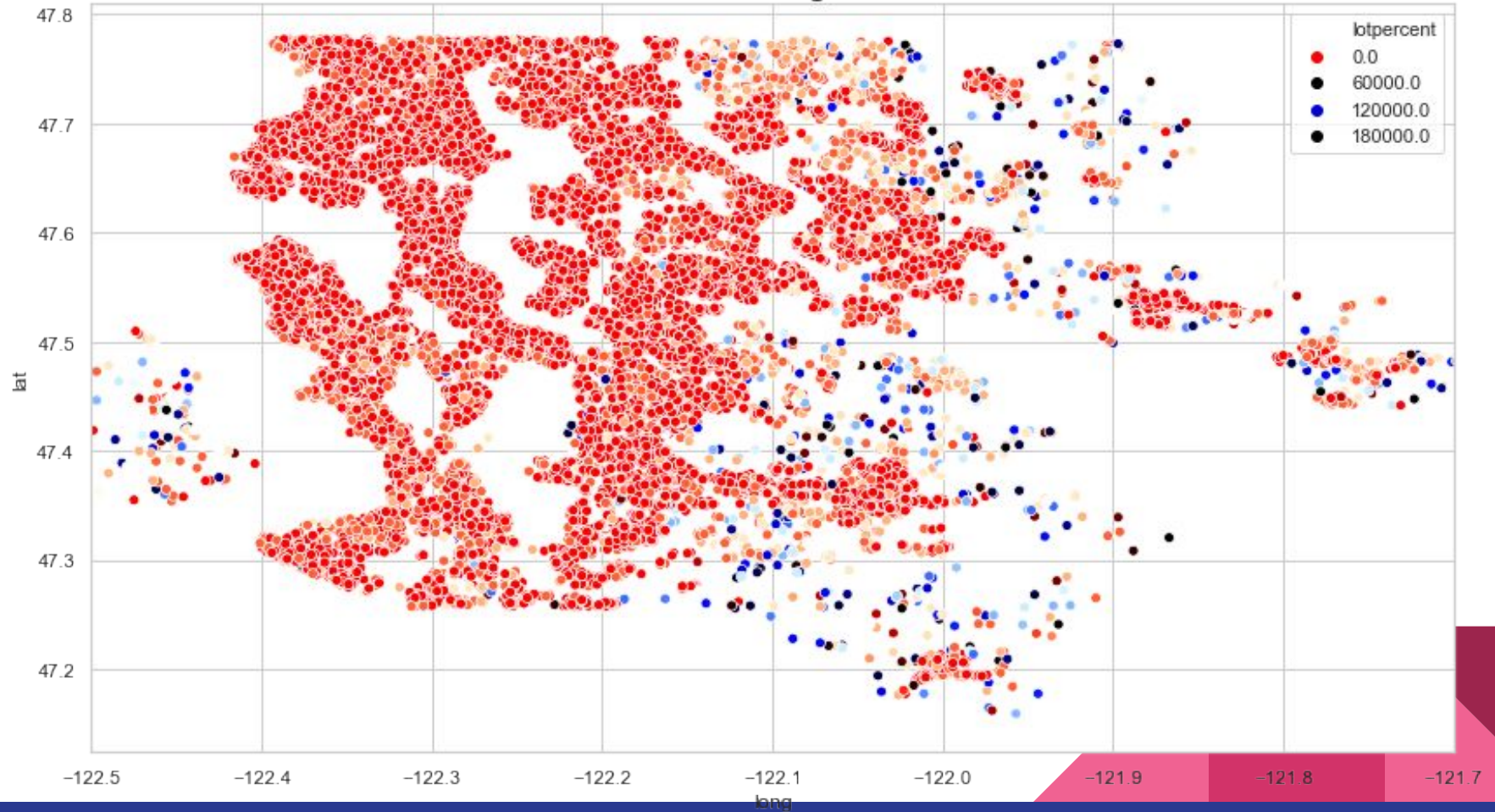- **By percentage of lot & living area**
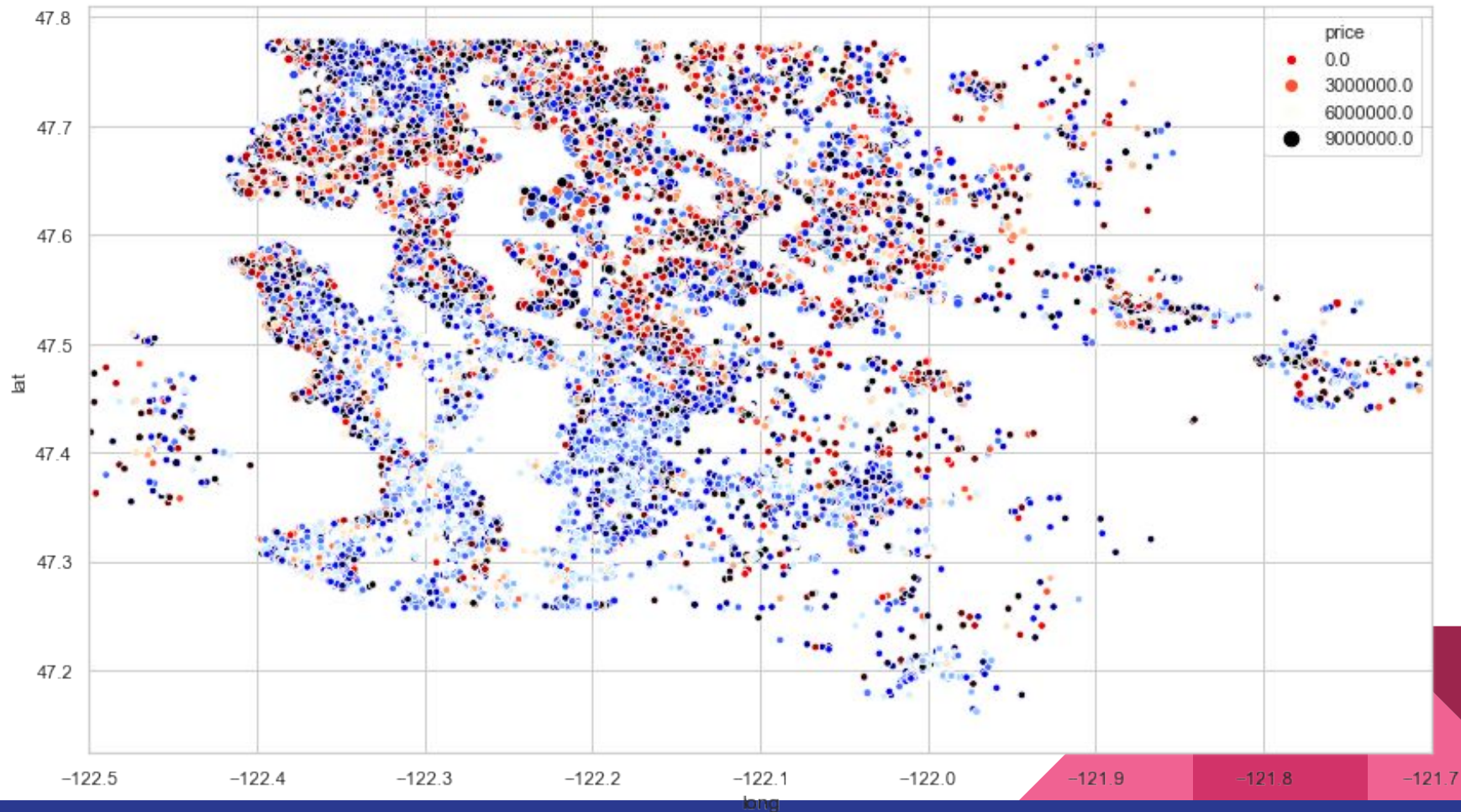- **By price**

Lot Size Distribution

Living Area Size Distribution

Lot Percent of the Living Area Distribution

Price Distribution

Data Analysis Highlights

- Bigger lot size of a house does not always increase the price of the house
- Larger living area is always increase the house price
- In east part of the county , larger lots are preferred by customers
- In the inner east part, living area of the houses are also generally large
- In west side of the county, moderate sizes are more popular
- In upper side of the county house prices are overall more expensive

Further Research

This analysis can be done by zip code to give better advice to the client.

Prediction Model

The house price prediction model uses following criterias:

- Location
- Number of bedrooms
- Living, lot and basement size
- View of the house
- Year of built
- Number of floors
- Condition

Accuracy

The accuracy of the model is:

81%

Most Significant Features

1. Latitude and longitude (upto 1.46 unit)
2. View (the most significant view is waterfront (0.63 unit))
3. Condition grade higher than 3 (3 :0.8,   4: 0.9,   5:1.0,)
4.   Living area size (0.48 unit)