

# Airbnb Price Prediction Modeling

Claire Sarraillé,  
Data Scientist



\$\$\$



# Background & Business Case



- **Background:**
  - Experience = Ability to set sustainable price
- **Problem:**
  - First-time hosts lack experience and new listings lack reviews!
- **Solution:**
  - Sustainable Pricing Tool





# Use Cases

## ————→ **Sustainable Price-Setting**

What price will consistently attract customers given the attributes of the listing.

## ————→ **New Listings – Enough Data?**

Do the variables that exist for new listings belonging to new hosts have enough “signal” for a predictive model?

# Data

**Source**

InsideAirbnb.com

**Instances**

Over 270,000 rental listings

**Target**

price

# Models:

# Ridge and Lasso Regression

Both Aim for Optimal Importance while also...

- **Ridge Regression:**
  - ...Preventing coefficients with extreme values
- **Lasso Regression:**
  - ...Eliminating any unnecessary coefficients

# Results

## Lasso Regression

- RMSE: **268**
- R-Squared: **0.35**

## Ridge Regression:

- RMSE: **265**
- R-Squared: **0.36**

## Final Model:

### Tuned Ridge Regression:

- $\alpha=10$
- solver=svd
- RMSE: **258**
- R-Squared: **0.38**



# Conclusions

01

My model's performance was drastically lower compared to studies I consulted.

02

There weren't great gains in model performance were achieved: Lasso->Ridge->Tuned Ridge

# Recommendations *and* Next Steps

01

Tweak Approach to Problem:

- Which features are most important for highest priced listings?
- Subset by attributes listers can't change like location and property type.
- Different models for each city or region.

02

More Data:

- Review Data
- Additional Amenities

03

Spatial Auto-correlation for lat/long values



# Thank You

Q U E S T I O N S ?



 [linkedin.com/in/claire-sarraille/](https://www.linkedin.com/in/claire-sarraille/)

 [github.com/claiesarraille](https://github.com/claiesarraille)