# London Airbnb Pricing Model

## By Inesa & Grace

# Goals

- **Build** a model to predict price of London Airbnb properties

- **Recommend** strategies to Airbnb hosts

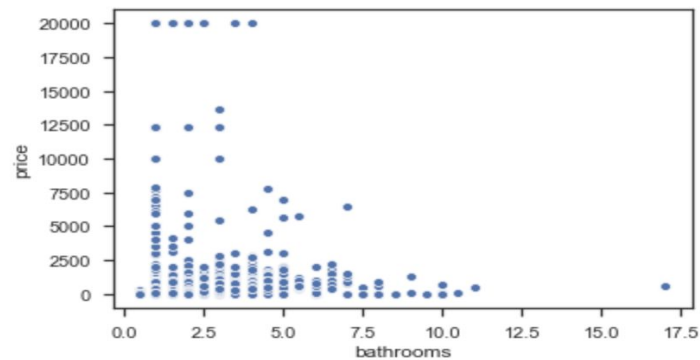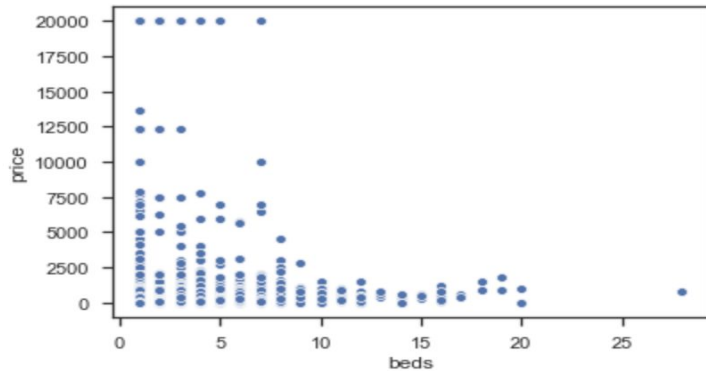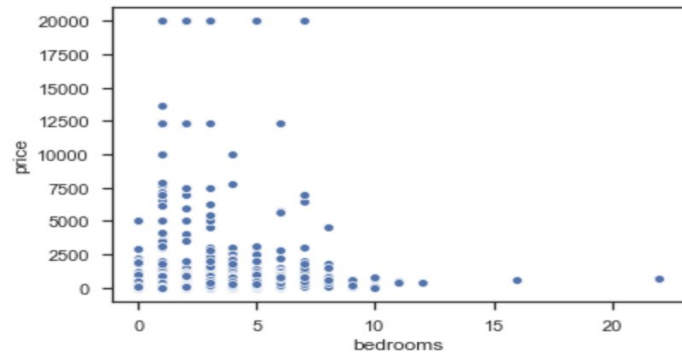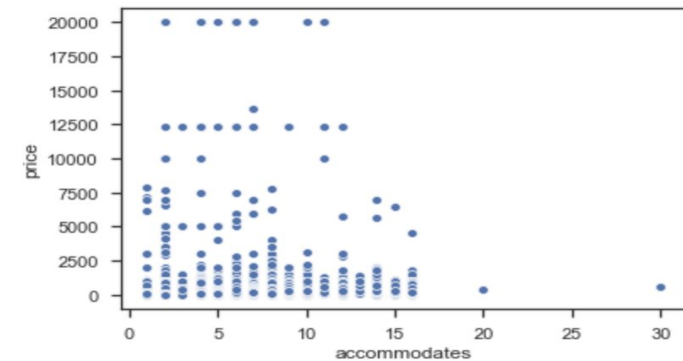    What are the main features that affect price, overall rating and Superhost status?

# Data

- Source: http://insideairbnb.com
- London Airbnb listings as of September 2019
- 85,273 listings: entire homes/apartments, private rooms, hotel rooms, shared rooms

**Inside Airbnb**
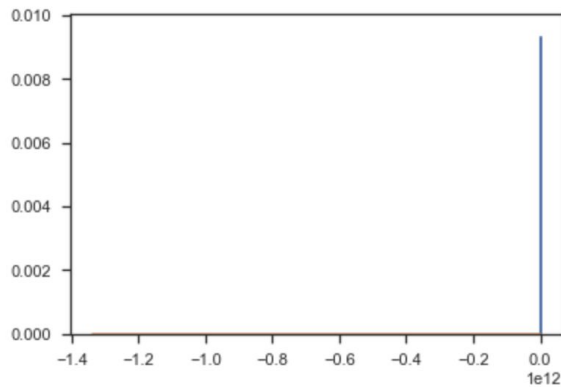Adding data to the debate



1541 properties > £500

# Data

# Initial Models

- Linear Regression:

  - 121 features

  - Negative $R^2$

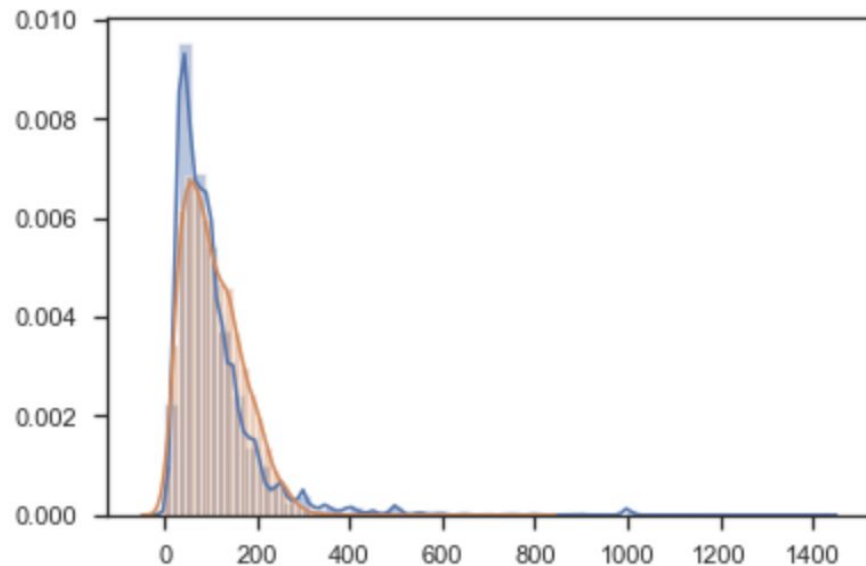- Polynomial Regression:

  - Optimal degree = 3

  - $R^2$ = 31%



Testing dataset: price vs predicted price

{'training_r^2': 0.40, 'mse_train': 5843, 'testing_r^2': -2.83e+16, 'mse_test': 2.67e+20, baseline:  -3.18e+19}
{'degree': 3, 'training_r^2': 0.31,  mse_train: 6464, 'testing_r^2': 0.31, 'mse_test': 6829}

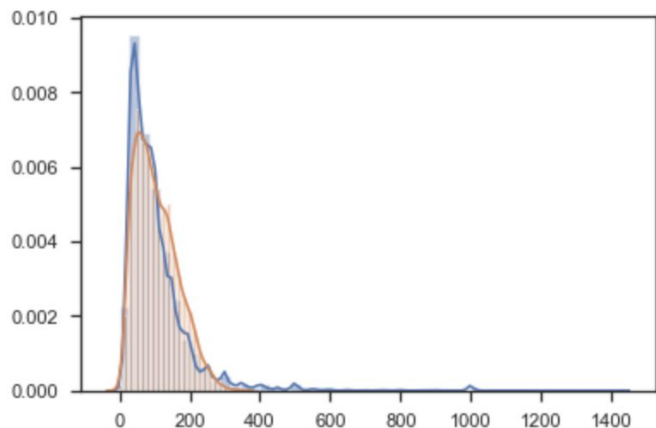# Ridge Regression

- 121 features

- $R^2$ = 40% at alpha = 0.20
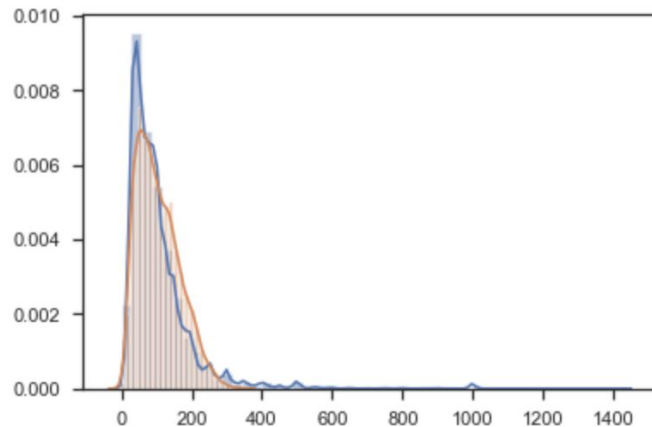


Testing dataset: price vs predicted price

{'alpha': 0.20, 'training_r^2': 0.40, 'mse_train': 5843, 'testing_r^2': 0.40, 'mse_test': 5637, 'baseline': 0.39}

# LASSO Regression

- α = ~0: 121 features, $R^2$ = 40%
- α = 0.20: 50 features, $R^2$ = 38%



Testing dataset: price vs predicted price
α = ~0



Testing dataset: price vs predicted price
α = 0.20

{'alpha': 0.0, 'training_r^2': 0.40, 'mse_train': 5843, 'testing_r^2': 0.40, 'mse_test': 5634, 'baseline': 0.39}
{'alpha': 0.20, 'training_r^2': 0.38, 'mse_train': 6000, 'testing_r^2': 0.38 'mse_test': 5867, 'baseline': 0.38}

airbnb

# Recommendations

- **Price:** increase capacity and number of beds and amenities offered
- **Rating:**
  - encourage guests to write *good* reviews
  - maintain cleanliness and improve accuracy, responsiveness, check-in process
- **Superhost status:** improve overall rating and increase number of amenities offered

# Thank you!

# Questions?