NEXIS DATA SCIENCE PROJECT: NYC AIRBNB LISTING PREDICTION
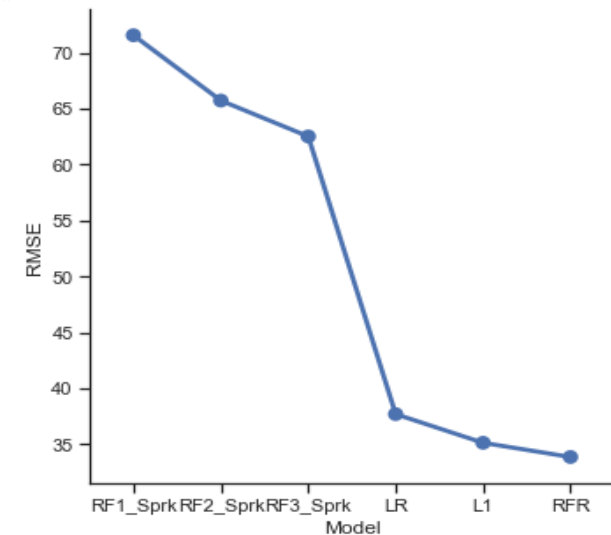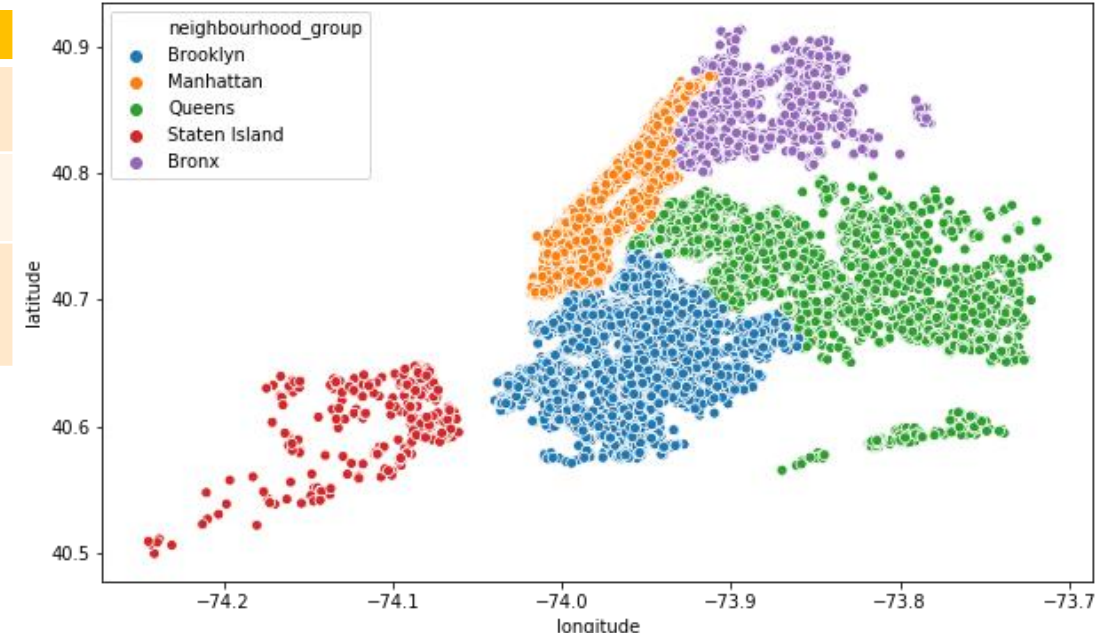## AMIT JADHAV
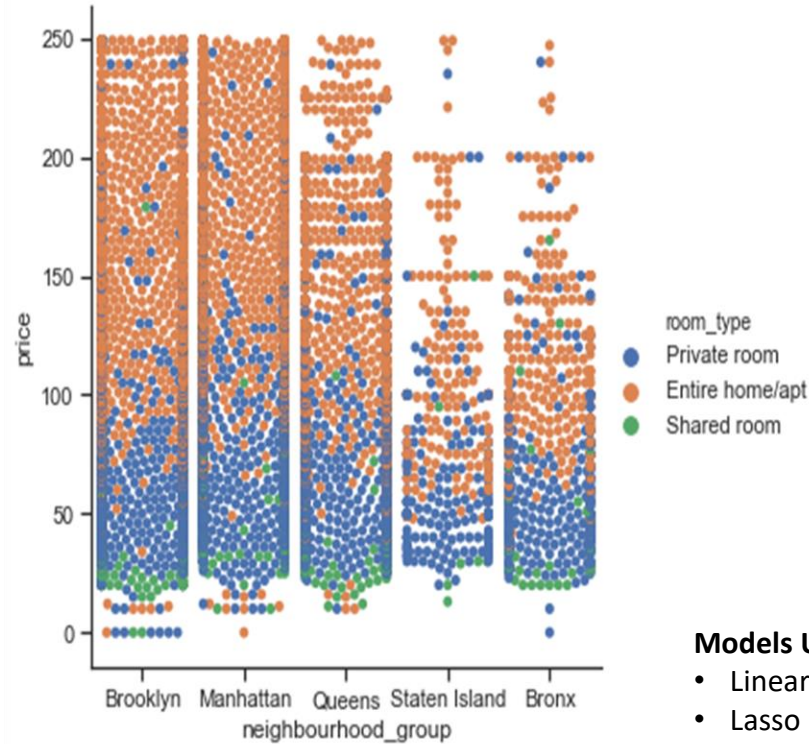
MODELS	RSQUARED	RMSE
LINEAR REGRESSION	50.3%	37.69
LASSO REGRESSION	56.7%	35.12
RANDOM FOREST REGRESSOR	78.14%	33.85

Objective:
The objective of this analysis is to predict the price of an Airbnb listing in NYC. This analysis can help those who are unsure about the price range of their listing.

Models Used:
- Linear Regression
- Lasso Regression
- Random Forest Regressor

Technologies Used:
- PySpark
- Scikit-Learn

Top Features:
- Rt_Entire
- Ng_Manhattan
- Reviews per month
- Nh_Inwood

Result:
Random Forest Regressor performed the best with Rsquared value of 78.14% and RMSE of 33.85

Interesting Insights:
- Entire home/Apt tend to be more expensive and Shared room the least expensive
- Manhattan and Brooklyn have the highest average price while Bronx has the lowest average price
- There are five major neighbourhood groups and 229 neighboorhoods in NYC
- The length of description of listing has no impact on user
- The maximum listing per host is 238
- The average Airbnb listing price in NYC per night is $107
- Most Expensive neighbourhood: Vinegar Hill
- Least Expensive neighbourhood: Inwood