

Study Area: New York city (Source: kaggle)

Dataset

- Describes the listing activity of homestays in New York City
- Each Airbnb is categorized with neighbourhood

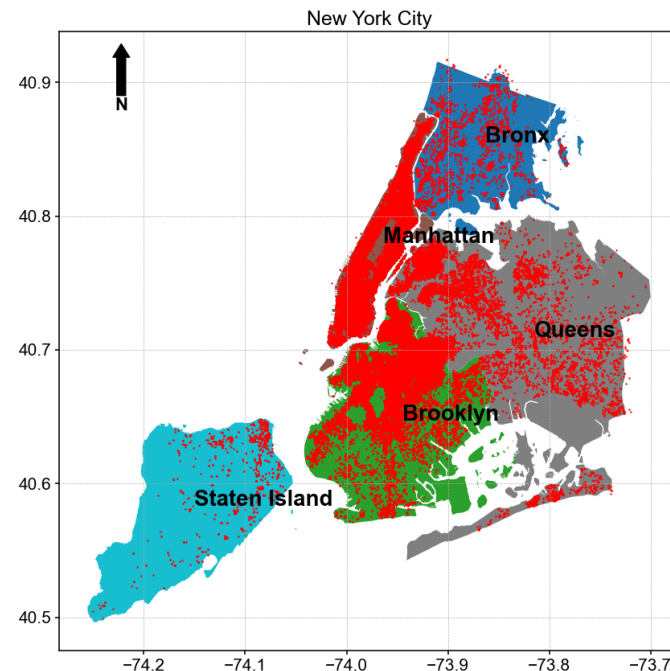
Objective

- Prediction of price based on other features
- Identification of important features
- How do reviews and ratings influence the popularity and booking rate of listings

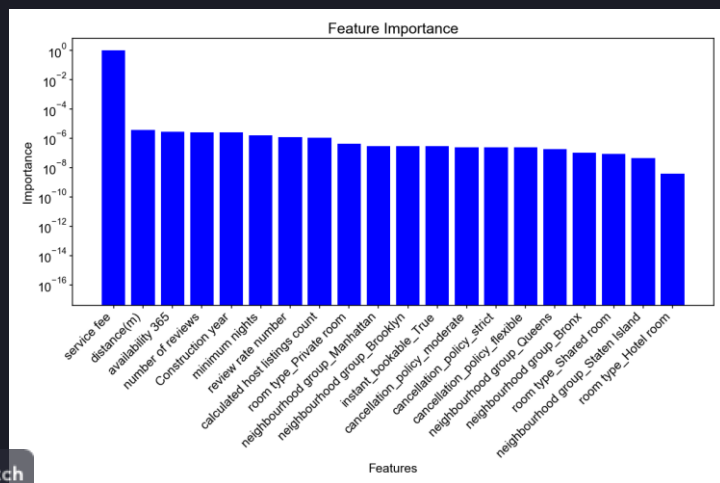
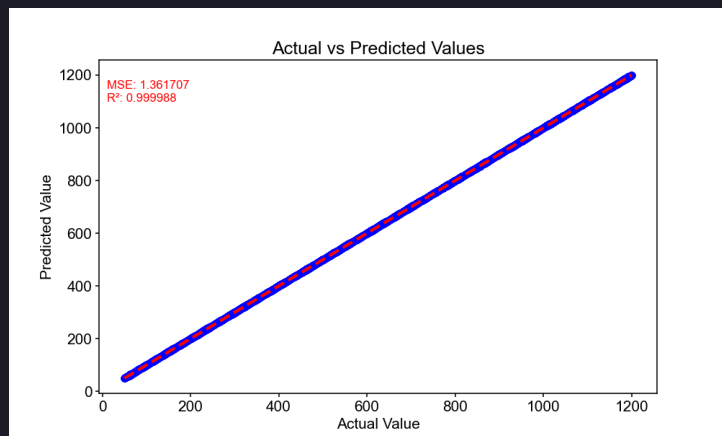
Methods

- Data collection, data cleaning, data transformation
- Machine learning: Random Forest

Why Random Forest?



Random Forest



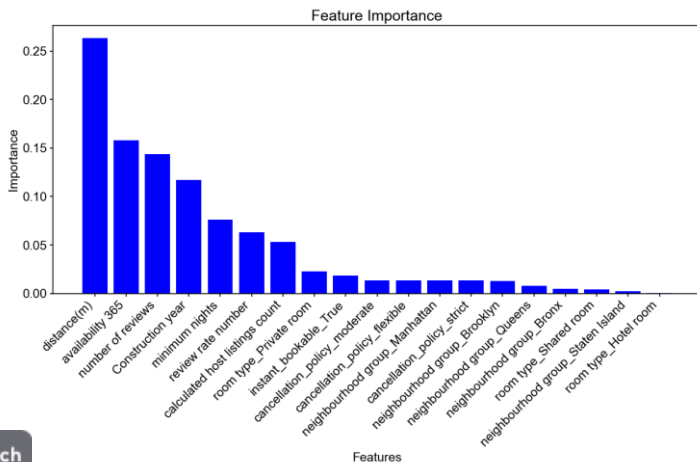
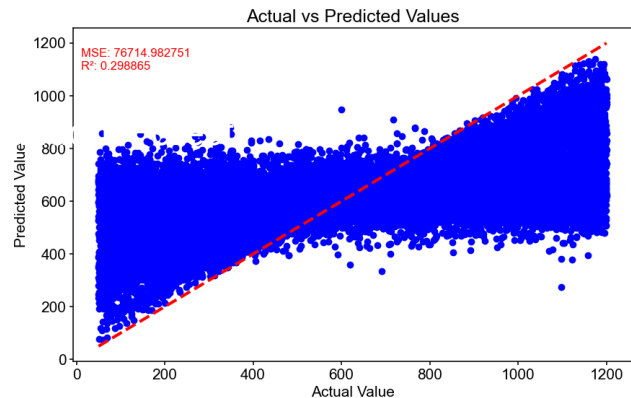
Mean Squared Error: 1.3617070706567749

R² value: 0.9999875658747559

Feature importances from the Random Forest model:

	Importance
service fee	9.999820e-01
distance(m)	3.826681e-06
availability_365	2.703783e-06
number of reviews	2.673413e-06
Construction year	2.466152e-06
minimum nights	1.578167e-06
review rate number	1.257907e-06
calculated host listings count	1.056798e-06
room type_Private room	4.248019e-07
neighbourhood group_Manhattan	3.073556e-07
neighbourhood group_Brooklyn	2.965780e-07
instant_bookable_True	2.814046e-07
cancellation_policy_moderate	2.396607e-07
cancellation_policy_strict	2.359052e-07
cancellation_policy_flexible	2.348185e-07
neighbourhood group_Queens	1.798479e-07
neighbourhood group_Bronx	1.034339e-07
room type_Shared room	8.297237e-08
neighbourhood group_Staten Island	4.387923e-08
room type_Hotel room	3.810702e-09

Random Forest (Without Service Fee)

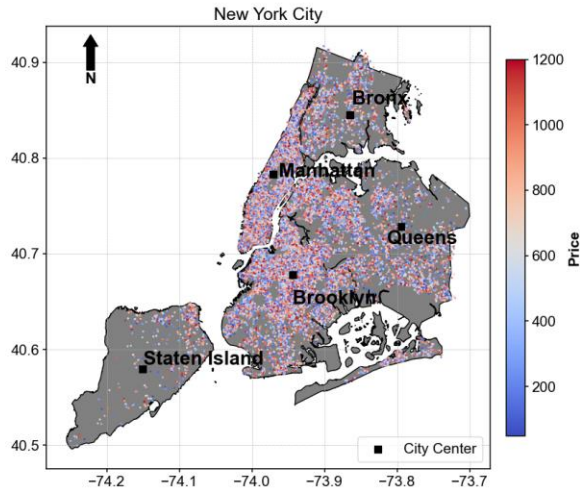


Mean Squared Error: 76714.9827508757

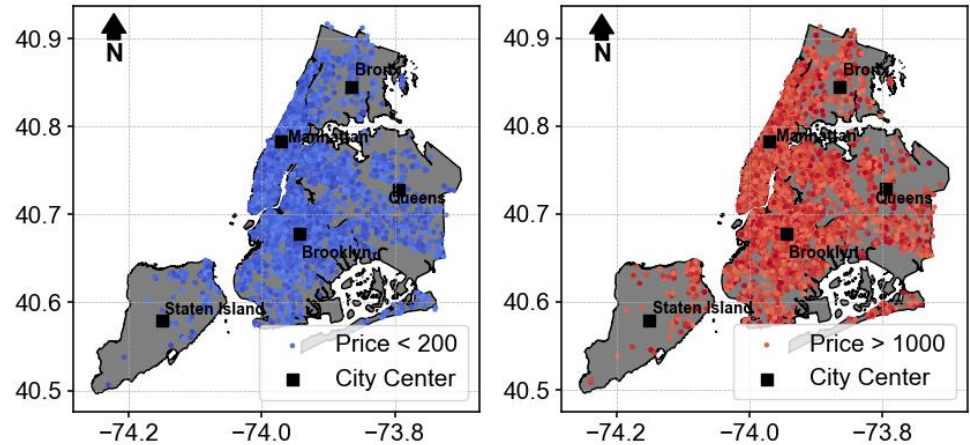
R² value: 0.29886452798830965

Feature importances from the Random Forest model:

	Importance
distance(m)	0.263409
availability_365	0.158274
number of reviews	0.143557
Construction year	0.117131
minimum nights	0.075848
review rate number	0.063116
calculated host listings count	0.053195
room type_Private room	0.022491
instant_bookable_True	0.018229
cancellation_policy_moderate	0.013523
cancellation_policy_flexible	0.013450
neighbourhood_group_Manhattan	0.013226
cancellation_policy_strict	0.013168
neighbourhood_group_Brooklyn	0.012719
neighbourhood_group_Queens	0.007779
neighbourhood_group_Bronx	0.004533
room type_Shared room	0.004099
neighbourhood_group_Staten Island	0.002002
room type_Hotel room	0.000250



New York City



- There is no specific hotspot for higher and lower prices
- There is no visible change of prices with the distance from center of the neighbourhood group

Thanks for your attention