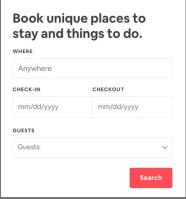


About Airbnb



- Airbnb is the most demanding online marketplace where people seek and offer lodging as homestay.
 - Over 7 million accommodations listed
 - Over half a billion guest check-ins
 - o In 191 countries & regions
- Renters list a variety of criteria to attract renters and increase their prices.
 - Neighborhood
 - Offered amenities like wifi or hair dryers
 - Number of bedrooms/bathrooms
 - Pictures



Research Question



What features predict airbnb prices in Florence, Italy?

Dataset Details



- Original Dataset
 - o 58MB
 - 97 variables
 - Pictures
 - Reviews
 - Location information
 - Price information
 - Amenities offered
 - Counts of bedrooms & bathrooms
 - Among others
 - Total Observations in dataset 12304
 - Dataset URL: http://insideairbnb.com/get-the-data.html

Data Preprocessing

Missing Values Table

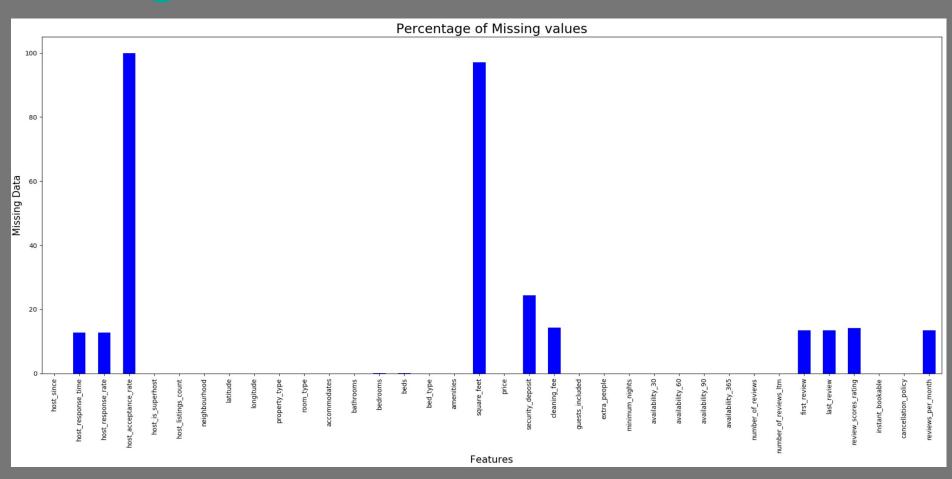


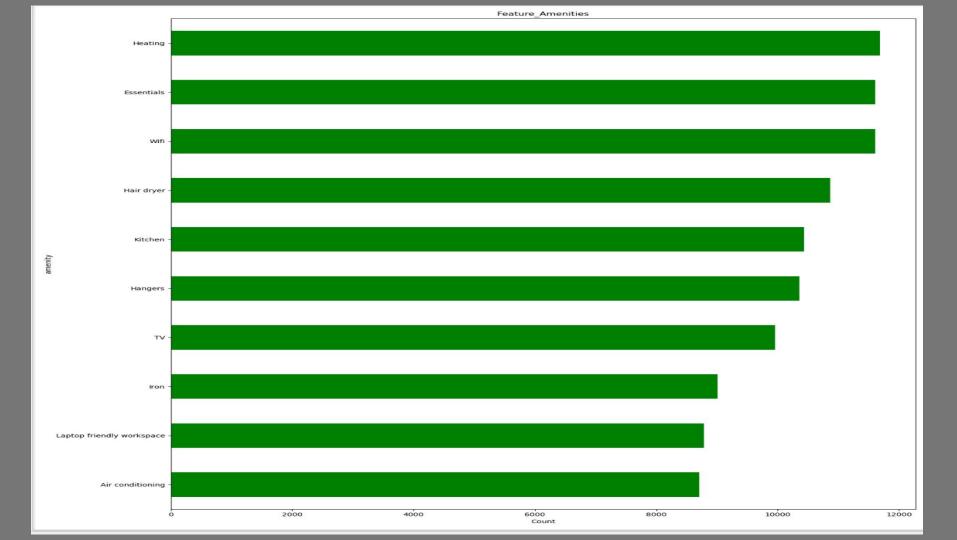
Table 1	
Missing values and Data Type of Variables in	Dataset

Target	NaNs	Dtype	Features	NaNs	DType				
price 0 object			host_since	5	datetime64[ns]				
			host response time	1566	object				
			host response rate	1566	object				
			host acceptance rate	12304	float64				
			host is superhost	5	object				
			host listings count	5	float64				
			neighbourhood	О	object				
			latitude	О	float64				
			longitude	О	float64				
			property type	О	object				
			room_type	О	object				
			accommodates	О	int64				
			bathrooms	5	float64				
			bedrooms	7	float64				
			beds	23	float64				
			bed_type	О	object				
			amenities	О	object				
			square feet	11940	float64				
			security deposit	3010	object				
			cleaning fee	1766	object				
			guests_included	О	int64				
			extra people	О	object				
			minimum_nights	О	int64				
			availability_30	О	int64				
			availability_60	О	int64				
			availability_90	О	int64				
			availability_365	О	int64				
			number of reviews	О	int64				
			number of reviews ltm	О	int64				
			first_review	1663	datetime64[ns]				
			last_review	1663	datetime64[ns]				
			review scores rating	1736	float64				
			instant bookable 0 object cancellation policy 0 object						
			cancellation policy	object					

Missing Values







Experimental Setup



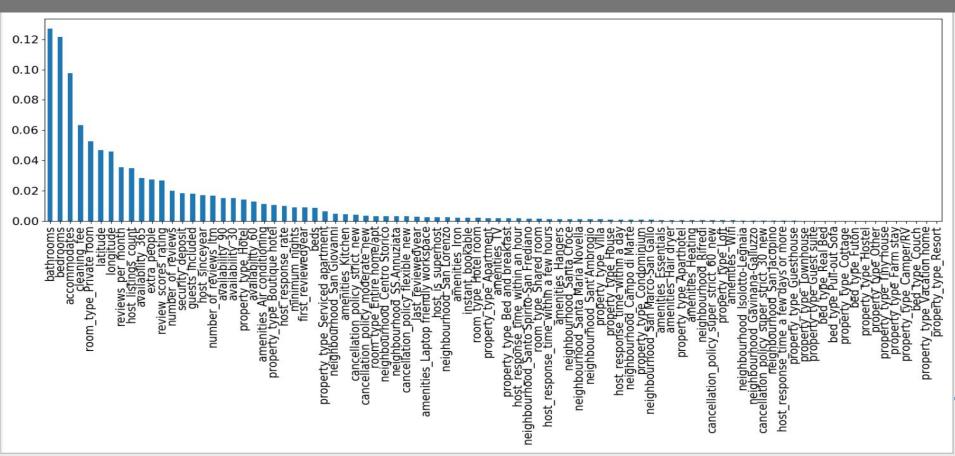
- Feature Selection was conducted using Random Forest
 - Results will be shown during the demo

- Once features were paired down, the research question was answered using linear regression.
 - Results will be shown during the demo

Feature Selection

Overall Features





Linear Regression

host_listings_count	-	-0.061	0.016	0.081	0.09	0.057	V	-0.044	-0.039	-0.13	-0.011	-0.045	-0.054	-0.028	-0.11	-0.14	-0.13	-0.16	-0.14	-0.0055	-0.096		1.00
latitude -	-0.061	1	-0.1	-0.059	-0.047	-0.029	-0.034	-0.12	-0.044	0.0034	0.032	0.014	0.0037	-0.052	-0.014	0.0021	0.015	0.02	0.099	0.022	0.14	-0.16	
longitude -	0.016	-0.1	1	0.00096	0.027	0.017	0.0082	0.04	0.0016	0.0016	-0.0049	-0.0048	-0.0072	0.0094	0.00068	-0.011	0.0014	-0.025	-0.052	0.017	-0.025	0.05	
accommodates -	0.081	-0.059	0.00096	1	0.6	0.75	0.15	0.44	0.55		0.03	0.024	0.019	0.065	-0.082	-0.07	-0.0092	-0.11	-0.031	-0.026	-0.38	0.52	
bathrooms -	0.09	-0.047	0.027	0.6	1	0.66	0.14	0.36	0.31	0.11	0.033	0.0058	-0.0089	0.011	-0.083	-0.082	0.017	-0.1	-0.034	-0.017	-0.1	0.48	- 0.75
bedrooms -	0.057	-0.029	0.017	0.75	0.66	1	0.13			0.13	0.03	-0.0016	-0.017	0.024	-0.11	-0.11	0.015	-0.14	-0.0065	-0.018	-0.2	0.51	
security_deposit -	0.17	-0.034	0.0082	0.15	0.14	0.13	1	0.17	0.097	0.07	-0.045	-0.053	-0.056	-0.038	-0.051	-0.07	0.039	-0.11	-0.098	-0.018	-0.16	0.19	
deaning_fee -	-0.044	-0.12	0.04	0.44	0.36	0.42	0.17	1	0.32	0.08	0.0046	0.0022	-0.0028	0.08	-0.1	-0.12	-0.0041	-0.15	-0.083	-0.0066	-0.32	0.39	
guests_included -	-0.039	-0.044	0.0016	0.55	0.31		0.097	0.32	1		0.0034	0.016	0.022	0.065	0.012	0.012	0.027	-0.018	-0.0053	-0.022	-0.22	0.27	- 0.50
extra_people -	-0.13	0.0034	0.0016	0.21	0.11	0.13	0.07	0.08	0.31	1	0.0027	0.031	0.042	0.05	0.054	0.033	0.11	0.0045	0.00076	-0.022	-0.058	0.055	
availability_30 -	-0.011	0.032	-0.0049	0.03	0.033	0.03	-0.045	0.0046	0.0034	0.0027	1	0.85	0.72	0.32	-0.13	-0.17	-0.11	-0.16	0.024	0.018	0.1	0.04	
availability_60 -	-0.045	0.014	-0.0048	0.024	0.0058	-0.0016	-0.053	0.0022	0.016	0.031	0.85	1	0.95		-0.034	-0.038	-0.057	-0.025	0.049	0.00099	0.062	0.041	
availability_90 -	-0.054	0.0037	-0.0072	0.019	-0.0089	-0.017	-0.056	-0.0028	0.022	0.042	0.72	0.95	1		0.018	0.032	-0.027	0.041	0.043	-0.0072	0.031	0.04	- 0.25
availability_365 -	-0.028	-0.052	0.0094	0.065	0.011	0.024	-0.038	0.08	0.065	0.05	0.32	0.48	0.57	1	0.077	0.029	-0.015	-0.011	-0.11	-0.0026	-0.072	0.12	
number_of_reviews -	-0.11	-0.014	0.00068	-0.082	-0.083	-0.11	-0.051	-0.1	0.012	0.054	-0.13	-0.034	0.018	0.077	1	0.72	0.087	0.6	0.22	-0.023	0.037	-0.13	
imber_of_reviews_ltm -	-0.14	0.0021	-0.011	-0.07	-0.082	-0.11	-0.07	-0.12	0.012	0.033	-0.17	-0.038	0.032	0.029	0.72	10	0.15	0.84	0.054	-0.02	0.0038	-0.14	
review_scores_rating -	-0.13	0.015	0.0014	-0.0092	0.017	0.015	0.039	-0.0041	0.027	0.11	-0.11	-0.057	-0.027	-0.015	0.087	0.15	1	0.15	0.073	-0.0022	-0.016	0.051	- 0.00
reviews_per_month -	-0.16	0.02	-0.025	-0.11	0.1	-0.14	-0.11	0.15	-0.018	0.0045	-0.16	0.025	0.041	-0.011	0.6	0.84	0.15	1	0.18	-0.0095	0.04	-0.18	
host_sinceyear -		0.099	-0.052	-0.031	-0.034	-0.0065	-0.098	-0.083	-0.0053	0.00076	0.024	0.049	0.043	-0.11	-0.22	0.054	0.073	0.18	1	0.018	0.037	-0.098	
property_type_Hotel -		0.022	0.017	-0.026	-0.017	-0.018	-0.018	-0.0066	-0.022	-0.022	0.018	0.00099	-0.0072	-0.0026	-0.023	-0.02	-0.0022	-0.0095	0.018	1	0.00073	0.1	0.25
m_type_Private room -	-0.096	0.14	-0.025	-0.38	-0.1	-0.2	-0.16	-0.32	-0.22	-0.058	0.1	0.062	0.031	-0.072	0.037	0.0038	-0.016	0.04	0.037	0.00073	1	-0.33	
price	0.39	0.16	0.05	0.52	0.48	0.53	0.20	0.39	0.27	0.055	0.04	0.002	0.031	0.22	-0.13	-0.14	0.051	-0.18	-0.098	0.00073	-0,33		
price	host_listings_count -	lar frude	kongitude	accommodates	bathrooms -	bedrooms -	security_deposit_	deaning_fee	guests_included -	extra_people	availability_30	availabilty_60	availability_90	availability_365	number_of_reviews	number of reviews Itm	review_scores_rating §	reviews per month a	host_sinceyear	property_type_Hotel	oom_type_Private room t	price -	

Results



• We'll now use the GUI to demonstrate our findings from our dataset



Limitations



Missing values

User self-reported data - may contain social desirability bias

