AIRBNB'S BUSINESS AND CUSTOMER ANALYSIS

MIS 6V99 – HIGH PERFORMANCE ANALYTICS

GROUP 4
SILVER TANKS

TEAM MEMBERS

ABHINAV SINGH TILWAR
PRAVEEN SUBRAMANIAN
MOHAMMAD ZAIN FARAZ

Contents

Executive Summary

Real estate market has been considered as one of the most profitable markets in US. As per Maslow's hierarchy of needs, shelter comes under a human being's most essential needs which are physiological needs and safety needs. The meaning of shelter has now taken the form of a condominium, an apartment, a mansion etc. While it is true that shelter was a Man's basic need, ownership of this shelter which could be a residential property has become one of the hottest means to provide monetary profitability to owners.

Our idea stands with the base of analyzing various factors that determine the price of the Listing. With data mining gaining tremendous popularity to make informed decisions, our project is based upon the factors such as location, availability, reviews, amenities, quality of living, places to visit etc. along with several others features such as ease of booking, host response, renter satisfaction and mutual verification that will help us to determine the Price of a Listing. This information will help customers to weigh the tradeoffs between different listing and will help the property owner to set the price of his/her property.

Our system aims at establishing connections between these fields and generating profitable solutions to both – The Property owner and the customer. Based on our analysis, another aim of the project is to help property owner to classify existing listings based on various user preferences and use predictive analysis to classify new hosts into these categories for future bookings. A customer can leverage the findings to make best possible cost-effective decision to accomplish the best deal. These predictions will assist The Renter and Customer to exploit the value of data to make informed decisions.

Project Background – Data and Formulation and Presentation

Data Description:

The dataset which we are using this project is a second hand data obtained from an online source http://insideairbnb.com/get-the-data.html

We are using 4 excel files namely listings.csv, calender.csv, reviews.csv, neighbourhood.csv and a geojson file named neighbourhood.geojson for each city as a dataset. Our Dataset consists of the above mentioned files for 5 major US cities. Usually the largest file for each city is the listings.csv file which has 67 columns and 5,000 to 20,000 rows (depending on the city).
The data is cleansed, analyzed, aggregated and utilizes public information compiled from Airbnb website. This dataset contains information about Airbnb listings used in computing target variables based on the project's objectives. Each variable has its own importance in determining the Target variable. We are sure that this project will help us in enhancing customer's experience when he looks forward to book an airbnb listing.

Data Partitioning

Training:-50
Validate:-30
Test:-20

Variables:

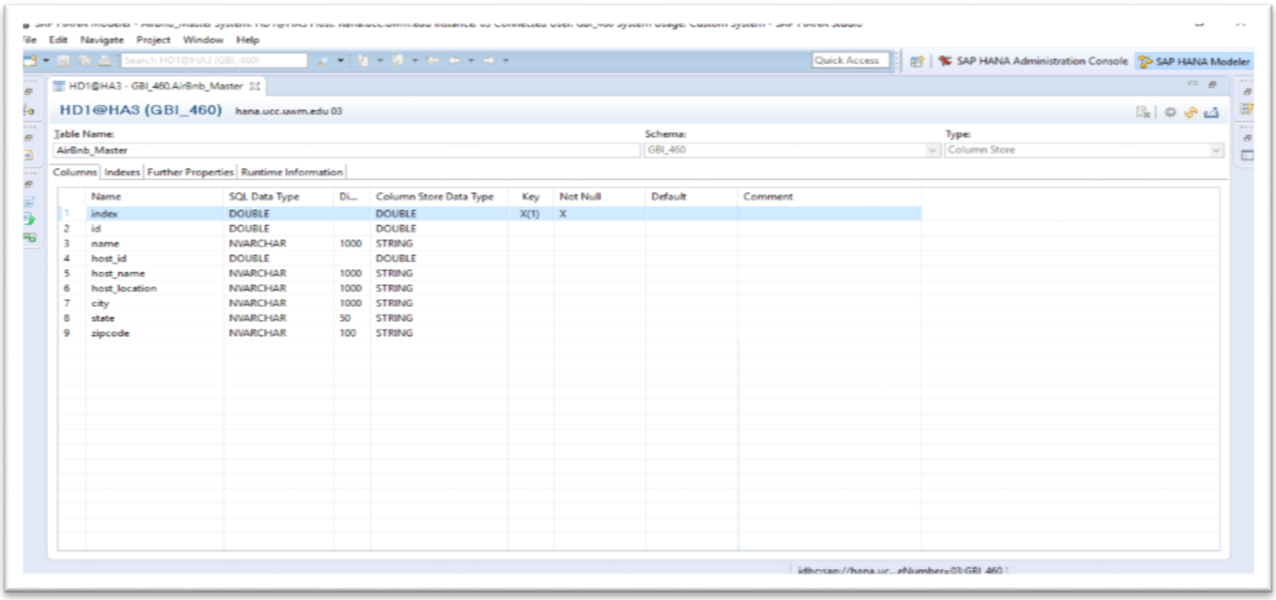
The variables used for an objective depends on the type of objective.

1) To classify existing listings based on various user preferences and to classify new hosts into these categories for future bookings implementing predictive analytics. For this objective we are using id, host_name, host_response_time, host_acceptance_rate, street, neighbourhood, zipcode, latitude, longitude, property_type, room_type, bathrooms, bedrooms, amenities, square feet, weekly_price, monthly_price, security_deposit, guests_included, extra_people, minimum_nights, maximum_nights, has_availability, availability_30, review_scores_rating, review_scores_cleanliness, review_scores_location, review_scores_value, cancellation_policy from listings.csv. Date, available, price from calender.csv file.

2) To suggest dynamic pricing for each listing by using the current availability and listing history. For this objective we are using id, street, neighbourhood, market, property_type, room_type, square feet, price, weekly_price, monthly_price, security_deposit, availability_30, availability_60, availability_90, availability_365, number_of_reviews, review_scores_rating, review_scores_accuracy, review_scores_cleanliness, review_scores_checkin, review_scores_value, review_scores_location, review_scores_communication, review_scores_value, cancellation_policy from listings.csv. date, available, price from calender.csv file.

3) To create a word-cloud from listing descriptions from which users can pull data to find a perfect match for a dream vacation. For this objective we are using id, name, summary, space, description, neighbourhood_overview, notes, transit, street, property_type from listings.csv.

- **Table Definition**

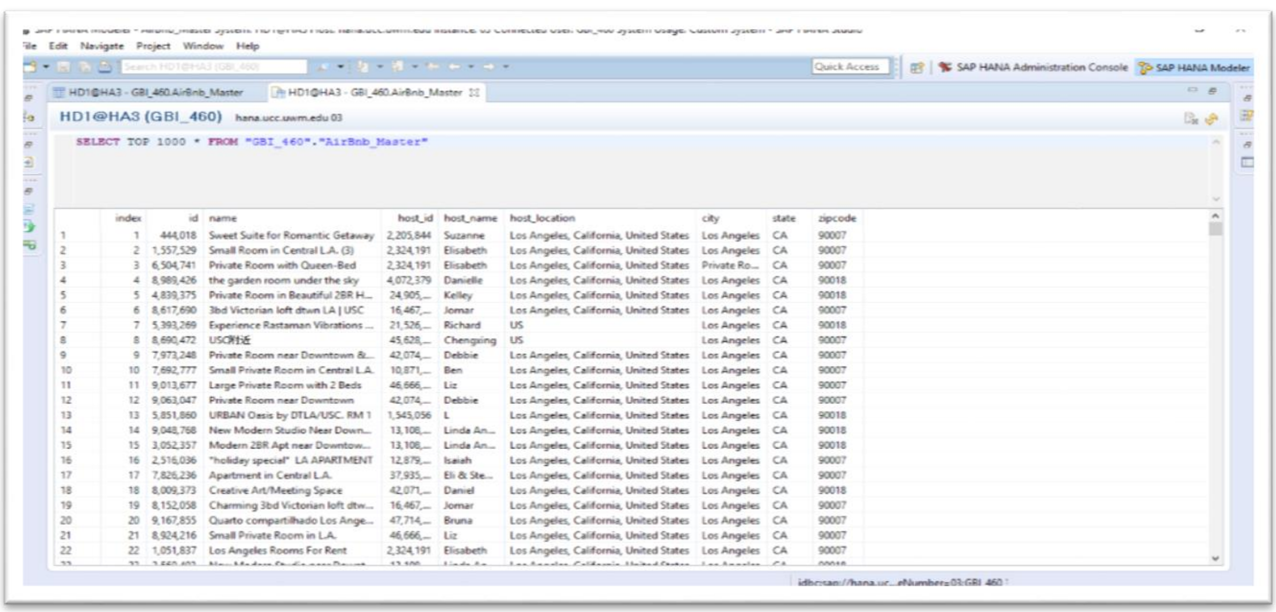


HD1@HA3 (GBI_460) hana.ucc.uwm.edu 03

Table Name: AirBnb_Master Schema: GBI_460 Type: Column Store

	Name	SQL Data Type	Di...	Column Store Data Type	Key	Not Null	Default	Comment
1	index	DOUBLE		DOUBLE	X(1)	X		
2	id	DOUBLE		DOUBLE				
3	name	NVARCHAR	1000	STRING				
4	host_id	DOUBLE		DOUBLE				
5	host_name	NVARCHAR	1000	STRING				
6	host_location	NVARCHAR	1000	STRING				
7	city	NVARCHAR	1000	STRING				
8	state	NVARCHAR	50	STRING				
9	zipcode	NVARCHAR	100	STRING				

- **Table Content**

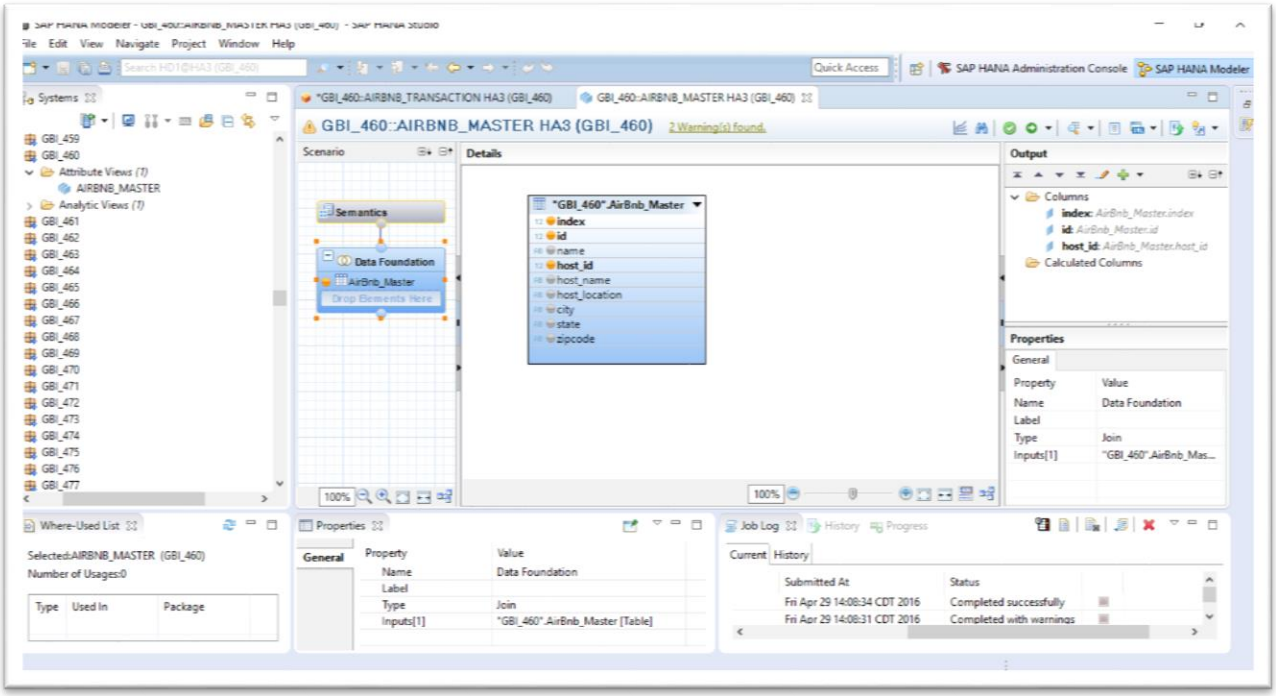


HD1@HA3 (GBI_460) hana.ucc.uwm.edu 03

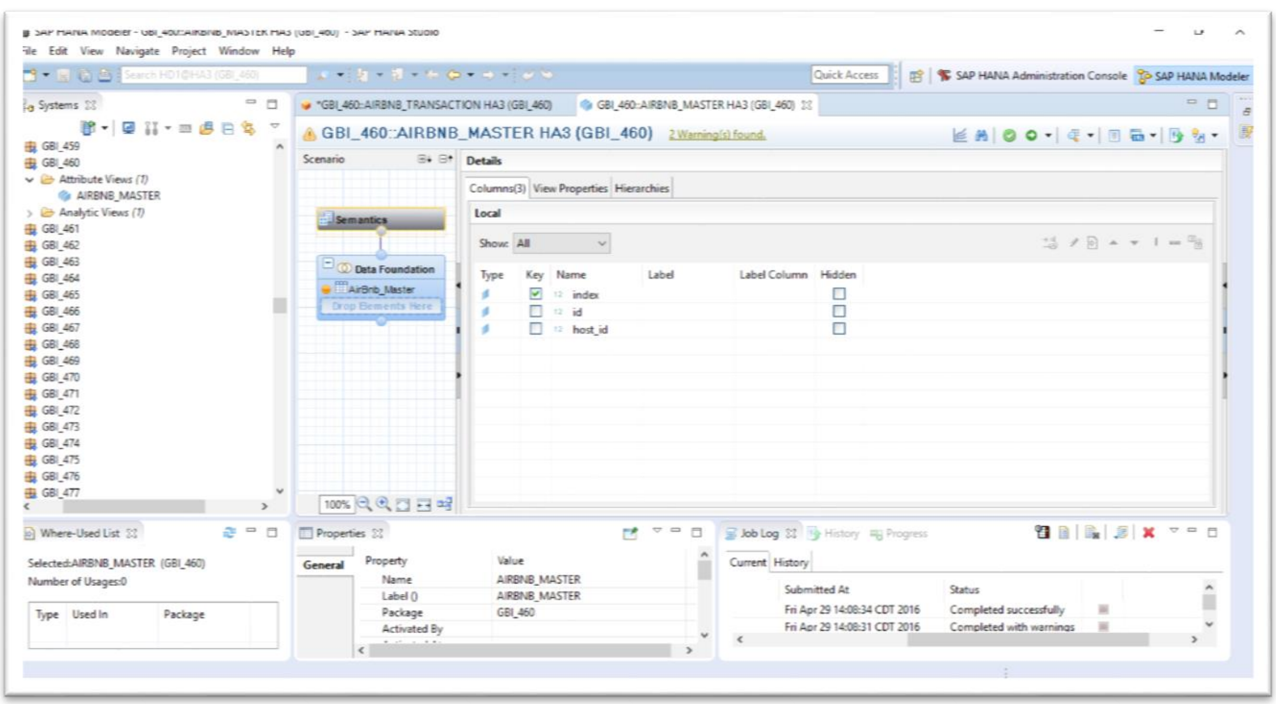
```
SELECT TOP 1000 * FROM "GBI_460"."AirBnb_Master"
```

	index	id	name	host_id	host_name	host_location	city	state	zipcode
1	1	444,018	Sweet Suite for Romantic Getaway	2,205,844	Suzanne	Los Angeles, California, United States	Los Angeles	CA	90007
2	2	1,557,529	Small Room in Central L.A. (3)	2,324,191	Elisabeth	Los Angeles, California, United States	Los Angeles	CA	90007
3	3	6,504,741	Private Room with Queen-Bed	2,324,191	Elisabeth	Los Angeles, California, United States	Private Ro...	CA	90007
4	4	8,989,426	the garden room under the sky	4,072,379	Danielle	Los Angeles, California, United States	Los Angeles	CA	90018
5	5	4,839,375	Private Room in Beautiful 2BR H...	24,905,...	Kelley	Los Angeles, California, United States	Los Angeles	CA	90018
6	6	8,617,690	3bd Victorian loft dtwn LA | USC	16,467,...	Jomar	Los Angeles, California, United States	Los Angeles	CA	90007
7	7	5,393,269	Experience Rastaman Vibrations ...	21,526,...	Richard	US	Los Angeles	CA	90018
8	8	8,690,472	USC附近	45,628,...	Chengxing	US	Los Angeles	CA	90007
9	9	7,973,248	Private Room near Downtown &...	42,074,...	Debbie	Los Angeles, California, United States	Los Angeles	CA	90007
10	10	7,692,777	Small Private Room in Central L.A.	10,871,...	Ben	Los Angeles, California, United States	Los Angeles	CA	90007
11	11	9,013,677	Large Private Room with 2 Beds	46,666,...	Liz	Los Angeles, California, United States	Los Angeles	CA	90007
12	12	9,063,047	Private Room near Downtown	42,074,...	Debbie	Los Angeles, California, United States	Los Angeles	CA	90007
13	13	5,851,860	URBAN Oasis by DTLA/USC. RM 1	1,545,056	L	Los Angeles, California, United States	Los Angeles	CA	90018
14	14	9,048,768	New Modern Studio Near Down...	13,108,...	Linda An...	Los Angeles, California, United States	Los Angeles	CA	90018
15	15	3,052,357	Modern 2BR Apt near Downtow...	13,108,...	Linda An...	Los Angeles, California, United States	Los Angeles	CA	90018
16	16	2,516,036	"holiday special" LA APARTMENT	12,879,...	Isaiah	Los Angeles, California, United States	Los Angeles	CA	90007
17	17	7,826,236	Apartment in Central L.A.	37,935,...	Eli & Ste...	Los Angeles, California, United States	Los Angeles	CA	90007
18	18	8,009,373	Creative Art/Meeting Space	42,071,...	Daniel	Los Angeles, California, United States	Los Angeles	CA	90018
19	19	8,152,058	Charming 3bd Victorian loft dtw...	16,467,...	Jomar	Los Angeles, California, United States	Los Angeles	CA	90007
20	20	9,167,855	Quarto compartilhado Los Ange...	47,714,...	Bruna	Los Angeles, California, United States	Los Angeles	CA	90007
21	21	8,924,216	Small Private Room in LA.	46,666,...	Liz	Los Angeles, California, United States	Los Angeles	CA	90007
22	22	1,051,837	Los Angeles Rooms For Rent	2,324,191	Elisabeth	Los Angeles, California, United States	Los Angeles	CA	90007

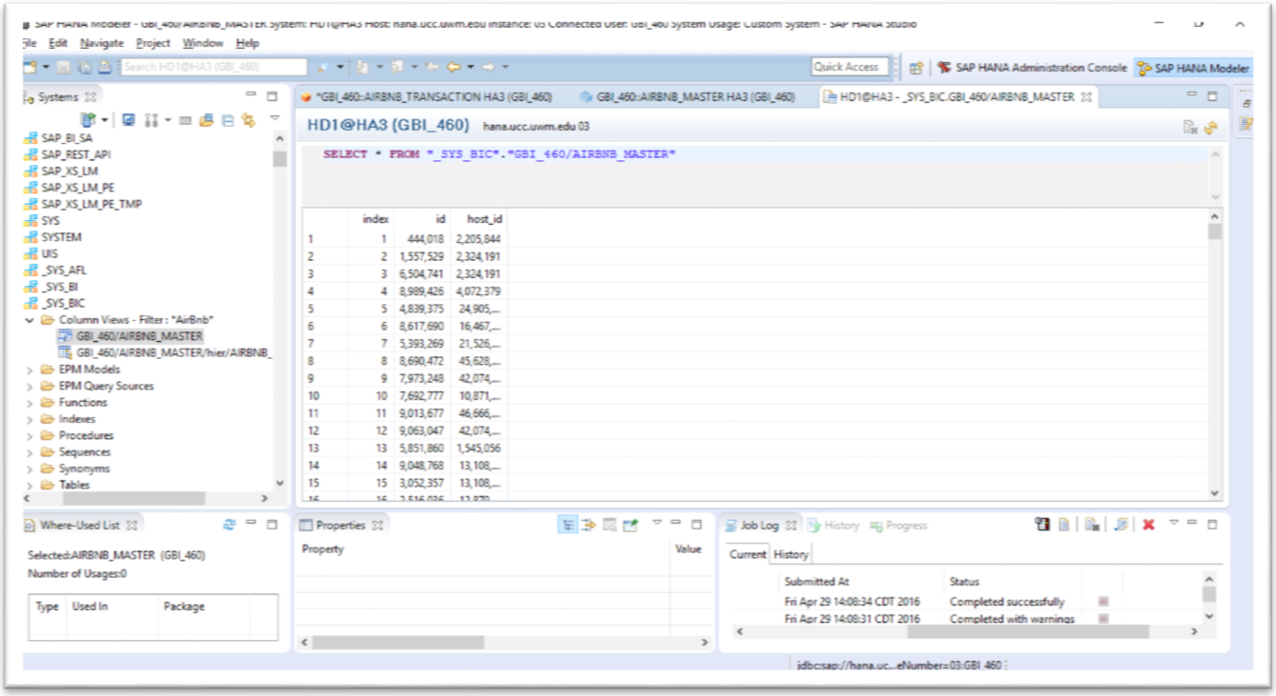
- **Data Foundation**



- **Semantics**



- **Column View**



Predicative Hypothesis:

- Airbnb is a website for people to list, find, and rent lodging. Customers choose a place to live based on criteria like location, availability, price, reviews, amenities, quality of living and places to visit.
- User experience can be enhanced by features like ease of booking, host response, renter satisfaction and mutual verification.
- Quality of the business is affected by factors like reach, demand and historical experience

Predictive Theorems:

- To classify existing listings based on various user preferences and use predictive analysis to classify new hosts into these categories for future bookings
- To suggest dynamic pricing for each listing by using the current availability and listing history.
- To create a word-cloud from listing descriptions from which users can pull data to find a perfect match for a dream vacation

Results and Process Presentation

Visualization of the Data set:

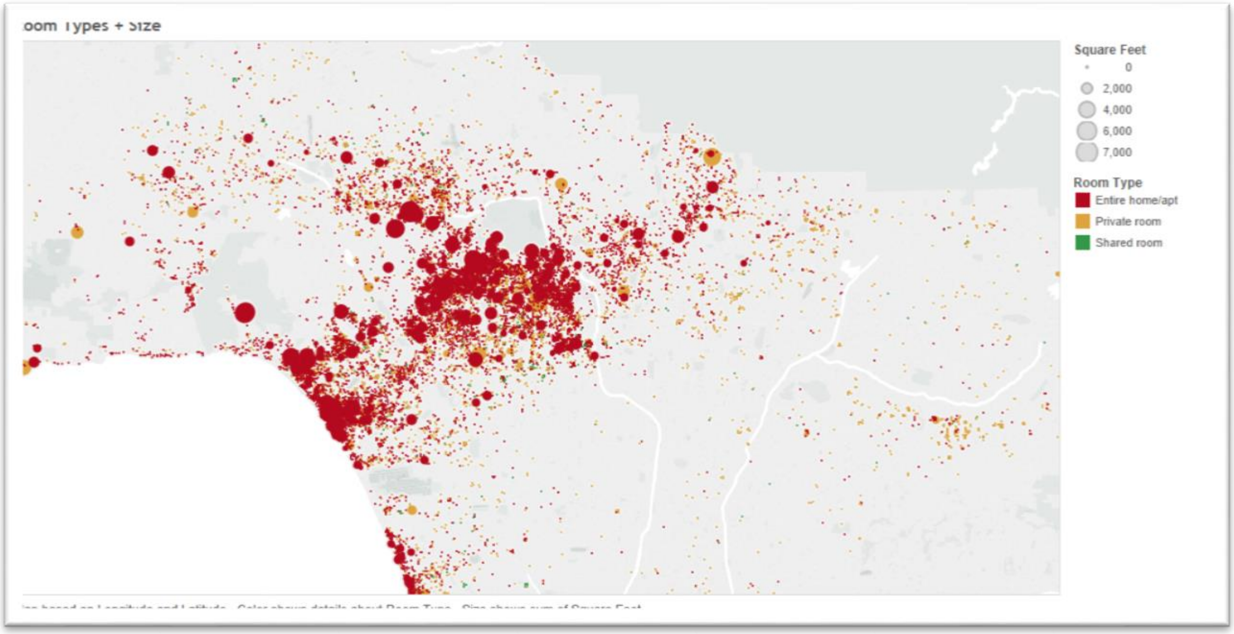
1) Top Listings

'op Listings

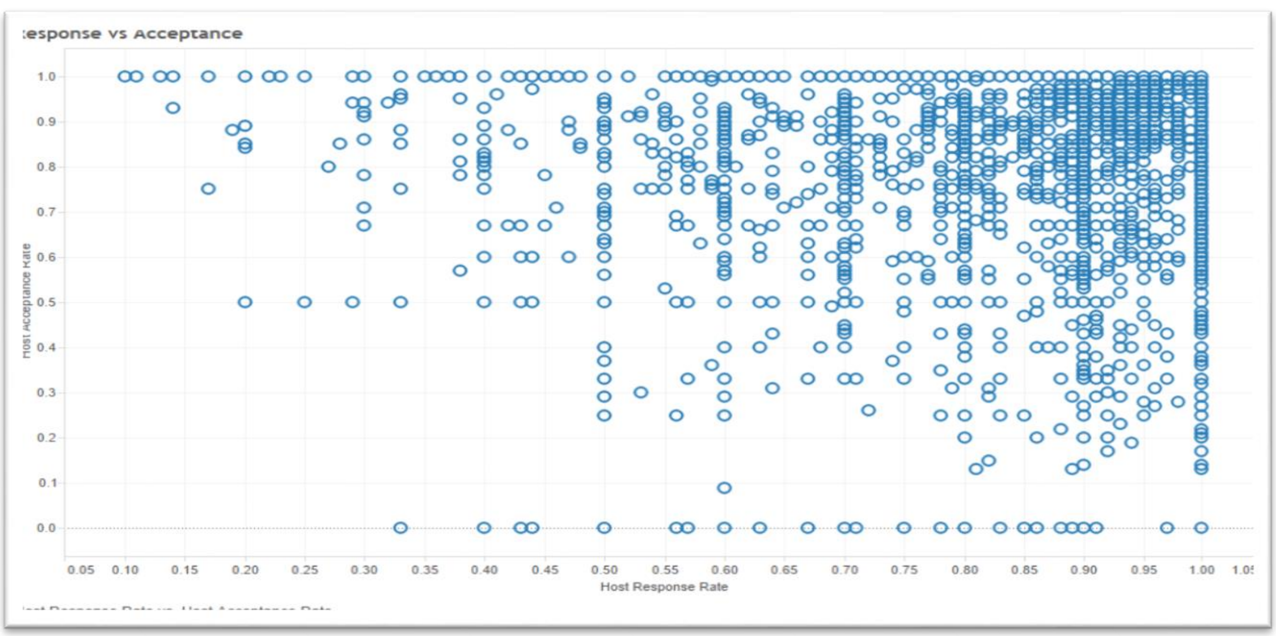
	Host Name									
Property Type	Michael	Shuki	David	Alex	Sarah	John	Chris	James	Lisa	Jennifer
Apartment	112		67	86	54	56	48	49	36	35
Bed & Breakfast	1	1				1	2			
Boat						2	1			
Bungalow	1		1	2	2	2	1		1	1
Cabin	1		1			1				1
Camper/RV	2	2						1	1	
Castle								2		3
Condominium	3		2			3	1		1	4
Dorm	1									
Earth House										1
House	59	159	76	33	57	40	32	35	56	46
Hut								1		
Loft	4	4	4	4	4	5	5	7		3
Other				1	1	1		1		
Townhouse	6		5	1		4	5	1		
Villa		1	1	1			2			

:ount of Host Total Listings Count broken down by Host Name vs. Property Type. The view is filtered on Host Name, which keeps 0 of 5,443 members.

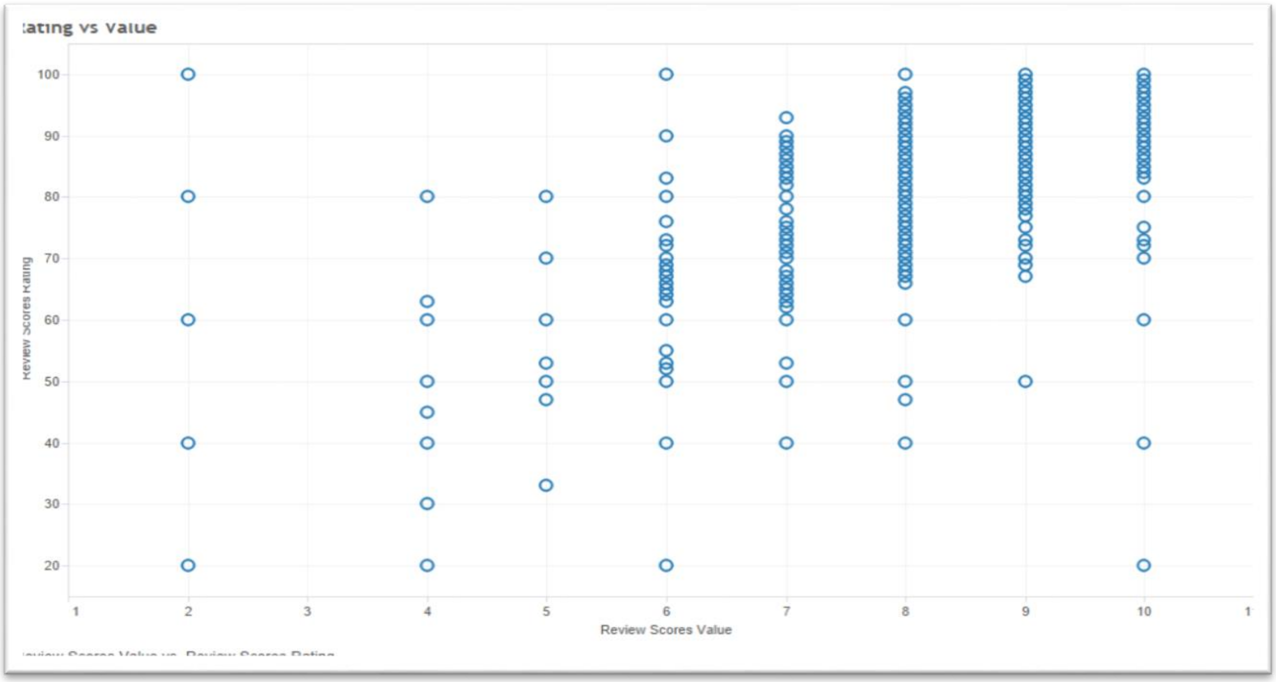
2) Room Types with their Sqft area.



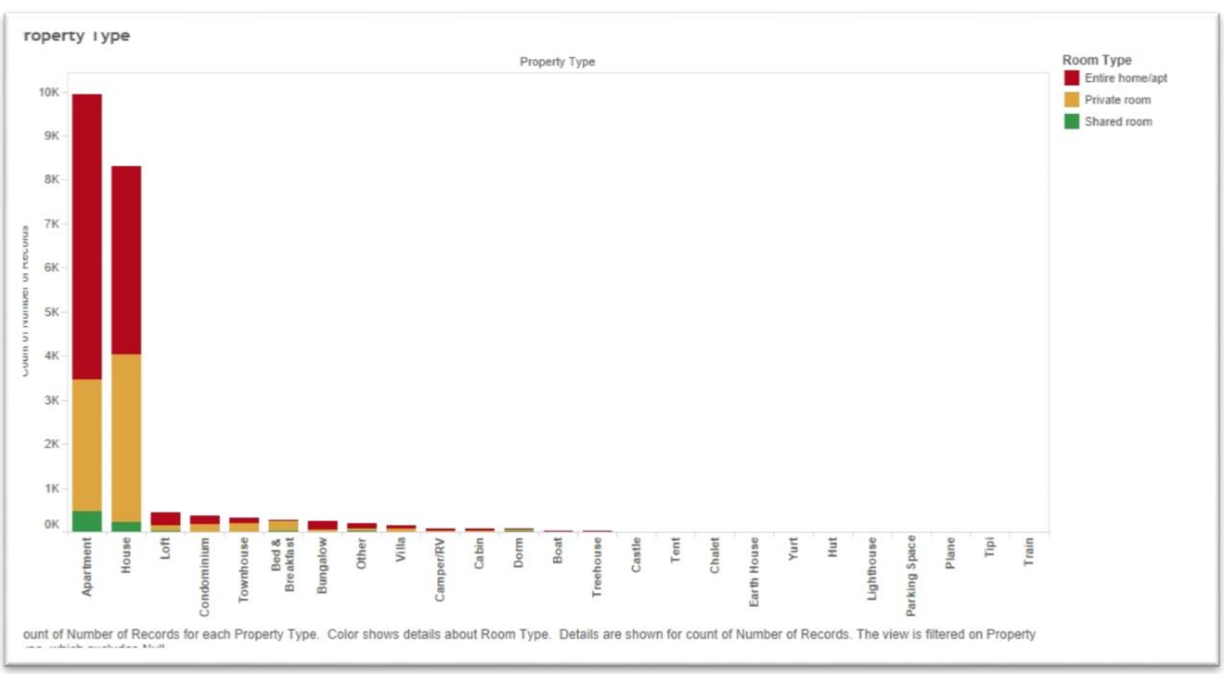
3) Response vs Acceptance



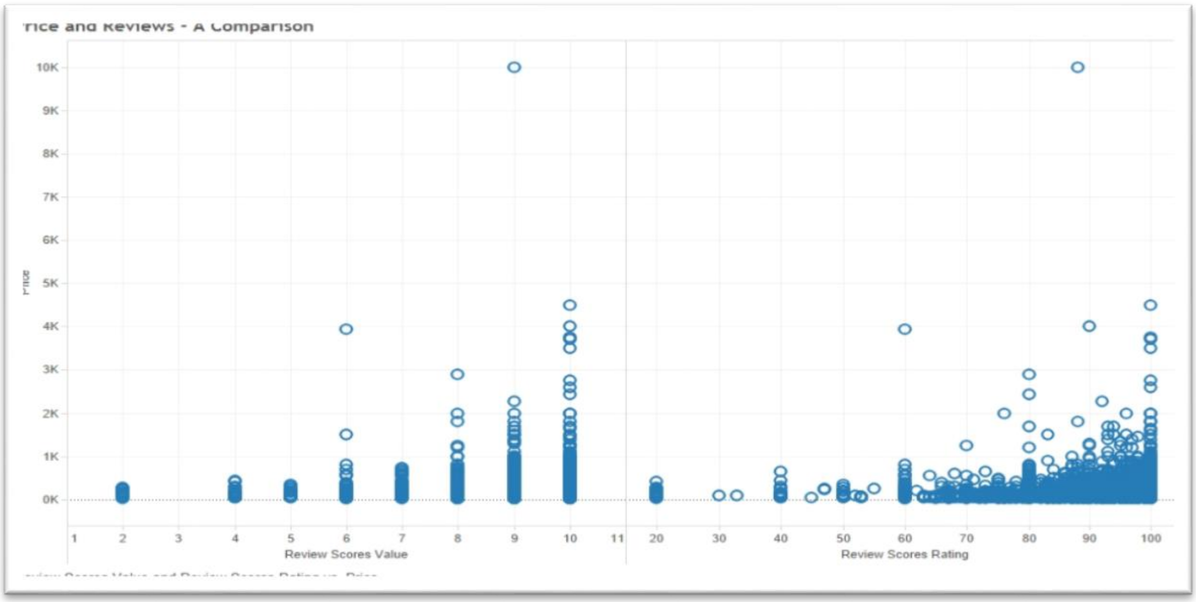
4) Listing's Rating vs Listing'sValue



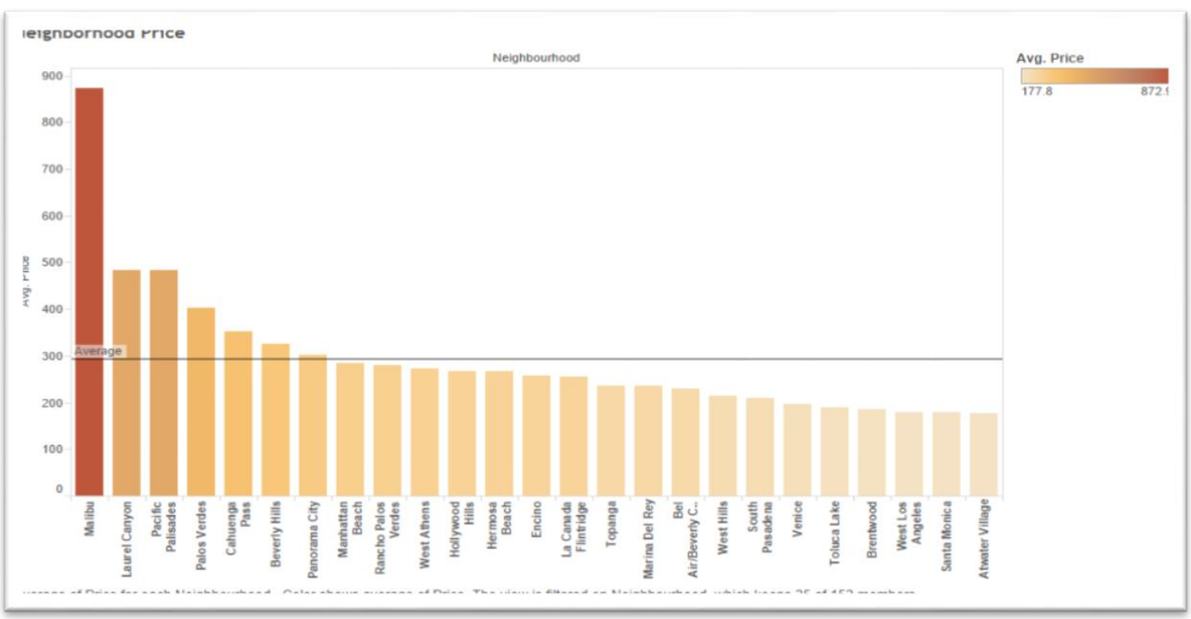
5) Count of Property type



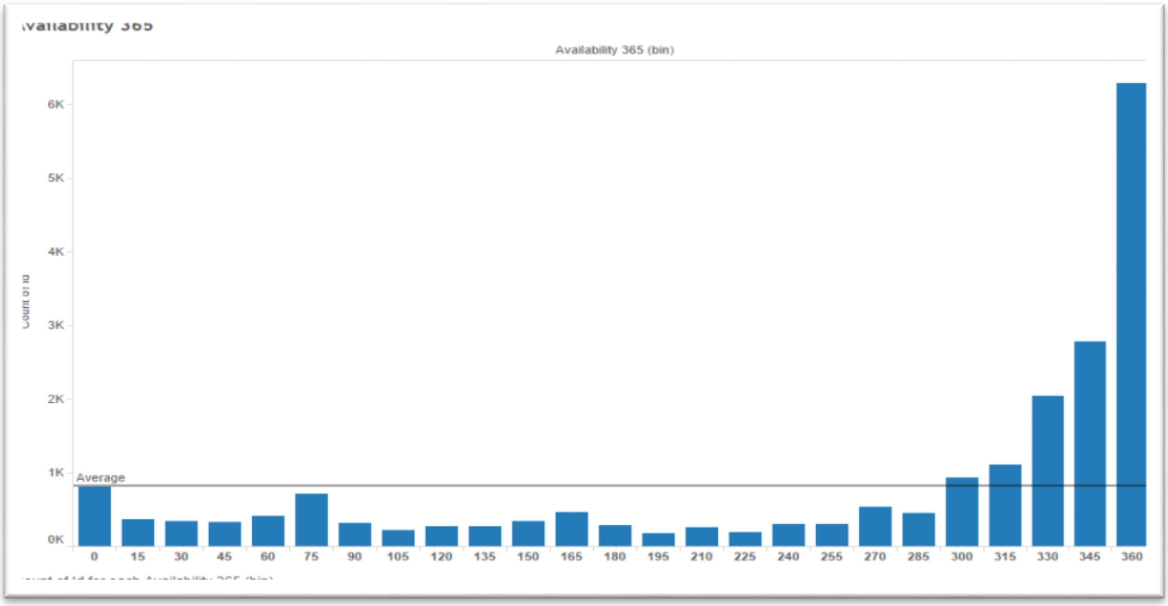
6) Price and Reviews – A Comparison



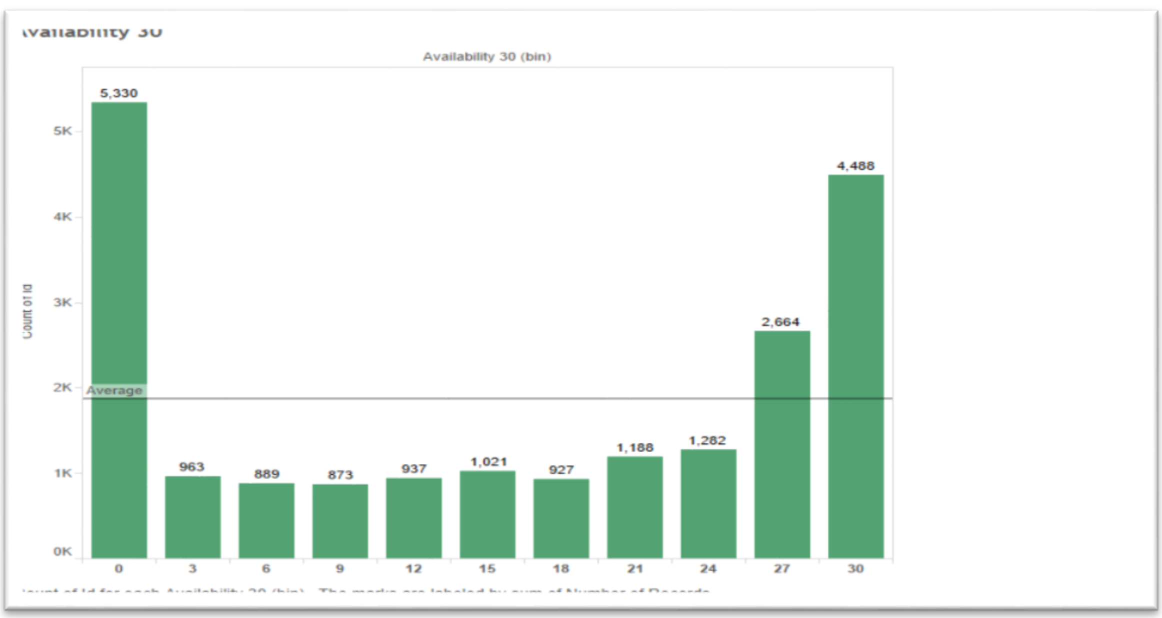
7) Variation in Price by Neighborhood.



8) Availability For 365 days



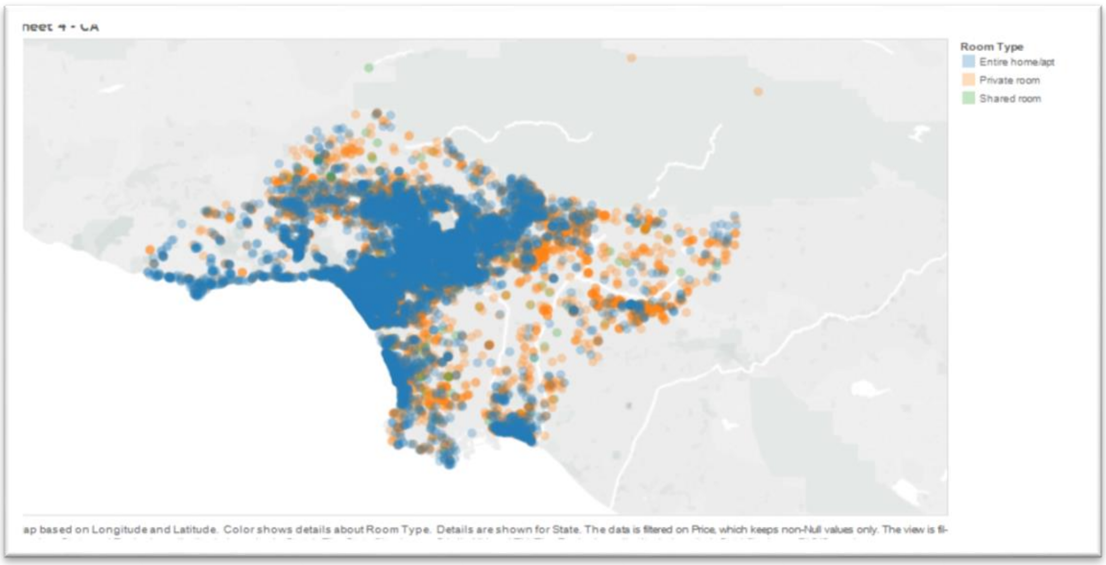
9) Availability For 30 Days



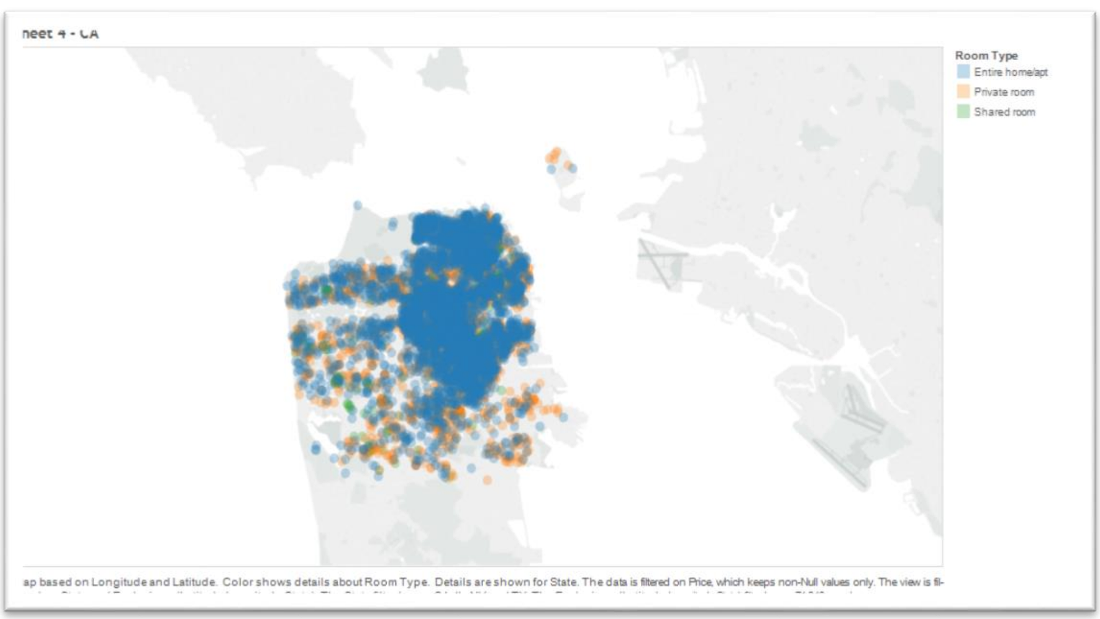
Analysis of Theorem I:

To classify existing listings based on various user preferences and use predictive analysis to classify new hosts into these categories for future bookings.

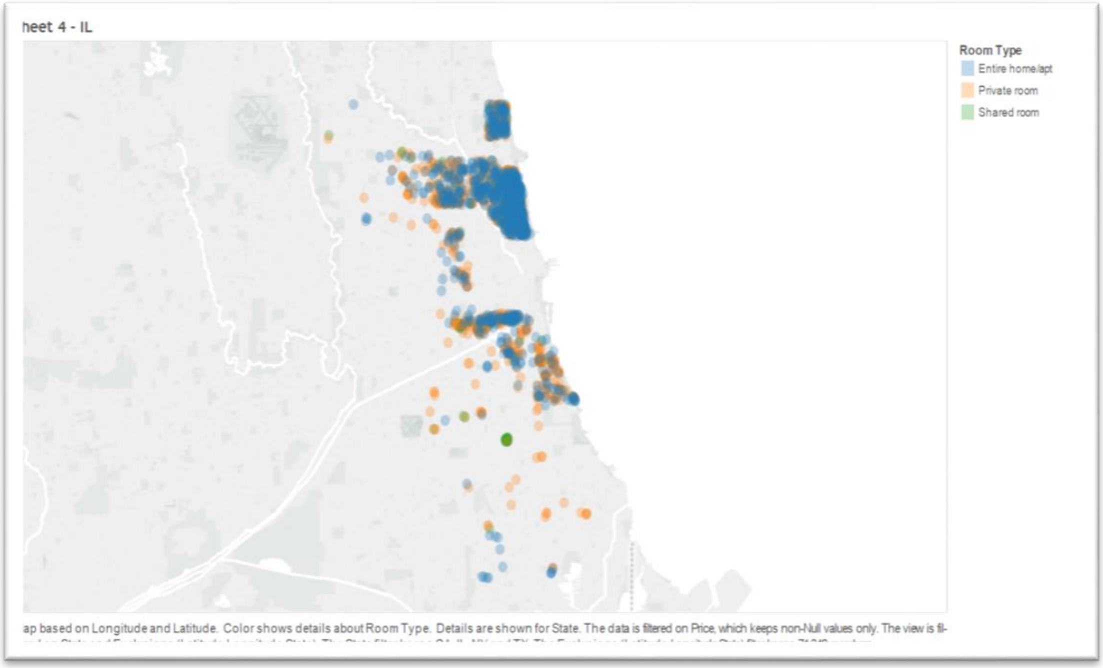
- **Los Angeles**



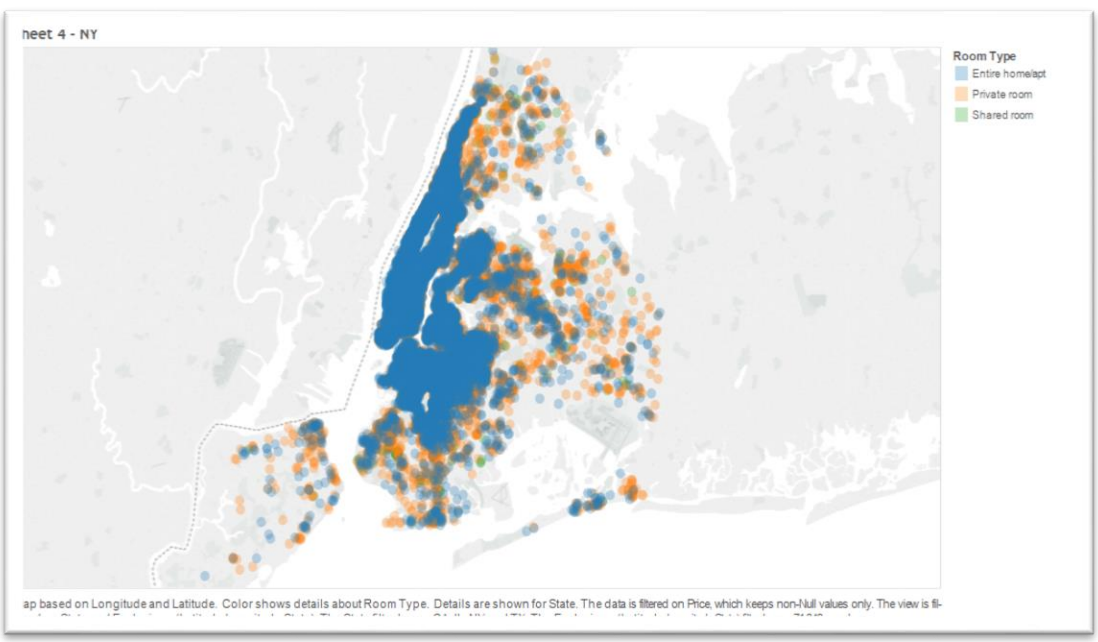
- **San Francisco**



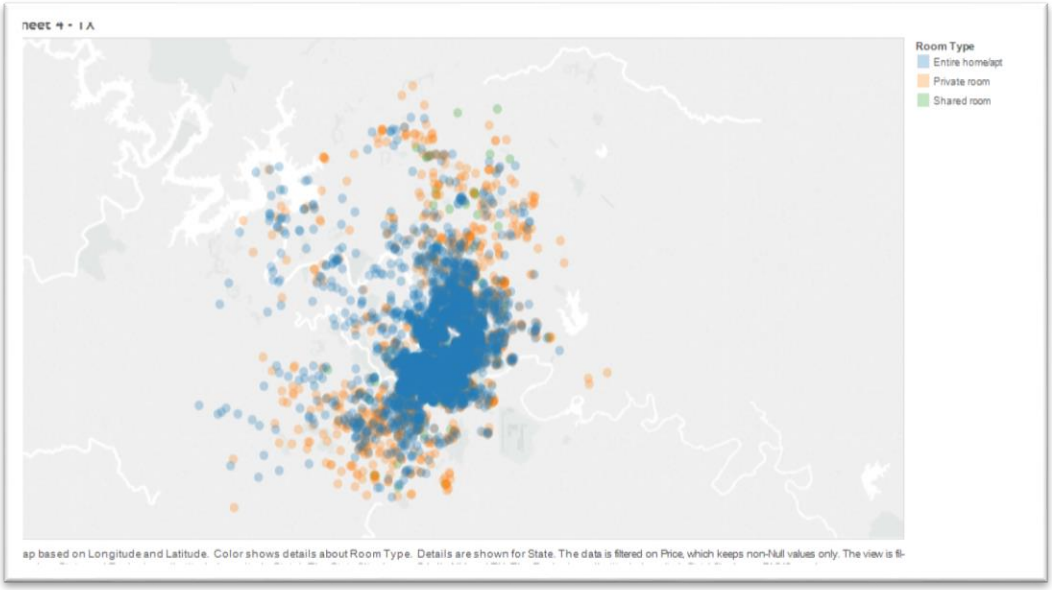
- **Chicago**



New York City



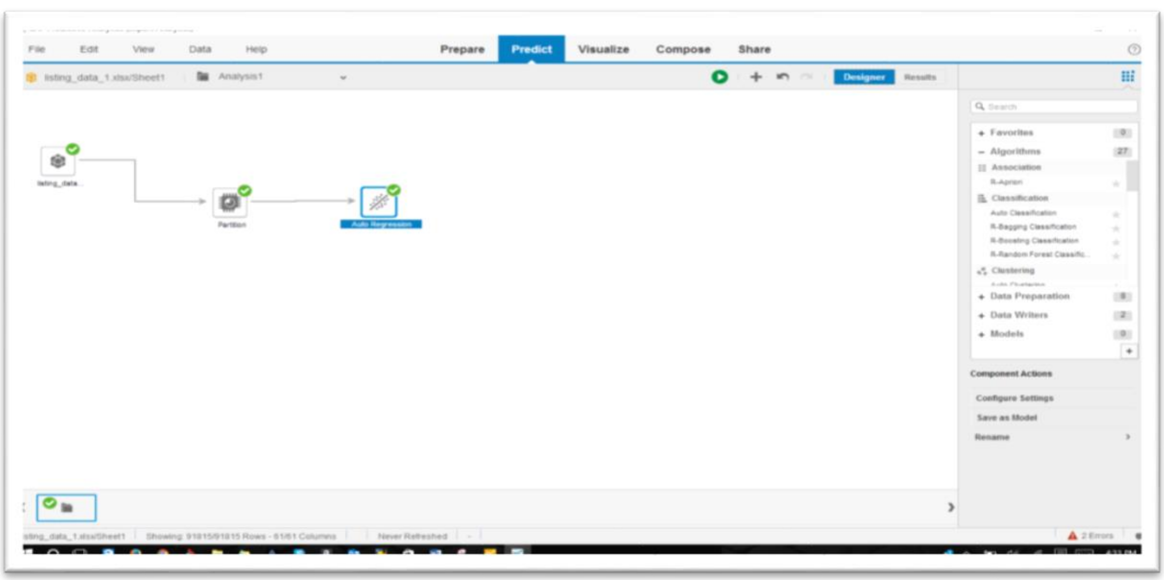
- **Austin**



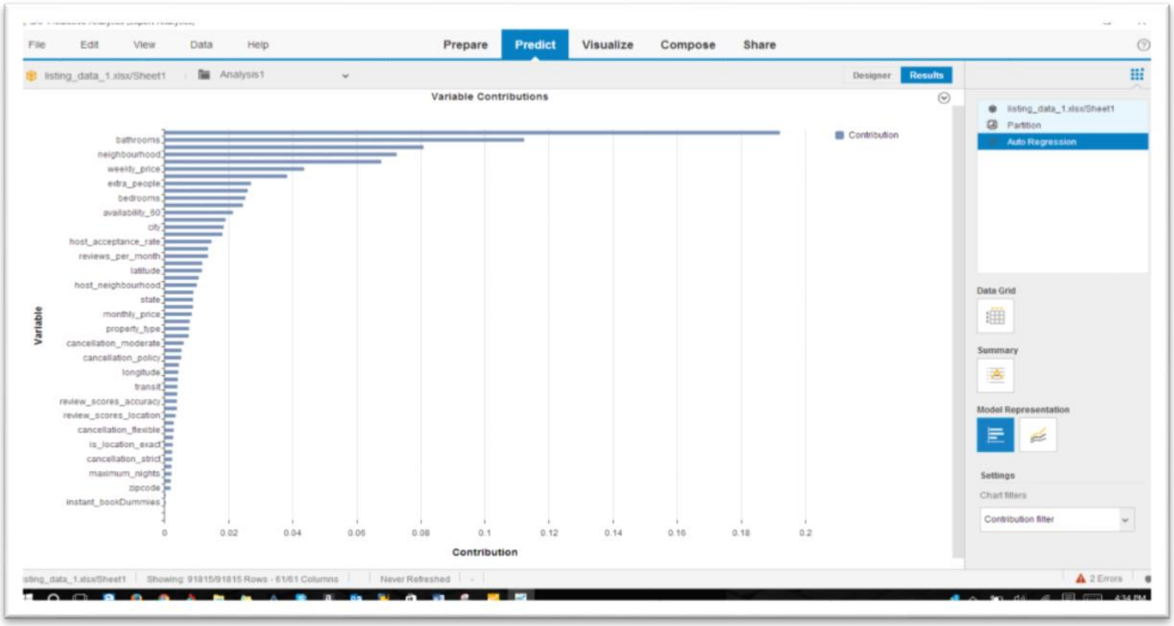
Analysis of Theorem II:

To suggest dynamic pricing for each listing by using the current availability and listing history.

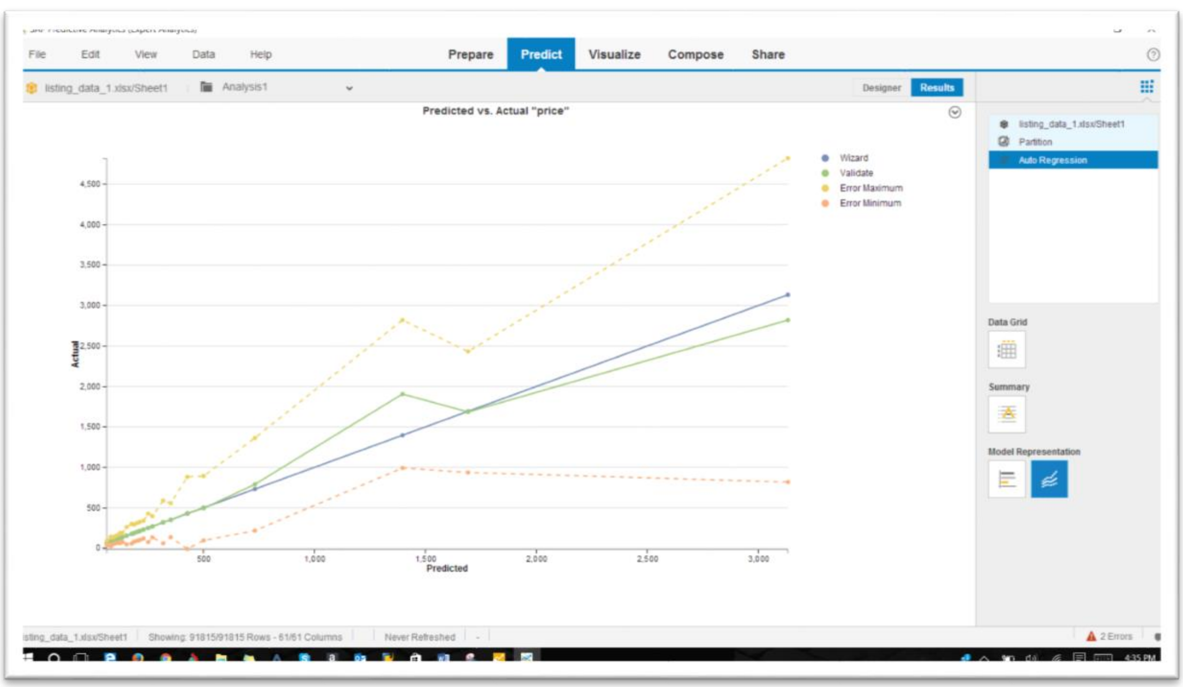
- Ran automatic regression to check for the best variable that has an impact on target variable. Our target variable here is price.



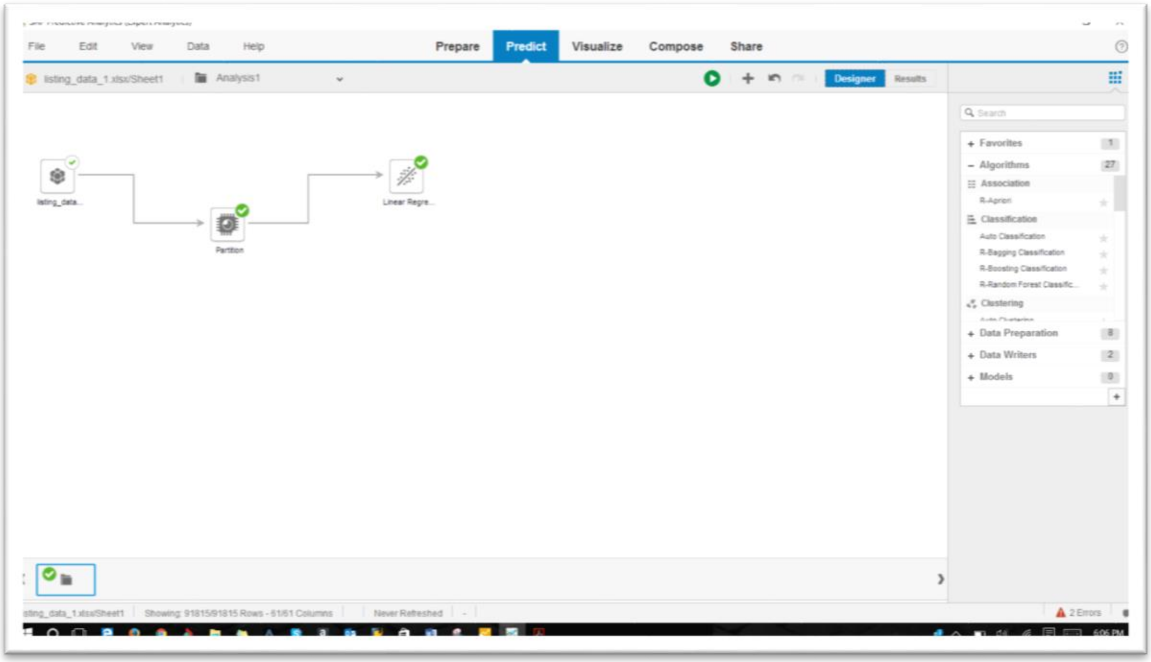
- Variable impact on Price



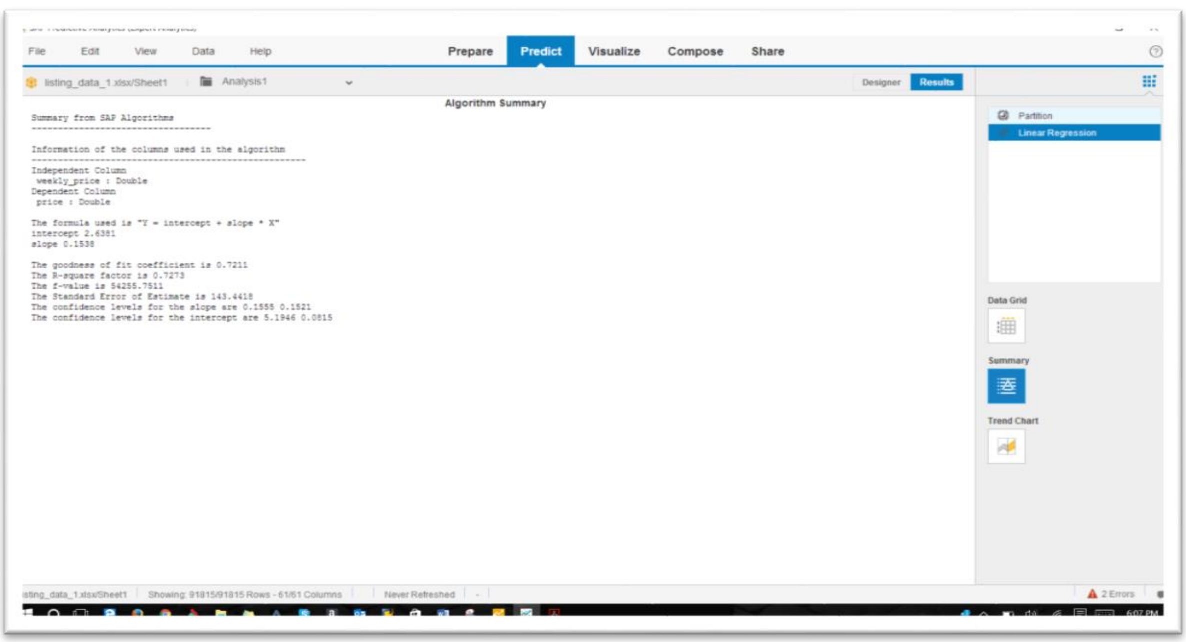
- Model Representation



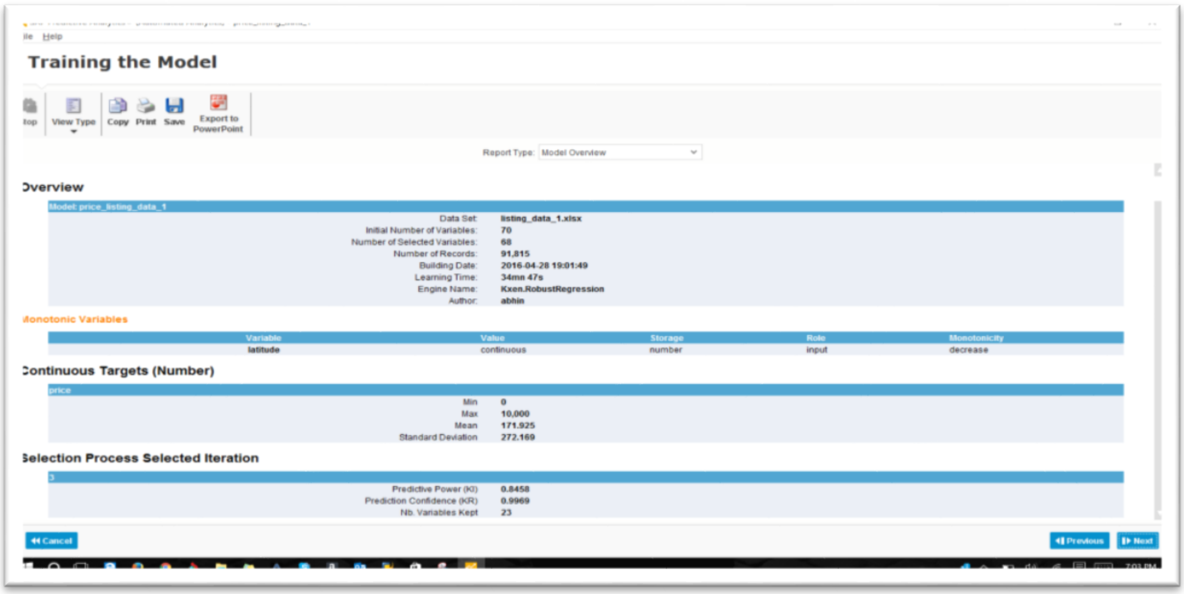
- Ran a linear regression on dataset after checking the impact of variables.



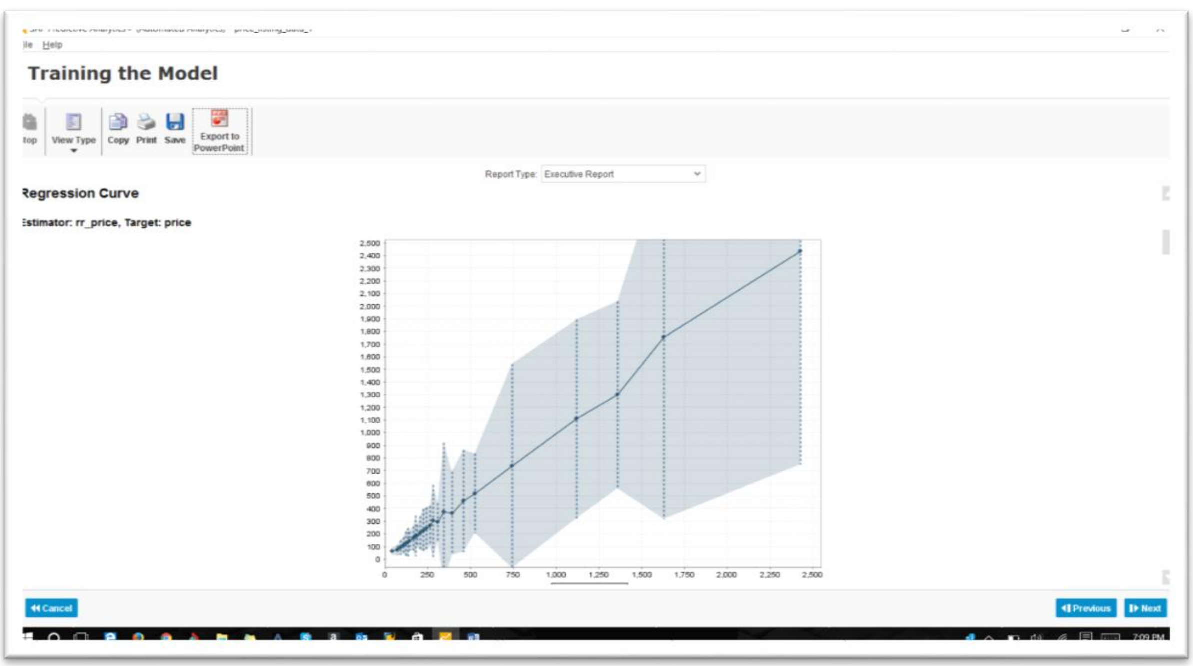
- Target variable is price and independent variable is weekly price. R square value = 0.72 which indicates that 72% of the target variable is explained by the input variable.



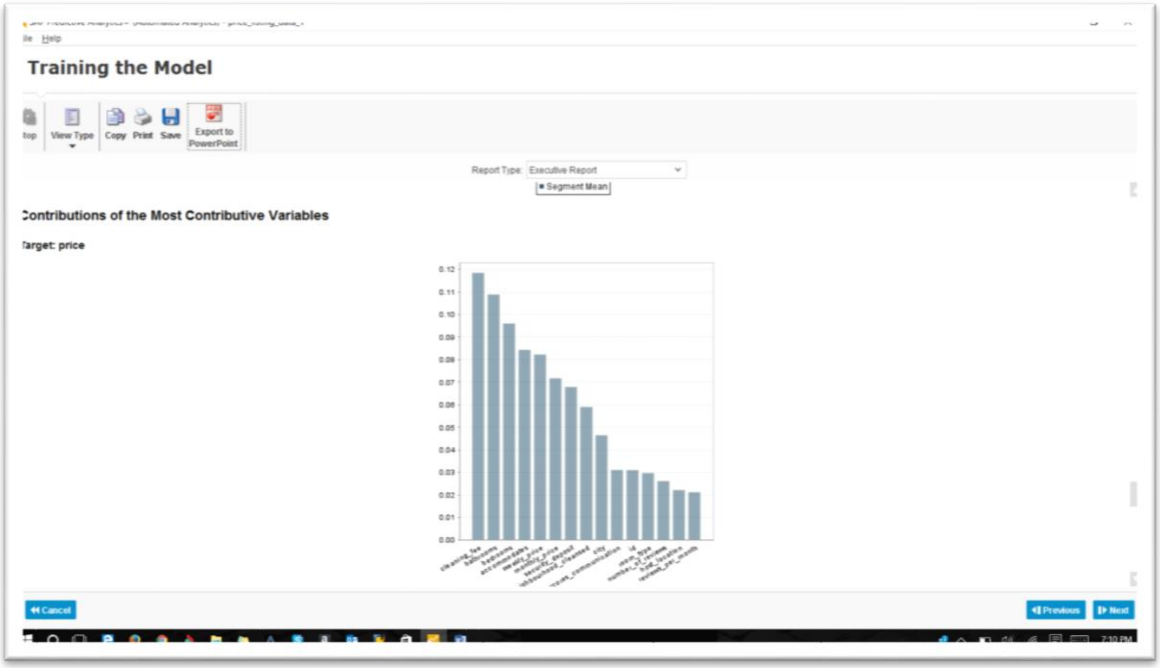
Automated tool Run classification and regression



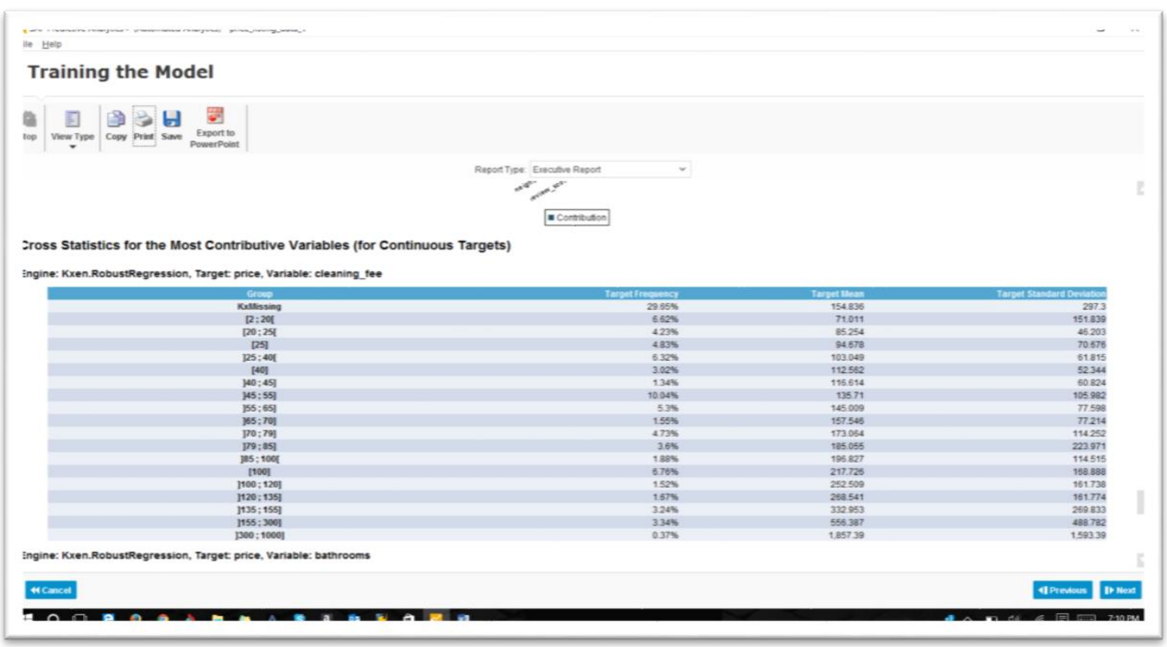
- Regression curve obtained



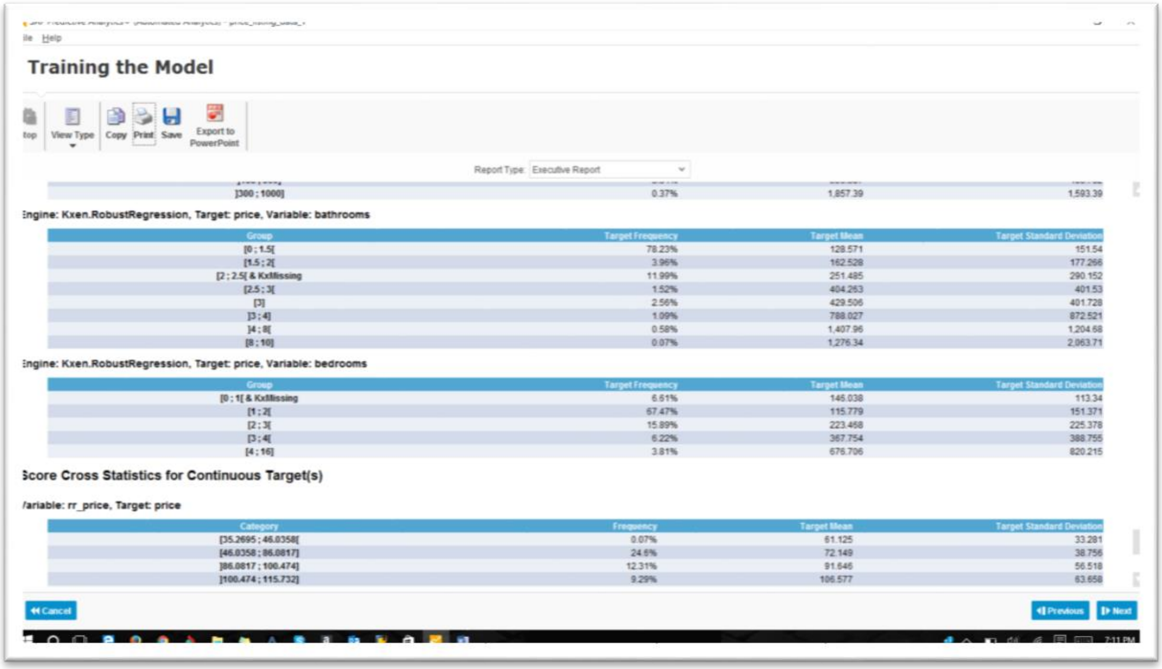
- Contribution of variables



- Training the model



Group	Target Frequency	Target Mean	Target Standard Deviation
KxMissing	29.65%	154.836	297.3
[2 ; 20[	6.62%	71.011	151.839
[20 ; 25[	4.23%	85.254	46.203
[25]	4.83%	94.678	70.676
]25 ; 40[	6.32%	103.049	61.815
[40]	3.02%	112.562	52.344
]40 ; 45]	1.34%	116.614	60.824
]45 ; 55]	10.04%	135.71	105.982
]55 ; 65]	5.3%	145.009	77.598
]65 ; 70]	1.55%	157.546	77.214
]70 ; 79]	4.73%	173.064	114.252
]79 ; 85]	3.6%	185.055	223.971
]85 ; 100[	1.88%	196.827	114.515
[100]	6.76%	217.726	168.888
]100 ; 120]	1.52%	252.509	161.738
]120 ; 135]	1.67%	268.541	161.774
]135 ; 155]	3.24%	332.953	269.833
]155 ; 300]	3.34%	556.387	488.782
]300 ; 1000]	0.37%	1,857.39	1,593.39



ile Help

Training the Model

top View Type Copy Print Save Export to PowerPoint

Report Type: Executive Report

]300 ; 1000]	0.37%	1,857.39	1,593.39

Engine: Kxen.RobustRegression, Target: price, Variable: bathrooms

Group	Target Frequency	Target Mean	Target Standard Deviation
[0 ; 1.5[	78.23%	128.571	151.54
[1.5 ; 2[	3.96%	162.528	177.266
[2 ; 2.5[& KxMissing	11.99%	251.485	290.152
[2.5 ; 3[	1.52%	404.263	401.53
[3]	2.56%	429.506	401.728
]3 ; 4]	1.09%	788.027	872.521
]4 ; 8[	0.58%	1,407.96	1,204.68
[8 ; 10]	0.07%	1,276.34	2,063.71

Engine: Kxen.RobustRegression, Target: price, Variable: bedrooms

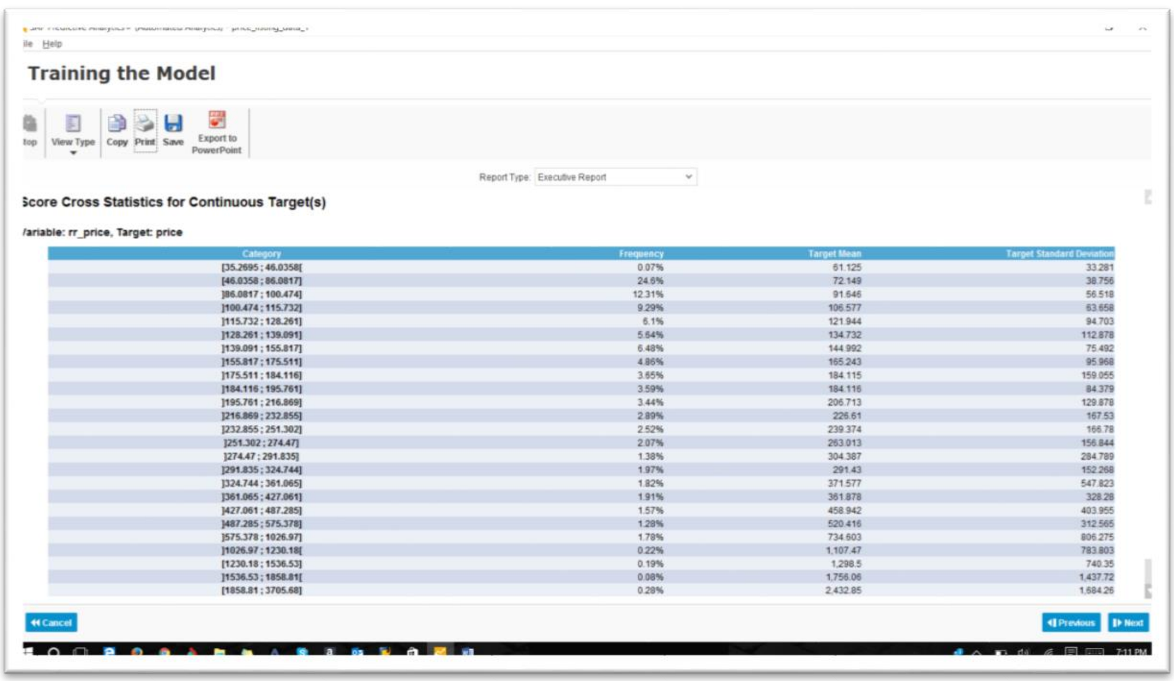
Group	Target Frequency	Target Mean	Target Standard Deviation
[0 ; 1[& KxMissing	6.61%	146.038	113.34
[1 ; 2[	67.47%	115.779	151.371
[2 ; 3[	15.89%	223.468	225.378
[3 ; 4[	6.22%	367.754	388.755
[4 ; 16]	3.81%	676.706	820.215

Score Cross Statistics for Continuous Target(s)

Variable: rr_price, Target: price

Category	Frequency	Target Mean	Target Standard Deviation
[35.2695 ; 46.0358[	0.07%	61.125	33.281
[46.0358 ; 86.0817]	24.6%	72.149	38.756
]86.0817 ; 100.474]	12.31%	91.646	56.518
]100.474 ; 115.732]	9.29%	106.577	63.658

Cancel Previous Next



ile Help

Training the Model

top View Type Copy Print Save Export to PowerPoint

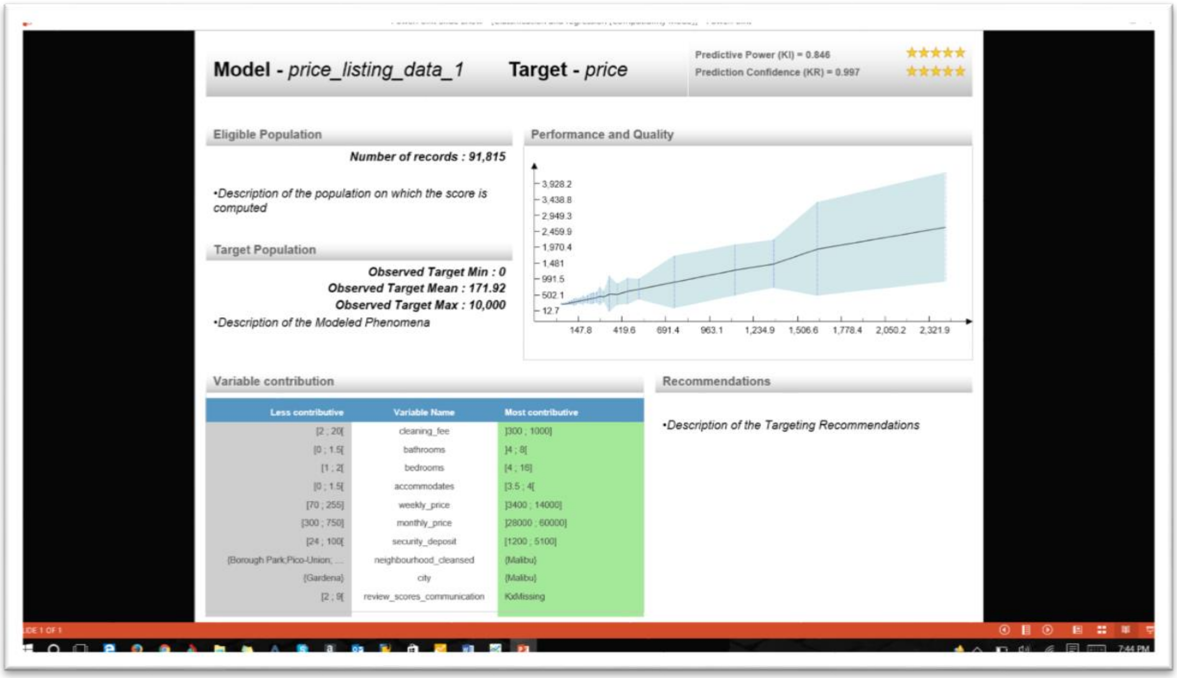
Report Type: Executive Report

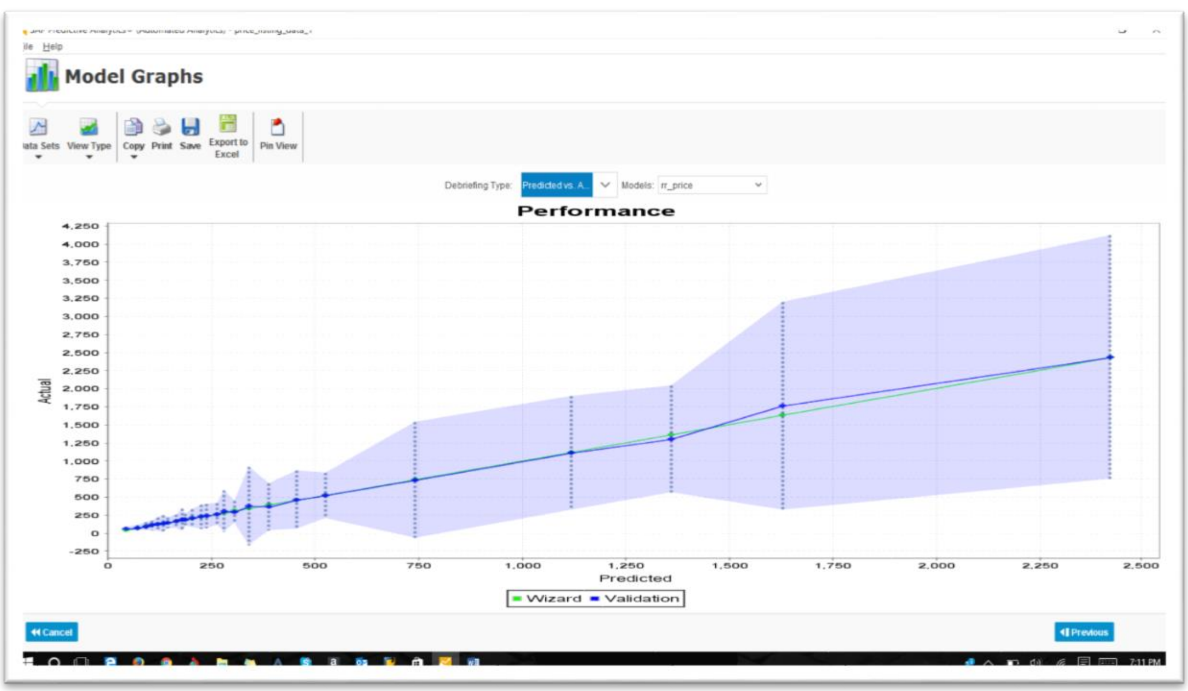
Score Cross Statistics for Continuous Target(s)

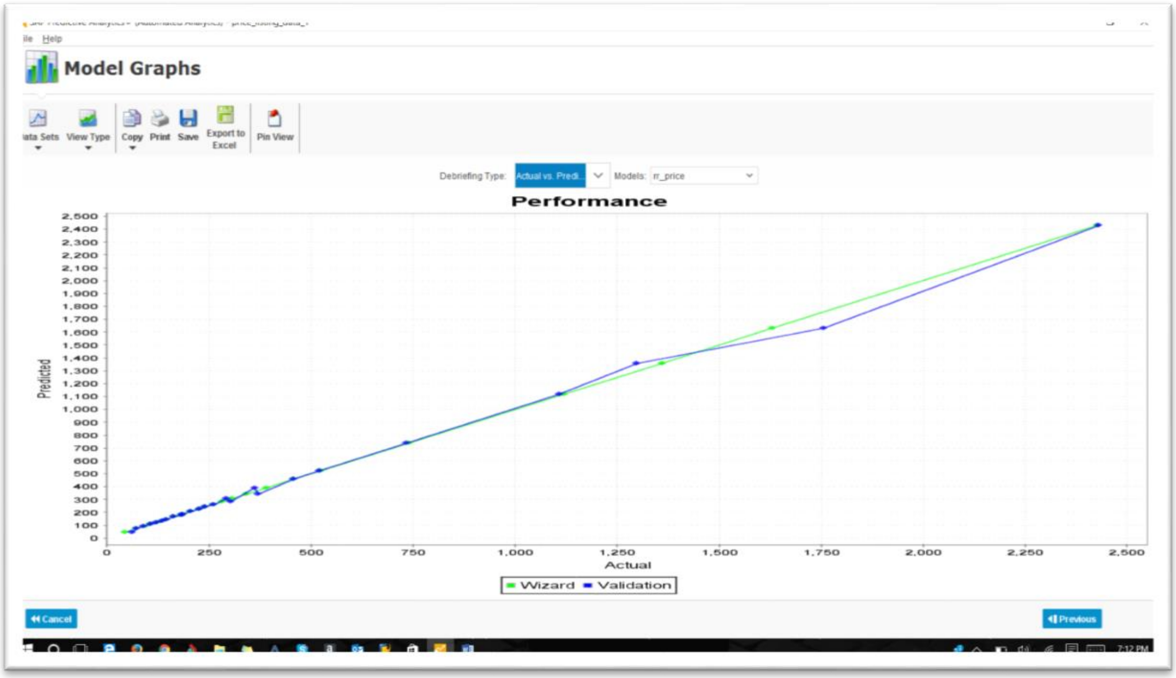
Variable: rr_price, Target: price

Category	Frequency	Target Mean	Target Standard Deviation
[35.2695 ; 46.0358[	0.07%	61.125	33.281
[46.0358 ; 86.0817]	24.6%	72.149	38.756
]86.0817 ; 100.474]	12.31%	91.646	56.518
]100.474 ; 115.732]	9.29%	106.577	63.658
]115.732 ; 128.261]	6.1%	121.944	94.703
]128.261 ; 139.091]	5.64%	134.732	112.878
]139.091 ; 155.817]	6.48%	144.992	75.492
]155.817 ; 175.511]	4.86%	165.243	95.968
]175.511 ; 184.116]	3.65%	184.115	159.055
]184.116 ; 195.761]	3.59%	184.116	84.379
]195.761 ; 216.869]	3.44%	206.713	129.878
]216.869 ; 232.855]	2.89%	226.61	167.53
]232.855 ; 251.302]	2.52%	239.374	166.78
]251.302 ; 274.47]	2.07%	263.013	156.844
]274.47 ; 291.835]	1.38%	304.387	284.789
]291.835 ; 324.744]	1.97%	291.43	152.268
]324.744 ; 361.065]	1.82%	371.577	547.823
]361.065 ; 427.061]	1.91%	361.878	328.28
]427.061 ; 487.285]	1.57%	458.942	403.955
]487.285 ; 575.378]	1.28%	520.416	312.565
]575.378 ; 1026.97]	1.78%	734.603	806.275
]1026.97 ; 1230.18[	0.22%	1,107.47	783.803
[1230.18 ; 1536.53]	0.19%	1,298.5	740.35
]1536.53 ; 1858.81[	0.08%	1,756.06	1,437.72
[1858.81 ; 3705.68]	0.28%	2,432.85	1,684.26

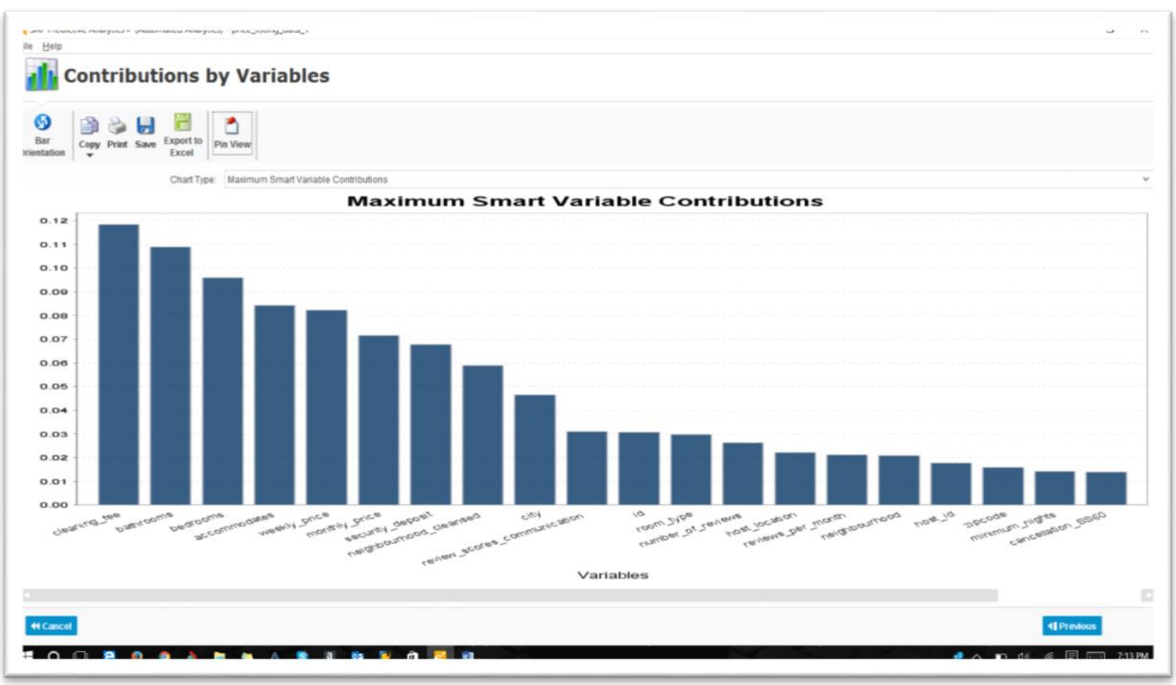
Cancel Previous Next


Model - price_listing_data_1
Target - price
Predictive Power (KI) = 0.846
Prediction Confidence (KR) = 0.997
Eligible Population
Number of records : 91,815
•Description of the population on which the score is computed
Target Population
Observed Target Min : 0
Observed Target Mean : 171.92
Observed Target Max : 10,000
•Description of the Modeled Phenomena
Performance and Quality
Variable contribution
Less contributive
Variable Name
Most contributive
[2 ; 20[cleaning_fee [300 ; 1000]
[0 ; 1.5[bathrooms]4 ; 8]
[1 ; 2[bedrooms [4 ; 10]
[0 ; 1.5[accommodates [3.5 ; 4[
[70 ; 255] weekly_price]3400 ; 14000]
[300 ; 750] monthly_price]28000 ; 60000]
[24 ; 100[security_deposit [1200 ; 5100]
{Borough Park,Pico-Union; ... neighbourhood_cleansed {Malibu}
{Gardena} city {Malibu}
[2 ; 9[review_scores_communication KxMissing
Recommendations
•Description of the Targeting Recommendations

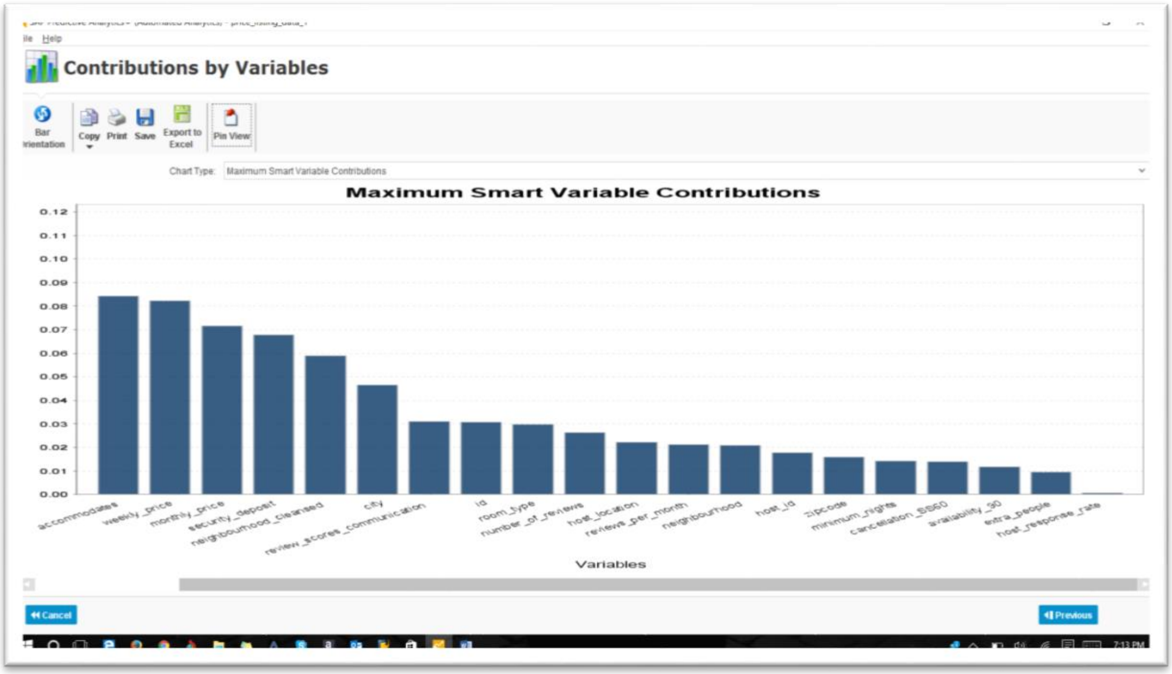

Model Graphs
Copy
Print
Save
Export to Excel
Pin View
Debriefing Type:
Models:
Performance
Actual
Predicted
Wizard
Validation
Cancel
Previous



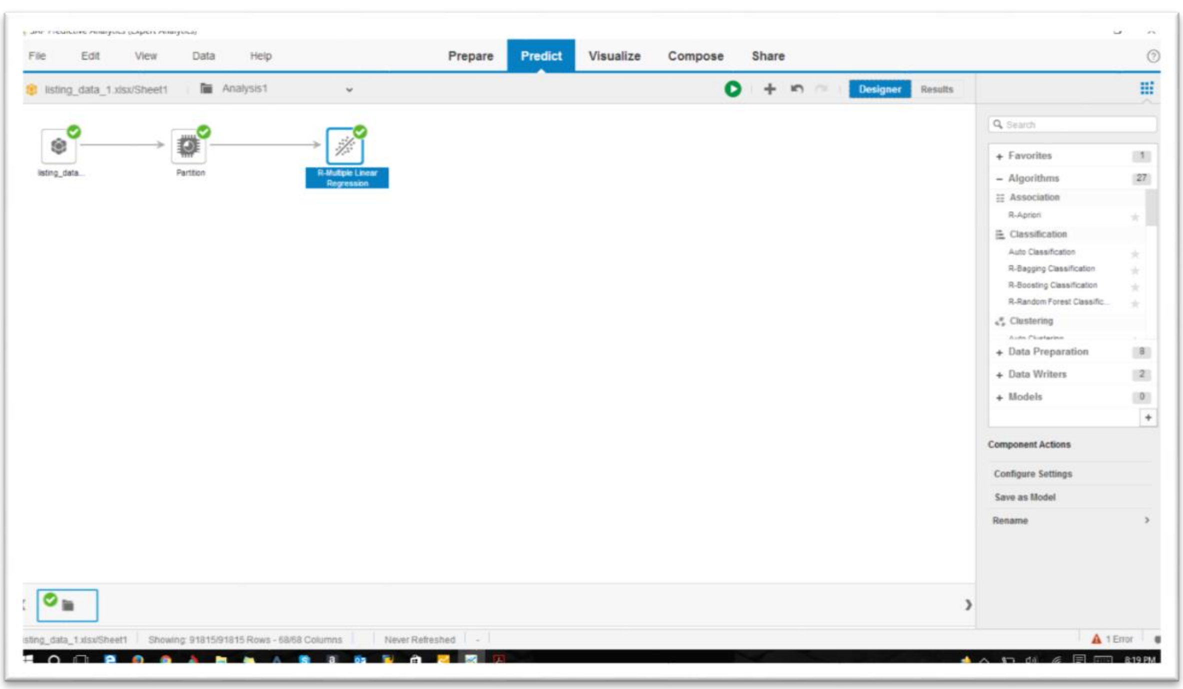
Model Graphs
Copy
Print
Save
Export to Excel
Pin View
Debriefing Type:
Models:
Performance
Predicted
Actual
Wizard
Validation
Cancel
Previous

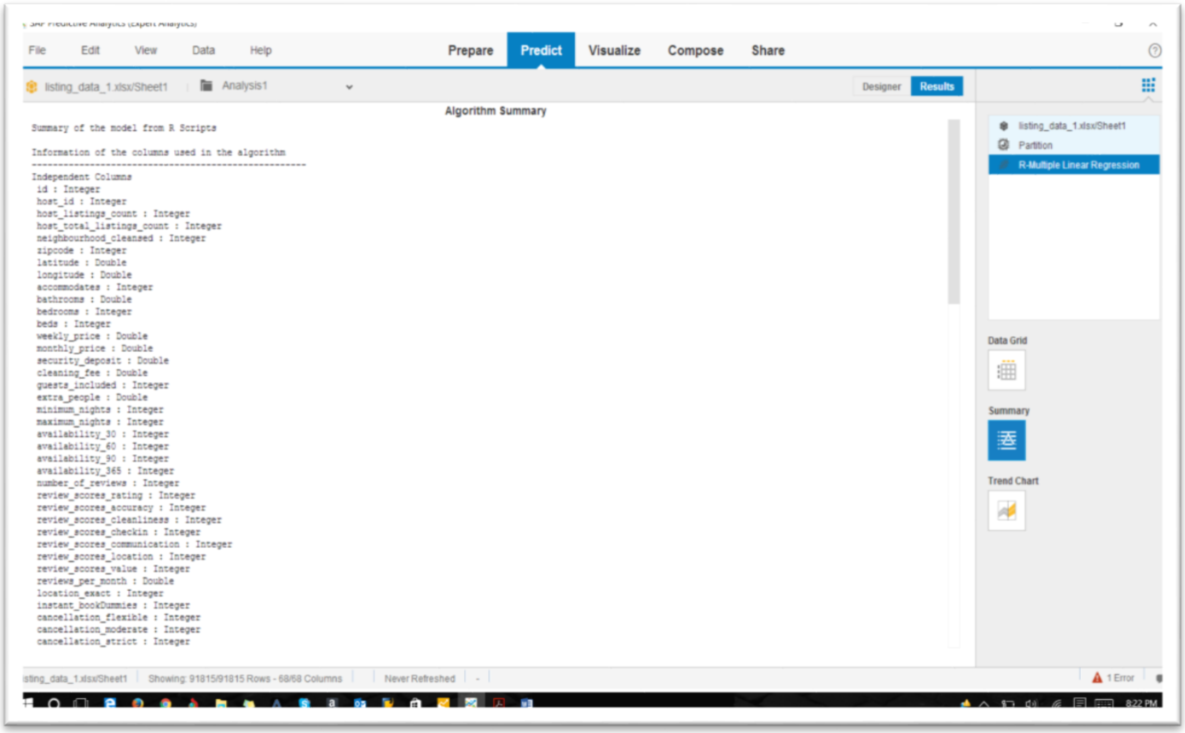


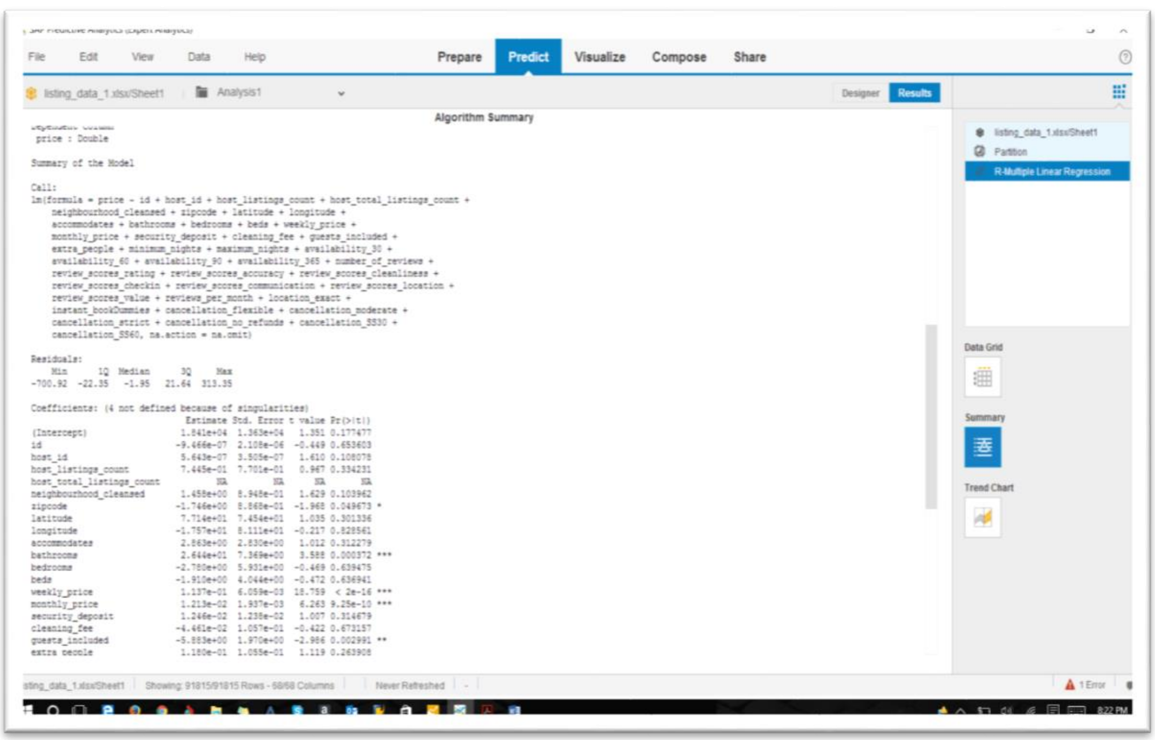
Contributions by Variables
Bar Orientation
Copy
Print
Save
Export to Excel
Pin View
Chart Type:
Maximum Smart Variable Contributions
Maximum Smart Variable Contributions
cleaning_fee
bathrooms
bedrooms
accommodates
weekly_price
monthly_price
security_deposit
neighbourhood_cleansed
review_scores_communication
city
id
room_type
number_of_reviews
host_location
reviews_per_month
neighbourhood
host_id
zipcode
minimum_nights
Variables
Cancel
Previous

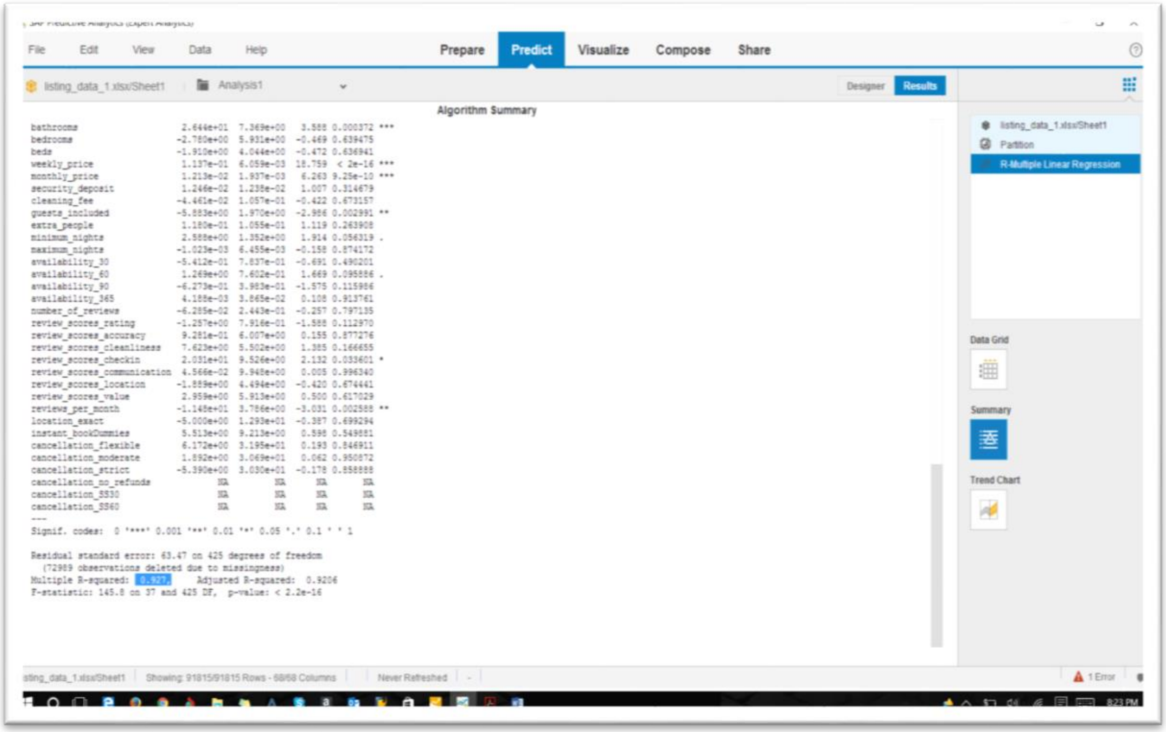


Multiple linear regression.




File
Edit
View
Data
Help
Prepare
Predict
Visualize
Compose
Share
listing_data_1.xlsx/Sheet1
Analysis1
Designer
Results
Algorithm Summary
Summary of the model from R Scripts
Information of the columns used in the algorithm
Independent Columns
id : Integer
host_id : Integer
host_listings_count : Integer
host_total_listings_count : Integer
neighbourhood_cleansed : Integer
zipcode : Integer
latitude : Double
longitude : Double
accommodates : Integer
bathrooms : Double
bedrooms : Integer
beds : Integer
weekly_price : Double
monthly_price : Double
security_deposit : Double
cleaning_fee : Double
guests_included : Integer
extra_people : Double
minimum_nights : Integer
maximum_nights : Integer
availability_30 : Integer
availability_60 : Integer
availability_90 : Integer
availability_365 : Integer
number_of_reviews : Integer
review_scores_rating : Integer
review_scores_accuracy : Integer
review_scores_cleanliness : Integer
review_scores_checkin : Integer
review_scores_communication : Integer
review_scores_location : Integer
review_scores_value : Integer
reviews_per_month : Double
location_exact : Integer
instant_bookDummies : Integer
cancellation_flexible : Integer
cancellation_moderate : Integer
cancellation_strict : Integer
listing_data_1.xlsx/Sheet1
Partition
R-Multiple Linear Regression
Data Grid
Summary
Trend Chart
isting_data_1.xlsx/Sheet1
Showing: 91815/91815 Rows - 68/68 Columns
Never Refreshed
1 Error


File
Edit
View
Data
Help
Prepare
Predict
Visualize
Compose
Share
listing_data_1.xlsx/Sheet1
Analysis1
Designer
Results
Algorithm Summary
price : Double
Summary of the Model
Call:
lm(formula = price ~ id + host_id + host_listings_count + host_total_listings_count +
neighbourhood_cleansed + zipcode + latitude + longitude +
accommodates + bathrooms + bedrooms + beds + weekly_price +
monthly_price + security_deposit + cleaning_fee + guests_included +
extra_people + minimum_nights + maximum_nights + availability_30 +
availability_60 + availability_90 + availability_365 + number_of_reviews +
review_scores_rating + review_scores_accuracy + review_scores_cleanliness +
review_scores_checkin + review_scores_communication + review_scores_location +
review_scores_value + reviews_per_month + location_exact +
instant_bookDummies + cancellation_flexible + cancellation_moderate +
cancellation_strict + cancellation_no_refunds + cancellation_SS30 +
cancellation_SS60, na.action = na.omit)
Residuals:
Min 1Q Median 3Q Max
-700.92 -22.35 -1.95 21.64 313.35
Coefficients: (4 not defined because of singularities)
Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.841e+04 1.363e+04 1.351 0.177477
id -9.466e-07 2.108e-06 -0.449 0.653603
host_id 5.643e-07 3.505e-07 1.610 0.108078
host_listings_count 7.445e-01 7.701e-01 0.967 0.334231
host_total_listings_count NA NA NA NA
neighbourhood_cleansed 1.458e+00 8.948e-01 1.629 0.103962
zipcode -1.746e+00 8.868e-01 -1.968 0.049673 *
latitude 7.714e+01 7.454e+01 1.035 0.301336
longitude -1.757e+01 8.111e+01 -0.217 0.828561
accommodates 2.863e+00 2.830e+00 1.012 0.312279
bathrooms 2.644e+01 7.369e+00 3.588 0.000372 ***
bedrooms -2.780e+00 5.931e+00 -0.469 0.639475
beds -1.910e+00 4.044e+00 -0.472 0.636941
weekly_price 1.137e-01 6.059e-03 18.759 < 2e-16 ***
monthly_price 1.213e-02 1.937e-03 6.263 9.25e-10 ***
security_deposit 1.246e-02 1.238e-02 1.007 0.314679
cleaning_fee -4.461e-02 1.057e-01 -0.422 0.673157
guests_included -5.883e+00 1.970e+00 -2.986 0.002991 **
extra people 1.180e-01 1.055e-01 1.119 0.263908
listing_data_1.xlsx/Sheet1
Partition
R-Multiple Linear Regression
Data Grid
Summary
Trend Chart
sting_data_1.xlsx/Sheet1
Showing: 91815/91815 Rows - 68/68 Columns
Never Refreshed
1 Error

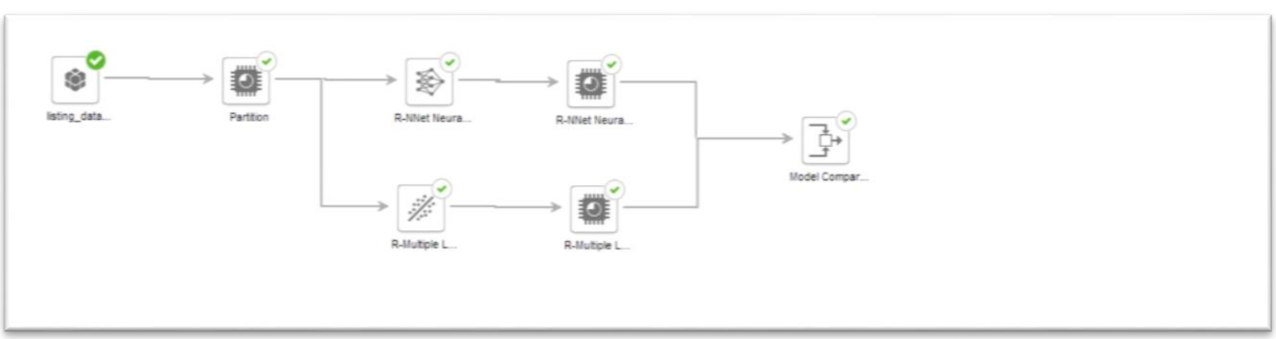


Target variable is price. R square value = 0.92 which indicates that 92% of the target variable is explained by the input variables.

Model Comparison

We have considered implementing two different models to precise check the correlation coefficient.

Neural network and multiple linear regression.

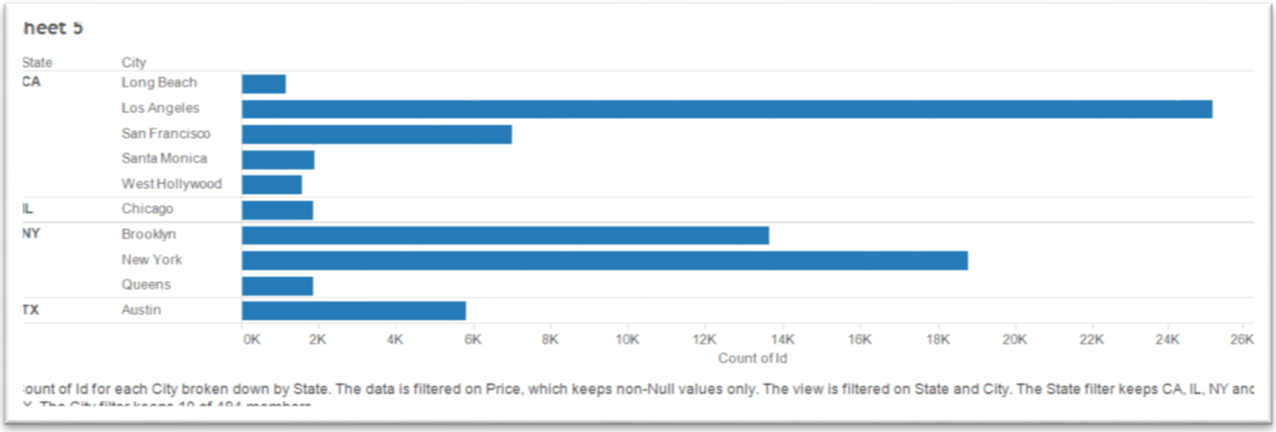


```
Summary of the model from R Scripts

Information of the columns used in the algorithm
------------------------------------------------------
Independent Columns
 host_total_listings_count : Integer
 accommodates : Integer
 bathrooms : Double
 bedrooms : Integer
 beds : Integer
 weekly_price : Double
 monthly_price : Double
 security_deposit : Double
 cleaning_fee : Double
 guests_included : Integer
 extra_people : Double
 minimum_nights : Integer
 maximum_nights : Integer
 availability_30 : Integer
 availability_60 : Integer
 availability_90 : Integer
 availability_365 : Integer
Dependent Column
 price : Double

Summary of the Model
a 17-5-1 network with 96 weights
options were - linear output units
  b->h1  i1->h1  i2->h1  i3->h1  i4->h1  i5->h1  i6->h1  i7->h1  i8->h1  i9->h1
  -0.59   19.46    2.38    8.47   18.46  851.10 -151.03  169.39 -306.89  -17.09
i10->h1 i11->h1 i12->h1 i13->h1 i14->h1 i15->h1 i16->h1 i17->h1
  19.80 -174.53   36.96 -127.70 -415.74 -889.34  -83.99 -306.70
  b->h2  i1->h2  i2->h2  i3->h2  i4->h2  i5->h2  i6->h2  i7->h2  i8->h2  i9->h2
   0.13   -0.37    0.20    0.53   -0.22   -0.10   -0.56    0.54   -0.55    0.33
i10->h2 i11->h2 i12->h2 i13->h2 i14->h2 i15->h2 i16->h2 i17->h2
  -0.05   -0.17   -0.37   -0.02   -0.15   -0.55    0.63    0.56
  b->h3  i1->h3  i2->h3  i3->h3  i4->h3  i5->h3  i6->h3  i7->h3  i8->h3  i9->h3
   0.50    0.62   -0.45    0.61   -0.22   -0.09    0.46    0.42   -0.45   -0.34
i10->h3 i11->h3 i12->h3 i13->h3 i14->h3 i15->h3 i16->h3 i17->h3
  -0.45    0.70   -0.18    0.12   -0.47    0.60   -0.16   -0.58
  b->h4  i1->h4  i2->h4  i3->h4  i4->h4  i5->h4  i6->h4  i7->h4  i8->h4  i9->h4
  -0.08    0.47    0.21    0.06   -0.40   -0.03   -1.22   -0.88   -0.20    0.09
i10->h4 i11->h4 i12->h4 i13->h4 i14->h4 i15->h4 i16->h4 i17->h4
```

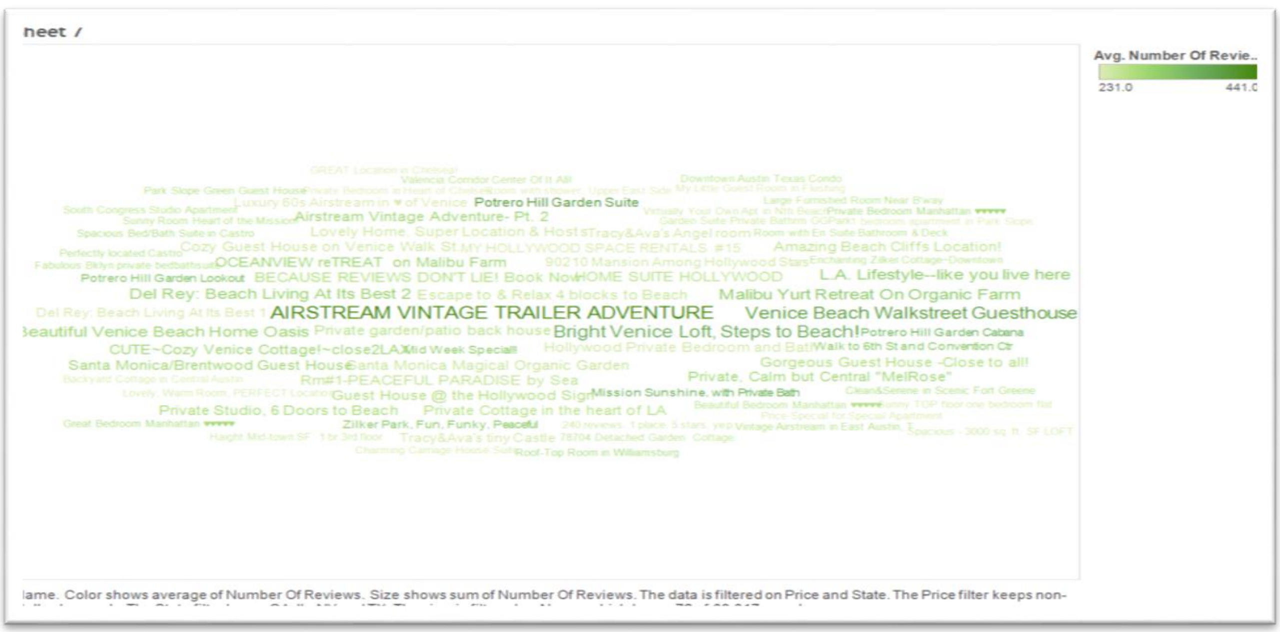
```
Summary of the Model
a 17-5-1 network with 96 weights
options were - linear output units
  b->h1  i1->h1  i2->h1  i3->h1  i4->h1  i5->h1  i6->h1  i7->h1  i8->h1  i9->h1
  -0.59   19.46    2.38    8.47   18.46  851.10 -151.03  169.39 -306.89  -17.09
i10->h1 i11->h1 i12->h1 i13->h1 i14->h1 i15->h1 i16->h1 i17->h1
  19.80 -174.53   36.96 -127.70 -415.74 -889.34  -83.99 -306.70
  b->h2  i1->h2  i2->h2  i3->h2  i4->h2  i5->h2  i6->h2  i7->h2  i8->h2  i9->h2
   0.13   -0.37    0.20    0.53   -0.22   -0.10   -0.56    0.54   -0.55    0.33
i10->h2 i11->h2 i12->h2 i13->h2 i14->h2 i15->h2 i16->h2 i17->h2
  -0.05   -0.17   -0.37   -0.02   -0.15   -0.55    0.63    0.56
  b->h3  i1->h3  i2->h3  i3->h3  i4->h3  i5->h3  i6->h3  i7->h3  i8->h3  i9->h3
   0.50    0.62   -0.45    0.61   -0.22   -0.09    0.46    0.42   -0.45   -0.34
i10->h3 i11->h3 i12->h3 i13->h3 i14->h3 i15->h3 i16->h3 i17->h3
  -0.45    0.70   -0.18    0.12   -0.47    0.60   -0.16   -0.58
  b->h4  i1->h4  i2->h4  i3->h4  i4->h4  i5->h4  i6->h4  i7->h4  i8->h4  i9->h4
  -0.08    0.47    0.21    0.06   -0.40   -0.03   -1.22   -0.88   -0.20    0.09
i10->h4 i11->h4 i12->h4 i13->h4 i14->h4 i15->h4 i16->h4 i17->h4
  -0.55   -2.70    0.38    0.52   -0.34   -0.77   -0.37   -0.10
  b->h5  i1->h5  i2->h5  i3->h5  i4->h5  i5->h5  i6->h5  i7->h5  i8->h5  i9->h5
  -0.61   -0.42   -0.02   -0.47    0.64   10.48   35.73    8.58    0.59    0.65
i10->h5 i11->h5 i12->h5 i13->h5 i14->h5 i15->h5 i16->h5 i17->h5
   1.07    0.69    0.60    0.98    0.45    3.42    0.17    0.44
  b->o   h1->o   h2->o   h3->o   h4->o   h5->o
  54.63  149.92 -127.10   55.41   59.30   -1.54


SSE
[1] 428620151


Model Converged (0-Yes,1-No)
[1] 0
```

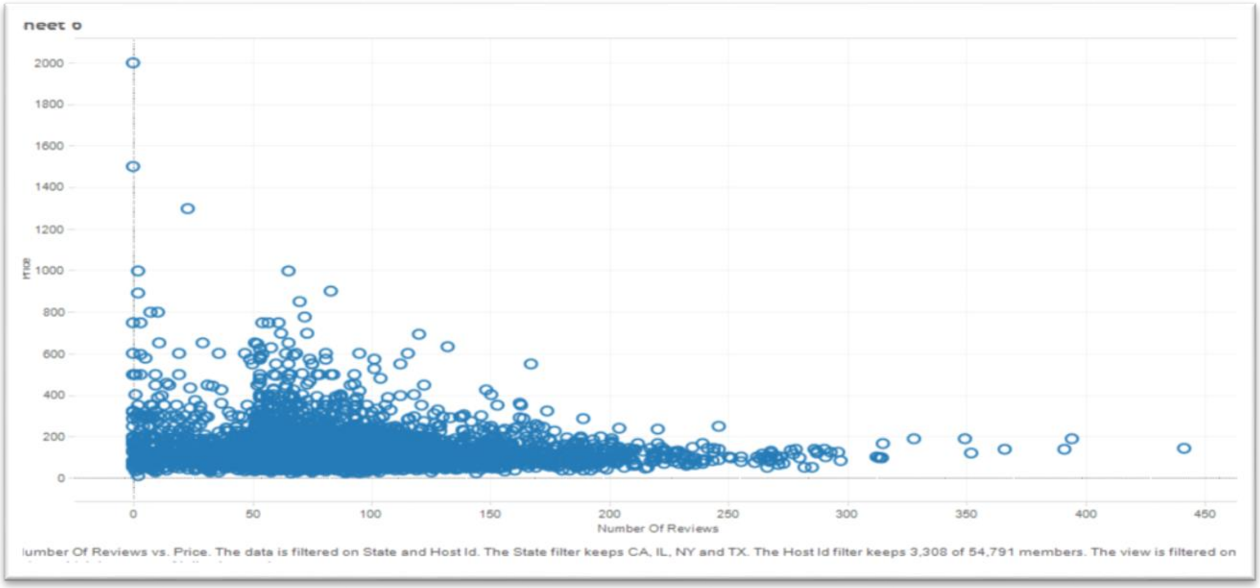
Analyzing Theorem III: To create a word-cloud from listing descriptions from which users can pull data to find a perfect match for a dream vacation

Listings by Locality


heet 5
State
City
CA
Long Beach
Los Angeles
San Francisco
Santa Monica
West Hollywood
IL
Chicago
NY
Brooklyn
New York
Queens
TX
Austin
0K 2K 4K 6K 8K 10K 12K 14K 16K 18K 20K 22K 24K 26K
Count of Id
ount of Id for each City broken down by State. The data is filtered on Price, which keeps non-Null values only. The view is filtered on State and City. The State filter keeps CA, IL, NY and

Word Cloud


heet 7
Avg. Number Of Revie..
231.0
441.0
Potrero Hill Garden Suite
Airstream Vintage Adventure- Pt. 2
Lovely Home: Super Location & Hosts
Cozy Guest House on Venice Walk St
Amazing Beach Cliffs Location!
OCEANVIEW reTREAT on Malibu Farm
BECAUSE REVIEWS DON'T LIE! Book Now
HOME SUITE HOLLYWOOD
L.A. Lifestyle--like you live here
Del Rey: Beach Living At Its Best 2
Escape to & Relax 4 blocks to Beach
Malibu Yurt Retreat On Organic Farm
AIRSTREAM VINTAGE TRAILER ADVENTURE
Venice Beach Walkstreet Guesthouse
Beautiful Venice Beach Home Oasis
Private garden/patio back house
Bright Venice Loft, Steps to Beach!
CUTE~Cozy Venice Cottage!~close2LAX
Hollywood Private Bedroom and Bath
Santa Monica/Brentwood Guest House
Santa Monica Magical Organic Garden
Gorgeous Guest House -Close to all!
Private, Calm but Central "MelRose"
Rm#1-PEACEFUL PARADISE by Sea
Guest House @ the Hollywood Sign
Mission Sunshine, with Private Bath
Private Studio, 6 Doors to Beach
Private Cottage in the heart of LA
Zilker Park, Fun, Funky, Peaceful
Tracy&Ava's tiny Castle
Roof-Top Room in Williamsburg
ame. Color shows average of Number Of Reviews. Size shows sum of Number Of Reviews. The data is filtered on Price and State. The Price filter keeps non-



Conclusion:

From the analysis we have set the benchmark to predict the price of listing based on these factors we provide the following recommendations:

1) How much of an impact will the variables like weekly price, monthly price, host response rate etc. have on the target variable price
2) Using multiple linear regression, we tried to predict the price of listings. These models will help the customer and Renter to weigh tradeoffs and identify the best possible solution.

   These predictions will assist The Renter and Customer to exploit the value of data to make informed decisions