Data Science Monsters Hotel

Hilary
Qi Rui
Ying Ki

Who, What, Why, Where, How

ReviewPro

- Manage & Improve Online Reputation
- Dataset includes -
  - Timeline: **2013**
  - 3 worksheets:
    - Sentiment Data;
    - Sentiment Mentions; &
    - STB Rating (Hotel Star Rating).

What is GRI™

- Global Review Index
- Benchmark for reputation management efforts
- A proprietary algorithm

Hypothesis

Based on a [regional hotel review report done by PwC](#)

1. Location affects customer satisfaction significantly
2. Five star hotels are rated more positively on 'rooms' than Three-star hotels.
3. Three-star hotels are rated more negatively on 'value' than Five-star hotels

Is this an oversimplified picture?

Highlighting important factors contributing to customer satisfaction to Singapore based hotels, thus allowing the hotels to prioritize operational and service enhancements for greater customer satisfaction

How 'Customer Satisfaction' was defined

For each hotel,

Customer Satisfaction Score = (% of positive reviews) - (% of negative reviews)

Assumptions:

Positive reviewers are highly satisfied customers who act as evangelists for your brand, while _negative reviewers_ are displeased customers who will also let everyone know why they are displeased with your business and its products and/or services. Assumptions similar to that of Net Promoter Score (NPS).

Location affects customer satisfaction significantly, but it's not one of the main driver

Bar chart illustrating distribution of the correlation against customer satisfaction
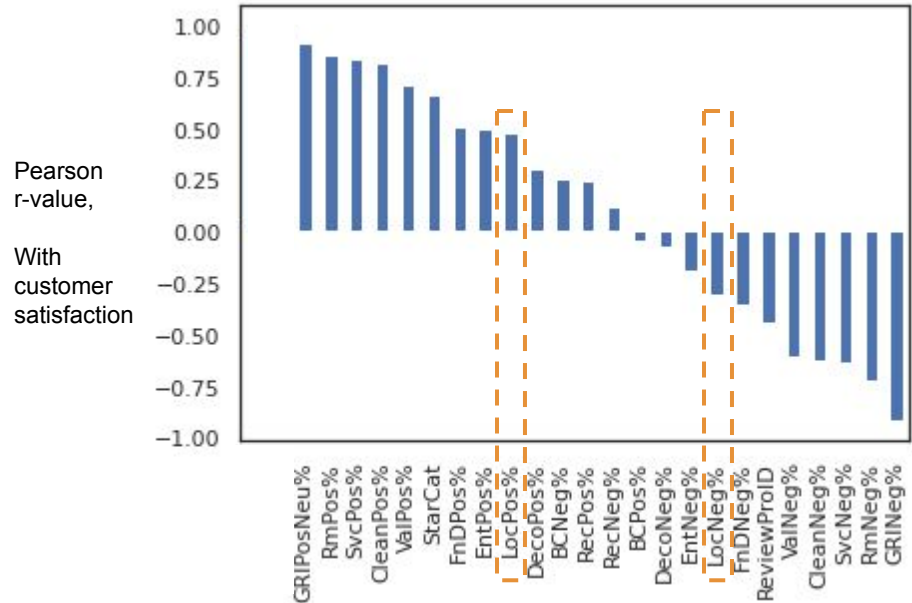
Pearson-r Test Result:

r= 0.476, p=0.000
'LocPos%' is significantly dependent to
Customer Satisfaction in the positive direction

r= -0.309, p=0.000
'LocNeg%' is significantly dependent to
Customer Satisfaction in the negative direction

Pearson r-value,

With customer satisfaction

Insights

- Location is not the most important feature that affects the customer satisfaction
- Top 3 features that has the highest correlation to 'Customer Satisfaction': 1.'Room', 2.'Service', 3.'Cleanliness''
- Does the top features differ between hotels of different star category?

Hotel customers care more about 'room satisfaction' for 5-star hotels, than 4-star and 3-star hotels

	sum_sq	df	mean_sq	F	PR(>F)	eta_sq	omega_sq
C(StarCat)	4.150548	2.0	2.075274	60.088714	7.399444e-21	0.392509	0.38472
Residual	6.423852	186.0	0.034537	NaN	NaN	NaN	NaN
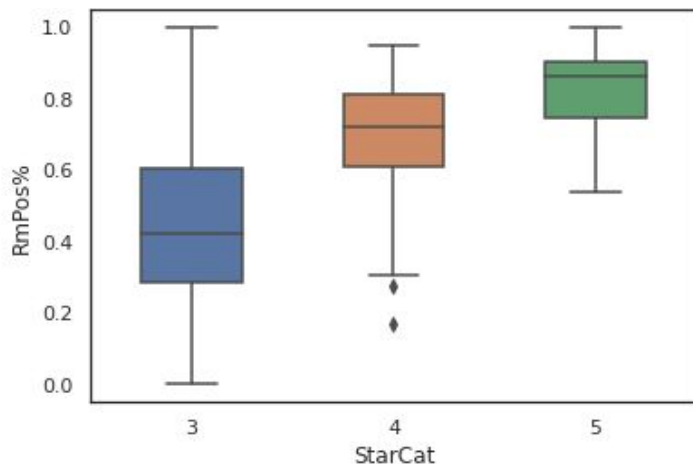
'RmPos%' =
the percentage of positive reviews hotels get about their 'rooms'

Multiple Comparison of Means - Tukey HSD,
FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
3	4	0.2482	0.001	0.176	0.3204	True
3	5	0.3787	0.001	0.2912	0.4662	True
4	5	0.1305	0.001	0.0456	0.2154	True

Test Multiple Comparison ttest_ind
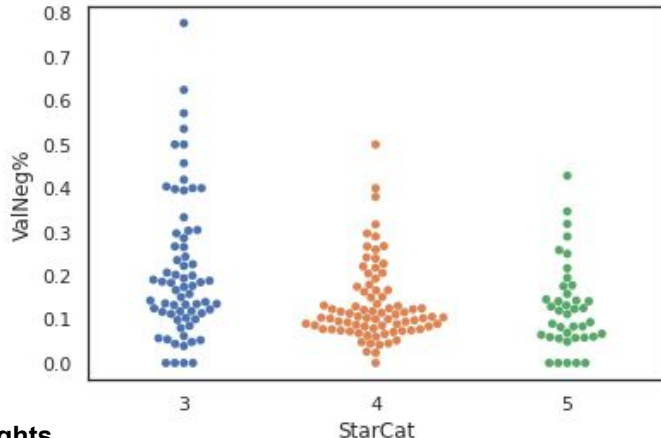FWER=0.05 method=bonf alphacSidak=0.02,
alphacBonf=0.017

group1	group2	stat	pval	pval_corr	reject
3	4	-7.5452	0.0	0.0	True
3	5	-9.5151	0.0	0.0	True
4	5	-4.4499	0.0	0.0001	True

Insights

- **Using a One-way ANOVA test + Posthoc tests - mean of 'RmPos%' is significantly different between all 3 groups.**
- **'RmPos%' is significantly higher for 5 star hotels, compared to 4 star hotels and 3-star hotels**
- What is contributing to room satisfaction? What aspects should hotels pay attention to relating to 'rooms'?

Hotel customers care more about 'value' for 3-star hotels, than 4-star and 5-star hotels

	sum_sq	df	mean_sq	F	PR(>F)	eta_sq	omega_sq
C(StarCat)	0.288789	2.0	0.144395	9.628889	0.000105	0.093822	0.083671
Residual	2.789251	186.0	0.014996	NaN	NaN	NaN	NaN

'ValNeg%' =
the percentage of negative reviews hotels get about their 'value''

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
3	4	-0.0788	0.001	-0.1264	-0.0312	True
3	5	-0.0862	0.0015	-0.1439	-0.0286	True
4	5	-0.0074	0.9	-0.0634	0.0485	False

Test Multiple Comparison ttest_ind FWER=0.05
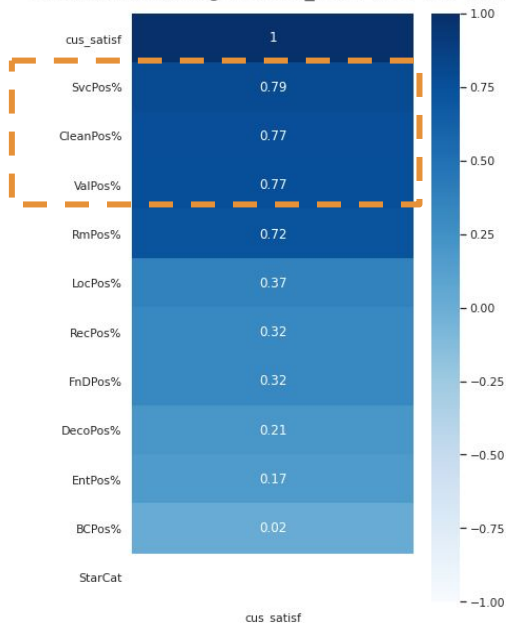method=bonf alphacSidak=0.02,
alphacBonf=0.017

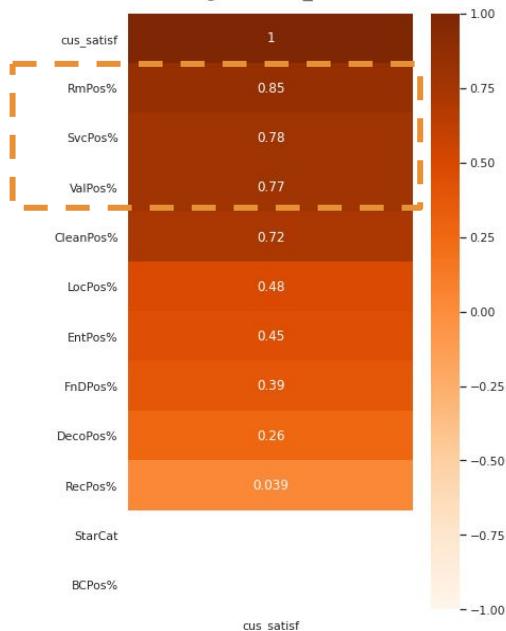| group1	group2	stat	pval	pval_corr	reject
3	4	3.7447	0.0003	0.0008	True
3	5	3.0345	0.003	0.0091	True
4	5	0.4159	0.6782	1.0	False

Insights

- **Using a One-way ANOVA test + Posthoc tests - mean of 'ValNeg%' is significantly different between 3-star hotels, as compared to 4 & 5-star hotels.**
- **'ValNeg%' is significantly higher for 3-star hotels, as compared to 4 & 5-star hotels.**
- Were the customers for the 5-star hotels actually satisfied with the value?
- What made the customers think that 5-star hotels are worth it? What are the factors that led to them coming to this conclusion?

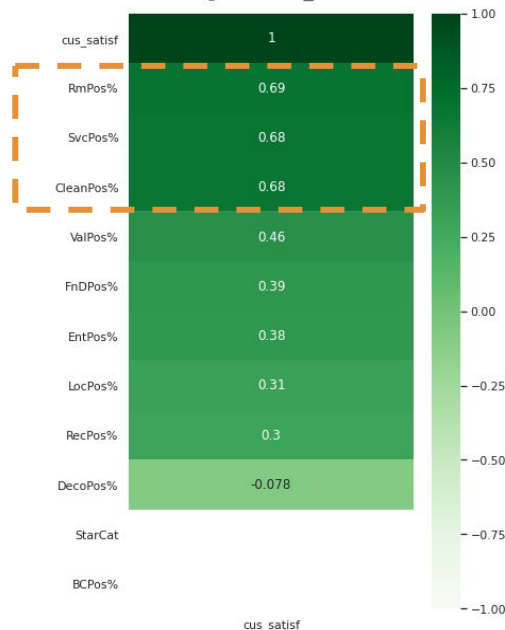# Top 3 Features highly Correlates with Customer Satisfaction by Hotel Star Rating

Features Correlating with cus_satisf for 3-Star Hotel

	cus_satisf
cus_satisf	1
SvcPos%	0.79
CleanPos%	0.77
ValPos%	0.77
RmPos%	0.72
LocPos%	0.37
RecPos%	0.32
FnDPos%	0.32
DecoPos%	0.21
EntPos%	0.17
BCPos%	0.02
StarCat	

Features Correlating with cus_satisf for 4-Star Hotel

	cus_satisf
cus_satisf	1
RmPos%	0.85
SvcPos%	0.78
ValPos%	0.77
CleanPos%	0.72
LocPos%	0.48
EntPos%	0.45
FnDPos%	0.39
DecoPos%	0.26
RecPos%	0.039
StarCat	
BCPos%	

Features Correlating with cus_satisf for 5-Star Hotel

	cus_satisf
cus_satisf	1
RmPos%	0.69
SvcPos%	0.68
CleanPos%	0.68
ValPos%	0.46
FnDPos%	0.39
EntPos%	0.38
LocPos%	0.31
RecPos%	0.3
DecoPos%	-0.078
StarCat	
BCPos%	

Top 3 features that hotels should focus on

- 3-star hotel customers care more about 'Service', 'Cleanliness' & 'Value'
- 4-star hotel customers care more about 'Room', 'Service' & 'Value'
- 5-star hotel customers care more about 'Room', 'Service' & 'Cleanliness'

Summary

Hypotheses

1. Location plays a significant role but it is not the most important for customer satisfaction
2. 5-star hotel customers care more about room satisfaction than 4-star and 3-star hotel customers
3. 5-star hotel customers care about value as well

Recommendations

1. Target Audience
2. Room Concept
3. Amenities
4. Services Rendered

Thank you

Feel free to give us feedback!