

Disneyland Reviews

Anthony DeNiro, Rastko Stojsin, Teddy Weaver



Motivation / Goal

1 - Which Disneyland is best to visit and during what time of the year

Plus Rastko has always wanted to go to Disneyland - where should Rastko go?

2 - On the disneyland website, park visitors can leave text reviews and star ratings of the parks. The ratings may be more useful when trying to make decisions about the park. Users are not required to enter star ratings when they write a review.

During analysis these missing star ratings reviews must be dropped. We built models to predict a star rating from the visitor text review to be able to impute these reviews for future strategic decisions.

Data Overview & Tools

Review_ID	Rating	Year_Month	Reviewer_Location	Review_Text	Branch
670772142	4	2019-4	Australia	If you've ever be...	Disneyland_HongKong
670682799	4	2019-5	Philippines	Its been a while ...	Disneyland_HongKong
670623270	4	2019-4	United Arab Emirates	Thanks God it was...	Disneyland_HongKong
670607911	4	2019-4	Australia	HK Disneyland is ...	Disneyland_HongKong
670607296	4	2019-4	United Kingdom	the location is n...	Disneyland_HongKong

only showing top 5 rows

42,000+ Reviews (~2,000 missing dates)

1-5 Rating Scale

Hong Kong, Paris, California

2010-2019 ('10, '19 were partial)



Development Notebook



Feature Engineering



Modeling



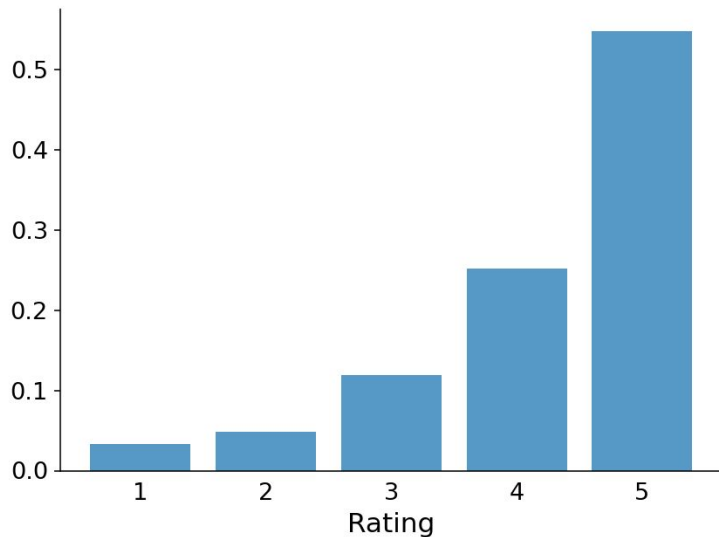
Formatting + Plotting



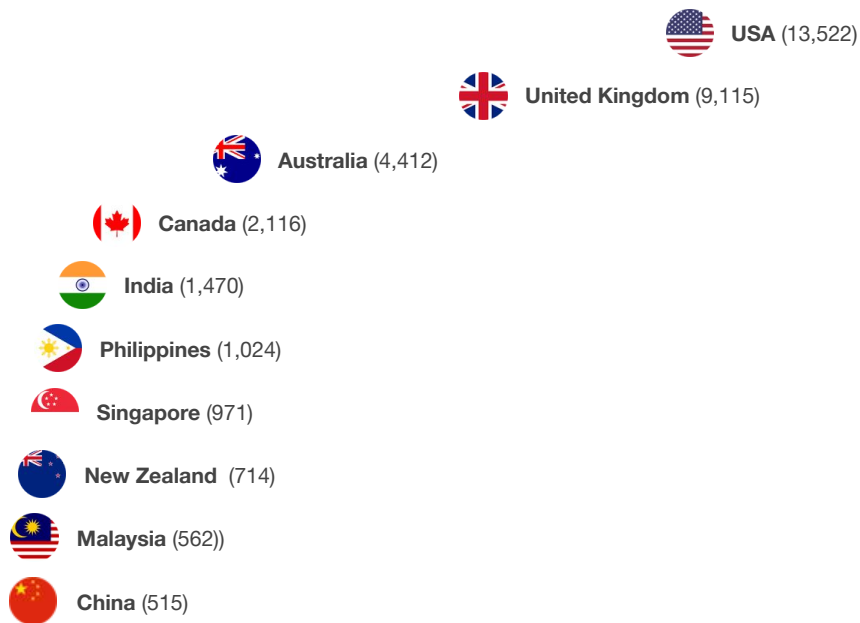
Distribution of Ratings

Over 40,000 reviews from 160 different countries

Heavy skew of 4 and 5 stars ratings



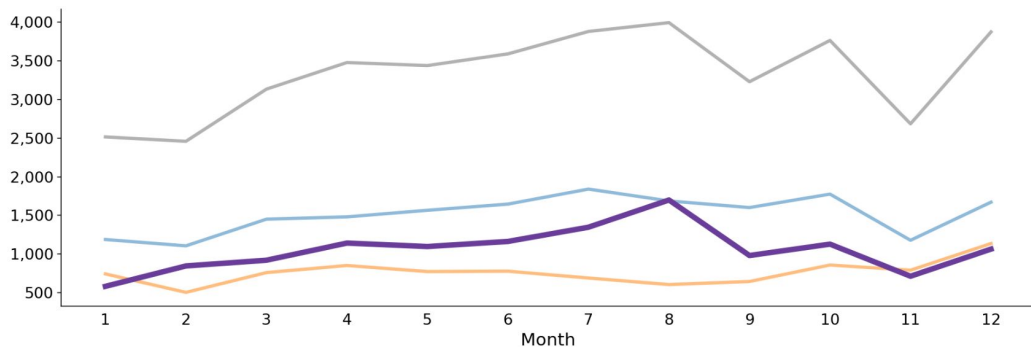
86% of reviews from 10 countries



Reviews by Calendar Month

Does the time of year impact reviews?

Total # of Reviews

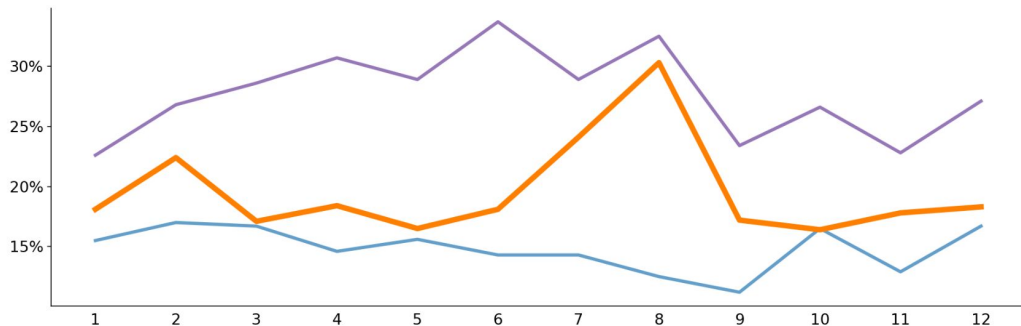


Overall: More review in Spring/Summer than Fall/Winter

Overall: Increase in December (holiday travel?)

Paris: August was the busiest month

Pct. Negative Reviews (1-3)



Paris: Regularly had bad reviews

Hong Kong: Huge spike in August!

California: Consistently well reviewed

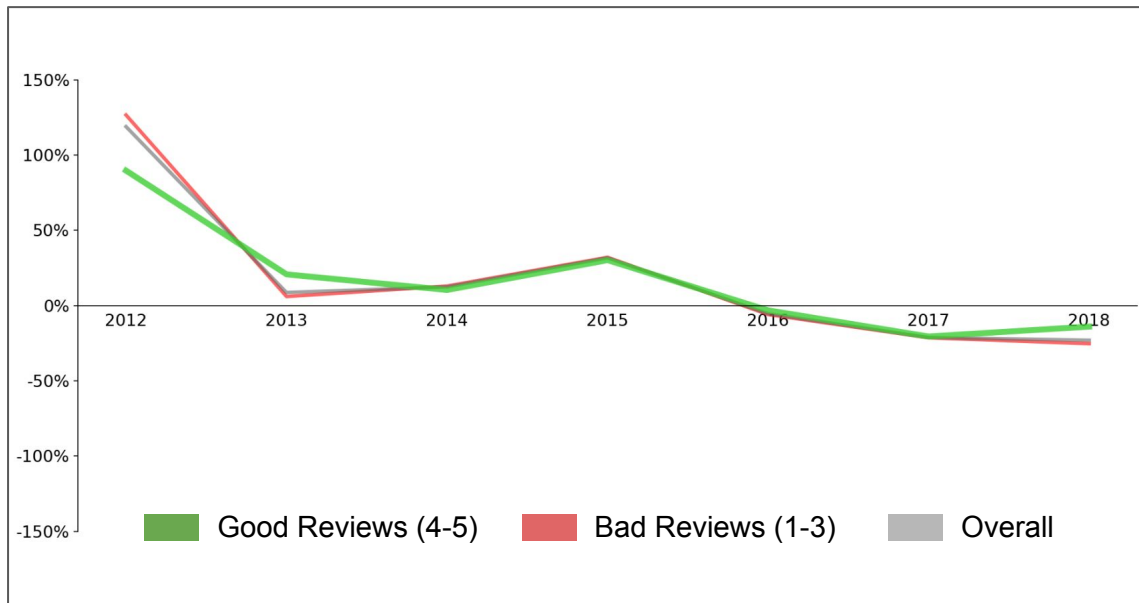
Year/Year Change

Did the number of new reviews change over time? If so, was this because of bad experiences?

Decline in the number of new reviews starting in 2016 across all locations.

No significant year/year difference between good reviews (green, 4-5) and bad reviews (red, 1-3).

2018 had a higher percentage of good reviews compared to previous years



What do Serbians Think?

11 reviews from 2012 to 2019 (10 Paris, 1 Hong Kong)



Disneyland Hong Kong

"This Disney park **is the least interesting of all Disneylands** in the world, and also the smallest one"

Disneyland Paris

Great for both, adults and kids [...] **Try not to miss parade** at 17.30 [...], **its really magical :)**

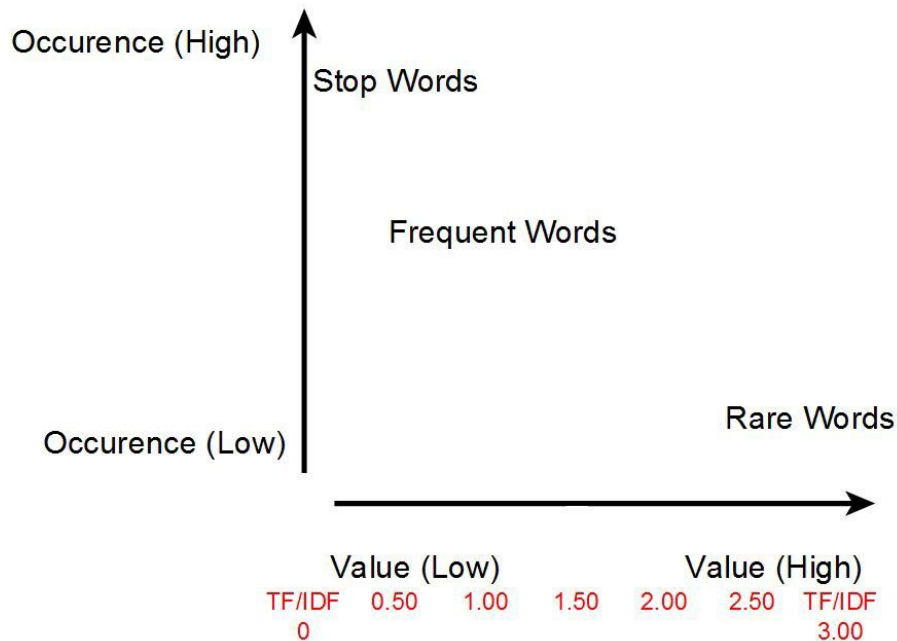
Totally amazing experience. I would really be so sorry if I missed that one day in this fabulous park

Well, in one word I can say **it's an ordinary Disneyland**

For picture with Mickey Mouse we have to **wait almost two hours**, that was crazy, but what is Disneyland visit without picture with Mickey Mouse?! [...] I suggest people not to go there. **The only positive thing about Disneyland ist feuerwerk at 11 o'clock!!!**

Feature Engineering

- TF-IDF
 - Tokenize words, remove stop words
- word embeddings (Word2Vec) to vectorize text input
- K-Means with 2 clusters using word embeddings



Predicting Missing Ratings

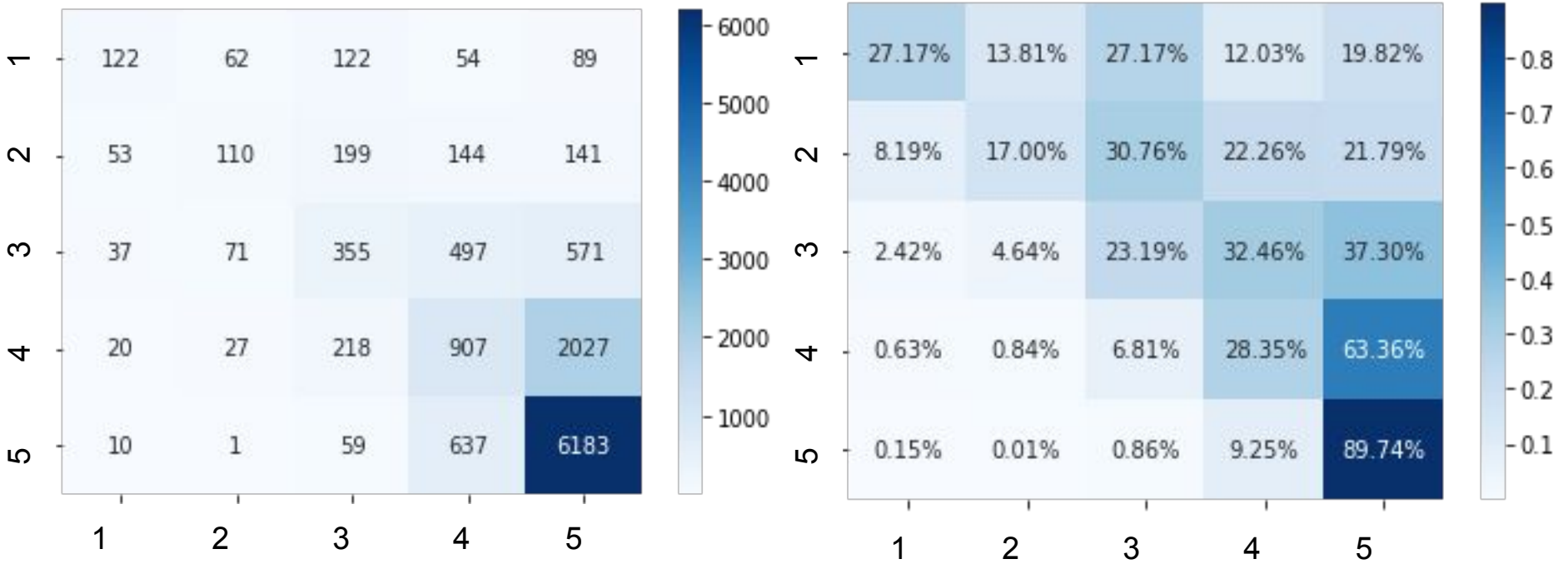
Inputs : feature engineered

Prediction : 1-5 star rating

Models	Accuracy
Logistic Regression	60.37%
Decision Tree	55.33%
Random Forest	54.18%

	precision	recall	f1-score	support
1	0.50	0.27	0.35	449
2	0.41	0.17	0.24	647
3	0.37	0.23	0.29	1531
4	0.41	0.28	0.33	3199
5	0.69	0.90	0.78	6890
accuracy			0.60	12716
macro avg	0.47	0.37	0.40	12716
weighted avg	0.56	0.60	0.56	12716

Best Model Performance



Wrap Up / Findings

What we learned about the different locations

California > Hong Kong > Paris

Hong Kong is the worst in August

California and Paris were consistently reviewed year-round

Where should Rastko go?

California for the best time!

If Hong Kong then not in august!

If Paris make sure to catch the feurwerks at 11!

Rating Prediction Findings

Not bad results - could work as imputation tool for Disneyland to not lose data for analysis

Recommendation

Use our tool to impute the already existing missing ratings for reviews and change website so that you have to input rating as well

Next Steps

Topic Modeling

- What aspects with the parks are highlighted the most in the reviews?

Sentiment Analysis

- Further NLP to help us identify what people liked or disliked about the parks



Ist feuerwerk