



Yelp Data Analysis

Group 9

Jingwen Yan, Chong Wei, Yuchen Zeng

What we've done so far ...

— — —

1. Text Processing

- Change all letters into lower case;
- Delete stop words;
- Split the sentences into words;
- Turn the tense into the present tense;
- Correct the spelling error.

2. Count the frequency of words

- Find high frequency words: Movie, seat, theater, ticket, ...



Sentiment Analysis

Extracting Key Attributes

— — —

Attributes	Key words
Price	Ticket, cost, deal, price, ...
Location	Location, hotel, restaurant, bar, car park, ...
Facility	Seat, movie, screen, sound, 3d, ...
Environment	Smell, style, wait time, ...
Food & Drinks	Concession stand, menu, soda, popcorn, ...
Promotion	Vip, free pass, discount, family, ...
Service	Service, staff, online, app, ...

Sentiment Analysis

— — —

1. Find the segments where key words appear

*“Reclining **chairs** with arm rests that rise are very good. Movie **prices** are rising more and more.”*

Key words	Segments used for sentiment analysis
Chair	Recline chair with arm rest that rise are very good
Price	Movie price are rise more and more



Sentiment Analysis

— — —

2. Compute the polarity for the key words:

Lexicon of Adjectives

```
<word form="very" pos="RB" sense="used as intensifier" polarity="0.2" subjectivity="0.3"  
intensity="1.3" />
```

- `pos` - the position of the word (verb., noun., ...)
- `sense` - the situation where the word is used
- `polarity` - 1 being the most positive, -1 being the most negative
- `intensity` - the effects of modifier words (eg. very, extremely, ...)



Sentiment Analysis

— — —

2. Compute the polarity for the key words:

- Find every word's average polarity of all senses
- If the adjective comes with a negation (eg. not, never, ...), we multiply the polarity by -0.5
- For modifier words (eg. very), we only consider its "intensity"
- See whether the adjective comes with a modifier word, if yes:
 - If the modifier word is together with a negation (eg. not very good), we divide the polarity by intensity
 - If not, we multiply the polarity by intensity



Sentiment Analysis

Example: “Recline chair with arm rest that rise are *very good*”

“Good”: polarity = 0.7

```
<word form="good" pos="JJ" sense="having desirable or positive qualities especially those  
suitable for a thing specified" polarity="0.7" subjectivity="0.4" intensity="1.0"/>  
<word form="good" pos="JJ" sense="tending to promote physical well-being" polarity="0.7"  
subjectivity="0.8" intensity="1.0"/>
```

“Very”: intensity = 1.3

```
<word form="very" pos="RB" sense="used as intensifier" polarity="0.2" subjectivity="0.3"  
intensity="1.3" />
```

“Very good”:

Polarity = $0.7 \times 1.3 = 0.91$



Sentiment Analysis

— — —

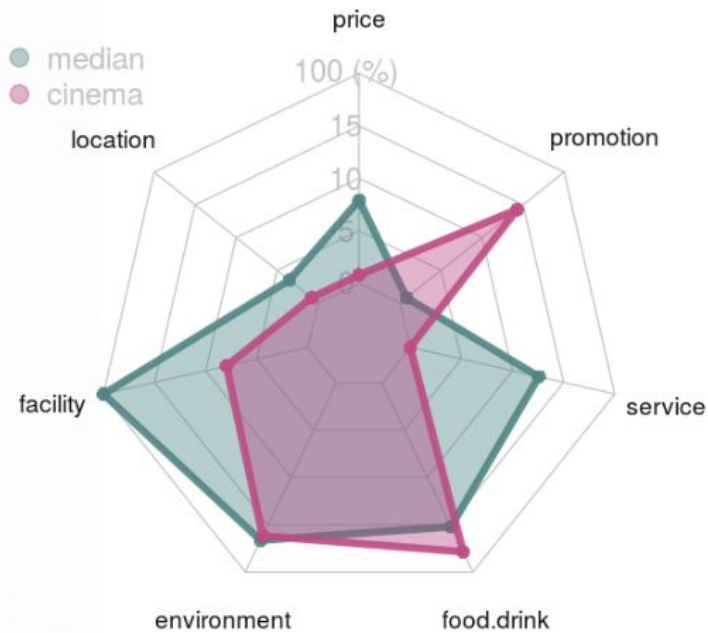
Score Table

id	name	price	location	facility	environment	food&drink	service	promotion
OEQrPxeku4BfHMCSi8UASQ	Chandler Cinemas	0.000	0.000	0.686	1.005	0.225	-0.063	0.313
zRV7bzP_CfTg-_R9U-VsVg	Visulite Theatre	0.009	0.8	9.596	5.817	6.180	3.115	0.000
pDA8NJUwGI1loLDeaVfo0Q	AMC Ridge Park Square Cinema 8	0.400	0.450	0.624	0.234	0.000	0.683	0.000

Sentiment Analysis

— — —

Radar Chart



Advice Generating

Comparison With Median

— — —

Price	Location	Facility	Environment	Food&drinks	Service	Promotion
0.80	0.45	2.11	1.24	0.74	0.99	0.13
0.23	1.30	2.78	1.71	1.77	0.41	0.50

“The rating of location, facility, environment, food&drinks, promotion of your cinema is higher than median. ”

“The rating of price, service of your cinema is lower than median. ”



Comparison Between Aspects

— — —

Price	Location	Facility	Environment	Food&drinks	Service	Promotion
0.80	0.45	2.11	1.24	0.74	0.99	0.13
0.23	1.30	2.78	1.71	1.77	0.41	0.50

“Your customers are complaining about the price. ”

“The customers are expecting better service. ”

“The facilities make your customers comfortable. ”

“Food and drinks here provides customers a great experience.”



Comparison Between Aspects

— — —

Your customers are complaining about the price.

Your cinema may not be in the best area.

You should improve the facilities.

It's helpful if your can create a better environment.

Better food and drinks can make your cinema more competitive.

The customers are expecting better service.

More promotion plans may attract more customers.

Your customers are pleased with the price.

The location of your cinema is great.

The facilities make your customers comfortable.

Your cinema has create a pleasant environment for the customers.

Food and drinks here provides customers a great experience.

The service is gratifying.

Customers like your promotion activities very much.



Price	Location	Facility	Environment	Food&drinks	Service	Promotion
0.80	0.45	2.11	1.24	0.74	0.99	0.13
2.61	1.11	8.46	0.88	3.76	2.23	0.38

The rating of price, location, facility, food&drinks, service, promotion of your cinema is higher than median.

The rating of environment of your cinema is lower than median.

Your cinema has done a great job, but there's still something you can improve. More promotion plans may attract more customers. It's helpful if you can create a better environment.

The facilities make your customers comfortable. Food and drinks here provides customers a great experience.



Strength and Weakness

Strength

1. Find spam users and exclude them in the following analysis.
2. Use very specific key words for cinemas to analyse the sentiment of certain review.
3. Generate a comfortable paragraph to the cinema owners to give advice and compliments.
4. Provide both vertical and horizontal comparison.

Weakness

1. Not enough reviews makes it hard to assess all aspects for all cinemas.
2. The sentiment analysing model may not give the best result in certain contexts.
3. The key words cover most of the reviews but there exists some using uncommon words which we can't analyse properly.

Shiny Application

Visit our shiny webpage: <https://reaperconan.shinyapps.io/theateradvisor/>