

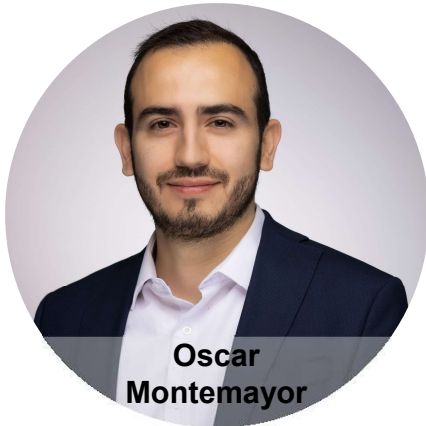
NLP Techniques with Disneyland Reviews

April 2023

 McGill |  DESAUTELS



The Team



Customer Satisfaction drives growth and brand reputation

Measuring and analyzing customer satisfaction is crucial for retaining customers, building a positive brand reputation and driving business growth

Customer Satisfaction Relevance



1. Customer Retention

Satisfied customer = Loyalty customers



>25%

Increase in profit due a 5% increase in customer retention^{/1}

2. Brand Reputation

Satisfied customers will share positive reviews creating positive brand perception



86%

Of customers stated that online negative reviews impact their choices^{/2}

3. Business Growth

Satisfied customers will become repeat customers leading to increased revenue



5 - 25%

More spending from customers when they get excellent customer service^{/3}

4. Customer Engagement

Satisfied customers will develop long-term relationships



5

Times more likely to repurchase from a company by loyalty customers^{/4}

Sources: /1 Harvard Business Review, 2/Dimensional Research, 3/ American Express /4 Temkin Group

But, how to measure and analyze customer satisfaction?

One common option to measure customer satisfaction is through surveys; however, they are expensive
The solution is to develop an NLP Sentiment Analysis in the customers reviews

The old way: Surveys



- **Small & biased sample**
- **Time lag** between data collection and analysis
- Rely on **close-ended questions** and may not capture the depth of feedback
- May face issues of **response bias** or social desirability bias
- **Associated costs** (survey design, data collection, and data entry)

The solution: NLP Sentiment Analysis



- **Large and diverse** sample of customer opinions
- **Real-time** and current feedback from actual customers
- **Rich and detailed feedback** that provide more qualitative insights
- **Authentic and trustworthy** reviews
- More **cost-effective** compared to conducting traditional surveys

The solution: Sentiment Analysis and Topic Modeling



The **goal of the project is to develop a Customer Satisfaction analytical tool to monitor, track and analyze visitor reviews** in Paris, California, and Hong Kong Disneyland branches



The solution is based on **42,000 reviews on 3 Disneyland branches** posted by visitors on **Trip Advisor**

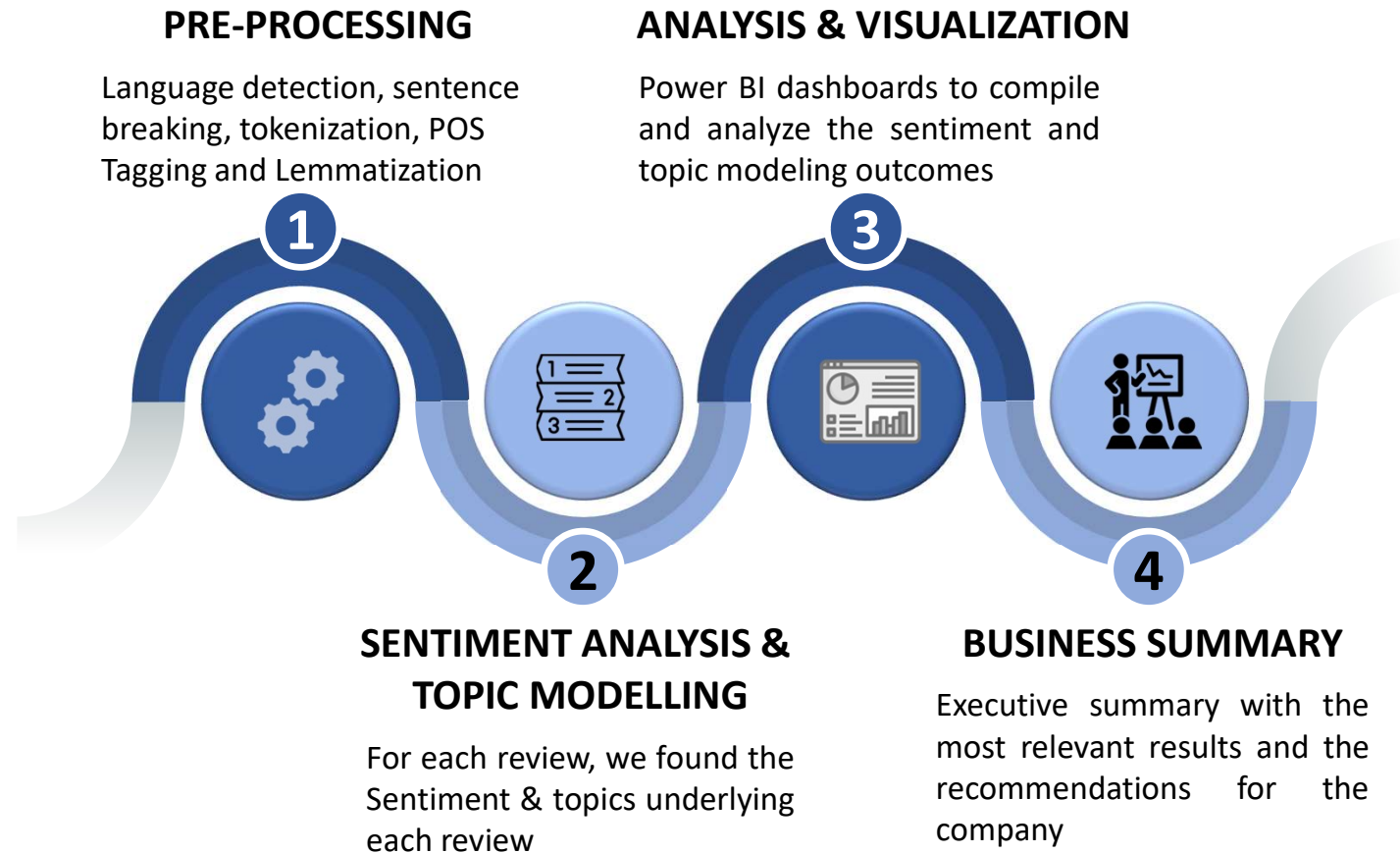


The Customer Satisfaction tool will allow Disneyland to:

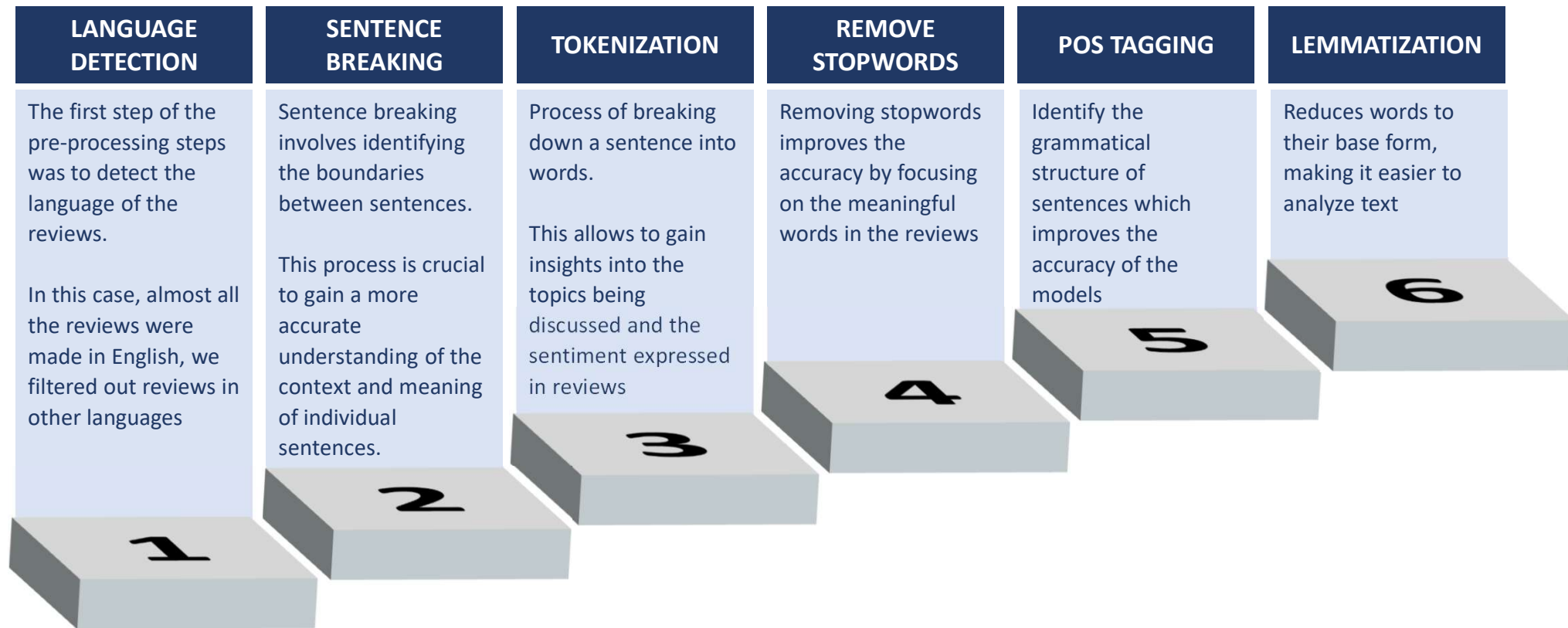
- **Analyze Sentiment Analysis by branch and a timeframe** selected by the user
- Understand how **customer satisfaction is related to the overall rating**
- Identify which **themes are underlying the positive and negative reviews** to:
 - Know the competitive advantages of Disneyland in the positive reviews and;
 - Take action to improve the underlying themes in negative reviews



The solution is composed of 4 modules



Pre-processing steps of the solution



Sentiment Analysis



After the pre-processing steps, sentiment analysis was developed by each review to see if it was positive, neutral, or negative using vaderSentiment

Sentiment Analysis category

Based on the sentiment score given by vaderSentiment, we categorized each review to be:

- Positive: above 0.05
- Neutral: between -0.05 and 0.05
- Negative: below -0.05

review_text_list	scores	sentiment
have ever anywhere will find hong kong similar...	0.7889	Positive
since last time visit hk .. yet time stay tomo...	0.9821	Positive
thanks god hot humid visit otherwise big issue...	0.9769	Positive
world anaheim tokyo feel hong kong small call ...	-0.3900	Negative



Positive



Neutral



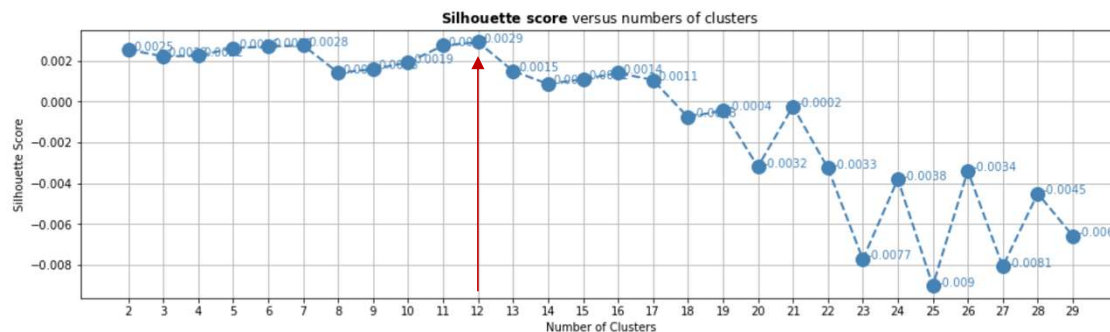
Negative

Topic Modeling



The Team selected the K-Means method to find the number of underlying topics in the reviews and then each review was categorized in only one of those topics

Topic Modeling Clusters and categorization



Review_ID	review_text_list	topic_number
0	670772142 have ever anywhere will find hong kong similar...	12
1	670682799 since last time visit hk .. yet time stay tomo...	4
2	670623270 thanks god hot humid visit otherwise big issue...	8
3	670607911 hk great compact unfortunately quite bit maint...	2
4	670607296 location city take around 1 hour kowlon kid li...	12

The team decided to have **12 clusters** since this number had the best silhouette score and relatively good distortion

Each review was categorized into **only of the 12 topics** based on the minimum Euclidean distance from the numerical representation of each review and the clusters center

Analysis and Visualization

The Power BI dashboards allow the company to visualize and analyze visitor's reviews and allow the company:

- Filter and **display information about a specific year, reviewer location and park**
- Reviews table** with review, sentiment and rating, **that allow user to sort by sentiment score and analyze the details of the review**
- Dashboard with **Sentiment analysis by topic** with the option to filter by park, year and topic
- Display information **about topics underlying negative reviews to find opportunity areas**



Business Summary

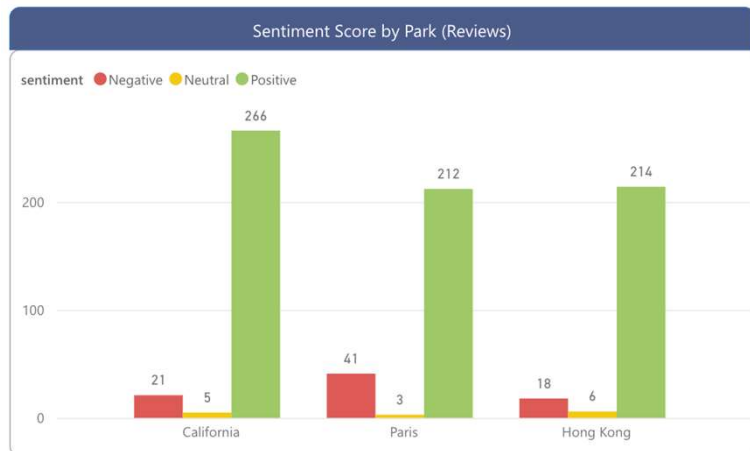


Main results from Sentiment Analysis

In 2019, California received the review with higher positive scores, while Paris the lower ones

Sentiment analysis score by review

- 1° California (0.72)
- 2° Hong Kong (0.69)
- 3° Paris (0.60)

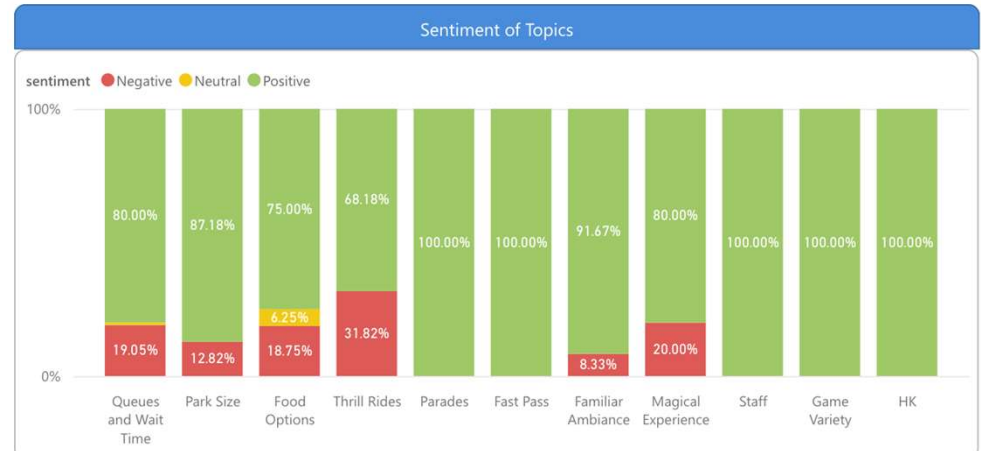


Main results from Topic Modeling

Given that Disneyland Paris has received negative reviews, it is important to identify the specific topics that visitors are complaining about so the company can take appropriate action to solve these issues

Comparison of topics with worst reviews in Disneyland Paris

Paris		California
1°. Thrill rides (32% negative)	vs.	14%
2°. Magical Experience (20%)	vs.	6%
3°. Queues & Wait times (19%)	vs.	11%



Estimated outcome and Business Value

Based on the statistics previously presented, is estimated that the software can create at least an impact of 2.5 M in profit and 200 M in revenue due to higher customer retention and greater spending

Business Value of the Solution

The **customer retention and visitor spending** in the parks, specifically for Disneyland Paris, **can be improved** if the company focuses on addressing and **solving the issues related to the topics with higher negative reviews** (presented previously)



Higher Customer Retention

- Disneyland Paris has an annual operating profit of USD 51 M^{/1}
- **Even if customer retention increase only by 1%, it is estimated that the profit would increase by 5%^{/2}, which represents USD 2.5 M**

Higher Visitor Spending

- Each year Disneyland Paris receives 12 M of visitors, each one spending an average of USD 110 per day and staying on average for 3 days^{/3}
- **Even if customer spending only increases by 5% due to customer satisfaction, the revenues (not profit) would increase by 200 M**

/1 Forbes

/2 Based on the statistic shown in the first slide, which states that an increase of 5% of customer retention represents a 25% increase in profit

/3 Themed Entertainment Association and AECOM