MTECH KE-5106 (DWBA) PROJECT REPORT

TripAdvisor – Japan Attraction Review Text Mining & Sentiment Analysis

TEAM MEMBERS

LOW KANG JIANG
HUANG FUXING
TAN HUI XIN
TEH PENG MOK
ZHANG HUAI PENG

MASTER OF TECHNOLOGY IN
KNOWLEDGE ENGINEERING
BATCH KE-29

1.0 EXECUTIVE SUMMARY

Going for an oversea trip to Japan is always an attractive option for Singaporean during the holidays. However, given that there are more than 100 of attractions across Japan, deciding which attractions is the best to visit is not as easy as it seems. They must spend a lot of time doing considerable amount of research before they arrive at the best decision.

TripAdvisor as the most popular travel review platform is always the primary source of information for prospective traveler. Given that there are more than 1,000 reviews for each attraction, it is always a challenge to understand the sentiment for each review without referring to the rating. It is always an interest for traveler to be able to extract a positive or negative sentiment reviews from the huge number of reviews without read through it.

Thus, we came up with the idea to conduct some data crawling and business analytics project to extract the key word and determine the sentiment score of each review submitted for attraction in Japan to shorten the time taken to carry out the research without compromise a good guidance on the decision-making process for traveler.

This report aim to provide a detailed analysis on the process of designing the data crawling and business analytics project. The first stage of the design process is data schema design in which several schema design considerations have been considered for one-to-one and one-to-many data relationship in the project. We used referenced and embedded method to store the data acquired depends on the frequency of data extraction and minimum storage space consideration. The schema design is important to ensure that it fulfills the scalability and flexibility best practices to ensure the effectiveness and optimality of data analysis in the later stage.

Then, we proceed with data crawling to extract all the useful data for each review of five attraction in Japan from TripAdvisor website according to the schema designed and store in MongoDB.

The next part of this report is to conduct business analytics on the unstructured data stored in MongoDB. The analysis can be separated into two parts in which the first part we have conducted text mining on the text extracted from all the reviews to determine the key phrases for each attraction. A word clouds for each attraction have been constructed using the *wordcloud* library available in Python to have a better visualization on the reviews.

In the second part of the analysis, we have conducted sentiment analysis for each of the comment and calculate their sentiment polarity score. The result shows that it is possible to extract the positive or negative sentiment reviews from the thousands of reviews with a high accuracy which is very valuable particularly for the traveler who are more interested in the negative sentiment reviews to gain more insight from it.

In conclusion, by constructing a well-defined data schema prior to the data crawling process and taking the data analysis methodology into account, it is efficient and flexible to analyze unstructured data compared to the traditional RDBMS.

2.0 BUSINESS OBJECTIVE

2.1 Current Challenges

Every backpacker must have heard about TripAdvisor. Almost every backpacker will use the platform to conduct travel research to decide which attraction to go. Many of the well-known attractions at Japan have been reviewed and rated by thousands of past travelers and it will be useful for prospective traveler who plan to visit Japan to be able to get some insights from the thousands rating and reviews submitted efficiently.

However, although the platform is resourceful, most users have difficulties to have an overview on the review comment or extract the negative comments from thousands of reviews submitted as there are far more positive comment as compared to negative comment.

In fact, many users agree that there are more preparation and consideration can be done from reading the negative sentiment reviews. As for the attraction management committee, negative sentiment reviews are valuable for them as it can provide a useful insight to assist in the decision-making process to improve the services and facilities of attractions. Given that there are thousands of reviews available and usually only a small amount of the reviews carried a negative sentiment polarity, it is always a challenge for both prospective travelers and attraction management committee to make use of the reviews for their interest.

Hence, the business objective of this project is to figure out the key phrases and understand the sentiment associated with traveler's satisfactions of the services provided by various attractions in Japan from their submitted reviews.

2.2 Proposed Solution

This project aim to scrape travelers' reviews from the TripAdvisor website and stored them as desired schema in MongoDB before proceed for further analysis. The reviews for five famous attractions in Japan for a period of a year has been scrapped. The five attractions are:

1. Fushimi Inari Shrine
2. Hiroshima Peace Memorial Park
3. Jigokudani Snow Monkey Park
4. Kiyomizu Dera
5. Tokyo Disneyland

Two parts of analysis have been conducted on the stored reviews for each attraction. The first part involves text mining on the reviews and extract the key phrases to construct a word cloud for each attraction. This allow prospective traveler to have a glimpse on the reviews on an attraction in Japan. Second part of the analysis involve the sentiment analysis in which a sentiment polarity score for each of the reviews has been calculated. This allow the prospective traveler to extract the negative sentiment reviews easily from thousands of reviews.

3.0 SCHEMA DESIGN AND DATA CRAWLING

3.1 Data Sources

The main data source for this project is TripAdvisor website.
(https://www.tripadvisor.com.sg/Attractions).

The review pages for five attractions have been used as seed url to scrape the necessary data as listed below.

1. https://www.tripadvisor.com.sg/ShowUserReviews-g298564-d321456-r516217425-Fushimi_Inari_taisha_Shrine-Kyoto_Kyoto_Prefecture_Kinki.html#REVIEWS
2. https://www.tripadvisor.com.sg/ShowUserReviews-g298561-d1165220-r514989321-Hiroshima_Peace_Memorial_Park-Hiroshima_Hiroshima_Prefecture_Chugoku.html#REVIEWS
3. https://www.tripadvisor.com.sg/ShowUserReviews-g1117904-d324924-r513787085-Jigokudani_Snow_Monkey_Park-Yamanouchi_machi_Shimotakai_gun_Nagano_Prefecture_Ch.html#REVIEWS
4. https://www.tripadvisor.com.sg/ShowUserReviews-g298564-d321401-r516024895-Kiyomizu_dera_Temple-Kyoto_Kyoto_Prefecture_Kinki.html#REVIEWS
5. https://www.tripadvisor.com.sg/ShowUserReviews-g298162-d320634-r516239476-Tokyo_Disneyland-Urayasu_Chiba_Prefecture_Kanto.html#REVIEWS

3.2 Schema Design

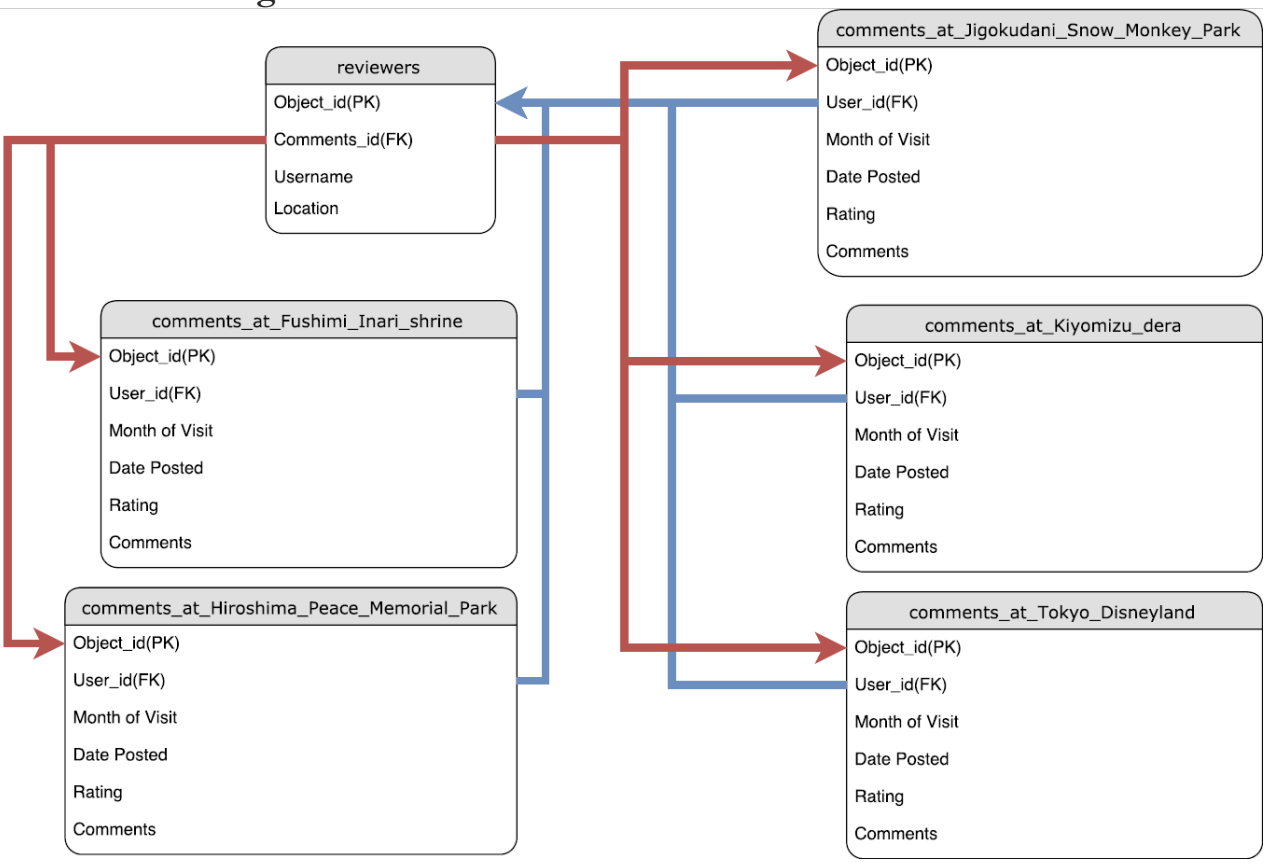


Figure 1: Data Schema and Relationship for TripAdvisor Japan Attraction Sentiment Analysis

Databases	Collection	Document	Field	Types
tripAdvisor	*reviewers*	reviewers	*_id*	String
			Comments_id	String
			location	String
			User name	String
	*comments_at_* + attraction attraction: • *Fushimi_Inari_Shrine* • *Hiroshima_Peace_Memorial_Park* • *Jigokudani_Snow_Monkey_Park* • *Kiyomizu_Dera* • *Tokyo_Disneyland*	comments	*_id*	String
			User id	String
			Title	String
			Month of Visit	String
			Date Posted	String
			Rating	String
			Comment	String

Table 1: Overview of MongoDB Database, Collections and Documents

Figure 1 show the schema design and the reference relationship between each of the collections in the *tripAdvisor* database. As shown in Table 1, tripAdvisor database contained two types of collections named *reviewers* and *comments_at_+attraction*, there are a total of five *comments_at_+attraction* collections as there are five attractions have been used for our analysis.

There are 4 fields in each of the document of *reviewers* collection. Other than the preassigned *object_id*, the *user name*, *location* and the *comments_id* which are used as a reference key to map to the *comments_at_+attraction* collection have also been stored in each of the document. There are a few schema design considerations that have been considered to come out with these designs to fulfill the data analytics requirement in the later part.

3.2.1 Referenced One-to-One Relationship from Comments to Reviewers


```
{
    "_id" : ObjectId("599bf8836253117c08bfe5c2"),
    "User Id" : ObjectId("599bf8836253117c08bfe5c3"),
    "Title" : "“Nice place to visit ”",
    "Month of Visit" : "August 2017",
    "Date Posted" : "20 August 2017",
    "Rating" : "5",
    "Comment" : "We got up early in the morning to avoid the crowd. The Senbon Torii is a good
        place to take pictures. Most of the places is covered by trees, so it is not too hot. We
        walked all the way to the top. As we left the shrine, there seems to be more people. The
        advice is to visit this place as early as possible when there are less people to make it
        easier to take pictures."

}
{
    "_id" : ObjectId("599bf8836253117c08bfe5c4"),
    "User Id" : ObjectId("599bf8836253117c08bfe5c5"),
    "Title" : "“Very nice”",
    "Month of Visit" : "August 2017",
    "Date Posted" : "20 August 2017",
    "Rating" : "4",
    "Comment" : "Kyoto is a bit hotter than tokyo . August may not be the best time to visit .The
        shrine is one of the main attractions in kyoto, taking a bus drom kyoto station is not
        hard, we took the express 105 bus, which takes you directly to the gate of the shrine and
        viceversa. Calm and beautiful environment. There are many kimono rentals nearby, costs
        around 5000~7000 yen. Be prepared to wear a comfortable shoes as it is a bit of a hike
        around ."
}
...
```


Figure 2: Example of documents in *comments_at_+attraction* collection

Figure 2 shows the bson view of the *comments_at_+attraction* collection. Each of the document in *comments_at_+attraction* collection can be referred to one reviewer in the *reviewers* collection using the stored *User Id* as a reference key.

The reason such design is for text mining and sentiment analysis, analyst need to frequently extract the review content related info such as *Comment*, *Rating*, *Title*, *Month of Visit* and *Date Posted* but not so for the reviewer related info such as *user name* and their base *location*.

Analyst can conduct text mining and sentiment analysis on *Title* or *Comment*. The *Rating* field for each document can be used to calculate the correlation between the overall sentiment polarity score and rating the *comments_at_+attraction* collection in the analytics stage.

Month of Visit and *Date Posted* are two imported filed which can allow analyst to select documents within a period to conduct the analysis.

Hence, in terms of optimality, it is more appropriate to store the reviewer info in another collection as shown in Figure 1.

3.2.2 Referenced & Embedded One-to-Many Relationship from Reviewers to Comments

```
{
    "_id" : ObjectId("599bf8836253117c08bfe5c3"),
    "User name" : "Brenda L",
    "Location" : "Calgary, Canada",
    "Comments_Id" : [ ObjectId("599bf8836253117c08bfe5c2") ]
}
{
    "_id" : ObjectId("599bf8846253117c08bfe5cf"),
    "User name" : "James H",
    "Location" : "New York City, New York",
    "Comments_Id" : [ ObjectId("599bf8846253117c08bfe5ce"), ObjectId("599bf8836253117c08bfe5c6")]
}
...
```

Figure 3: Example of documents in *reviewers* collection

Figure 3 shows the bson view of the *reviewer*s collection. Reviewer related data such *user name* and *location* for each review have been extracted from TripAdvisor website and stored in another collection named *reviewer*s instead of *comments_at_+attraction* collection. *comments_id* for each document in *reviewers* collection have been used as a reference key to map the reviewers back to their comment in *comments_at_+attraction* collection.

In terms of scalability and flexibility, storage of reviewers in another collection allow analyst to conduct analysis on the reviewer level. For instance, sentiment analysis can be conducted on a reviewer with a *User name* of "*James H*" on all the comments he/she submitted for different attraction. The used of reference *Comments_id* in each of the reviewer document ensure that the comments do not have repetition storage in the same database as the comments has already been stored in *comments_at_+attraction* collection.

Notice that the *Comments_id* is stored as an embedded field value pair in the documents. The reason is because during the sentiment analysis of a reviewer, analyst must frequently retrieve all the comments that he/she submitted, hence it is more optimal to embed each comment id in the *Comments_id* field. The schema design of reviewer collection has taken the benefit of both referenced and embedded methodology according to analytics requirement.

3.2.3 Reviewers Collection Mutable Array Problem Consideration

Nevertheless, one of the fallback of such schema design is the mutable array problem. When the number of attraction to be analyzed in the future increase and hence increase the number of comments stored in database, the *comment_id* of *reviewers* document may grow inevitably. However, our discussion concluded that the number of comments per reviewer is small with limiting growth. As we use reference mythology to store the *Comments_id*, the storage does not consume a lot of memory and it carried more business value to conduct sentiment analysis for each reviewer based on their base location in the future. In consideration of the future analytics, the comments for each reviewer then be stored in the *reviewers* documents with an embedded reference *Comments_id*.

In conclusion, the scheme design allow analyst to conduct text mining or sentiment analysis on both attractions level (across different *comments_at_+attraction* collection) and reviewers level (across different *reviewers* document). Analyst can simply add in a new seed url for other attraction in the crawler code to extract the reviews and store as a new collection in the database. *comments_at_+attraction* collection can grow unbounded whereas there is only one *reviewers* collection with clear naming as shown in Figure 1 to ensure scalability of the schema. The storage of reviewer related data in another *reviewers* collection allow analyst to conduct sentiment analysis on reviewer based ensure the flexibility of the schema to meet different analytic objectives.

3.3 Data Crawling

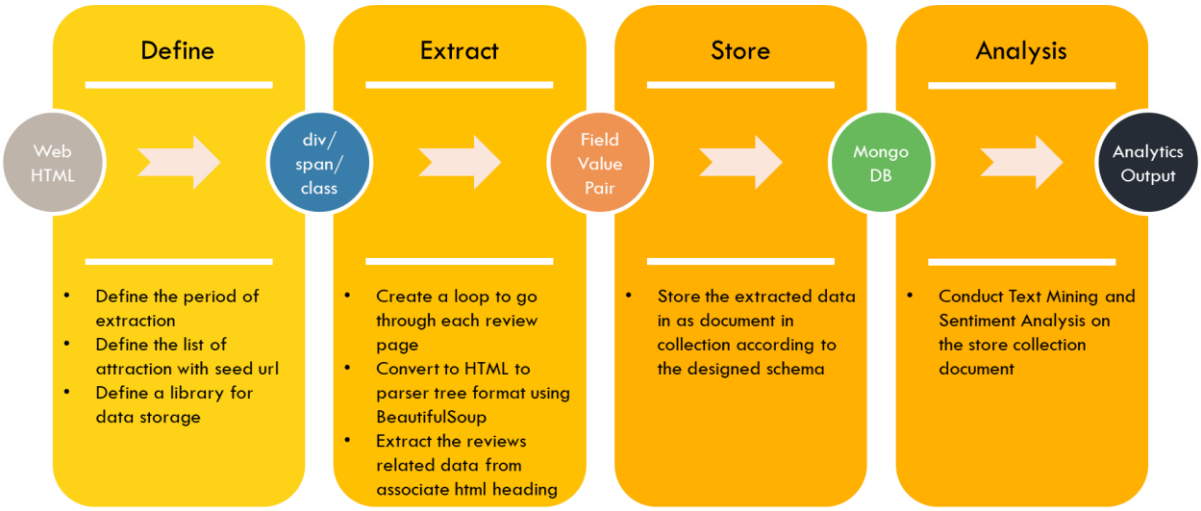


Figure 4: Data Scraping, Storage & Analysis for TripAdvisor Japan Attraction Review Sentiment Analysis

Using the schema designed as a reference, we proceed with the data scraping. The scraping process has been summarized in Figure 4. We defined the range of the date to be extracted to be *365* days and define a list (*JapanAttractionListAndSeedURL*) to store the seed *url* for each attraction review. The definition provides flexibility and scalability for analyst in which the period of reviews extracted and attractions of interest can be modified depends on the business requirement. Analyst can simply add in the attraction with the seed *url* into the predefined list to construct new collection in the database.

During the crawling process, one of the main challenge is to seek for the next page of the review page as the reviews usually span across a couple of pages. Hence, we locate the next *url* heading from the html file and use it to update the *url* at the end of each scraping cycle until the one year period reviews condition was met.

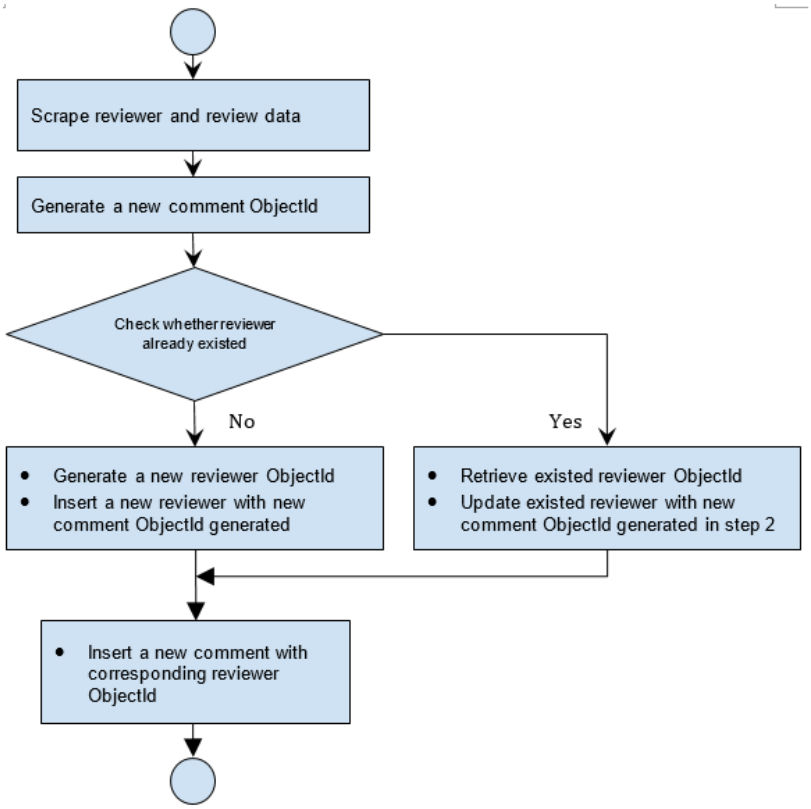


Figure 5: Data Scraping flow chart for TripAdvisor Japan Attraction Review

3.3.1 Reviewer Username and Location Extraction

Reviewer *User Name* and *Location* can be extracted directly from the class "*member_info*". However, some reviews have missing *User Name* and *Location*. Hence, we use a condition check to detect such missing value fields and assigned *anonymous* and *unavailable* tag for *User Name* and *Location* fields in the *reviewers* collection accordingly. Then, the *User Name* and *Location* can be extracted from class "*username mo*" and class "*location*" within the class "*member_info*".

3.3.2 Review Title Extraction

Review *Title* can be extracted from class "*quote*" directly without extra process.

3.3.3 Review Rating Extraction

Review *Rating* was stored indirectly in the class "*rating reviewItemInline*" as "*ui_bubble_rating bubble_x0*" where *x* represent the integer value of the rating. Thus, *Rating* can be extracted by extracting the *x* string and stored as an integer value.

3.3.4 Review Date Posted Extraction

Date Posted is a very imported field as the value is used as a variable for condition check of period of extraction. One of the main challenge when extracting *Date Posted* attribute is the html class heading that store the data changed from "*ratingDate relativeDate*" to "*ratingDate*" when the review is older than a certain period. Hence, a condition check is used to extract the necessary *Date Posted* data.

3.3.5 Review comments extraction

Review *Comment* can be extracted from header "*p*" directly without extra process.

3.3.6 Reviewer Visited Month

Reviewer *Month of Visit* to the attraction can be extracted from the class '*recommend-titleInline noRatings*' without extra process.

3.4 Data Backup & Data Restore

Finally, to ensure the maintainability and recoverability of database, we backup all the data scraped from the website using '*mongodump'* package available in MongoDB console. The dumped file can be used as a backup for all the data collected as well as to transfer the date from one server to another server.

Data Backup

Mongodump --db tripAdvisor --out dump\tripAdvisorDatabase

Data Restore

Mongorestore --db tripAdvisor --dir dump\tripAdvisorDatabase

4.0 BUSINESS ANALYICS

4.1 TripAdvisor Attraction Review Text Mining

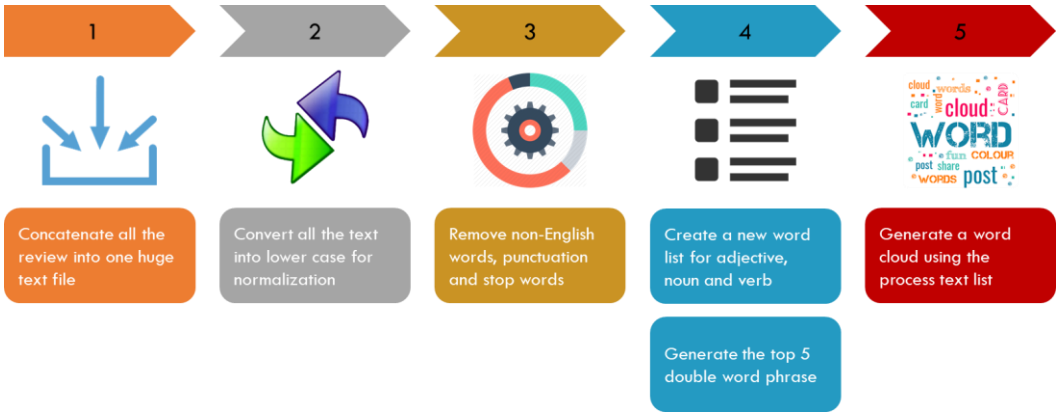


Figure 5: Text Mining data processing on the Review Comments for each comments collection

In the first stage of analysis, several information has been extracted using the reviews data stored in the TripAdvisor database as listed below:

1. The rating distribution of the attraction for the past one year
2. The top 5 common word used in the reviews (overall, adjective, noun, verb)
3. The top 5 doubles word used in the reviews
4. The top 5 doubles word used in the reviews with 5-star rating
5. The word cloud of the word used in the reviews

The processes to conduct the text mining and construct the word cloud using the comments for each *comments_at_+attraction* collection have been illustrated in Figure 5. The detailed result can be referred to Appendix A.

As the rating and comments data used in the analytics has been stored in the field value pair in MongoDB, it is relatively easy and directive to conduct the necessary analysis on the comments field. The common word, double words and word cloud generated can provide insights for traveler on the attraction such as:

1. Which tourist spot in the attraction is worth to visit? (from the 5-star rating double words)
   Eg. Crystal Palace for Tokyo Disneyland, Shiga Koen for Jigokudani Snow Monkey Park
2. What are the common object that can be seen when visit the attraction? (common noun)
   Eg. Gates and Shrine for Fushimi Inari Shrine, Museum for Hiroshima Peace Memorial Park
3. What is the common feeling associated with that tourist attraction? (common adjective)
   Eg. Peaceful and beautiful for Hiroshima Peace Memorial Park
4. Interesting findings:
   Eg. May need to bring mosquito repellant when visit Fushimi Inari Shrine and be aware of the mobility issue when visit to Jigokudani Snow Monkey Park.

4.2 TripAdvisor Attraction Review Sentiment Analysis

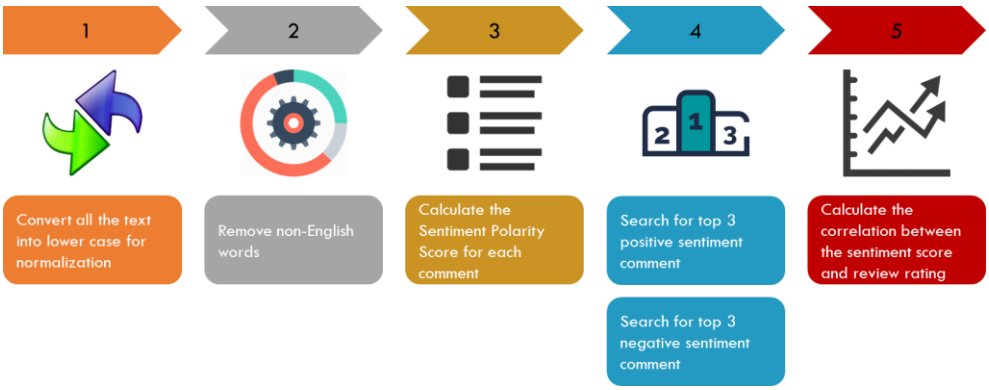


Figure 6: Sentiment Polarity Score calculation process for each comment document

In the second part of the analysis, the sentiment polarity score for each of the comment has been calculated as illustrated in the Figure 6. The main difference between the analysis in part 2 as compared to part 1 is we do not remove the punctuation and stop words as these elements have been used as a factor to calculate the sentiment polarity score.

A Python library named "*textblob*" has been used to calculate the sentiment polarity for each comment and the detailed result can refer to Appendix A. Besides, the correlation of sentiment polarity and review rating also have been calculated for each attraction and the result shows that other than the Hiroshima Peace Memorial Park, the rest of the attraction analyzed have a positive correlation between these two metrics although the correlation is not significant. The reason of low correlation score is due to the fact the user who rated 1-star or 2-star most probably have a negative sentiment toward the attraction and hence the rating need to be normalized from 1 to 5 scale to -1 to 1 for better verification on the exact correlation in the future.

	Comment	Date Posted	Month of Visit	Rating	Title	User Id	_id	Processed Comment	Sentiment Polarity
5	If you love Disney you will really love Tokyo ...	13 August 2017	August 2017	5	"Awesome fun in Tokyo"	599bfb756253117c08c0005a	599bfc356253117c08c0058e	if you love disney you will really love tokyo ...	0.346894
6	ดีมากกกกกกกกกกก ฟินนน เติมเต็มความฝัน Disney l...	12 August 2017	May 2017	5	"อยู่ตั้งแต่พาร์เปิดจนปิด"	599bfc356253117c08c00590	599bfc356253117c08c0058f	disney lover buzz light year winni...	0.300000
7	Of course it's the happiest place on earth! I ...	12 August 2017	July 2017	5	"So fun!"	599bf8896253117c08bfe603	599bfc356253117c08c00591	of course its the happiest place on earth i lo...	0.400000

Figure 7: Sentiment Polarity Score calculation for each comment

Figure 7 shows the two new fields with processed text and sentiment polarity for each of the comment. The not English word with index 6 has been removed after processed to ensure the accuracy for sentiment polarity. By calculating the sentiment polarity score for each of the comment, analysts can easily filter and extracted the comments with certain polarity score to be further analyzed.

As shown in Appendix A, the top 3 negative comments from Tokyo Disneyland indicated that there was a long queue in the theme park and hence the traveler who do not like to queue can avoid to visit this extraction. Besides, it also provides insights for the theme park management committee to improve the queuing process and crowd control management of theme park to enhance the service satisfaction among the visitors.

Nevertheless, the sentiment polarity score does not always accurate as can be seen in the top negative sentiment comment for Fushimi Inari Shrine and Hiroshima Peace Memorial Park, the reason lied in the fact that these comments consist of negative sentiment terms such as *bad*, *forget*, *horrifying* and *devastated* but it does not describe the attraction. However, as most of the real negative sentiment comment can be correctly classified, it is still can provide analyst with valuable insights.

6.0 CONCLUSION

In conclusion, in an era where there are a lot of unstructured data available on the open source website such as TripAdvisor, it is worth to leverage on the flexibility and scalability of NoSQL data modelling methodology to deal with these unstructured data as compared to traditional RDBMS.

By designing the data schema per our analytics usage, analysts can conduct text mining and sentiment analysis on the reviews submitted by reviewers for five Japan attraction for a period of one year. The analysis provides insights for prospective traveller on the sentiment of the attraction and allow them to extract the negative sentiment comment easily from the reviews.

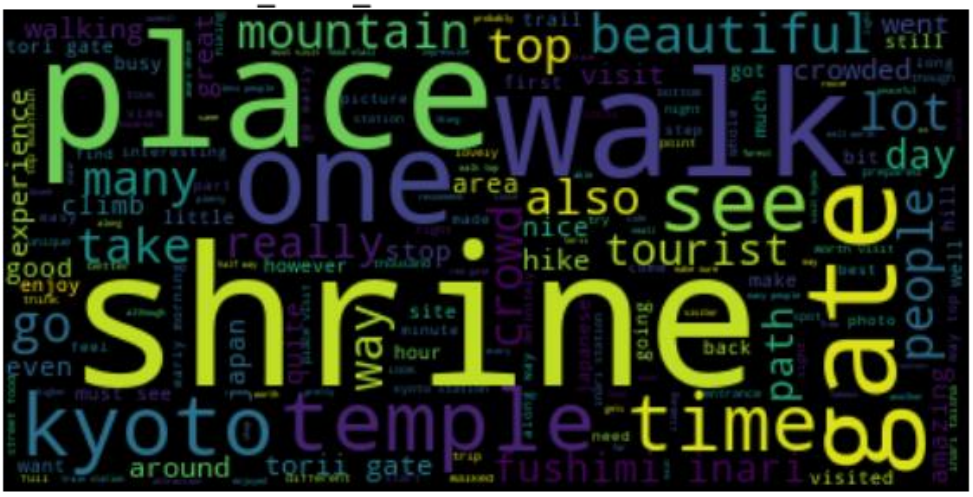
The TripAdvisor review text mining and sentiment analysis shows that it is very simple to extract and store the unstructured data from web. The unstructured data can then be used for analysis easily with a proper schema design to bring valuable insights for end user and decision maker. As such, the data schema need to be well defined in consider with the frequency of use and storage availability before extracting and storing the data into MongoDB.

APPENDIX A: BUSINESS ANALYTICS OUTPUT

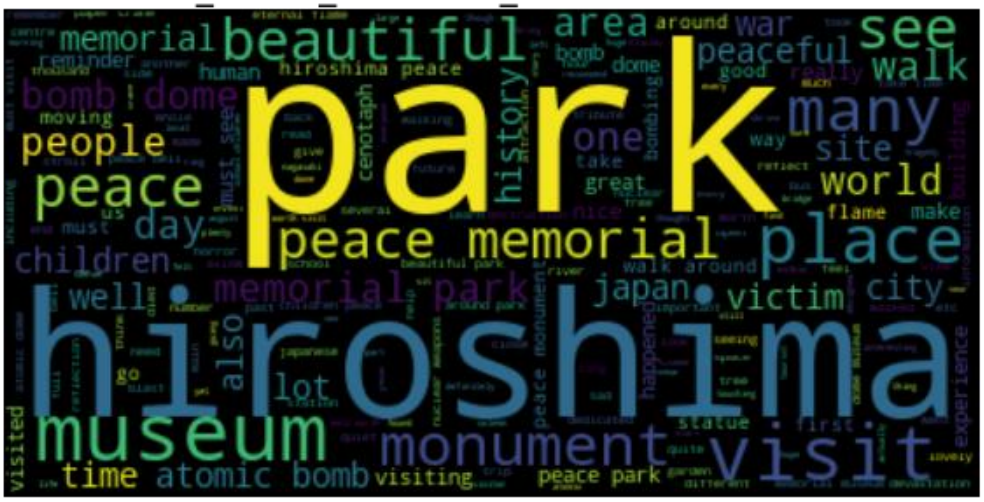
1. TripAdvisor Attraction Review Text Mining

Attractions	Rating Distributions	Top 5 double words phrase	Top 5 double words phrase with Rating 5	Top 5 Common Word	Top 5 Common Adjective	Top 5 Common Noun	Top 5 Common Verb
Fushimi Inari Shrine	1 star: 8 2 stars: 21 3 stars: 142 4 stars: 729 5 stars: 1885	Oinari San Tofu Steak Mosquito Repellant Public Transportation Faint Heated	Oinari San Tofu Steak Mosquito Repellant Public Transportation Faint Heated	Shrine Gates Kyoto Top Walk	Top Many Beautiful Great Good	Place Way Gates Shrine People	Go Get See Take Walk
Hiroshima Peace Memorial Park	1 star: 0 2 stars: 0 3 stars: 14 4 stars: 73 5 stars: 253	Sister Cities Mobilized Students Cherry Blossom Rest House Cherry Blossoms	Hand Accounts Cherry Blossom Rest House Cherry Blossoms Jr Pass	Park Peace Memorial Hiroshima Museum	Many Atomic Beautiful Peaceful Good	Park Peace Museum Bomb Place	See Walk Visiting Moving Visit
Jigokudani Snow Monkey Park	1 star: 7 2 stars: 6 3 stars: 18 4 stars: 78 5 stars: 266	Yaen Koen Mobility Issues Zen Hotel Shiga Koen Roman Museum	Shiga Koen Mobility Issues Bucket List Roman Museum Leisurely Pace	Monkeys Park Snow Walk Monkey	Hot Great Good Many Worth	Monkeys Park Bus Walk Monkey	Get See Walk Go Took
Kiyomizu Dera	1 star: 3 2 stars: 11 3 stars: 64 4 stars: 296 5 stars: 556	Shita Kiri Kodai Ji Selfie Sticks Tainai Meguri Pitch Black	Shita Kiri Pitch Black Golden Week Single Nail Public Transport	Temple Kyoto Walk Visit Beautiful	Beautiful Great Many Nice Worth	Temple Kyoto Place View Time	See Go Walk Crowded Take
Tokyo Disneyland	1 star: 18 2 stars: 16 3 stars: 33 4 stars: 148 5 stars: 310	Donald Duck Crystal Palace Astro Blasters Indiana Jones Universal Studio	Crystal Palace Ice Cream Indiana Jones Toon Town Peter Pan	Disneyland Disney Rides Park Time	Japanese Fast Great Many Good	Disney Rides Disneyland Time Park	Go Get Went See Take

Fushimi_Inari_Shrine Reviews Word Cloud

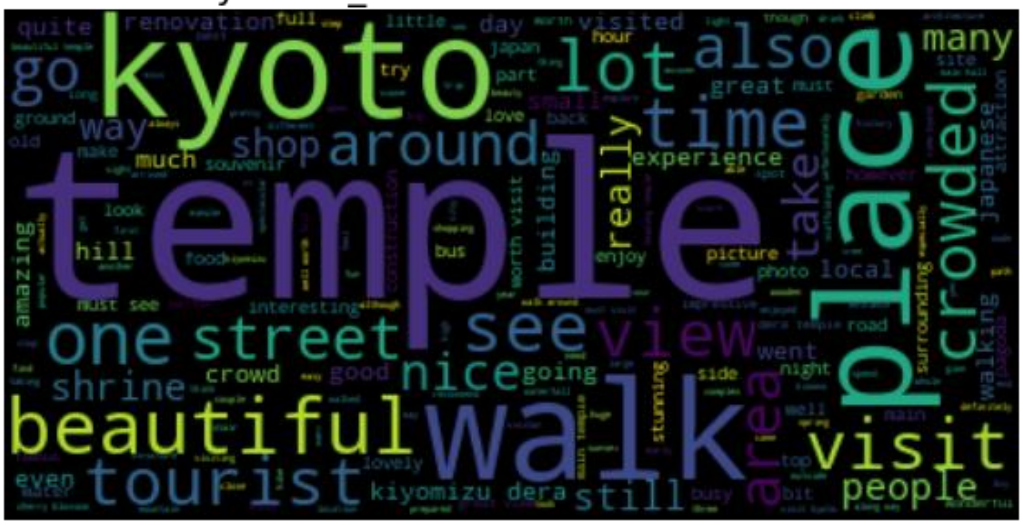


Hiroshima_Peace_Memorial_Park Reviews Word Cloud

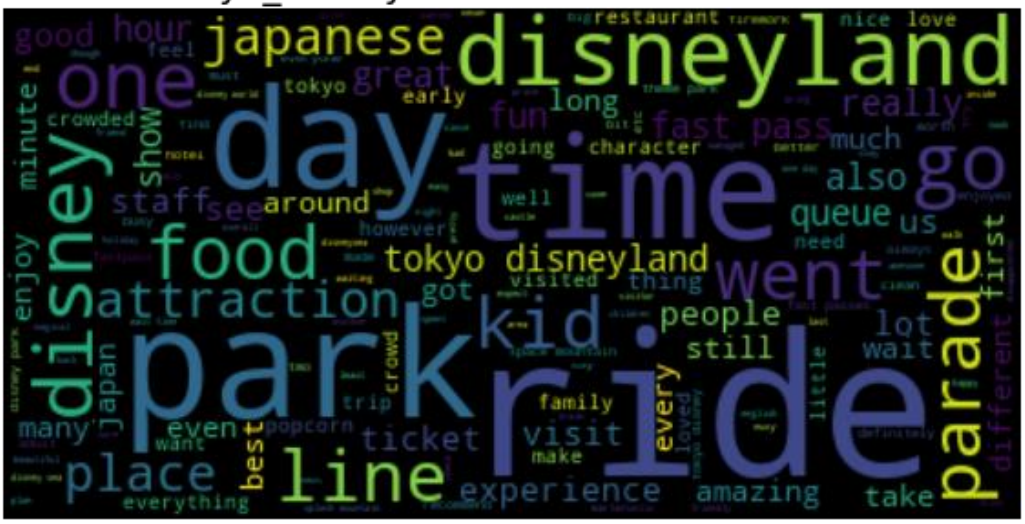


Jigokudani_Snow_Monkey_Park Reviews Word Cloud

Kiyomizu_Dera Reviews Word Cloud



Tokyo_Disneyland Reviews Word Cloud



TripAdvisor Attraction Review Sentiment Analysis Overview

Attractions	Rating Distributions	Sentiment Polarity
Fushimi Inari Shrine	1 star: 8 2 stars: 21 3 stars: 142 4 stars: 729 5 stars: 1885	Max: 1.0 Min: -0.7 Mean: 0.28 Median: 0.27 Rating Correlation: 0.14
Hiroshima Peace Memorial Park	1 star: 0 2 stars: 0 3 stars: 14 4 stars: 73 5 stars: 253	Max: 0.85 Min: -0.3 Mean: 0.26 Median: 0.24 Rating Correlation: 0
Jigokudani Snow Monkey Park	1 star: 7 2 stars: 6 3 stars: 18 4 stars: 78 5 stars: 266	Max: 1.0 Min: -0.31 Mean: 0.29 Median: 0.28 Rating Correlation: 0.19
Kiyomizu Dera	1 star: 3 2 stars: 11 3 stars: 64 4 stars: 296 5 stars: 556	Max: 1.0 Min: -0.31 Mean: 0.30 Median: 0.28 Rating Correlation: 0.19
Tokyo Disneyland	1 star: 18 2 stars: 16 3 stars: 33 4 stars: 148 5 stars: 310	Max: 0.9 Min: -0.3 Mean: 0.26 Median: 0.22 Rating Correlation: 0.39

TripAdvisor Attraction Review Sentiment Analysis(Positive)

Attractions	1st Most Positive Sentiment Review	2nd Most Positive Sentiment Review	3rd Most Positive Sentiment Review
Fushimi Inari Shrine	Really a wonderful attraction, you decide if to spend 10 mins or 2 hours walkind in neverending torii tunnels up on the hill.	The best place to visit!!! Prepare your feet! You'll be walking a lot! Lots of inclining up toward the Mt. Lots of food vendors and souvenirs shop.	The number of Torii gates that's spread throughout the shrine is impressive to see and to walk through. It could be quite a hike to walk through all the gates to reach the end and return..
Hiroshima Peace Memorial Park	The memorial park is a must if you are visiting the museum or the dome. It has been put together beautifully to allow you to remember those impacted and reflect on the impact of the a-bomb.	It's solemn and a good place to contemplate. Autumn is the best time to visit, seeing reds and yellows in the area. Afterwards you can visit the shopping street - Hondori, it's just walking distance from the park.	walking from the museum to the dome through the park brings about flashes showing leaders signing documents on TV and visiting the location for memorials etc.feels GREAT to be park of this walk and park.it brings you to your senses.
Jigokudani Snow Monkey Park	Its a great experience to have when you go to Nagano... walking up the trail during winter was a calming experience for me... would recommend this activity...	A beautiful nature walk through the forest that climaxed upon arrival to the park where the snow monkeys rule. An awesome experience! Remember not to get too close and stare at the monkeys when taking pictures. By doing this you may be a threat to them. I couldn't stop laughing when my daughter's hand was swatted by one of the snow monkeys with a smelly hand! Be sure to visit this park.	Very nice - great place to visit. Fascinating lifetime experience. We have been there for two days and enjoyed every minute there.
Kiyomizu Dera	Salut, hi!Trop de monde !!! Sinon, oui, c'est pas mal du tout...Take a breath... 1 billion people between you and the temple... Here you are; it is lovely.PS: Not the best to me here in Kyoto	We took a taxi from Rengeo-in temple to Kiyomizu-dera. The taxi saved us a lot of time because the road is steep. The temple was crowed. The temple was wonderful and the views were beautiful. Recommended.	One of the must-do places to visit in Kyoto. Take bus or subway from Kyoto station and its a 10 minute walk from the bus stop. The gallery view is great and is best during cherry blossom season and winter, during snowfall.
Tokyo Disneyland	This is my fifth time in Tokyo Disneyland. And it is always a place that will bring out the child in everyone. I think Springtime is the best time to go to Tokyo.Better to check-in at one of their hotel to avail of their "happy 15". This will give you an advance 15 minutes to enter the park.	If you are visiting Tokyo, don't miss the Disneyland. It's a happy place for everyone! Staff are nice, rides are awesome!	Is very nice of a place.....if have a time i will think go again❄i will choose winter season if go

TripAdvisor Attraction Review Sentiment Analysis(Negative)

Attractions	1st Most Negative Sentiment Review	2nd Most Negative Sentiment Review	3rd Most Negative Sentiment Review
Fushimi Inari Shrine	Don't forget to think about your bad habits and behaviour during climbing the hill. And don't forget to think about your positve way to improve it! This is an order to the visitors!!	This place is a must see in Kyoto and in Japan! Spirituality and nature merge for an experience difficult to forget!	For all the hype, and the insane amount of other people there I was disappointed with the experience. Maybe I've become numb to the amount of shrines in Kyoto since I left it last.
Hiroshima Peace Memorial Park	Going around this park gave me goosebumps. To think this was totally devastated during WWII. Praying that such horrifying event will never happen again!	It is part of the peace complex which includes the museum, the park, the A-bomb dome and the cenotaph.	If you visit Japan, you must travel to Hiroshima and this Memorial Park. To walk there on this little island, the museums and statues is moving.
Jigokudani Snow Monkey Park	Sadly the snow had melted and the monkeys were not interested in bathing but there were plenty to see. Slightly touristy. Much quicker to get express tourist bus rather than the train-train-bus route which is slow.	What an experience. What I was not aware of was the length of the hike in. It was about one and a half kilometres with both uphill and downhill on an earth path. Majestic cedar forest.	it\'s a main tourist attraction, hence the crowds. It\'s a 4 km hike from base point. There is a distinct "aroma". Must go to tick it off the list.
Kiyomizu Dera	To Kiyomizu-Gojo, Take Keihan Line to Kiyomizu-Gojo [KH38] Exit 4 and walk to the Temple for 1.5 KM or Take Taxi but it's quite expensive due to traffic jam. If you want to wear Kimono, you can stop by shops along the way to the temple.	Never ever drive here. Drove around the narrow crowded streets for a bit looking for parking lots which we didn't know which we could or could not park, ended up parking at one which cost us 1600 yen for some 30-40 minutes we went! That's crazy.Super Crowded. Lots of slopes and stairs to climb, not suitable with elderly or toddlers.	Unfortunately we arrived and they were doing work on the outside of the structure. The road leading up to the temple is narrow and steep and look out for cars coming down. Have loads of small shops on the way to have a look at.
Tokyo Disneyland	Terrible experience wait 2 hours for every attraction! Then lunch time you wait for over an hour for chicken nuggets, to find out you cant get a seat! The food is horrible on top of that..... honestly i dont understand why people go here. Sat on the stage but got kicked off, so ate standing up.I reccomend you spare yourself and your kids from this horror!	Visited with our 2 kids today, 28 December. Terrible. Queues were ridiculous. We only got 3 rides in. Even queues for popcorn were 30 minutes. There were clearly too many people in the park. Very disappointed with Disney who are clearly prioritising revenue over customer service. Avoid.	At about 1pm you are told that there is no more fast pass. You then have to queue for hours for rides that wouldn't tame a puppy. Millions of people all dressed in Disney crap. Food is appalling. Stores are all the same. Products are poor and highly priced. However, did not see any obese people in wheelchairs advancing to front of queue as in LA. Why oh why did we bother? Walt would turn in his grave! Dissapointed visitor