

## Kuredu TripAdvisor Review Analysis

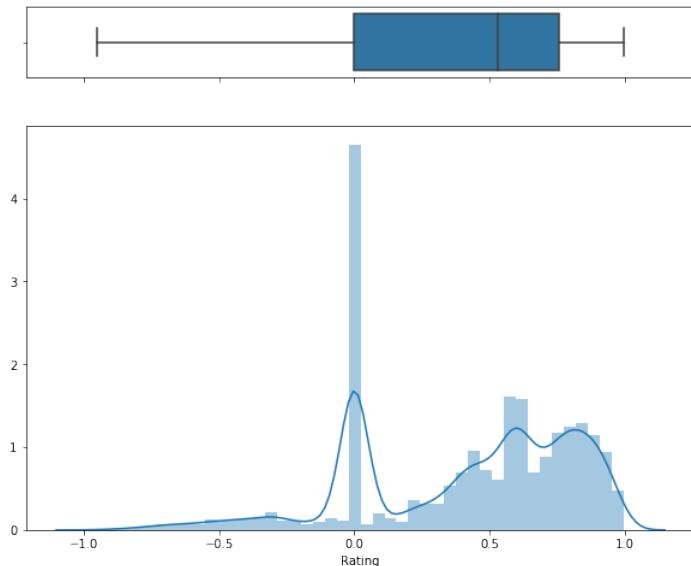
Reviews were scraped from TripAdvisor on 20/04/2020, and a total of 7,475 reviews were retrieved. Utilising aspect keywords, sentences within each review were separated out into aspects – Value, Rooms, Location, Cleanliness, Service, Facilities/Activities and Other. Sentiment Analysis was then used to generate a sentiment score for each of these sentences. Scores are within the range of [-1, 1], being:

- [-1 to 0]: Negative
- [0]: Neutral
- (0 to +1]: Positive

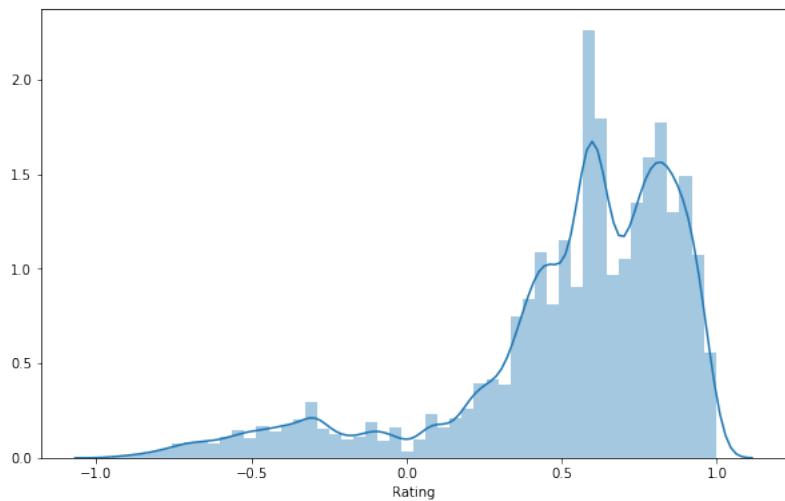
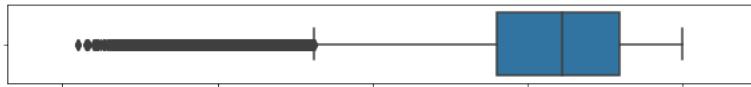
### 1 Reviews overall

Using this method, the average rating overall is 0.417 (positive). When transformed into the range 1-5 (as on TripAdvisor), the average score is 3.543; this is below the average rating of 4.5 as shown on TripAdvisor, suggesting refinements to the sentiment scoring algorithm are required to account for things like double negatives, negative words used in a non-negative manner (e.g. ‘to die for’) etc. However, it could also be considered that where reviewers may provide a single score of 5 to represent their overall holiday, they may include sentences within their reviews that are negative (or not 5 star in nature) that did not warrant a lowering of the overall score by the reviewer, but could provide suggestions for improvements to the resort, operations, staffing or other element of the overall experience.

The distribution of reviews overall is as follows:



This shows a concentration of sentences within reviews classed as neutral; this makes sense as many reviewers will include informative sentences to flesh out their reviews, which do not convey any sort of opinion/sentiment. These neutral sentences will be removed from the dataset when analysing aspects, as the purpose is to identify, analyse and glean lessons from guest opinions, rather than identify what guests consider as factual information. Removing these sentences leaves a total of 31,653 sentences, with a mean of 0.526 (3.815 in the range 1-5), and the following sentiment distribution:



## **2 Aspect Analysis**

For each aspect, the mean and distribution of sentiment will be reported, followed by the results of topic modelling. Topics generated do not equate to sentiment (i.e. topics generated are based on keywords that appear in the sentences, so will not necessarily be all positives or all negatives; each topic will likely contain positive and negative reviews on the same theme).

Topics are helpful in understanding what guests' opinions are of different themes related to their holiday. Topic modelling of the results of review sentence aspect rating analysis enables the drawing out and analysis of themes within aspects – value, location, rooms, facilities/activities, service, other – for more in-depth insight into guest experience and opinion. Further analysis could entail isolating the negative reviews from each topic within each aspect in order to gain very specific insight into areas for improvement.

For topic modelling, the following information and visualisations are provided:

1. Table containing keywords characterising each topic and a review sentence that is representative of each topic. Representative text is of the processed reviews, so is simplified with common words removed and words in list form
2. Topic keyword tree: displays the most important keywords per topic, with intensity of colour indicating the weights of keywords for each topic (higher the intensity, higher the weight/importance of a given keyword).
3. Word cloud: displays the keywords per topic, with word size indicating keyword weight per topic
4. Bar plot: plots frequency of the keywords within the reviews and the weights of those keywords within topics.
5. Interactive topic plot (included as separate html files): the circles in the figure represent different topics, the size of the circle represents the importance of the topic, and the larger the circle indicates that the topic is more important for the data. If the circles overlap each other, the topics they represent have similarities. Hovering over the circle on the interactive version brings up the important keywords that are representative of that topic on the bar chart on the right, indicating the overall term frequency (blue) and the estimated term frequency within the selected topic (red); the more red the bar, the heavily concentrated that term is within reviews that fall into the topic being investigated. The numbers given to topics in the plot below are not the same as those for topics given in other diagrams.
6. Based on topic visualisations and keywords, each topic is characterised in summary, stating subjectively what the review sentences for each topic within that aspect relate to. These summaries can then be used to selectively target certain themes within an aspect (if, for example, guest opinions on waiting staff are of interest, the service aspect would be the focus, with service aspect topic 0 identified as most relevant to waiting staff, and so those would be the reviews that could be isolated)

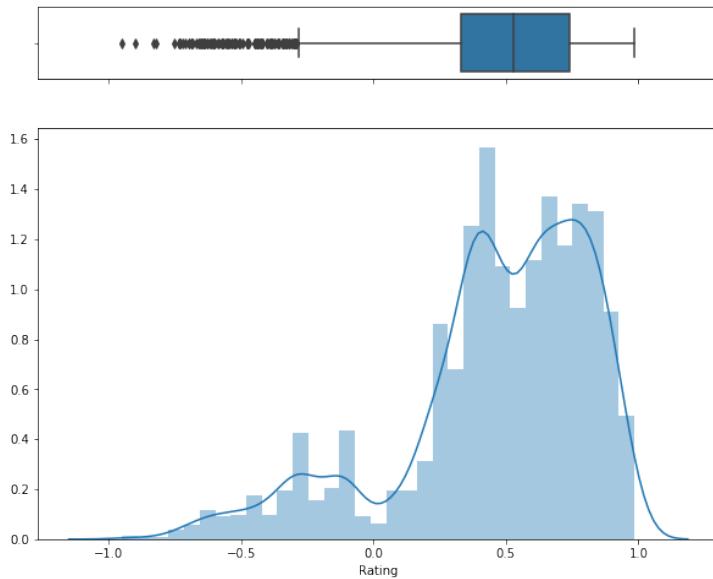
## 2.1 Value

### 2.1.1 Sentiment

Total sentences categorised as 'Value': 2,656

Mean sentiment rating (mean in 1-5 range): 0.461 (3.654)

Distribution of sentence sentiment ratings for value:



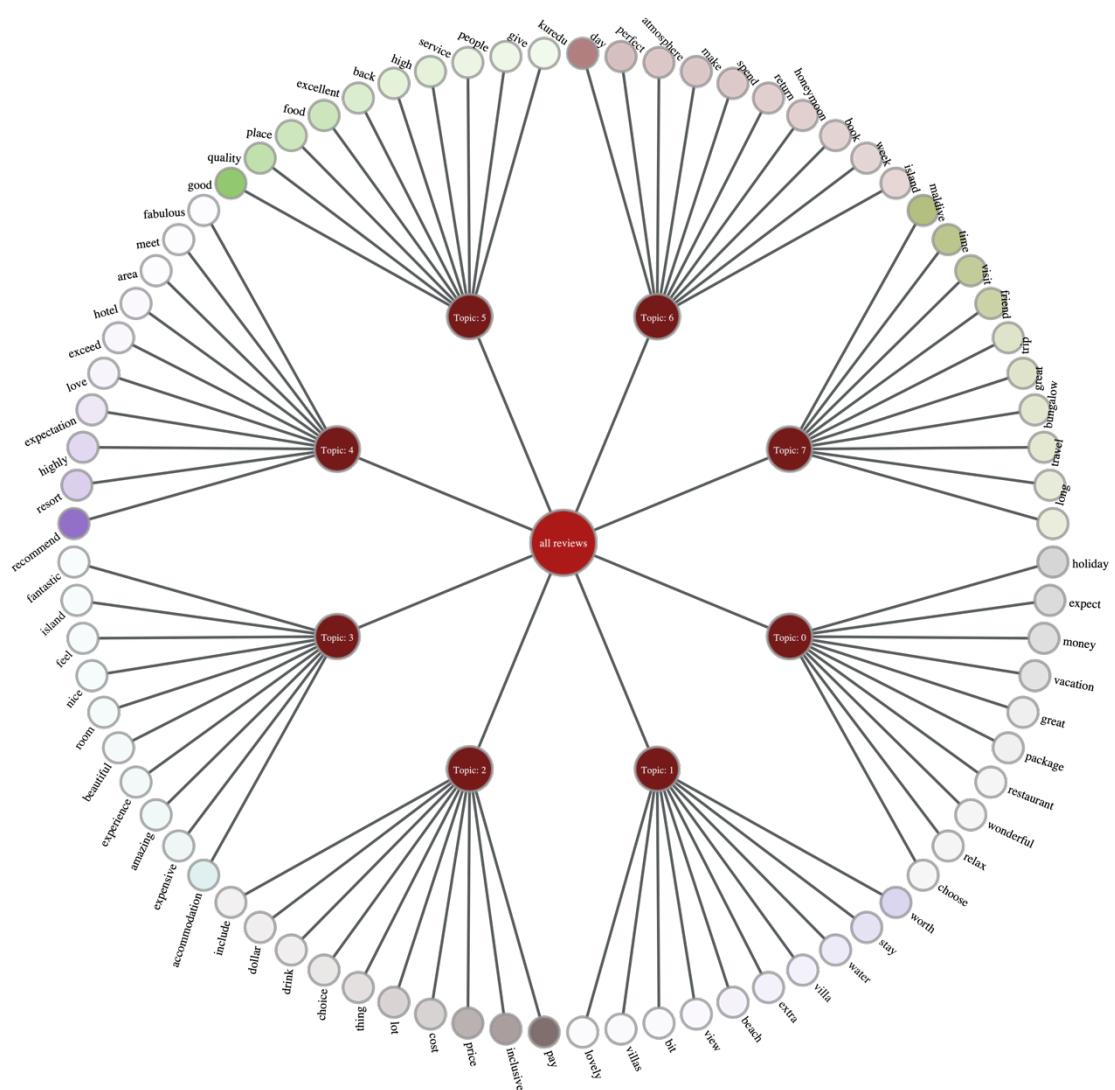
### 2.1.2 Topic modelling

Number of topics: 8

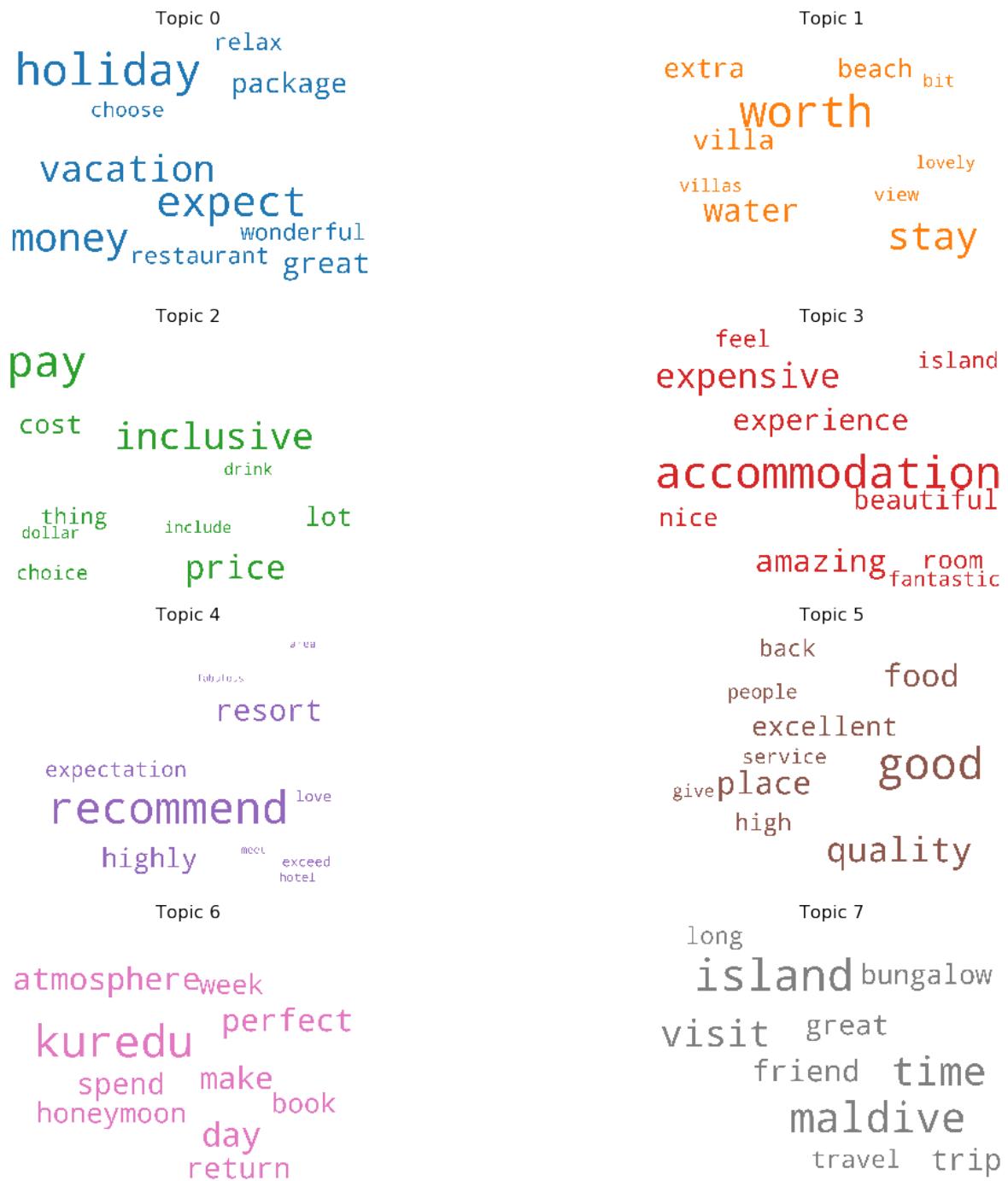
Reviews representative of each topic:

Topic Num	Topic Perc Contrib	Keywords	Representative Text (processed review)
0.0	0.2012	holiday, expect, money, vacation, great, package, restaurant, wonderful, relax, choose	['unfortunately', 'seem', 'monsoon', 'season', 'start', 'little', 'early', 'year', 'however', 'resort', 'offer', 'many', 'activity', 'rain', 'get', 'way', 'relax', 'enjoyable', 'vacation']
1.0	0.25	worth, stay, water, villa, extra, beach, view, bit, villas, lovely	['stay', 'water', 'lovely', 'definately', 'recommend', 'pay', 'bit', 'extra', 'water', 'villa', 'villa']
2.0	0.2992	pay, inclusive, price, cost, lot, thing, choice, drink, dollar, include	['beach', 'bungalow', 'basic', 'need', 'need', 'lot', 'update', 'outside', 'toilet', 'shower', 'crawlie', 'night', 'get', 'dark', 'lot', 'bulde', 'work', 'go', 'inclusive', 'leave', 'lot', 'desire', 'pay', 'gold', 'inclusive', 'option', 'resteraunt', 'rest', 'pay']
3.0	0.221	accommodation, expensive, amazing, experience, beautiful, room, nice, feel, island, fantastic	['info', 'apology', 'length', 'accommodation', 'stay', 'adequate', 'comfortable', 'describe', 'luxurious', 'say', 'least', 'expensive', 'island', 'yet', 'facility', 'available', 'activity', 'return', 'sleep', 'time', 'evening', 'nee', 'pay', 'expensive', 'accommodation']
4.0	0.2227	recommend, resort, highly, expectation, love, exceed, hotel, area, meet, fabulous	['would', 'highly', 'recommend', 'pack', 'pair', 'short', 'hand', 'luggage', 'change']
5.0	0.2244	good, quality, place, food, excellent, back, high, service, people, give	['good', 'quality', 'well', 'prepare', 'varied', 'food', 'bulking', 'platter', 'chicken', 'thigh', 'wing', 'good', 'lean', 'breast']
6.0	0.2244	kuredu, day, perfect, atmosphere, make, spend, return, honeymoon, book, week	['go', 'home', 'day', 'hope', 'see', 'future', 'thank', 'driver', 'never', 'forget', 'accommodation', 'number', 'especially', 'cleaner', 'else', 'make', 'husband', 'birthday', 'special']
7.0	0.2422	island, maldives, time, visit, friend, trip, great, bungalow, travel, long	['take', 'notice', 'man', 'moan', 'could', 'take', 'cheap', 'plonk', 'country', 'slate', 'alnclusive', 'island', 'perfect', 'romantic', 'place', 'world', 'cheap', 'price', 'compare', 'maldives', 'island', 'will', 'better', 'alnclusive', 'perfect', 'food', 'staff', 'good', 'world', 'have', 'twice', 'would', 'go', 'island', 'ever', 'place', 'go', 'go', 'maldives', 'get', 'better', 'go', 'never', 'stop', 'go', 'back', 'make', 'dream', 'come', 'true', 'think', 'best', 'take', 'notice', 'extra', 'go', 'elsewhere', 'afford']

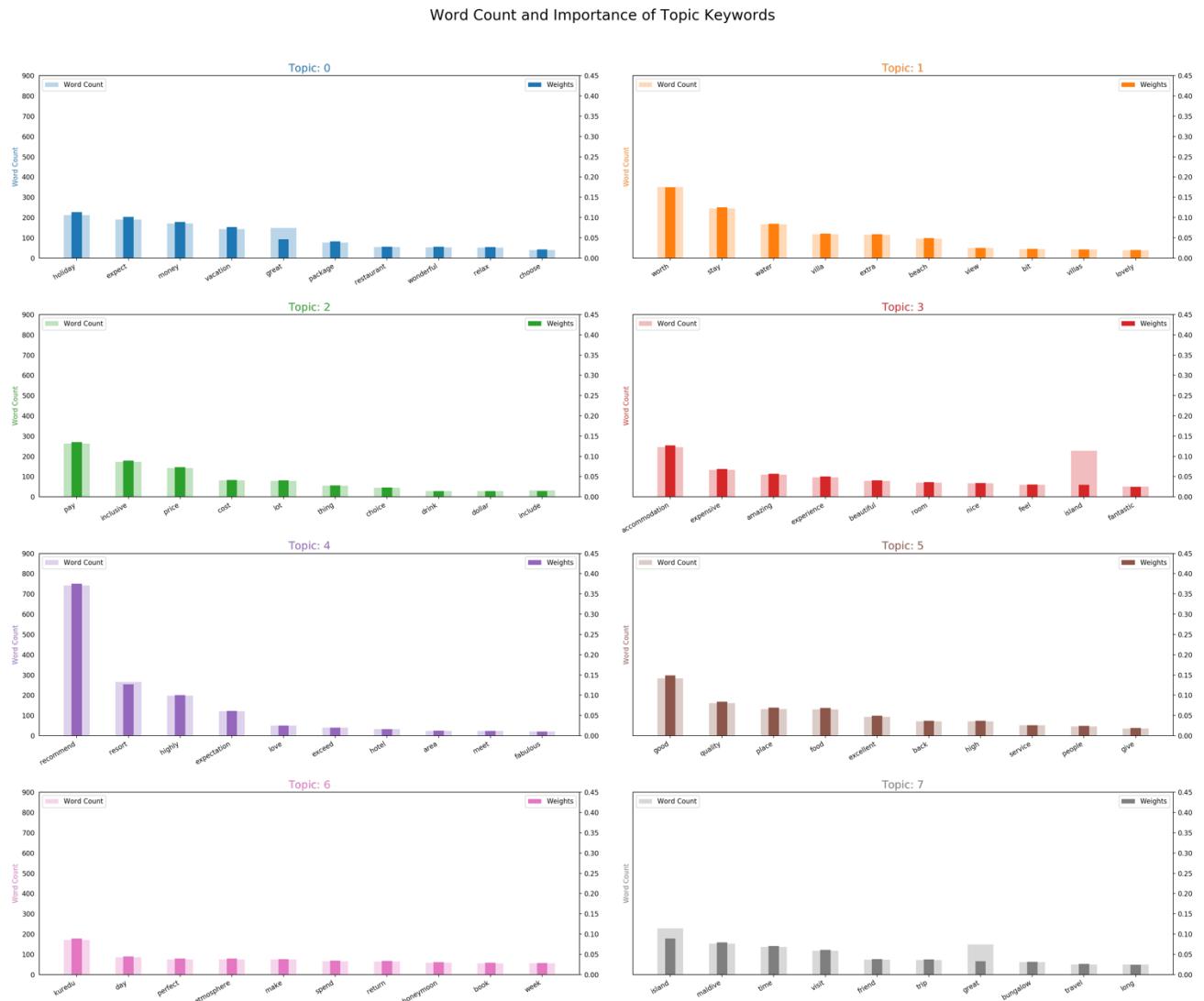
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



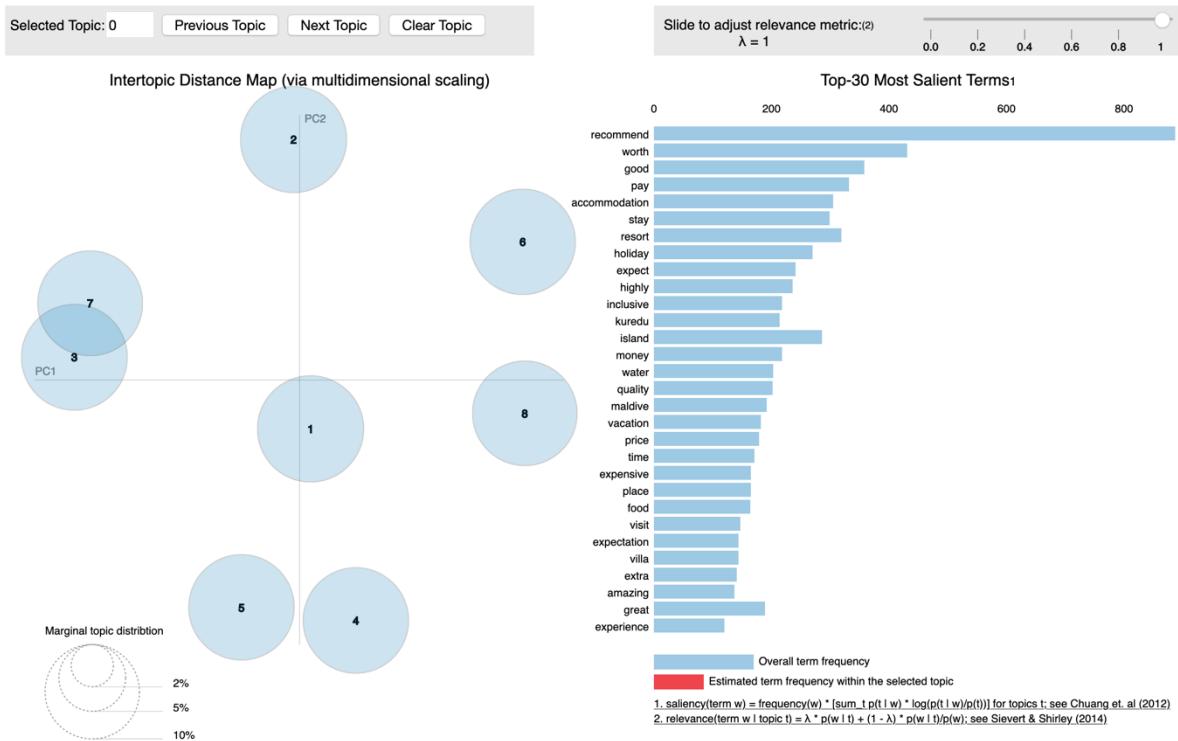
**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



## Example of interactive plot as static screen shot:



### Value topics:

- Topic 0: mentions of package suggest this topic is of guests who put value on what is included in their package
- Topic 1: mentions of worth and recommend, so reviewers who consider their holiday good value for money; they consider value as monetary
- Topic 2: mentions of inclusive, gold and discount suggest these sentences would be worth analysing to find out more about guest views of packages currently offered by the resort; similar to topic 0, but more focused on all-inclusive guests than other packages
- Topic 3: mentions of expensive but also positive words such as amazing and beautiful suggest reviews from guests for which this is a once in a lifetime trip so do not necessarily consider it good value, but worth the expenditure; they consider value as the location
- Topic 4: words of exceed and expectation suggest reviewers who consider it very good value for money; they consider value as something that fulfils promises (what they believe a holiday in the Maldives should provide)
- Topic 5: words of quality, food and service suggest reviewers who consider value as what is provided by the resort itself; useful reviews to analyse to gain insight on staffing practices and provisions
- Topic 6: honeymoon and atmosphere mentioned a lot, suggesting that they would put value on the 'paradise' theme portrayed by the idea of a holiday in the Maldives
- Topic 7: family and friend amongst keywords, as well as mentions of children in the reviews, suggesting they put value on the company they have with them and what is available for them to do together

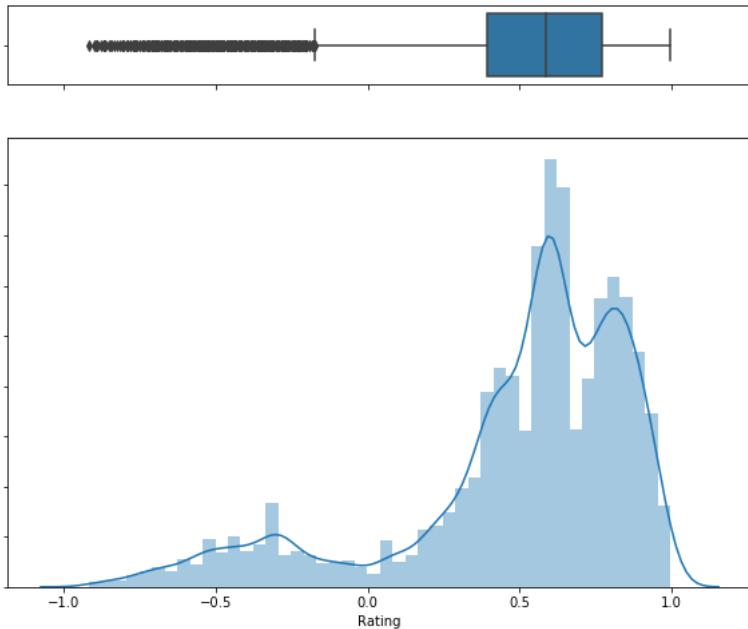
## 2.2 Rooms

### 2.2.1 Sentiment

Total sentences categorised as 'Rooms: 5,712

Mean sentiment rating (mean in 1-5 range): 0.460 (3.650)

Distribution of sentence sentiment ratings for room:



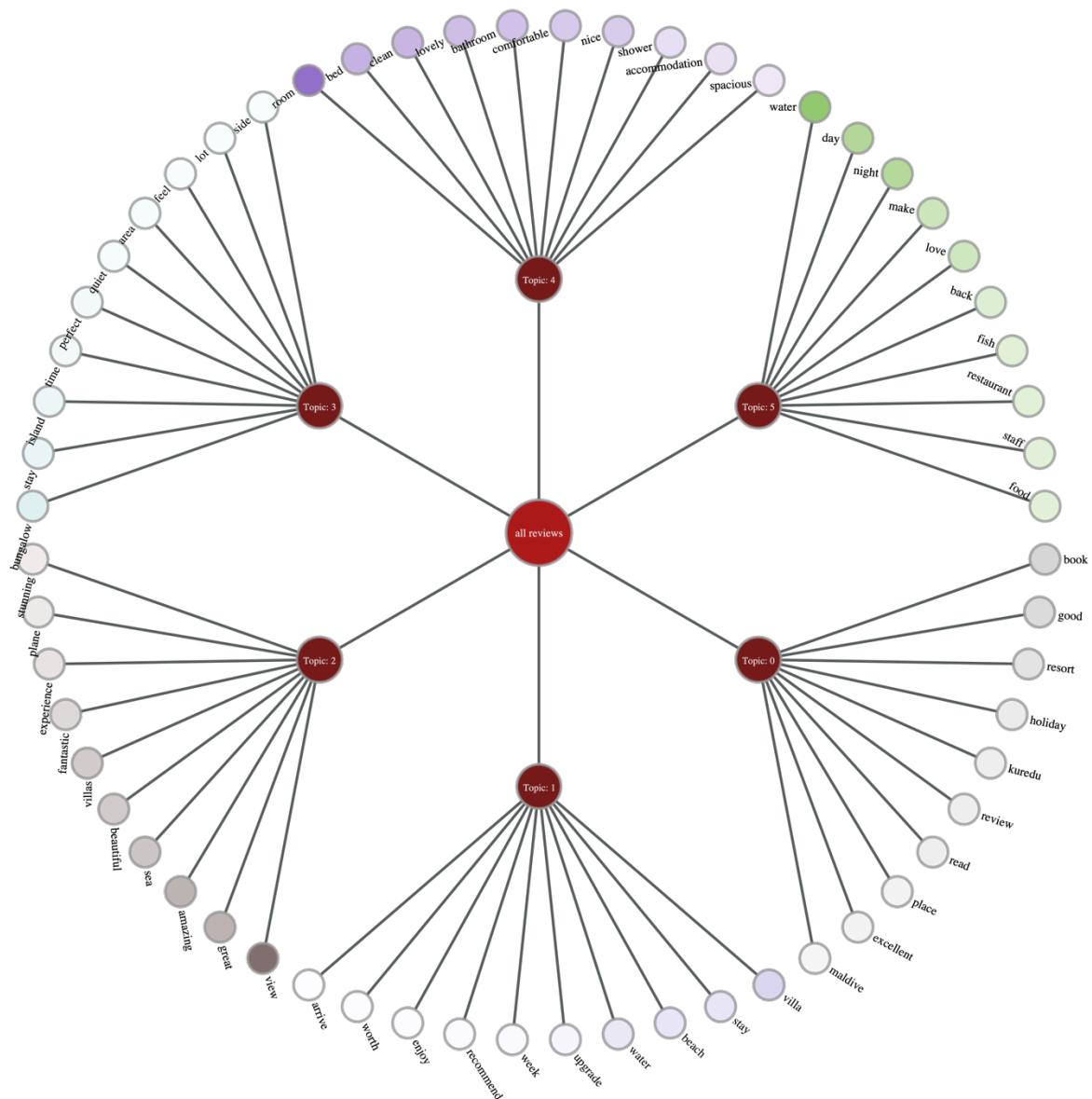
### 2.2.2 Topic modelling

Number of topics: 6

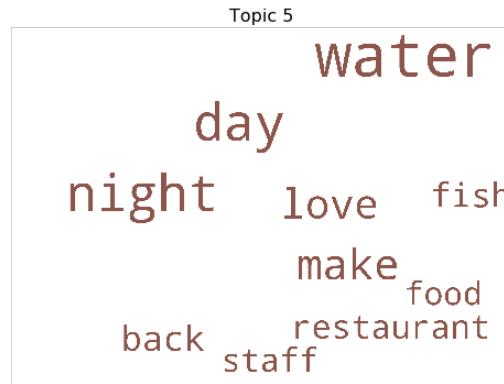
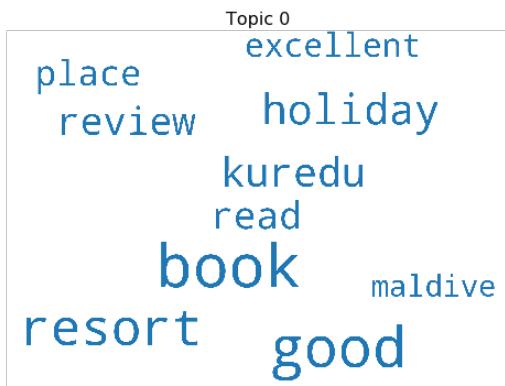
Reviews representative of each topic:

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text (processed review)
0.0	0.36	book, good, resort, holiday, kuredu, review, read, place, excellent, maldive	['know', 'want', 'make', 'reservation', 'honeymoon', 'look', 'best', 'good', 'week', 'marriage', 'read', 'review', 'decide', 'go', 'book', 'point', 'decide', 'send', 'direct', 'mail', 'kuredu', 'ask', 'regard', 'room', 'request', 'direct', 'booking', 'guest', 'apply', 'agent', 'booking']
1.0	0.2857	villa, stay, beach, water, upgrade, week, recommend, enjoy, worth, arrive	['closer', 'end', 'week', 'also', 'offer', 'free', 'upgrade', 'water', 'villa', 'decide', 'stay', 'grateful', 'offer', 'case']
2.0	0.3134	view, great, amazing, sea, beautiful, villas, fantastic, experience, plane, stunning	['reach', 'transfer', 'plane', 'male', 'kuredu', 'experience', 'little', 'noisy', 'take', 'earplug', 'help', 'unbelievable', 'prepare', 'wow', 'forget', 'camera', 'easy', 'reach', 'capture', 'moment']
3.0	0.3005	bungalow, stay, island, time, perfect, quiet, area, feel, lot, side	['prefer', 'meet', 'people', 'find', 'other', 'main', 'area', 'prefer', 'quiet', 'time', 'easily', 'find']
4.0	0.3123	room, bed, clean, lovely, bathroom, comfortable, nice, shower, accommodation, spacious	['strange', 'use', 'outside', 'bathroom', 'first', 'time', 'soon', 'get', 'use', 'benefit', 'keep', 'steam', 'bedroom', 'other', 'speak', 'problem', 'issue', 'tile', 'shower', 'area', 'screen', 'water', 'soap', 'get', 'splash', 'rest', 'bathroom', 'floor', 'tile', 'non', 'slip', 'become', 'skate', 'rink', 'solution']
5.0	0.287	water, day, night, make, love, back, fish, restaurant, staff, food	['day', 'birthday', 'bed', 'decorate', 'fresh', 'flower', 'well', 'table', 'dinner', 'birthday', 'cake']

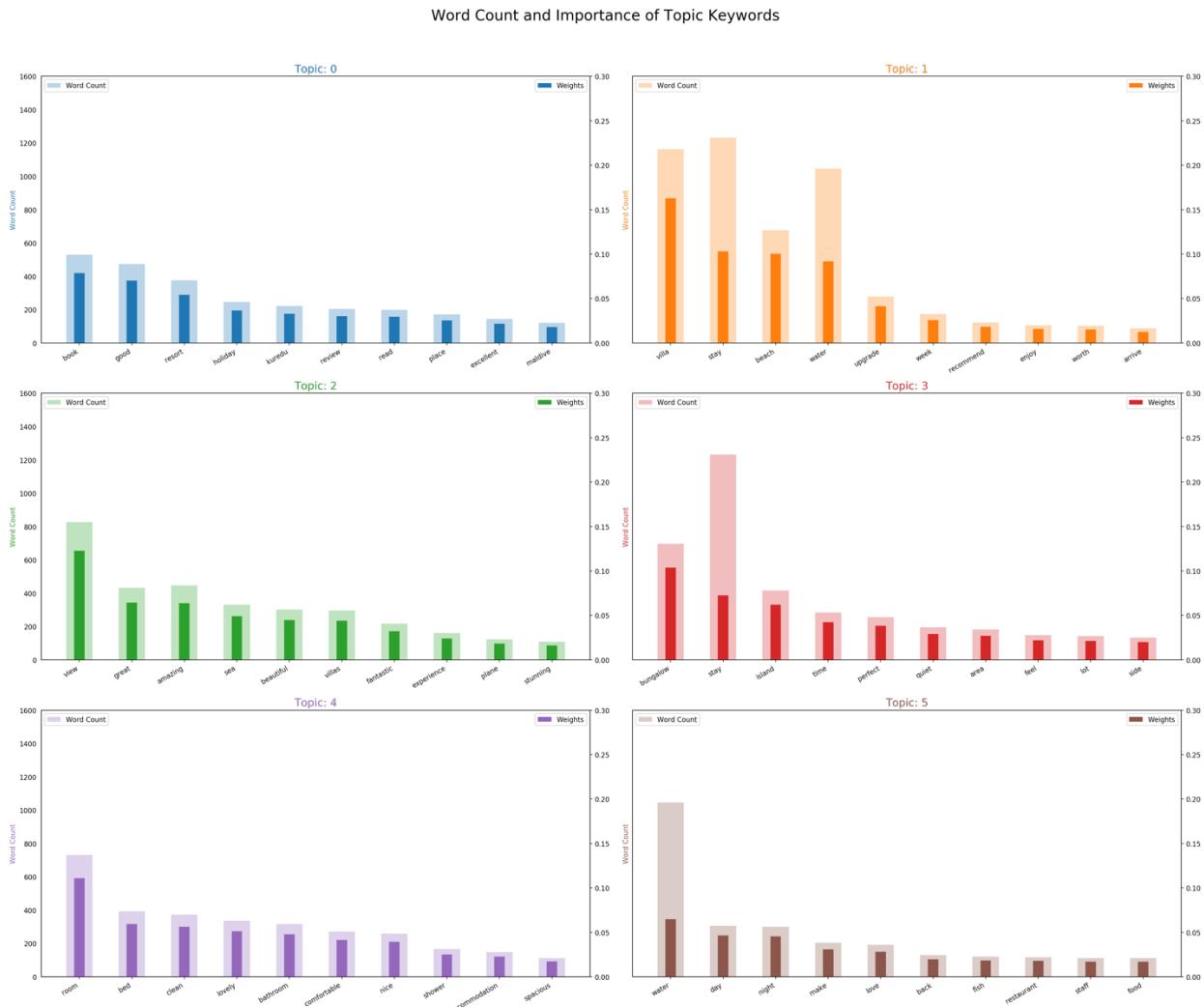
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



The interactive plot for rooms is saved as rooms\_topics\_pyldavis.html

### Room topics

- Topic 0: sentences that appear linked to making reservations and booking
- Topic 1: review sentences related to upgrade
- Topic 2: review sentences related to the view
- Topic 3: review sentences related to bungalows
- Topic 4: review sentences linked to room quality, room cleanliness and room provisions and facilities
- Topic 5: this topic doesn't appear to be very related to the rooms, so should perhaps be disregarded

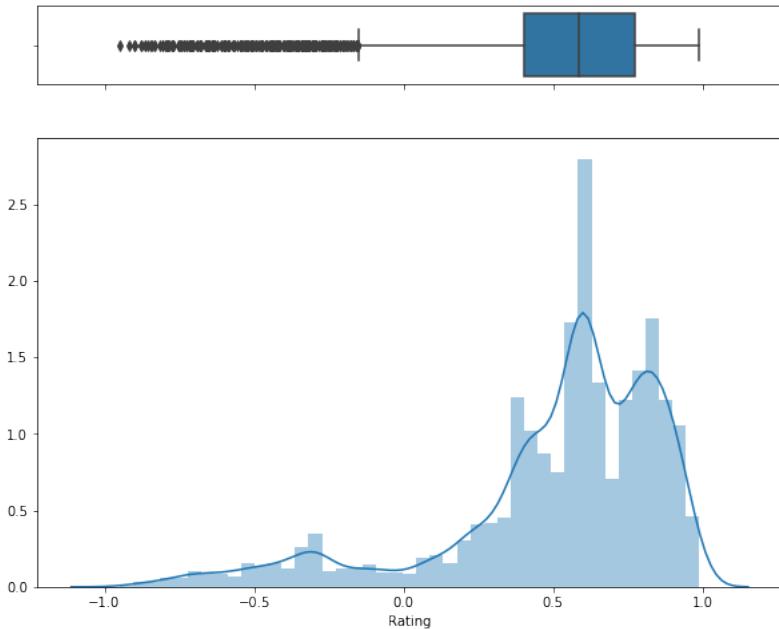
## 2.3 Location

### 2.3.1 Sentiment

Total sentences categorised as ‘Location: 4,369

Mean sentiment rating (mean in 1-5 range): 0.506 (3.765)

Distribution of sentence sentiment ratings for location:



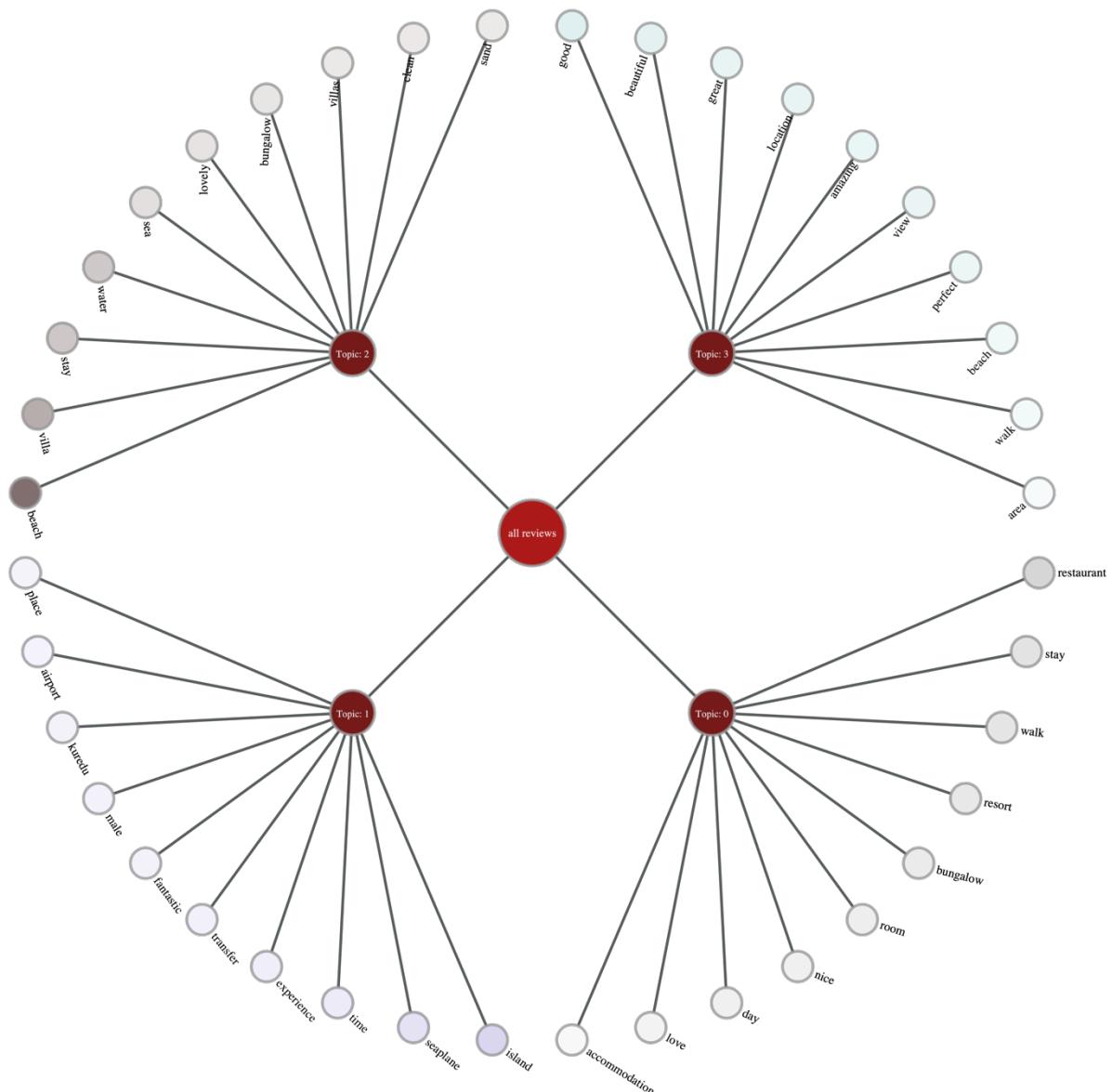
### 2.3.2 Topic modelling

Number of topics: 8

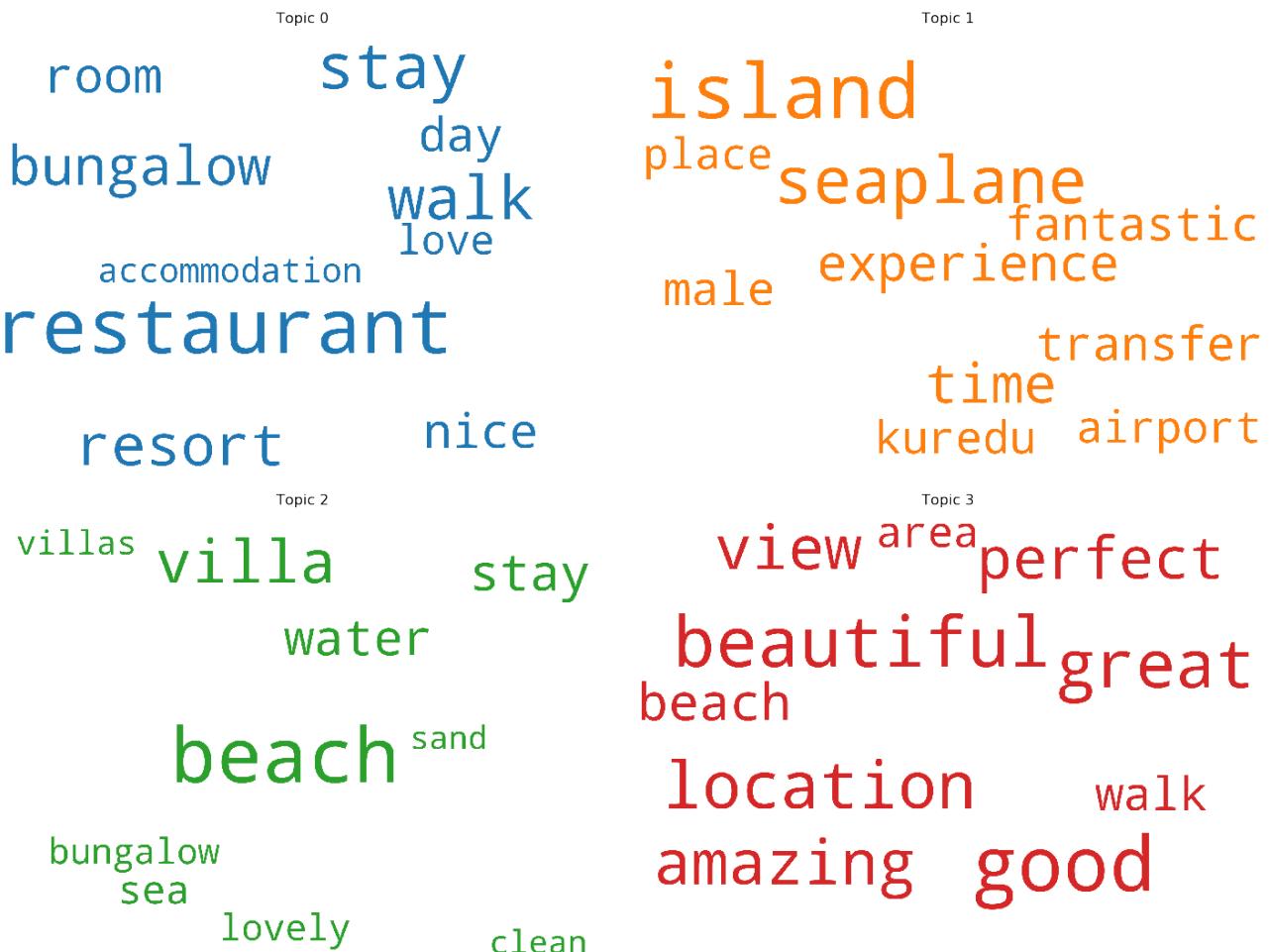
Reviews representative of each topic:

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text (processed review)
0.0	0.3786	restaurant, stay, walk, resort, bungalow, room, nice, day, love, accommodation	['explain', 'title', 'take', 'pair', 'flip_flops', 'pair', 'flat', 'shoe', 'wear', 'day', 'ever', 'wear', 'pair', 'shoe', 'still', 'take', 'soon', 'sit', 'eat', 'even', 'restaurant', 'beautiful', 'soft', 'white', 'sand', 'floor']
1.0	0.4538	island, seaplane, time, experience, transfer, fantastic, male, kuredu, airport, place	['partner', 'get', 'bit', 'dodgy', 'start', 'connect', 'flight', 'hour', 'late', 'therefore', 'go', 'one', 'leave', 'behind', 'least', 'couple', 'know', 'would', 'arrive', 'late', 'ahead', 'least', 'would', 'aware', 'could', 'make', 'decision', 'fly', 'pm', 'even', 'make', 'arrangement', 'evening', 'hook', 'couple', 'strand', 'stay', 'horrendous', 'hotel', 'male', 'even', 'star', 'sleep', 'wink', 'imagine', 'great', 'start', 'change', 'get', 'seaplane', 'next', 'day', 'fantastic', 'start']
2.0	0.3797	beach, villa, stay, water, sea, lovely, bungalow, villas, clean, sand	['accommodation', 'beachside', 'bungalow', 'spacious', 'comfortable', 'clean', 'metre', 'sea', 'large', 'firm', 'bed', 'beach', 'simply', 'lovely', 'white', 'soft', 'sand', 'lot', 'sea', 'creature', 'observe', 'fish', 'quite', 'large', 'brightly', 'coloured', 'everywhere', 'water', 'exotic', 'tree', 'flower', 'shrub', 'walk', 'right', 'round', 'island', 'actually', 'leave', 'beach']
3.0	0.3861	good, beautiful, great, location, amazing, view, perfect, beach, walk, area	['accommodation', 'choose', 'bungalow', 'perfect', 'seclude', 'minute', 'stroll', 'reception', 'area', 'restaurant', 'facility', 'electric', 'buggy', 'could', 'allway', 'request', 'transport', 'need', 'near', 'beach', 'minute', 'walk', 'away', 'view', 'stunning', 'quality', 'food', 'restaurant', 'high', 'quality', 'variety', 'seemingly', 'endless', 'whole']

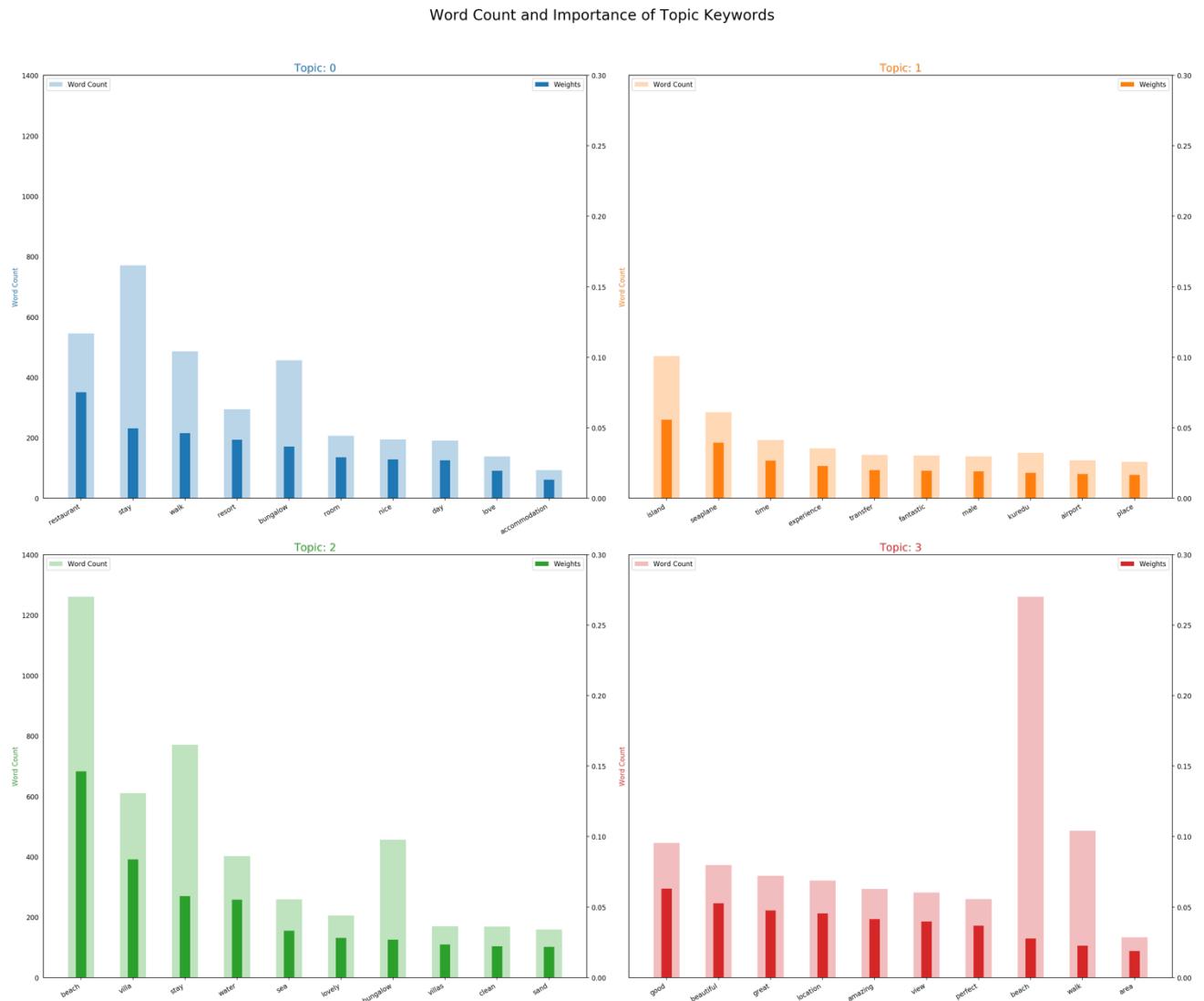
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



The interactive plot for location is saved as `location_vis_pyldavis.html`

### Location topics

Topic 0: location facilities – accommodation, restaurant etc.

Topic 1: travel to and from the resort (seaplane and plane)

Topic 2: beach, water, sand and other natural environment elements

Topic 3: referring to the character of the location (beautiful, amazing)

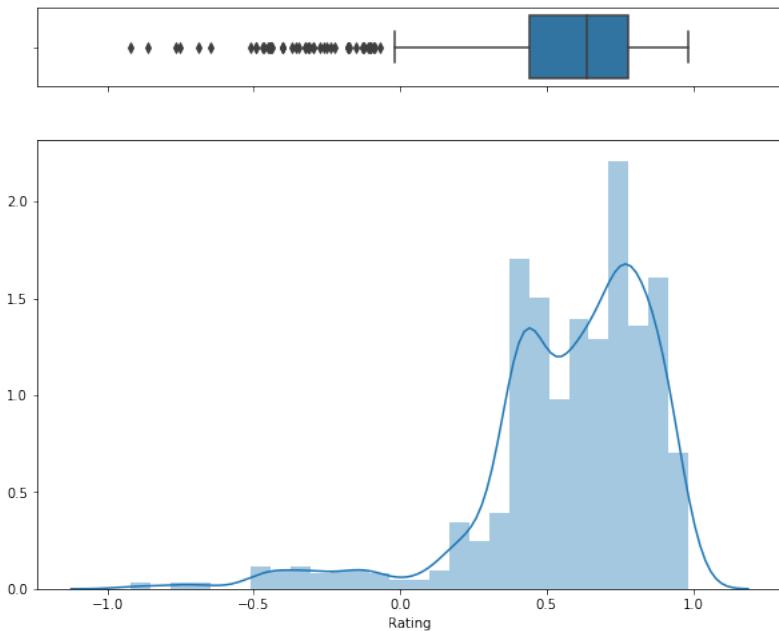
## 2.4 Cleanliness

### 2.4.1 Sentiment

Total sentences categorised as ‘Cleanliness: 899

Mean sentiment rating (mean in 1-5 range): 0.578 (3.944)

Distribution of sentence sentiment ratings for cleanliness:



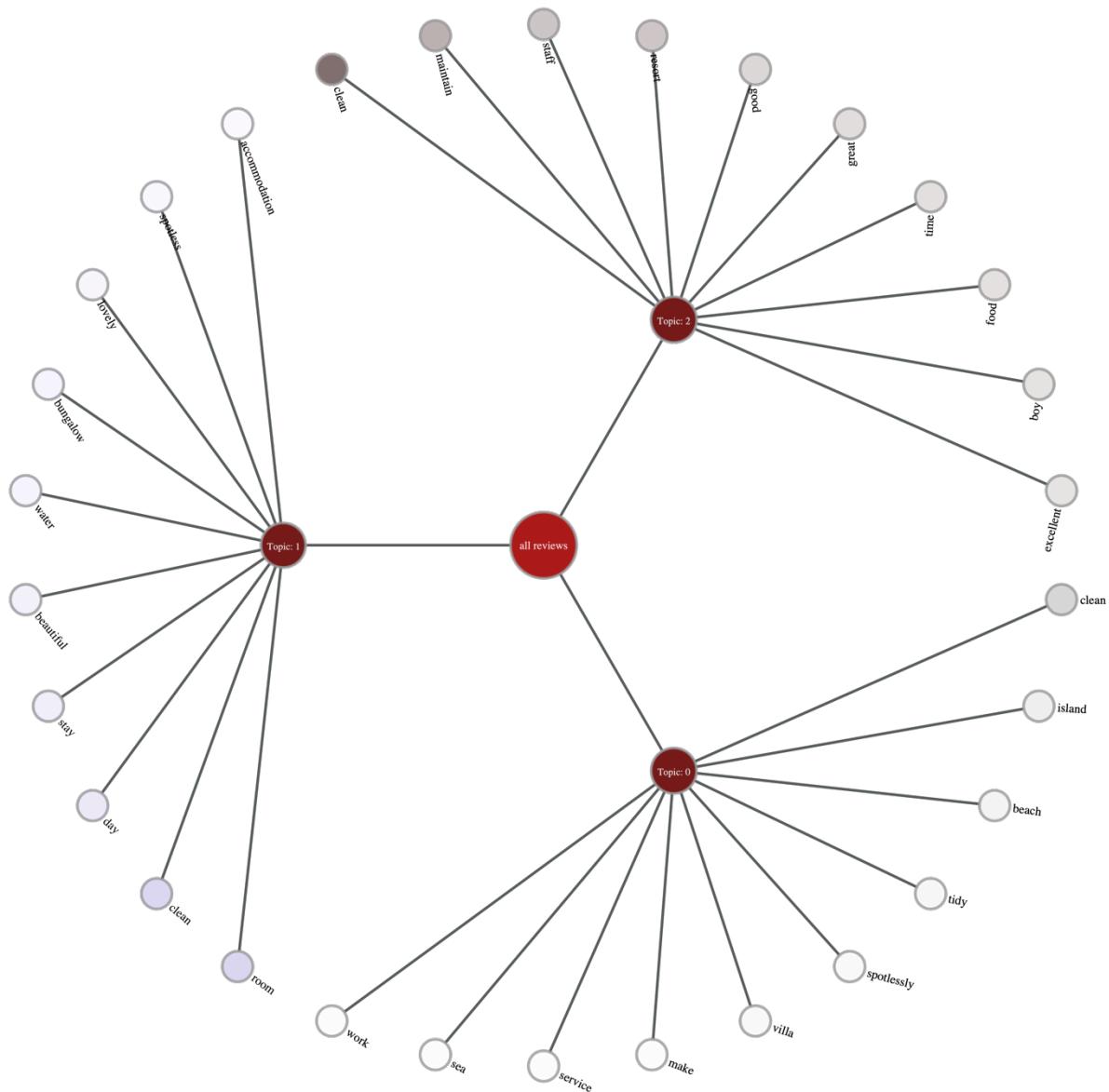
### 2.4.2 Topic modelling

Number of topics: 3

Reviews representative of each topic:

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text
0.0	0.4037	clean, island, beach, tidy, spotlessly, villa, make, service, sea, work	['always', 'keep', 'clean', 'public', 'area', 'smoke', 'permit', 'restaurant', 'permit', 'bar', 'spoilt', 'none', 'smoker']
1.0	0.419	room, clean, day, stay, beautiful, water, bungalow, lovely, spotless, accommodation	['return', 'night', 'stay', 'beautiful', 'kuredu', 'couple', 'week', 'ago', 'still', 'miss', 'stay', 'water', 'simply', 'lovely', 'ismail', 'look', 'water', 'villa', 'twice', 'day', 'check', 'need', 'keep', 'spotless', 'thankyou']
2.0	0.4226	clean, maintain, staff, resort, good, great, time, food, boy, excellent	['course', 'stun', 'white', 'sandy', 'beach', 'clean', 'turquoise', 'water', 'clear', 'blue', 'sky', 'amazing', 'sunset', 'plan', 'trip', 'maldive', 'already', 'expect']

**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:

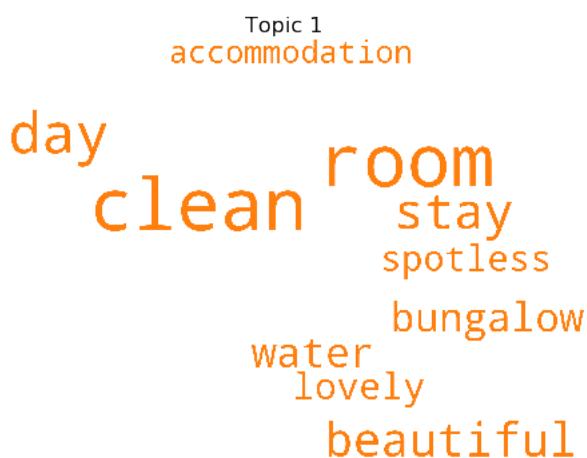
Topic 0



A word cloud for Topic 0. The most prominent word is "clean" in large blue font. Other visible words include "tidy", "sea", "work", "service", "make", "beach", "spotlessly", "island", and "villa".

clean  
tidy  
sea      work  
service    make  
beach    spotlessly  
island  
villa

Topic 1



A word cloud for Topic 1. The most prominent word is "clean" in large orange font. Other visible words include "day", "room", "stay", "spotless", "bungallow", "water", "lovely", and "beautiful".

day      clean    room  
              stay  
              spotless  
              bungallow  
water    lovely  
              beautiful

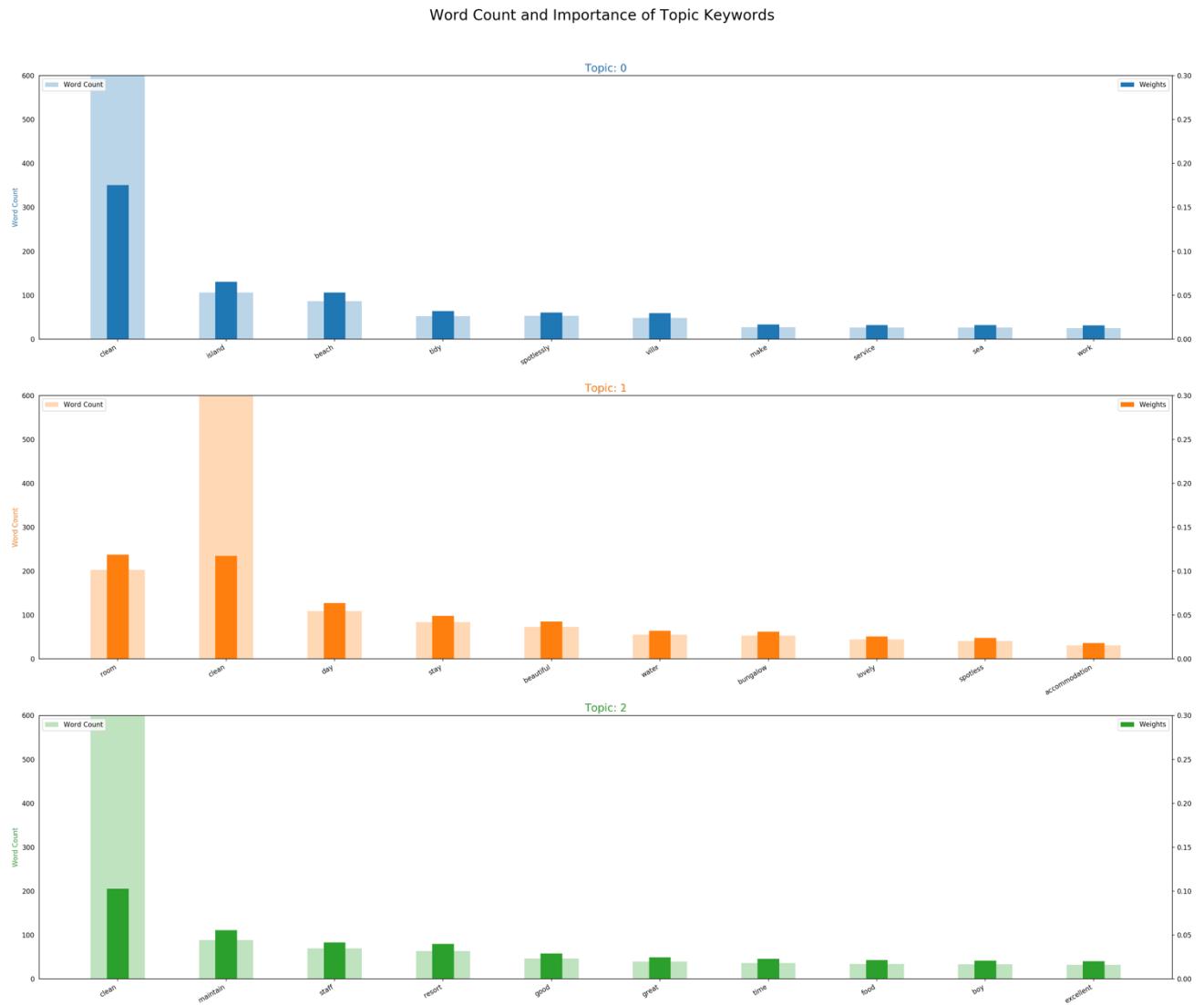
Topic 2



A word cloud for Topic 2. The most prominent word is "clean" in large green font. Other visible words include "good", "great", "boy", "maintain", "staff", "resort", "time", and "food".

good great clean  
              boy  
maintain  
staff  
resort  
time  
              food  
excellent

## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



### Cleanliness topics:

- Topic 0: keywords such as island, beach, clean and tidy and related reviews suggest this topic focuses on the cleanliness of the island, including the sea and beaches
- Topic 1: this is more focused on the cleanliness of the guest accommodation
- Topic 2: this appears similar to topic 0, with a greater focus on the cleaning staff

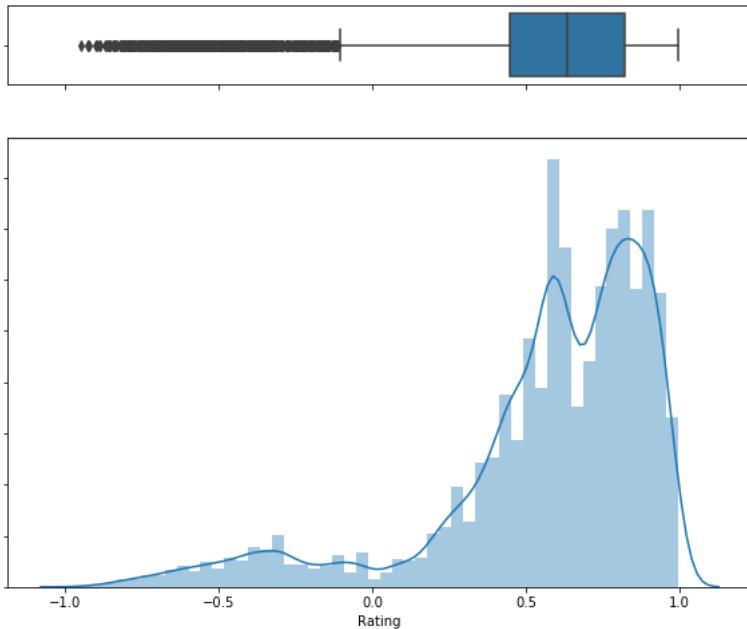
## 2.5 Service

### 2.5.1 Sentiment

Total sentences categorised as ‘Service: 13,807

Mean sentiment rating (mean in 1-5 range): 0.565 (3.913/)

Distribution of sentence sentiment ratings for service:



### 2.5.2 Topic modelling

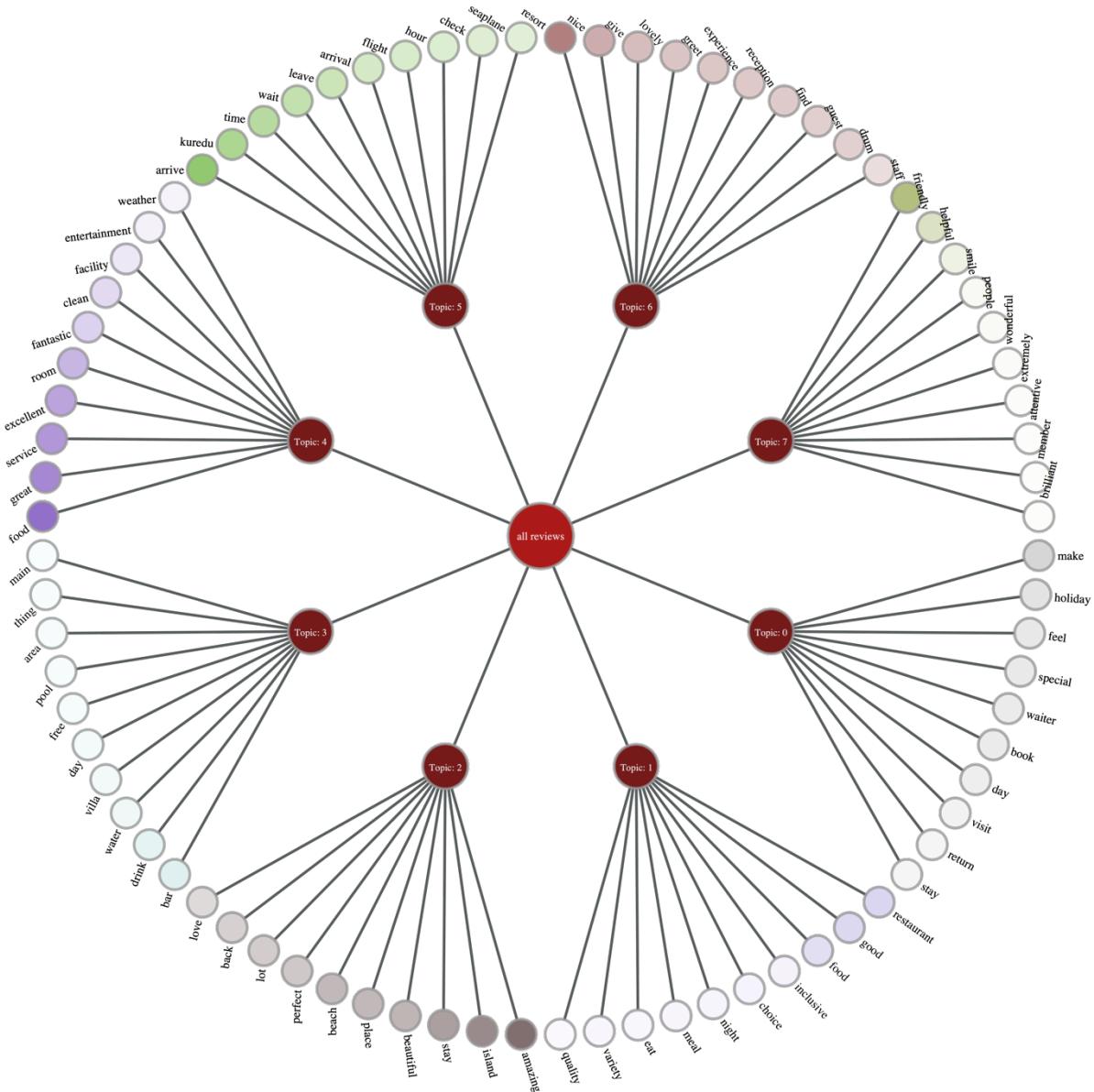
Number of topics: 8

Reviews representative of each topic:

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text (processed review)
0.0	0.2612	make, holiday, feel, special, waiter, book, day, visit, return, stay	['home', 'day', 'thought', 'still', 'wonderful', 'island', 'long', 'way', 'away', 'amazing', 'holiday', 'way', 'positive', 'review', 'read', 'go', 'true', 'service', 'second', 'none', 'make', 'feel', 'genuinely', 'welcome', 'would', 'love', 'return', 'day', 'contemplate', 'go', 'regret', 'enjoy']
1.0	0.2715	restaurant, good, food, inclusive, choice, night, meal, eat, variety, quality	['food', 'drink', 'bar', 'restaurant', 'high', 'quality', 'night', 'always', 'wide', 'variety', 'fresh', 'delicious', 'well', 'present', 'dish', 'meat', 'fish', 'veggie']
2.0	0.3229	amazing, island, stay, beautiful, place, beach, perfect, lot, back, love	['return', 'kuredu', 'day', 'ago', 'beautiful', 'place', 'stay', 'bungalow', 'perfect', 'restaurant', 'beautiful', 'food', 'good', 'waiter', 'make', 'experience', 'fun', 'everyday', 'see', 'amazing', 'fish', 'shark', 'turtle', 'love', 'wear', 'shoe', 'holiday', 'would', 'recommend', 'sea', 'shoe', 'time', 'coral', 'sharp', 'lot', 'sea', 'edge', 'wonderful', 'meal', 'really', 'make', 'special', 'friendly', 'would', 'definitely', 'go', 'back', 'island', 'truly', 'think', 'negative', 'island', 'side', 'long', 'take', 'get', 'afraid', 'want', 'go', 'stunning', 'place', 'cost', 'pay', 'big', 'thankyou', 'kuredo']
3.0	0.3157	bar, drink, water, villa, day, free, pool, area, thing, main	['meal', 'include', 'select', 'drink', 'water', 'sport', 'actually', 'hardly', 'use', 'free', 'let', 'loose', 'windsurf', 'purchase', 'lesson', 'even', 'windsurf', 'drink', 'espresso', 'pretty', 'basic', 'include', 'wifi', 'limited', 'mb', 'person', 'day', 'run', 'flash']
4.0	0.2424	food, great, service, excellent, room, fantastic, clean, facility, entertainment, weather	['scene', 'tour', 'impressive', 'go', 'kitchen', 'staff', 'live', 'spotless', 'brave', 'see', 'food', 'prepare', 'beach', 'clean', 'day', 'room', 'good', 'standard', 'staff', 'well', 'management', 'proud', 'first', 'class', 'make', 'holiday', 'wish', 'thank', 'great']
5.0	0.373	arrive, kuredu, time, wait, leave, arrival, flight, hour, check, seaplane	['arrive', 'male', 'today', 'take', 'private', 'lounge', 'give', 'different', 'time', 'departure', 'leave', 'plane', 'unbeknown', 'suit', 'case', 'finally', 'get', 'room', 'clock', 'telephone', 'reception', 'find', 'case', 'finally', 'around', 'find', 'still', 'male', 'promise', 'shortly', 'moment', 'really', 'bad', 'storm', 'suppose', 'wait', 'rain', 'pass', 'meanwhile', 'sit', 'room', 'dressing', 'gown', 'clothe', 'one', 'travel', 'eat', 'breakfast', 'plane', 'long', 'time', 'ago', 'maldives', 'time']

6.0	0.3261	resort, nice, give, lovely, greet, experience, reception, find, guest, drum	['great', 'get', 'seaplane', 'meet', 'shook', 'hand', 'approx', 'people', 'assume', 'management', 'would', 'nice', 'introduce', 'handshake', 'take', 'resort', 'reception', 'greet', 'mervin', 'give', 'wrong', 'accommodation', 'pay', 'ojbv', 'give', 'happy', 'start', 'put', 'low', 'category', 'pay', 'lucky', 'full', 'detail', 'destinology', 'problem', 'sort', 'original', 'book', 'accommodation', 'sort', 'hour', 'kuredu', 'say', 'destinology', 'fault', 'however', 'destinology', 'say', 'kuredu', 'fault', 'lovely', 'accommodation', 'ojbv', 'apology', 'would', 'nice', 'none', 'receive', 'loved', 'fact']
7.0	0.2539	staff, friendly, helpful, smile, people, wonderful, extremely, attentive, member, brilliant	['would', 'like', 'compliment', 'staff', 'include', 'lady', 'sweeper', 'friendly', 'way', 'build', 'crew', 'always', 'wave', 'gardener', 'always', 'smile', 'driver', 'staff', 'meet', 'path', 'always', 'friendly', 'word', 'competent', 'people', 'shop', 'people']

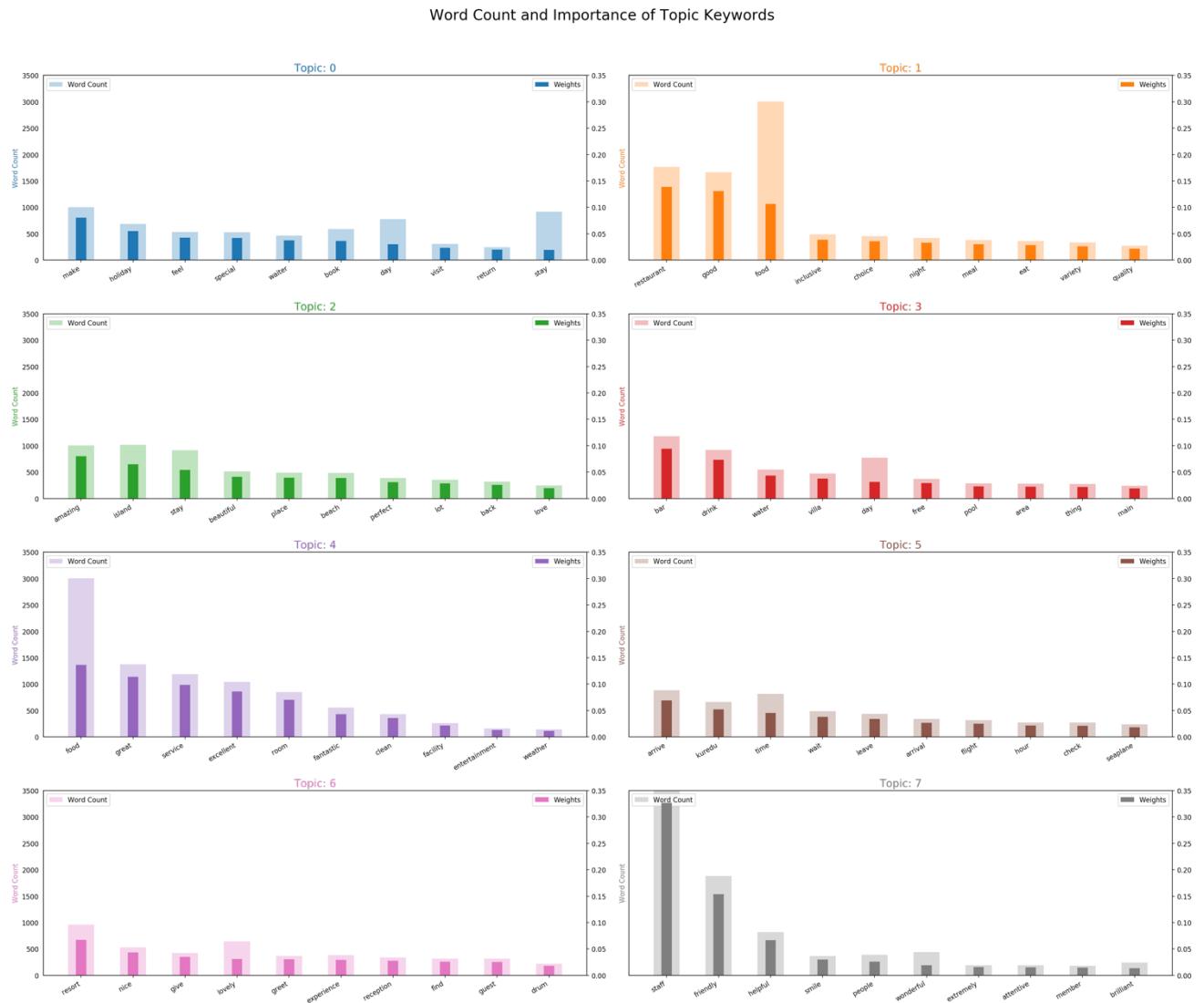
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



### Service topics:

- Topic 0: waiting staff
- Topic 1: food and meals
- Topic 2: staff mention in overall (generally positive) summary context of review as accompanied with amazing, beautiful, perfect
- Topic 3: service around bars and pool areas
- Topic 4: generally overall summary of customer service experience
- Topic 5: travel-related service
- Topic 6: arrival and welcome
- Topic 7: level of helpfulness, attentiveness and friendliness of staff

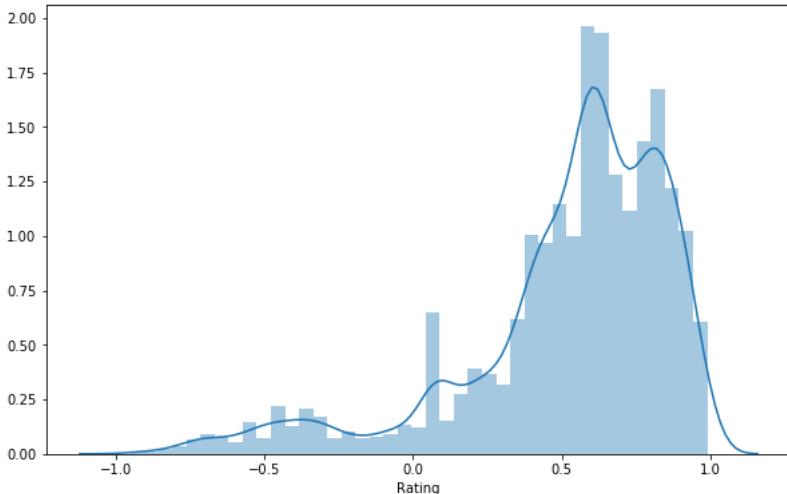
## 2.6 Facilities/Activities

### 2.6.1 Sentiment

Total sentences categorised as 'Facilities/Activities': 3,894

Mean sentiment rating (mean in 1-5 range): 0.519 (3.798)

Distribution of sentence sentiment ratings for cleanliness:



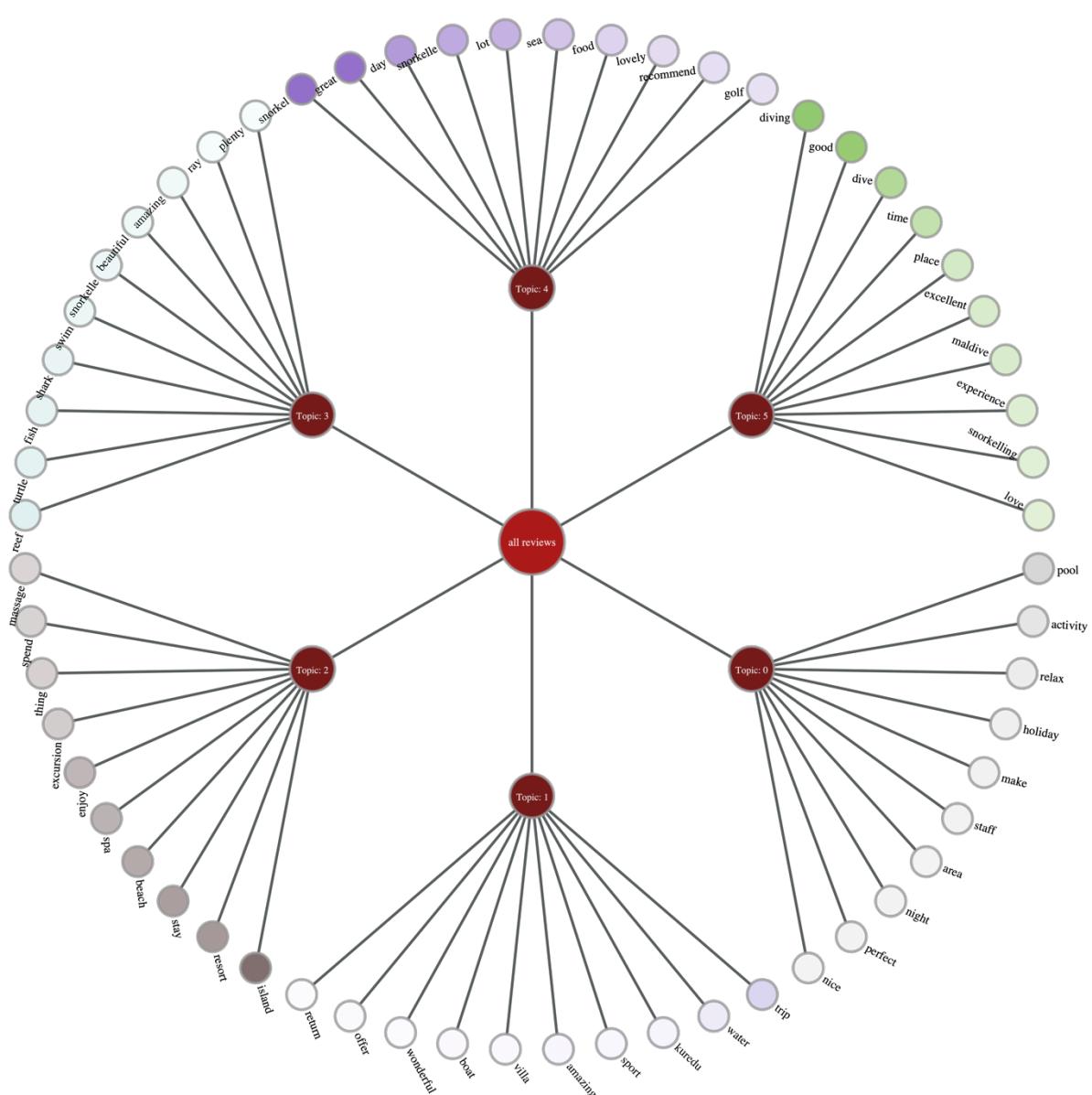
### 2.6.2 Topic modelling

Number of topics: 6

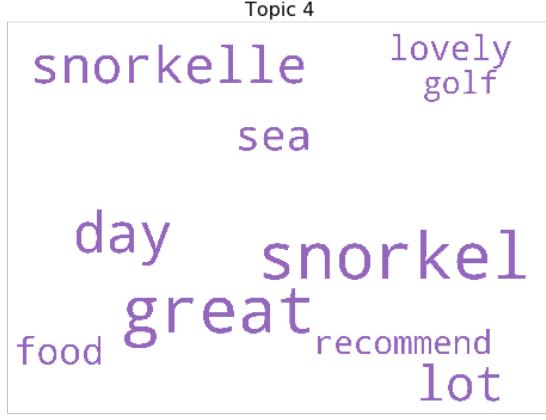
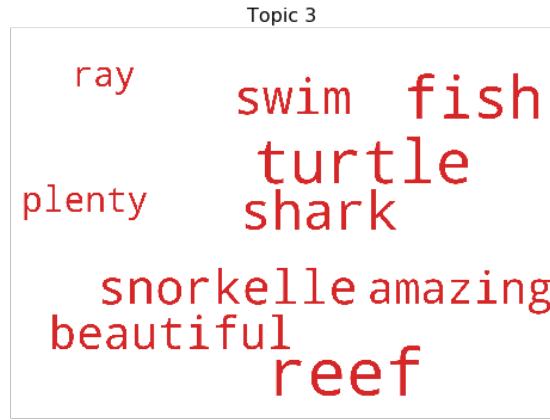
Reviews representative of each topic:

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text
0.0	0.3469	pool, activity, relax, holiday, make, staff, area, night, perfect, nice	['pay', 'week', 'resort', 'child', 'think', 'couple', 'would', 'nice', 'arrival', 'advise', 'pool', 'leak', 'could', 'use', 'resort', 'pool', 'result', 'busy', 'pool', 'lot', 'kid', 'noise', 'offer', 'pay', 'facility', 'refund', 'staff', 'also', 'admit', 'leak', 'almost', 'refit', 'pool', 'would', 'action', 'time', 'therefore', 'feel', 'sell', 'upgrade']
1.0	0.2885	trip, water, kuredu, sport, amazing, villa, boat, wonderful, offer, return	['fishing', 'trip', 'could', 'go', 'wrong', 'cancel', 'notify', 'management', 'fantastic', 'listening', 'issue', 'arrange', 'boat', 'go', 'crew']
2.0	0.3348	island, resort, stay, beach, spa, enjoy, excursion, thing, spend, massage	['choose', 'fly', 'stop', 'advice', 'would', 'fly', 'direct', 'go', 'fine', 'layover', 'hour', 'come', 'back', 'layover', 'hour', 'male', 'airport', 'hour', 'alot', 'wait', 'tell', 'hotel', 'end', 'runway', 'pool', 'spa', 'away', 'time', 'wait', 'flight']
3.0	0.3182	reef, turtle, fish, shark, swim, snorkelle, beautiful, amazing, ray, plenty	['snorkeling', 'trip', 'reef', 'lagoon', 'never', 'fail', 'impress', 'shark', 'lion', 'eel', 'turtle', 'turtle', 'hundred', 'colorful', 'reef', 'fish']
4.0	0.2881	snorkel, great, day, snorkelle, lot, sea, food, lovely, recommend, golf	['great', 'location', 'snorkelle', 'side', 'get', 'lot', 'sea', 'plane', 'noise', 'day']
5.0	0.2889	diving, good, dive, time, place, excellent, maldive, experience, snorkelling, love	['dive', 'master', 'almost', 'age', 'make', 'unable', 'understand', 'concept', 'enjoyable', 'dive', 'make', 'new', 'rule', 'dive', 'padi', 'naui', 'diver', 'would', 'ever', 'learn', 'way', 'pay', 'money', 'dive', 'computer']

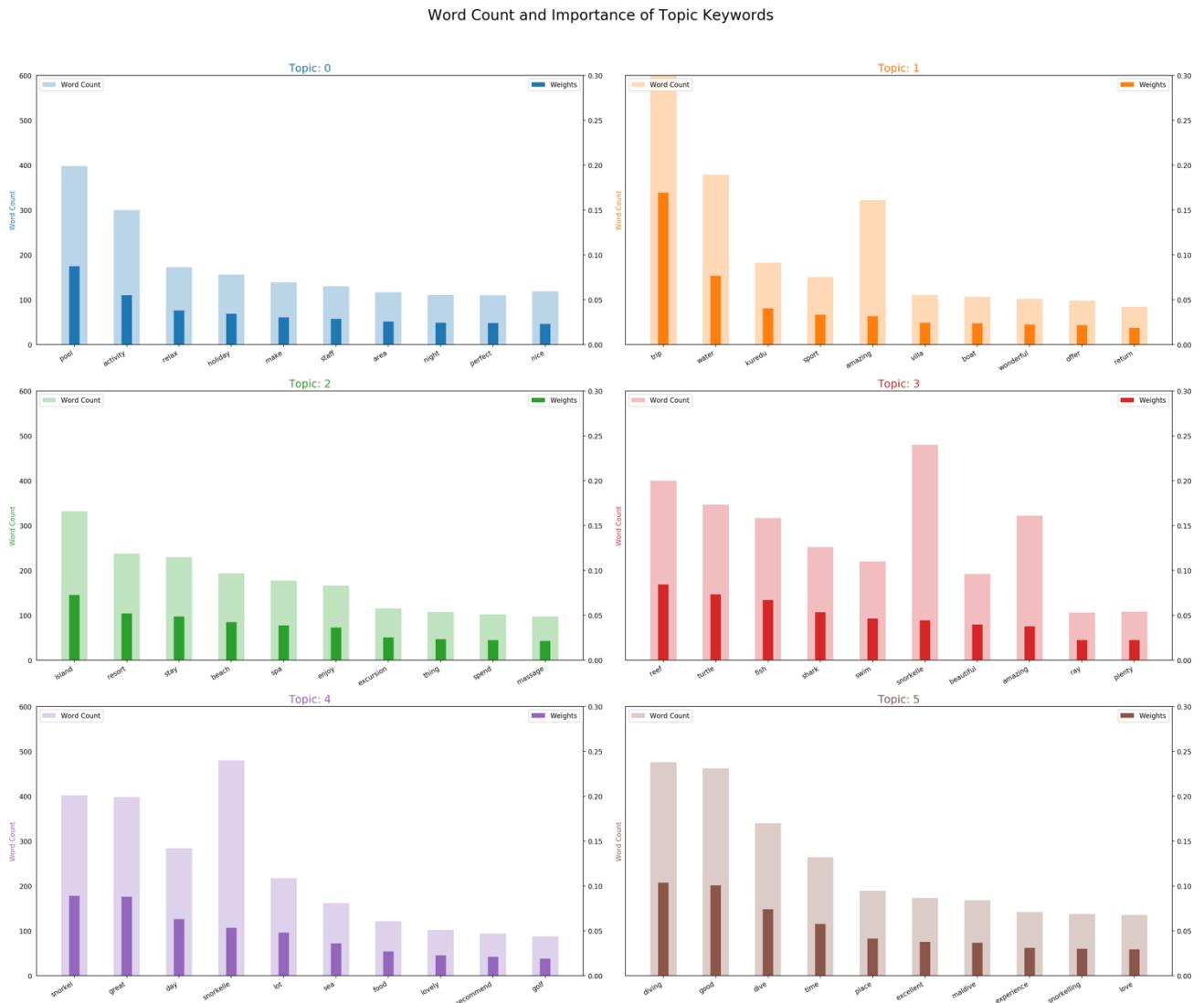
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



### Facilities/Activities topics:

Some repetition between topics is likely especially within this aspect as groups of guests may undertake the same or similar combinations of activities during their holiday (e.g. one group may dive and snorkel, another group may golf and snorkel so there will be crossover with snorkelling), but these topics may be useful in, for example, targeted marketing of guest cohorts with certain offers for activity combinations.

- Topic 0: the pool and relaxing
- Topic 1: sport, boats, and mentions of 'trip' but the latter may more be in reference to their trip to the Maldives rather than any excursion
- Topic 2: the spa and massage
- Topic 3: snorkelling and the reef
- Topic 4: snorkelling, golf
- Topic 5: diving and snorkelling

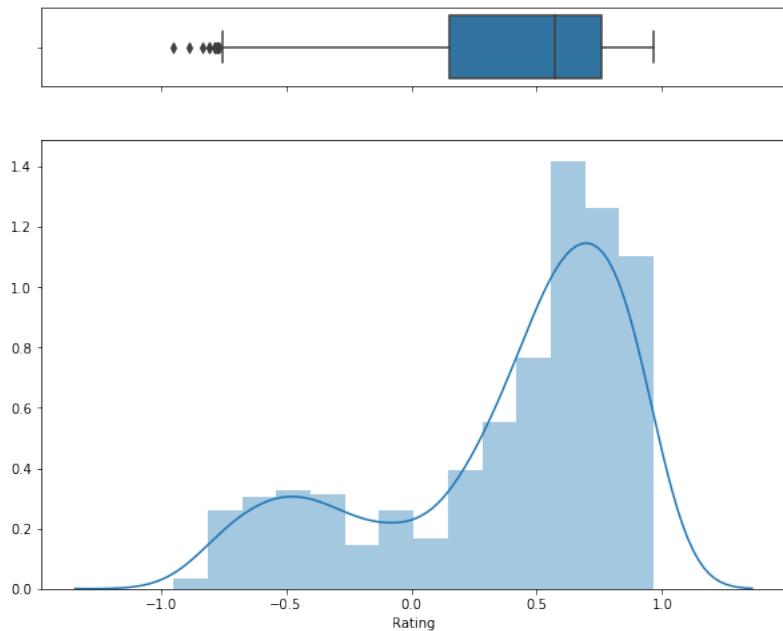
## 2.7 Other

### 2.7.1 Sentiment

Total sentences categorised as ‘Other’: 649

Mean sentiment rating (mean in 1-5 range): 0.379 (3.446)

Distribution of sentence sentiment ratings for other:



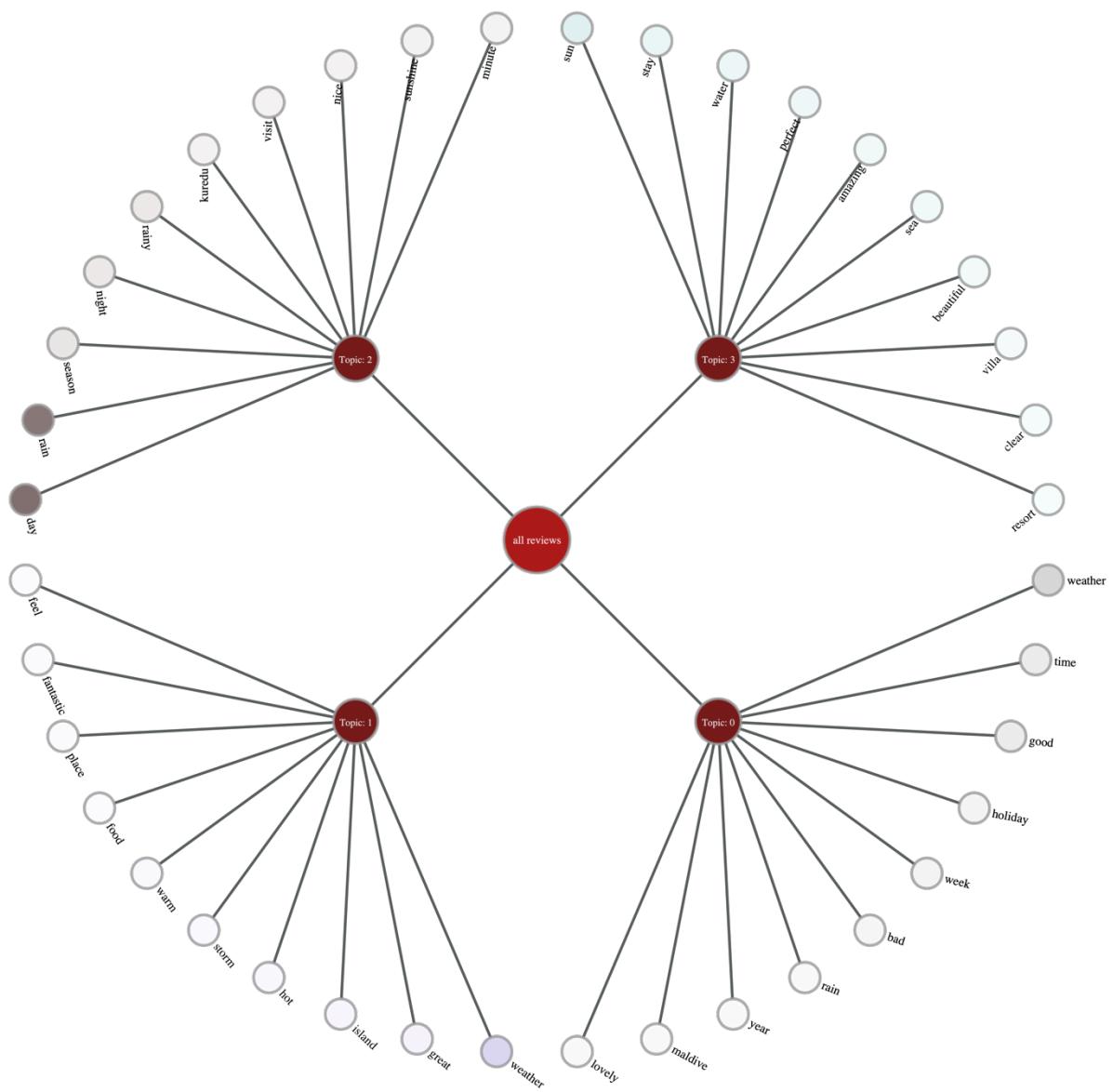
### 2.7.2 Topic modelling

**Number of topics:** 4

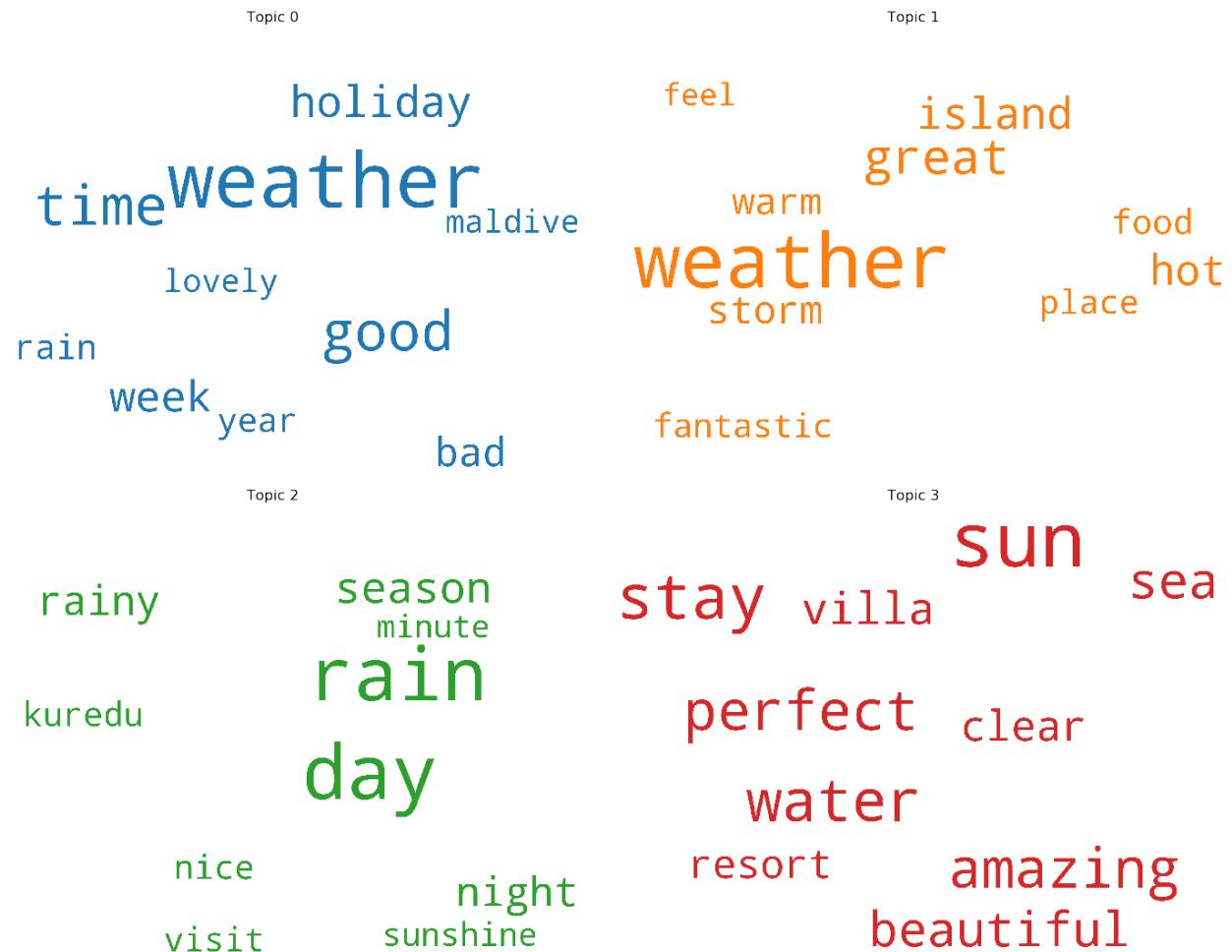
**Reviews representative of each topic:**

Topic_Num	Topic_Perc_Contrib	Keywords	Representative Text (processed review)
0.0	0.3325	weather, time, good, holiday, week, bad, rain, year, maldives, lovely	['water', 'stunning', 'atmosphere', 'lovely', 'weather', 'bit', 'windy', 'go', 'made', 'heat', 'bearable', 'good', 'mean', 'kayaking', 'bit', 'difficult']
1.0	0.3395	weather, great, island, hot, storm, warm, food, place, fantastic, feel	['place', 'want', 'relax', 'cell', 'phone', 'ring', 'get', 'back', 'touch', 'reality', 'leave', 'feel', 'world', 'time', 'hold']
2.0	0.3607	day, rain, season, night, rainy, kuredu, visit, nice, sunshine, minute	['turn', 'thankfully', 'day', 'mix', 'weather', 'usually', 'sunny', 'morning', 'lunch', 'rain', 'cloudy', 'windy', 'day', 'rain', 'day', 'still', 'manage', 'thing', 'want', 'dinner', 'cloudy', 'day', 'want', 'sunset', 'cruise', 'watch', 'see', 'sunset', 'see', 'lovely', 'sunset', 'well']
3.0	0.3828	sun, stay, water, perfect, amazing, sea, beautiful, villa, clear, resort	['problem', 'beach', 'villas', 'right', 'unobstructed', 'sea', 'view', 'direct', 'access', 'mean', 'continually', 'camp', 'villa', 'even', 'use', 'designate', 'sun', 'lounger']

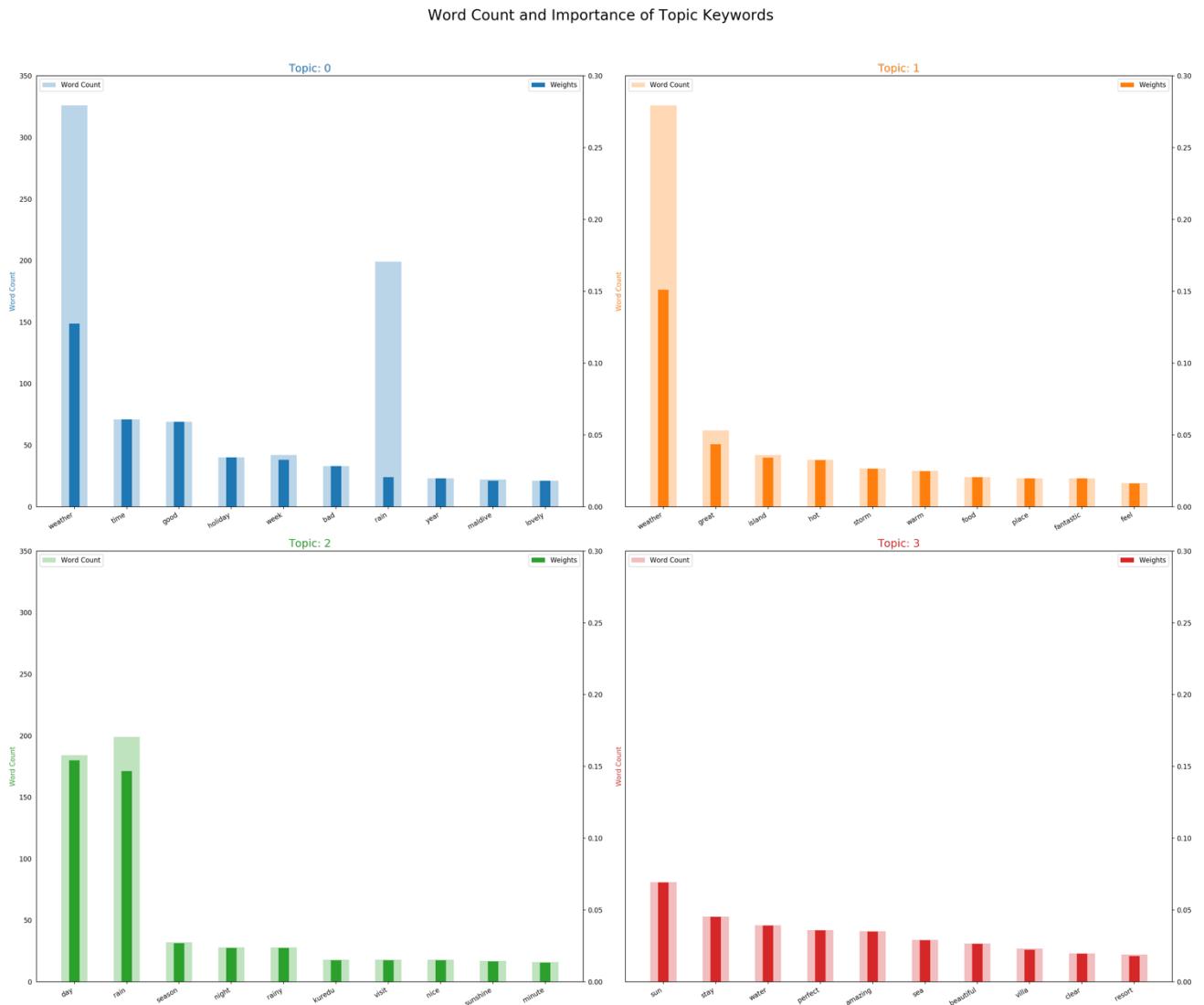
**Tree diagram of the 10 most important words per topic, with colour intensity of each word indicating weight (importance):**



**Word cloud of most important words (highest weights) per topic, with size indicating relative importance of the word:**



## Bar plot of how frequently keywords appear in the reviews and their importance (weight) within the topic that is characterised by it:



### Topics:

- Topic 0: bad weather and rain
- Topic 1: hot weather and storm
- Topic 2: fluctuating weather
- Topic 3: sun and clear weather

## 3 Comments

Improvements:

- Scraping reviews from other platforms and combining would provide more data to work with and improve the robustness of topics generated
- The algorithm used here for sentiment scoring of sentences is an out-of-the-box, predeveloped algorithm – as this dataset is relatively small, a sentiment analysing algorithm developed from it is unlikely to be very robust, hence relying on a predefined one. For improved performance a sentiment analysis algorithm could be developed from a larger dataset of Maldives hotel reviews (e.g. scraping reviews from multiple hotels in the Maldives and either using reviewer scores, or manually labelling review sentences to train a model that can be used in future exercises such as this) to help it in recognising and correctly identifying ambiguous sentences (such as double negative sentences), and making it more specific to the types of wording used in reviews
- Refining the aspect keywords on which aspect identification and separation is based to improve categorisation of review segments. This would require staff input to identify particular keywords specific to the resort in question.
- Negative reviews only could be isolated for further analysis, potentially applying topic modelling again to all negative reviews together and to those within topics within aspects (may require a larger dataset) to provide clearer points for improvement within each aspect
- Investigating with different numbers of topics per aspect may improve topic content
- Topics identified are based on a subjective idea of what is included in each topic, and would benefit from input from those with more domain knowledge

Accompanying files:

- Raw review data scraped from TripAdvisor (includes user name, date of review, raw text review, and manager response if any)
- Tables for each aspect, each including the raw reviews, processed reviews, dominant sentiment rating, topic number, and topic percentage contribution
  - o Raw reviews: KureduReviews\_19\_04\_2020.csv
  - o Processed, categorised and rated review sentences:
  - o Review sentences by aspect: df\_breakdown.csv
    - Value – df\_value\_topics.csv
    - Location – df\_location\_topics.csv
    - Rooms – df\_rooms\_topics.csv
    - Cleanliness – df\_cleanliness\_topics.csv
    - Service – df\_service\_topics.csv
    - Facilities/activities – df\_facilities\_topics.csv
    - Other – df\_other\_topics.csv
- HTML's for the visualisations for each aspect (opening launches a webpage)
  - o Value - value\_topics\_pyldavis.html
  - o Location - location\_vis\_pyldavis.html
  - o Rooms - rooms\_topics\_pyldavis.html
  - o Cleanliness - cleanliness\_vis\_pyldavis.html
  - o Facilities/activities - facilities\_vis\_pyldavis.html
  - o Service - service\_vis\_pyldavis.html
  - o Other - other\_vis\_pyldavis.html