

HOTEL REVIEWS



TABLE OF CONTENTS

BUSINESS
UNDERSTANDING

01

DATA
UNDERSTANDING

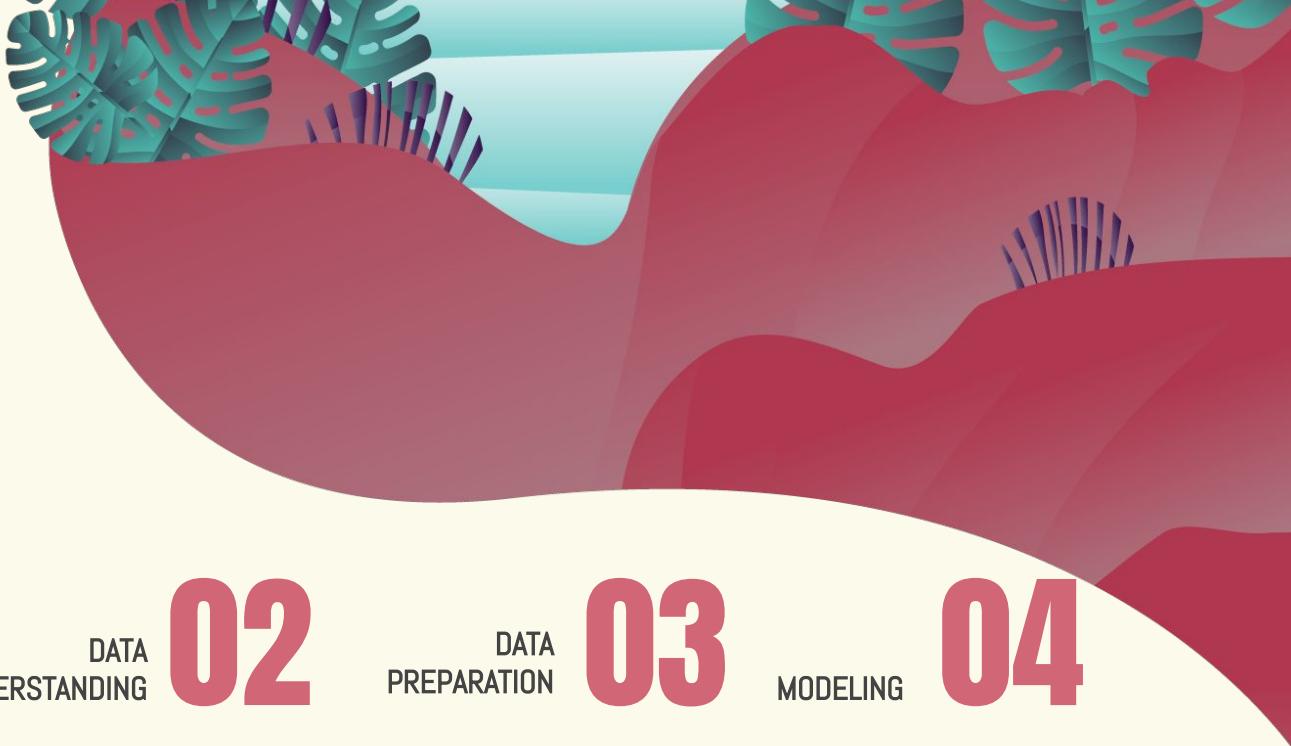
02

DATA
PREPARATION

03

MODELING

04



01. BUSINESS OVERVIEW



BUSINESS PROBLEM

How can hotels increase number of customers?

- Increase number of return customers
- Attract more new customers



Fully assess the current situation

understanding the analytics capability



Basic structure

common features



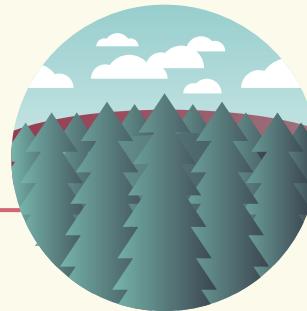
Predictive analytics could help increasing number of customers

1. Which type of customer were most likely to NOT return?
2. What changes/ offers would a particular customer best respond to?
3. What factors were likely to affect the hotel booking rate?



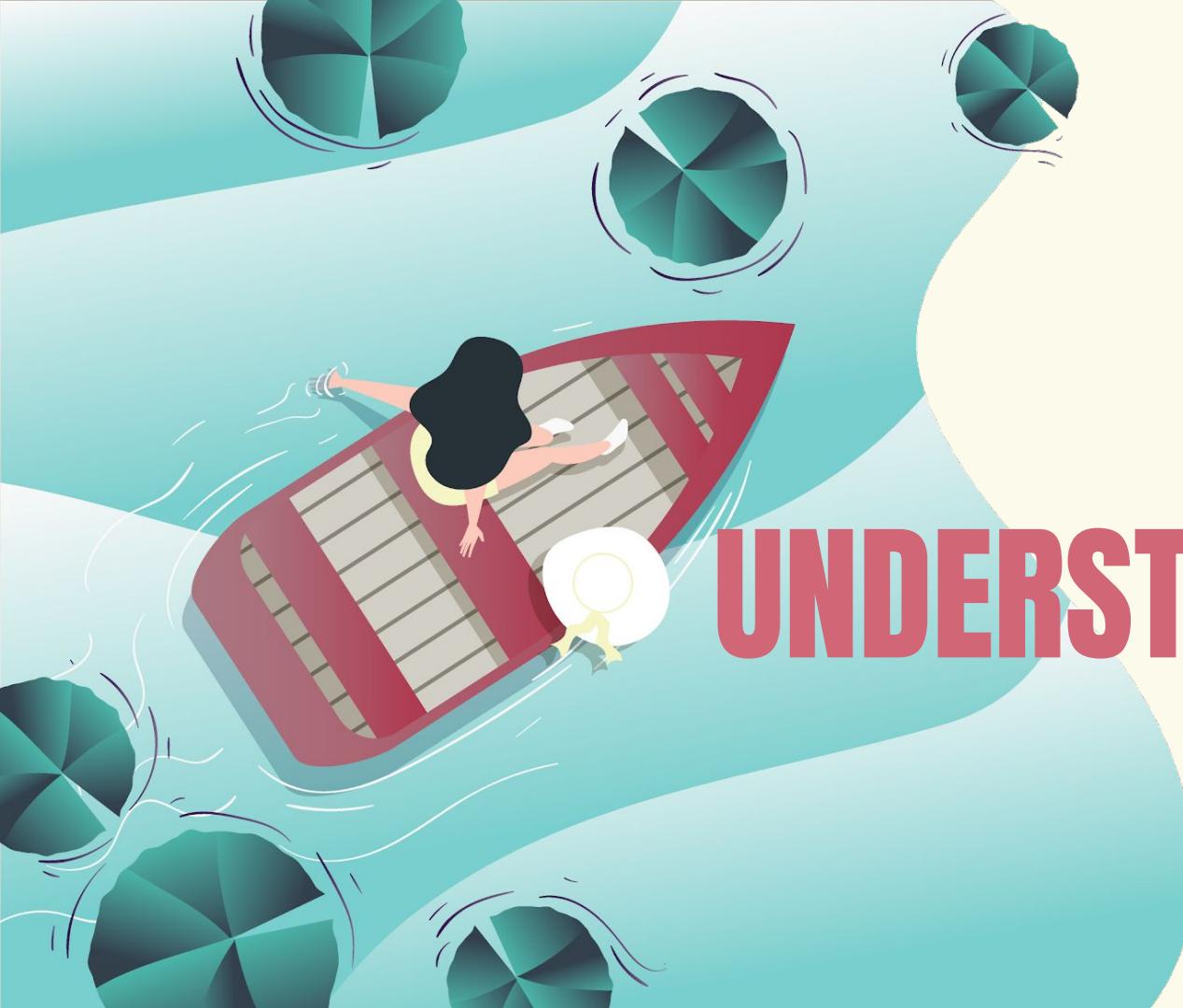
Predicting problem

What factors were likely to affect the hotel booking rate?



Analytics Solution

1. Data availability and reasonably easily accessible
2. Prediction model could be easily integrated.
3. Explained main drivers behind gaining customers



02. DATA UNDERSTANDING

	Review	Rating
0	nice hotel expensive parking got good deal sta...	4
1	ok nothing special charge diamond member hilton...	2
2	nice rooms not 4* experience hotel monaco seat...	3
3	unique, great stay, wonderful time hotel monac...	5
4	great stay great stay, went seahawk game aweso...	5

```
reviews.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20491 entries, 0 to 20490
Data columns (total 2 columns):
 #   Column   Non-Null Count  Dtype  
--- 
 0   Review    20491 non-null  object 
 1   Rating    20491 non-null  int64  
dtypes: int64(1), object(1)
memory usage: 320.3+ KB
```

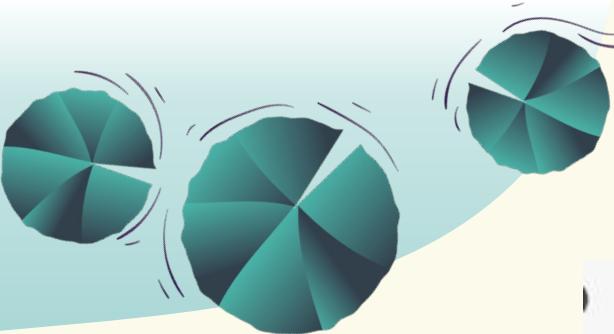
Data Source:

tripadvisor_hotel_reviews.csv

<https://www.kaggle.com/andrewmvd/trip-advisor-hotel-reviews>

20,491

number of instances
included in dataset



20,491

1. RATING

number of instances included in dataset

levels: 1, 2, 3, 4, 5

2 Features:

Review (number)
Rating (string)

2. REVIEW

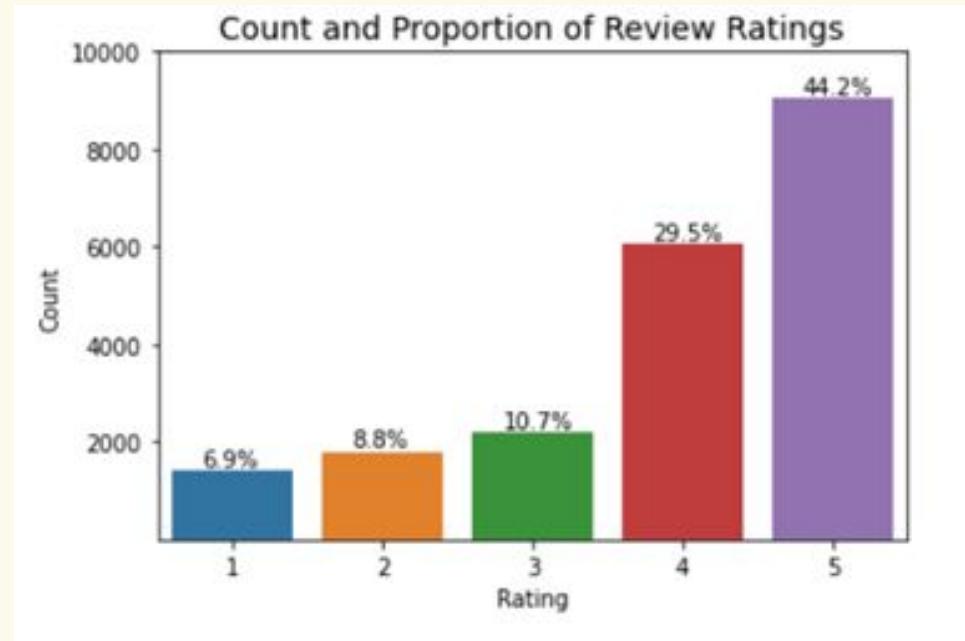
```
reviews.Rating.describe()
```

```
count      20491.000000
mean       3.952223
std        1.233030
min        1.000000
25%        3.000000
50%        4.000000
75%        5.000000
max        5.000000
Name: Rating, dtype: float64
```

```
reviews.iloc[1342]['Review']
```

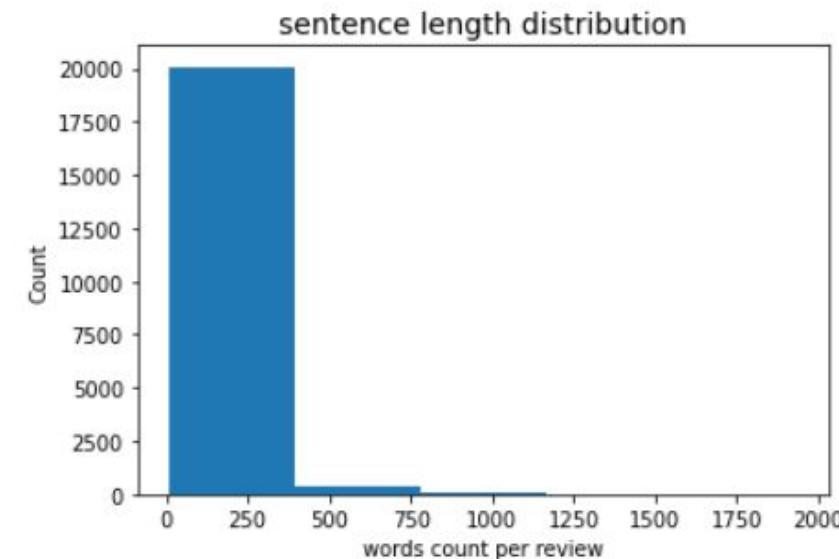
```
'great resort area hote price, 1. hotel not beach, prices reflec
cony resort style pool not stay here. 2, breakfast included room p
lines start right 8 a.m 3, safe secure immaculately clean hotel s
y 8th floor.4, want space ask 8th floor mountain view rooms rooms
e no real view parking garage, great spot drink friends dinner.5,
ght street,
```





**Rating:
Exponential Distribution**

1st mode: 5 star (44.2%)
2nd mode: 4 star (29.5%)



Review: Exponential Distribution

Longest sentence: 1933 words.
Shortest sentence: 9 words.



“ great resort area hote price, 1. hotel not beach, prices reflect, want beach large balcony resort style pool not stay here.2, breakfast included room price excellent, early, lines start right 8 a.m.3, safe secure immaculately clean hotel small rooms, coin laundry 8th floor.4, want space ask 8th floor mountain view rooms rooms hotel lanai, lanai huge no real view parking garage, great spot drink friends dinner.5, public access beach right street,”

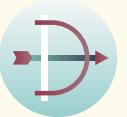
03. DATA PREPROCESSING



OUR SERVICES



Missing Value



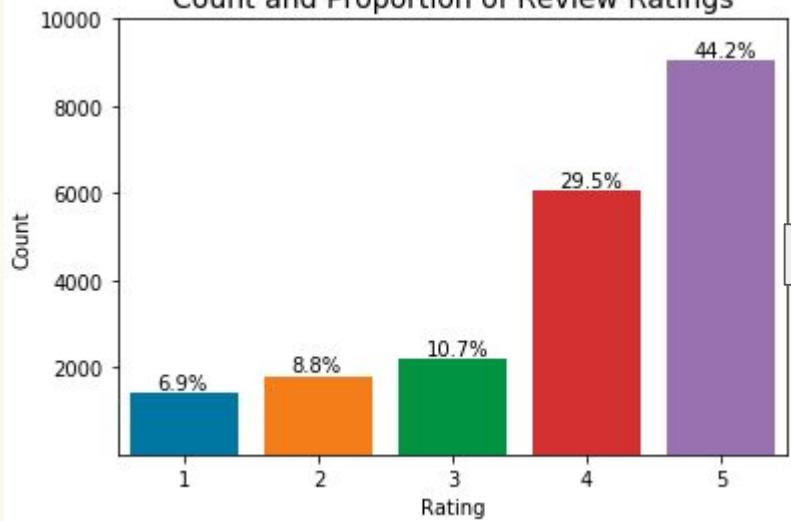
Remove Duplication



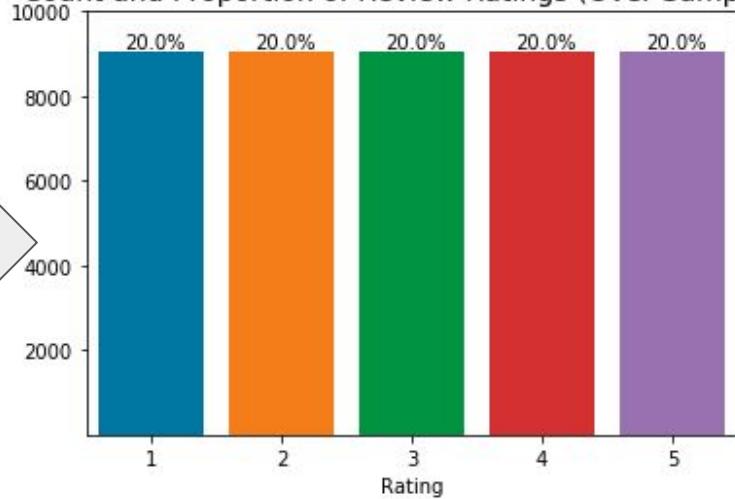
Sampling

Over Sampling

Count and Proportion of Review Ratings



Count and Proportion of Review Ratings (Over Sampling)





Normalization

(1) stemming, (2) lemmatization, and (3)
lowercase



Tokenization

Split strings of text into smaller pieces



Stopwords

Remove stopwords

Word Count & Word Cloud

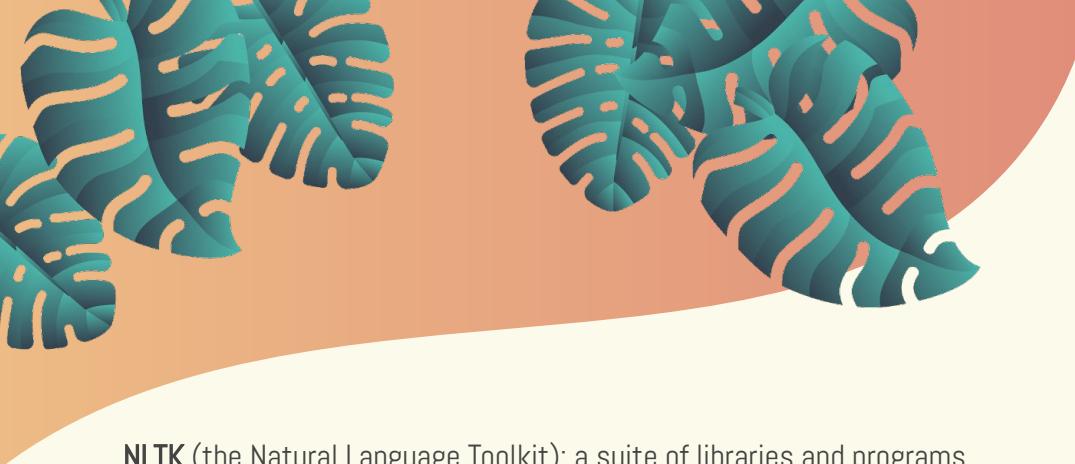
```
words.most_common(50)
```

```
[('room', 47747),  
 ('great', 21478),  
 ('good', 17484),  
 ('staff', 16702),  
 ('stay', 15599),  
 ('night', 14562),  
 ('day', 13470),  
 ('time', 12969),  
 ('nice', 12644),  
 ('location', 11558),  
 ('service', 10940),  
 ('beach', 10522),  
 ('stayed', 10502),  
 ('restaurant', 10476),  
 ('breakfast', 10134),  
 ('food', 9675),  
 ('clean', 9606),  
 ('place', 9514),  
 ('resort', 9162),  
 ('pool', 8680),
```



A vibrant, stylized illustration of a tropical beach at sunset. A woman with dark hair, wearing a yellow bikini, stands on a sandy beach holding a surfboard with a green and blue leaf pattern. She is looking towards the ocean. The background features large, colorful palm trees in shades of teal, orange, and red. The sky is a gradient from light blue to a warm orange and yellow near the horizon. The overall style is graphic and modern.

04. Modeling



TRY 1: Find common words

common_words_by_ratings

	1	2	3	4	5
0	(hotel, 3588)	(room, 4250)	(hotel, 5058)	(hotel, 14125)	(hotel, 21822)
1	(room, 3267)	(hotel, 4166)	(room, 4424)	(room, 10172)	(room, 12126)
2	(stay, 1207)	(good, 1484)	(good, 2453)	(great, 7414)	(great, 10665)
3	(staff, 958)	(stay, 1367)	(great, 1751)	(good, 6697)	(staff, 7808)
4	(rooms, 893)	(rooms, 1354)	(nice, 1739)	(nice, 4826)	(stay, 6823)
5	(night, 874)	(staff, 1283)	(rooms, 1480)	(staff, 4665)	(good, 5705)
6	(service, 810)	(service, 1217)	(staff, 1468)	(stay, 4331)	(stayed, 4817)
7	(day, 807)	(night, 1141)	(stay, 1409)	(location, 3865)	(location, 4807)
8	(like, 713)	(food, 1075)	(location, 1402)	(rooms, 3615)	(rooms, 4652)
9	(told, 707)	(day, 1059)	(beach, 1260)	(clean, 3452)	(nice, 4407)
10	(time, 686)	(nice, 1049)	(night, 1219)	(beach, 3438)	(breakfast, 4142)
11	(desk, 637)	(time, 1044)	(food, 1159)	(breakfast, 3064)	(service, 4138)
12	(place, 634)	(beach, 1034)	(clean, 1151)	(stayed, 3049)	(time, 3958)
13	(got, 629)	(like, 1001)	(like, 1121)	(time, 2999)	(clean, 3840)
14	(resort, 617)	(resort, 980)	(time, 1113)	(food, 2886)	(excellent, 3670)
15	(stayed, 611)	(great, 908)	(service, 1101)	(night, 2829)	(day, 3586)
16	(food, 608)	(stayed, 889)	(stayed, 1099)	(day, 2789)	(beach, 3380)
17	(good, 604)	(got, 843)	(day, 1080)	(service, 2661)	(friendly, 3378)
18	(people, 534)	(people, 839)	(resort, 1041)	(resort, 2523)	(night, 3342)
19	(went, 499)	(really, 769)	(breakfast, 1001)	(really, 2517)	(place, 3302)

TRY 2: Improve common words

common_nouns_by_ratings					
	1	2	3	4	5
0	(hotel, 3436)	(room, 4250)	(hotel, 4845)	(hotel, 13542)	(hotel, 20949)
1	(room, 3267)	(hotel, 3979)	(room, 4424)	(room, 10172)	(room, 12126)
2	(staff, 958)	(rooms, 1354)	(rooms, 1480)	(staff, 4665)	(staff, 7808)
3	(rooms, 893)	(staff, 1283)	(staff, 1468)	(location, 3751)	(location, 4662)
4	(night, 874)	(service, 1206)	(location, 1344)	(rooms, 3615)	(rooms, 4652)
5	(day, 807)	(night, 1141)	(night, 1219)	(time, 2999)	(service, 4081)
6	<u>(service, 801)</u>	(food, 1075)	(food, 1159)	(food, 2886)	(time, 3958)
7	(time, 686)	(day, 1059)	(time, 1113)	(night, 2829)	(day, 3586)
8	(stay, 614)	(time, 1044)	(service, 1081)	(day, 2789)	(breakfast, 3551)
9	(food, 608)	(resort, 880)	(day, 1080)	(beach, 2632)	(stay, 3446)
10	(place, 582)	(people, 839)	(beach, 971)	(service, 2624)	(night, 3342)
11	(resort, 553)	(beach, 794)	(resort, 909)	(breakfast, 2602)	(food, 3232)
12	(people, 534)	(stay, 661)	(breakfast, 868)	(stay, 2353)	(place, 3118)
13	(desk, 484)	(location, 654)	(people, 855)	(resort, 2257)	(beach, 2568)
14	(water, 425)	(water, 632)	(area, 797)	(place, 2082)	(people, 2464)
15	(days, 365)	(place, 616)	(pool, 770)	(people, 2058)	(resort, 2444)
16	(experience, 351)	(breakfast, 577)	(place, 767)	(area, 2015)	(pool, 2370)
17	(bathroom, 346)	(desk, 564)	(stay, 765)	(pool, 1891)	(area, 2324)
18	(manager, 335)	(pool, 550)	(water, 633)	(restaurants, 1722)	(trip, 2313)
19	(beach, 335)	(floor, 507)	(bathroom, 609)	(bar, 1631)	(restaurants, 2064)

common_adjective_by_ratings					
	1	2	3	4	5
0	(good, 602)	(good, 1473)	(good, 2436)	(great, 7414)	(great, 10665)
1	(bad, 435)	(nice, 954)	(great, 1751)	(good, 6670)	(good, 5670)
2	(small, 363)	(great, 908)	(nice, 1601)	(nice, 4479)	(nice, 4085)
3	(nice, 329)	(small, 628)	(clean, 957)	(clean, 2870)	(excellent, 3209)
4	(great, 319)	(clean, 519)	(small, 944)	(small, 2349)	(clean, 3195)
5	(old, 303)	(bad, 502)	(little, 741)	(little, 2074)	(best, 2586)
6	(worst, 284)	(little, 472)	(bad, 523)	(comfortable, 1611)	(wonderful, 2583)
7	<u>(terrible, 256)</u>	(old, 359)	(free, 421)	(helpful, 1591)	(helpful, 2478)
8	(clean, 236)	(beautiful, 351)	(best, 394)	(excellent, 1472)	(little, 2261)
9	(new, 235)	(best, 335)	(old, 387)	(free, 1328)	(small, 2233)
10	(horrible, 220)	(new, 328)	(comfortable, 368)	(large, 1280)	(comfortable, 2166)
11	(stay, 204)	(big, 281)	(overall, 368)	(friendly, 1248)	(beautiful, 2029)
12	(little, 204)	(hot, 281)	(ok, 365)	(beautiful, 1199)	(fantastic, 2020)
13	(available, 185)	(stay, 277)	(friendly, 356)	(best, 1172)	(friendly, 1982)
14	(open, 181)	(free, 274)	(beautiful, 355)	(big, 1068)	(free, 1825)
15	(best, 174)	(sure, 269)	(large, 354)	(quiet, 1000)	(large, 1766)
16	(hot, 167)	(poor, 267)	(helpful, 354)	(new, 994)	(new, 1605)
17	(poor, 164)	(open, 262)	(big, 352)	(wonderful, 910)	(stay, 1529)
18	(free, 150)	(ok, 242)	(new, 344)	(overall, 856)	(quiet, 1383)
19	(beautiful, 148)	(available, 226)	(main, 336)	(main, 844)	(perfect, 1370)

TRY 2: Improve common words

Feedback:

1. Seems improved, get something useful.

2. Can't produce n-grams in the same time, e.g. 'great location' and 'good room service' can't be displayed at the same time.

3. Some useless words like 'san juan' and 'punta cana' are included, should try to remove them.

common_bigrams_by_ratings

	1	2	3	4	5
0	((punta, cana), 150)	((punta, cana), 198)	((great, location), 248)	((great, location), 808)	((staff, friendly), 968)
1	((credit, card), 110)	((room, service), 140)	((staff, friendly), 235)	((staff, friendly), 726)	((great, location), 957)
2	((customer, service), 103)	((staff, friendly), 124)	((punta, cana), 201)	((hotel, great), 557)	((great, hotel), 934)
3	((stay, hotel), 101)	((star, hotel), 121)	((good, location), 166)	((punta, cana), 548)	((hotel, great), 866)
4	((star, hotel), 95)	((air, conditioning), 120)	((room, clean), 136)	((walking, distance), 512)	((recommend, hotel), 818)
5	((5, star), 88)	((5, star), 116)	((room, service), 132)	((friendly, helpful), 464)	((friendly, helpful), 735)
6	((hotel, staff), 85)	((stay, hotel), 102)	((walking, distance), 131)	((great, hotel), 450)	((highly, recommend), 727)
7	((room, service), 83)	((san, juan), 97)	((hotel, great), 130)	((good, value), 405)	((walking, distance), 704)
8	((worst, hotel), 82)	((great, location), 96)	((san, juan), 125)	((nice, hotel), 400)	((hotel, staff), 618)
9	((stay, away), 75)	((stayed, hotel), 90)	((stayed, hotel), 124)	((room, clean), 396)	((punta, cana), 582)
10	((hotel, room), 75)	((make, sure), 87)	((good, value), 120)	((great, time), 387)	((stayed, hotel), 550)
11	((stayed, hotel), 71)	((hotel, room), 85)	((hotel, good), 114)	((staff, helpful), 382)	((place, stay), 532)
12	((air, conditioning), 70)	((hotel, staff), 79)	((make, sure), 112)	((minute, walk), 373)	((stay, hotel), 506)
13	((got, room), 69)	((room, small), 79)	((location, hotel), 110)	((recommend, hotel), 361)	((staff, helpful), 504)
14	((travel, agent), 66)	((good, location), 77)	((star, hotel), 109)	((room, service), 360)	((hotel, stayed), 495)
15	((san, juan), 61)	((room, clean), 76)	((5, star), 104)	((hotel, good), 358)	((definitely, stay), 494)
16	((hotel, stayed), 60)	((hotel, good), 74)	((location, great), 103)	((stayed, hotel), 358)	((new, york), 477)
17	((booked, hotel), 59)	((customer, service), 72)	((room, small), 103)	((hotel, staff), 348)	((minute, walk), 471)
18	((hot, water), 58)	((room, ready), 68)	((minute, walk), 99)	((good, location), 340)	((room, service), 469)
19	((desk, staff), 56)	((4, star), 66)	((nice, hotel), 98)	((place, stay), 322)	((great, place), 454)

TRY 3: Get keywords

Gensim: an open-source library for unsupervised topic modeling and natural language processing.

bad_features_1_star.most_common(20)	bad_features_2_stars.most_common(20)	bad_features_3_stars.most_common(20)	good_features_4_stars.most_common(20)	good_features_5_stars.most_common(20)
[('hotel', 566), ('room', 506), ('rooms', 219), ('hotels', 173), ('stay', 161), ('staff', 108), ('night', 100), ('service', 86), ('place', 83), ('day', 72), ('resort', 69), ('food', 62), ('like', 60), ('staying', 60), ('stayed', 59), ('time', 58), ('bed', 52), ('desk', 50), ('days', 47), ('good', 46)]	[('room', 681), ('hotel', 659), ('rooms', 303), ('hotels', 200), ('stay', 199), ('night', 147), ('good', 128), ('service', 122), ('resort', 110), ('staff', 104), ('beach', 98), ('day', 92), ('food', 84), ('time', 77), ('stayed', 73), ('nice', 73), ('like', 70), ('location', 65), ('days', 59), ('staying', 57)]	[('hotel', 878), ('room', 730), ('rooms', 370), ('good', 296), ('stay', 228), ('hotels', 219), ('nice', 195), ('great', 186), ('location', 173), ('beach', 139), ('night', 137), ('staff', 113), ('resort', 104), ('service', 96), ('clean', 88), ('stayed', 87), ('small', 86), ('breakfast', 86), ('pool', 84), ('time', 81)]	[('hotel', 2556), ('room', 1868), ('great', 1040), ('good', 920), ('stay', 825), ('rooms', 727), ('nice', 611), ('hotels', 607), ('location', 483), ('staff', 454), ('beach', 380), ('night', 335), ('stayed', 279), ('breakfast', 270), ('resort', 268), ('clean', 259), ('service', 248), ('little', 205), ('time', 203), ('walk', 202)]	[('hotel', 4142), ('room', 2329), ('great', 1728), ('stay', 1428), ('hotels', 998), ('rooms', 917), ('staff', 842), ('good', 684), ('stayed', 576), ('location', 512), ('nice', 494), ('service', 438), ('excellent', 433), ('night', 403), ('beach', 384), ('breakfast', 369), ('day', 351), ('place', 349), ('staying', 329), ('resort', 325)]

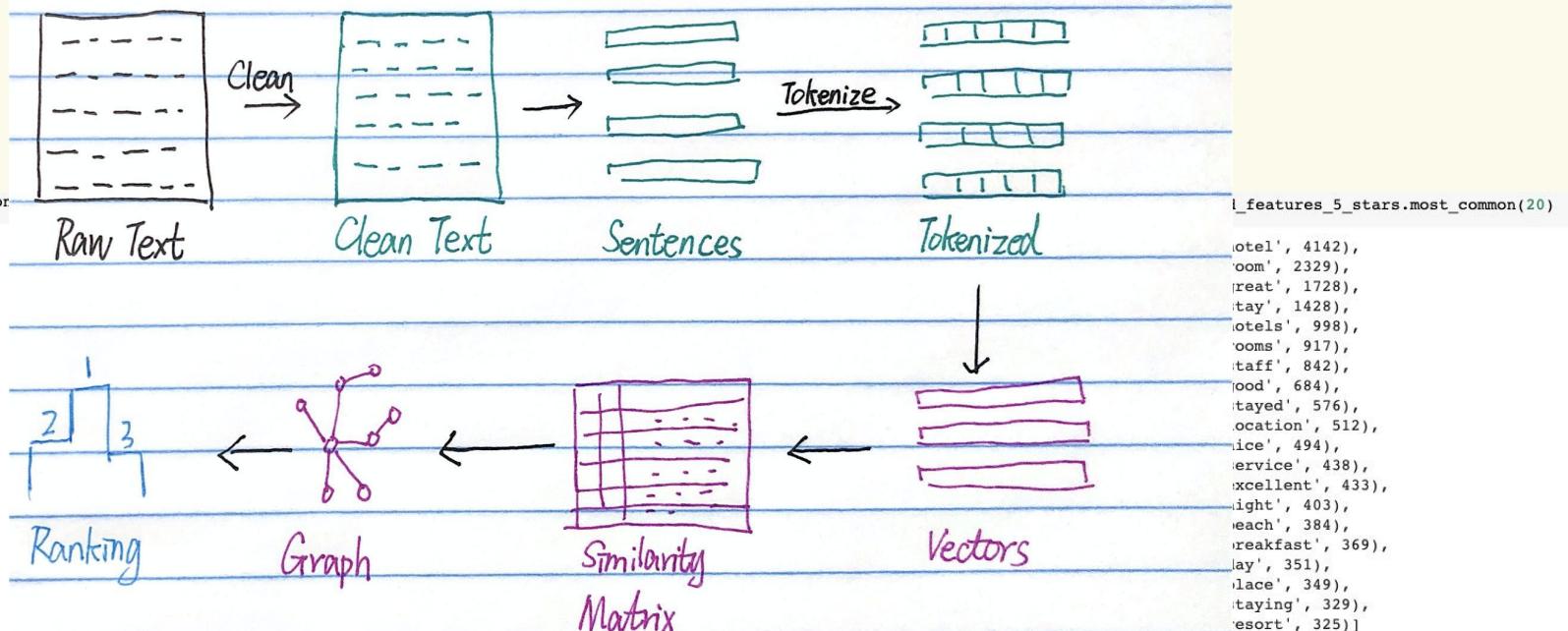
TRY 3: Get keywords

I love dog.



{ I: 1 love: 2 dog: 3 }

for unsupervised
language processing.



TRY 3: Get keywords

Gensim: an open-source library for unsupervised topic modeling and natural language processing.

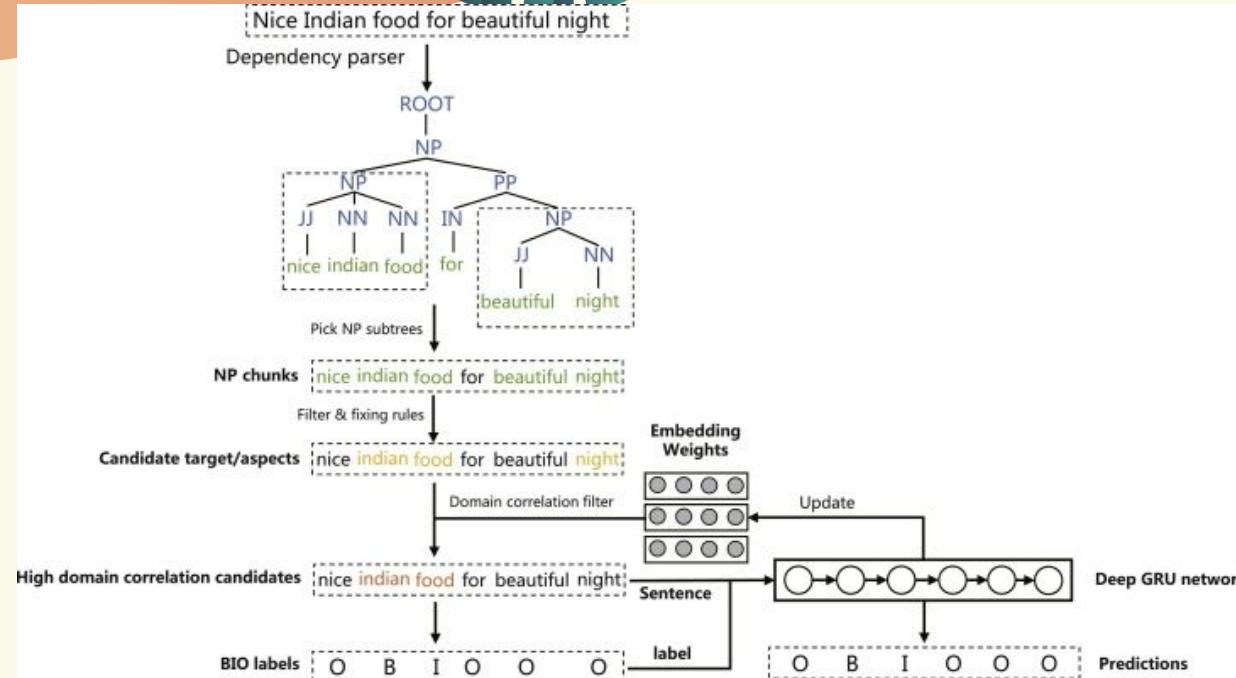
Feedback:

1. Remove stopwords.
 2. Smarter.
 3. Lemmatization
-
4. Still not accurate.
 5. Slower.

bad_features_1_star.most_common(20)	bad_features_2_stars.most_common(20)	bad_features_3_stars.most_common(20)	good_features_4_stars.most_common(20)	good_features_5_stars.most_common(20)
[('hotel', 566), ('room', 506), ('rooms', 219), ('hotels', 173), ('stay', 161), ('staff', 108), ('night', 100), ('service', 86), ('place', 83), ('day', 72), ('resort', 69), ('food', 62), ('like', 60), ('staying', 60), ('stayed', 59), ('time', 58), ('bed', 52), ('desk', 50), ('days', 47), ('good', 46)]	[('room', 681), ('hotel', 659), ('rooms', 303), ('hotels', 200), ('stay', 199), ('night', 147), ('good', 128), ('service', 122), ('resort', 110), ('staff', 104), ('beach', 98), ('day', 92), ('food', 84), ('time', 77), ('stayed', 73), ('nice', 73), ('like', 70), ('location', 65), ('days', 59), ('staying', 57)]	[('hotel', 878), ('room', 730), ('rooms', 370), ('good', 296), ('stay', 228), ('hotels', 219), ('nice', 195), ('great', 186), ('location', 173), ('beach', 139), ('night', 137), ('staff', 113), ('resort', 104), ('service', 96), ('clean', 88), ('stayed', 87), ('small', 86), ('breakfast', 86), ('pool', 84), ('time', 81)]	[('hotel', 2556), ('room', 1868), ('great', 1040), ('good', 920), ('stay', 825), ('rooms', 727), ('nice', 611), ('hotels', 607), ('location', 483), ('staff', 454), ('beach', 380), ('night', 335), ('stayed', 279), ('breakfast', 270), ('resort', 268), ('clean', 259), ('service', 248), ('little', 205), ('time', 203), ('walk', 202)]	[('hotel', 4142), ('room', 2329), ('great', 1728), ('stay', 1428), ('hotels', 998), ('rooms', 917), ('staff', 842), ('good', 684), ('stayed', 576), ('location', 512), ('nice', 494), ('service', 438), ('excellent', 433), ('night', 403), ('beach', 384), ('breakfast', 369), ('day', 351), ('place', 349), ('staying', 329), ('resort', 325)]

Future work

Aspect term and opinion target extraction.



-- [A hybrid unsupervised method for aspect term and opinion target extraction](#)

OUR TEAM



JAMES PATTERSON

You can replace the image on the screen with your own



JOHN JAMES

You can replace the image on the screen with your own



JIMMY TIMMY

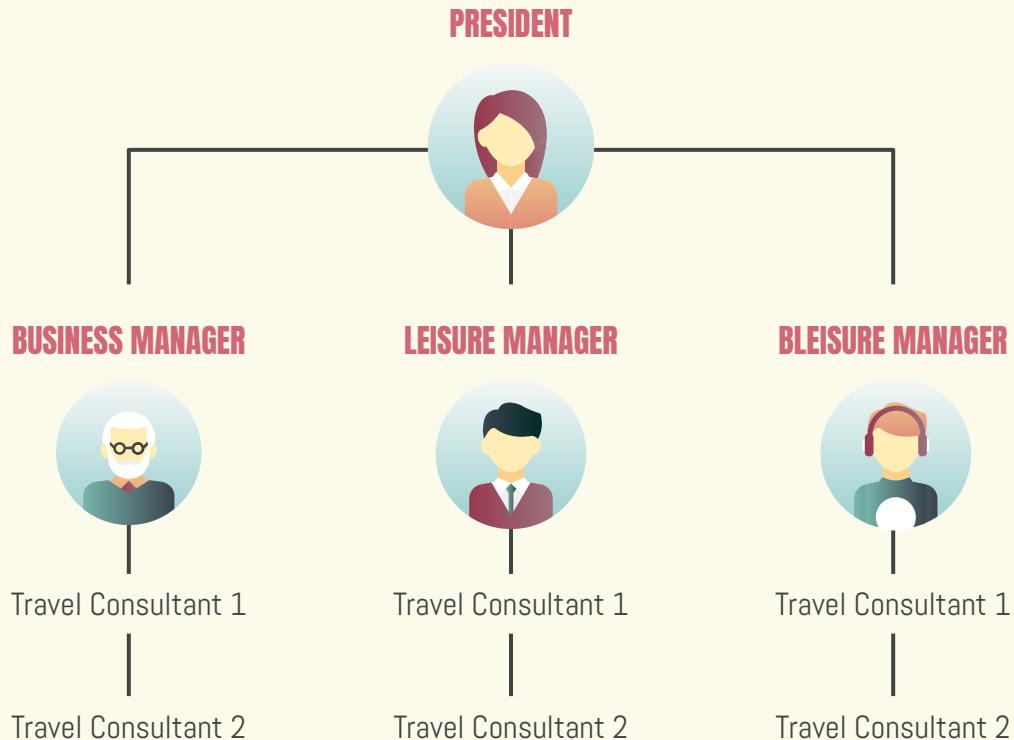
You can replace the image on the screen with your own



JENNA DOE

You can replace the image on the screen with your own

ORGANIZATIONAL CHART





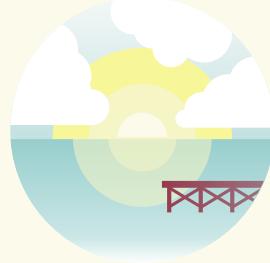
JUPITER

Yes, this is the ringed one. It's composed mostly of hydrogen and helium



VENUS

Venus has a beautiful name, but it's terribly hot, even hotter than Mercury



MARS

Despite being red, Mars is a cold place, not hot. It's full of iron oxide dust

05. OPERATIONAL PLAN



OUR PROCESS

MARS



Despite being red,
Mars is a cold place,
not hot

JUPITER



Jupiter is the biggest
planet in our Solar
System

NEPTUNE



Neptune is the
fourth-largest planet
in our Solar System

SATURN

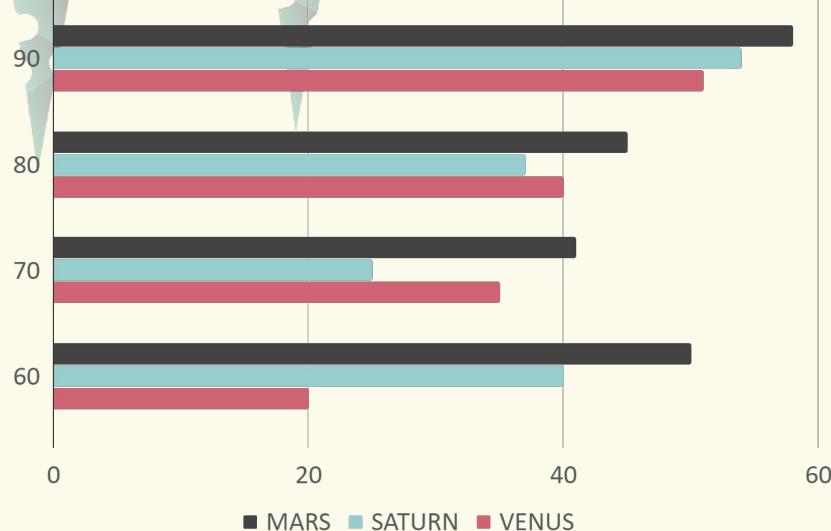


Saturn is composed
mostly of hydrogen
and helium

06. FINANCIAL PLAN

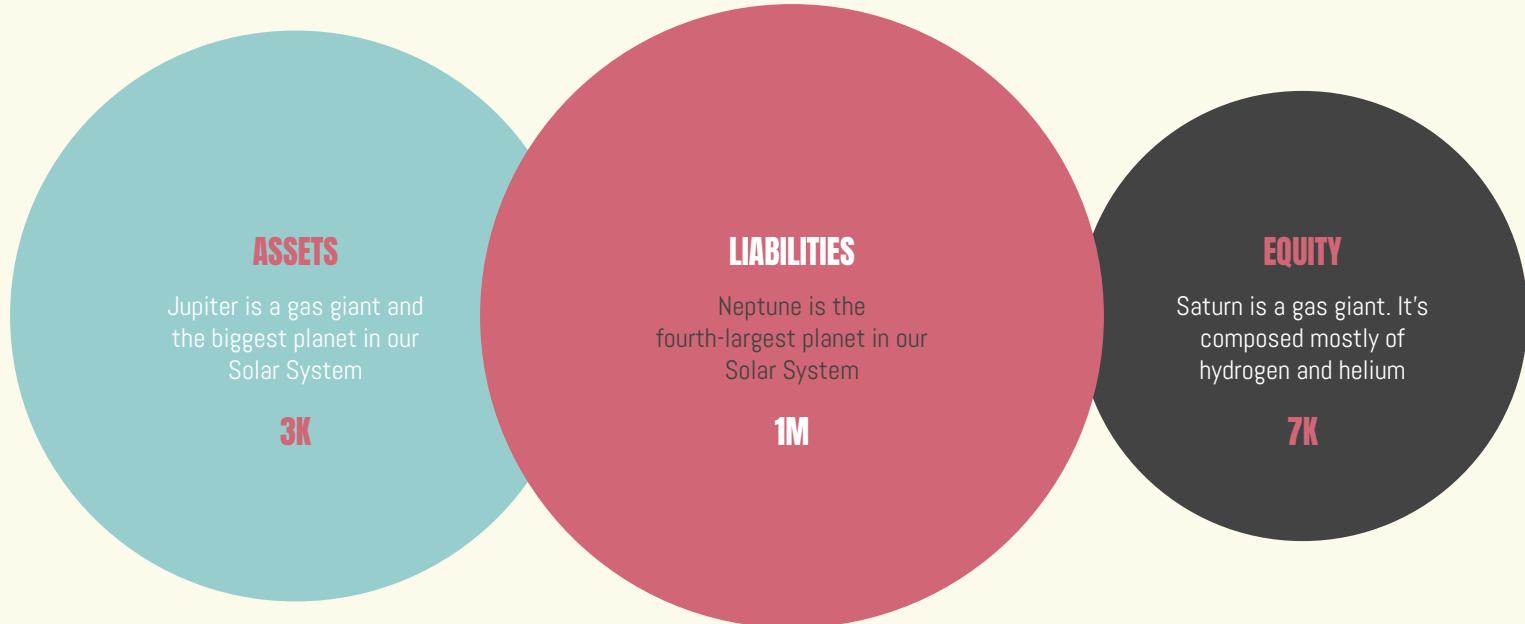


PREDICTED GROWTH



If you want to modify this graph, click on it, follow the link, change the data and replace it

BALANCE SHEETS





THANKS

THANKS

Does anyone have any
questions?

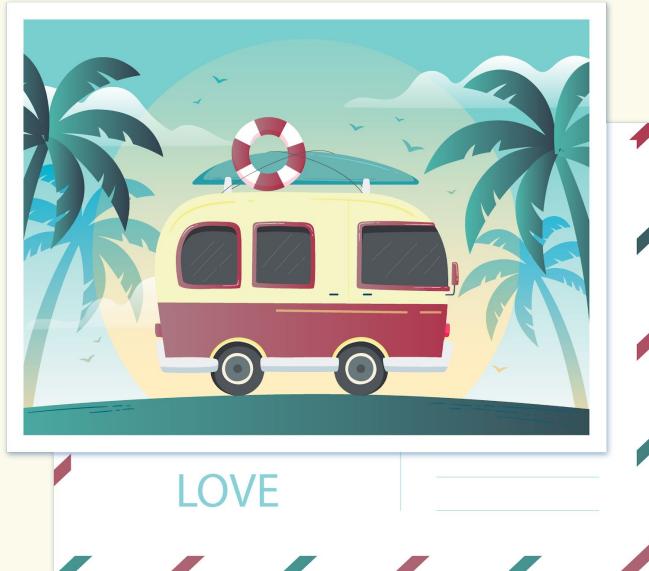
addyouremail@freepik.com
+91 620 421 838
yourcompany.com



CREDITS

This is where you give credit to the ones who are part of this project.

- ◀ Presentation template by [Slidesgo](#)
- ◀ Icons by [Flaticon](#)
- ◀ Infographics and images by [Freepik](#)
- ◀ Author introduction slide photo created by Freepik
- ◀ Text & Image slide photo created by Freepik.com



Did you like the resources on this template? Get them for free at our other websites.

- Tent Icon
- Shack Icon
- Car Icon
- Equality Icon
- Motivation Icon Pack
- Aviation Icon Pack
- Landscapes collection Icon Pack
- Marketing Icon Pack
- Outdoor Activities Icon Pack
- Profession avatars Icon Pack
- Flat vintage travel logo Free Vector
- Travel landing page flat design Free Vector
- Flat tropical summer holiday postcard Free Vector
- Flat tropical summer holiday postcard Free Vector
- Outdoor activities landing page template Free Vector
- Hand drawn camping landing page template Free Vector

Instructions for use

In order to use this template, you must credit **Slidesgo** by keeping the Credits slide.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs:

<https://slidesgo.com/faqs>

Fonts & colors used

This presentation has been made using the following fonts:

Anton

(<https://fonts.google.com/specimen/Anton>)

Abel

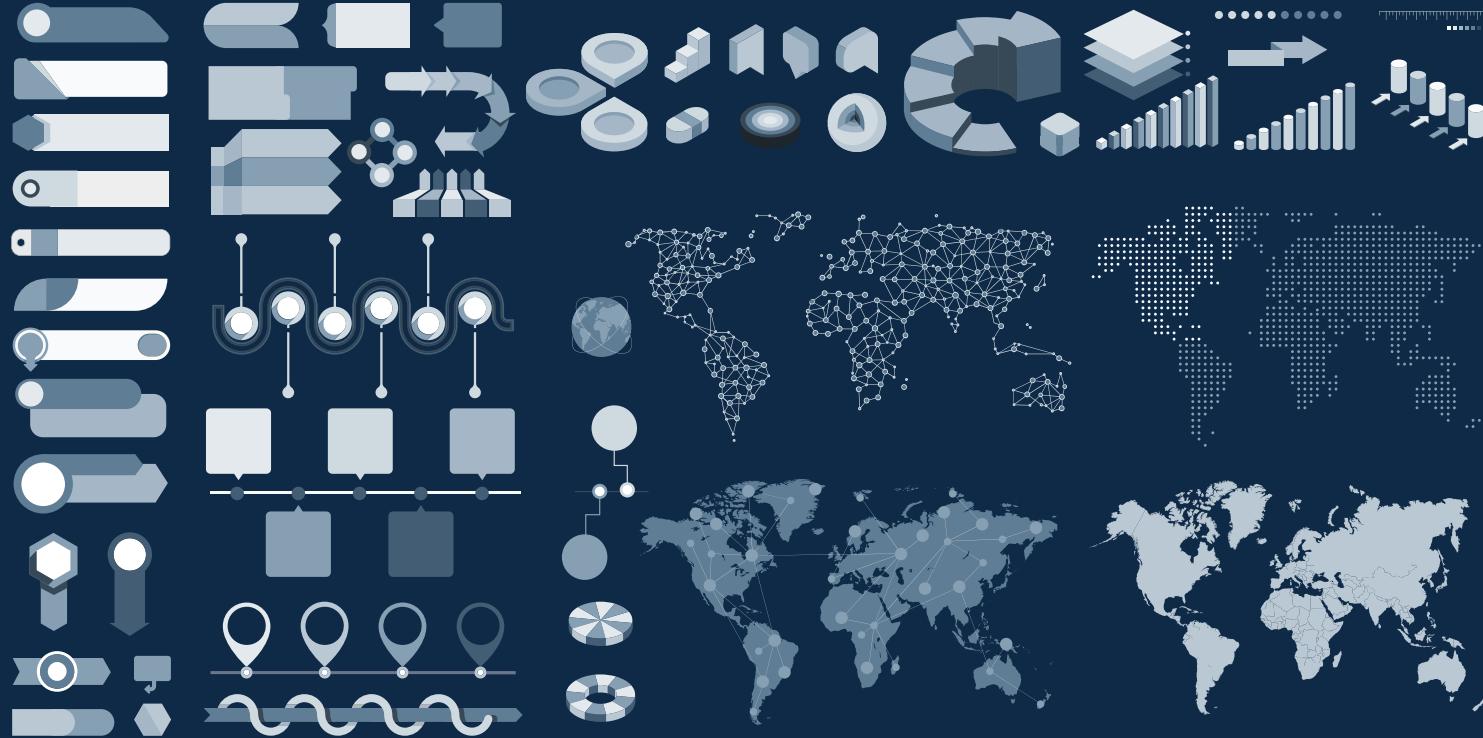
(<https://fonts.google.com/specimen/Abel>)

#d16676

#97cecd

#434343

#f5f5c6



Use our editable graphic resources...

You can easily resize these resources, keeping the quality. To change the color, just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want. Don't forget to group the resource again when you're done.



...and our set of editable icons

You can resize these icons, keeping the quality.

You can change the stroke and fill color; just select the icon and click on the paint-bucket/pen



Business Icons



Avatar Icons



Creative Process Icons



Educational Process Icons



Help & Support Icons



Medical Icons



Nature Icons



Performing Arts Icons



SEO & Marketing Icons



Teamwork Icons



