

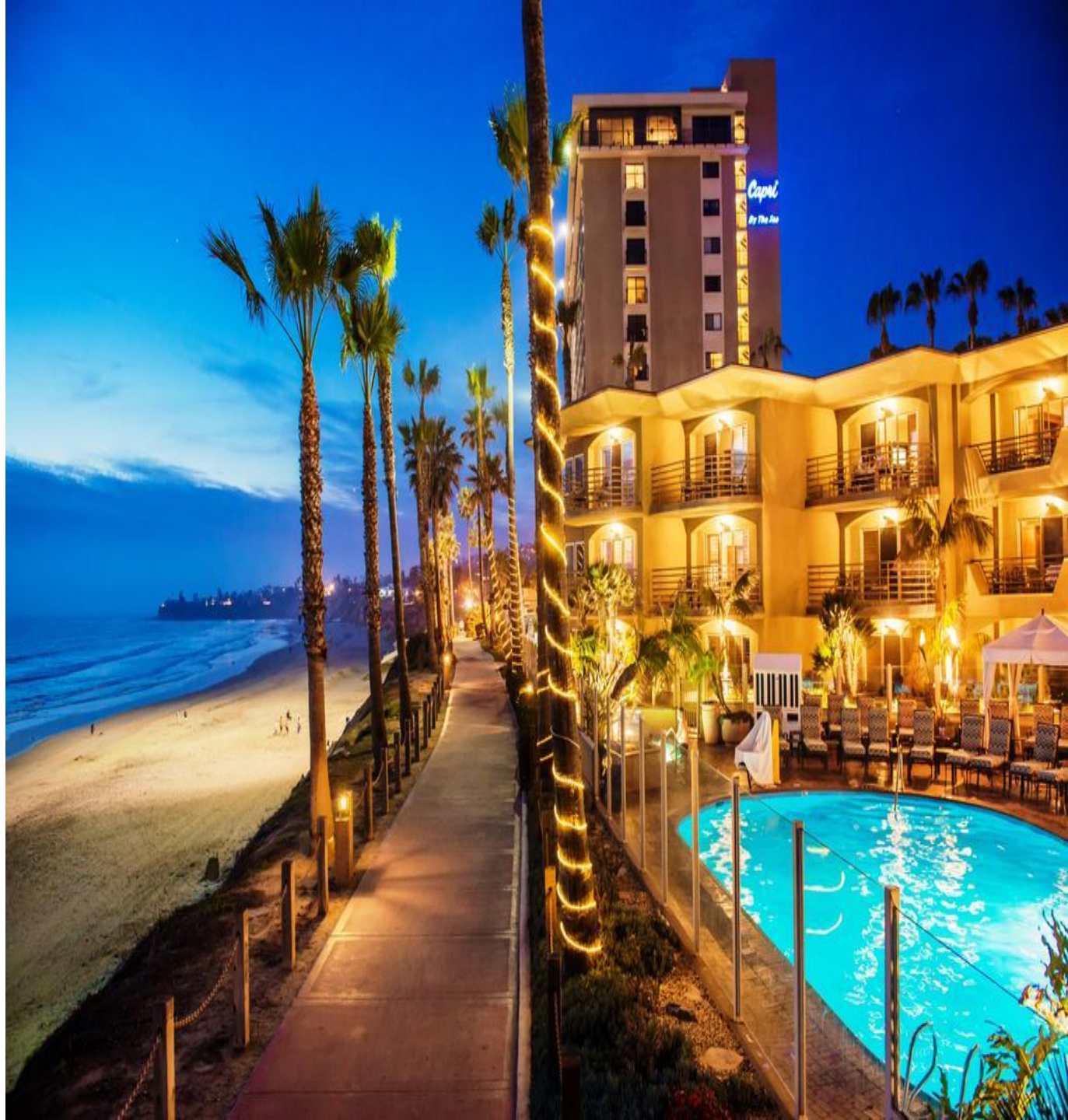
ניתוח דירוגי מלונות בסן דייגו, ארה"ב

כריית טקסט

רוני רמון-גורן

22.1.2021

יונתן הופמן 304867930






מבוא

- חוות דעת אורחים כבר מזמן לא מסתכמות רק בדירוג של מלון ע"י סימון מספר ★. בשנים האחרונות נכנסו סקרי חוות דעת ואתרי שביעות רצון כמו **TripAdvisor**, שבאמצעותם אורחי המלון משאירים חוות דעת מפורטות כדי לסייע לאנשים אחרים לקבל החלטות על בחירת המלון.
- לא רק אנשים המחפשים מקום לינה יכולים להפיק תועלת מחוות הדעת האלו, אלא גם מנהלי בתי המלון עצמם.
- באמצעות חוות הדעת – בין אם שליליות ובין אם חיוביות, מנהלי בתי המלון יכולים לזהות את נקודות החולשה ולפעול בהתאם על מנת לשפרן. בנוסף, ניתן לאתר את נקודות החוזק של המלון ולמצב את המלון בהתאם.

בתי מלון

דירוג הכוכבים או ביקורת גולשים: מה משפיע על בחירת המלון?

שר התיירות יריב לוין משוכנע כי מלון שאינו מדורג נמצא בעמדת נחיתות, אולם בעלי המלונות טוענים כי חוות-הדעת של הגולשים ברשת משפיעות על הנופשים יותר ממספר הכוכבים

    | 

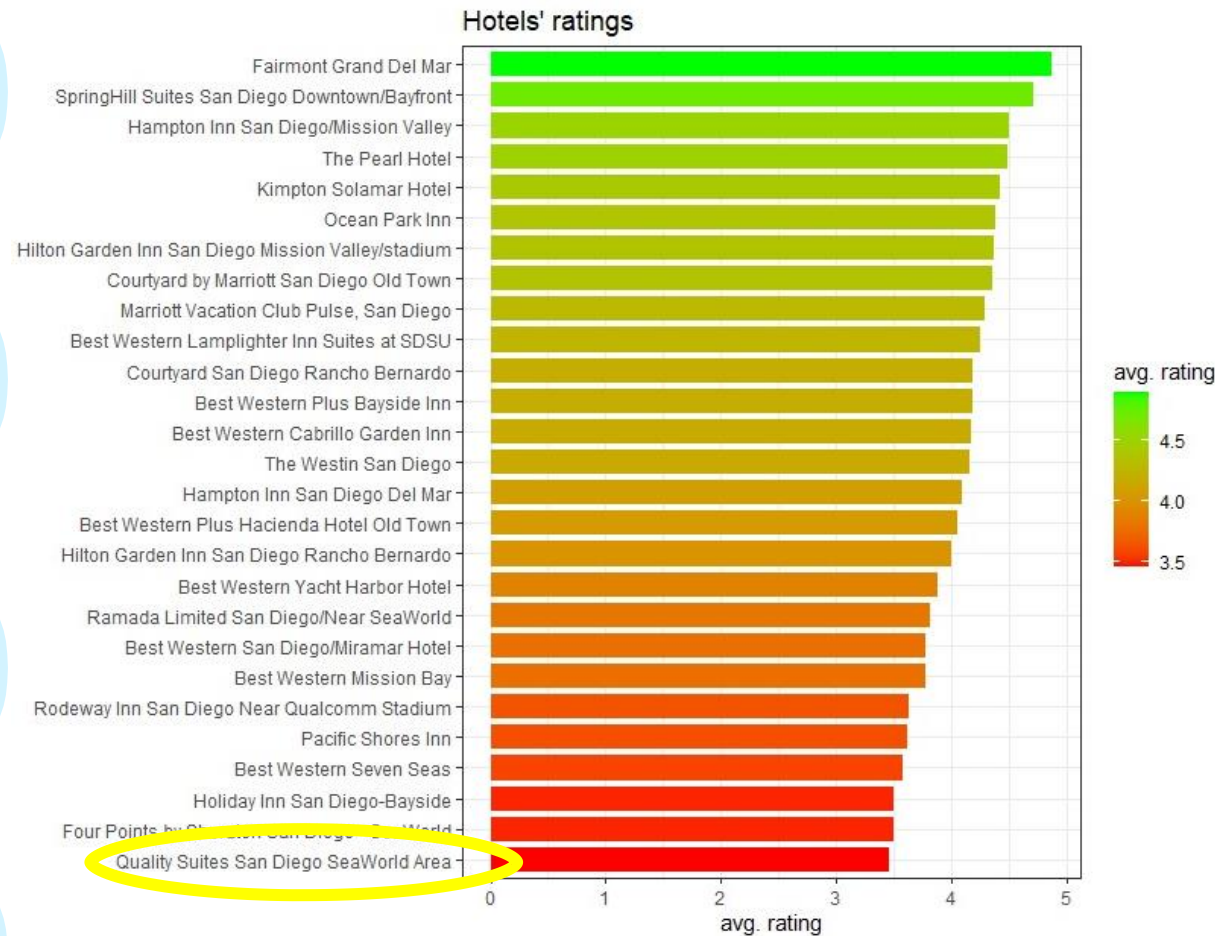
23.09.2016 מיכל רוזניימוביץ'

הגדרת הצורך והמטרה

○ הצורך – קיים מידע שימושי רב הטמון בביקורות ברשת. האתגר הוא להפיק ידע שימושי בצורה שיטתית ומאורגנת.

○ המטרה – להפיק ידע שימושי למנהל המלון ע"י שימוש בטכניקות מעולם כריית הטקסט.

○ מקרה בוחן – המלון עם ממוצע דירוג הכוכבים הנמוך ביותר בעיר: Quality Suites San Diego Seaworld Area

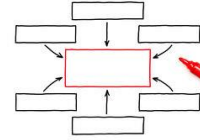




סקירת ספרות

○ במאמרם של Riyanarto Sarno & Reza Amalia Priyantina (2019), העוסק בנושא ה- **Sentiment Analysis**, נמצא כי למלון יש 5 היבטים שמשפיעים על דירוגו- מיקום, ארוחה, שירות, נוחות וניקיון. מחקרם מציג שיטות לקביעת סנטימנט הביקורת בהתאם להיבטי המלונות. ביקורות על מלונות מעובדות מראש לרשימת מונחים. החוקרים משתמשים ב- **LDA** ואז בחישוב מדידת הדמיון, כמו כן, רשימת המונחים מורחבת על ידי שימוש בשיטת **Term Frequency-Inverse Cluster Frequency (TF-ICF)**. לבסוף, סיווג של סנטימנט הלקוח (מרוצה/לא מרוצה) מתבצע באמצעות השילוב של הטמעת מילים וזיכרון לטווח ארוך-קצר (**LSTM**). התוצאות מראות שהשיטה המוצעת יכולה לסווג את הביקורות בחמשת ההיבטים המלונאים.

○ במאמרם של Ian Sutherland, Youngseok Sim, Seul Ki Lee, Jaemun Byun and Kiattipoom Kiatkawsin (2020), העוסק בנושא ה- **topic modeling**, מיישמים גישה אינדוקטיבית על ידי שימוש בנתוני טקסט גדולים ולא מובנים של 104,161 ביקורות מקוונות של לקוחות אירוח קוריאניים כדי להבין אילו קטגוריות מעניינות את האורחים. באמצעות **LDA**, מחלצים ומאמתים נושאים בעלי עניין. התוצאות מאששות ספרות קיימת בכך שממדים, כמו מיקום ואיכות השירות, הם חשובים. עם זאת, המחקר הנ"ל מרחיב את הממדים ומוצא כי יש חשיבות לסוג הלינה והאם מקום האירוח נמצא בכפר/עיר.



סט הנתונים

- אתר **kaggle** – "**Hotel Reviews**". שני מערכי נתונים המכילים יחדיו 3,400 בתי מלון בארה"ב ו-20,000 ביקורות מאתר **Trip Advisor** ממאגר המידע של **Datafiniti**. המאגר התעדכן בין דצמבר 2018 עד מאי 2019.
- מערך הנתונים כולל מיקום בתי מלון, שם, דירוג, נתוני ביקורת ועוד.
- מיקוד במלונות השוכנים בסן דייגו (1195 ביקורות, 27 מלונות)

דוגמאות לביקורות

Hampton Inn San Diego/Mission Valley

views.title	views.text
Wind chill in bed	The A/C unit under the window blows air directly on the bed. I like it cold when I sleep!
Excellent Stay!	We really enjoyed our stay at the Hampton Inn Mission Valley. The staff were all very friendly.
Nice Property	This is a pretty new property that is conveniently located to Seaworld, downtown, and I
Parking	The hotel is ok but considering you pay 19 a day for parking it should be better. There
Awesome stay	I have to say that I was very pleased with my stay at this hotel. The customer service was
Excelent Hotel	We enjoy a lot our stay, I went with my family and it was perfect. We ate a delicious breakfast
Stay in San Diego	The hotel is clean and the staff was excellent. There is no shuttle service but you can take
Very Good Stay	Very good hotel with easy access. Parking can be tight for 14/nights when coming back
Great value and even better location	Stayed here for a week (Sat. to Sat.) while visiting San Diego. The hotel was clean and
Nicest Hampton EVER	Likely the fanciest Hampton I have ever stayed in. Very large property, very modern and
Family Vacation	The location of the hotel was central to all of the surrounding attractions. We definitely
Nice rooms but horrible breakfast	This hotel is in a great location and the rooms are quite nice. The large was extremely
Great Hotel and Perfect Location	- Great Hotel and Staff - Perfect location - Right off the Freeway, close to all restaurants
Meeting Room	Location is easy to get to and even though the room is small it met our needs. Staff is
Amazing hotel n very cleaning. Staff are so generous willing to help	Best hotel ever. Right by mission bay beach. Rooms are beautiful extremely clean. Pa
hampton inn mission valley san diego	super friendly staff, average breakfast, free wifi, 16.00 per night parking, good AC in ro
Excellent service	Excellent service.. Very welcoming staff. Always smiling and willing to assist in any way
Great location and Very Clean property	Stayed at this Hampton with my family for College Tour at UCSD and SDSU. It's a great
Excellent service	I was traveling to see my son graduate from Marine boot camp and I couldn't be more

Quality Suites San Diego Seaworld Area

views.title	views.text
Not what was in the images or described by Hotel staff	I basically live out of hotels and this is the WORST I have ever seen. Starting with parking you are paying 12.00 for a spot that you have
Not the best, not the worse	We stayed for two nights and chose this location because it was close to things we wanted to do. My first issue is the parking and
Our birthday stay	Staff was friendly and professional. Rooms were nice and clean. Pool and jacuzzi was great. Breakfast was delicious. We enjoyed
Lower your expectations and you may enjoy your stay....	We stayed at this hotel last year for 10 nights and while the rooms weren't completely up to date, it was clean, had extra room, breakfast
Great Breakfasts	Great price for a 2 room suite with fridge microwave + TV's in each. Sliding doors can separate second room from bedroom with kitchen
This place gets a bad rap, but not that bad.	Ok people, before the bashing begins just remember you are paying roughly 129/night and you get free WiFi and a hot breakfast to
Good for the money	We picked this hotel due to its proximity to Sea World. The room was outdated but large and clean. The only complaint we had was
Very impressed with Quality Suites!	My wife and I recently spent a night at this property. We had never stayed at a Quality Suites property before. We came away very impressed
No ice	We stayed here on October 9-12 and the entire weekend the ice machines and soda machine on our floor was out of order. I did not
A family experience	The King suite with 2 double beds is great, although very tight space. The beds are bunkbeds in a separate room with door. Each
Bay area	The TV and refrigerator was not working. Excellent customer service to fix the problem. The location is close to restaurants and shopping
wonderful experience	we had an amazing time at this hotel definitely coming back for our next family vacation room was nice and cold and clean and for a decent
Business Trip 3/2/2016	Great Stay-I was provided an upgraded suite room. Very nice. Room suite very spacious. All hotel amenities were readily available
Great	It was an excellent stay! Perfect location for my conference. The person at the front desk rather serious. The maid service excellent, the
Very Reasonable Family Hotel	We stayed here on a 9 night vacation in San Diego. We had a tremendous deal that could not be beat for the amenities this hotel provided
Has potential but needs some work	Very unique room configuration especially if you are a family. Doors open to the outside which all face inward to the open courtyard
Decent hotel, decent location	Spent the first three nights of a family trip here, using it as a base to see Legoland and Seaworld. The hotel itself is fine (we had a
Enjoyed our time here.	I think it has been said previously if you are expecting a Hilton at this price, you'll be disappointed. Overall this is a very good property
Hotel in good area	Hotel is ok, but need upgrades in the rooms, the carpet, tv. Is small the bathroom can use up grade tile in floor. When you call from
Not what the pictures showed	I anticipated this resort having interior corridors like the pictures led me to believe however it did not. The rooms were very old and
Great and Affordable	The Hampton Inn Mission Valley was clean and the amenities were great. The staff were very friendly and the location was perfect.



ארגון הטקסט

```
# create cleaning function
clean.vec<-function(text.vec){
  custom.stopwords <- c(stopwords("english"),'san diego', 'hotel', 'stay','regards','regarding')
  text.vec <- tolower(text.vec)
  text.vec<- gsub("rooms","room",text.vec)
  text.vec <- removeWords(text.vec, custom.stopwords)
  text.vec <- removePunctuation(text.vec)
  text.vec <- removeNumbers(text.vec)
  text.vec <- stripwhitespace(text.vec)
  return(text.vec)
}
```

```
... -----
# google translate
reviews.df$reviews.text[grepl("con mi familia",reviews.df$reviews.text, fixed = TRUE)] <- "Excellent 3
reviews.df$reviews.text[grepl("Ho soggiornato qui con la mia",reviews.df$reviews.text, fixed = TRUE)] <
reviews.df$reviews.text[grepl("calidad precio es ptima",reviews.df$reviews.text, fixed = TRUE)] <- "The
reviews.df$reviews.text[grepl("Desayuno incluido habitaciones",reviews.df$reviews.text, fixed = TRUE)]
```

○ מציאת הערים עם הביקורות הרבות ביותר

○ צמצום הביקורות לעיר סן דייגו עם הקטגוריות הרלוונטיות

○ הסרת מלונות עם פחות מ-10 ביקורות

○ יצירת גרף המציג את דירוגי בתי המלון

○ ניקוי הטקסט-

○ הורדת stopwords ומילים נוספות שראינו כי חוזרות על עצמן

○ הפיכת כל האותיות לאותיות קטנות

○ הורדת סימני פיסוק

○ הורדת מספרים

○ צמצום רווחים

○ תרגום ביקורות שלא בשפה האנגלית

○ מחיקה של תגובות המלונות לביקורות

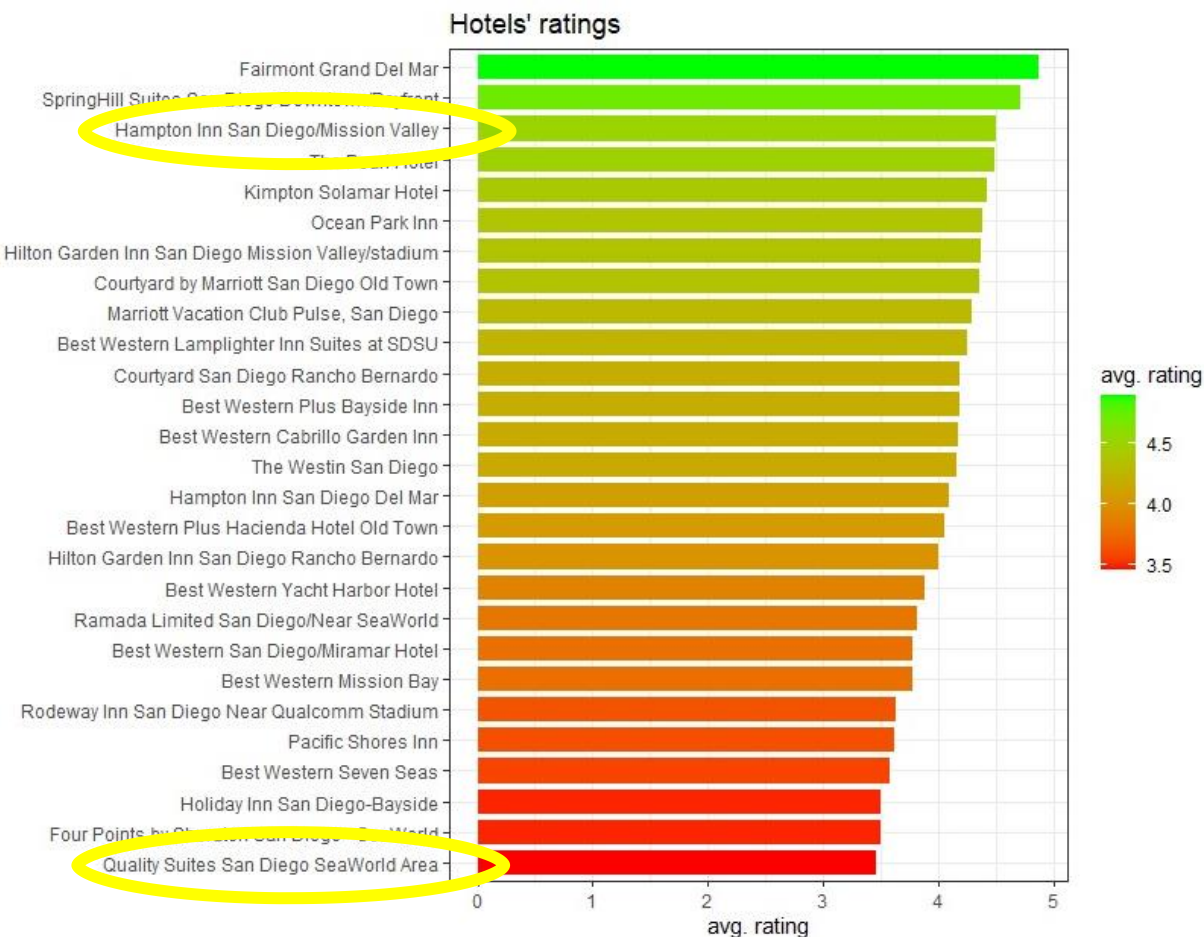


שיטות הניתוח

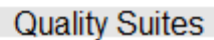
○ השוואה למלון עם דירוג מוצלח יותר: Hampton Inn San Diego/Mission Valley

○ ניתוח סנטימנט והשוואה בין ביקורות חיוביות לשליליות

○ מציאת נקודות חוזק וחולשה ע"י פירוק כל ביקורת למשפטים, ביצוע חלוקה לנושאים וניתוח סנטימנט

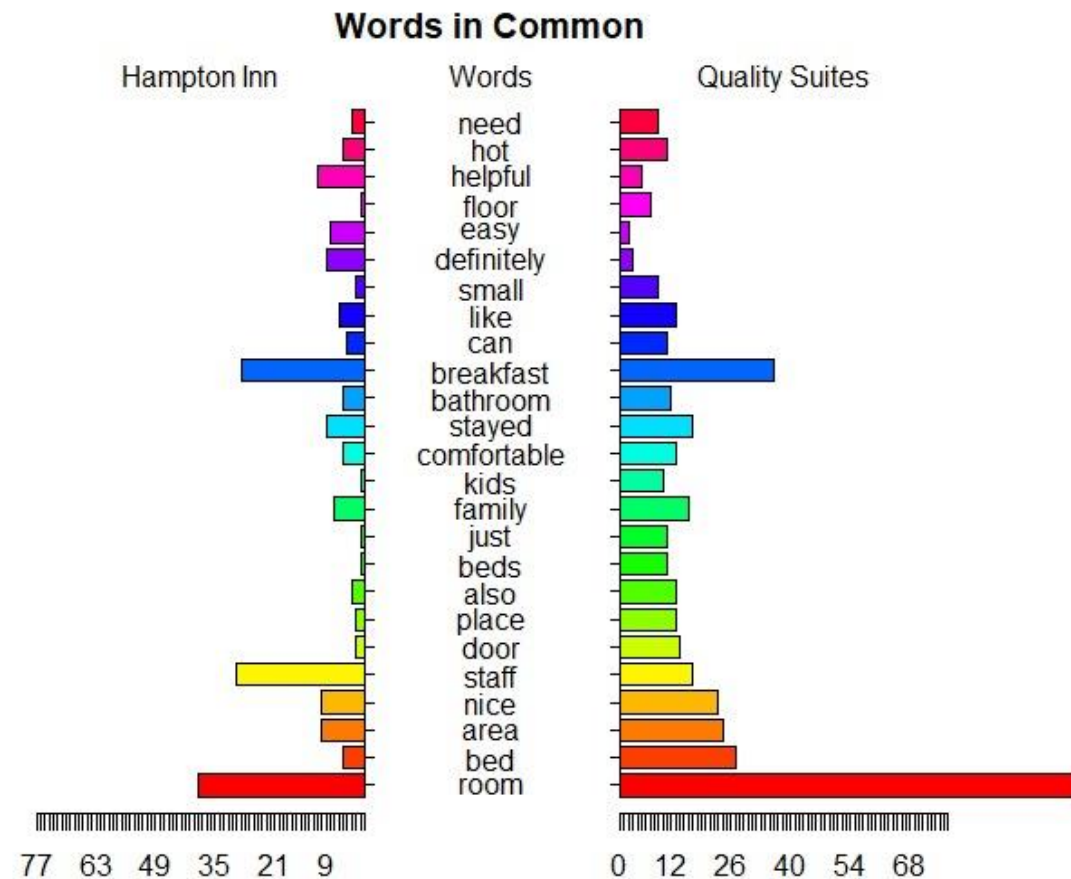


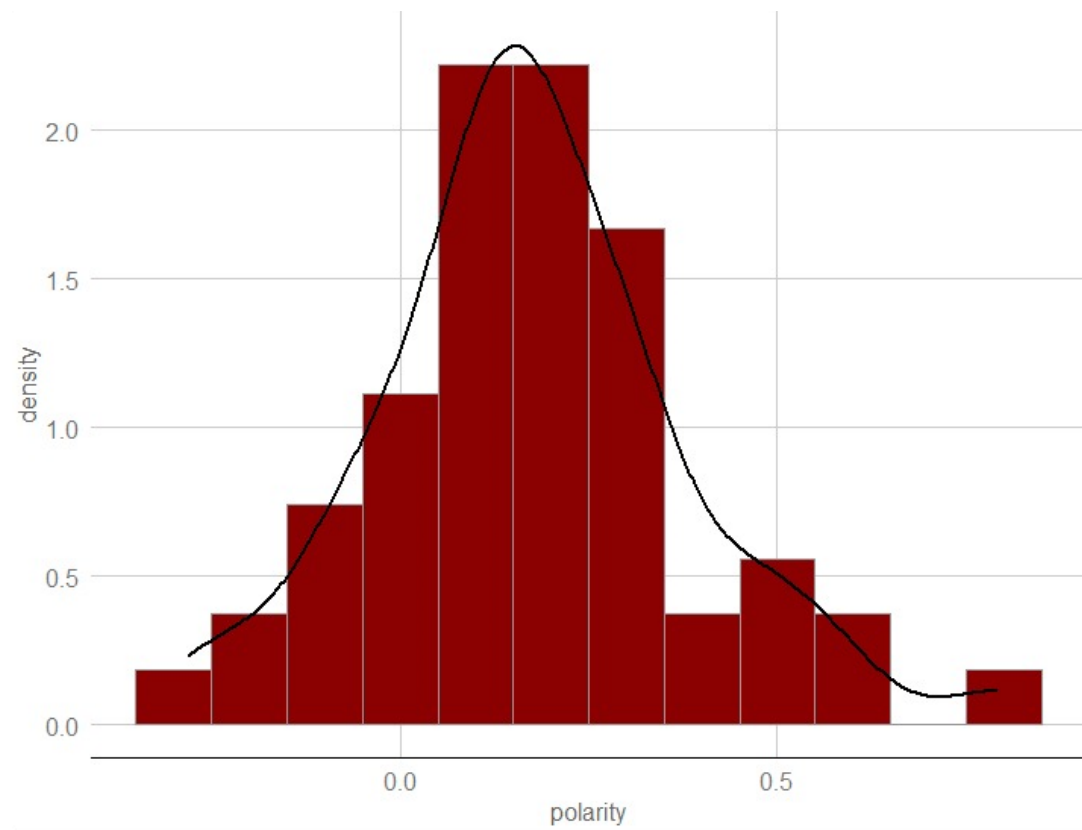
Hampton Inn



○ בסיס לבחינה בהמשך

פערי שכיחויות





נִיתּוּחַ סֵנְטִימֵנְט

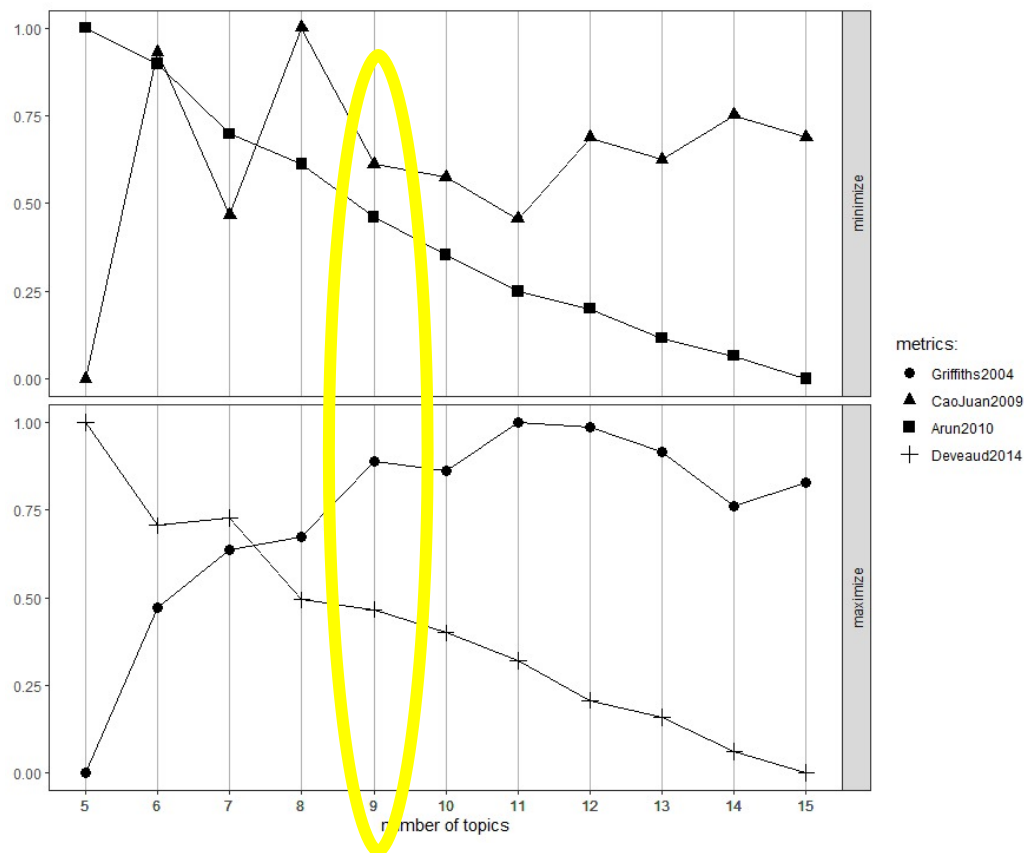


סבטימנט לפי נושאים

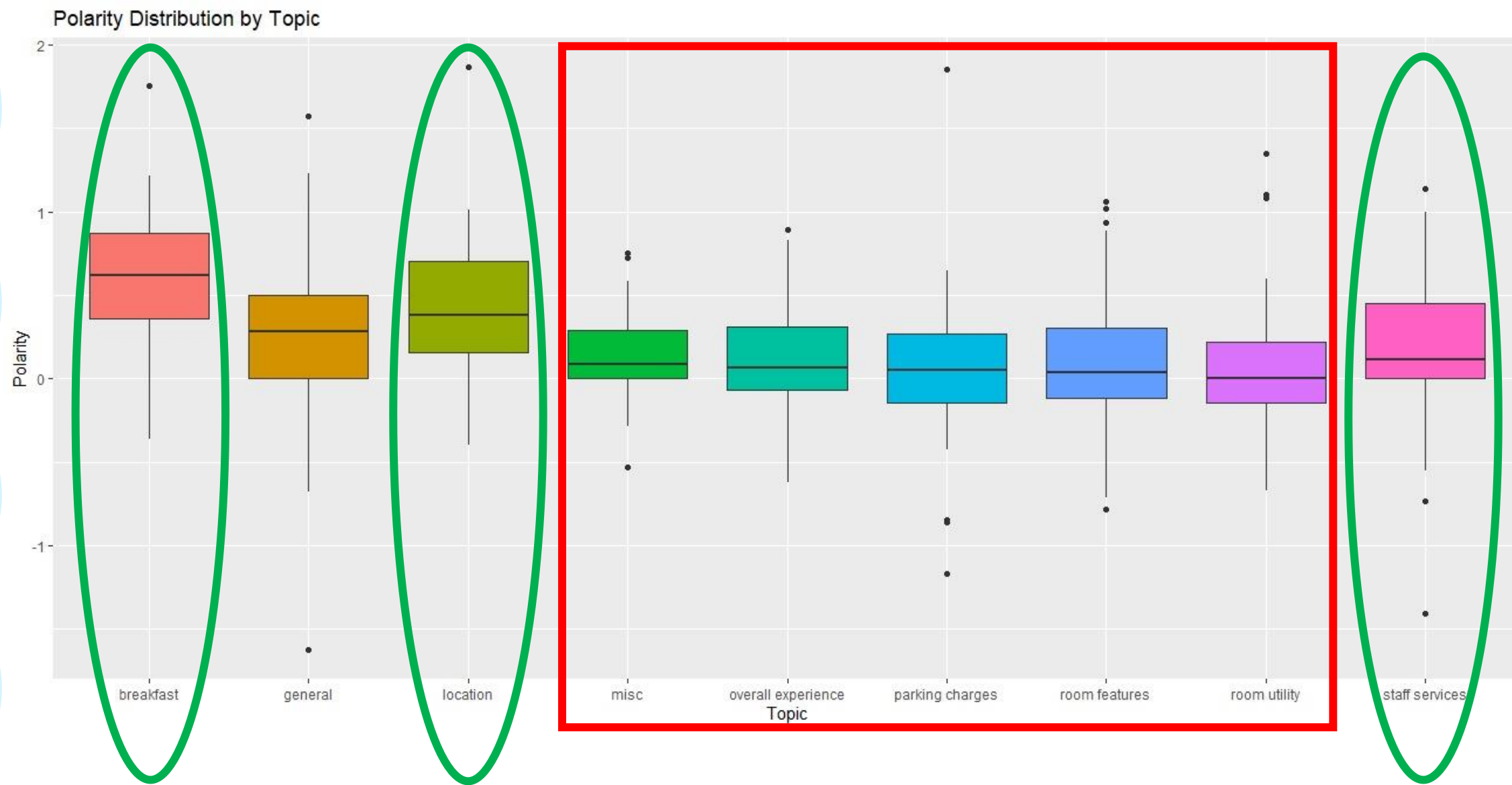
○ חלוקה של כל הביקורות למשפטים

○ אלגוריתם LDA-GIBBS

○ חלוקה ל 9 נושאים



תוצאות





סיכום ומסקנות

○ ניתן לייצר ערך למלון בעזרת הביקורות ברשת

○ נקודות חוזק – מיקוד בשיווק ובמיצוב

○ נקודות תורפה – לטפל

תודה על ההקשבה

