Comparative Analysis

Digital Company for Holidaymakers

Summary

This report presents a descriptive analysis of a digital company catering to holidaymakers for 24 months(Link), focusing on the market performance of hotels booked in Turkey. The report provides a comparative analysis of the company's performance against its core competitors in Spain, Egypt, and Greece. Through a detailed examination of various metrics, such as booking volumes, grouping by hotel stars, user pictures, session, and customer recommendation scores, the report provides insights into the digital company's market position. Additionally, the report utilizes a linear regression model to analyze the relationship between different variables and identify areas of opportunity for growth for hotels. Overall, the report offers valuable information for businesses operating in the digital holiday industry and highlights the importance of leveraging data analytics, including linear regression, to drive business growth.

Data Set

The Dataset in this project contain information about market performance of booking hotels in the 4 countries (Turkey, Spain, Egypt and Greece) in 4 Excel worksheets

Tools

R
KNIME
Tableau
Microsoft Excel
Microsoft SQL Server

Analytical Skills

Descriptive analysis
Writing SQL queries (Link)
Linear Regression
Data Visualization

Prepared by
Hamed Movahedizadeh

Overview

Data Structures

Data Structures

Divisions	Metrics	Values
size	observations	24,795
size	variables	18
size	values	446,310
size	memory size (MB)	7
duplicated	duplicate observation	0
missing	complete observation	15,497
missing	missing observation	9,298
missing	missing variables	4
missing	missing values	15,342

Data Types

Divisions	Metrics	Values
data type	numerics	10
data type	integers	0
data type	factors/ordered	0
data type	characters	7
data type	Dates	0
data type	POSIXcts	0
data type	others	1

Job Informations

Divisions	Metrics	Values
dataset	dataset	Holiday
dataset	dataset type	tbl_df
job	samples	24,795 / 24,795 (100%)
job	created	2023-04-15 19:43:19
job	created by	dlookr

Missing Values

List of Missing Values

Variables that include missing values

Variables	Missing	Missing (%)	Status	Recommand
stars	8,485	34.2%	● Bad	Model based Imputation
address_line	4,659	18.8%	● NotBad	Model based Imputation
latitude	1,099	4.4%	● Good	Delete or Imputation
longitude	1,099	4.4%	● Good	Delete or Imputation

Although it seems data need to be cleaned in some columns, there is no need because they are aligned with the purpose of this analysis report.

Warnings

	Checks	Judgements	Removes
	9	13	1

Warnings	Types	Recommands
stars has 8,485 (34.2%) missing values	missing	judgement
address_line has 4,659 (18.8%) missing values	missing	judgement
latitude has 1,099 (4.4%) missing values	missing	judgement
longitude has 1,099 (4.4%) missing values	missing	judgement
name has high(1.00) cardinality, Maybe identifier	cardinality	check
hc_country has constant value "Türkei"	cardinality	remove
number_of_bookings_last24m has 23,075 (93.06%) zeros	zero	check
ttv_last24m has 23,075 (93.06%) zeros	zero	check
hotelier_pictures has 22,476 (90.65%) zeros	zero	check
countactiveaccounts has 22,024 (88.82%) zeros	zero	check
user_pictures has 20,889 (84.25%) zeros	zero	check
totalavgrecommendation has 19,255 (77.66%) zeros	zero	check
overall_reviews has 18,597 (75%) zeros	zero	check
sessions_last_24m has 4,325 (17.44%) zeros	zero	check
overall_reviews has 6,198 (25%) outliers	outlier	judgement
totalavgrecommendation has 5,540 (22.34%) outliers	outlier	judgement
sessions_last_24m has 4,296 (17.33%) outliers	outlier	judgement
user_pictures has 3,906 (15.75%) outliers	outlier	judgement
countactiveaccounts has 2,771 (11.18%) outliers	outlier	judgement
hotelier_pictures has 2,319 (9.35%) outliers	outlier	judgement
longitude has 2,126 (8.57%) outliers	outlier	judgement
number_of_bookings_last24m has 1,720 (6.94%) outliers	outlier	judgement
ttv_last24m has 1,720 (6.94%) outliers	outlier	judgement

General Overview of market performance in Turkey
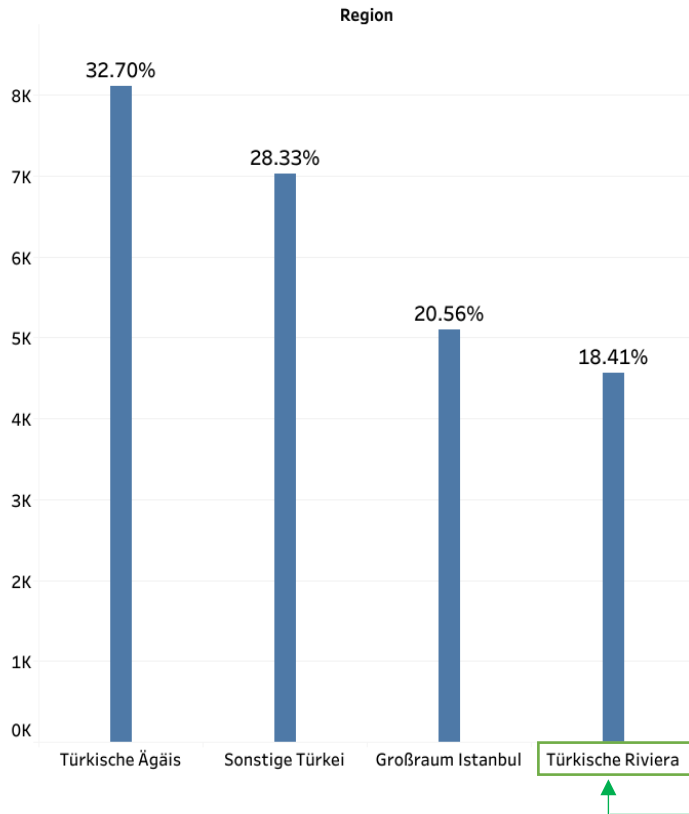
Figure 1
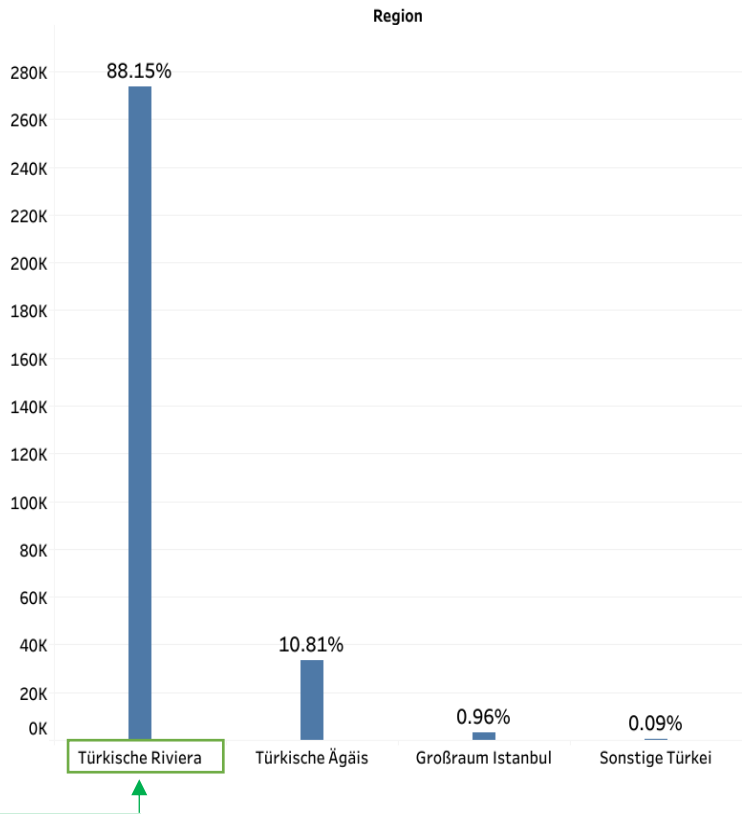Number of hotels in region

Figure 2
Number of bookings

Table 1. Ratios based on regions in Turkey

Region	Percentage of total Hotels	Percentage of total bookings
Türkische Ägäis	32.7 %	10.81 %
Sonstige Türkei	28.83 %	0.09 %
Großraum Istanbul	20.56 %	0.096 %
Türkische Riviera	18.41%	88.15

Table 1 states that although only 18.41% of total hotels are in Türkische Riviera region, 88.15 % of total bookings related to this region. But why?

1

General Overview of market performance in Turkey
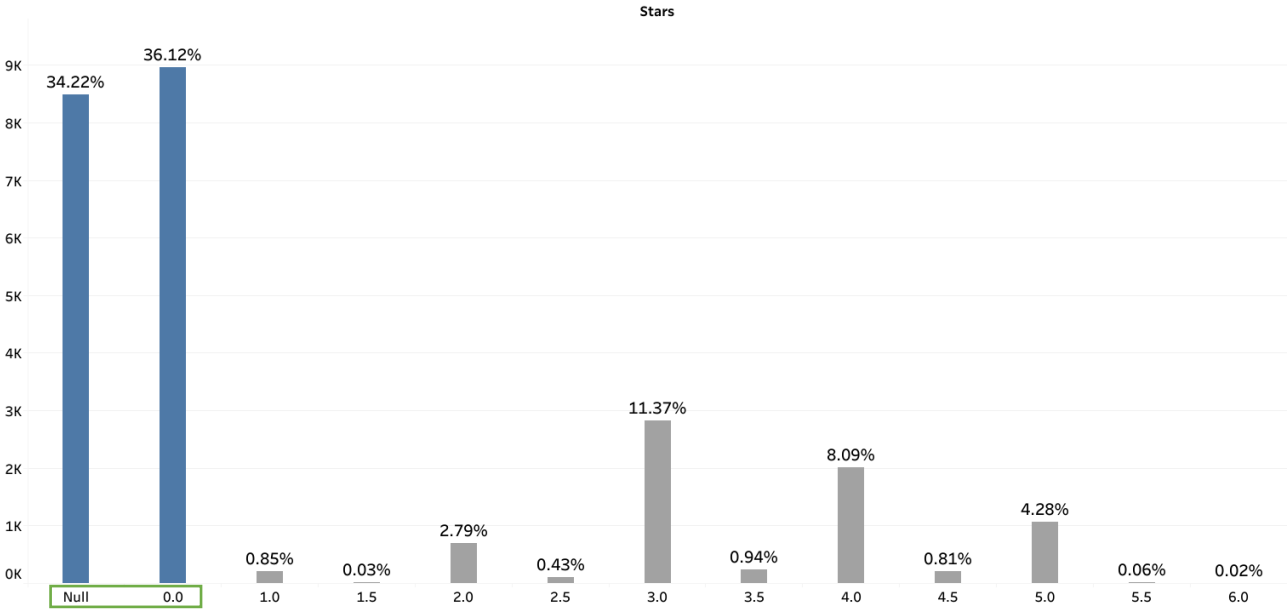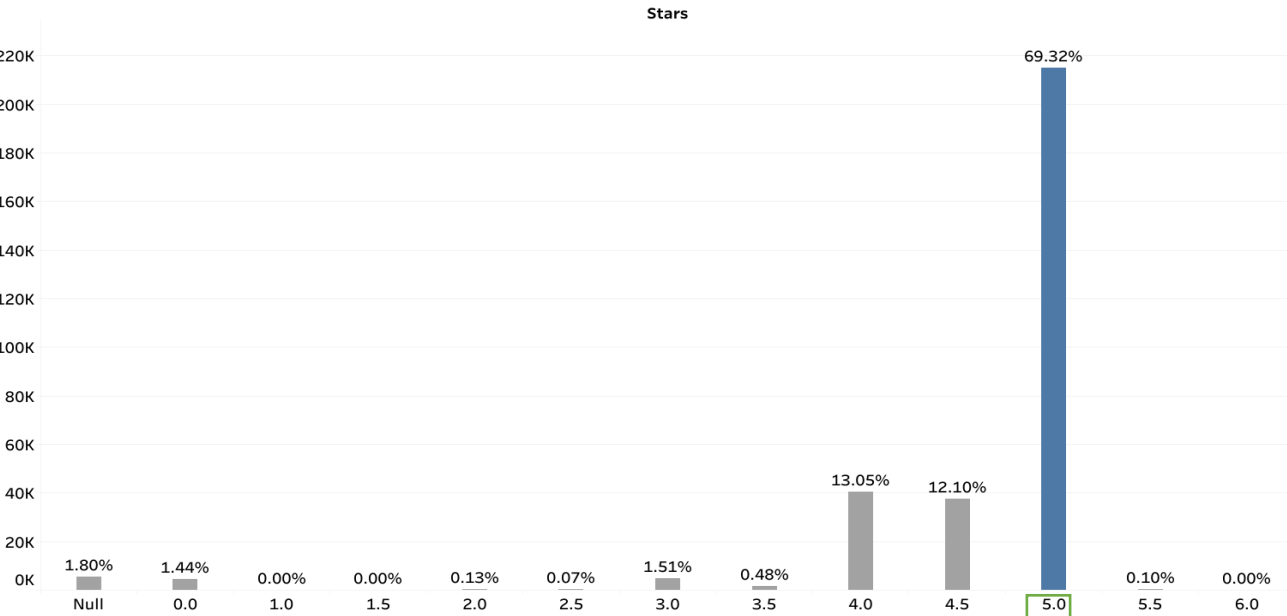
Figure 3

Number of hotels by star

Figure 3 shows that around 69% of total hotels in Turkey have Null or Zero stars, while Figure 4 illustrates that around 69% of total bookings in Turkey relate to 5-star hotels.
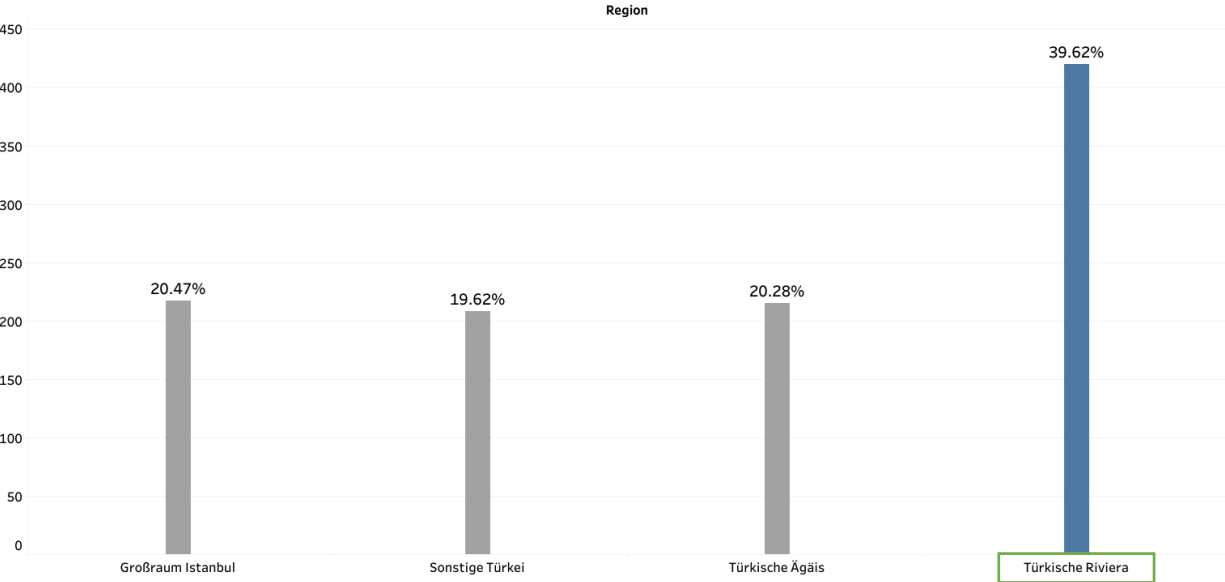
Figure 4

Number of bookings by stars

General Overview of market performance in Turkey

Figure 5

Number of 5 star hotels in regions

Based on Figure 5, around 40% of total hotels in Turkey are in Türkische Riviera, while Figures 6 & 7 show that this number for Null and Zero-star hotels are lower than those in the other regions.
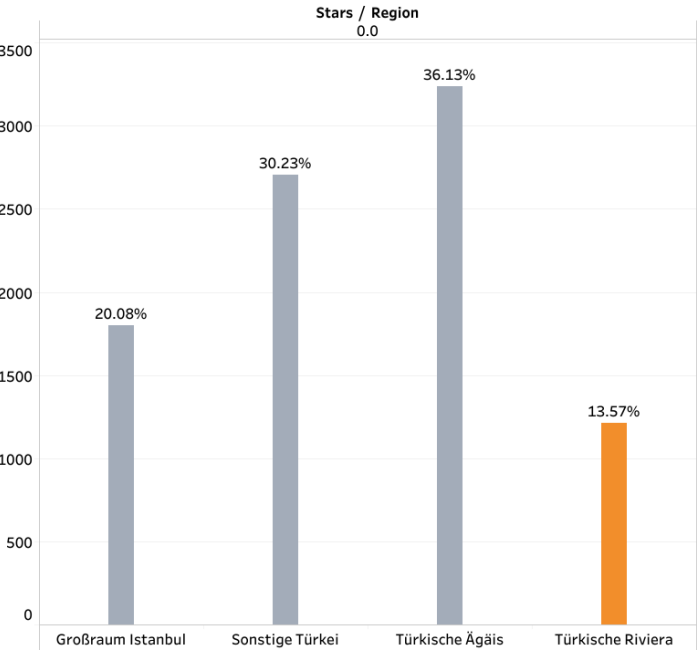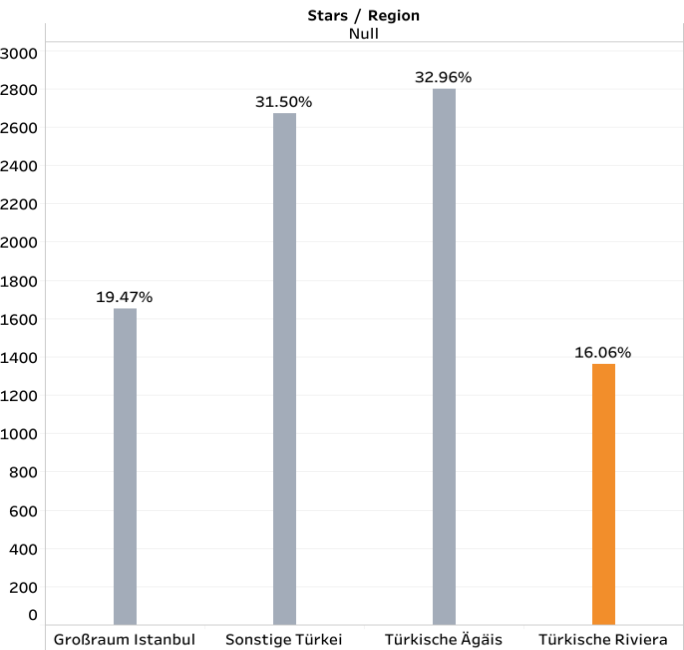
Figure 6

Number of zero-star hotels in regions

Figure 7

Number of null-star hotels in regions

These trends (on page 2) are almost similar in other countries as well.

Figure 8

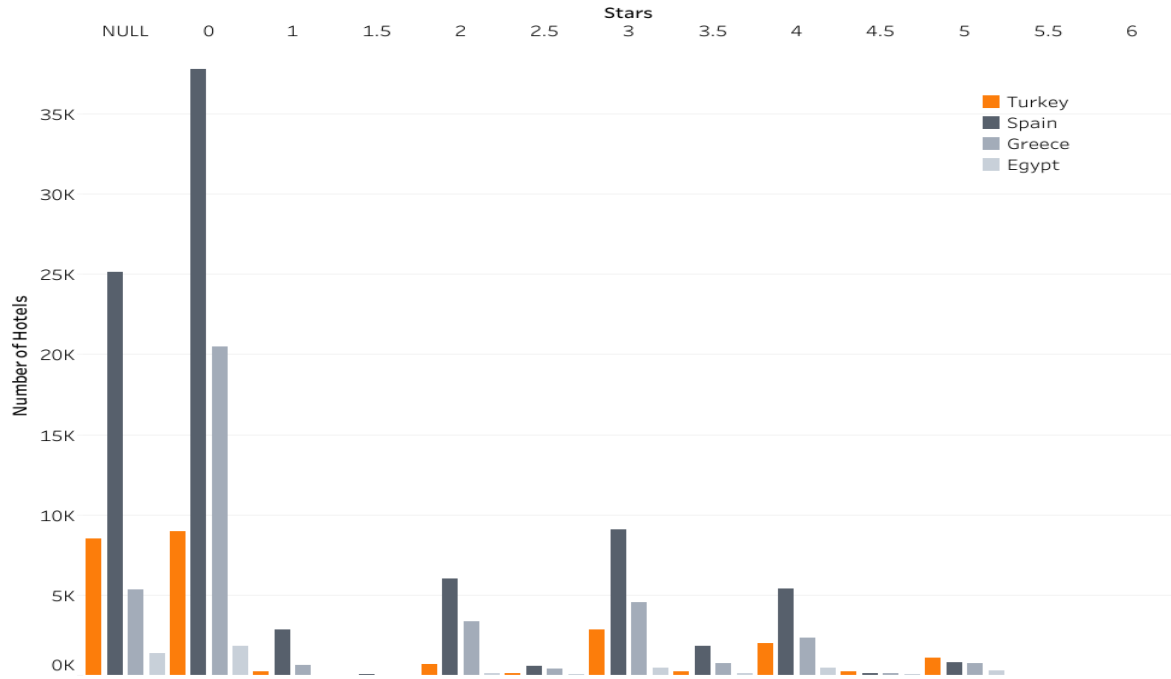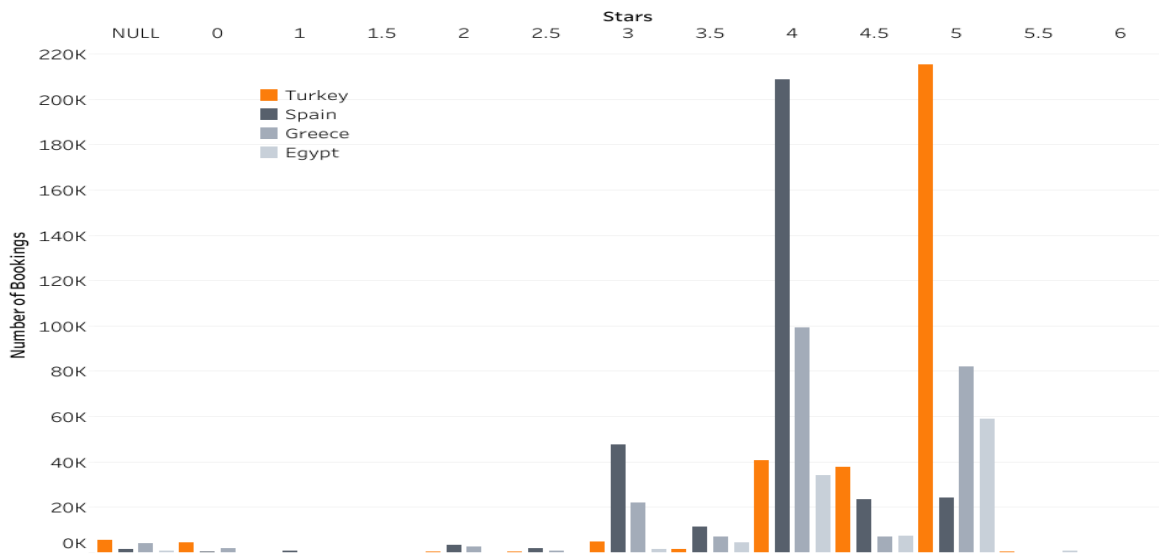Number of Hotels by star in the 4 counires

Figure 9

Number of Booking by Star in the 4 countries

Some Descriptive statistics and charts based on stars

Table 2

Percentage of hotels by star

Country / Star	Null & Zero	1	2	3	3.5	4	4.5	5
Turkey	**70%**	0.85%	3%	**11.00%**	0.94%	**8%**	1%	4%
Spain	**70%**	3%	7%	**10%**	2%	**6%**	0.15%	0.89%
Greece	**66%**	1.62%	9%	**12%**	2%	**6%**	0.32%	2%
Egypt	**66%**	0.58%	3%	**9.40%**	2%	**9%**	1.50%	6%

Table 3

Percentage of Review by star

Country / Star	Null & Zero	1	2	3	3.5	4	4.5	5
Turkey	2%	0.03%	0.16%	4%	1%	**21%**	18.35%	**53%**
Spain	2%	0.6%	2%	**19.6%**	5%	**58%**	6%	6%
Greece	6.4%	0.15%	2%	14%	5%	**43%**	3%	**26%**
Egypt	2%	0%	0.05%	2%	4%	**43%**	9%	**39%**

Table 4

Percentage of bookings by star

Country / Star	Null & Zero	1	2	3	3.5	4	4.5	5
Turkey	3.20%	0	0.13%	1.50%	0.50%	**13%**	12%	**69.30%**
Spain	0.50%	0.25%	1%	**15%**	3.50%	**65%**	7%	7%
Greece	2.60%	0.03%	1%	10%	3%	**44%**	3%	**36%**
Egypt	1%	0	0.01%	1%	4%	**31%**	7%	**55%**

Table 5

Percentage of transaction by star

Country / Star	Null & Zero	1	2	3	3.5	4	4.5	5
Turkey	3%	0%	0.05%	1%	0.3%	11%	12%	**73%**
Spain	0.4%	0.1%	0.65%	**11%**	3%	**66%**	7.6%	10.5%
Greece	0.61%	0.01%	0.6%	6%	2%	**41%**	3%	**44%**
Egypt	1%	0%	0%	1%	3%	**28%**	7%	**59%**

An overall overview of Turkish market performance compared to its core competitors

Figure 10
Total numbers
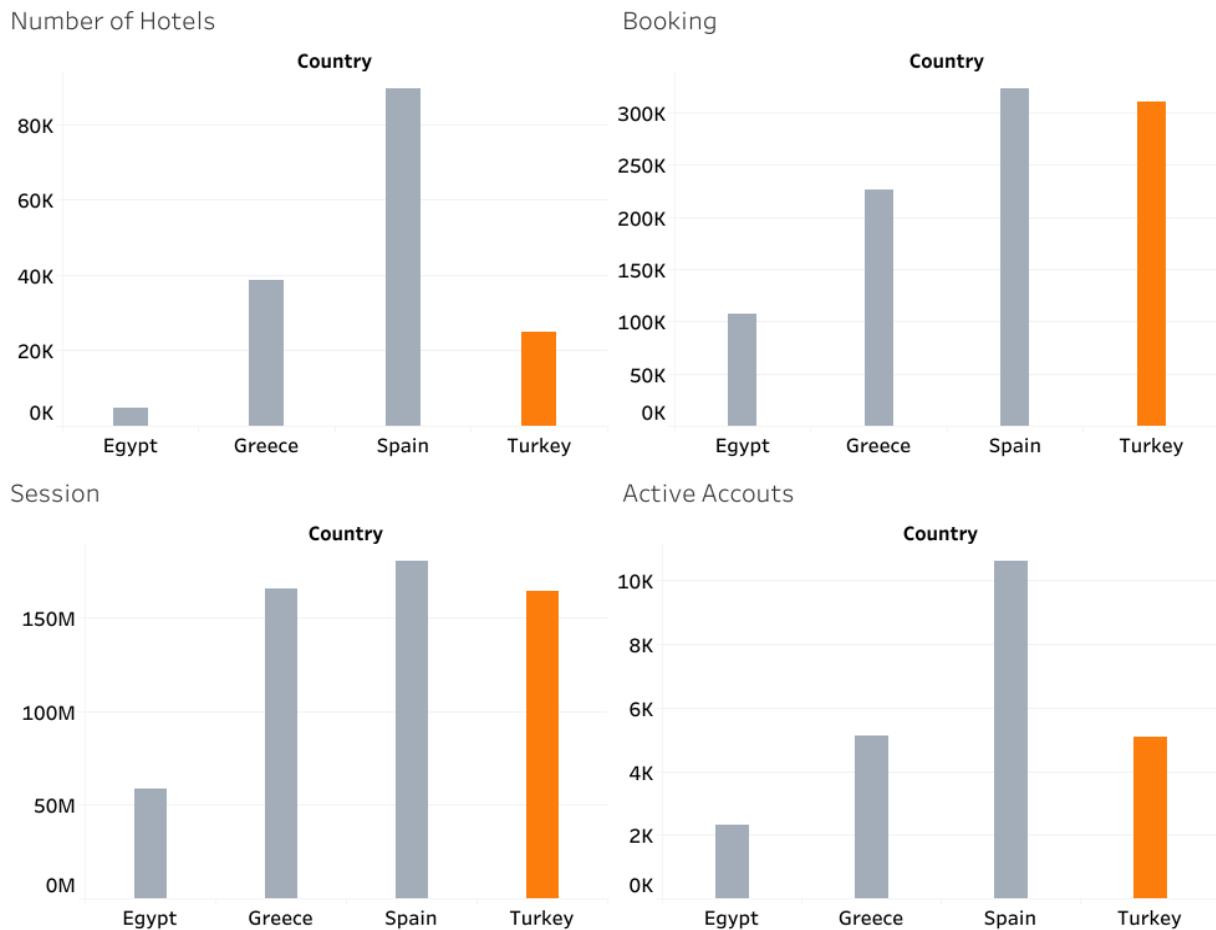
Total number in these bar charts state that:

- Although number of hotels in Turkey is significantly lower than that of in Spain, it seems that Turkish market performance in terms of total bookings is in a good position compared to Spain. However, number of active users in Turkey is considerably lower than that of in Spain.
- In addition, total amount of sessions in Spain, Turkey and Greece seems to be in a close range.

To compare market performance in the four countries, we can evaluate ratios based on the number of hotels.
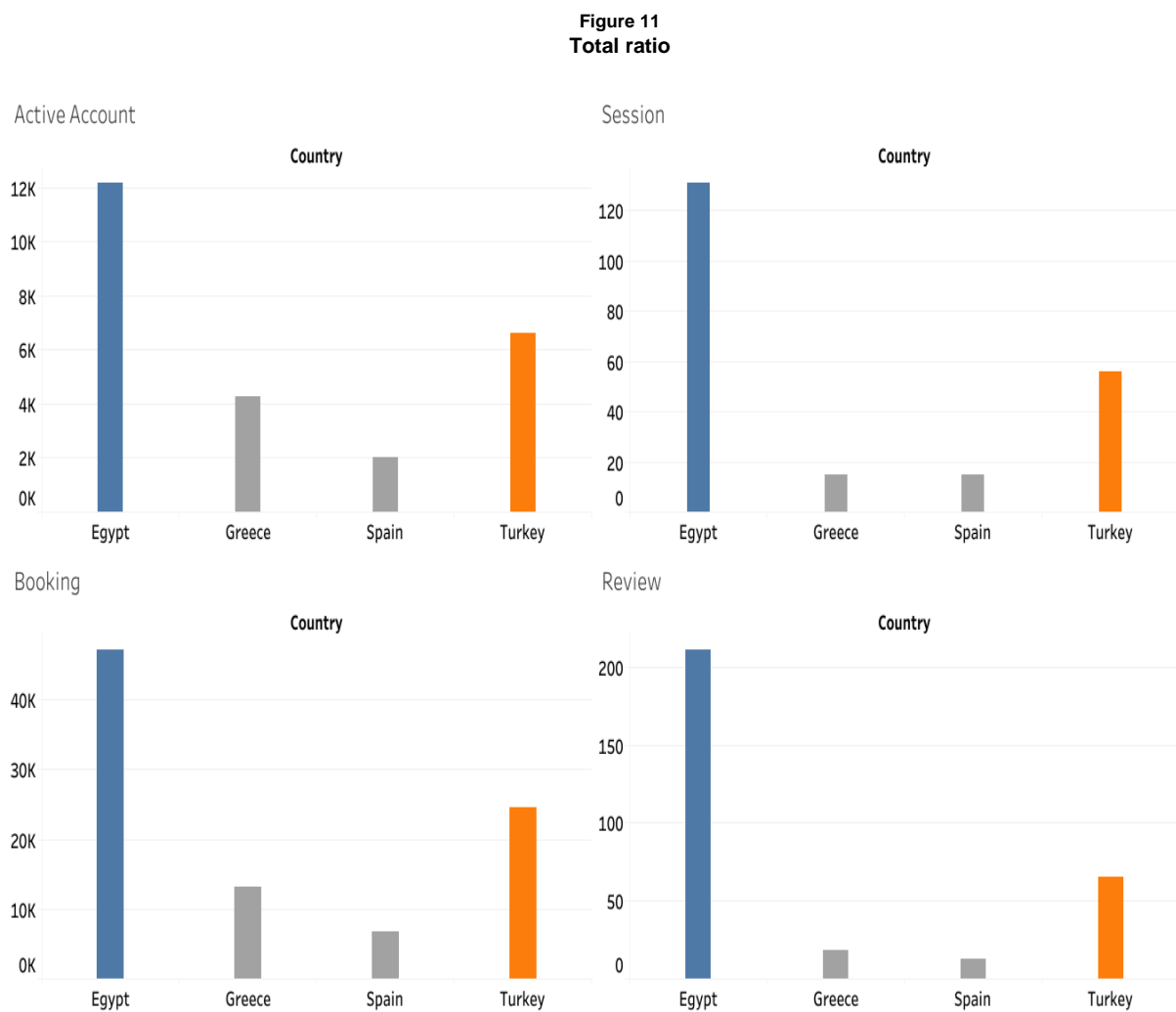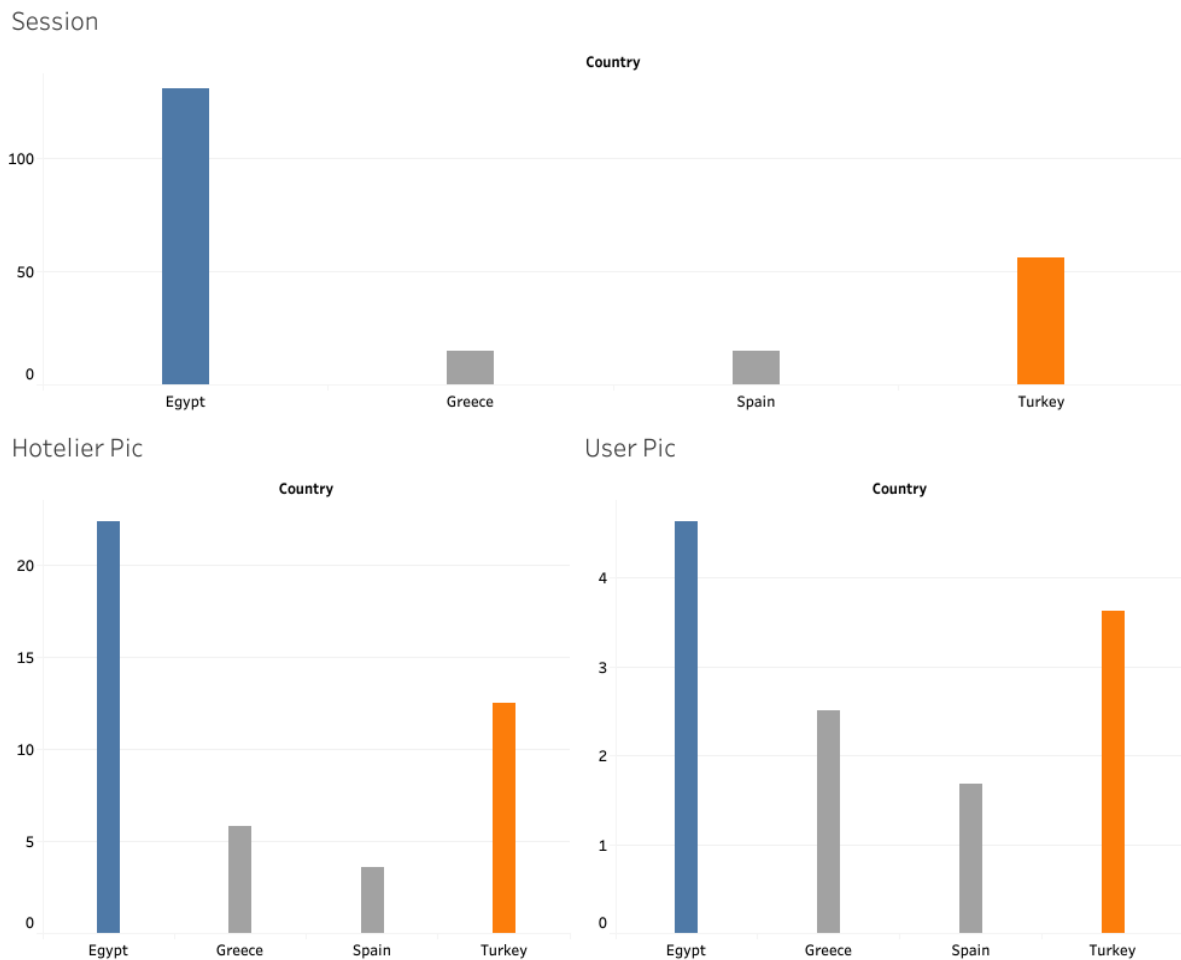
Figure 11
Total ratio

Figure 11 illustrates that Egypt has a better performance in terms of above ratios. However, Turkey still has a better performance compared to Spain and Greece.

Is this because of a higher number of sessions?
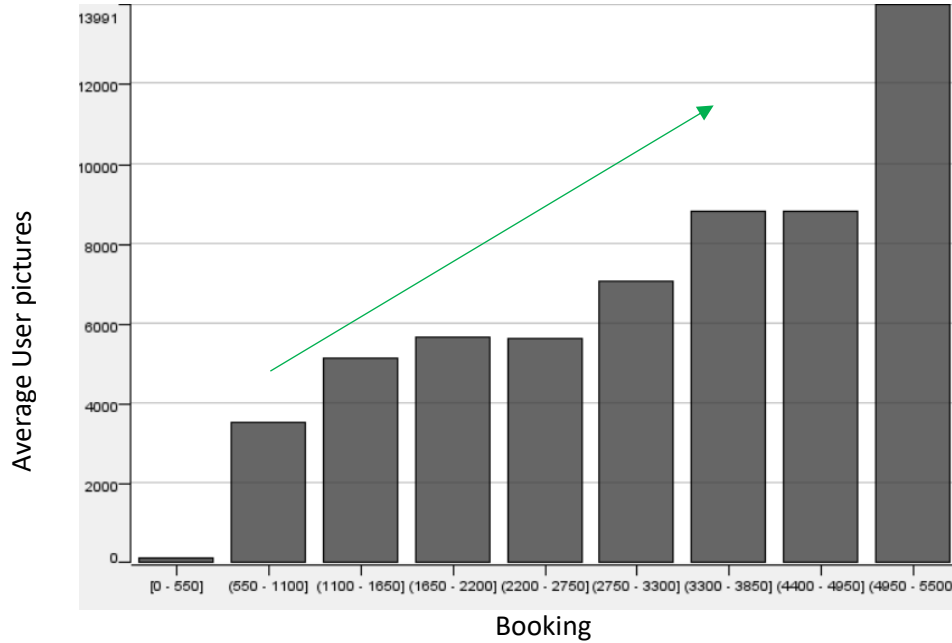
We can go further into detail to have a better view

Figure 12
Total ratio

Bar charts in figure 12 show that average number of hotelier and user pictures in Egypt are higher than those in other countries.
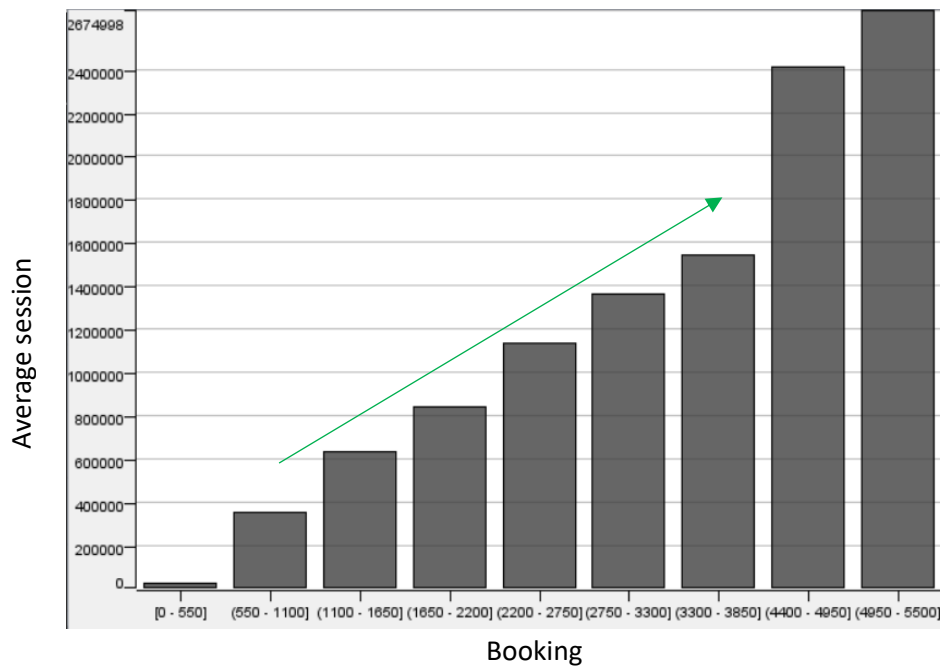
User pictures, sessions, and Booking in Turkey

Figure 13   Interactive histogram Chart

- Figure 13 shows that there is positive pattern between average user pictures and bookings.
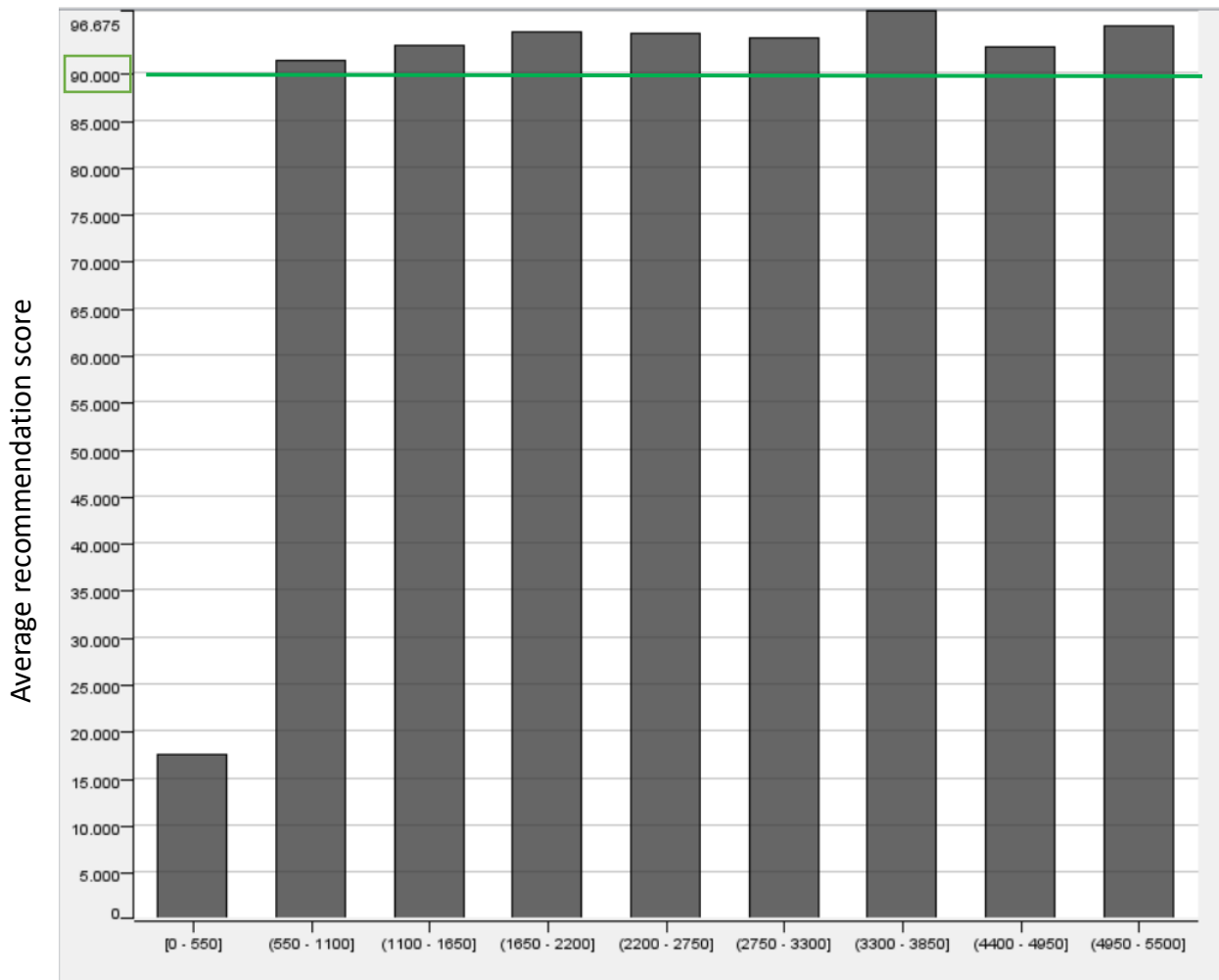
Figure 14   Interactive histogram Chart

- Figure 14 shows that there is positive pattern between average session and bookings.

Recommendation and Booking in Turkey
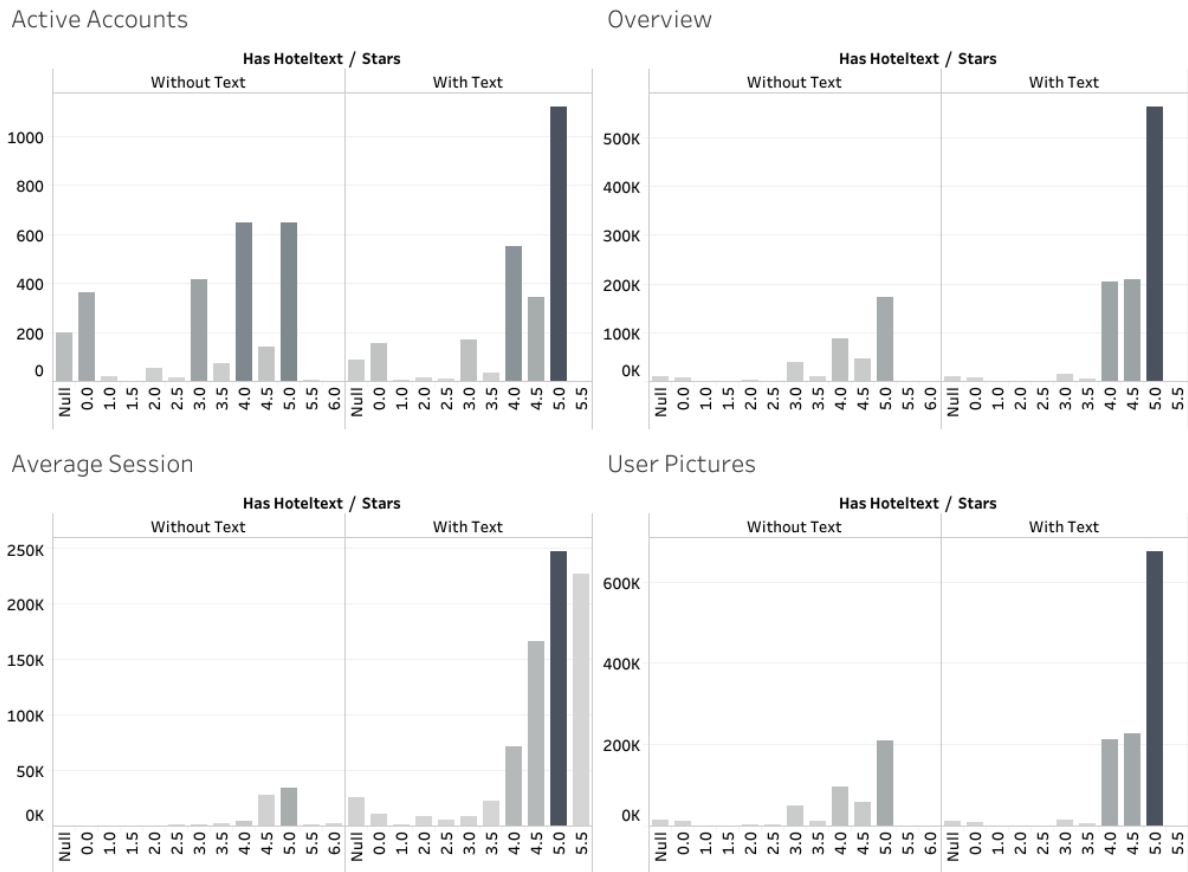
Figure 15

Interactive Histogram Chart

As it is illustrated, most of the bookings related to those hotels with average recommendation score above 90

Hotels with & without having text in Turkey

Another categorical factors we should evaluate is whether a hotel has a text on its profile or not.

Figure 16
Overall view categorized by text and stars

As it is seen above, overall ratios in hotels with text are significantly better that those in hotels without having text

A comparison of hotels with and without text in Turkey

Table 6
Percentage ratios for hotels with & without text in Turkey

Text	Active Accounts	Hotels	Transactions	Bookings	Hotelier Pictures	User Pictures	Reviews	Sessions
Without	51%	96%	22%	24%	39%	28%	27%	26%
With	49%	4%	78%	76%	61%	72%	73%	74%

Numbers on the table 6 show that only 4% hotels have text, however, 78% of total transaction and 76% of bookings belong to them.

- In addition, other ratios on hotels with text are considerably higher.

Number of Bookings in hotels with and without having text in Turkey

Figure 17
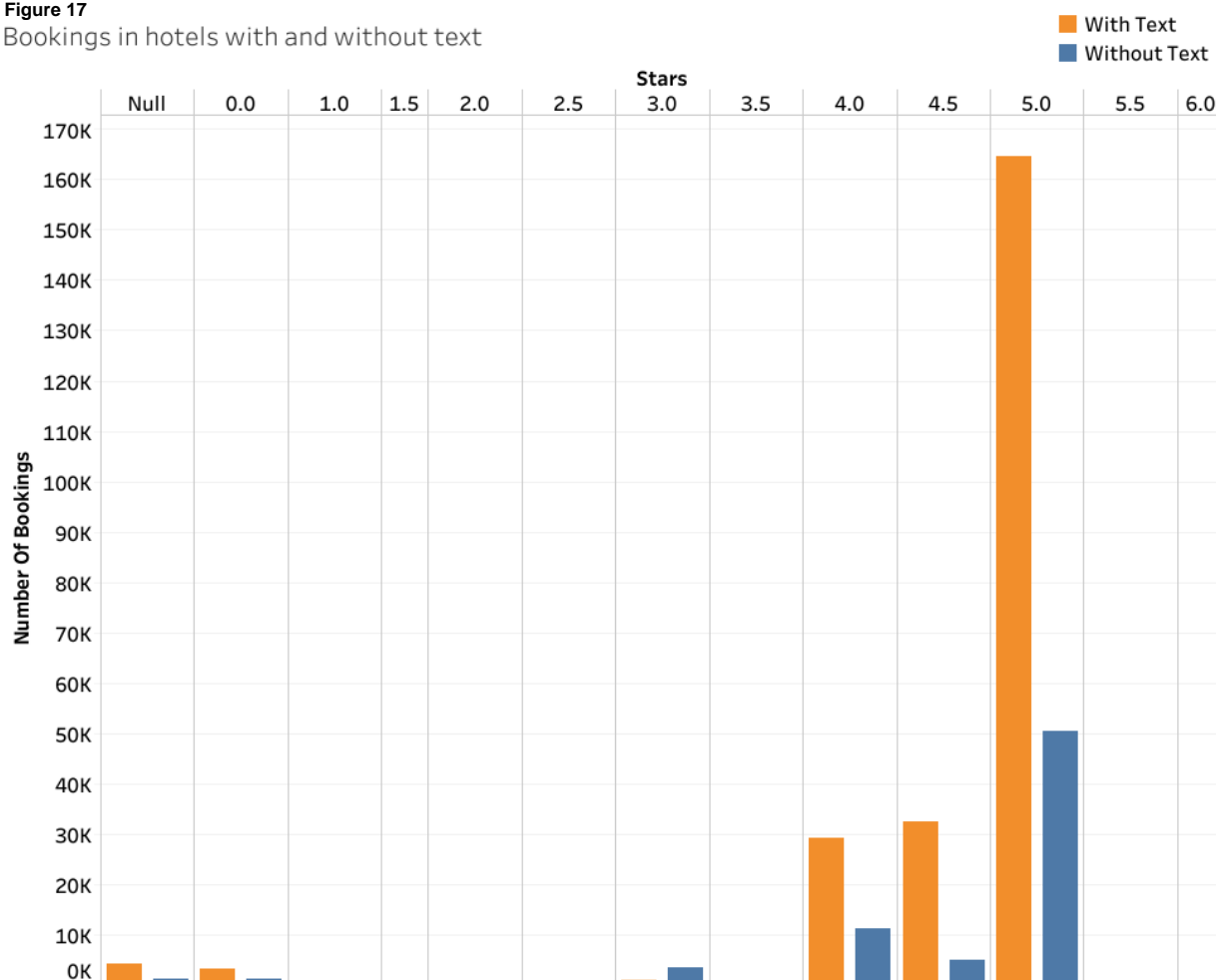Bookings in hotels with and without text

Table 7
Hotels with having text compared to without having text

Star	Booking ratio
Null and 0	6 times more
4	3 times more
4.5	6 times more
5	3 times more
5.5	9 times more

Average recommendation for hotels with and without having text in Turkey

Figure 18

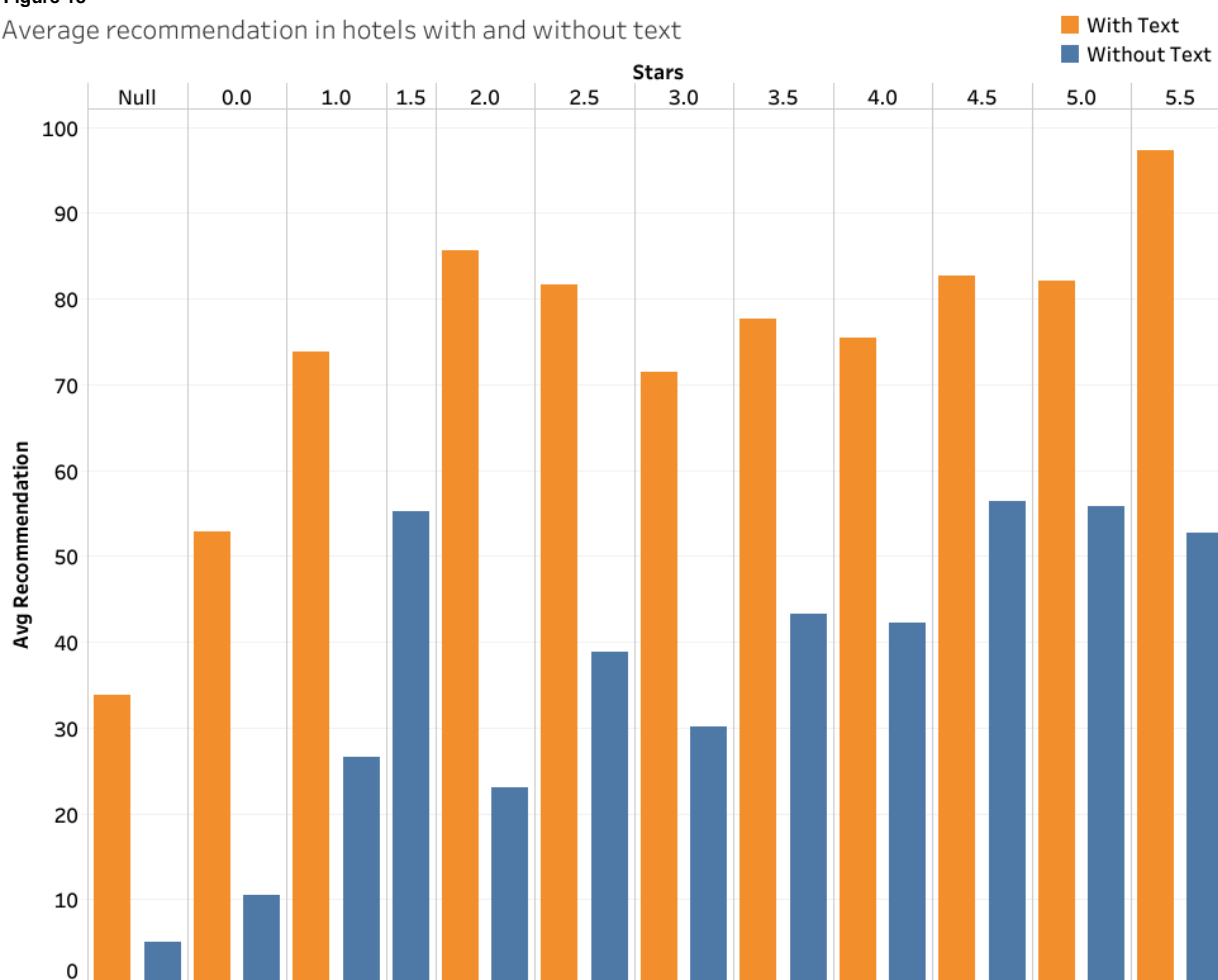Average recommendation in hotels with and without text

Table 8
Hotels with having text compared to without having text

Star	Recommendation ratio
Null and 0	6 times more
1 to 3	2.7 times more
3.5 to 5.5	1.7 times more

Regression Analysis Outcomes

Table 9
Regression result - Booking

Target Variable	Predictor	R square	P-Value
Booking	User Picture	0.6587	2.20E-16
Booking	Hotelier Picture	0.1233	2.20E-16
Booking	Average Recommendation	0.04085	2.20E-16
Booking	Average Rating	0.1177	2.20E-16
Booking	Session	0.9454	2.20E-16
Booking	Active Accounts	0.1586	2.20E-16

- Statistically, there is a significant relationship between number of bookings as our dependant variable and sessions and user pictures as our independent variables.

Table 10
Regression result - Session

Target Variable	Predictor	R square	P-Value
Session	User Picture	0.715	2.20E-16
Session	Hotelier Picture	0.1397	2.20E-16
Session	Average Recommendation	0.1292	2.20E-16
Session	Average Rating	0.04693	2.20E-16
Session	Active Accounts	0.186	2.20E-16

- However, it seems there is, also, a considerable relationship between sessions and number of user pictures

- A part of it may be caused by the times users have spent uploading photos of the hotels on the hotel's page. However, it is a clue for more investigation.

Summary

- Türkische Riviera region has the highest performance.
- About 40% of total 5-star hotels in Turkey are in Türkische Riviera region.
- Around 70% of total booked hotels in Turkey are 5-star, followed by 4-star (13%) and 4.5-star (12%)
- 70 % total hotels in Turkey are null or zero-star. Other competitors have similar numbers.
- Market performance ratios in Turkey are better than those in Spain and Greece. However, Egypt has better ratios.
- It seems there is a statistically significant relationship between number of uploaded pictures by users and sessions supported by interactive histogram chart and outcome of regression analysis
- Data show that although only 4% hotels have text, 78% of total transaction and 76% of bookings belong to them. Therefore, probably it has positive impact on customers' purchasing decisions.

Data Limitation

Clearly, by having access to more information we can evaluate the data in more detail and consequently, we can continue further to do data mining and predictive modelling.