# EDA CAPSTONE PROJECT HOTEL BOOKING ANALYSIS

SHIYAS ALI T M

# CONTENTS

- **OVERVIEW**
- **DATA SUMMARY**
- **CORRELATION HEAT MAP**
- **GENERAL FINDINGS**
- **BOOKING VS TIME OF THE YEAR**
- **MARKET SEGMENTS**
- **CANCELLATION**
- **AVERAGE DAILY RATE**
- **SPECIAL REQUESTS**
- **CONCLUSION**

# OVERVIEW

- The data gives information about bookings made to certain hotels in the year 2015 to 2017
- Contains 119390 rows and 32 columns
- Data contain general informations pertaining to the hotel bookings,personal records are not disclosed.
- Data is explored to find the key features governing the bookings.
- Focused on features like type of hotel prefered,number of bookings at different time of the year,market segments through which the bookings are made,bookings made from different countries,average daily rate,number of special requests etc.
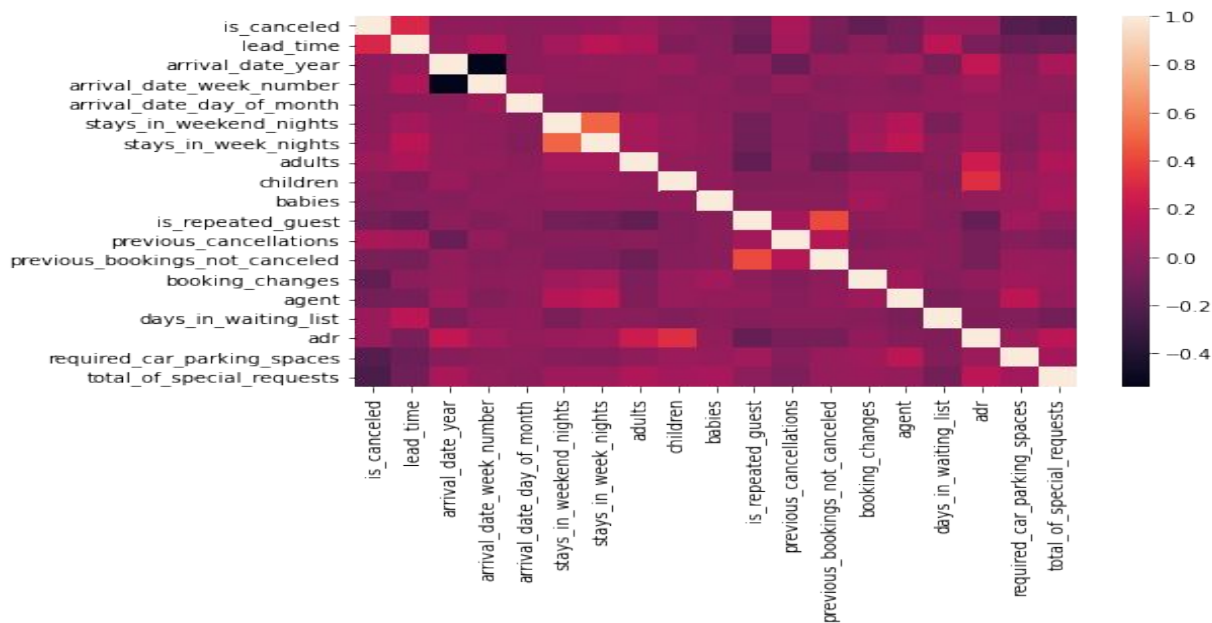
# DATA SUMMARY

## COLUMNS IN THE DATA SET

- hotel -    Categorical
- is_canceled -    Categorical
- arrival_date_year -    Categorical
- arrival_date_month -    Categorical
- arrival_date_day_of_month -    Categorical
- stays_in_weekend_nights -    Continuous
- stays_in_week_nights -    Continuous
- adults -    Continuous
- children -    Continuous

- babies -    Continuous
- country -    Categorical
- market_segment -    Categorical
- reserved_room_type -    Categorical
- assigned _room_type -    Categorical
- agent -    Categorical
- company
- adr -    Continuous
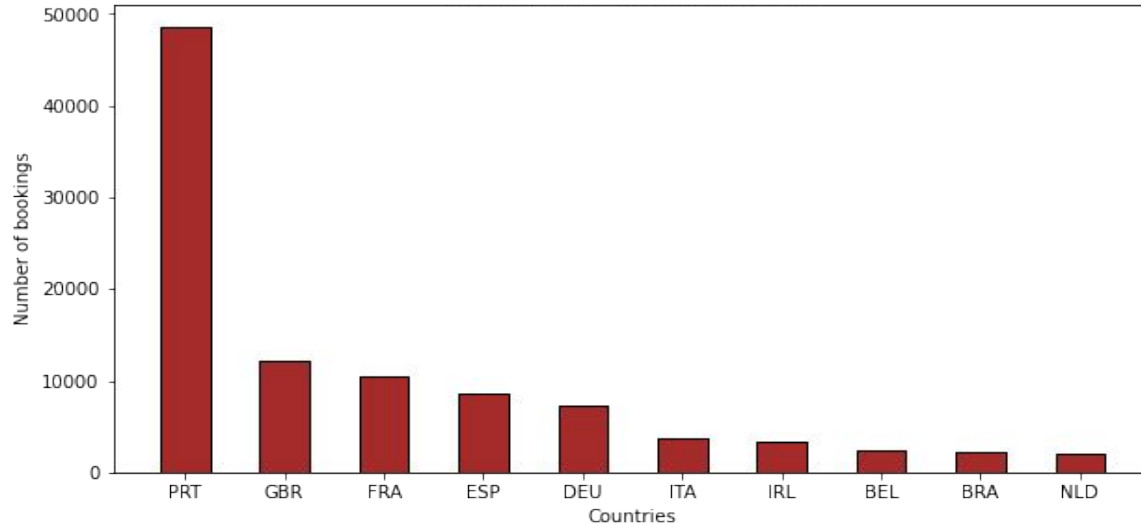- total_of_special_requests -    Continuous

# CORRELATION HEAT MAP

# GENERAL FINDINGS



Number of booking according to hotel type
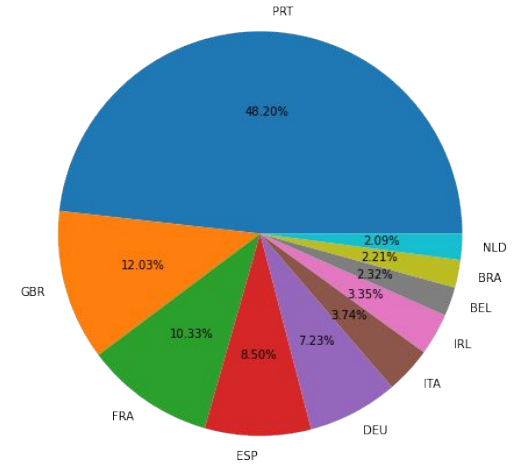


percentage of hotel types
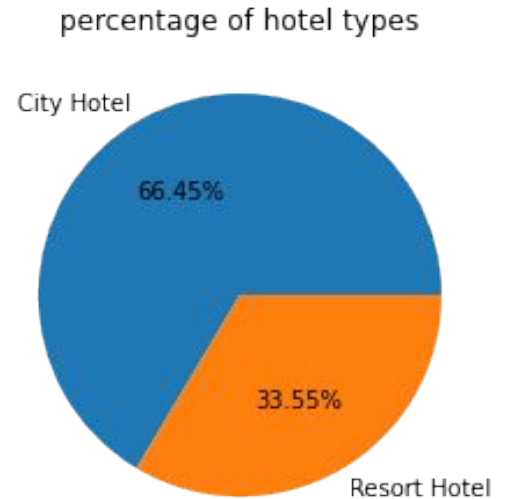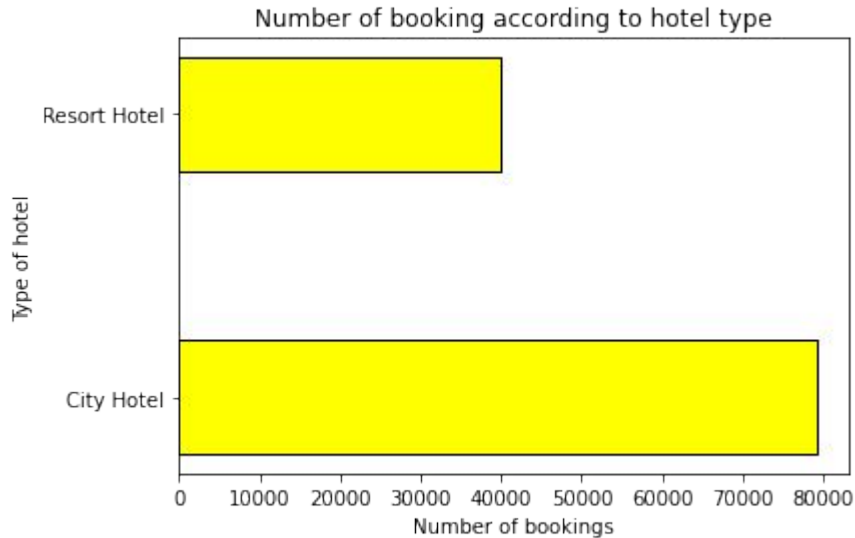
❖ Bookings to city hotel exceeds resort hotel by far(almost double)

# GENERAL FINDINGS  CONTINUED...

**Bookings from different countries**



**Percentage of bookings from different countries**



❖ The hotels had disproportionately high number of bookings from Portugal followed by Great Britain,France,Spain and Germany.

# GENERAL FINDINGS CONTINUED...

- ❖ Almost 75% of the total bookings have been made to room type 'A'
- ❖ Room type 'D' comes at the second position with less than one-fifth of the total bookings.
- ❖ Bookings to other types of rooms are marginal when compared to these two.

| room type | bookings |
|:---:|---:|
| A | 85994 |
| D | 19201 |
| E | 6535 |
| F | 2897 |
| G | 2094 |
| B | 1118 |
| C | 932 |
| H | 601 |
| P | 12 |
| L | 6 |

# BOOKING VS TIME OF THE YEAR

| | arrival_date_year | arrival_date_month | 2015 | 2016 | 2017 | total |
|---|---|---|---|---|---|---|
| 0 | | January | 3404.5 | 2248.0 | 3681.0 | 9333.5 |
| 1 | | February | 3404.5 | 3891.0 | 4177.0 | 11472.5 |
| 2 | | March | 3404.5 | 4824.0 | 4970.0 | 13198.5 |
| 3 | | April | 3404.5 | 5428.0 | 5661.0 | 14493.5 |
| 4 | | May | 3404.5 | 5478.0 | 6313.0 | 15195.5 |
| 5 | | June | 3404.5 | 5292.0 | 5647.0 | 14343.5 |
| 6 | | July | 2776.0 | 4572.0 | 5313.0 | 12661.0 |
| 7 | | August | 3889.0 | 5063.0 | 4925.0 | 13877.0 |
| 8 | | September | 5114.0 | 5394.0 | 5141.5 | 15649.5 |
| 9 | | October | 4957.0 | 6203.0 | 5141.5 | 16301.5 |
| 10 | | November | 2340.0 | 4454.0 | 5141.5 | 11935.5 |
| 11 | | December | 2920.0 | 3860.0 | 5141.5 | 11921.5 |

# BOOKING VS TIME OF THE YEAR CONTINUED...
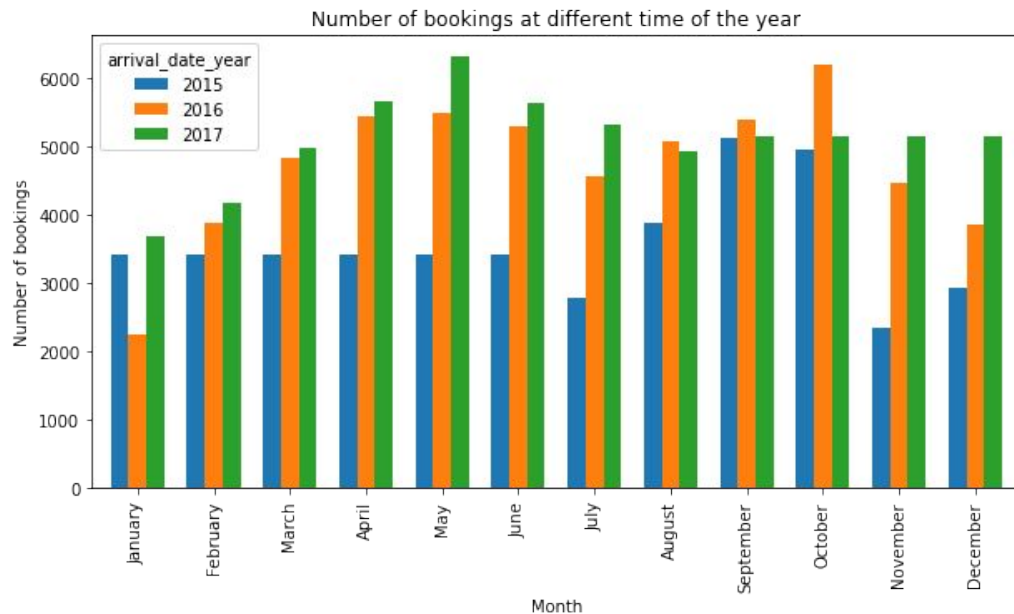


Number of bookings at different time of the year

# BOOKING VS TIME OF THE YEAR CONTINUED...



Number of bookings at different time of the year

❖ Number of bookings at the same time of the year has been on a rise with each passing year.

❖ Each year sees a high number of booking around may and october.

# BOOKING VS TIME OF THE YEAR CONTINUED...

Bookings in different months



- ❖ October is the busiest month.
- ❖ April,June,May,August and September receives high number of bookings.
- ❖ The period from November to February see a low in terms of bookings.

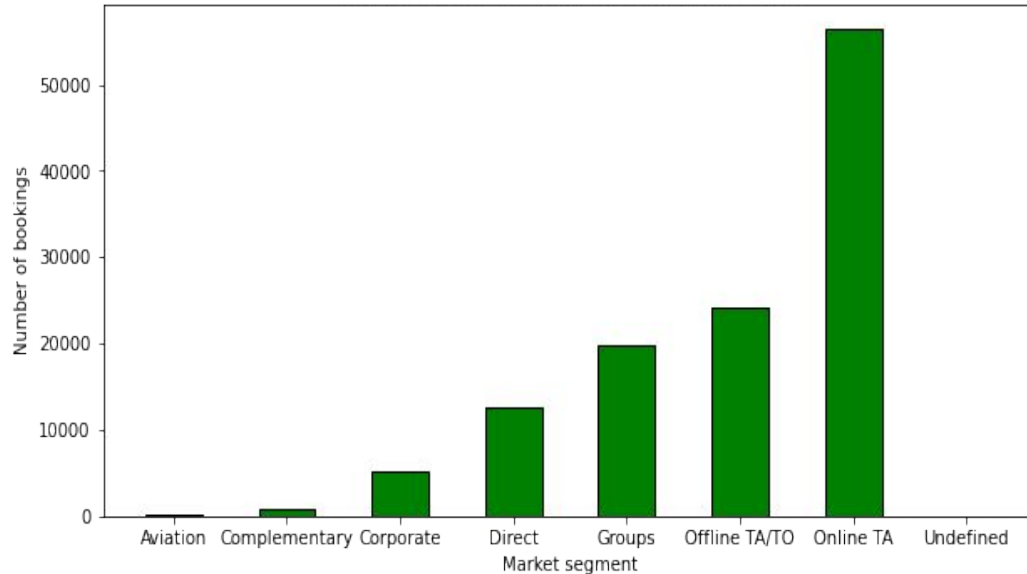# BOOKING VS TIME OF THE YEAR CONTINUED...



reserved room types in different months
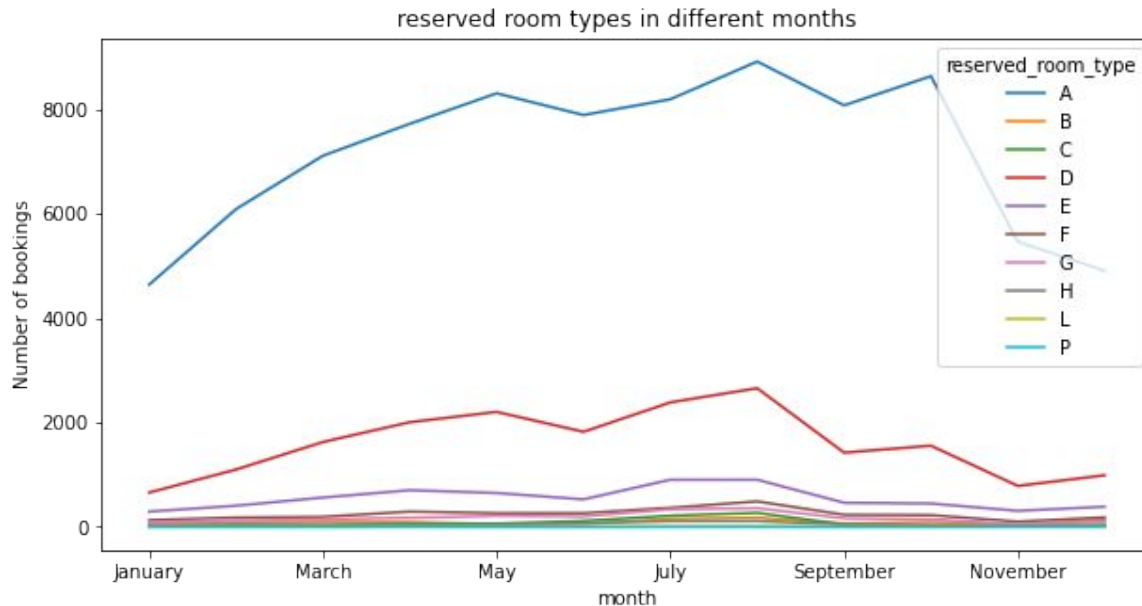
❖ No particular correlation between prefered room type and booking period.

# MARKET SEGMENTS

Bookings made by different market segments



❖ Most of the bookings were made through online travel agencies

❖ Offline travel agencies,groups and direct bookings come in second,third and fourth position respectively in terms of number of bookings made.

# MARKET SEGMENTS CONTINUED...



Bookings made by different market segments according to countries
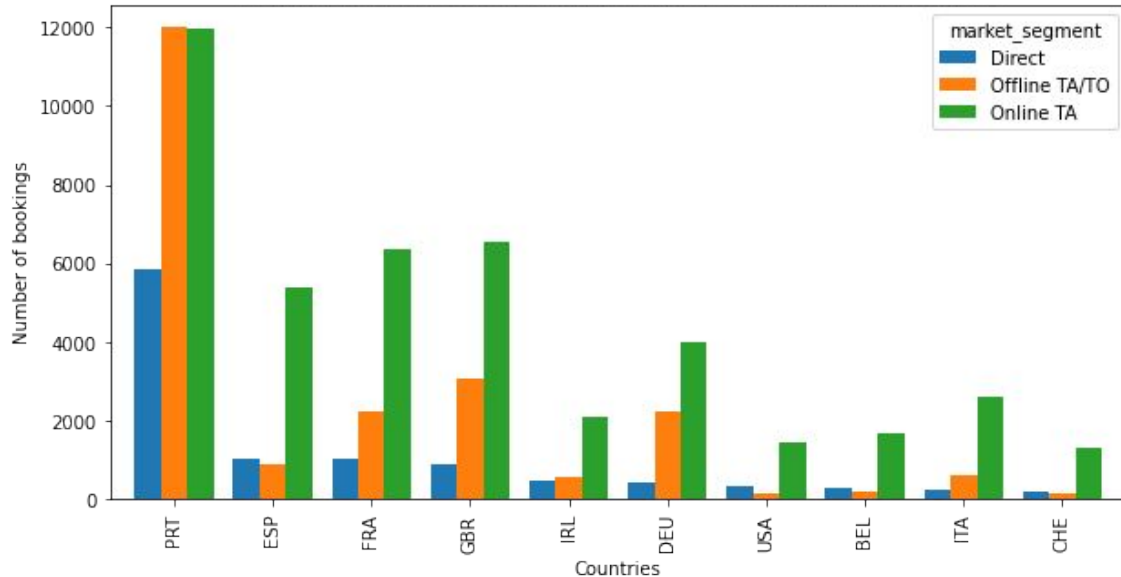
❖ Every countries ,except Portugal ,rely more on online travel agencies for booking.

❖ In case of Portugal ,bookings made through online and offline travel agencies are almost equal in number.

❖ Direct booking is close to,or even exceeds offline travel agencies in many cases.

# CANCELLATION

| cancellation status | number |
|---|---|
| 0 | 75166 |
| 1 | 44224 |

| Hotel | Bookings | cancellations | cancellation_percentage |
|---|---|---|---|
| City Hotel | 79330 | 33102 | 42.0 |
| Resort Hotel | 40060 | 11122 | 28.0 |
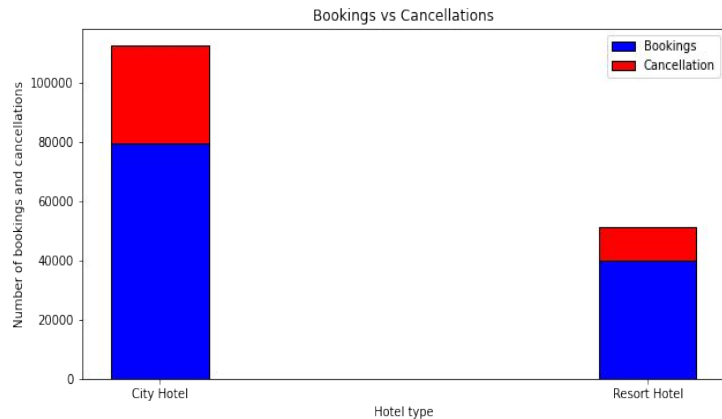


Bookings vs Cancellations

❖ 37.04 percent of the total bookings made were cancelled.

❖ Cancellation percentage in city hotel is slightly more than resort hotel.

# CANCELLATION CONTINUED...

| market_segment | not cancelled | cancelled | bookings | percentage_cancellation |
|---|---|---|---|---|
| Aviation | 185.0 | 52.0 | 237 | 21.94 |
| Complementary | 646.0 | 97.0 | 743 | 13.06 |
| Corporate | 4303.0 | 992.0 | 5295 | 18.73 |
| Direct | 10672.0 | 1934.0 | 12606 | 15.34 |
| Groups | 7714.0 | 12097.0 | 19811 | 61.06 |
| Offline TA/TO | 15908.0 | 8311.0 | 24219 | 34.32 |
| Online TA | 35738.0 | 20739.0 | 56477 | 36.72 |
| Undefined | 0.0 | 2.0 | 2 | 100.00 |
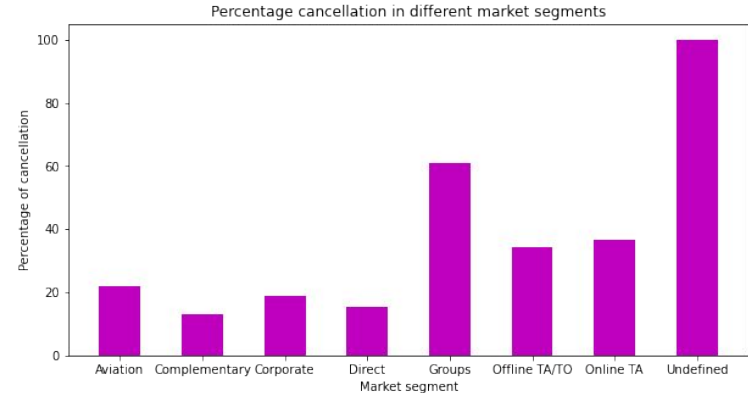


Percentage cancellation in different market segments

- ❖ Cancellation rate is high in group segment.
- ❖ Rate of cancellation in online and offline travel agencies are close to total average cancellation rate.

# CANCELLATION CONTINUED...

| agent | bookings | cancellations | percentage_cancellation |
|-------|----------|---------------|-------------------------|
| 9.0 | 31961 | 13264.0 | 41.50 |
| 240.0 | 13922 | 5484.0 | 39.39 |
| 1.0 | 7191 | 5280.0 | 73.43 |
| 14.0 | 3640 | 652.0 | 17.91 |
| 7.0 | 3539 | 474.0 | 13.39 |
| 6.0 | 3290 | 1025.0 | 31.16 |
| 250.0 | 2870 | 513.0 | 17.87 |
| 241.0 | 1721 | 236.0 | 13.71 |
| 28.0 | 1666 | 110.0 | 6.60 |
| 8.0 | 1514 | 427.0 | 28.20 |



Booking and cancellation for different agents
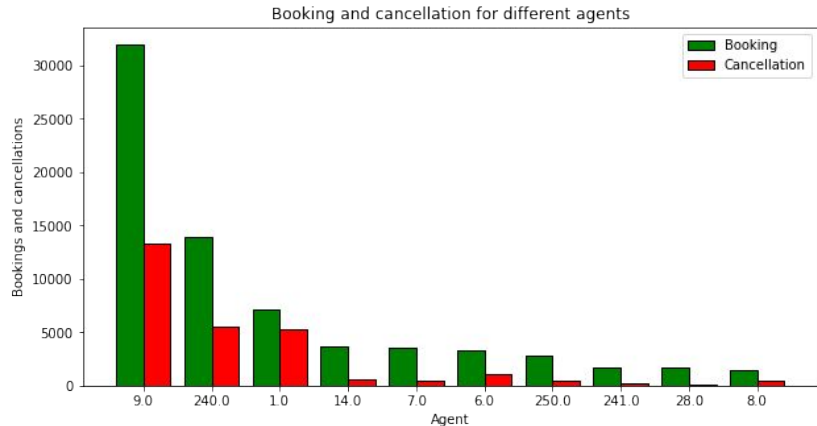
❖ There is a high rate of cancellation for certain agents.

# AVERAGE DAILY RATE

| City Hotel | 105.237722 |
|---|---|
| Resort Hotel | 94.952930 |

| reserved_room_type | average adr | adults |
|---|---|---|
| A | 90.734918 | 1.781098 |
| B | 90.360447 | 1.569767 |
| C | 160.216792 | 2.039700 |
| D | 120.682882 | 2.088902 |
| E | 124.540179 | 1.980260 |
| F | 167.689651 | 1.992406 |
| G | 175.996027 | 2.077841 |
| H | 188.223111 | 2.712146 |
| L | 124.666667 | 2.166667 |
| P | 0.000000 | 0.000000 |

❖ The difference in average ADR for city hotel and resort hotel is only marginal.
❖ Room type 'A' and 'B' have the lowest average ADR

# AVERAGE DAILY RATE CONTINUED...



Average adr and population for room types
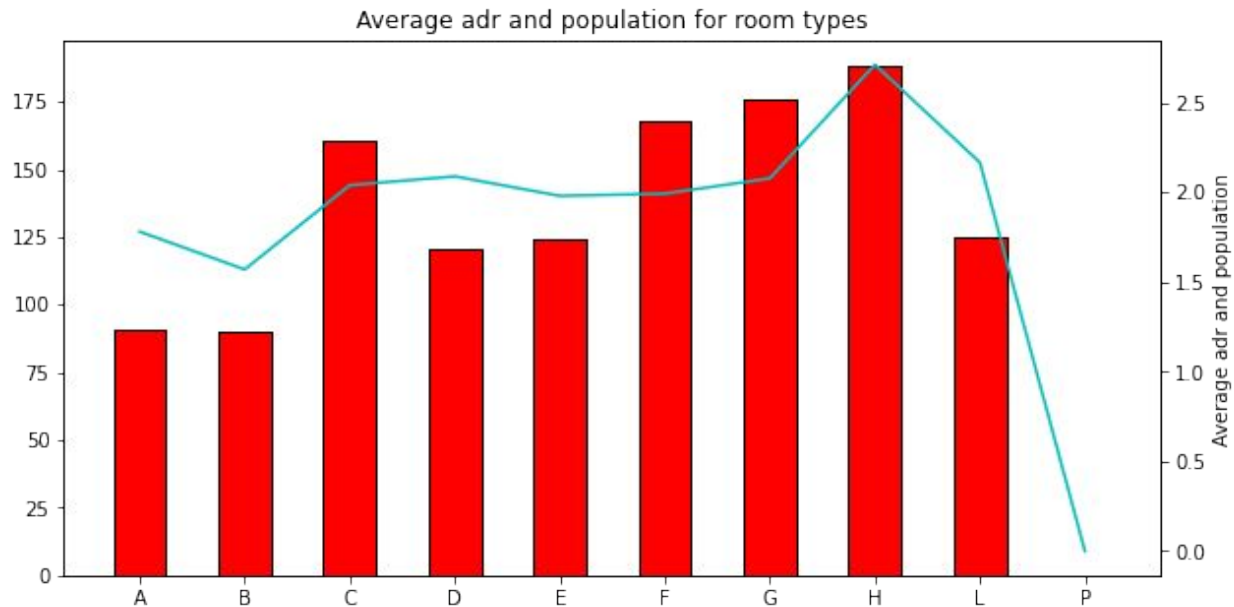
# AVERAGE DAILY RATE CONTINUED...

| arrival_date_month | average adr |
|---|---|
| January | 66.440646 |
| February | 67.999475 |
| March | 74.236730 |
| April | 91.429340 |
| May | 100.068101 |
| June | 107.256526 |
| July | 106.897709 |
| August | 116.025839 |
| September | 97.405611 |
| October | 82.747949 |
| November | 71.784608 |
| December | 73.601693 |

# AVERAGE DAILY RATE CONTINUED...

| total_stay | count | average adr A | average adr D |
|---|---|---|---|
| 1 | 17057 | 90.414349 | 88.537443 |
| 2 | 24165 | 90.304824 | 107.858305 |
| 3 | 23726 | 96.797271 | 114.552112 |
| 4 | 14707 | 94.293790 | 113.713503 |
| 5 | 6260 | 100.788754 | 119.649867 |
| 6 | 2982 | 104.490790 | 128.757158 |
| 7 | 6451 | 88.465057 | 104.382276 |
| 8 | 831 | 95.763541 | 115.155242 |
| 9 | 557 | 97.914642 | 120.336971 |
| 10 | 792 | 88.615695 | 107.270907 |
| 11 | 261 | 89.443711 | 104.052157 |
| 12 | 163 | 89.901524 | 105.440862 |
| 14 | 662 | 80.783681 | 92.172650 |
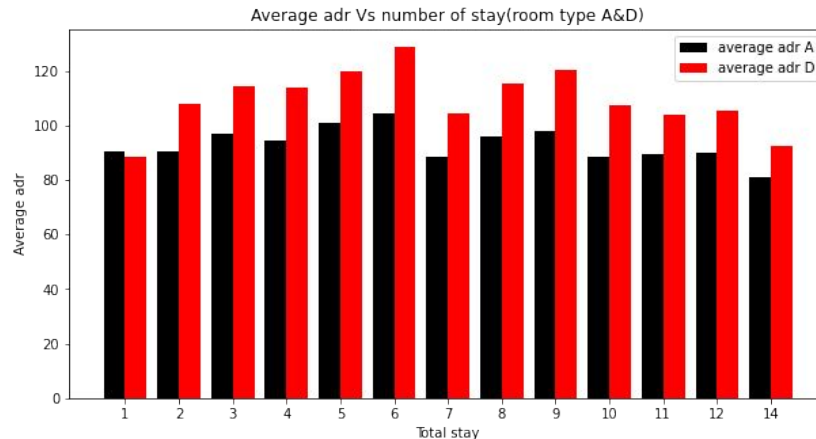


Average adr Vs number of stay(room type A&D)

- ❖ Optimal length of stay for room type 'A' is 14 nights.
- ❖ Optimal length of stay for room type 'D' is 1 night followed by 14 nights.
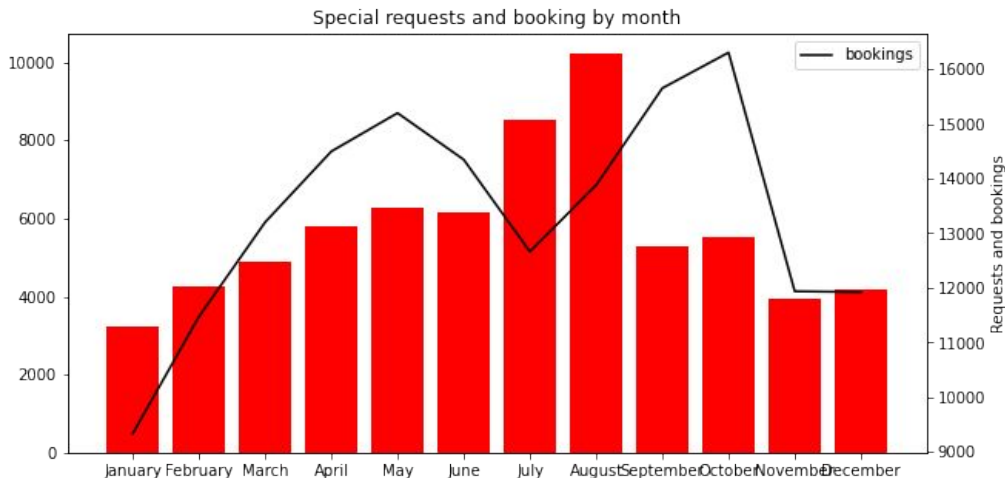
# SPECIAL REQUESTS

| babies | total_of_special_requests | bookings | request per booking |
|--------|---------------------------|----------|---------------------|
| 0 | 66807 | 118473 | 0.56 |
| 1 | 1379 | 900 | 1.53 |
| 2 | 28 | 15 | 1.87 |
| 9 | 0 | 1 | 0.00 |
| 10 | 1 | 1 | 1.00 |

❖ Special requests are more likely in bookings with babies.

# SPECIAL REQUESTS CONTINUED...

| arrival_date_month | total_of_special_requests |
|---|---|
| January | 3246 |
| February | 4241 |
| March | 4872 |
| April | 5783 |
| May | 6280 |
| June | 6169 |
| July | 8505 |
| August | 10223 |
| September | 5268 |
| October | 5516 |
| November | 3949 |
| December | 4163 |



Special requests and booking by month

❖ July and August receives disproportionate number of special requests.

# CONCLUSION

- Bookings to city hotel exceeds resort hotel by far(almost double)
- The hotels had disproportionately high number of bookings from Portugal followed by Great Britain,France,Spain and Germany.
- Almost 75% of the total bookings have been made to room type 'A'
- Room type 'D' comes at the second position with less than one-fifth of the total bookings.
- Bookings to other types of rooms are marginal when compared to these two.
- Number of bookings at the same time of the year has been on a rise with each passing year.
- Each year sees a high number of booking around may and october.

# CONCLUSION CONTINUED...

- Most of the bookings were made through online travel agencies
- Offline travel agencies,groups and direct bookings come in second,third and fourth position respectively in terms of number of bookings made.
- Every countries ,except Portugal ,rely more on online travel agencies for booking.
- In case of Portugal ,bookings made through online and offline travel agencies are almost equal in number.
- Direct booking is close to,or even exceeds offline travel agencies in many cases.
- Special requests are more likely in bookings with babies.
- July and August receives disproportionate number of special requests.

# CONCLUSION CONTINUED...

- 37.04 percent of the total bookings made were cancelled.
- Cancellation percentage in city hotel is slightly more than resort hotel.
- Cancellation rate is high in group segment.
- Rate of cancellation in online and offline travel agencies are close to total average cancellation rate.
- There is a high rate of cancellation for certain agents.
- The difference in average ADR for city hotel and resort hotel is only marginal.
- Room type 'A' and 'B' have the lowest average ADR
- Optimal length of stay for room type 'A' is 14 nights.
- Optimal length of stay for room type 'D' is 1 night followed by 14 nights.

THANK YOU.