hotel_reservations

May 12, 2023

Data Dictionary

- Booking_ID: unique identifier of each booking
- no_of_adults: Number of adults
- no_of_children: Number of Children
- no_of_weekend_nights: Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
- no_of_week_nights: Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
- type_of_meal_plan: Type of meal plan booked by the customer
- required_car_parking_space: Does the customer require a car parking space? (0 - No, 1- Yes)
- room_type_reserved: Type of room reserved by the customer. The values are ciphered (encoded) by INN Hotels.
- lead_time: Number of days between the date of booking and the arrival date
- arrival_year: Year of arrival date
- arrival_month: Month of arrival date
- arrival_date: Date of the month
- market_segment_type: Market segment designation
- repeated_guest: Is the customer a repeated guest? (0 - No, 1- Yes)
- no_of_previous_cancellations: Number of previous bookings that were canceled by the customer prior to the current booking
- no_of_previous_bookings_not_canceled: Number of previous bookings not canceled by the customer prior to the current booking
- avg_price_per_room: Average price per day of the reservation; prices of the rooms are dynamic (in euros)
- no_of_special_requests: Total number of special requests made by the customer (e.g. high floor, view from the room, etc)
- booking_status: Flag indicating if the booking was canceled or not

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights \
0	INN00001	2	0	1
1	INN00002	2	0	2
2	INN00003	1	0	2
3	INN00004	2	0	0
4	INN00005	2	0	1
...	...	...	...	...
36270	INN36271	3	0	2

```
36271    INN36272                2               0                    1
36272    INN36273                2               0                    2
36273    INN36274                2               0                    0
36274    INN36275                2               0                    1

       no_of_week_nights type_of_meal_plan  required_car_parking_space  \
0                      2       Meal Plan 1                            0
1                      3      Not Selected                            0
2                      1       Meal Plan 1                            0
3                      2       Meal Plan 1                            0
4                      1      Not Selected                            0
...                  ...               ...                          ...
36270                  6       Meal Plan 1                            0
36271                  3       Meal Plan 1                            0
36272                  6       Meal Plan 1                            0
36273                  3      Not Selected                            0
36274                  2       Meal Plan 1                            0

       room_type_reserved  lead_time  arrival_year  arrival_month  \
0              Room_Type 1        224          2017             10
1              Room_Type 1          5          2018             11
2              Room_Type 1          1          2018              2
3              Room_Type 1        211          2018              5
4              Room_Type 1         48          2018              4
...                    ...        ...           ...            ...
36270          Room_Type 4         85          2018              8
36271          Room_Type 1        228          2018             10
36272          Room_Type 1        148          2018              7
36273          Room_Type 1         63          2018              4
36274          Room_Type 1        207          2018             12

       arrival_date market_segment_type  repeated_guest  \
0                 2             Offline               0
1                 6              Online               0
2                28              Online               0
3                20              Online               0
4                11              Online               0
...             ...                 ...             ...
36270             3              Online               0
36271            17              Online               0
36272             1              Online               0
36273            21              Online               0
36274            30             Offline               0

       no_of_previous_cancellations  no_of_previous_bookings_not_canceled  \
0                                 0                                     0
1                                 0                                     0
```

2

2	0	0
3	0	0
4	0	0
...	...	...
36270	0	0
36271	0	0
36272	0	0
36273	0	0
36274	0	0

	avg_price_per_room	no_of_special_requests	booking_status
0	65.00	0	Not_Canceled
1	106.68	1	Not_Canceled
2	60.00	0	Canceled
3	100.00	0	Canceled
4	94.50	0	Canceled
...	...	...	...
36270	167.80	1	Not_Canceled
36271	90.95	2	Canceled
36272	98.39	2	Not_Canceled
36273	94.50	0	Canceled
36274	161.67	0	Not_Canceled

[36275 rows x 19 columns]

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36275 entries, 0 to 36274
Data columns (total 19 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   Booking_ID                            36275 non-null  object
 1   no_of_adults                          36275 non-null  int64
 2   no_of_children                        36275 non-null  int64
 3   no_of_weekend_nights                  36275 non-null  int64
 4   no_of_week_nights                     36275 non-null  int64
 5   type_of_meal_plan                     36275 non-null  object
 6   required_car_parking_space            36275 non-null  int64
 7   room_type_reserved                    36275 non-null  object
 8   lead_time                             36275 non-null  int64
 9   arrival_year                          36275 non-null  int64
 10  arrival_month                         36275 non-null  int64
 11  arrival_date                          36275 non-null  int64
 12  market_segment_type                   36275 non-null  object
 13  repeated_guest                        36275 non-null  int64
 14  no_of_previous_cancellations          36275 non-null  int64
 15  no_of_previous_bookings_not_canceled  36275 non-null  int64
 16  avg_price_per_room                    36275 non-null  float64
 17  no_of_special_requests                36275 non-null  int64
```

```
 18  booking_status                       36275 non-null  object
dtypes: float64(1), int64(13), object(5)
memory usage: 5.3+ MB

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36275 entries, 0 to 36274
Data columns (total 19 columns):
 #   Column                         Non-Null Count  Dtype
---  ------                         --------------  -----
 0   Booking_ID                     36275 non-null  object
 1   no_of_adults                   36275 non-null  int64
 2   no_of_children                 36275 non-null  int64
 3   no_of_weekend_nights           36275 non-null  int64
 4   no_of_week_nights              36275 non-null  int64
 5   type_of_meal_plan              36275 non-null  object
 6   required_car_parking_space     36275 non-null  int64
 7   room_type_reserved             36275 non-null  object
 8   lead_time                      36275 non-null  int64
 9   arrival_year                   36275 non-null  int64
 10  arrival_month                  36275 non-null  int64
 11  arrival_date                   36275 non-null  int64
 12  market_segment_type            36275 non-null  object
 13  repeated_guest                 36275 non-null  int64
 14  no_of_previous_cancellations   36275 non-null  int64
 15  no_of_previous_bookings_not_canceled  36275 non-null  int64
 16  avg_price_per_room             36275 non-null  float64
 17  no_of_special_requests         36275 non-null  int64
 18  booking_status                 36275 non-null  object
dtypes: float64(1), int64(13), object(5)
memory usage: 5.3+ MB

 array(['Meal Plan 1', 'Not Selected', 'Meal Plan 2', 'Meal Plan 3'],
       dtype=object)

 Index(['Booking_ID', 'no_of_adults', 'no_of_children',
        'no_of_weekend_nights_booked', 'no_of_week_nights_booked',
        'meal_plan_type', 'has_parking_space', 'reserved_room_type',
        'lead_time', 'arrival_year', 'arrival_month', 'arrival_date',
        'market_segment', 'repeated_guest', 'total_cancellations',
        'total_successful_bookings', 'average_room_price',
        'total_special_requests', 'booking_status', 'arrival_date_time'],
       dtype='object')
```

	no_of_adults	no_of_children	no_of_weekend_nights_booked \
count	36275.000000	36275.000000	36275.000000
mean	1.844962	0.105279	0.810724
std	0.518715	0.402648	0.870644
min	0.000000	0.000000	0.000000

	no_of_week_nights_booked	has_parking_space	lead_time
25%	2.000000	0.000000	0.000000
50%	2.000000	0.000000	1.000000
75%	2.000000	0.000000	2.000000
max	4.000000	10.000000	7.000000

	no_of_week_nights_booked	has_parking_space	lead_time
count	36275.000000	36275.000000	36275.000000
mean	2.204300	0.030986	85.232557
std	1.410905	0.173281	85.930817
min	0.000000	0.000000	0.000000
25%	1.000000	0.000000	17.000000
50%	2.000000	0.000000	57.000000
75%	3.000000	0.000000	126.000000
max	17.000000	1.000000	443.000000

	arrival_year	arrival_month	arrival_date	repeated_guest
count	36275.000000	36275.000000	36275.000000	36275.000000
mean	2017.820427	7.423653	15.596995	0.025637
std	0.383836	3.069894	8.740447	0.158053
min	2017.000000	1.000000	1.000000	0.000000
25%	2018.000000	5.000000	8.000000	0.000000
50%	2018.000000	8.000000	16.000000	0.000000
75%	2018.000000	10.000000	23.000000	0.000000
max	2018.000000	12.000000	31.000000	1.000000

	total_cancellations	total_successful_bookings	average_room_price
count	36275.000000	36275.000000	36275.000000
mean	0.023349	0.153411	103.423539
std	0.368331	1.754171	35.089424
min	0.000000	0.000000	0.000000
25%	0.000000	0.000000	80.300000
50%	0.000000	0.000000	99.450000
75%	0.000000	0.000000	120.000000
max	13.000000	58.000000	540.000000

	total_special_requests
count	36275.000000
mean	0.619655
std	0.786236
min	0.000000
25%	0.000000
50%	0.000000
75%	1.000000
max	5.000000

	Booking_ID	reserved_room_type	market_segment	booking_status
count	36275	36275	36275	36275

```
unique       36275                    7              5              2
top      INN00001        Room_Type 1         Online    Not_Canceled
freq            1               28130          23214          24390

0

Booking_ID                          0
no_of_adults                        0
no_of_children                      0
no_of_weekend_nights_booked         0
no_of_week_nights_booked            0
meal_plan_type                      0
has_parking_space                   0
reserved_room_type                  0
lead_time                           0
arrival_year                        0
arrival_month                       0
arrival_date                        0
market_segment                      0
repeated_guest                      0
total_cancellations                 0
total_successful_bookings           0
average_room_price                  0
total_special_requests              0
booking_status                      0
arrival_date_time                  37
dtype: int64
```
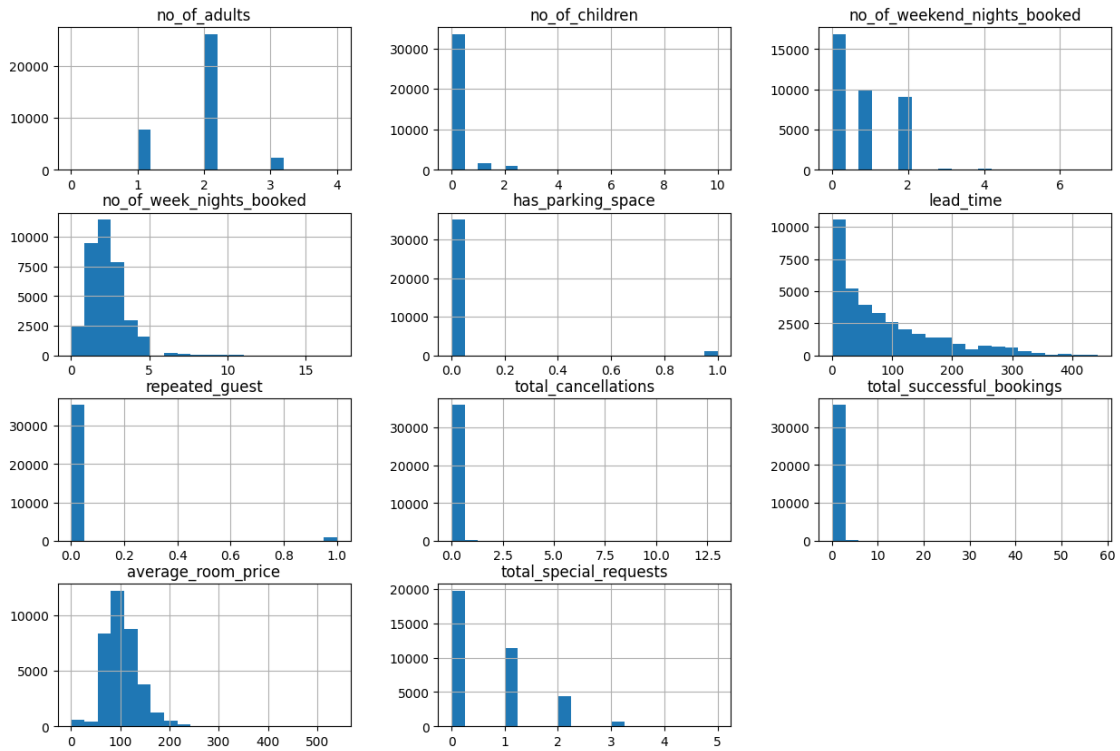
Distribution of Numerical Variables


```
Index(['Booking_ID', 'no_of_adults', 'no_of_children',
       'no_of_weekend_nights_booked', 'no_of_week_nights_booked',
       'meal_plan_type', 'has_parking_space', 'reserved_room_type',
       'lead_time', 'arrival_year', 'arrival_month', 'arrival_date',
       'market_segment', 'repeated_guest', 'total_cancellations',
       'total_successful_bookings', 'average_room_price',
       'total_special_requests', 'booking_status', 'arrival_date_time',
       'month_year'],
      dtype='object')
```

Data source:https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset