```
import numpy as np
```

```
df_airline_booking_passenger_data.shape(50000, 14)
```

```
df_airline_booking_passenger_data.head()
```

	booking_id	property_id	booking_date	check_in_date	checkout_date	no_guests	room_category	booking_platform	selling_price	booking_status	review_generated	review_received	review_rating
0	May2022/18889/17	18889	30-Apr-22	1/05/2022	3/05/2022	2.0	RT1	others	7500	Checked Out	0	0	0.0
1	May2022/18889/18	18889	30-Apr-22	1/05/2022	3/05/2022	2.0	RT1	others	7500	Checked Out	1	1	5.0
2	May2022/18889/19	18889	30-Apr-22	1/05/2022	3/05/2022	2.0	RT1	others	7500	Checked Out	0	0	0.0
3	May2022/18889/20	18889	30-Apr-22	1/05/2022	3/05/2022	2.0	RT1	others	7500	Checked Out	0	0	0.0
4	May2022/18889/21	18889	30-Apr-22	1/05/2022	3/05/2022	2.0	RT1	direct offline	7500	Checked Out	1658	1658	5.0

```
df_airline_booking_passenger_data.sample(5)
```

	property_id	no_guests	room_category	booking_status	review_generated	review_received
14006	17842.0	2.0000000	RT1	Cancelled	0.0000000	0.0000000
4811	17558.0000000	1.0000000	1.0000000	0.0000000	0.0000000	0.0000000
14006	17842.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000
3991	17558.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000
3991	17558.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000

```
df_airline_booking_passenger_data.describe()
```

	property_id	no_guests	room_category	review_generated	review_received
count	50000.000000	49997.000000	50000.000000	50000.000000	50000.000000
mean	17842.000000	2.000000	2.000000	0.000000	0.000000
std	1118.034077	1.414214	1.414214	0.000000	0.000000
min	16558.000000	1.000000	1.000000	0.000000	0.000000
25%	17558.000000	1.000000	1.000000	0.000000	0.000000
50%	17842.000000	2.000000	2.000000	0.000000	0.000000
75%	18889.000000	3.000000	3.000000	0.000000	0.000000
max	19556.000000	4.000000	4.000000	1.000000	1.000000

```
df_airline_booking_passenger_data.isnull().sum()
```

```
booking_id          0
property_id         0
booking_date        0
check_in_date       0
checkout_date       0
no_guests           3
room_category       0
booking_platform    0
selling_price       0
booking_status      0
review_generated    0
review_received     0
review_rating       0
dtype: int64
```

```
df_airline_booking_passenger_data.duplicated().sum()
```

```
0
```

```
df_airline_booking_passenger_data.nunique()
```

```
booking_id          50000
property_id         10000
booking_date          365
check_in_date         365
checkout_date         365
no_guests               4
room_category           4
booking_platform        5
selling_price       14006
booking_status          2
review_generated        2
review_received         2
review_rating           6
dtype: int64
```

```
df_airline_booking_passenger_data.no_guests.value_counts()
```

```
no_guests
2.0    25000
1.0    12500
3.0     7500
4.0     5000
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.room_category.value_counts()
```

```
room_category
RT1    25000
RT2    12500
RT3     7500
RT4     5000
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.booking_platform.value_counts()
```

```
booking_platform
others          25000
logtrip         12500
direct offline   7500
direct online    4999
tripster            1
Name: count, dtyp: int64
```

```
df_airline_booking_passenger_data.booking_status.value_counts()
```

```
booking_status
Checked Out    25000
Cancelled      25000
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.review_generated.value_counts()
```

```
review_generated
0    50000
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.review_received.value_counts()
```

```
review_received
0    50000
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.review_rating.value_counts()
```

```
review_rating
0.0    25000
5.0    12500
4.0     7500
3.0     5000
2.0        1
1.0        1
Name: count, dtype: int64
```

```
df_airline_booking_passenger_data.shape
```

```
(50000, 14)
```

```
df_airline_booking_passenger_data.dtypes
```

```
booking_id          object
property_id          int64
booking_date        object
check_in_date       object
checkout_date       object
no_guests           float64
room_category        int64
booking_platform    object
selling_price        int64
booking_status      object
review_generated     int64
review_received      int64
review_rating        int64
dtype: object
```

```
df_airline_booking_passenger_data.info()
```

```
<class>pandas.core.frame.DataFrame>
RangeIndex: 50000 entries, 0 to 49999
Data columns (total 14 columns):
 booking_id          50000 non-null  object
 property_id         50000 non-null  int64
 booking_date        50000 non-null  object
 check_in_date       50000 non-null  object
 checkout_date       50000 non-null  object
 no_guests           49997 non-null  float64
 room_category       50000 non-null  int64
 booking_platform    50000 non-null  object
 selling_price       50000 non-null  int64
 booking_status      50000 non-null  object
 review_generated    50000 non-null  int64
 review_received     50000 non-null  int64
 review_rating       50000 non-null  int64
dtype: int64, memory usage: 5.3MB
```

```
df_airline_booking_passenger_data.sample(5)
```

	booking_id	property_id	check_in_date	check_out_date	no_guests	room_category	booking_platform	selling_price	booking_status	review_generated	review_received	review_rating
14006	17842.0	2.0000000	RT1	Cancelled	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
4811	17558.0000000	1.0000000	1.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
14006	17842.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
3991	17558.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
3991	17558.0000000	2.0000000	2.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0