
**To investigate the relationship between reserved and assigned room types and its
impact on income**

HOTEL DATA

DATA

Raw Data

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_weekend_nights	stays_in_week_nights	adults	...	deposit_type	agent	company	days_in_waiting_list	customer_type	adr
0	Resort Hotel	0	342	2015	July	27	1	0	0	0	2	...	No Deposit	NaN	NaN	0	Transient	0.0
1	Resort Hotel	0	737	2015	July	27	1	0	0	0	2	...	No Deposit	NaN	NaN	0	Transient	0.0
2	Resort Hotel	0	7	2015	July	27	1	0	0	1	1	...	No Deposit	NaN	NaN	0	Transient	75.0
3	Resort Hotel	0	13	2015	July	27	1	0	0	1	1	...	No Deposit	304.0	NaN	0	Transient	75.0
4	Resort Hotel	0	14	2015	July	27	1	0	0	2	2	...	No Deposit	240.0	NaN	0	Transient	98.0

Clean Data

	hotel	reserved_room_type	assigned_room_type	reserved_assigned	total_stay	adr	revenue	potential_revenue	revenue_difference	potential_profit
2	Resort Hotel	A	C	False	1	75.00	75.00	93.142347	18.142347	True
12	Resort Hotel	D	E	False	4	97.00	388.00	429.814531	41.814531	True
15	Resort Hotel	D	E	False	4	97.00	388.00	429.814531	41.814531	True
17	Resort Hotel	A	E	False	1	88.20	88.20	93.142347	4.942347	True
18	Resort Hotel	A	G	False	1	107.42	107.42	93.142347	-14.277653	False



Removed irrelevant columns



Created new columns by extrapolating new information from existing data

DATA LIMITATIONS

- **Lack of information:**
 - **Why are some customers not assigned their reserved rooms?**
 - **What are the differences in each room type?**
 - **Why are some room types more popular than others?**
 - **Quantity of each room type**
 - **ADR (Average Day Rate)**
 - **Can't analyse daily rate changes and fluctuations**
-

DATA EXTRAPOLATION

	hotel	reserved_room_type	assigned_room_type	reserved_assigned	total_stay	adr	revenue	potential_revenue	revenue_difference	potential_profit
2	Resort Hotel	A	C	False	1	75.00	75.00	93.142347	18.142347	True
12	Resort Hotel	D	E	False	4	97.00	388.00	429.814531	41.814531	True
15	Resort Hotel	D	E	False	4	97.00	388.00	429.814531	41.814531	True
17	Resort Hotel	A	E	False	1	88.20	88.20	93.142347	4.942347	True
18	Resort Hotel	A	G	False	1	107.42	107.42	93.142347	-14.277653	False

Average rate by room type

A	93.142347
B	94.450264
C	113.423583
D	107.453633
E	117.704224
F	151.889931
G	166.530309
H	171.380772
I	40.843774
K	53.698889
L	8.000000
P	0.000000

- **reserved_assigned** : compares reserved and assigned room type
- **revenue** : $\text{total_stay} * \text{average assigned room type rate}$
- **potential_revenue** : $\text{total_stay} * \text{average reserved room type rate}$
- **revenue_difference**: $\text{potential_revenue} - \text{revenue}$
- **potential_profit** : check if revenue_difference is positive

CHI - SQUARE TEST: GOODNESS OF FIT (RESERVED ROOMS)

Value count of each reserved room type

```
A      85994
D      19201
E       6535
F       2897
G       2094
B       1118
C        932
H        601
P         12
L          6
Name: reserved_room_type, dtype: int64
```

> Hypothesis:

- > Test whether the reserved room types are equally distributed or not
- > H0: Distribution is not equal
- > H1: Distribution is equal

> Result:

- > Test statistic: 4.48732
- > P-vale: 0.87652

> Conclusion:

- > At 0.5 level of significance, we **FAILED TO REJECT** null hypotheses.
 - > Distribution is likely unequal
-

CHI - SQUARE TEST: GOODNESS OF FIT (ASSIGNED ROOMS)

Value count of each assigned room type

```
A      74053
D      25322
E       7806
F      3751
G      2553
C      2375
B      2163
H       712
I       363
K       279
P        12
L         1
Name: assigned_room_type, dtype: int64
```

➤ Hypothesis:

➤ Test whether the assigned room types are equally distributed or not

➤ H0: Distribution is not equal

➤ H1: Distribution is equal

➤ Result:

➤ Test statistic: 3.36297

➤ P-value: 0.94815

➤ Conclusion:

➤ At 0.5 level of significance, we **FAILED TO REJECT** null hypotheses.

➤ Distribution is likely unequal

HYPOTHESIS: REVENUE OF HOTELS WOULD INCREASE IF CUSTOMERS WERE ASSIGNED RESERVED ROOM TYPE

Current Revenue	3,510,294.4
Potential Revenue	4,021,700.69
Difference	511,405.29

- Majority of customers received their reserved room type
- Of the people who didn't, 63% of them would have resulted in a higher revenue if they were assigned their requested room type
- 15% increase in revenue if customers were assigned their requested room type

Customers who were assigned their reserved room type

Reserved - Assign	Frequency	Percentage
True	104473	87.50%
False	14917	12.50%

Of those who did not received their reserved room, if they were assigned their requested room

Potential of Profit		
True	9515	63.79%
False	5402	36.21%

RESEARCH: ROOM TYPE D

```
assigned_room_type
P      0.000000
L      8.000000
I     40.843774
K     53.698889
A     93.142347
B     94.450264
D    107.453633
C    113.423583
E    117.704224
F    151.889931
G    166.530309
H    171.380772
Name: adr, dtype: float64
```

- **Left: Average room rate of each assigned room type in ascending order**
- **Median room type is Room D**
- **Using type D as a benchmark to compare different rates depending on if customer reserved D or a different room**
- **For customer who booked a different room, compare rates if they originally booked a cheaper or more expensive room**

-
- **Group customers who were assigned room D:**
 - **1. Reserved room D (DD)**
 - **2. Reserved a higher rate room (HD)**
 - **3. Reserved a lower rate room (LD)**
 - **If you booked a higher rate room, you pay a higher rate for room D and vice versa**
 - **The difference in mean between DD and HD is more higher than between DD and LD.**
 - **The hotel can get a higher revenue for this room type value when reassigning HD customers**
 - **However, more research need to be conducted on the original revenue of these HD customers.**
-

```
Number of values in dd: 104473
Number of values in hd: 12138
Number of values in ld: 92335
```

DD

```
Mean: 104.46886104543756
Median: 96.0
Min: 0.0
Max: 5400.0
Standard Deviation: 50.363316628974104
Variance: 2536.4636618702993
```

HD

```
Mean: 149.89091613115835
Median: 148.35000000000002
Min: 0.0
Max: 451.5
Standard Deviation: 69.47852220958265
Variance: 4827.265048427469
```

LD

```
Mean: 98.4978543347593
Median: 93.6
Min: 0.0
Max: 5400.0
Standard Deviation: 43.91418340216929
Variance: 1928.455503879361
```

ONE SAMPLE Z-TEST

➤ Hypothesis:

- Test that the mean adr of customers who reserved and were assigned room type D is equal to those who reserved a different room (with a higher average adr) and were assigned to type D

➤ Null Hypothesis: $\mu=\mu_0$ Alternative: $\mu\neq\mu_0$

➤ Result:

- The test statistic is: **72.023** The p-value is: **0.0**
- At **0.05** level of significance, we **REJECT** null hypothesis.

➤ Hypothesis:

- Test that the mean adr of customers who reserved and were assigned room type D is equal to those who reserved a different room (with a lower average adr) and were assigned to type D

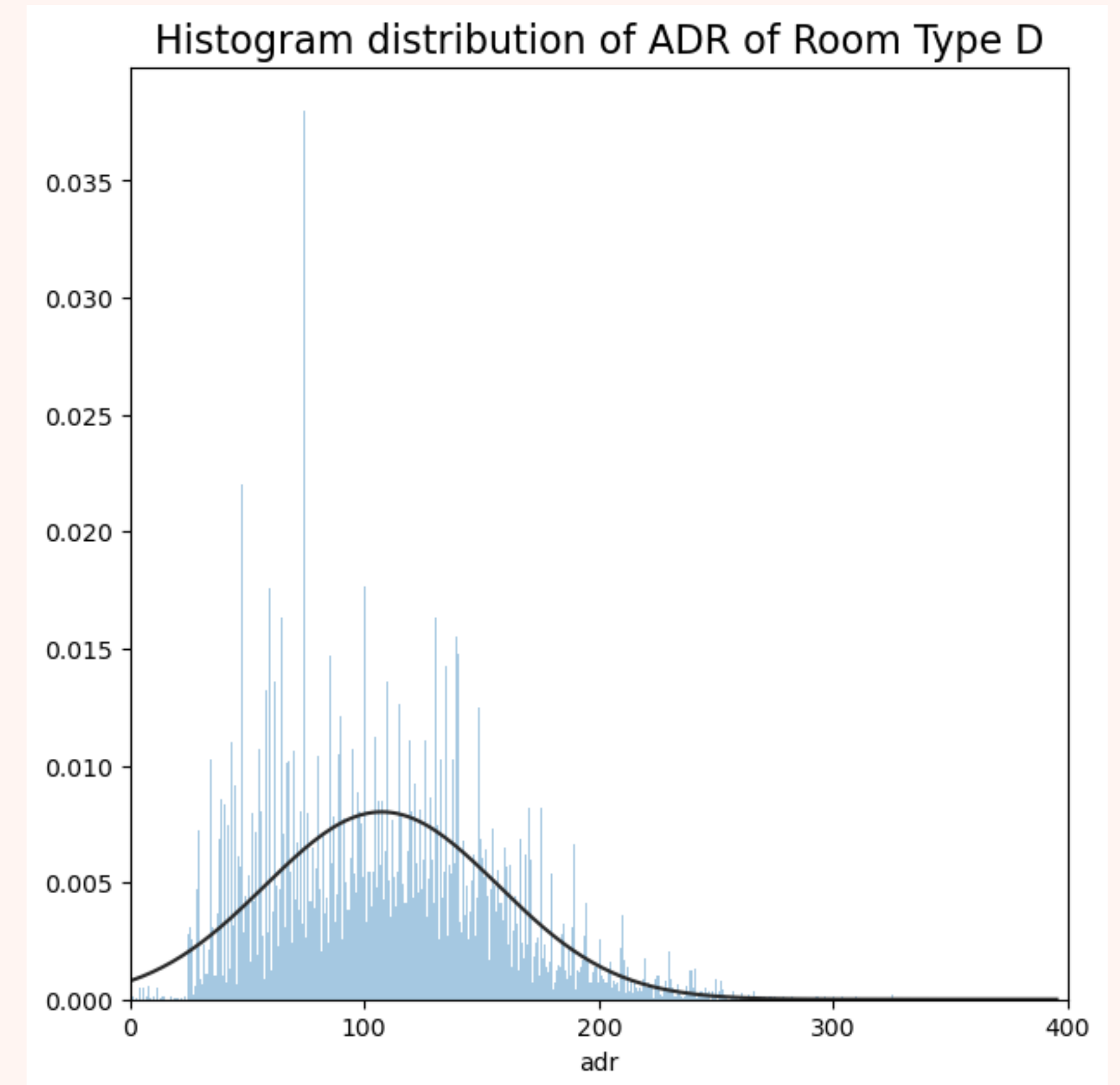
➤ Null Hypothesis: $\mu=\mu_0$ Alternative: $\mu\neq\mu_0$

➤ Result:

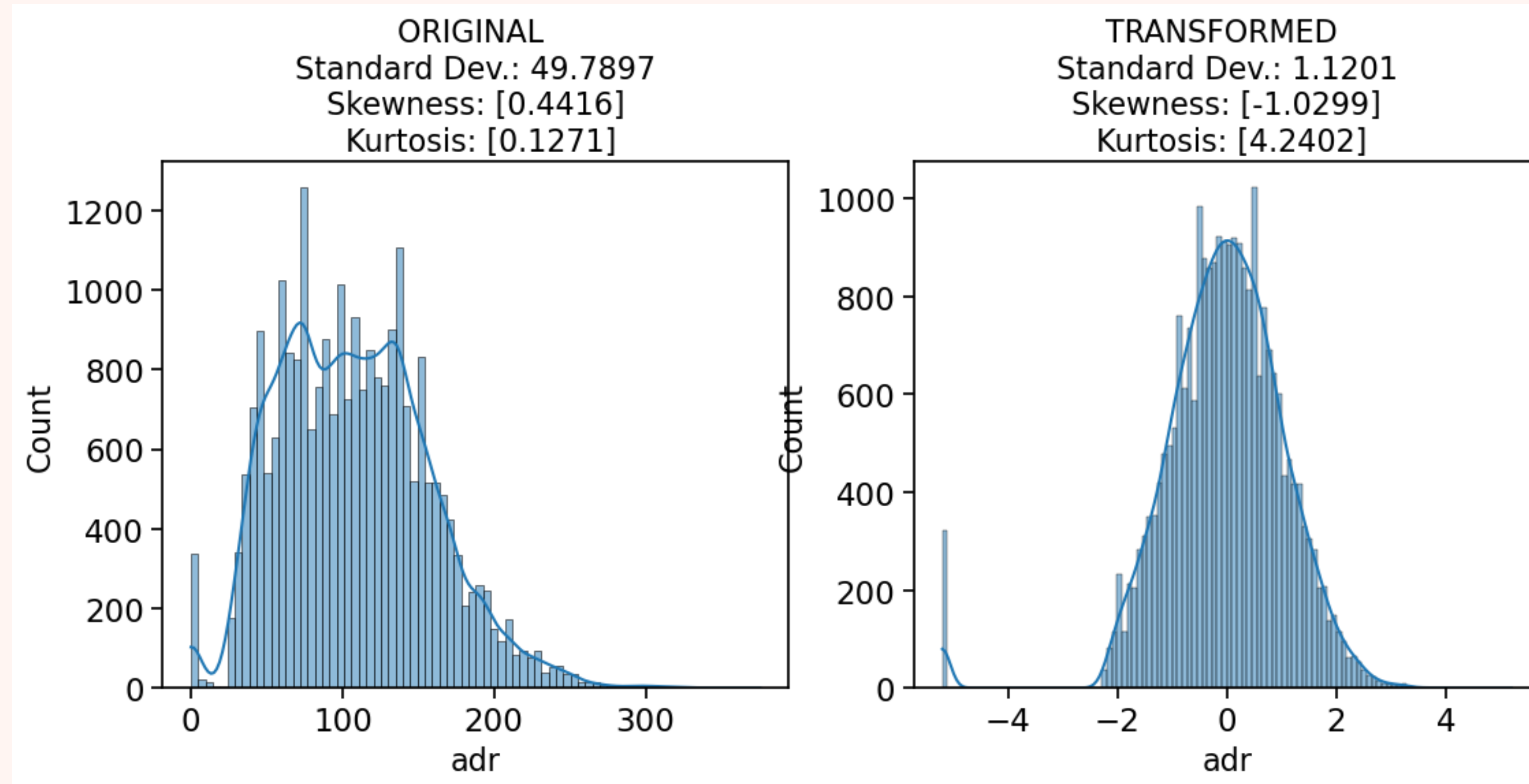
- The test statistic is: **-41.317** The p-value is: **0.0**
 - At **0.05** level of significance, we **REJECT** null hypothesis.
-

DISTRIBUTION OF ROOM D RATES

- **Kurtosis: 0.1270851623589988**
- **Mesokurtic distribution (normal)**
- **Skewness: 0.44156044404420286**
- **Distribution is mildly skewed**



TRANSFORMED DISTRIBUTION



- **Transformed data has a more normal distribution**
- **Standard deviation has the highest impact after transformation**

CONCLUSION

- **Assigning customers their reserved rooms increases revenue**
 - **However, in some cases, reassigning rooms result in above average room rates**
 - **Further investigation:**
 - **Further exploration in other reassigned rooms and room rates**
 - **e.g. comparing assigned room type G with other reserved rooms**
-