

---



# Hotel Bookings Analysis



---

# Problem Statement

The online travel platform specializing in hotel reservations seeks to understand customer booking behaviors and cancellation patterns in order to improve profitability, optimize channel strategy, and enhance customer retention. The provided dataset (Hotel\_bookings\_final.csv) contains detailed booking transactions that capture information such as booking channels, room types, star ratings, stay lengths, and cancellations.

# Objective

1

Identify key booking and cancellation patterns across channels, room types, and star ratings.

2

Conduct root cause analysis to uncover factors driving cancellations, channel performance, and temporal variations.

3

Provide actionable business recommendations to minimize cancellations, improve profitability, and optimize pricing and promotion strategies.

# Dataset

## Overview

1

Hotel\_bookings\_final.csv

2

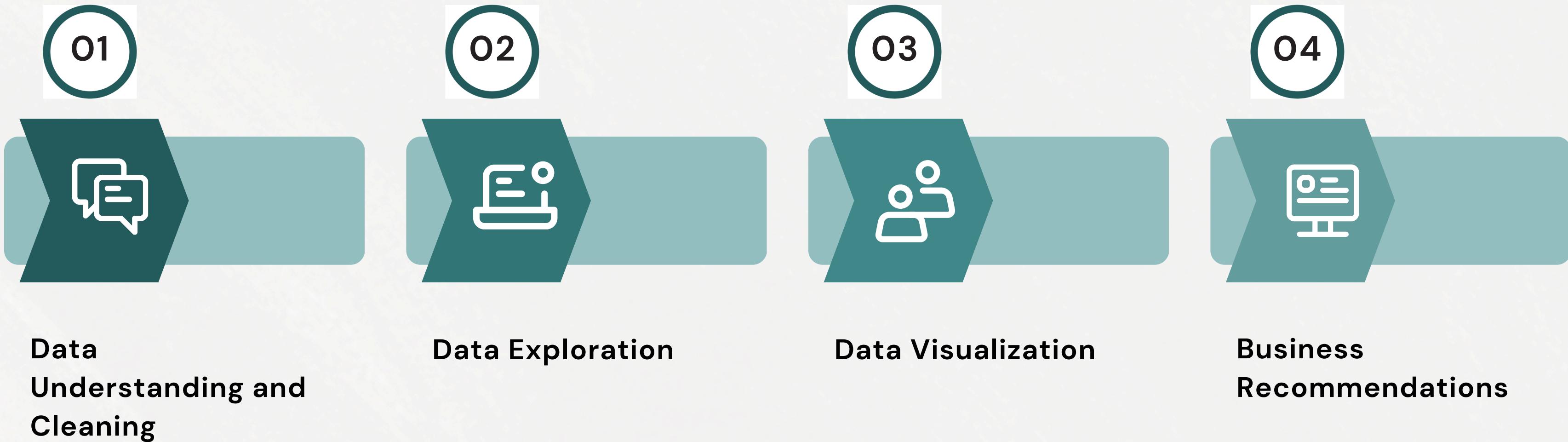
Contains booking transactions from an online hotel reservation platform

---

# Tools and Techniques Used

For this analysis, **Python** was used for data cleaning, exploration, and visualization with the help of libraries such as Pandas and Matplotlib. Python enabled efficient handling of large datasets and generation of insights through charts and trend analysis. In addition, **Excel** was utilized for quick data validation, preliminary exploration, and creating summary tables, supporting easier interpretation and presentation of results. Together, these tools ensured both analytical depth and clarity in reporting findings.

# Methodology



# — Understanding and cleaning data

- Dataset contained 30,000 rows × 24 columns.
- Date columns were converted from object to datetime format for proper time-based analysis.
- No duplicate rows were found, ensuring data reliability.
- 5,468 missing values were identified in “check\_in\_date” and “check\_out\_date” columns.
- Further analysis revealed that these missing values correspond to cancelled or failed bookings, since no actual check-in/check-out took place.

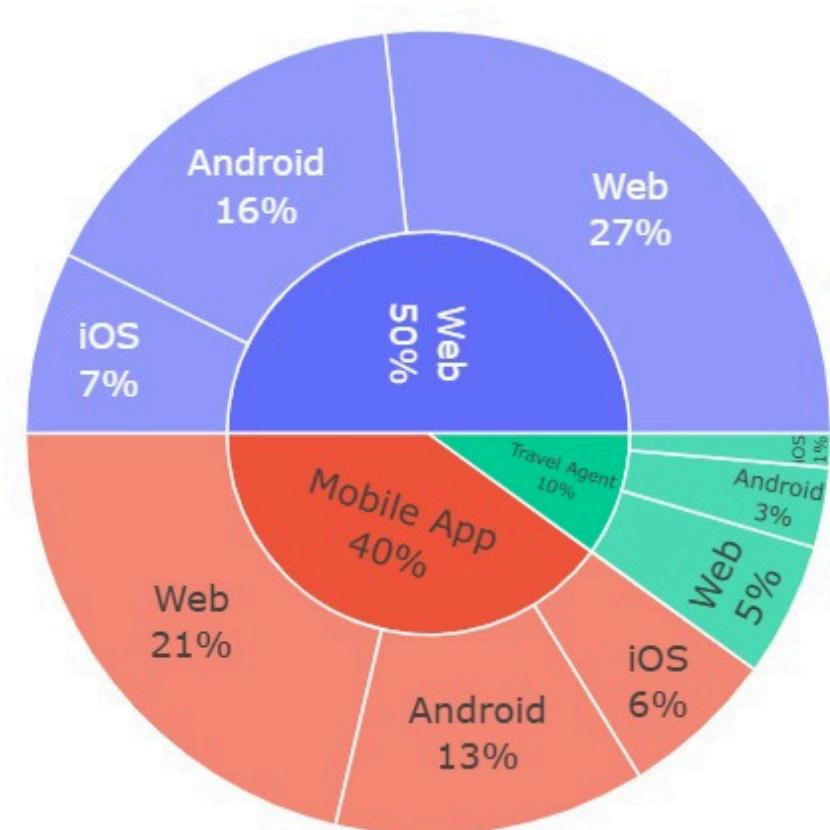
# Dataset Overview

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	customer_id	property_id	city	star_rating	booking_date	check_in_date	check_out_date	room_type	num_rooms_booked	stay_type	booking_channel	booking_value	costprice	markup	selling_price	payment_method	refund_status	refund_amount	channel_of_booking	booking_status	travel_date	cashback	coupon_redeem	Coupon Used?
2	492	3	San Francisco	4	01-04-2024	24-05-2024	26-05-2024	Standard	1	Leisure	Mobile App	19361	19361	5981	25342	PayPal	Yes	369.6489946	Web	Confirmed	04-03-2024	5.37469396	0	No
3	180	3	Dallas	3	01-04-2024	10-05-2024	17-05-2024	Deluxe	1	Leisure	Mobile App	6137	6137	1896	8033	Bank Transfer	Yes	492.505903	Web	Confirmed	19-07-2024	7.16103261	0	No
4	50	5	Dallas	3	01-04-2024	31-05-2024	05-06-2024	Deluxe	1	Business	Web	22702	22702	7013	29715	Debit Card	Yes	0	iOS	Confirmed	22-03-2024	0	0	No
5	294	3	Orlando	4	01-04-2024	18-04-2024	24-04-2024	Deluxe	3	Leisure	Web	34068	34068	10524	44592	Bank Transfer	Yes	545.5415163	Android	Confirmed	24-11-2024	7.93217008	24.5004415	Yes
6	40	5	Seattle	5	01-04-2024	18-04-2024	21-04-2024	Deluxe	1	Leisure	Mobile App	12127	12127	3746	15873	Debit Card	Yes	211.3676019	Web	Cancelled	02-03-2024	0	0	No
7	50	3	New York	2	01-04-2024	18-04-2024	21-04-2024	Standard	1	Leisure	Web	4452	37704	11647	49351	Credit Card	Yes	585.9048548	Web	Confirmed	22-12-2024	8.51905275	27.08844078	Yes
8	415	7	Los Angeles	4	01-04-2024	11-04-2024	14-04-2024	Deluxe	1	Leisure	Web	8429	8429	2604	11033	PayPal	Yes	111.449291	Android	Confirmed	02-03-2024	0	0	No
9	336	1	Dallas	4	01-04-2024			Suite	1	Leisure	Travel Agent	10414	10414	3217	13631	Credit Card	Yes	180.1193334	Web	Cancelled	19-11-2024	2.61893393	4.123233708	Yes
10	461	3	San Francisco	4	01-04-2024	12-05-2024	14-05-2024	Standard	1	Business	Mobile App	15064	15064	4654	19718	Credit Card	No	0	Android	Failed	03-03-2024	1.77618472	0	No
11	5	5	Los Angeles	5	01-04-2024	15-04-2024	19-04-2024	Deluxe	1	Leisure	Web	17396	17396	5374	22770	Bank Transfer	Yes	155.6501024	Web	Confirmed	28-03-2024	0	0	No
12	463	5	Orlando	3	01-04-2024	24-04-2024	30-04-2024	Standard	1	Business	Web	21845	21845	6748	28593	Debit Card	No	0	Web	Confirmed	22-03-2024	0	0	No
13	448	7	Boston	3	01-04-2024	29-05-2024	30-05-2024	Standard	1	Leisure	Web	30804	30804	9515	40319	Bank Transfer	Yes	534.3023604	Android	Confirmed	10-03-2024	0	0	No
14	497	3	Miami	3	01-04-2024	13-05-2024	15-05-2024	Standard	1	Leisure	Web	10158	10158	3138	13296	Debit Card	Yes	613.7492657	Web	Confirmed	22-08-2024	8.9239103	22.47262229	Yes
15	271	7	San Francisco	5	01-04-2024	14-05-2024	18-05-2024	Standard	2	Leisure	Web	34076	34076	10526	44602	Debit Card	Yes	442.7009385	Android	Confirmed	17-07-2024	0	0	No
16	493	5	Los Angeles	2	01-04-2024	02-05-2024	08-05-2024	Deluxe	1	Leisure	Mobile App	30645	30645	9466	40111	Credit Card	No	0	Web	Confirmed	30-03-2024	0	0	No
17	269	5	New York	4	01-04-2024	12-05-2024	16-05-2024	Standard	1	Business	Web	12963	12963	4005	16968	Bank Transfer	Yes	278.9459413	iOS	Failed	20-10-2024	0	13.08562943	Yes
18	179	3	Orlando	4	01-04-2024	17-05-2024	21-05-2024	Standard	2	Business	Mobile App	24554	24554	7585	32139	Credit Card	Yes	493.3679224	Web	Confirmed	16-03-2024	7.17356636	0	No
19	267	5	Orlando	3	01-04-2024	03-05-2024	10-05-2024	Standard	1	Business	Mobile App	9885	9885	3054	12939	Bank Transfer	Yes	81.47770028	Android	Confirmed	10-03-2024	0	0	No
20	12	3	Boston	4	01-04-2024	27-05-2024	29-05-2024	Standard	1	Leisure	Mobile App	21976	21976	6789	28765	Debit Card	Yes	455.3935384	Android	Confirmed	03-03-2024	6.62141907	0	No
21	119	5	Dallas	3	01-04-2024	19-05-2024	22-05-2024	Standard	1	Business	Web	20040	20040	6191	26231	Debit Card	Yes	428.2751871	Web	Confirmed	22-03-2024	0	0	No
22	189	7	Boston	5	01-04-2024	26-04-2024	01-05-2024	Standard	2	Leisure	Mobile App	22919	22919	7080	29999	PayPal	Yes	0	iOS	Confirmed	27-11-2024	0	2.600076337	Yes
23	241	5	Orlando	4	01-04-2024	15-04-2024	16-04-2024	Suite	1	Business	Mobile App	27164	27164	8391	35555	PayPal	Yes	0	iOS	Confirmed	17-09-2024	0	9.442209458	Yes
24	448	3	Las Vegas	4	01-04-2024	18-05-2024	25-05-2024	Deluxe	3	Business	Web	24677	18223	5629	23852	Credit Card	No	0	Android	Confirmed	22-03-2024	3.67490811	0	No
25	151	3	Dallas	3	01-04-2024	18-04-2024	23-04-2024	Standard	1	Business	Mobile App	8393	8393	2593	10986	Credit Card	Yes	0	Android	Confirmed	26-10-2024	1.29115742	3.958299314	Yes
26	362	3	Boston	5	01-04-2024			Deluxe	1	Business	Travel Agent	24185	24185	7471	31656	Bank Transfer	Yes	302.2249737	Android	Cancelled	05-03-2024	4.39434914	0	No
27	242	5	Dallas	5	01-04-2024	18-05-2024	22-05-2024	Standard	2	Business	Web	18341	18341	5666	24007	Bank Transfer	No	0	Web	Confirmed	21-03-2024	0	0	No
28	179	3	Miami	3	01-04-2024	07-04-2024	08-04-2024	Deluxe	1	Business	Mobile App	12584	22950	7090	30040	Debit Card	Yes	0	Web	Confirmed	24-12-2024	5.87648025	15.64937971	Yes
29	110	1	Las Vegas	4	01-04-2024	21-04-2024	26-04-2024	Standard	1	Leisure	Mobile App	20689	20689	6391	27080	Credit Card	Yes	0	Web	Confirmed	26-12-2024	8.56367335	23.63223745	Yes
30	389	5	New York	4	01-04-2024			Standard	1	Business	Mobile App	35874	12583	3887	16470	Credit Card	Yes	0	Web	Cancelled	25-11-2024	0	13.80123436	Yes
31	192	3	Dallas	2	01-04-2024	08-05-2024	12-05-2024	Standard	1	Leisure	Web	32610	18133	5602	23735	PayPal	Yes	0	Web	Confirmed	04-10-2024	15.4335341	45.45202621	Yes
32	375	5	Boston	3	01-04-2024	01-05-2024	02-05-2024	Deluxe	1	Business	Web	12494	12494	3860	16354	Credit Card	Yes	0	Web	Confirmed	30-08-2024	0	9.524903328	Yes

# Insights and Visualizations

## Booking pattern across different channels

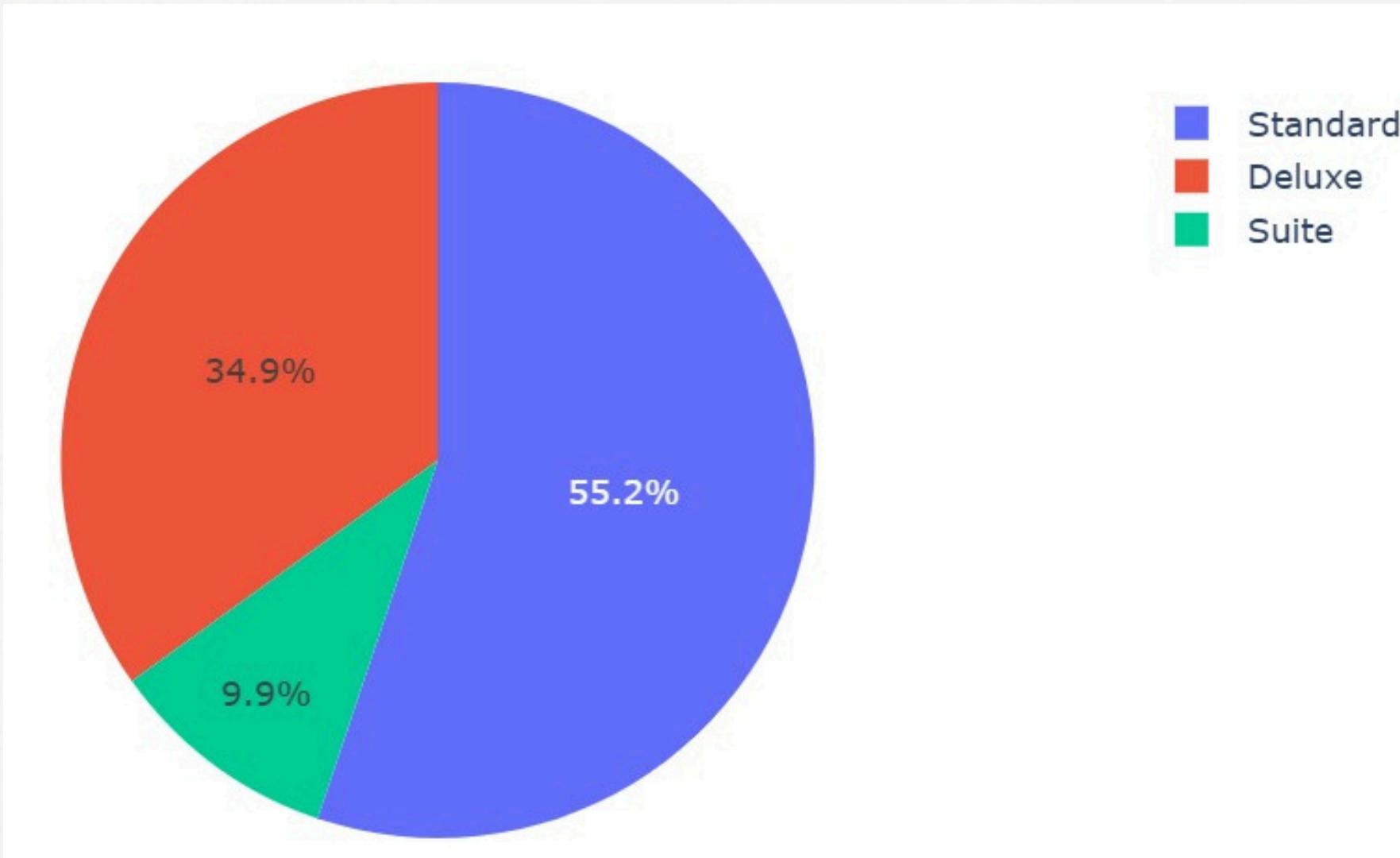
Booking Patterns across Channels



- Direct digital channels dominate (90%): Web (50%) and Mobile App (40%).
- Travel Agents low share (10%), least preferred option.
- Web is the most used channel, followed by Android; iOS shows the lowest usage.
- Highlights a Web-first preference and stronger Android adoption over iOS.

# Insights and Visualizations

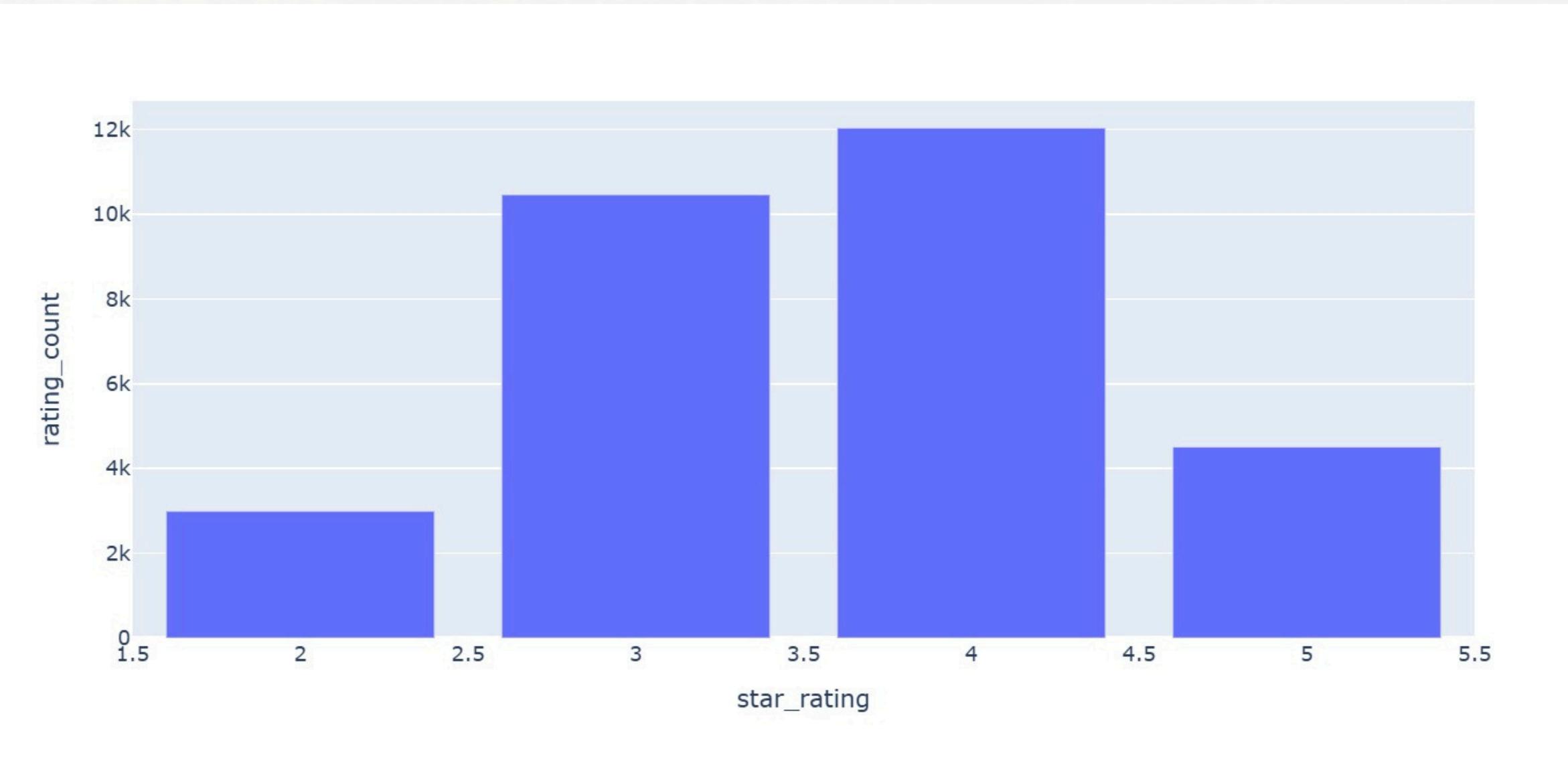
## Booking pattern across different room type



Standard rooms account for the highest bookings (16,552), followed by Deluxe (10,478), while Suites record the least (2,970). This suggests Standard rooms are the most preferred due to affordability and wider appeal, Deluxe rooms attract customers seeking extra comfort, and Suites cater to a smaller niche focused on luxury or special occasions.

# Insights and Visualizations

## Booking pattern across star ratings



Most customers rated their stays 3 or 4 (10,460 and 12,034), showing experiences were largely satisfactory to good. Fewer guests gave perfect 5 ratings (4,511), while low ratings of 2 (2,995) were least common, indicating overall customer sentiment leans average to positive.

# Insights and Visualizations

## Cancellation Behaviour

### *Cancellations by Room Type*

room_type	
Standard	63.542010
Deluxe	27.660626
Suite	8.797364

### *Cancellations by Booking channel*

booking_channel	
Web	43.591433
Mobile App	42.652389
Travel Agent	13.756178

### *Cancellations by Payment method*

payment_method	
PayPal	25.996705
Debit Card	25.485997
Credit Card	25.387150
Bank Transfer	23.130148

### *Effect of Coupons on Cancellation*

Coupon Used?	
No	78.78089
Yes	21.21911

### *Cancellation by Star Rating*

star_rating	
4	39.637562
3	34.810544
5	15.799012
2	9.752883

### *Cancellation vs. Refund*

refund_status	
Yes	77.940692
No	22.059308

- Overall cancellations: 20% of bookings.
- Standard rooms most affected (63%)
- Channel-wise: Web (43.6%), Mobile App (42.7%), Travel Agents lower (13.8%)
- Payment method impact minimal.
- No coupons = higher cancellations (79%) vs. With coupons (21%).
- 78% cancellations linked to refunds, encouraging casual behavior

Cancellations are driven more by pricing and refund policies than by channels or payment methods

# Root Cause Analysis

What might be the reasons for the observed cancellation patterns?

Cancellations may be happening due to flexible refund policies, customers booking multiple options and then dropping some, or changes in travel plans. In addition, cashback, and coupons may encourage tentative bookings that are not always completed.

Further analysis was done to check this...

# Root Cause Analysis

## Multiple Bookings for the Same Travel Date

customer_id	travel_date	num_bookings
8	1 2024-04-07	2
10	1 2024-04-21	2
18	1 2024-06-02	3
19	1 2024-06-05	2
27	1 2024-06-26	2
...	...	...
26083	499 2024-05-15	2
26086	499 2024-05-23	2
26101	499 2024-08-31	2
26107	499 2024-09-14	2
26109	499 2024-09-23	2

3400 rows × 3 columns

Customers making 2-3 bookings for the same travel date appear frequently.

This shows duplicate or comparison booking behavior (booking multiple rooms/hotels, then cancelling later).

# Root Cause Analysis

## Multi-booker cancellation behavior

Multi-booker cancellation behavior:

	customer_id	total_bookings	total_cancellations	cancellation_rate
114	115	48	20	41.666667
207	208	20	8	40.000000
12	13	52	18	34.615385
292	293	58	20	34.482759
115	116	80	27	33.750000
84	85	69	23	33.333333
498	499	57	19	33.333333
424	425	58	19	32.758621
364	365	41	13	31.707317
112	113	51	16	31.372549

Some customers cancel frequently, but none cross the “extreme” 60% cancellation threshold. Still, I found a set of multi-bookers with cancellation rates around 30-40%, which is relatively high and can hurt business revenue.

# Root Cause Analysis

## Channels most affected by cancellations

Channels most affected by cancellations:		
	booking_channel	cancellations
0	Mobile App	2589
1	Travel Agent	835
2	Web	2646

In terms of channels, the Mobile App and Web show the highest number of cancellations (2,589 and 2,646 respectively), while Travel Agents account for fewer cancellations (835). This shows that online, self-service platforms are more prone to cancellation behavior compared to mediated bookings.

# Root Cause Analysis

## Cancellation rate with vs without coupons



The cancellation rate for bookings with a coupon (20.8%) is almost the same as those without a coupon (20.1%). This shows that coupons are not a major driver of cancellations; people are cancelling regardless of coupon use.

# Root Cause Analysis

## Cashback impact

booking_status	Cancelled	Confirmed	Failed
<b>cashback</b>			
0.000000	0.198137	0.727142	0.074721
0.507662	0.500000	0.500000	0.000000
0.537409	0.000000	1.000000	0.000000
0.538850	0.000000	0.500000	0.500000
0.555062	0.500000	0.500000	0.000000
...	...	...	...
20.552876	0.000000	1.000000	0.000000
21.124024	0.000000	1.000000	0.000000
21.133242	0.000000	1.000000	0.000000
21.161835	0.250000	0.500000	0.250000
21.268013	0.000000	0.750000	0.250000
[7798 rows x 3 columns]			

For cashback, cancellation rates vary widely across different cashback values. At 0 cashback, cancellation is around 19.8%, but at some cashback levels, it jumps to 50%. This suggests that certain cashback offers may encourage trial bookings that don't convert.

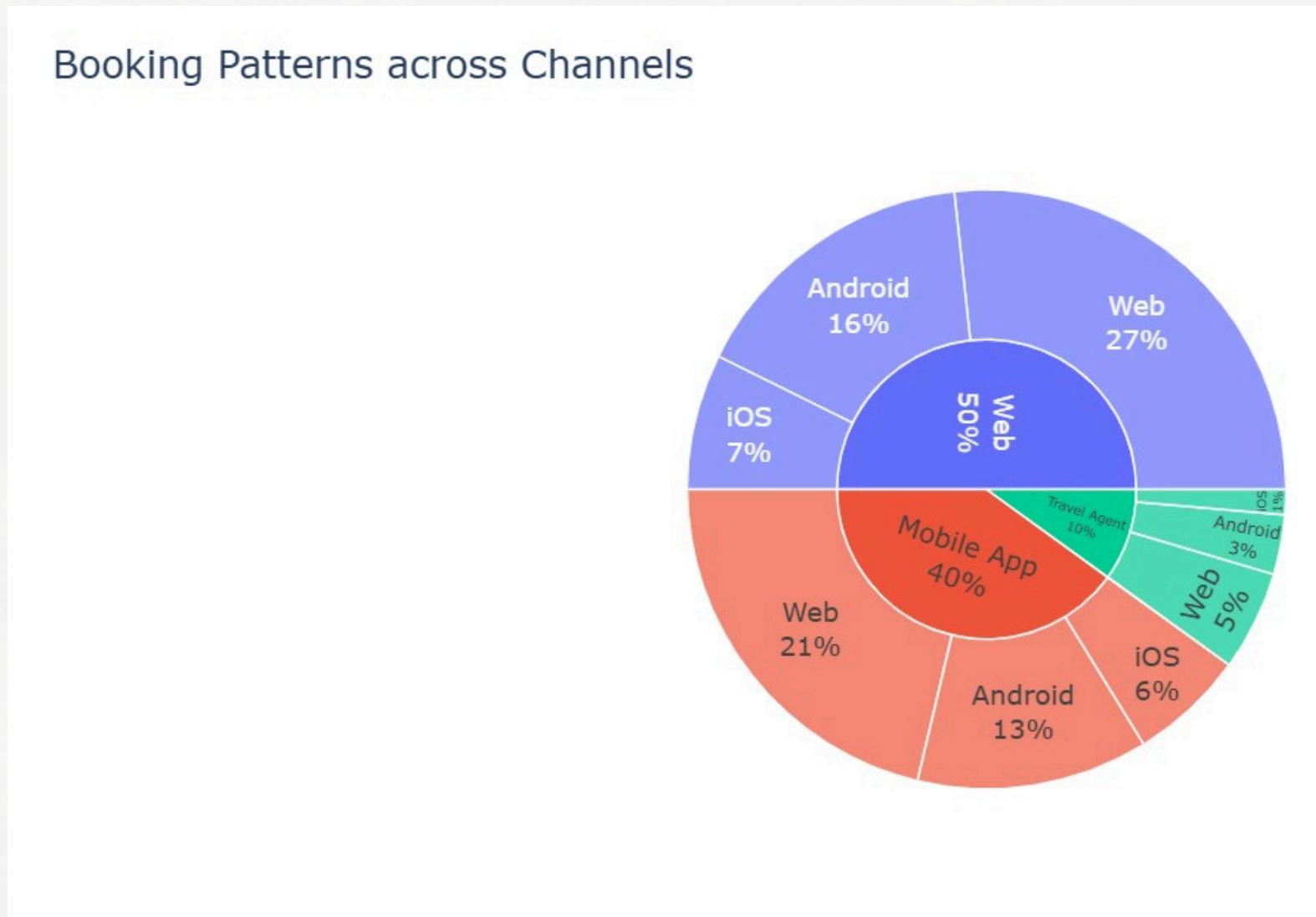
# Root Cause Analysis

## Conclusion

- Standard room customers are more price-sensitive, leading to higher cancellations.
- Refund policies encourage casual cancellations; coupons reduce them by increasing commitment.
- Cashback offers attract more non-serious customers.
- Easy cancellations via app/website add to rates, while travel agent bookings face fewer cancellations due to stricter policies and advance payments.
- Repeat customers often book multiple options and cancel later.
- Overall, cancellations are influenced more by pricing, customer intent, and flexibility than by channel or payment method.

# Root Cause Analysis

Why do some booking channels or property types perform better than others?

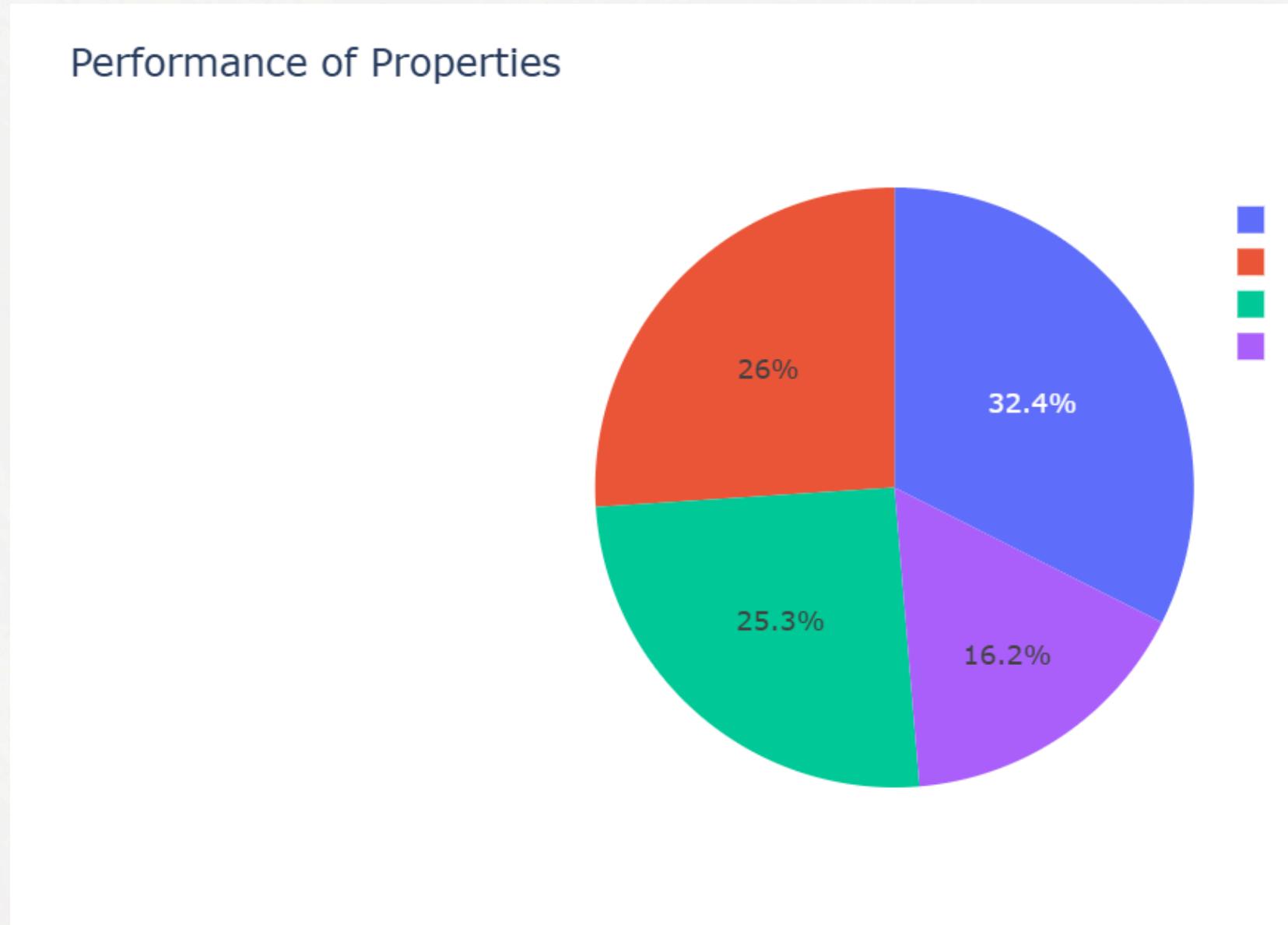


- Easier navigation, clearer information, and more familiar experience on Web.
- Websites trusted more for transparent pricing & detailed reviews.
- Apps may have limited features, glitches, or slower performance.
- Desktop/laptop habits naturally favor website usage.
- Higher Android usage reflects larger user base vs iOS.
- Skepticism about travel agents leads customers to prefer direct booking via Web/App.

Web is the most trusted and convenient booking channel, making it the key focus for customer engagement.

# Root Cause Analysis

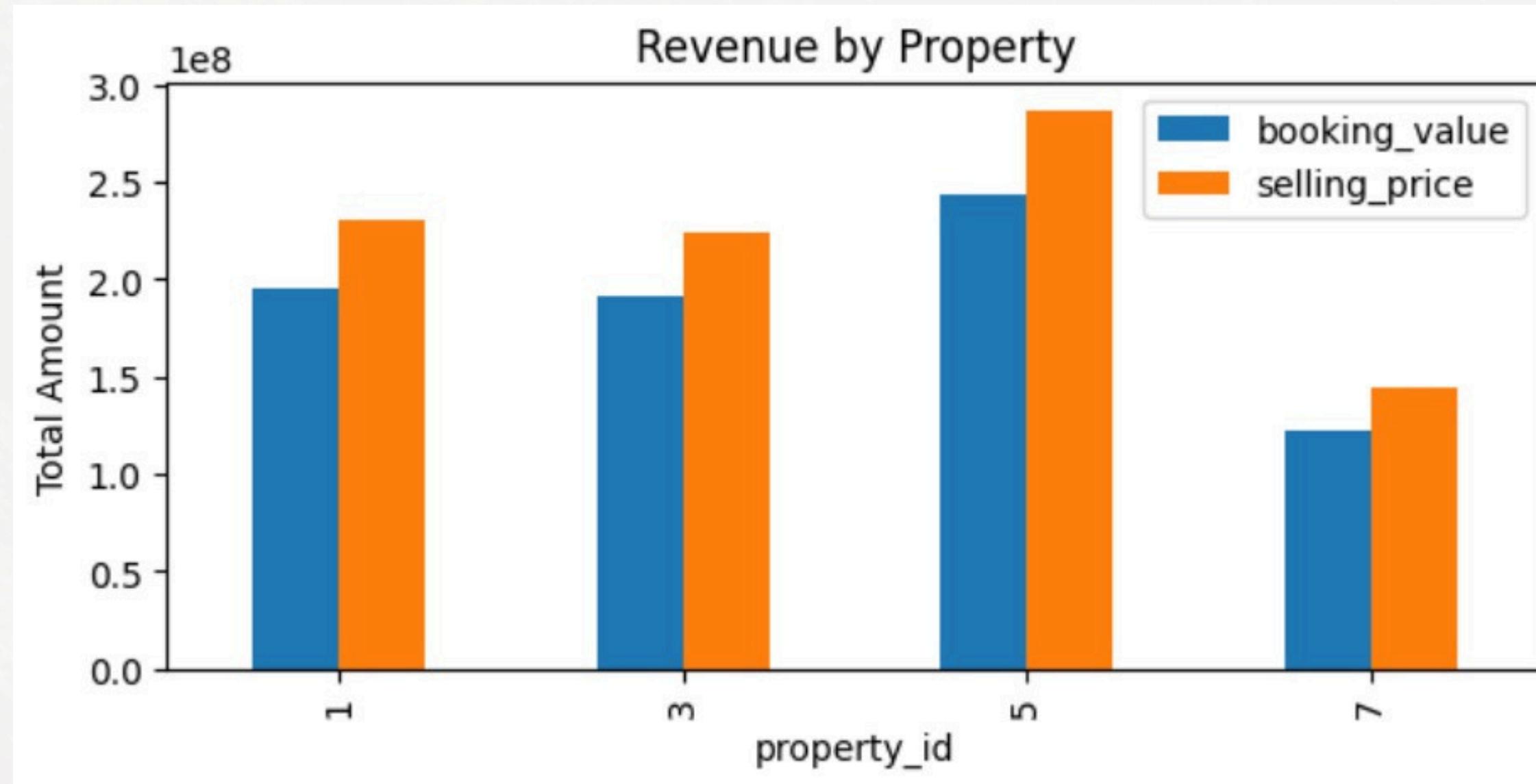
## Booking volume per property



Property 5 has the highest number of bookings and generates the maximum revenue, followed by Properties 1 and 3, while Property 7 consistently underperforms.

# Root Cause Analysis

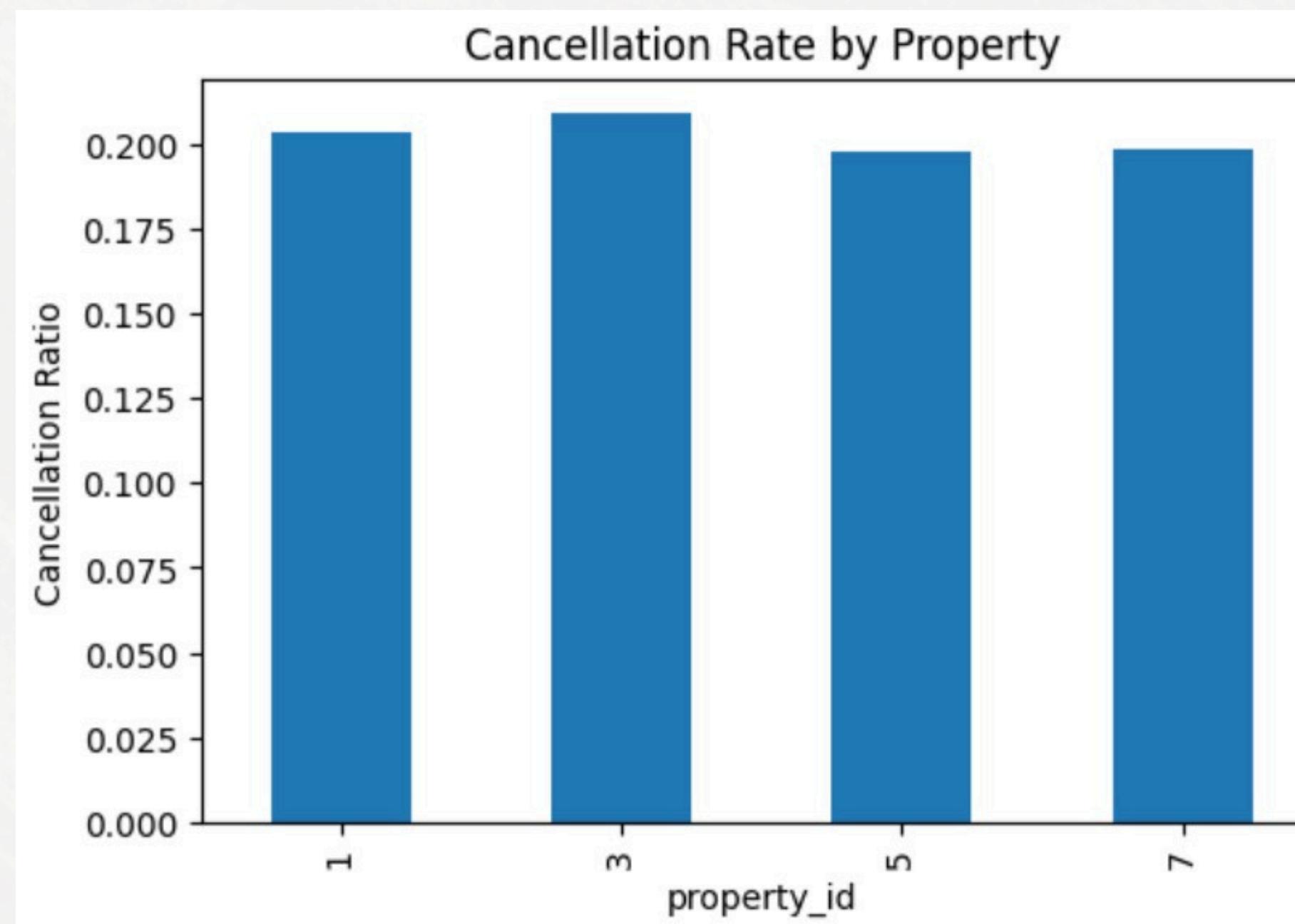
Revenue and profitability of each property type



Property 5 not only drives the most revenue but also has the highest markup contribution, showing strong profitability.

# Root Cause Analysis

## Cancellation & refund impact



All four properties (1, 3, 5, 7) show a similar cancellation rate of around 20%, with Property 3 being slightly higher. This means cancellations occur at almost the same frequency across properties.

# Root Cause Analysis

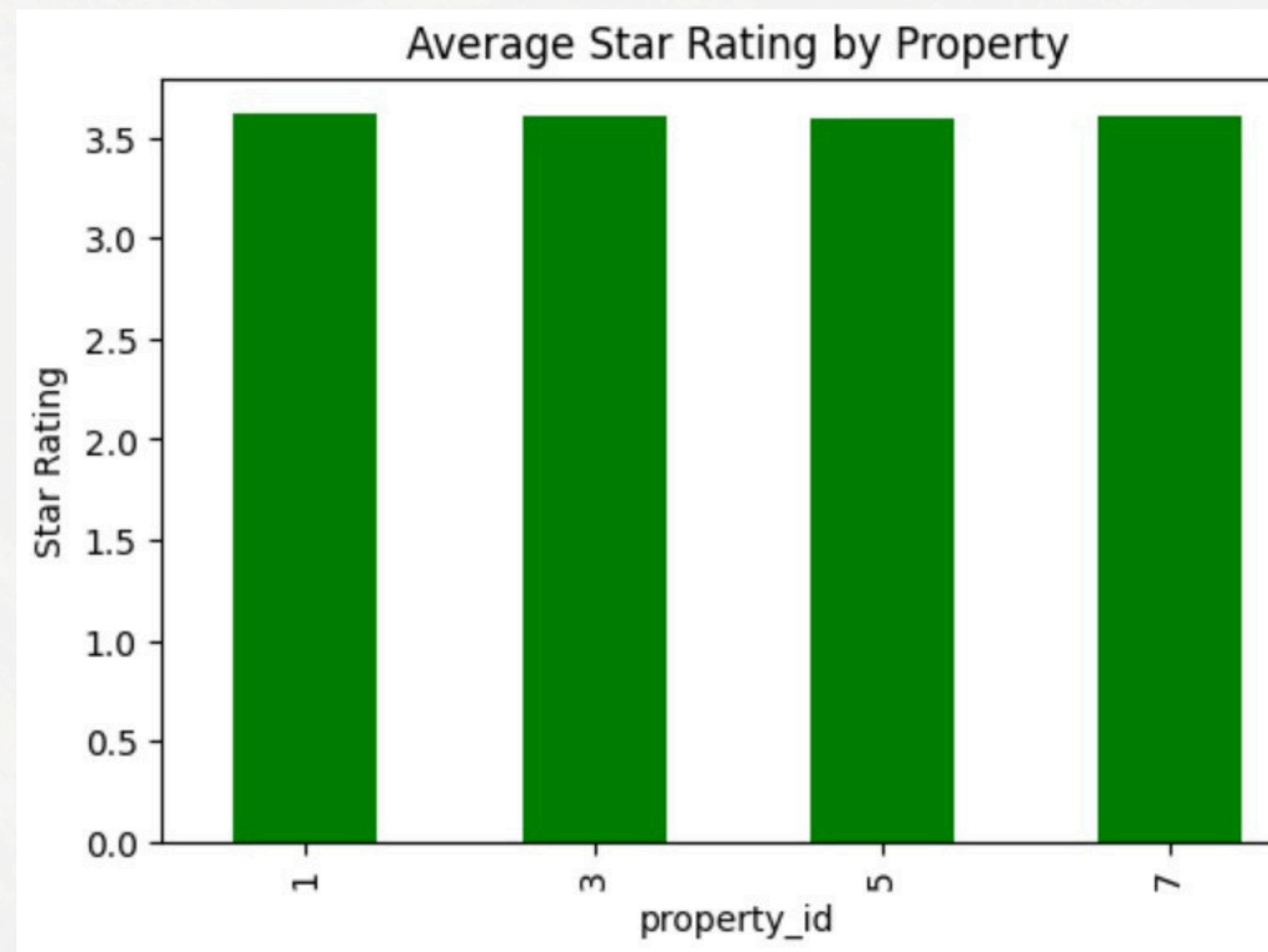
## Cancellation & refund impact



- Property 5 has the highest refund amount, suggesting it handles higher-value bookings that lead to costlier cancellations.
- Properties 1 and 3 also have substantial refund amounts, though lower than Property 5.
- Property 7 has the lowest refund amount, implying its bookings are either fewer in value or lower-priced.

# Root Cause Analysis

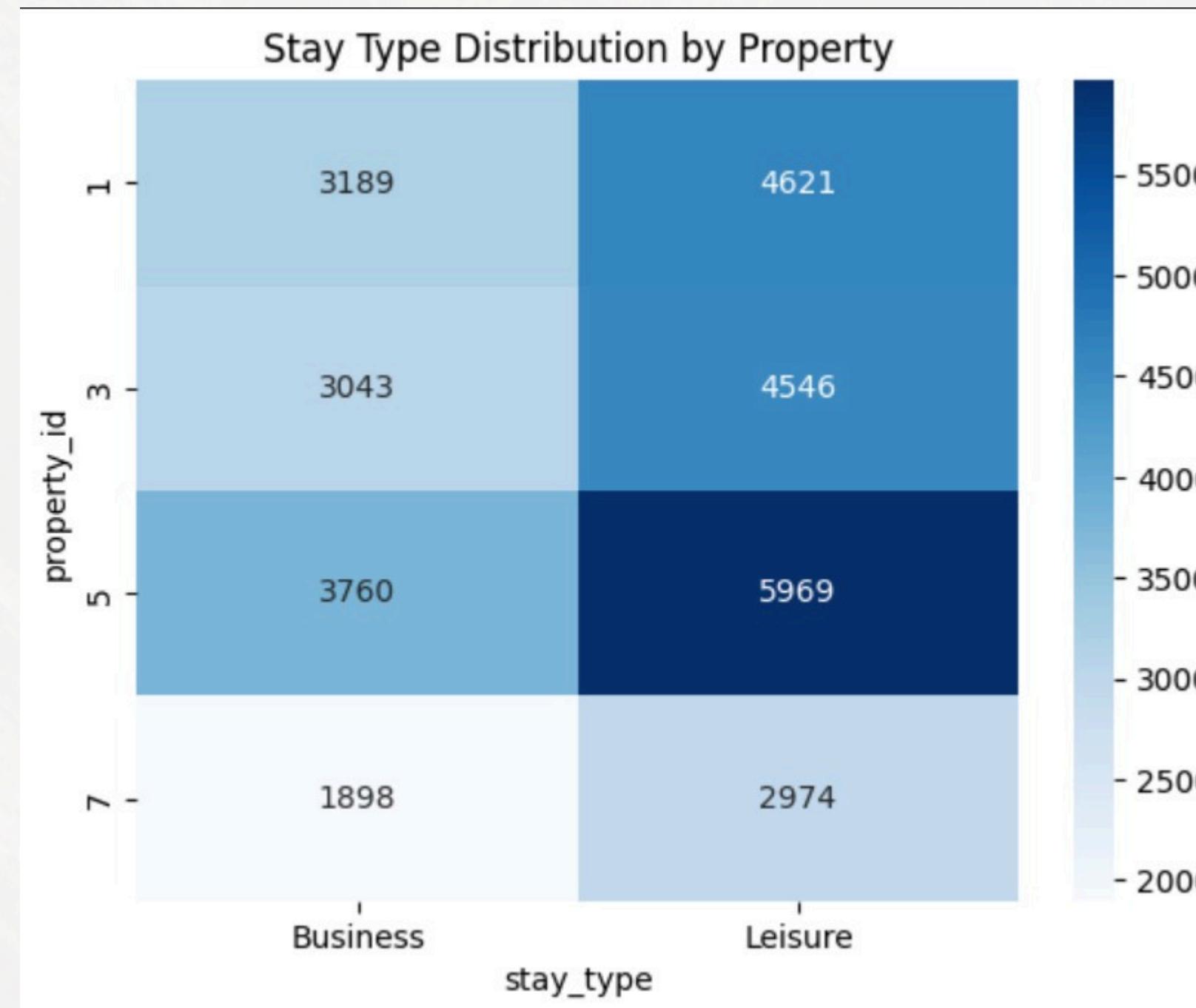
## Star rating performance



Star ratings are fairly stable (between 3.58-3.61) across all properties, suggesting customer experience is relatively uniform.

# Root Cause Analysis

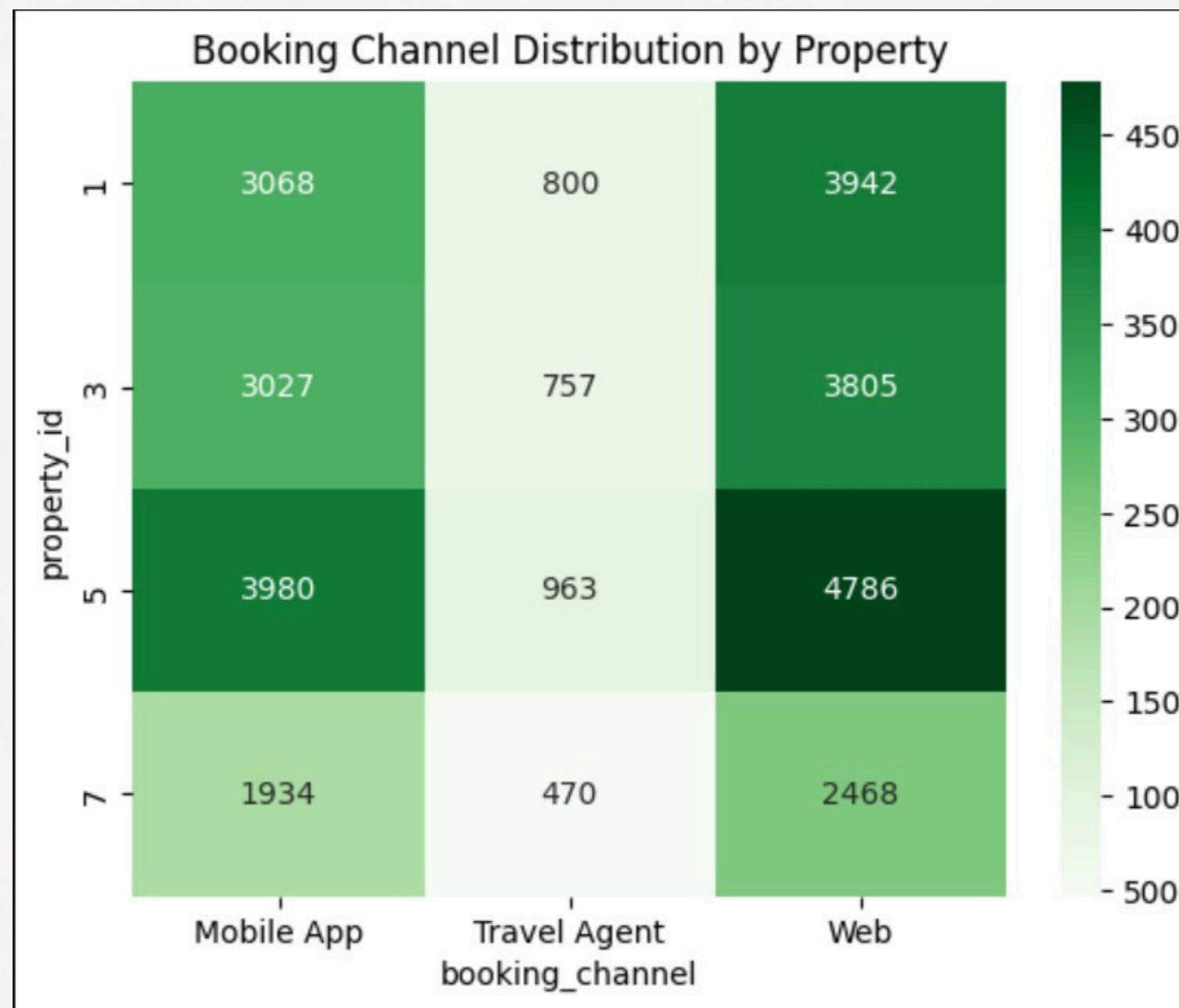
## Stay\_type distribution



Property 5 dominates in both business and leisure stays, making it attractive to diverse customer segments, while Property 7 is weaker in both segments

# Root Cause Analysis

## Booking channel distribution



Property 5 captures the largest share across web and mobile app channels, which likely boosts its overall performance. Property 7 is again the weakest across all booking channels.

# Root Cause Analysis

## Possible Reasons Why Some Properties Perform Better

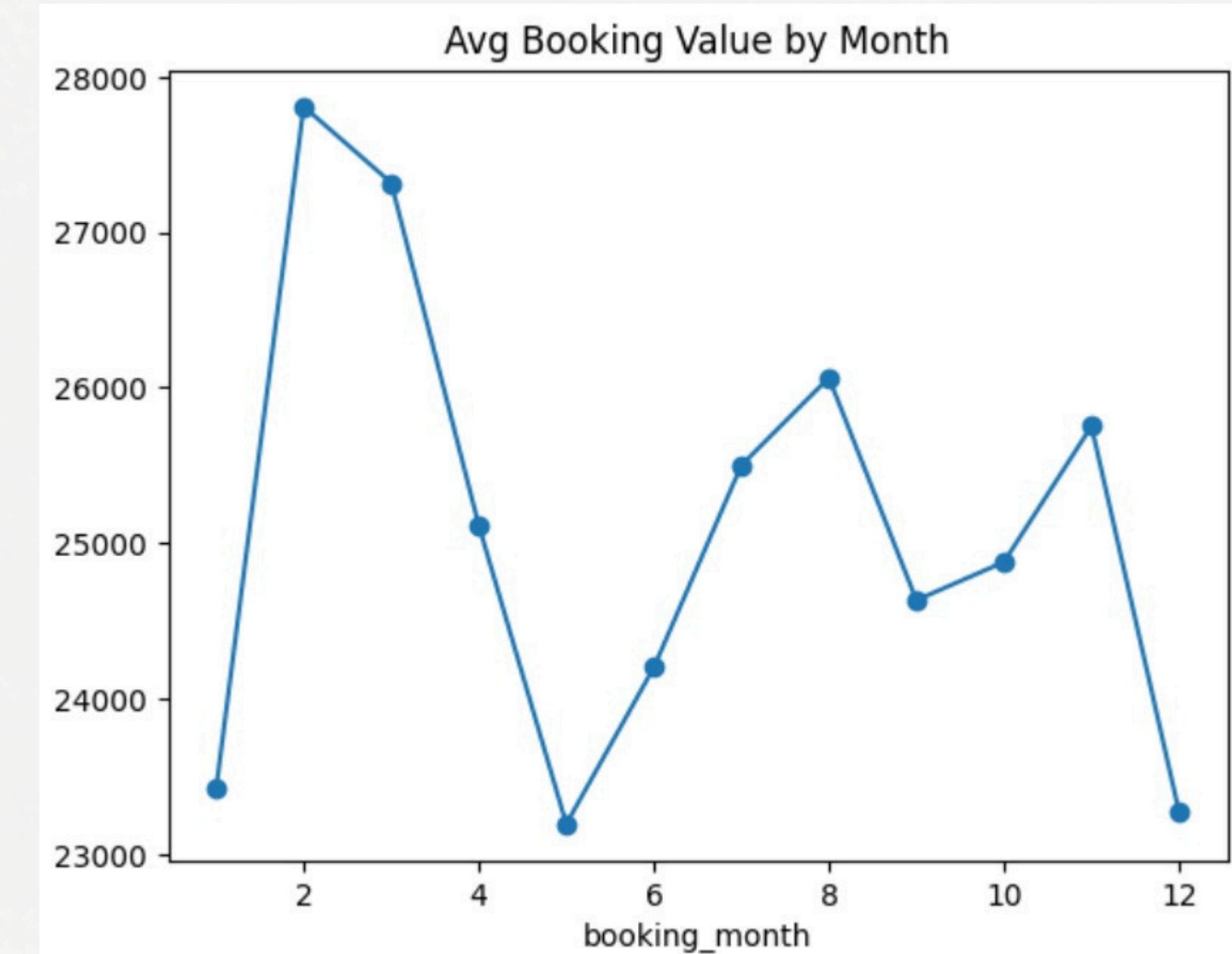
- Property 5 performs strongly across both business and leisure stays, unlike Property 7, which has weaker appeal in both segments.
- Stronger distribution across web and mobile channels gives Properties 5, 1, and 3 better visibility and access to customers.
- Property 5 shows effective pricing and markup strategies, driving higher profitability without significantly higher cancellation rates.
- Properties with better connectivity or central locations (possibly 1, 3, 5) tend to outperform (can be checked further).
- Higher refunds in Property 1 and 3 suggest operational inefficiencies or customer dissatisfaction compared to Property 5.

# Root Cause Analysis

Are there seasonal or temporal trends influencing booking values or stay lengths?

## Booking Value by Month

- Highest booking values are seen in February (~27,806) and March (~27,316).
- The lowest booking values appear in December (~23,276) and May (~23,197).
- This indicates peak demand in late winter and early spring, with a dip towards year-end and early summer.

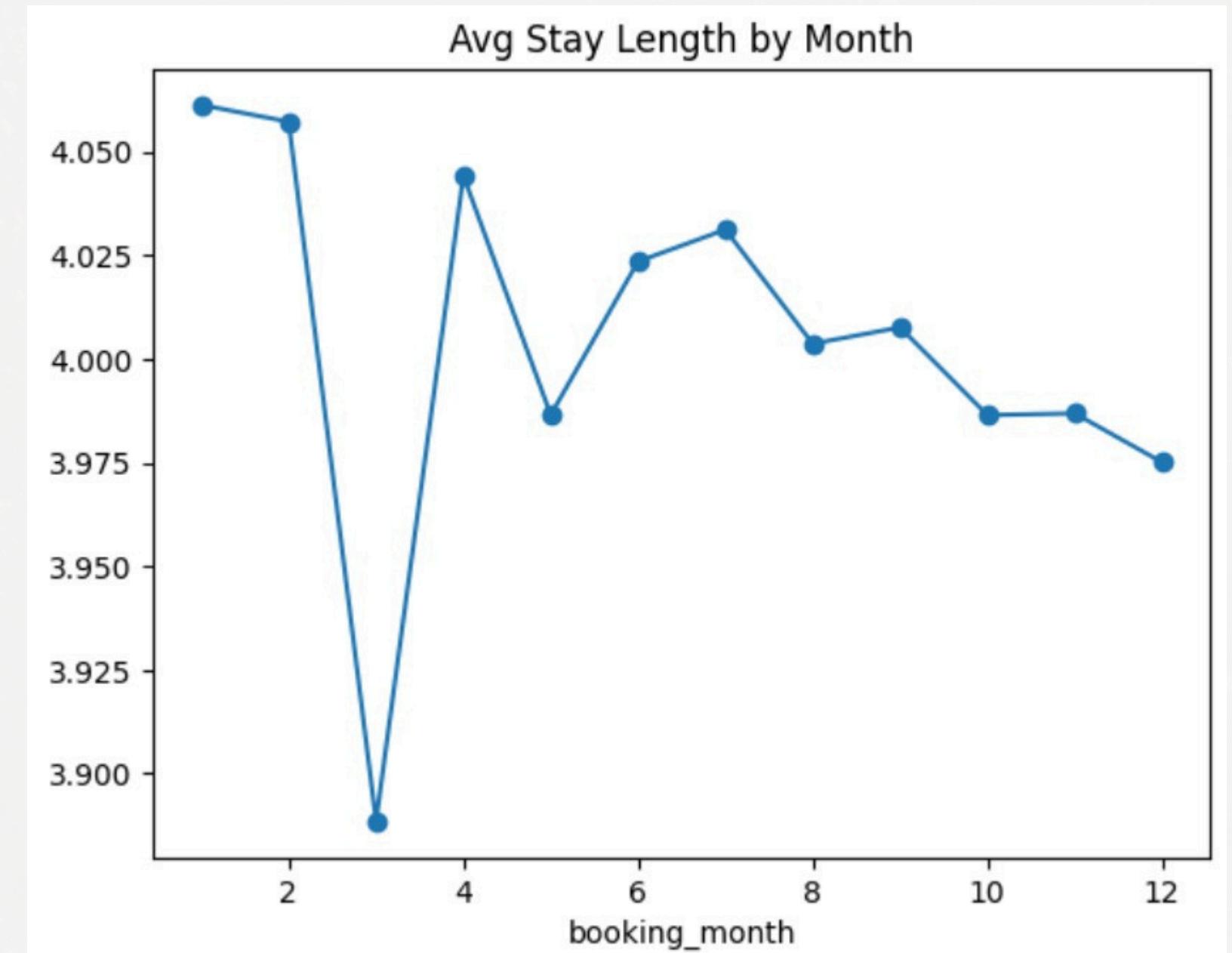


# Root Cause Analysis

Are there seasonal or temporal trends influencing booking values or stay lengths?

Average stay length by month

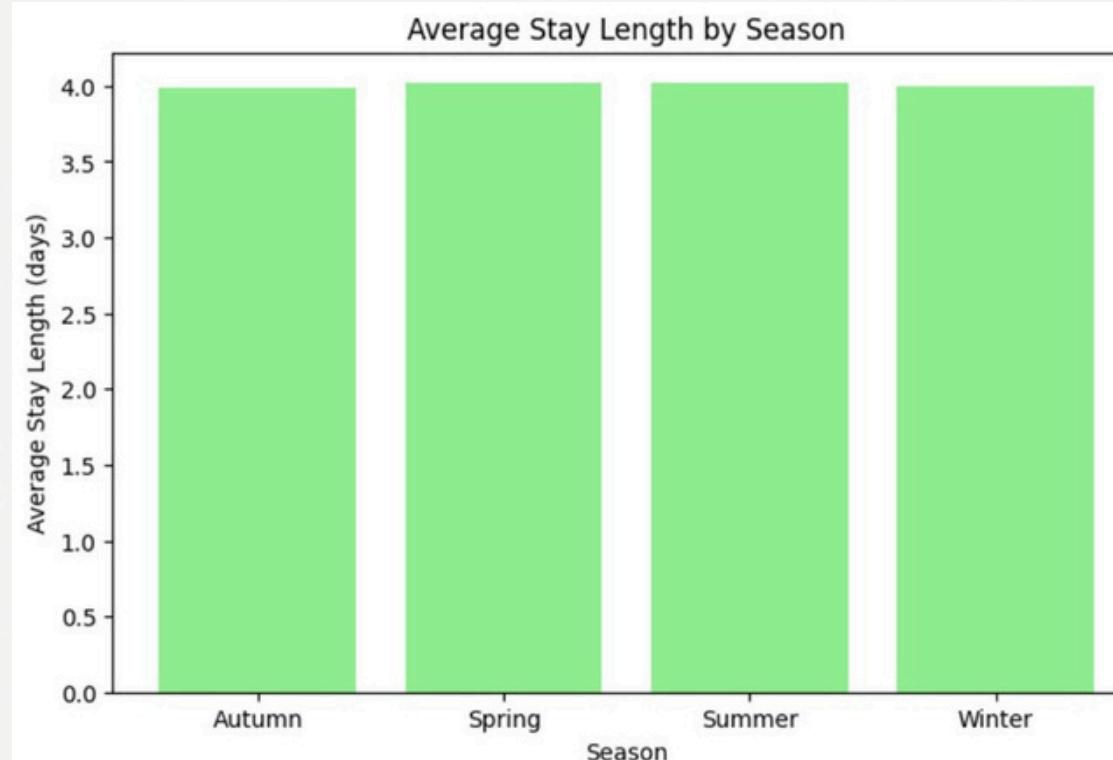
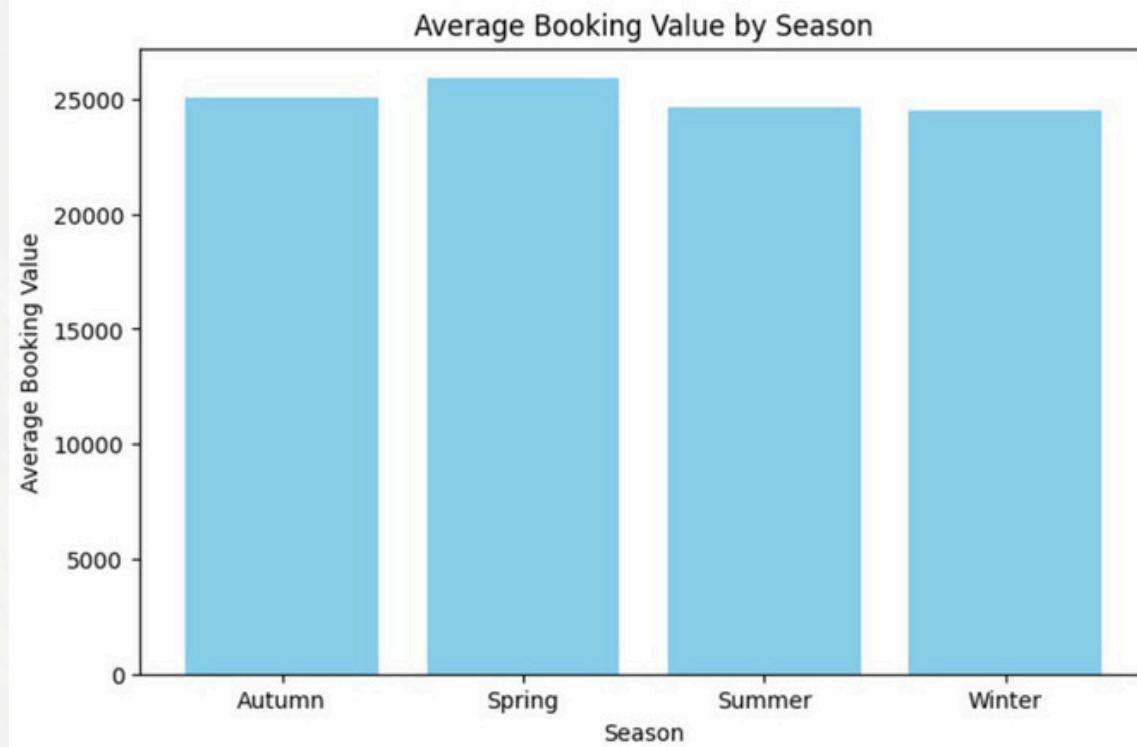
- Average stay length remains fairly consistent (around 3.9-4.1 days).
- Slightly longer stays occur in April (4.04 days) and Thursday check-ins (4.07 days).
- Shorter stays are seen in March (3.88 days) and Fridays (3.94 days).



# Root Cause Analysis

Are there seasonal or temporal trends influencing booking values or stay lengths?

Average Booking Value by Season



- Spring has the highest booking value (~25,911) with an average stay length of 4.02 days.
- Winter records the lowest booking value (~24,499), though stay length (~4.0 days) is consistent.
- Seasonal differences in stay length are marginal, but booking values vary by nearly ₹1,500-1,800 between high and low seasons

# Root Cause Analysis

Are there seasonal or temporal trends influencing booking values or stay lengths?

## Day-of-Week Patterns

check_in_day	booking_value	stay_length
Friday	25113.914726	3.946959
Monday	25173.290968	3.991691
Saturday	25002.874219	4.055571
Sunday	25045.322666	4.000296
Thursday	24956.448471	4.071489
Tuesday	25164.392726	3.988701
Wednesday	25408.240237	3.994581

- Highest booking values are on Wednesdays (~25,408) and Mondays (~25,173).
- Lowest values are on Thursdays (~24,956).
- Guests checking in on Thursdays and Saturdays tend to stay longer (4.07 and 4.05 days) compared to those arriving on Fridays (3.94 days).

# Root Cause Analysis

## Conclusion

**Yes**, there are seasonal and temporal trends influencing booking values and stay lengths: Booking values peak in late winter and spring (Feb-Mar, Spring season), while they are lowest in December and May (Winter and early summer).

Stay lengths are relatively stable across months and seasons (~4 days), with minor variations depending on day of check-in (slightly longer on Thursdays/Saturdays).

Day-of-week trends show that mid-week (Wednesdays) attracts higher-value bookings, while weekend stays are slightly longer but lower in booking value.

# Business Recommendations

## Strategies to reduce cancellations

- Revise refund policies: partial/tiered refunds instead of full refunds.
- Promote coupon-based bookings to increase commitment.
- Introduce discounted non-refundable rates for standard rooms.
- Offer flexible upgrades to shift demand to higher-value rooms.
- Add rescheduling options on Web/App to reduce drop-offs.
- Incentivize travel agent bookings (lower cancellation tendency).
- Use loyalty programs & reminders to engage repeat customers.

# Business Recommendations

## Ways to improve profitability and increase repeat bookings

- Promote Deluxe rooms with bundled add-ons (breakfast, transport) since they already show strong demand.
- Encourage Suite bookings through targeted offers for luxury travelers/special occasions to tap into niche profitability.
- Use coupon-based loyalty programs to reduce cancellations and increase repeat bookings.
- Provide exclusive discounts for repeat customers who book directly via Web/App, lowering dependency on travel agents.
- Focus on personalized offers for Android users, as they represent the largest booking base.
- Incentivize customers who frequently give 3-4 star ratings with upgrades/discounts to convert them into highly satisfied repeat guests.

# Business Recommendations

## Opportunities to optimize pricing, promotions, or channel strategy.

- Maintain competitive pricing for Standard rooms but add upsell nudges toward Deluxe to improve margins.
- Launch seasonal pricing strategies (higher in peak months, discounts in off-season) to balance demand.
- Offer coupon-based promotions over cashback, since coupons reduce cancellations and increase commitment.
- Strengthen direct booking channels (Web/App) with better UI/UX, exclusive deals, and transparent pricing to reduce travel agent dependency.
- Provide early-bird discounts with stricter refund rules to lock in serious bookings.
- Design personalized promotions for frequent Android users, given their larger booking share.



*Thank you*

