

TUGAS AKHIR DATA SCIENCE PADA DOMAIN SPESIFIK

STUDI KASUS PADA DATASET HOTEL



KELOMPOK: 03

DISUSUN OLEH:

Erwin Darsono / 6182001002

Neil Christopher / 6182001010

Wilson / 6182001039

KELAS: A

Dosen Kelas: Lydia Mutiara Dewi, MA

UNIVERSITAS KATOLIK PARAHYANGAN

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI DAN SAINS

BANDUNG

2024

I. LATAR BELAKANG

Pariwisata merupakan aktivitas yang melibatkan berbagai aspek seperti transportasi, akomodasi, dan konsumsi yang dilakukan oleh masyarakat baik secara individu maupun berkelompok dengan melakukan perjalanan jangka pendek ke suatu tempat di luar tempat tinggal dan tempat bekerja mereka dengan tujuan memenuhi kebutuhan kesejahteraan atau kesehatan fisik maupun mental.

a. Peran Pariwisata secara umum

- **Penggerak Ekonomi**

Pariwisata merupakan salah satu sektor yang memberikan sumber pendapatan terbesar baik bagi negara maupun daerah. Sektor pariwisata menghasilkan sumber pendapatan melalui pajak para wisatawan dan pengeluaran wisatawan seperti transportasi, akomodasi, dan konsumsi. Sumber pendapatan ini yang kemudian berkontribusi untuk meningkatkan penerimaan devisa.

- **Pembangunan Infrastruktur**

Infrastruktur suatu daerah didorong melalui investasi dari pemerintah, swasta, maupun asing untuk meningkatkan pariwisata dalam negeri dengan membangun berbagai infrastruktur seperti bandara, jalan raya, transportasi umum maupun fasilitas akomodasi lainnya seperti hotel dan restoran yang bertujuan untuk menarik wisatawan serta meningkatkan kemampuan negara untuk menyambut tamu.

- **Pemajuan Budaya**

Budaya lokal dapat diperkenalkan dan dipromosikan melalui sektor pariwisata. Sehingga melalui sektor pariwisata, budaya daerah dapat diperkenalkan ke mancanegara untuk memperoleh wawasan ataupun nilai-nilai yang terkandung dalam budaya lokal. Salah satu tujuan diperkenalkan budaya lokal ke mancanegara adalah untuk melestarikan budaya-budaya lokal.

- **Diplomasi antar negara atau bangsa**

Melalui sektor pariwisata, negara-negara juga dapat menjalin hubungan antarbangsa dan pertukaran budaya. Hal ini dilakukan melalui pertemuan dengan wisatawan asing sehingga persahabatan antar negara akan lebih erat dan mempermudah administrasi di sektor lain yang dapat memicu kunjungan wisatawan secara lebih intens.

b. Karakteristik Pariwisata Secara Umum

- Sektor yang membutuhkan koordinasi intensif dari berbagai produk/jasa lain seperti transportasi, akomodasi, dll.
- Pariwisata merupakan sektor yang rapuh (*perishable*): yang berarti tidak dapat disimpan untuk dikonsumsi di kemudian hari.
- Sering menghadapi ketidakpastian permintaan yang tinggi (*demand uncertainty*) dan dinamika yang lebih kompleks (*complex dynamics*) karena persaingan yang intensif di antara penyedia layanan.
- Sektor pariwisata sangat rentan terhadap kejadian-kejadian yang global, seperti contohnya wabah penyakit COVID - 19.

c. Definisi Hotel

Hotel merupakan suatu bangunan yang dikelola oleh perusahaan atau badan usaha akomodasi yang bertujuan untuk memfasilitasi sektor akomodasi, konsumsi, atau bahkan transportasi seperti jasa pelayanan penginapan, penyedia makanan, dan jasa antar jemput bagi tamu yang berkunjung.

d. Klasifikasi Hotel

1. Penentuan Standar Hotel menurut Ukuran (Size) Hotel.

- **Hotel Besar**
Hotel yang memiliki 300 kamar atau lebih.
- **Hotel menengah**
Hotel yang memiliki 100 - 299 kamar.
- **Hotel Kecil**
Hotel yang memiliki 25 - 99 kamar.

2. Penentuan Hotel menurut Operasinya

- ***Around the year operation hotel***
Hotel yang beroperasi sepanjang tahun.
- ***Seasonal Hotel***
Hotel yang beroperasi pada musim-musim tertentu.

3. Penentuan Hotel menurut Lokasinya

- ***City Hotel***
Hotel yang berlokasi di daerah perkotaan.
- ***Residential Hotel***
Hotel yang berlokasi di pinggiran yang cukup jauh dari keramaian.
- ***Resort Hotel***
Hotel yang berlokasi di daerah wisata seperti pegunungan (*mountain hotel*) atau di tepi pantai (*beach hotel*).
- ***Motel***
Hotel yang berlokasi di pinggir atau di sepanjang jalan raya yang menghubungkan satu kota dengan kota lainnya.

4. Penentuan Hotel menurut Bintangnya

Klasifikasi Hotel Bintang	Jumlah Unit Kamar	Luas Standar Minimum Kamar
Bintang 1	<ul style="list-style-type: none">• Unit standard minimum 15 kamar• Kamar mandi di dalam	<ul style="list-style-type: none">• Luas unit standard minimum $20m^2$
Bintang 2	<ul style="list-style-type: none">• Unit standard minimum 20 kamar• Unit <i>suite</i> minimum 2 kamar• Kamar mandi di dalam	<ul style="list-style-type: none">• Luas unit standard minimum $22m^2$• Luas unit <i>suite</i> minimum $44m^2$
Bintang 3	<ul style="list-style-type: none">• Unit standard minimum 30 kamar• Unit <i>suite</i> minimum 2 kamar• Kamar mandi di dalam	<ul style="list-style-type: none">• Luas unit standard minimum $24m^2$• Luas unit <i>suite</i> minimum $48m^2$
Bintang 4	<ul style="list-style-type: none">• Unit standard minimum 50 kamar• Unit <i>suite</i> minimum 3 kamar• Kamar mandi di dalam	<ul style="list-style-type: none">• Luas unit standard minimum $24m^2$• Luas unit <i>suite</i> minimum $48m^2$
Bintang 5	<ul style="list-style-type: none">• Unit standard minimum 100 kamar• Unit <i>suite</i> minimum 4 kamar• Kamar mandi di dalam	<ul style="list-style-type: none">• Luas standard minimum $26m^2$• Luas minimum <i>suite</i> $52m^2$

II. DATASET Hotel Reservation Dataset

Eksplorasi umum dataset

a. Narasi singkat mengenai dataset

Reservasi hotel secara *online* telah mengubah kemungkinan pemesanan dan perilaku pelanggan. Sejumlah besar reservasi hotel dibatalkan karena pembatalan atau ketidakhadiran tamu. Alasan umum pembatalan termasuk perubahan rencana, konflik jadwal, dan sebagainya. Hal ini seringkali lebih mudah dilakukan karena adanya opsi untuk membatalkan secara gratis atau dengan biaya rendah yang menguntungkan tamu hotel tetapi kurang diinginkan dan mungkin mengurangi pendapatan bagi hotel.

b. Link sumber dataset: [Hotel Reservation Dataset](#)

c. Jumlah kolom: 20 Kolom

d. Jumlah baris: 36.275 Baris

e. Identifikasi null values dan Outliers

Nama Kolom	Jenis Data	Isi Data	Null Values	Banyak null values	Outliers	Nilai Kisaran
No	Numerical	Nomor 1-36275 untuk menandakan urutan rekord data	Tidak ada	Tidak ada	Tidak ada	Tidak ada
Booking_ID	Text	Identifier untuk setiap booking	Tidak ada	Tidak ada	Tidak ada	Tidak ada
no_of_adults	Numerical	Jumlah orang dewasa	Tidak ada	Tidak ada	Ada	0-1 dan 3-4
no_of_children	Numerical	Jumlah anak-anak	Tidak ada	Tidak ada	Ada	1-10
no_of_weekend_nights	Numerical	Banyak malam akhir pekan yang direservasi oleh pengunjung	Tidak ada	Tidak ada	Ada	0-7
no_of_week_nights	Numerical	Banyak malam hari biasa yang direservasi oleh pengunjung	Tidak ada	Tidak ada	Ada	0-17
type_of_meal_plan	Text	Jenis paket makanan yang dipesan saat reservasi.	Tidak ada	Tidak ada	Tidak ada	Tidak ada
required_car_parking_space	Numerical	Jumlah tempat parkir yang dibutuhkan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
room_type_reserved	Text	Jenis kamar yang ditempati	Tidak ada	Tidak ada	Tidak ada	Tidak ada
lead_time	Numerical	Jarak antar hari pemesanan dengan hari penginapan	Tidak ada	Tidak ada	Ada	289.5 - 443
arrival_year	Numerical	Tahun penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
arrival_month	Numerical	Bulan penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
arrival_date	Numerical	Tanggal penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
market_segment_type	Text	Penggolongan segmentasi pasar	Tidak ada	Tidak ada	Tidak ada	Tidak ada
repeated_guest	Numerical	Tamu pengunjungan kembali	Tidak ada	Tidak ada	Tidak ada	Tidak ada
no_of_previous_cancellations	Numerical	Jumlah reservasi yang pernah dibatalkan.	Tidak ada	Tidak ada	Tidak ada	Tidak ada
no_of_previous_bookings_not_canceled	Numerical	Jumlah reservasi yang tidak dibatalkan.	Tidak ada	Tidak ada	Tidak ada	Tidak ada

d						
avg_price_per_room	Numerical	Harga rata-rata per kamar	Tidak ada	Tidak ada	Ada	0 - 20.74 dan 179.56 - 540
no_of_special_requests	Numerical	Jumlah permintaan khusus	Tidak ada	Tidak ada	Tidak ada	Tidak ada
booking_status	Text	Menandakan kamar yang dipeservasi dibatalkan atau tidak.	Tidak ada	Tidak ada	Tidak ada	Tidak ada

- f. Sample 5 baris data pertama dan 5 baris data terakhir untuk 6 kolom pertama saja dari Hotel Reservation Dataset

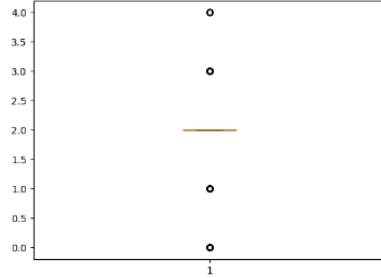
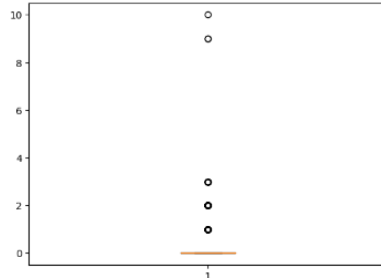
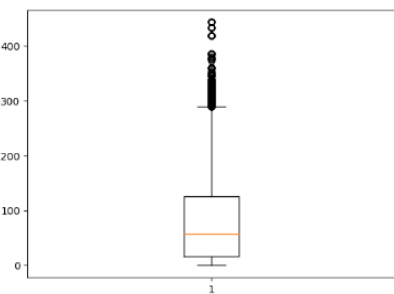
5 Baris Pertama

No	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights
36275	INN36013	2	1	2	2
36274	INN35965	2	0	0	1
36273	INN35924	2	1	2	2
36272	INN35878	2	0	0	3
36271	INN35682	1	0	0	1

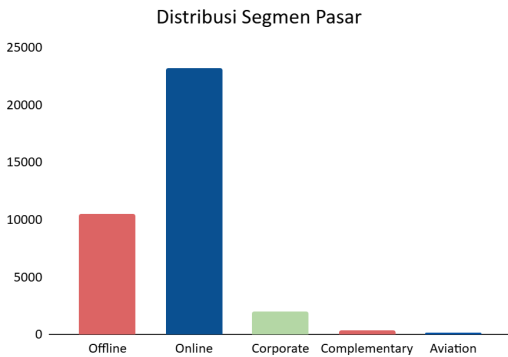
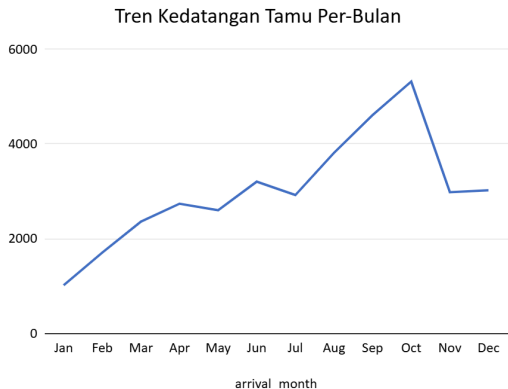
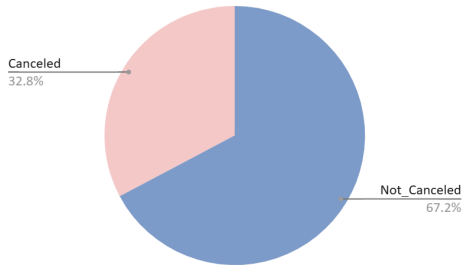
5 Baris Terakhir

No	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights
5	INN00146	1	0	0	2
4	INN00121	2	0	0	3
3	INN00081	2	2	0	2
2	INN00049	2	0	0	4
1	INN00041	2	0	0	4

g. Descriptive Analysis

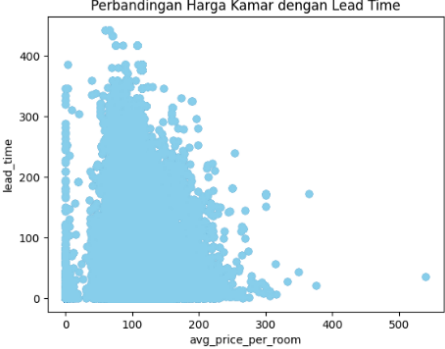
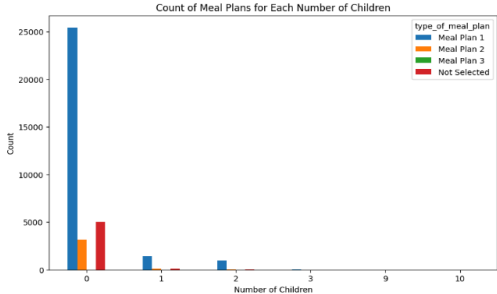
No	Kolom	Analisis	Interpretasi	Keterangan																								
1	no_of_adults	<table><tr><td>Mean</td><td>2</td></tr><tr><td>Median</td><td>2</td></tr><tr><td>Modus</td><td>2</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>2</td></tr><tr><td>Q2</td><td>2</td></tr><tr><td>Q3</td><td>2</td></tr><tr><td>Maximum</td><td>4</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>2</td></tr><tr><td>Upper Fence</td><td>2</td></tr></table>	Mean	2	Median	2	Modus	2	Five-number summary		Minimum	0	Q1	2	Q2	2	Q3	2	Maximum	4	Interquartile range		Lower Fence	2	Upper Fence	2	<p>Boxplot Jumlah Tamu Dewasa</p> 	Boxplot ini menunjukkan bahwa mayoritas jumlah tamu dewasa adalah 2, dengan nilai 0, 1, 3, dan 4 sebagai outlier. Median berada di 2, dan tidak ada whiskers yang terlihat, mengindikasikan data sangat terpusat pada nilai tersebut.
Mean	2																											
Median	2																											
Modus	2																											
Five-number summary																												
Minimum	0																											
Q1	2																											
Q2	2																											
Q3	2																											
Maximum	4																											
Interquartile range																												
Lower Fence	2																											
Upper Fence	2																											
2	no_of_children	<table><tr><td>Mean</td><td>0</td></tr><tr><td>Median</td><td>0</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>0</td></tr><tr><td>Q2</td><td>0</td></tr><tr><td>Q3</td><td>0</td></tr><tr><td>Maximum</td><td>10</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>0</td></tr><tr><td>Upper Fence</td><td>0</td></tr></table>	Mean	0	Median	0	Modus	0	Five-number summary		Minimum	0	Q1	0	Q2	0	Q3	0	Maximum	10	Interquartile range		Lower Fence	0	Upper Fence	0	<p>Boxplot Jumlah Tamu Anak-anak</p> 	Boxplot ini menunjukkan distribusi jumlah tamu anak-anak. Median berada di 0, menunjukkan bahwa lebih dari setengah data memiliki nilai 0. Tidak ada whiskers yang terlihat, menandakan data sangat terpusat pada nilai rendah.
Mean	0																											
Median	0																											
Modus	0																											
Five-number summary																												
Minimum	0																											
Q1	0																											
Q2	0																											
Q3	0																											
Maximum	10																											
Interquartile range																												
Lower Fence	0																											
Upper Fence	0																											
3	lead_time	<table><tr><td>Mean</td><td>85</td></tr><tr><td>Median</td><td>57</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>17</td></tr><tr><td>Q2</td><td>57</td></tr><tr><td>Q3</td><td>126</td></tr><tr><td>Maximum</td><td>443</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>-146.5</td></tr><tr><td>Upper Fence</td><td>289.5</td></tr></table>	Mean	85	Median	57	Modus	0	Five-number summary		Minimum	0	Q1	17	Q2	57	Q3	126	Maximum	443	Interquartile range		Lower Fence	-146.5	Upper Fence	289.5	<p>Boxplot Lead Time</p> 	Visualisasi ini merupakan boxplot dari Lead Time yang menunjukkan distribusi data. Mayoritas nilai Lead Time berada antara 50 hingga 100, dengan beberapa outlier yang mencapai lebih dari 400. Garis tengah kotak menunjukkan median, sementara whisker dan titik-titik di atasnya menunjukkan variasi dan outlier dalam data.
Mean	85																											
Median	57																											
Modus	0																											
Five-number summary																												
Minimum	0																											
Q1	17																											
Q2	57																											
Q3	126																											
Maximum	443																											
Interquartile range																												
Lower Fence	-146.5																											
Upper Fence	289.5																											

4	avg_price_per_room	<table><tr><td>Mean</td><td>103</td></tr><tr><td>Median</td><td>99</td></tr><tr><td>Modus</td><td>65</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>80.3</td></tr><tr><td>Q2</td><td>99</td></tr><tr><td>Q3</td><td>120</td></tr><tr><td>Maximum</td><td>540</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td></td><td>39.7</td></tr><tr><td>Lower Fence</td><td>20.75</td></tr><tr><td>Upper Fence</td><td>179.55</td></tr></table>	Mean	103	Median	99	Modus	65	Five-number summary		Minimum	0	Q1	80.3	Q2	99	Q3	120	Maximum	540	Interquartile range			39.7	Lower Fence	20.75	Upper Fence	179.55	<div>Boxplot Rata-rata Harga Kamar</div>	Visualisasi ini adalah boxplot dari rata-rata harga kamar yang menunjukkan bahwa mayoritas harga berkisar antara 50 hingga 100, dengan median di sekitar 75. Ada beberapa outlier yang mencapai hingga lebih dari 500, menunjukkan variasi signifikan di antara harga-harga tersebut. Whisker dan titik-titik di atasnya menggambarkan distribusi harga dan keberadaan outlier.
Mean	103																													
Median	99																													
Modus	65																													
Five-number summary																														
Minimum	0																													
Q1	80.3																													
Q2	99																													
Q3	120																													
Maximum	540																													
Interquartile range																														
	39.7																													
Lower Fence	20.75																													
Upper Fence	179.55																													
5	type_of_meal_plan	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>type_of_meal_plan</td><td>Count of type_of_meal_plan</td></tr><tr><td>Meal Plan 1</td><td>27835</td></tr><tr><td>Meal Plan 2</td><td>3305</td></tr><tr><td>Meal Plan 3</td><td>0</td></tr><tr><td>Not Selected</td><td>5130</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>type_of_meal_plan</td><td>Proportion of type_of_meal_plan</td></tr><tr><td>Meal Plan 1</td><td>76.74%</td></tr><tr><td>Meal Plan 2</td><td>9.11%</td></tr><tr><td>Meal Plan 3</td><td>0.00%</td></tr><tr><td>Not Selected</td><td>14.14%</td></tr></table>	Frekuensi		type_of_meal_plan	Count of type_of_meal_plan	Meal Plan 1	27835	Meal Plan 2	3305	Meal Plan 3	0	Not Selected	5130	Proporsi		type_of_meal_plan	Proportion of type_of_meal_plan	Meal Plan 1	76.74%	Meal Plan 2	9.11%	Meal Plan 3	0.00%	Not Selected	14.14%	<div>Distribusi Tipe Meal Plan</div>	Visualisasi ini memvisualisasikan distribusi tipe meal plan yang dipesan saat melakukan reservasi. Dari visualisasi tersebut dapat dilihat meal plan 1 merupakan meal plan yang paling banyak dipesan dan lebih banyak reservasi yang tidak memesan meal plan dibandingkan dengan reservasi yang memesan meal plan 2. Dan tidak ada reservasi yang memesan meal plan 3.		
Frekuensi																														
type_of_meal_plan	Count of type_of_meal_plan																													
Meal Plan 1	27835																													
Meal Plan 2	3305																													
Meal Plan 3	0																													
Not Selected	5130																													
Proporsi																														
type_of_meal_plan	Proportion of type_of_meal_plan																													
Meal Plan 1	76.74%																													
Meal Plan 2	9.11%																													
Meal Plan 3	0.00%																													
Not Selected	14.14%																													
6	required_car_parking_space	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>required_parking_space</td><td>count</td></tr><tr><td>1</td><td>35151</td></tr><tr><td>0</td><td>1124</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>required_parking_space</td><td>proportion</td></tr><tr><td>Butuh</td><td>96.90%</td></tr><tr><td>Tidak Butuh</td><td>3.10%</td></tr></table>	Frekuensi		required_parking_space	count	1	35151	0	1124	Proporsi		required_parking_space	proportion	Butuh	96.90%	Tidak Butuh	3.10%	<div>Proporsi Kebutuhan Tempat Parkir</div>	Dari visualisasi pie chart yang merepresentasikan proporsi kebutuhan tempat parkir, dapat dilihat bahwa 96.9% dari reservasi membutuhkan tempat parkir, dan 3.1% reservasi tidak membutuhkan tempat parkir.										
Frekuensi																														
required_parking_space	count																													
1	35151																													
0	1124																													
Proporsi																														
required_parking_space	proportion																													
Butuh	96.90%																													
Tidak Butuh	3.10%																													

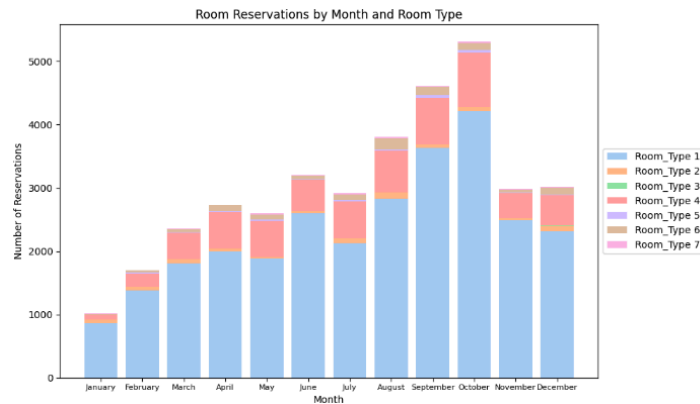
7	market_segment_type	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>market_segment_type</td><td>Count of market_segment_type</td></tr><tr><td>Online</td><td>23214</td></tr><tr><td>Complementary</td><td>391</td></tr><tr><td>Corporate</td><td>2017</td></tr><tr><td>Offline</td><td>10528</td></tr><tr><td>Aviation</td><td>125</td></tr><tr><td colspan="2"> </td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>market_segment_type</td><td>Proportion of market_segment_type</td></tr><tr><td>Online</td><td>63.99%</td></tr><tr><td>Complementary</td><td>1.08%</td></tr><tr><td>Corporate</td><td>5.56%</td></tr><tr><td>Offline</td><td>29.02%</td></tr><tr><td>Aviation</td><td>0.34%</td></tr></table>	Frekuensi		market_segment_type	Count of market_segment_type	Online	23214	Complementary	391	Corporate	2017	Offline	10528	Aviation	125			Proporsi		market_segment_type	Proportion of market_segment_type	Online	63.99%	Complementary	1.08%	Corporate	5.56%	Offline	29.02%	Aviation	0.34%	<p>Distribusi Segmen Pasar</p>  <table><tr><th>Segment</th><th>Count</th></tr><tr><td>Offline</td><td>10528</td></tr><tr><td>Online</td><td>23214</td></tr><tr><td>Corporate</td><td>2017</td></tr><tr><td>Complementary</td><td>391</td></tr><tr><td>Aviation</td><td>125</td></tr></table>	Segment	Count	Offline	10528	Online	23214	Corporate	2017	Complementary	391	Aviation	125	Dari visualisasi distribusi segmen pasar, dapat dilihat bahwa segmen pasar yang paling banyak melakukan reservasi adalah segmen pasar online. Dan kedua terbanyak adalah segmen pasar offline, diikuti oleh segmen pasar corporate, complementary, dan aviation.												
Frekuensi																																																										
market_segment_type	Count of market_segment_type																																																									
Online	23214																																																									
Complementary	391																																																									
Corporate	2017																																																									
Offline	10528																																																									
Aviation	125																																																									
Proporsi																																																										
market_segment_type	Proportion of market_segment_type																																																									
Online	63.99%																																																									
Complementary	1.08%																																																									
Corporate	5.56%																																																									
Offline	29.02%																																																									
Aviation	0.34%																																																									
Segment	Count																																																									
Offline	10528																																																									
Online	23214																																																									
Corporate	2017																																																									
Complementary	391																																																									
Aviation	125																																																									
8	arrival_month	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>arrival_month</td><td>Count of arrival_month</td></tr><tr><td>Jan</td><td>1014</td></tr><tr><td>Feb</td><td>1704</td></tr><tr><td>Mar</td><td>2358</td></tr><tr><td>Apr</td><td>2736</td></tr><tr><td>May</td><td>2598</td></tr><tr><td>Jun</td><td>3203</td></tr><tr><td>Jul</td><td>2920</td></tr><tr><td>Aug</td><td>3813</td></tr><tr><td>Sep</td><td>4611</td></tr><tr><td>Oct</td><td>5317</td></tr><tr><td>Nov</td><td>2980</td></tr><tr><td>Dec</td><td>3021</td></tr></table>	Frekuensi		arrival_month	Count of arrival_month	Jan	1014	Feb	1704	Mar	2358	Apr	2736	May	2598	Jun	3203	Jul	2920	Aug	3813	Sep	4611	Oct	5317	Nov	2980	Dec	3021	<p>Tren Kedatangan Tamu Per-Bulan</p>  <table><tr><th>Month</th><th>Count</th></tr><tr><td>Jan</td><td>1014</td></tr><tr><td>Feb</td><td>1704</td></tr><tr><td>Mar</td><td>2358</td></tr><tr><td>Apr</td><td>2736</td></tr><tr><td>May</td><td>2598</td></tr><tr><td>Jun</td><td>3203</td></tr><tr><td>Jul</td><td>2920</td></tr><tr><td>Aug</td><td>3813</td></tr><tr><td>Sep</td><td>4611</td></tr><tr><td>Oct</td><td>5317</td></tr><tr><td>Nov</td><td>2980</td></tr><tr><td>Dec</td><td>3021</td></tr></table>	Month	Count	Jan	1014	Feb	1704	Mar	2358	Apr	2736	May	2598	Jun	3203	Jul	2920	Aug	3813	Sep	4611	Oct	5317	Nov	2980	Dec	3021	Visualisasi ini menunjukkan tren kedatangan tamu per bulan sepanjang tahun. Jumlah tamu meningkat secara bertahap dari Januari, mencapai puncaknya pada bulan September, kemudian menurun tajam pada bulan Oktober sebelum sedikit meningkat kembali pada bulan Desember.
Frekuensi																																																										
arrival_month	Count of arrival_month																																																									
Jan	1014																																																									
Feb	1704																																																									
Mar	2358																																																									
Apr	2736																																																									
May	2598																																																									
Jun	3203																																																									
Jul	2920																																																									
Aug	3813																																																									
Sep	4611																																																									
Oct	5317																																																									
Nov	2980																																																									
Dec	3021																																																									
Month	Count																																																									
Jan	1014																																																									
Feb	1704																																																									
Mar	2358																																																									
Apr	2736																																																									
May	2598																																																									
Jun	3203																																																									
Jul	2920																																																									
Aug	3813																																																									
Sep	4611																																																									
Oct	5317																																																									
Nov	2980																																																									
Dec	3021																																																									
9	booking_status	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>booking_status</td><td>count</td></tr><tr><td>Not_Canceled</td><td>24390</td></tr><tr><td>Canceled</td><td>11885</td></tr><tr><td colspan="2"> </td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>booking_status</td><td>proportion</td></tr><tr><td>Not_Canceled</td><td>67.24%</td></tr><tr><td>Canceled</td><td>32.76%</td></tr></table>	Frekuensi		booking_status	count	Not_Canceled	24390	Canceled	11885			Proporsi		booking_status	proportion	Not_Canceled	67.24%	Canceled	32.76%	<p>Proporsi Status Pemesanan</p>  <table><tr><th>Status</th><th>Count</th><th>Proportion</th></tr><tr><td>Not_Canceled</td><td>24390</td><td>67.24%</td></tr><tr><td>Canceled</td><td>11885</td><td>32.76%</td></tr></table>	Status	Count	Proportion	Not_Canceled	24390	67.24%	Canceled	11885	32.76%	Visualisasi proporsi status pemesanan merepresentasikan seberapa banyak reservasi yang dibatalkan dan yang tidak dibatalkan. 67.2% dari reservasi yang ada tidak dibatalkan, dan 32.8% dari reservasi yang ada dibatalkan.																											
Frekuensi																																																										
booking_status	count																																																									
Not_Canceled	24390																																																									
Canceled	11885																																																									
Proporsi																																																										
booking_status	proportion																																																									
Not_Canceled	67.24%																																																									
Canceled	32.76%																																																									
Status	Count	Proportion																																																								
Not_Canceled	24390	67.24%																																																								
Canceled	11885	32.76%																																																								

10	no_of_week_nights	<table><tr><td>Mean</td><td>2</td></tr><tr><td>Median</td><td>2</td></tr><tr><td>Modus</td><td>2</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>1</td></tr><tr><td>Q2</td><td>2</td></tr><tr><td>Q3</td><td>3</td></tr><tr><td>Maximum</td><td>17</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>-2</td></tr><tr><td>Upper Fence</td><td>6</td></tr></table>	Mean	2	Median	2	Modus	2	Five-number summary		Minimum	0	Q1	1	Q2	2	Q3	3	Maximum	17	Interquartile range		Lower Fence	-2	Upper Fence	6	<p>Boxplot Durasi Menginap (Weekday)</p> <p>Visualisasi tersebut adalah boxplot yang menunjukkan durasi menginap selama hari kerja (weekday). Dari boxplot tersebut, terlihat bahwa median durasi menginap sekitar 2 hari, dengan kuartil pertama (Q1) sekitar 1.5 hari dan kuartil ketiga (Q3) sekitar 3 hari. Terdapat beberapa outlier yang menunjukkan durasi menginap yang jauh lebih lama, bahkan ada yang mencapai lebih dari 15 hari.</p>
Mean	2																										
Median	2																										
Modus	2																										
Five-number summary																											
Minimum	0																										
Q1	1																										
Q2	2																										
Q3	3																										
Maximum	17																										
Interquartile range																											
Lower Fence	-2																										
Upper Fence	6																										
11	no_of_weekend_nights	<table><tr><td>Mean</td><td>1</td></tr><tr><td>Median</td><td>1</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>0</td></tr><tr><td>Q2</td><td>1</td></tr><tr><td>Q3</td><td>2</td></tr><tr><td>Maximum</td><td>7</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>-3</td></tr><tr><td>Upper Fence</td><td>5</td></tr></table>	Mean	1	Median	1	Modus	0	Five-number summary		Minimum	0	Q1	0	Q2	1	Q3	2	Maximum	7	Interquartile range		Lower Fence	-3	Upper Fence	5	<p>Boxplot Durasi Menginap (Weekend)</p> <p>Visualisasi tersebut adalah boxplot yang menunjukkan durasi menginap selama akhir pekan (weekend). Median durasi menginap adalah 1 hari, dengan kuartil pertama (Q1) sekitar 0.5 hari dan kuartil ketiga (Q3) sekitar 2 hari. Terdapat beberapa outlier yang menunjukkan durasi menginap yang lebih lama, dengan nilai tertinggi mencapai sekitar 7 hari.</p>
Mean	1																										
Median	1																										
Modus	0																										
Five-number summary																											
Minimum	0																										
Q1	0																										
Q2	1																										
Q3	2																										
Maximum	7																										
Interquartile range																											
Lower Fence	-3																										
Upper Fence	5																										
12	repeated_guests	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>repeated_guests</td><td>count</td></tr><tr><td>0</td><td>35345</td></tr><tr><td>1</td><td>930</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>booking_status</td><td>proportion</td></tr><tr><td>0</td><td>97.44%</td></tr><tr><td>1</td><td>2.56%</td></tr></table>	Frekuensi		repeated_guests	count	0	35345	1	930	Proporsi		booking_status	proportion	0	97.44%	1	2.56%	<p>Proporsi Tamu Langganan</p> <p>Visualisasi proporsi tamu langganan menunjukkan proporsi dari seberapa banyak tamu yang melakukan reservasi berulang. 97.4% dari tamu merupakan tamu yang baru satu kali melakukan reservasi, dan 2.6% dari tamu yang ada telah melakukan reservasi lebih dari satu kali.</p>								
Frekuensi																											
repeated_guests	count																										
0	35345																										
1	930																										
Proporsi																											
booking_status	proportion																										
0	97.44%																										
1	2.56%																										

g. Pola Data

No	Pola Data	Analisis
1	 <p>The scatter plot shows a dense cluster of blue data points. The x-axis, labeled 'avg_price_per_room', ranges from 0 to 500. The y-axis, labeled 'lead_time', ranges from 0 to 400. The points are most concentrated at lower prices and shorter lead times, with a few outliers at higher prices and longer lead times.</p>	<p>Dari visualisasi perbandingan harga kamar dengan lead time dari pemesanan kamar, dapat dilihat bahwa semakin mahal harga kamar maka durasi lead time akan semakin berkurang. Hal ini menunjukkan bahwa orang yang memesan kamar dengan harga yang tinggi biasanya memesan dalam waktu dekat. Akan tetapi, hal tersebut juga dapat mengindikasikan bahwa kamar yang dipesan dalam waktu dekat dapat menyebabkan harga kamar menjadi lebih mahal. Sebaliknya, kamar yang dipesan jauh sebelum waktu reservasi memiliki harga yang lebih murah dan terjadi peningkatan pemesanan di rentang waktu tersebut</p>
2	 <p>The bar chart displays the count of meal plan selections for 0, 1, 2, 3, 9, and 10 children. The y-axis, labeled 'Count', ranges from 0 to 25,000. The legend indicates four categories: Meal Plan 1 (blue), Meal Plan 2 (orange), Meal Plan 3 (green), and Not Selected (red). For 0 children, Meal Plan 1 has the highest count (around 25,000), followed by Not Selected (around 5,000) and Meal Plan 2 (around 3,000). For 1 child, Meal Plan 1 is around 1,000 and Not Selected is around 500. For 2 children, Meal Plan 1 is around 1,000. For 3 children, Meal Plan 1 is around 500. For 9 and 10 children, the counts are very low, near zero.</p>	<p>Dari visualisasi pengaruh jumlah anak terhadap anak terhadap meal plan yang dipesan, kamar yang ada anak-anak dan yang tidak ada anak-anak cenderung untuk memesan meal plan 1. Dan untuk setiap kamar yang ada anak-anaknya cenderung untuk memesan meal plan. Dari seluruh kamar yang dipesan dengan keberadaan anak-anak, hanya 112 kamar yang tidak memesan meal plan saat reservasi. Namun untuk kamar yang ada anak-anak, tidak ada pemesanan meal plan 3 pada saat reservasi.</p>

3



Dari visualisasi tren reservasi tipe kamar setiap bulan, dapat dilihat bahwa reservasi kamar paling banyak dilakukan pada bulan 10 atau bulan Oktober. Dan dapat dilihat juga dari visualisasi tersebut, reservasi paling sedikit ada di bulan 1 atau bulan Januari. Jika dilihat dari visualisasi tersebut, disetiap bulannya, jenis kamar yang paling sering dipesan atau direservasi adalah Room_Type 1.

Beberapa penyebab terjadinya peningkatan di September dan Oktober

- Musim Gugur

<https://web-holidays.com/holiday-calendar/september-holidays-celebrations/>

- Festival Halloween

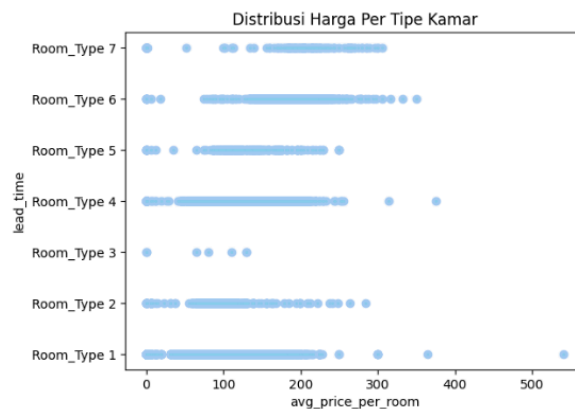
<https://web-holidays.com/holiday-calendar/october-holidays-celebrations/>

Beberapa penyebab terjadinya penurunan di Januari dan Februari

- Puncak Musim Salju

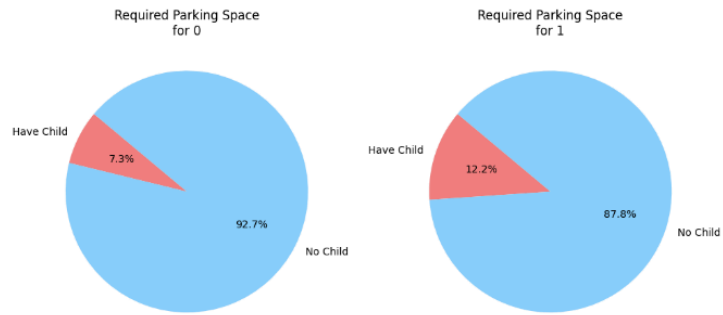
<https://www.foxweather.com/learn/which-is-the-snowiest-month-of-the-year>

4



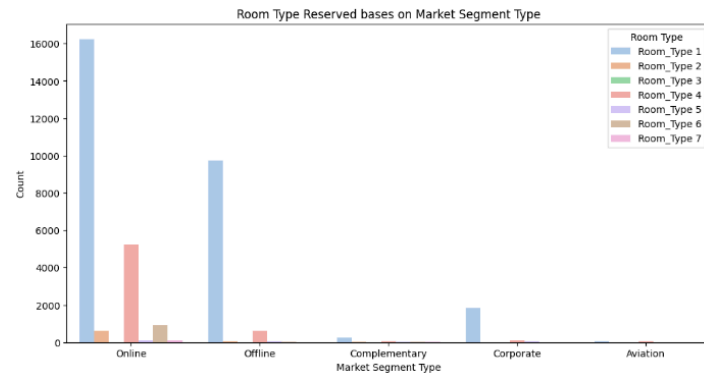
Dari visualisasi data distribusi harga per tipe kamar, dapat disimpulkan bahwa Kamar Tipe 1 dan 4 memiliki kecenderungan rata-rata harga yang lebih murah. Hal ini sekaligus menjawab pola yang terjadi pada analisis nomor 3 alasan mengapa Kamar Tipe 1 dan 4 memiliki jumlah pemesanan tertinggi. Pemesan cenderung memilih tipe kamar dengan harga yang lebih murah.

5



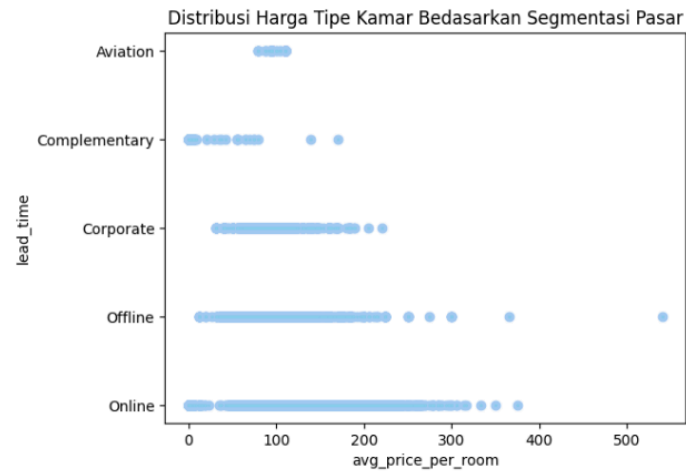
Dari visualisasi di samping, terjadi perbedaan sebanyak 4.9% terhadap kebutuhan tempat parkir apabila pemesan memiliki anak atau tidak. Pada kategori ini, jika `no_of_children = 0` maka No Child dan jika `no_of_children > 1` maka masuk ke dalam kategori Have Child. Hal ini menunjukkan bahwa pemesan yang memiliki anak cenderung lebih membutuhkan tempat parkir, yang berarti bahwa pemesan membawa kendaraan pribadi. Akan tetapi pengaruh ada atau tidak adanya anak, tidak terlalu signifikan terhadap kebutuhan tempat parkir

6



Dari visualisasi Tipe kamar berdasarkan segmen pasar, dapat disimpulkan bahwa kamar tipe 1 merupakan kamar yang paling diminati oleh semua segmen. Sedangkan kamar tipe 3, 5 dan 7 merupakan kamar yang paling tidak diminati oleh tamu. Dari visualisasi data ini, dapat juga disimpulkan bahwa segmen pasar terbesar berasal dari online.

7

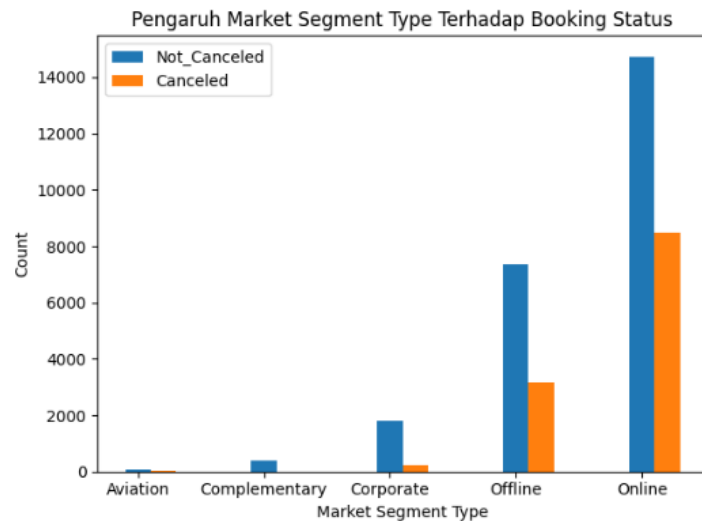


Berdasarkan visualisasi di samping, dapat dilihat bahwa pasar Aviation memiliki rentang rata-rata harga kamar yang sangat sempit, berbeda dengan pasar Online yang memiliki rentang harga kamar yang sangat luas. Hal ini menunjukkan pasar Aviation memiliki segmentasi yang tidak terlalu beragam.

Sedangkan rentang rata-rata harga kamar pada pasar Offline tidak seluas Online, akan tetapi pemesanan harga kamar tertinggi hanya terjadi pada pasar Offline.

Pasar Complementary memiliki kecenderungan untuk memilih harga kamar yang lebih rendah (memiliki upper fence lebih rendah dibandingkan pasar lain), sedangkan pasar Corporate memiliki kecenderungan harga kamar yang kompetitif (persebaran merata dan tidak memiliki banyak outliers)

8



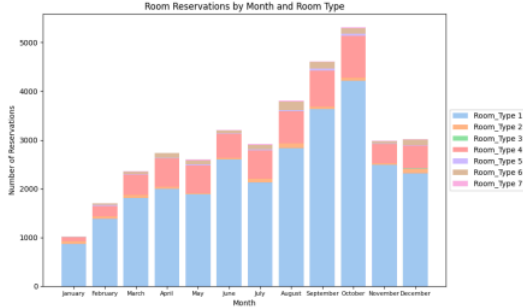
Dari visualisasi pengaruh market segment type terhadap booking status, dapat dilihat bahwa untuk setiap golongan segmen pasar yang ada banyaknya pemesanan atau reservasi yang tidak dibatalkan lebih banyak dibandingkan dengan yang dibatalkan. Namun, untuk segmen pasar online, jumlah pemesanan yang dibatalkan cukup tinggi nilainya dibandingkan dengan segmen pasar lainnya.

III. Dashboard Hotel Reservation Dataset

Dashboard 1

Personalisasi Penjualan Berdasarkan Perilaku Konsumen

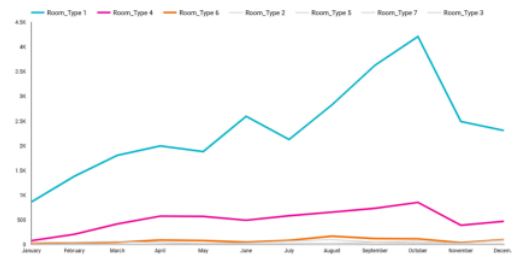
Visualisasi Total Penjualan Kamar di Setiap Bulan



Informasi Chart 1

Penjualan tertinggi terjadi di Bulan Oktober dan Kamar yang paling sering dipesan adalah Kamar Tipe 1. Sedangkan penurunan penjualan terjadi dari bulan November hingga Januari

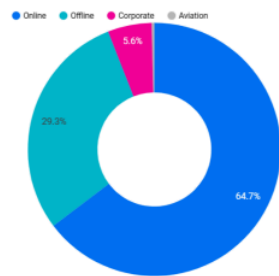
Tren Penjualan Kamar Setiap Kamar selama Bulan Januari-Desember



Informasi Chart 2

Kamar Tipe 1 merupakan tipe kamar yang paling diminati oleh tamu dengan tren menaik, yang dimana puncaknya berada di bulan Oktober sebanyak ± 4000 kamar. Sedangkan kamar Tipe lain cenderung stabil namun dengan jumlah penjualan yang jauh lebih sedikit

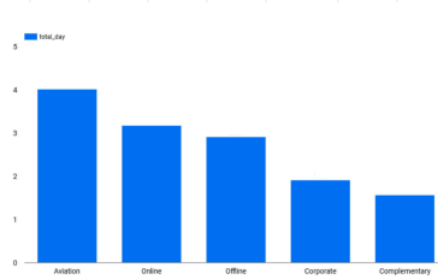
Proporsi Penjualan Kamar Berdasarkan Segmentasi pasar



Informasi Chart 3

Konsumen paling sering melakukan reservasi secara online dengan market share sebesar 64,7%. Selanjutnya, konsumen lebih sering melakukan reservasi secara Offline sebesar 29.3%. Di posisi ketiga dan keempat yaitu Corporate sebesar 5.6% dan Aviation dengan proporsi sebesar 0.4%.

Lama Waktu Menginap Berdasarkan Segmentasi Pasar



Informasi Chart 4

Aviation memiliki rata-rata durasi menginap tertinggi sekitar 4 hari. Sementara itu, segmen online dan offline memiliki durasi menginap ± 3 hari. Walaupun aviation merupakan segmen pasar terkecil, namun memiliki durasi menginap paling lama

Insights:

Berdasarkan informasi yang diperoleh, konsumen lebih suka untuk melakukan reservasi secara online dibandingkan secara offline. Segmen lain yaitu corporate dan aviation hanya memegang total market share sebesar 6%. Akan tetapi, walaupun Aviation merupakan segmen pasar terkecil, pasar aviation memiliki durasi menginap paling lama yaitu sekitar 4 hari. Untuk tipe kamar, kamar Tipe 1 merupakan favorit utama tamu dengan penjualan puncak mencapai sekitar ± 4000 kamar di bulan Oktober. Untuk tren penjualan, distribusi yang terjadi adalah *left skewed*, menanjak dari januari hingga oktober, dan menurun kembali hingga januari

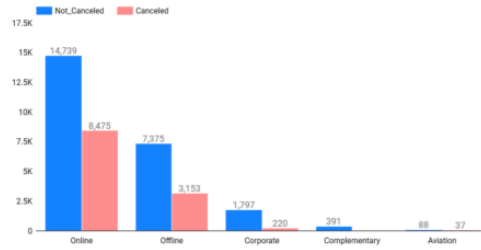
Actionable Insights:

1. Penyesuaian Strategi Penjualan: Perlu diperhatikan faktor yang menyebabkan terjadinya penjualan yang rendah di kuartal pertama (seperti faktor liburan, musim, dsb) sehingga dapat digunakan strategi pemasaran yang lebih efektif. Perusahaan dapat mengoptimalkan strategi penjualan dan pemasaran untuk memanfaatkan puncak penjualan di bulan Oktober, seperti menawarkan promo atau paket khusus jika menginap di bulan-bulan berikutnya.
2. Okupansi Kamar: Karena penjualan didominasi oleh kamar Tipe 1, perusahaan dapat mengalokasikan sumber daya lebih untuk mempromosikan dan meningkatkan penjualan kamar ini. Namun, dengan tetap mempertimbangkan penjualan kamar Tipe lainnya (seperti tipe 4 dan tipe 6 karena termasuk Top 3 penjualan) agar lebih diminati oleh konsumen.
3. Segmentasi Pasar: Karena konsumen cenderung melakukan reservasi secara online, perusahaan dapat fokus pada pengembangan dan peningkatan pengalaman pengguna melalui platform reservasi online, seperti investasi dalam teknologi dan pemasaran digital. Perusahaan juga tetap harus menjaga kualitas pemasaran secara offline, karena segmen offline masih cukup besar.
4. Segmen Aviation: Meskipun segmen Aviation memiliki pangsa pasar yang kecil, durasi menginap yang lebih lama menunjukkan potensi untuk meningkatkan loyalitas pelanggan. Perusahaan dapat menerapkan strategi bisnis khusus untuk meningkatkan layanan dan pengalaman bagi pelanggan Aviation, guna meningkatkan kualitas penjualan.

Dashboard 2

Pola dan Faktor Penyebab Pembatalan Reservasi Kamar Hotel

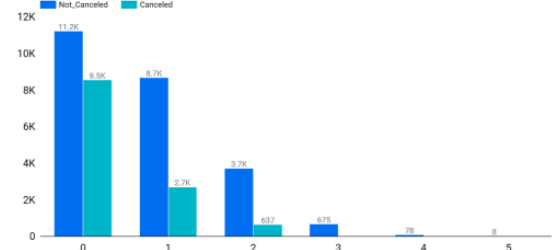
Status Reservasi Berdasarkan Segmentasi Pasar



Informasi Chart 1

Pemesanan kamar yang paling sering dibatalkan adalah pemesanan kamar secara online (karena jumlah pesanan secara online paling sering dilakukan)

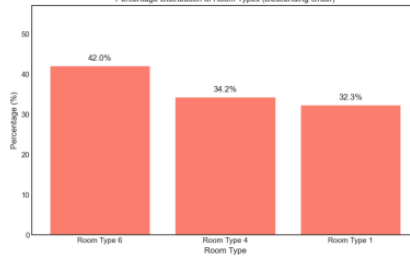
Visualisasi Status Reservasi Berdasarkan Jumlah Special Request



Informasi Chart 2

Pemesanan tanpa special request cenderung untuk dibatalkan. Dapat dilihat dari jumlah pemesanan tanpa special request memiliki jumlah yang hampirimbang antar yang dibatalkan dan yang tidak dibatalkan. Pemesanan tanpa special request memiliki persentase pembatalan sebesar 42%, sedangkan pemesanan dengan satu special request memiliki persentase pembatalan sebesar 27%

Persentase Pembatalan Reservasi Berdasarkan Tipe Kamar yang Dipesan
Percentage Distribution of Room Types (Descending Order)



Informasi Chart 3

Kamar dengan tipe ruangan 1, 4, dan 6 merupakan tipe kamar yang paling banyak dibatalkan.

Room Type 1

9072 canceled
dari 28130

Informasi 4

Walaupun menempati urutan ketiga dalam persentase pembatalan, akan tetapi kamar Tipe 1 merupakan kamar yang memiliki jumlah pembatalan terbanyak. Dari 28130 pemesanan, 9072 pemesanannya dibatalkan.

Insight:

Berdasarkan informasi di atas, terdapat beberapa informasi yang memperlihatkan kecenderungan pembatalan reservasi. Konsumen yang melakukan reservasi secara online, tanpa special request, dan yang memesan tipe kamar 1 adalah contoh penyebab tingginya pembatalan reservasi. Melalui informasi terkait kecenderungan tersebut, dapat dilakukan penelusuran informasi secara lebih mendalam untuk menemukan pola pembatalan reservasi.

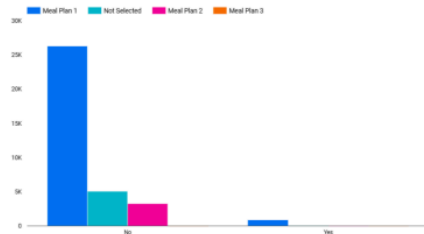
Actionable Insight:

Dari insight yang diperoleh, ada beberapa hal yang dapat dilakukan untuk meminimalisir pembatalan pemesanan. Pertama, menerapkan kebijakan *no cancellation refund* untuk tipe pemesanan tertentu agar konsumen yang sudah memesan tidak membatalkan reservasi. Kedua, menghubungi pelanggan yang tidak memiliki special request, sehingga dengan menghubungi pelanggan, pelanggan dapat merasa sangat terlayani dan memilih untuk tidak membatalkan reservasi.

Dashboard 3

Strategi Penjualan Meal Plan Berdasarkan Perilaku Konsumen

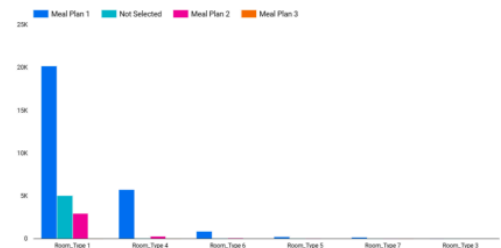
Pemilihan Meal Plan Berdasarkan Status Kehadiran Tamu Baru dan Tamu Lama



Informasi Chart 1

Baik tamu yang belum dan sudah pernah menginap, cenderung akan memilih Meal Plan 1. Tamu yang belum pernah menginap akan memilih Meal Plan 1 dengan peluang sebesar 72%

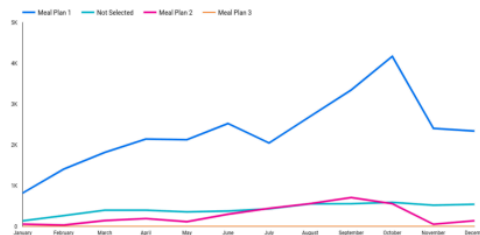
Pemilihan Meal Plan Berdasarkan Tipe Kamar yang Direservasi



Informasi Chart 2

Konsumen yang memesan kamar Tipe 1, lebih banyak memilih tipe meal plan 1 dan tidak memilih. Namun, konsumen yang memesan kamar Tipe 4, lebih banyak memilih meal plan 2 dibandingkan dengan yang tidak memilih

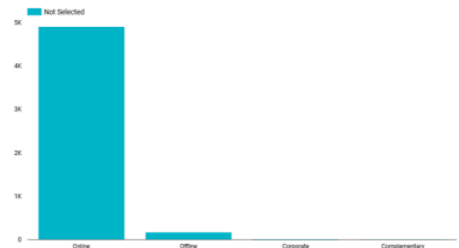
Jumlah Pemilihan Meal Plan di Bulan Januari-Desember



Informasi Chart 3

Meal Plan 1 paling sering dipilih oleh konsumen dan meningkat dari bulan Januari hingga Oktober. Sedangkan di bulan Juli, Agustus, dan September, meal plan tipe 2 lebih sering dipilih dibandingkan yang tidak memilih meal plan

Jumlah Konsumen yang Tidak Memilih Meal Plan Berdasarkan Segmen Pasar



Informasi Chart 4

Tamu dengan segmen pasar aviation memiliki durasi penginapan terlama dibandingkan segmen pasar lainnya

Insight:

Berdasarkan informasi yang diperoleh, dapat disimpulkan bahwa penjualan Meal Plan 1 selalu meningkat bersamaan dengan penjualan kamar. Namun, tipe meal plan lain seperti Meal Plan 2 dan 3 tidak memiliki performa penjualan yang baik (lebih sedikit dibandingkan yang tidak memilih meal plan). Selain itu, konsumen yang melakukan reservasi secara online memiliki jumlah terbanyak untuk tidak memilih meal plan.

Actionable Insight:

Dari Insight yang diperoleh, dapat dilakukan perencanaan pangan untuk mempersiapkan bahan baku di bulan-bulan tertentu agar penjualan dapat lebih efektif. Dari Chart 4, terlihat market Online paling banyak untuk tidak memilih Meal Plan. Oleh karena itu, dapat diterapkan sebuah strategi pemasaran (seperti foto makanan yang lebih estetik, promo menarik, review, dsb) agar konsumen dapat tertarik untuk memilih Meal Plan. Karena bisa saja konsumen tidak mengetahui informasi dari meal plan.

IV. DATASET Hotel Booking Demand

Eksplorasi umum dataset:

- Hotel Booking Demand merupakan data yang berasal dari artikel Hotel Booking Demand Datasets, yang ditulis oleh Nuno Antonio, Ana Almeida, dan Luis Nunes untuk Data in Brief, Volume 22, Februari 2019. Data tersebut diunduh dan dibersihkan oleh Thomas Mock dan Antoine Bichat. Data Hotel Booking Demand berisi informasi pemesanan untuk hotel kota dan hotel resor yang mencakup informasi seperti waktu pemesanan, lama menginap, jumlah tamu, dan jumlah tempat parkir yang tersedia.
- Link sumber dataset: [Hotel Booking Dataset](#)
- Jumlah baris: 119390
- Jumlah kolom: 32
- Identifikasi null values dan Outliers

Nama Kolom	Jenis Data	Isi Data	Null Values	Banyak null values	Outliers	Nilai Outliers
hotel	Text	Jenis hotel, Resort Hotel atau City Hotel	Tidak ada	Tidak ada	Tidak ada	Tidak ada
is_canceled	Integer (Category)	Menunjukkan apakah pemesanan dibatalkan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
lead_time	Integer	Jarak antar hari pemesanan dengan hari penginapan	Tidak ada	Tidak ada	Ada	373-737
arrival_date_year	Integer	Tahun penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
arrival_date_month	Text	Bulan penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
arrival_date_week_number	Integer	Penginapan Minggu ke sekian	Tidak ada	Tidak ada	Tidak ada	Tidak ada
arrival_date_day_of_month	Integer	Tanggal penginapan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
stays_in_weekend_nights	Integer	Jumlah hari penginapan di akhir pekan	Tidak ada	Tidak ada	Ada	5-19
stays_in_week_nights	Integer	Jumlah hari di akhir pekan yang dipesan oleh pengunjung	Tidak ada	Tidak ada	Ada	6-50
adults	Integer	Jumlah hari biasa yang dipesan oleh pengunjung	Tidak ada	Tidak ada	Ada	2-55
children	Integer	Jumlah anak-anak	Ada	4	Ada	0-10
babies	Integer	Jumlah bayi	Tidak ada	Tidak ada	Ada	0-10
meal	Text	Jenis paket makanan yang dipesan saat reservasi.	Tidak ada	Tidak ada	Tidak ada	Tidak ada
country	Text	Asal negara pengunjung	Ada	488	Tidak ada	Tidak ada

market_segment	Text	Penggolongan segmentasi pasar	Tidak ada	Tidak ada	Tidak ada	Tidak ada
distribution_channel	Text	Perantara reservasi	Tidak ada	Tidak ada	Tidak ada	Tidak ada
is_repeated_guest	Integer (Category)	Tamu pengunjung kembali	Tidak ada	Tidak ada	Tidak ada	Tidak ada
previous_cancellations	Integer	Jumlah reservasi yang pernah dibatalkan.	Tidak ada	Tidak ada	Ada	0-26
previous_bookings_not_canceled	Integer	Jumlah reservasi yang tidak dibatalkan.	Tidak ada	Tidak ada	Ada	0-72
reserved_room_type	Text	Jenis kamar yang ditempati	Tidak ada	Tidak ada	Tidak ada	Tidak ada
assigned_room_type	Text	Jenis kamar yang diberikan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
booking_changes	Integer	Jumlah perubahan yang dibuat setelah pemesanan	Tidak ada	Tidak ada	Ada	0-21
deposit_type	Text	Jenis pembayaran deposit	Tidak ada	Tidak ada	Tidak ada	Tidak ada
agent	Integer	id dari agensi yang melakukan reservasi	Ada	16340	Tidak ada	Tidak ada
company	Integer	id dari perusahaan yang melakukan reservasi	Ada	112593	Tidak ada	Tidak ada
days_in_waiting_list	Integer	Jumlah hari pemesanan dalam daftar tunggu pra-konfirmasi ke pelanggan	Tidak ada	Tidak ada	Ada	0-391
customer_type	Text	Jenis pelanggan	Tidak ada	Tidak ada	Tidak ada	Tidak ada
adr	Integer	Tarif harian rata-rata	Tidak ada	Tidak ada	Tidak ada	Tidak ada
required_car_parking_spaces	Integer	Jumlah tempat parkir yang dibutuhkan	Tidak ada	Tidak ada	Ada	0-8
total_of_special_requests	Integer	Jumlah permintaan khusus	Tidak ada	Tidak ada	Ada	2.5-5.0
reservation_status	Text	Status reservasi	Tidak ada	Tidak ada	Tidak ada	Tidak ada
reservation_status_date	Date	Tanggal terakhir status reservasi dibuat	Tidak ada	Tidak ada	Tidak ada	Tidak ada

- Sample 5 baris data pertama dan 5 baris data terakhir untuk 6 kolom pertama saja dari Hotel Booking Demand Dataset

5 Baris Pertama

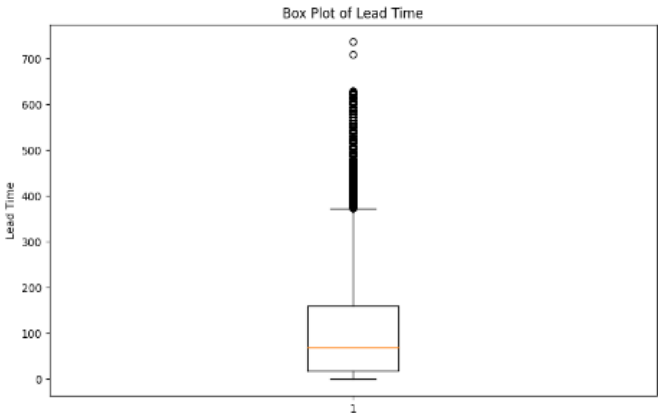
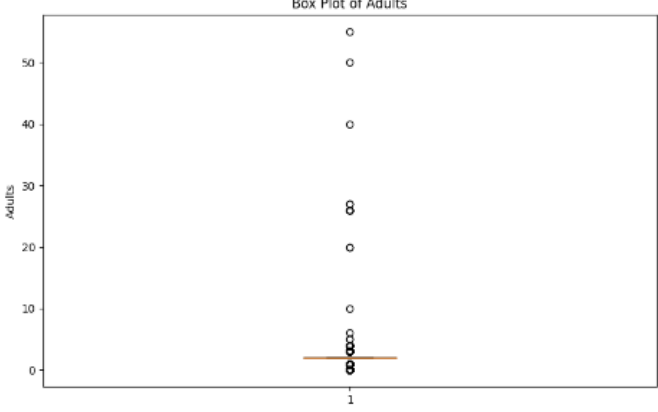
hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number
Resort Hotel	0	342	2015	July	27
Resort Hotel	0	737	2015	July	27
Resort Hotel	0	7	2015	July	27
Resort Hotel	0	13	2015	July	27
Resort Hotel	0	14	2015	July	27

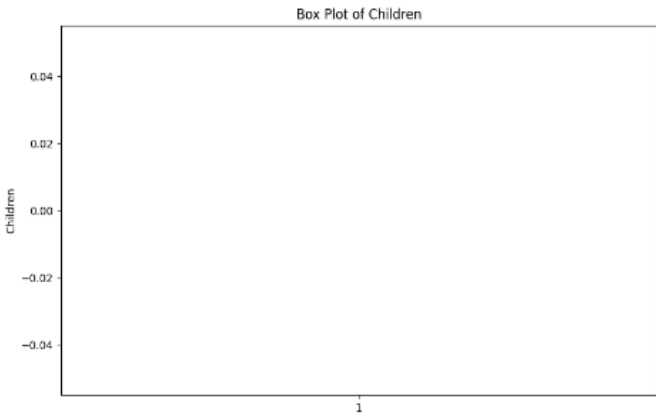

5 Baris Terakhir

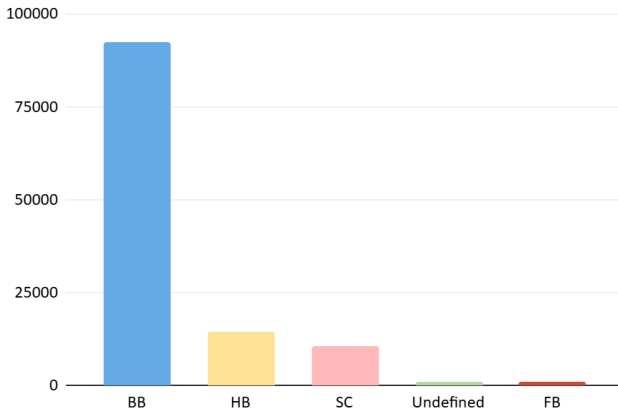
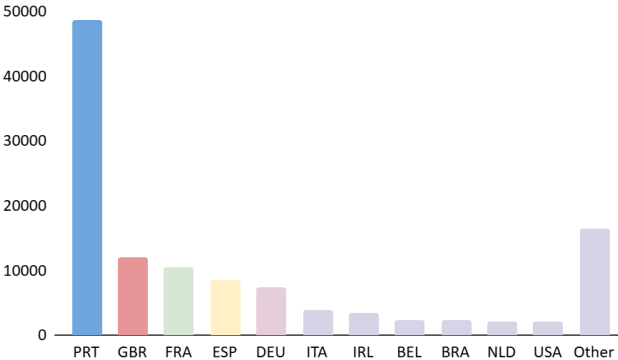
hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number
City Hotel	0	23	2017	August	35
City Hotel	0	102	2017	August	35
City Hotel	0	34	2017	August	35
City Hotel	0	109	2017	August	35
City Hotel	0	205	2017	August	35

○ Descriptive Analysis

No	Kolom	Analisis	Visualisasi	Interpretasi																				
1	hotel	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>hotel_count</td><td>count</td></tr><tr><td>City Hotel</td><td>79330</td></tr><tr><td>Resort Hotel</td><td>40060</td></tr><tr><td>Total</td><td>119390</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>hotel_count</td><td>proportion</td></tr><tr><td>City Hotel</td><td>66.45%</td></tr><tr><td>Resort Hotel</td><td>33.55%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		hotel_count	count	City Hotel	79330	Resort Hotel	40060	Total	119390	Proporsi		hotel_count	proportion	City Hotel	66.45%	Resort Hotel	33.55%	Total	100.00%	<p>Proporsi Hotel</p> <p>Resort Hotel 33.6%</p> <p>City Hotel 66.4%</p>	Berdasarkan visualisasi di samping, terlihat bahwa keberadaan City Hotel memiliki jumlah yang lebih banyak dibandingkan Resort Hotel. Jumlah City Hotel adalah sebanyak 66.4% (79.330) dan Resort Hotel adalah sebanyak 33.55% (40.060)
Frekuensi																								
hotel_count	count																							
City Hotel	79330																							
Resort Hotel	40060																							
Total	119390																							
Proporsi																								
hotel_count	proportion																							
City Hotel	66.45%																							
Resort Hotel	33.55%																							
Total	100.00%																							
2	is_canceled	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>is_canceled_count</td><td>count</td></tr><tr><td>0</td><td>75166</td></tr><tr><td>1</td><td>44224</td></tr><tr><td>Total</td><td>119390</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>is_canceled_count</td><td>proportion</td></tr><tr><td>not canceled</td><td>62.96%</td></tr><tr><td>canceled</td><td>37.04%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		is_canceled_count	count	0	75166	1	44224	Total	119390	Proporsi		is_canceled_count	proportion	not canceled	62.96%	canceled	37.04%	Total	100.00%	<p>Proporsi Reservasi Dibatalkan</p> <p>canceled 37.0%</p> <p>not canceled 63.0%</p>	Berdasarkan visualisasi di samping, terlihat bahwa reservasi yang tidak dibatalkan (0) lebih banyak dibandingkan reservasi yang dibatalkan (1). Jumlah reservasi yang tidak dibatalkan adalah sebanyak 62.96% dan reservasi yang dibatalkan adalah sebanyak 37.04%
Frekuensi																								
is_canceled_count	count																							
0	75166																							
1	44224																							
Total	119390																							
Proporsi																								
is_canceled_count	proportion																							
not canceled	62.96%																							
canceled	37.04%																							
Total	100.00%																							

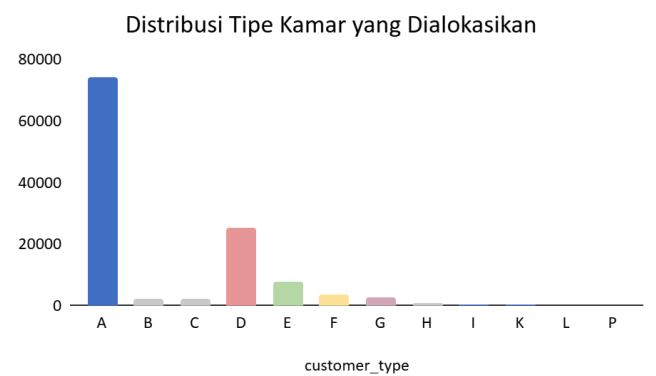
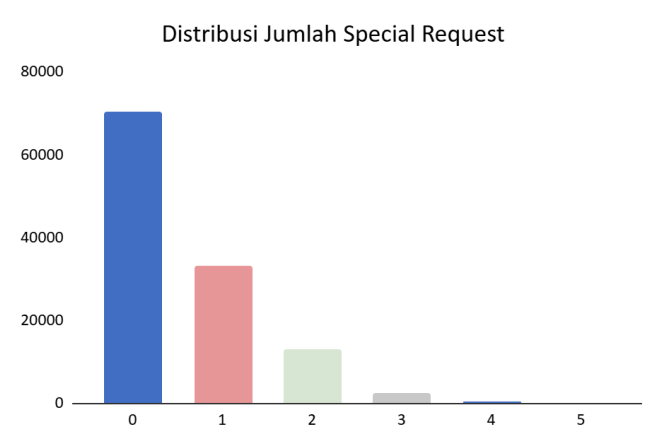
3	lead_time	<table><tr><td>Mean</td><td>104.011</td></tr><tr><td>Median</td><td>69</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>18</td></tr><tr><td>Q2</td><td>69</td></tr><tr><td>Q3</td><td>160</td></tr><tr><td>Maximum</td><td>737</td></tr><tr><td colspan="2">Interquartile range</td><td>142</td></tr><tr><td colspan="2">Lower Fence</td><td>-195</td></tr><tr><td colspan="2">Upper Fence</td><td>373</td></tr></table>	Mean	104.011	Median	69	Modus	0	Five-number summary		Minimum	0	Q1	18	Q2	69	Q3	160	Maximum	737	Interquartile range		142	Lower Fence		-195	Upper Fence		373		Dari boxplot di samping, terlihat bahwa median waktu reservasi adalah \pm 69hari, sedangkan 25% dari pemesanan dilakukan kurang dari 20 hari dan 75% pemesanan dilakukan kurang dari 160 hari. Selain itu, terdapat banyak outliers di atas 400 hari, bahkan hingga lebih dari 700 hari.
Mean	104.011																														
Median	69																														
Modus	0																														
Five-number summary																															
Minimum	0																														
Q1	18																														
Q2	69																														
Q3	160																														
Maximum	737																														
Interquartile range		142																													
Lower Fence		-195																													
Upper Fence		373																													
4	adults	<table><tr><td>Mean</td><td>1.856</td></tr><tr><td>Median</td><td>2</td></tr><tr><td>Modus</td><td>2</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>2</td></tr><tr><td>Q2</td><td>2</td></tr><tr><td>Q3</td><td>2</td></tr><tr><td>Maximum</td><td>55</td></tr><tr><td colspan="2">Interquartile range</td><td>0</td></tr><tr><td colspan="2">Lower Fence</td><td>2</td></tr><tr><td colspan="2">Upper Fence</td><td>2</td></tr></table>	Mean	1.856	Median	2	Modus	2	Five-number summary		Minimum	0	Q1	2	Q2	2	Q3	2	Maximum	55	Interquartile range		0	Lower Fence		2	Upper Fence		2		Dari boxplot di samping, terlihat bahwa mayoritas tamu adalah berjumlah 2 orang dewasa. Namun, juga terdapat berbagai variasi jumlah tamu dewasa, baik kurang dari 2 dan lebih dari 2, bahkan ada yang hingga berjumlah 55 orang dewasa
Mean	1.856																														
Median	2																														
Modus	2																														
Five-number summary																															
Minimum	0																														
Q1	2																														
Q2	2																														
Q3	2																														
Maximum	55																														
Interquartile range		0																													
Lower Fence		2																													
Upper Fence		2																													

5	children	<table><tr><td>Mean</td><td>0.104</td></tr><tr><td>Median</td><td>0</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>0</td></tr><tr><td>Q2</td><td>0</td></tr><tr><td>Q3</td><td>0</td></tr><tr><td>Maximum</td><td>10</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>0</td></tr><tr><td>Upper Fence</td><td>0</td></tr></table>	Mean	0.104	Median	0	Modus	0	Five-number summary		Minimum	0	Q1	0	Q2	0	Q3	0	Maximum	10	Interquartile range		Lower Fence	0	Upper Fence	0	 <p>A box plot titled "Box Plot of Children". The y-axis is labeled "Children" and ranges from -0.04 to 0.04. The x-axis has a single category "1". The plot shows a box from 0 to 0 with a median line at 0. There are no whiskers or outliers visible.</p>	Dari boxplot di samping, terlihat bahwa mayoritas tamu adalah berjumlah 2 orang dewasa. Namun, juga terdapat berbagai variasi jumlah tamu dewasa, baik kurang dari 2 dan lebih dari 2, bahkan ada yang hingga berjumlah 55 orang dewasa
Mean	0.104																											
Median	0																											
Modus	0																											
Five-number summary																												
Minimum	0																											
Q1	0																											
Q2	0																											
Q3	0																											
Maximum	10																											
Interquartile range																												
Lower Fence	0																											
Upper Fence	0																											
6	babies	<table><tr><td>Mean</td><td>0.008</td></tr><tr><td>Median</td><td>0</td></tr><tr><td>Modus</td><td>0</td></tr><tr><td colspan="2">Five-number summary</td></tr><tr><td>Minimum</td><td>0</td></tr><tr><td>Q1</td><td>0</td></tr><tr><td>Q2</td><td>0</td></tr><tr><td>Q3</td><td>0</td></tr><tr><td>Maximum</td><td>10</td></tr><tr><td colspan="2">Interquartile range</td></tr><tr><td>Lower Fence</td><td>0</td></tr><tr><td>Upper Fence</td><td>0</td></tr></table>	Mean	0.008	Median	0	Modus	0	Five-number summary		Minimum	0	Q1	0	Q2	0	Q3	0	Maximum	10	Interquartile range		Lower Fence	0	Upper Fence	0	 <p>A box plot titled "Box Plot of Babies". The y-axis is labeled "Babies" and ranges from 0 to 10. The x-axis has a single category "1". The plot shows a box from 0 to 0 with a median line at 0. There are two outliers at approximately 1 and 2. There are also two points at approximately 9 and 10.</p>	Jumlah rata-rata tamu balita adalah 0, akan tetapi ada konsumen yang terdiri dari 10 tamu balita
Mean	0.008																											
Median	0																											
Modus	0																											
Five-number summary																												
Minimum	0																											
Q1	0																											
Q2	0																											
Q3	0																											
Maximum	10																											
Interquartile range																												
Lower Fence	0																											
Upper Fence	0																											

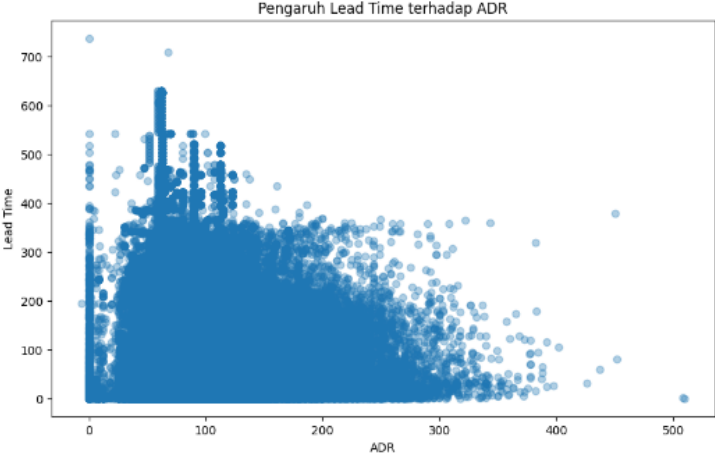
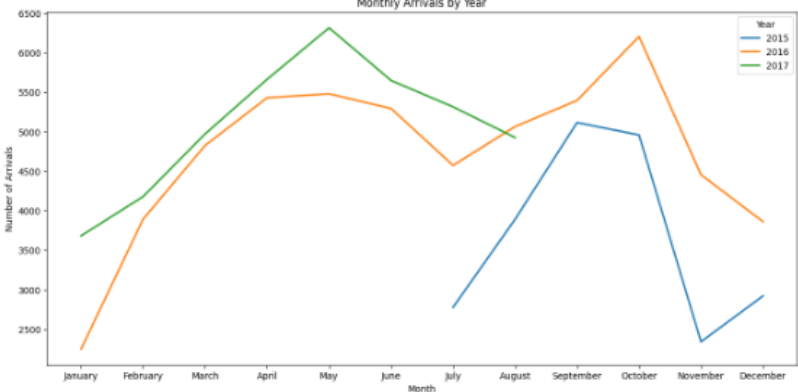
7	meal	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>meal_count</th><th>count</th></tr><tr><td>BB</td><td>92310</td></tr><tr><td>HB</td><td>14463</td></tr><tr><td>SC</td><td>10650</td></tr><tr><td>Undefined</td><td>1169</td></tr><tr><td>FB</td><td>798</td></tr><tr><td>Total</td><td>119390</td></tr><tr><th colspan="2">Proporsi</th></tr><tr><th>meal_count</th><th>proportion</th></tr><tr><td>BB</td><td>77.32%</td></tr><tr><td>HB</td><td>12.11%</td></tr><tr><td>SC</td><td>8.92%</td></tr><tr><td>Undefined</td><td>0.98%</td></tr><tr><td>FB</td><td>0.67%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		meal_count	count	BB	92310	HB	14463	SC	10650	Undefined	1169	FB	798	Total	119390	Proporsi		meal_count	proportion	BB	77.32%	HB	12.11%	SC	8.92%	Undefined	0.98%	FB	0.67%	Total	100.00%	<p>Distribusi Meal Plan</p>  <table><tr><th>Meal Plan</th><th>Count</th></tr><tr><td>BB</td><td>92310</td></tr><tr><td>HB</td><td>14463</td></tr><tr><td>SC</td><td>10650</td></tr><tr><td>Undefined</td><td>1169</td></tr><tr><td>FB</td><td>798</td></tr></table>	Meal Plan	Count	BB	92310	HB	14463	SC	10650	Undefined	1169	FB	798	Berdasarkan visualisasi di samping, terlihat bahwa meal plan BB merupakan meal plan yang paling banyak dipesan. 77.32% pemesanan meal plan merupakan meal plan BB.												
Frekuensi																																																												
meal_count	count																																																											
BB	92310																																																											
HB	14463																																																											
SC	10650																																																											
Undefined	1169																																																											
FB	798																																																											
Total	119390																																																											
Proporsi																																																												
meal_count	proportion																																																											
BB	77.32%																																																											
HB	12.11%																																																											
SC	8.92%																																																											
Undefined	0.98%																																																											
FB	0.67%																																																											
Total	100.00%																																																											
Meal Plan	Count																																																											
BB	92310																																																											
HB	14463																																																											
SC	10650																																																											
Undefined	1169																																																											
FB	798																																																											
8	country	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>country_count</th><th>count</th></tr><tr><td>PRT</td><td>48590</td></tr><tr><td>GBR</td><td>12129</td></tr><tr><td>FRA</td><td>10415</td></tr><tr><td>ESP</td><td>8568</td></tr><tr><td>DEU</td><td>7287</td></tr><tr><td>ITA</td><td>3766</td></tr><tr><td>IRL</td><td>3375</td></tr><tr><td>BEL</td><td>2342</td></tr><tr><td>BRA</td><td>2224</td></tr><tr><td>NLD</td><td>2104</td></tr><tr><td>USA</td><td>2097</td></tr><tr><td>Other</td><td>16493</td></tr><tr><td>Total</td><td>119390</td></tr></table>	Frekuensi		country_count	count	PRT	48590	GBR	12129	FRA	10415	ESP	8568	DEU	7287	ITA	3766	IRL	3375	BEL	2342	BRA	2224	NLD	2104	USA	2097	Other	16493	Total	119390	<p>Distribusi Negara (Top 11)</p>  <table><tr><th>Country</th><th>Count</th></tr><tr><td>PRT</td><td>48590</td></tr><tr><td>GBR</td><td>12129</td></tr><tr><td>FRA</td><td>10415</td></tr><tr><td>ESP</td><td>8568</td></tr><tr><td>DEU</td><td>7287</td></tr><tr><td>ITA</td><td>3766</td></tr><tr><td>IRL</td><td>3375</td></tr><tr><td>BEL</td><td>2342</td></tr><tr><td>BRA</td><td>2224</td></tr><tr><td>NLD</td><td>2104</td></tr><tr><td>USA</td><td>2097</td></tr><tr><td>Other</td><td>16493</td></tr></table>	Country	Count	PRT	48590	GBR	12129	FRA	10415	ESP	8568	DEU	7287	ITA	3766	IRL	3375	BEL	2342	BRA	2224	NLD	2104	USA	2097	Other	16493	Berdasarkan visualisasi di samping, terlihat bahwa 5 negara yang memiliki penjualan tertinggi berasal dari benua Eropa, yaitu Portugal, UK, Perancis, Spanyol, dan Jerman yang jika ditotal berjumlah 72.8%. Sedangkan negara-negara lain yang tidak berada dalam Top 11 memegang penjualan sebesar 13.81%
Frekuensi																																																												
country_count	count																																																											
PRT	48590																																																											
GBR	12129																																																											
FRA	10415																																																											
ESP	8568																																																											
DEU	7287																																																											
ITA	3766																																																											
IRL	3375																																																											
BEL	2342																																																											
BRA	2224																																																											
NLD	2104																																																											
USA	2097																																																											
Other	16493																																																											
Total	119390																																																											
Country	Count																																																											
PRT	48590																																																											
GBR	12129																																																											
FRA	10415																																																											
ESP	8568																																																											
DEU	7287																																																											
ITA	3766																																																											
IRL	3375																																																											
BEL	2342																																																											
BRA	2224																																																											
NLD	2104																																																											
USA	2097																																																											
Other	16493																																																											

9	deposit_type	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>deposit_type</td><td>count</td></tr><tr><td>No Deposit</td><td>104641</td></tr><tr><td>Non Refund</td><td>14587</td></tr><tr><td>Refundable</td><td>162</td></tr><tr><td>Total</td><td>119390</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>deposit_type</td><td>count</td></tr><tr><td>No Deposit</td><td>87.65%</td></tr><tr><td>Non Refund</td><td>12.22%</td></tr><tr><td>Refundable</td><td>0.14%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		deposit_type	count	No Deposit	104641	Non Refund	14587	Refundable	162	Total	119390	Proporsi		deposit_type	count	No Deposit	87.65%	Non Refund	12.22%	Refundable	0.14%	Total	100.00%	<p>Proporsi Tipe Deposit</p> <p>Detailed description: A pie chart titled 'Proporsi Tipe Deposit'. The largest slice, representing 'No Deposit', is light blue and accounts for 87.6% of the total. A smaller green slice represents 'Non Refund' at 12.2%. A very thin purple slice represents 'Refundable' at 0.1%.</p> <table><thead><tr><th>deposit_type</th><th>count</th><th>proporsi</th></tr></thead><tbody><tr><td>No Deposit</td><td>104641</td><td>87.6%</td></tr><tr><td>Non Refund</td><td>14587</td><td>12.2%</td></tr><tr><td>Refundable</td><td>162</td><td>0.1%</td></tr></tbody></table>	deposit_type	count	proporsi	No Deposit	104641	87.6%	Non Refund	14587	12.2%	Refundable	162	0.1%	<p>Dari pie chart di samping, dapat disimpulkan bahwa penjualan tanpa deposit merupakan tipe deposit yang paling banyak dijual dengan proporsi sebesar 87.6%. Sedangkan penjualan tanpa refund menduduki posisi kedua sebesar 12.2%, terakhir adalah penjualan yang dapat di refund namun hanya sebesar 0.1% dari keseluruhan penjualan</p>							
Frekuensi																																															
deposit_type	count																																														
No Deposit	104641																																														
Non Refund	14587																																														
Refundable	162																																														
Total	119390																																														
Proporsi																																															
deposit_type	count																																														
No Deposit	87.65%																																														
Non Refund	12.22%																																														
Refundable	0.14%																																														
Total	100.00%																																														
deposit_type	count	proporsi																																													
No Deposit	104641	87.6%																																													
Non Refund	14587	12.2%																																													
Refundable	162	0.1%																																													
10	customer_type	<table><tr><td colspan="2">Frekuensi</td></tr><tr><td>customer_type</td><td>count</td></tr><tr><td>Contract</td><td>4076</td></tr><tr><td>Group</td><td>577</td></tr><tr><td>Transient</td><td>89613</td></tr><tr><td>Transient-Party</td><td>25124</td></tr><tr><td>Total</td><td>119390</td></tr><tr><td colspan="2">Proporsi</td></tr><tr><td>customer_type</td><td>count</td></tr><tr><td>Contract</td><td>3.41%</td></tr><tr><td>Group</td><td>0.48%</td></tr><tr><td>Transient</td><td>75.06%</td></tr><tr><td>Transient-Party</td><td>21.04%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		customer_type	count	Contract	4076	Group	577	Transient	89613	Transient-Party	25124	Total	119390	Proporsi		customer_type	count	Contract	3.41%	Group	0.48%	Transient	75.06%	Transient-Party	21.04%	Total	100.00%	<p>Distribusi Tipe Kustomer</p> <p>Detailed description: A bar chart titled 'Distribusi Tipe Kustomer'. The x-axis is labeled 'customer_type' and has four categories: Contract, Group, Transient, and Transient-Party. The y-axis represents the count, ranging from 0 to 100,000. The 'Transient' bar is the tallest, reaching approximately 89,613. The 'Transient-Party' bar is the second tallest, reaching approximately 25,124. The 'Contract' bar is much shorter, reaching approximately 4,076. The 'Group' bar is the shortest, reaching approximately 577.</p> <table><thead><tr><th>customer_type</th><th>count</th><th>proporsi</th></tr></thead><tbody><tr><td>Contract</td><td>4076</td><td>3.41%</td></tr><tr><td>Group</td><td>577</td><td>0.48%</td></tr><tr><td>Transient</td><td>89613</td><td>75.06%</td></tr><tr><td>Transient-Party</td><td>25124</td><td>21.04%</td></tr></tbody></table>	customer_type	count	proporsi	Contract	4076	3.41%	Group	577	0.48%	Transient	89613	75.06%	Transient-Party	25124	21.04%	<p>Dari diagram batang di samping, dapat dilihat bahwa tipe konsumen yang paling banyak adalah Transient dengan proporsi sebesar 75.06%, sedangkan di posisi kedua adalah konsumen tipe Transient-Party dengan proporsi 21.04%. Sedangkan tipe contract dan grup hanya berjumlah 3.41% dan 0.48%</p>
Frekuensi																																															
customer_type	count																																														
Contract	4076																																														
Group	577																																														
Transient	89613																																														
Transient-Party	25124																																														
Total	119390																																														
Proporsi																																															
customer_type	count																																														
Contract	3.41%																																														
Group	0.48%																																														
Transient	75.06%																																														
Transient-Party	21.04%																																														
Total	100.00%																																														
customer_type	count	proporsi																																													
Contract	4076	3.41%																																													
Group	577	0.48%																																													
Transient	89613	75.06%																																													
Transient-Party	25124	21.04%																																													

11	distribution_channel	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>distribution_channel</th><th>count</th></tr><tr><td>Corporate</td><td>6677</td></tr><tr><td>Direct</td><td>14645</td></tr><tr><td>GDS</td><td>193</td></tr><tr><td>TA/TO</td><td>97870</td></tr><tr><td>Undefined</td><td>5</td></tr><tr><td>Total</td><td>119390</td></tr><tr><th colspan="2">Proporsi</th></tr><tr><th>distribution_channel</th><th>count</th></tr><tr><td>Corporate</td><td>5.59%</td></tr><tr><td>Direct</td><td>12.27%</td></tr><tr><td>GDS</td><td>0.16%</td></tr><tr><td>TA/TO</td><td>81.98%</td></tr><tr><td>Undefined</td><td>0.00%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		distribution_channel	count	Corporate	6677	Direct	14645	GDS	193	TA/TO	97870	Undefined	5	Total	119390	Proporsi		distribution_channel	count	Corporate	5.59%	Direct	12.27%	GDS	0.16%	TA/TO	81.98%	Undefined	0.00%	Total	100.00%	<p>Distribusi Channel</p>	Dari diagram batang di samping, dapat disimpulkan bahwa distribution Travel Agent baik online dan offline memiliki market terbesar, disusul pemesanan langsung ke hotel, dan melalui corporate
Frekuensi																																				
distribution_channel	count																																			
Corporate	6677																																			
Direct	14645																																			
GDS	193																																			
TA/TO	97870																																			
Undefined	5																																			
Total	119390																																			
Proporsi																																				
distribution_channel	count																																			
Corporate	5.59%																																			
Direct	12.27%																																			
GDS	0.16%																																			
TA/TO	81.98%																																			
Undefined	0.00%																																			
Total	100.00%																																			
12	reserved_room_type	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>reserved_room_type</th><th>count</th></tr><tr><td>A</td><td>85994</td></tr><tr><td>B</td><td>1118</td></tr><tr><td>C</td><td>932</td></tr><tr><td>D</td><td>19201</td></tr><tr><td>E</td><td>6535</td></tr><tr><td>F</td><td>2897</td></tr><tr><td>G</td><td>2094</td></tr><tr><td>H</td><td>601</td></tr><tr><td>L</td><td>6</td></tr><tr><td>P</td><td>12</td></tr><tr><td>Total</td><td>119390</td></tr></table>	Frekuensi		reserved_room_type	count	A	85994	B	1118	C	932	D	19201	E	6535	F	2897	G	2094	H	601	L	6	P	12	Total	119390	<p>Distribusi Tipe Kamar yang Direservasi</p>	Berdasarkan visualisasi di samping, terlihat bahwa kamar tipe A paling diminati konsumen, disusul kamar tipe D, E, F, dan G (5 tipe kamar penjualan tertinggi)						
Frekuensi																																				
reserved_room_type	count																																			
A	85994																																			
B	1118																																			
C	932																																			
D	19201																																			
E	6535																																			
F	2897																																			
G	2094																																			
H	601																																			
L	6																																			
P	12																																			
Total	119390																																			

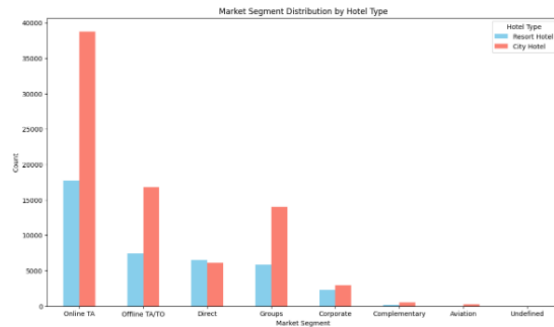
13	assigned_room_type	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>reserved_room_type</th><th>count</th></tr><tr><td>A</td><td>74053</td></tr><tr><td>B</td><td>2163</td></tr><tr><td>C</td><td>2375</td></tr><tr><td>D</td><td>25322</td></tr><tr><td>E</td><td>7806</td></tr><tr><td>F</td><td>3751</td></tr><tr><td>G</td><td>2553</td></tr><tr><td>H</td><td>712</td></tr><tr><td>I</td><td>363</td></tr><tr><td>K</td><td>279</td></tr><tr><td>L</td><td>1</td></tr><tr><td>P</td><td>12</td></tr><tr><td>Total</td><td>119390</td></tr></table>	Frekuensi		reserved_room_type	count	A	74053	B	2163	C	2375	D	25322	E	7806	F	3751	G	2553	H	712	I	363	K	279	L	1	P	12	Total	119390	<p>Distribusi Tipe Kamar yang Dialokasikan</p>  <table><caption>Data for Distribusi Tipe Kamar yang Dialokasikan</caption><thead><tr><th>customer_type</th><th>count</th></tr></thead><tbody><tr><td>A</td><td>74053</td></tr><tr><td>B</td><td>2163</td></tr><tr><td>C</td><td>2375</td></tr><tr><td>D</td><td>25322</td></tr><tr><td>E</td><td>7806</td></tr><tr><td>F</td><td>3751</td></tr><tr><td>G</td><td>2553</td></tr><tr><td>H</td><td>712</td></tr><tr><td>I</td><td>363</td></tr><tr><td>K</td><td>279</td></tr><tr><td>L</td><td>1</td></tr><tr><td>P</td><td>12</td></tr></tbody></table>	customer_type	count	A	74053	B	2163	C	2375	D	25322	E	7806	F	3751	G	2553	H	712	I	363	K	279	L	1	P	12	Berdasarkan visualisasi di samping, terlihat bahwa kamar tipe A paling sering ditempati, disusul kamar tipe D, E, F, dan G (5 tipe kamar penjualan tertinggi)
Frekuensi																																																												
reserved_room_type	count																																																											
A	74053																																																											
B	2163																																																											
C	2375																																																											
D	25322																																																											
E	7806																																																											
F	3751																																																											
G	2553																																																											
H	712																																																											
I	363																																																											
K	279																																																											
L	1																																																											
P	12																																																											
Total	119390																																																											
customer_type	count																																																											
A	74053																																																											
B	2163																																																											
C	2375																																																											
D	25322																																																											
E	7806																																																											
F	3751																																																											
G	2553																																																											
H	712																																																											
I	363																																																											
K	279																																																											
L	1																																																											
P	12																																																											
14	total_of_special_requests	<table><tr><th colspan="2">Frekuensi</th></tr><tr><th>total_of_special_request</th><th>count</th></tr><tr><td>0</td><td>70318</td></tr><tr><td>1</td><td>33226</td></tr><tr><td>2</td><td>12969</td></tr><tr><td>3</td><td>2497</td></tr><tr><td>4</td><td>340</td></tr><tr><td>5</td><td>40</td></tr><tr><td>Total</td><td>119390</td></tr></table> <table><tr><th colspan="2">Proporsi</th></tr><tr><th>distribution_channel</th><th>count</th></tr><tr><td>0</td><td>58.90%</td></tr><tr><td>1</td><td>27.83%</td></tr><tr><td>2</td><td>10.86%</td></tr><tr><td>3</td><td>2.09%</td></tr><tr><td>4</td><td>0.28%</td></tr><tr><td>5</td><td>0.03%</td></tr><tr><td>Total</td><td>100.00%</td></tr></table>	Frekuensi		total_of_special_request	count	0	70318	1	33226	2	12969	3	2497	4	340	5	40	Total	119390	Proporsi		distribution_channel	count	0	58.90%	1	27.83%	2	10.86%	3	2.09%	4	0.28%	5	0.03%	Total	100.00%	<p>Distribusi Jumlah Special Request</p>  <table><caption>Data for Distribusi Jumlah Special Request</caption><thead><tr><th>count</th><th>percentage</th></tr></thead><tbody><tr><td>0</td><td>58.90%</td></tr><tr><td>1</td><td>27.83%</td></tr><tr><td>2</td><td>10.86%</td></tr><tr><td>3</td><td>2.09%</td></tr><tr><td>4</td><td>0.28%</td></tr><tr><td>5</td><td>0.03%</td></tr></tbody></table>	count	percentage	0	58.90%	1	27.83%	2	10.86%	3	2.09%	4	0.28%	5	0.03%	Berdasarkan visualisasi di samping, terlihat bahwa pemesanan tanpa special request memiliki jumlah tertinggi, disusul dengan satu, dua, dan tiga special request secara berurutan						
Frekuensi																																																												
total_of_special_request	count																																																											
0	70318																																																											
1	33226																																																											
2	12969																																																											
3	2497																																																											
4	340																																																											
5	40																																																											
Total	119390																																																											
Proporsi																																																												
distribution_channel	count																																																											
0	58.90%																																																											
1	27.83%																																																											
2	10.86%																																																											
3	2.09%																																																											
4	0.28%																																																											
5	0.03%																																																											
Total	100.00%																																																											
count	percentage																																																											
0	58.90%																																																											
1	27.83%																																																											
2	10.86%																																																											
3	2.09%																																																											
4	0.28%																																																											
5	0.03%																																																											

○ Pola Data

No	Pola Data	Analisis
1		<p>Berdasarkan visualisasi di samping, terlihat bahwa semakin rendah lead time mempengaruhi ADR yang semakin tinggi. Sedangkan semakin tinggi leadtime, maka ADR menjadi semakin rendah (terpusat di rentang 20 hingga 150)</p>
2		<p>Berdasarkan visualisasi di samping, terlihat bahwa penjualan di setiap tahun cenderung memiliki pola yang sama. Di tahun 2016 dan 2017, penjualan terendah terjadi di bulan Januari, namun tahun 2016 memiliki penjualan yang lebih rendah. Selain itu, Pada bulan Juli terjadi penurunan di setiap tahun dengan penjualan terendah terjadi di tahun 2015. Pola yang sama terjadi di bulan Oktober dan November yaitu terjadi penurunan penjualan baik di tahun 2015 dan 2016.</p>

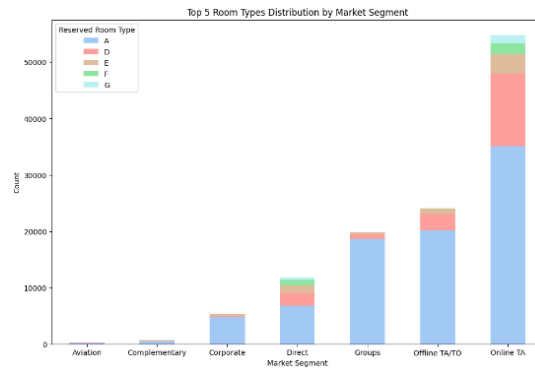
3		<p>City Hotel: Tipe hotel ini menunjukkan peningkatan yang signifikan dari Januari hingga Juni, lalu mengalami penurunan pada bulan Juli dan kembali meningkat hingga mencapai puncaknya pada bulan Oktober, diikuti oleh penurunan tajam hingga Desember. Oleh karena itu, perlu dicari penyebab mengapa hal berikut dapat terjadi (seperti musim, liburan dsb)</p> <p>Resort Hotel: Tipe hotel ini menunjukkan fluktuasi yang lebih rendah dan relatif lebih stabil sepanjang tahun, meskipun ada puncak kecil pada bulan April dan Oktober. Sama seperti City Hotel, penjualan terendah terjadi di bulan Januari</p>
4		<p>City Hotel memiliki tingkat pembatalan yang paling tinggi dengan jumlah 33.102 pembatalan dan probabilitas tingkat pembatalan sebesar 41.7%. Sedangkan City Hotel memiliki probabilitas pembatalan sebesar 27.7% dengan jumlah pembatalan sebesar 11.122</p>
5		<p>City Hotel: Memiliki distribusi yang lebih luas dan merata, mayoritas pemesanan terpusat pada lead_time yang lebih rendah hingga menengah, namun terdapat beberapa pemesanan yang menyebar hingga 700 hari</p> <p>Resort Hotel: Memiliki distribusi yang lebih padat pada lead_time yang lebih pendek dibandingkan City Hotel, hanya sedikit pemesanan yang menyebar hingga 500 hari atau lebih. Tamu memiliki kecenderungan memesan dalam rentang waktu lebih dekat</p>

6



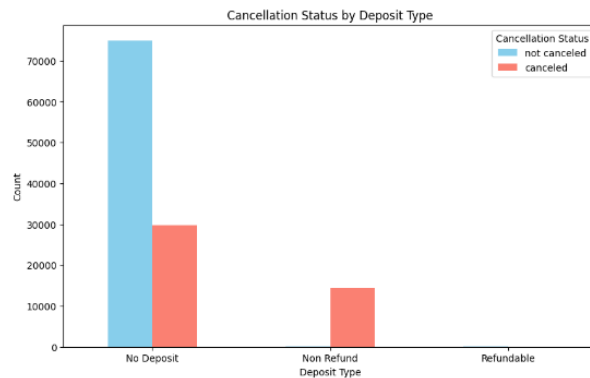
Berdasarkan visualisasi disamping, hampir seluruh market segment lebih memilih untuk melakukan reservasi di city hotel dibandingkan dengan resort hotel. Hanya ada 1 market segment yang lebih memilih untuk melakukan reservasi di resort hotel yaitu market segment direct. Walaupun lebih unggul pemesanan resort hotel, keunggulan pemesanan resort hotel dapat dianggap sedikit karena jumlah perbandingan antar pemesana resort hotel dan city hotel kecil.

7



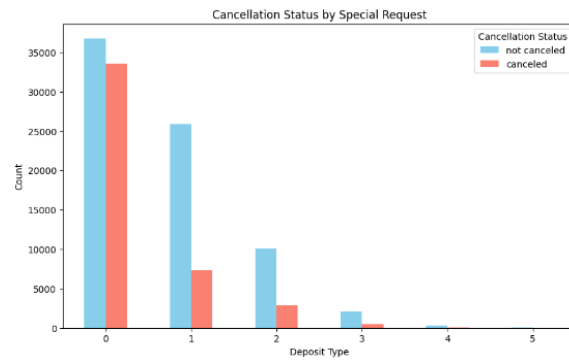
Dari keseluruhan segmen pasar, tipe kamar yang paling sering dipesan (diminati) adalah kamar Tipe A, disusul kamar tipe D di posisi kedua. Pada segmen lain seperti Direct dan Online TA, kamar Tipe E juga merupakan kamar yang paling diminati dengan menduduki peringkat ke-3

8



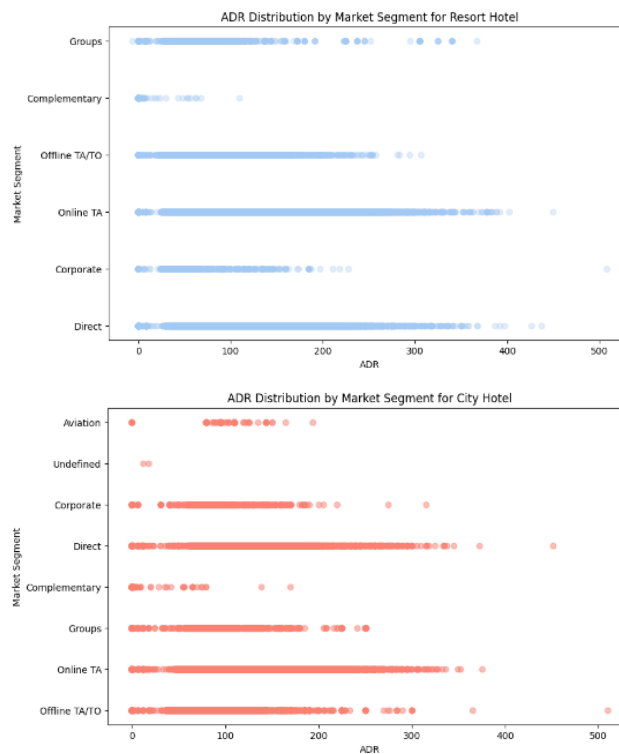
Berdasarkan visualisasi di samping, dapat disimpulkan bahwa pemesanan dengan tidak deposit cenderung memiliki cancellation yang lebih besar, dibanding dengan tipe deposit yang lain. Dikarenakan tamu merasa mereka sudah melakukan pembelian dan memiliki rasa tidak ingin rugi, sehingga tindakan cancellation jauh lebih kecil selain tipe deposit 'No Deposit'

9



Berdasarkan visualisasi di samping, dapat disimpulkan bahwa Special Request memiliki pengaruh yang signifikan terhadap status pembatalan reservasi. Konsumen yang memesan kamar tanpa special request, memiliki persentase sebesar 91% dengan jumlah pembatalan sebesar 33.556. Sedangkan konsumen yang memesan kamar dengan minimal satu buah special request, memiliki peluang yang jauh lebih kecil yaitu hanya sebesar 28.2%

10



Berdasarkan visualisasi di samping, tipe hotel Resort tidak memiliki segmen market aviation dan segmen yang tidak diketahui (undefined). Secara keseluruhan, hampir semua segmen seperti Direct, Online TA, TO, dan Group pada Resort Hotel memiliki distribusi harga yang lebih luas/lebih mahal dibandingkan City Hotel. Sedangkan corporate dan complementray, distribusi ADR justru lebih tinggi pada City Hotel

Catatan: Eksplorasi detail dapat dilihat pada file Tugas 3 versi final di folder UASDSDS_03

- Enam model yang sudah dibuat dari dataset Hotel Booking

No	Nama File	Pengukuran	Informasi yang diperoleh dari model
1	M1_Classification_Forest_Weekday_03.py	Akurasi = 66%	fitur stays in week dan stays in weekend dapat melakukan prediksi terhadap pembatalan reservasi dengan tingkat keberhasilan sebesar 66%. Dengan memasukan data stays in week dan stays in weekend ke dalam model, dapat diprediksi apakah reservasi tersebut akan dibatalkan atau tidak.
2	M2_Classification_NeuralNetwork_LeadTime_03.py	Akurasi = 75.11%	fitur lead time , deposit type , dan reserved room type , dapat melakukan prediksi terhadap pembatalan reservasi dengan tingkat keberhasilan sebesar 75.1%. Dengan memasukan data lead time, deposit type, dan reserved room type, akan diprediksi melalui data yang tersedia apakah reservasi tersebut akan dibatalkan atau tidak.
3	M3_Classification_NeuralNetwork_Segment_03	Akurasi = 68%	Fitur market segment , adults , dan arrival date month dapat melakukan prediksi terhadap pembatalan reservasi dengan tingkat keberhasilan sebesar 68%. Dengan memasukan fitur tersebut, akan diprediksi melalui data yang tersedia apakah reservasi akan dibatalkan atau tidak.
4	M4_Pivot_TrendMonth_03.py	-	City Hotel jumlah tamu yang lebih banyak dari Resort hotel. Di Bulan Juni, City Hotel mengalami kenaikan sedangkan Resort Hotel mengalami penurunan. Pada bulan Desember-Januari kedua tipe Hotel mengalami penjualan terendah
5	M5_Pivot_SpecialRequest_03.py	-	Special request memiliki pengaruh yang signifikan terhadap status pembatalan reservasi. Konsumen yang memesan kamar tanpa special request, memiliki persentase pembatalan 48%, sedangkan konsumen yang memesan kamar dengan minimal satu buah special request hanya sebesar 28.2%
6	M6_Pivot_RoomType_03.py	-	Dari visualiasi tersebut, dapat dilihat tipe-tipe kamar yang memiliki penjualan tertinggi di setiap bulannya

- Empat dari enam model terbaik yang diusulkan beserta alasannya (lihat dari tugas 4 yang sudah diperbaiki):

No	Nama File	Pengukuran	Informasi yang diperoleh dari model
1	M1_Classification_Forest_Weekday_03.py	Akurasi = 66%	fitur stays in week dan stays in weekend dapat melakukan prediksi terhadap pembatalan reservasi dengan tingkat keberhasilan sebesar 66%. Dengan memasukan data stays in week dan stays in weekend ke dalam model, dapat diprediksi apakah reservasi tersebut akan dibatalkan atau tidak.
2	M2_Classification_NeuralNetwork_LeadTime_03.py	Akurasi = 75.11%	fitur lead time , deposit type , dan reserved room type , dapat melakukan prediksi terhadap pembatalan reservasi dengan tingkat keberhasilan sebesar 75.1%. Dengan memasukan data lead time, deposit type, dan reserved room type, akan diprediksi melalui data yang tersedia apakah reservasi tersebut akan dibatalkan atau tidak.
3	M4_Pivot_TrendMonth_03.py	-	City Hotel jumlah tamu yang lebih banyak dari Resort hotel. Di Bulan Juni, City Hotel mengalami kenaikan sedangkan Resort Hotel mengalami penurunan. Pada bulan Desember-Januari kedua tipe Hotel mengalami penjualan terendah
4	M5_Pivot_SpecialRequest_03.py	-	Special request memiliki pengaruh yang signifikan terhadap status pembatalan reservasi. Konsumen yang memesan kamar tanpa special request, memiliki persentase pembatalan 48%, sedangkan konsumen yang memesan kamar dengan minimal satu buah special request hanya sebesar 28.2%

Catatan: keseluruhan model yang sudah dibuat dapat dilihat pada folder tugas 4 versi final di folder UASDSDS_03

- Insights yang diperoleh dari 4 Model di atas beserta actionable insights

Insight Model 1:

Berdasarkan visualisasi di atas, City Hotel cenderung memiliki angka kedatangan tamu yang lebih banyak dari Resort hotel. Di Bulan Juni, City Hotel mengalami kenaikan sedangkan Resort Hotel mengalami penurunan. Sedangkan bulan Desember-Januari kedua tipe Hotel mengalami penurunan penjualan terendah

Actionable insight Model 1:

Penyesuaian Strategi Penjualan: Perlu diperhatikan faktor yang menyebabkan terjadinya penjualan yang rendah di kuartal pertama (seperti faktor liburan, musim, dsb) sehingga dapat mengoptimalkan strategi pemasaran. Perusahaan dapat mengoptimalkan strategi penjualan dengan memanfaatkan puncak penjualan di bulan Juni dan Oktober, seperti menawarkan promo atau paket khusus jika menginap di bulan-bulan berikutnya.

Insight Model 2:

Berdasarkan visualisasi di atas, special request memiliki pengaruh yang signifikan terhadap status pembatalan reservasi. Konsumen yang memesan kamar tanpa special request, memiliki persentase sebesar 48% dengan jumlah pembatalan sebesar 33.556. Sedangkan konsumen yang memesan kamar dengan minimal satu buah special request hanya sebesar 28.2%

Actionable insight Model 2:

Mendorong tamu untuk menambahkan special requests saat melakukan reservasi dengan menawarkan keuntungan. Misalnya dengan menawarkan paket bundling yang mencakup special request sebagai bagian dari penawaran atau mengirim email yang dipersonalisasi setelah adanya reservasi untuk menawarkan special request.

Insight Model 3:

Jumlah reservasi di weekday dan weekend memiliki pengaruh terhadap dibatalkan atau tidaknya sebuah reservasi. Reservasi dengan jumlah pemesanan pada weekday dan weekend di rentang 0-10 hari cenderung untuk tidak dibatalkan. Dan untuk rentang 10-25 hari cenderung untuk dibatalkan. Dan pemesanan diatas 25 hari cenderung untuk tidak dibatalkan.

Actionable insight Model 3:

Memberikan fleksibilitas kebijakan pembatalan untuk reservasi dengan durasi 10-25 hari untuk mengurangi pembatalan, seperti memberikan diskon tambahan atau fasilitas gratis bagi tamu yang melakukan reservasi dalam rentang waktu tersebut. Dapat juga dilakukan follow up untuk reservasi 10-25 hari dengan harapan ketika dihubungi tamu dapat memastikan reservasinya dan tidak dibatalkan.

Insight Model 4:




Jika menggunakan fitur lead time, deposit type, dan reserved room type secara bersamaan, maka dapat dilakukan prediksi terhadap pembatalan sebuah kamar dengan tingkat keberhasilan sebesar 75.1%

Actionable insight Model 4:

Dengan adanya machine learning, pemilik hotel dapat melakukan tindakan preventif jika ada konsumen yang diduga kuat akan membatalkan pesanan melalui fitur lead time, deposit type, dan reserved room type secara bersamaan. Setelah dicek menggunakan machine learning dan hasilnya pelanggan diduga akan membatalkan, maka hotel dapat menghubungi agar dipastikan reservasi yang dibuat dibatalkan atau tidak.

V. Lampiran

- Tabel kontribusi tugas akhir

No	NPM	Nama	Kontribusi Pekerjaan (%)	Tanda Tangan	Kelas
1	6182001002	Erwin Darsono	100	 Erwin Darsono	A
2	6182001010	Neil Christopher	100		
3	6182001039	Wilson	100		

- Lampiran script coding terlampir di dalam folder folder UASDSDS_03

VI. Referensi

LSPR Communication and Business Institute. (n.d.). Jenis-jenis Pariwisata. Retrieved from <https://www.lspr.ac.id/jenis-jenis-pariwisata>

Yosephine, Y. M. (2012). Peranan Dinas Pariwisata dan Kebudayaan dalam Pengembangan Objek Wisata Kabupaten Samosir (Bachelor's thesis, Universitas Atma Jaya Yogyakarta). Retrieved from <https://e-journal.uajy.ac.id/2412/3/2TA12294.pdf>

Subakti, A. G. (2016, August 1). Hotel dan Sejarahnya. Hotel Management BINUS. Retrieved from <https://hotel-management.binus.ac.id/2016/08/01/hotel-dan-sejarahnya/>

Niko, I. (2024, March 5) Technology Insights in The Hotel Industry. Retrieved from Slide Kuliah Data Sciene pada Domain Spesifik