

Data Analytics for Business 2024

Project Based Learning

Study Case

[RedDoorz]

CS 17 - 01 :

- | | | |
|----|-------------------------------|------------|
| 1. | Aurelia Andry Xaviera - Ketua | KM-CS17038 |
| 2. | Najwa Safinabudhi - Anggota | KM-CS17059 |
| 3. | Koko - Anggota | KM-CS17185 |

Kata Pengantar

Puji syukur kami panjatkan kehadiran Allah SWT berkat rahmat dan karunia-Nya kami dapat menyelesaikan tugas laporan *Project Based Learning* ini dengan baik. Proyek ini merupakan bagian dari studi kasus perusahaan RedDoorz yang berfokus untuk meningkatkan wawasan dalam membuat keputusan berbasis data dan mengidentifikasi solusi atas tantangan bisnis yang dihadapi.

Dalam penyusunan laporan ini, kami tidak lepas dari pihak-pihak yang telah memberikan bantuan, bimbingan, dan dukungan dari awal hingga laporan dapat terselesaikan dengan baik. Oleh sebab itu, kami mengucapkan terima kasih kepada perusahaan RedDoorz atas data dan studi kasus yang memberikan kami pengalaman belajar yang berharga mengenai bidang analisis data, Kak Chlaudiah selaku mentor yang telah memberikan arahan, saran, dan wawasan yang berharga selama proses pengerjaan proyek, Kak Hani selaku *student success* yang mendampingi kami selama program berlangsung, kedua orangtua kami yang senantiasa memberikan doa dan dukungan yang menjadi semangat untuk terus belajar dan berkembang, serta teman-teman atas kerja sama dan menjadi rekan diskusi selama proses pengerjaan proyek ini, serta seluruh pihak yang tidak dapat kami sebutkan satu per satu tetapi turut memberikan kontribusi baik secara langsung maupun tidak langsung dalam penyelesaian laporan ini.

Laporan ini disusun sebagai bagian dari persyaratan dalam program Studi Independen. Laporan ini membahas mengenai tantangan yang dihadapi oleh perusahaan bisnis dan mengidentifikasi *insight* strategis menggunakan Python, SQL, dan Tableau yang dapat digunakan untuk merekomendasikan promosi guna mengoptimalkan kinerja bisnis secara keseluruhan.

Kami menyadari bahwa laporan ini masih terdapat kekurangan sehingga kami terbuka dalam menerima kritik dan saran dari pembaca untuk perbaikan di masa depan. Kami berharap laporan ini dapat bermanfaat, tidak hanya bagi kami sebagai penyusun, tetapi juga bagi pihak lain yang membacanya.

Daftar Isi

Daftar Isi

Kata Pengantar	2
BAB I	4
A. Latar Belakang	4
B. Identifikasi Masalah	6
BAB II	7
A. Data dan Sumber Data	7
B. Metode Analisis Data	7
1. Alat Analisis Data	7
2. Proses Analisis Data	7
BAB III	9
A. Gambaran Umum Objek Penelitian	9
B. Temuan Hasil Penelitian	9
C. Pembahasan	31
BAB IV	34
A. Simpulan	34
B. Saran	34
Lampiran	35

BAB I

Pendahuluan

A. Latar Belakang

RedDoorz adalah perusahaan yang bergerak di bidang perhotelan berbasis teknologi yang menawarkan akomodasi berkualitas dengan harga yang terjangkau. RedDoorz berdiri pada tahun 2015 di Singapura dan memasuki Indonesia sebagai pasar pertama dan terus berkembang di berbagai kota di Asia Tenggara, seperti Singapura, Filipina, hingga Vietnam. Saat ini, RedDoorz mengelola lebih dari 1.500 properti yang tersebar di berbagai kota di Indonesia, Singapura, Filipina, dan Vietnam dengan rencana ekspansi berkelanjutan untuk terus memperluas jangkauannya di bidang ini.

RedDoorz sebagai perusahaan di bidang teknologi perhotelan (*hospitality-tech company*) menghadapi tantangan dalam mengambil keputusan berbasis data. Tantangan ini menyebabkan kesulitan dalam mengoptimalkan operasional perusahaan dan meningkatkan kepuasan pelanggan sehingga diperlukan pemahaman yang lebih mendalam mengenai performa bisnis, perilaku pelanggan, dan tren industri untuk mengatasi hambatan ini.

Dataset yang digunakan untuk menganalisis performa bisnis, perilaku pelanggan, dan tren industri terdiri dari tiga data utama, yaitu data transaksi, data pelanggan, dan data properti.

- Data transaksi: data ini berisi informasi, seperti *series_date*, *revenue*, *room/nights (RNs)*, *occupancy*, dan *Average Daily Rate* yang digunakan untuk membantu dalam memahami tren pendapatan, tingkat hunian, dan tarif rata-rata di setiap properti.
- Data pelanggan: data ini berisi informasi, seperti *new user*, *repeat user*, *gender*, *business/leisure*, dan *age*. Data ini membantu dalam menganalisis segmentasi pelanggan berdasarkan perilaku, preferensi, dan demografi.
- Data properti: data ini berisi informasi, seperti *city*, *district*, *inventory*, *date joined*, dan *brand type*. Data ini membantu dalam mengidentifikasi performa properti di berbagai lokasi dan kategori merek serta memberikan wawasan tentang ekspansi bisnis dan segmentasi properti.

Melalui tantangan yang ada, masalah-masalah bisnis yang harus diselesaikan dan akan menjadi fokus utama dalam analisis ini adalah mengoptimalkan kegiatan operasional dengan melakukan analisis menggunakan data, meningkatkan kepuasan pelanggan dengan mempelajari perilaku dan preferensi mereka, dan mengidentifikasi tren industri untuk merumuskan strategi bisnis lebih lanjut. Tahapan-tahapan analisis data yang dilakukan untuk menyelesaikan masalah bisnis ini adalah Business Modeling, Data Preparation, Exploratory Data Analysis (EDA), dan Data Visualization. Melalui tahapan-tahapan ini, analisis data akan menghasilkan *output*, antara lain:

- Distribusi *stay duration* dan *lead time*.
- Pelanggan-pelanggan dengan *total booking revenue* tertinggi.
- Tingkat hunian properti dengan performa terbaik.
- Korelasi antara *occupancy rate* dan *total revenue per property*.
- Analisis *revenue*, *room nights* dan ADR per kota setiap bulannya di tahun 2024.
- Analisis pertumbuhan *room nights* tahun 2023 dan 2024.
- Segmentasi pelanggan berdasarkan perilaku pembelian (frekuensi pemesanan, demografi, dan tujuan perjalanan).
- Analisis strategi harga normal dan musim puncak tahun 2023 dan 2024.
- Penilaian kualitas properti berdasarkan *revenue* dan *occupancy rate*.
- SQL Views untuk menyoroti *key metrics* selama enam bulan terakhir di tingkat kota.
- Dashboard visualisasi interaktif menggunakan Tableau beserta rekomendasi rencana tindakan untuk area-area yang memerlukan perbaikan dalam *sales*, *pricing*, dan *property acquisition*, serta strategi *marketing campaign* yang dapat mendukung pertumbuhan dan efisiensi operasional perusahaan.

B. Identifikasi Masalah

Masalah-masalah yang dihadapi RedDoorz sebagai perusahaan *hospitality-tech* adalah sebagai berikut:

Tabel 1. Identifikasi Masalah Bisnis

Masalah	Akibat
Rendahnya tingkat retensi pelanggan	Pelanggan tidak memiliki keinginan untuk melakukan pemesanan kembali sehingga kontribusi pendapatan jangka panjang kurang maksimal.
Kurangnya pemahaman dalam mengevaluasi tren <i>seasonal</i> .	Kehilangan peluang untuk memaksimalkan pendapatan pada <i>peak season</i> .
Kurangnya pemahaman dalam memahami segmen pelanggan.	Tidak dapat membuat strategi promosi sesuai dengan segmentasi berdasarkan perilaku pembelian pelanggan.
Minimnya pemahaman dalam upaya meningkatkan kualitas properti.	Menurunnya tingkat kepuasan pelanggan akibat kurang optimalnya perawatan properti.

Berdasarkan tabel di atas, masalah-masalah yang teridentifikasi memberikan dampak yang signifikan terhadap pengoptimalan operasional perusahaan dan kepuasan pelanggan sehingga diperlukan wawasan yang lebih baik dalam menganalisis data dan memberikan strategi yang tepat sesuai data untuk mengembangkan bisnis RedDoorz. Masalah-masalah seperti rendahnya retensi pelanggan dalam pemesanan, penetapan harga sesuai dengan tren *seasonal*, minimnya pemahaman dalam upaya meningkatkan kualitas properti, dan kurangnya strategi pemasaran yang sesuai dengan segmentasi pelanggan; misalnya antara segmen *business* dan *leisure*, akan membuat perusahaan tidak dapat memaksimalkan peluang untuk meningkatkan pendapatan. Melalui peningkatan wawasan dalam meningkatkan keputusan berbasis data dan perbaikan pada masalah-masalah yang ada, perusahaan diharapkan dapat mengoptimalkan operasional perusahaan dan meningkatkan kepuasan pelanggan sehingga dapat memberikan dampak positif terhadap pendapatan perusahaan.

BAB II

Metodologi

A. Data dan Sumber Data

Jenis data yang menjadi fokus penelitian adalah data transaksi yang digunakan untuk menganalisis kinerja keuangan dan dan operasional properti, data pelanggan yang digunakan untuk melihat data demografi dan perilaku pelanggan, serta data properti yang berisi data lokasi, jenis properti, dan kapasitas yang dikelola oleh perusahaan. Data-data ini diperoleh langsung dari pihak RedDoorz dan Bitlabs.

B. Metode Analisis Data

1. Alat Analisis Data

Perangkat lunak atau *tools* yang digunakan untuk memproses dan menganalisis data adalah Google Colab dengan Python dan SQL DBeaver serta *software* berbasis visualisasi menggunakan Tableau. Penggunaan alat lain seperti Online Visual Paradigm untuk pembuatan BPMN dan Canva untuk mendesain *data product canvas* dan *slides* sebagai media presentasi.

Alasan pemilihan alat-alat ini adalah kemudahan penggunaan karena Google Colab, DBeaver, dan Tableau merupakan alat yang telah dipelajari dan digunakan selama program studi independen di Bitlabs sehingga penggunaannya akan lebih optimal dan efisien. Selain itu, Online Visual Paradigm dan Canva merupakan alat yang telah familiar digunakan dalam memodelkan proses bisnis dan menyajikan hasil analisis dengan format yang menarik dan mudah dipahami.

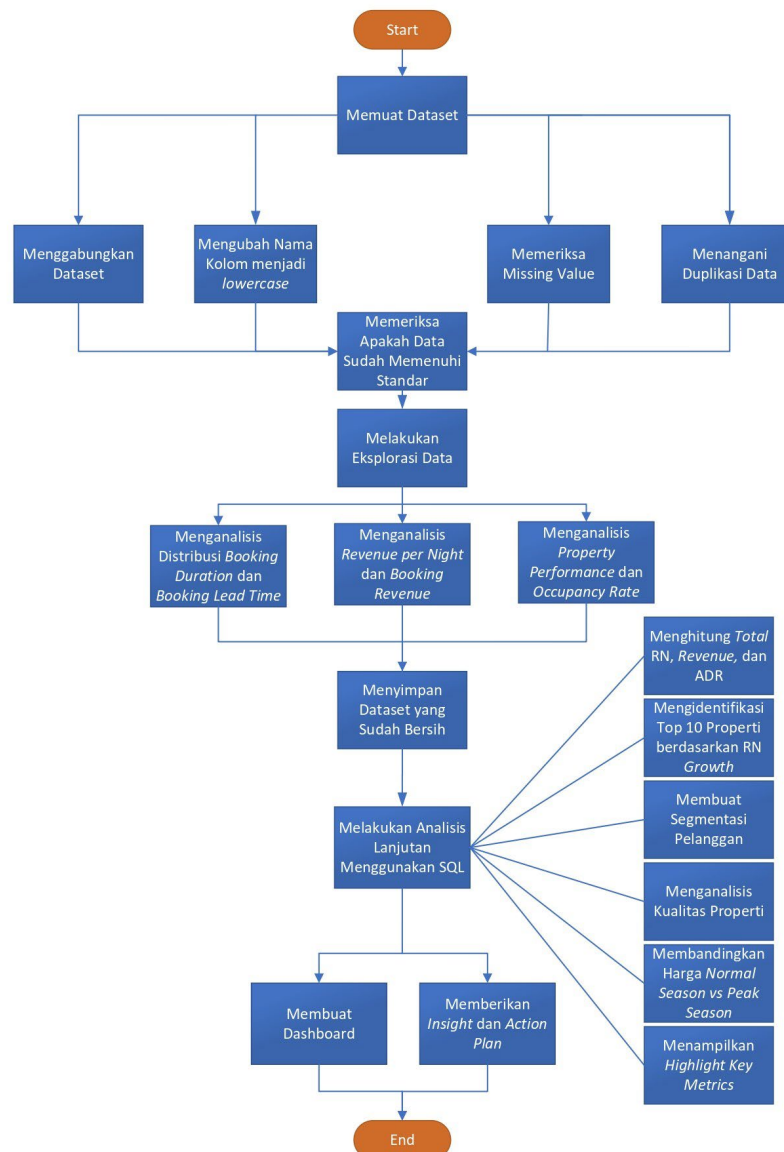
2. Proses Analisis Data

Proses analisis data dimulai dari tahap *data preparation* dan transformasi data untuk mengidentifikasi dan memperbaiki data yang hilang, inkonsisten, format data yang tidak tepat, dan *outlier* data. Rincian proses *data preparation* meliputi beberapa hal berikut:

- Menggabungkan Data dan Memperbaiki Nama Kolom
- Memeriksa Tipe Data

- Memeriksa *Missing Value* dan Menghapus Duplikasi Data
- Memeriksa *Outliers* Data

Setelah tahap *data preparation*, tahap berikutnya adalah eksplorasi data (*exploratory data analysis* - EDA), di mana pola, tren, dan hubungan antar variabel dianalisis menggunakan statistik deskriptif dan visualisasi. Setelah data dipahami, dilakukan analisis lanjutan pada SQL dan visualisasi di Tableau.



Gambar 1. Flowchart pengerjaan study case

BAB III

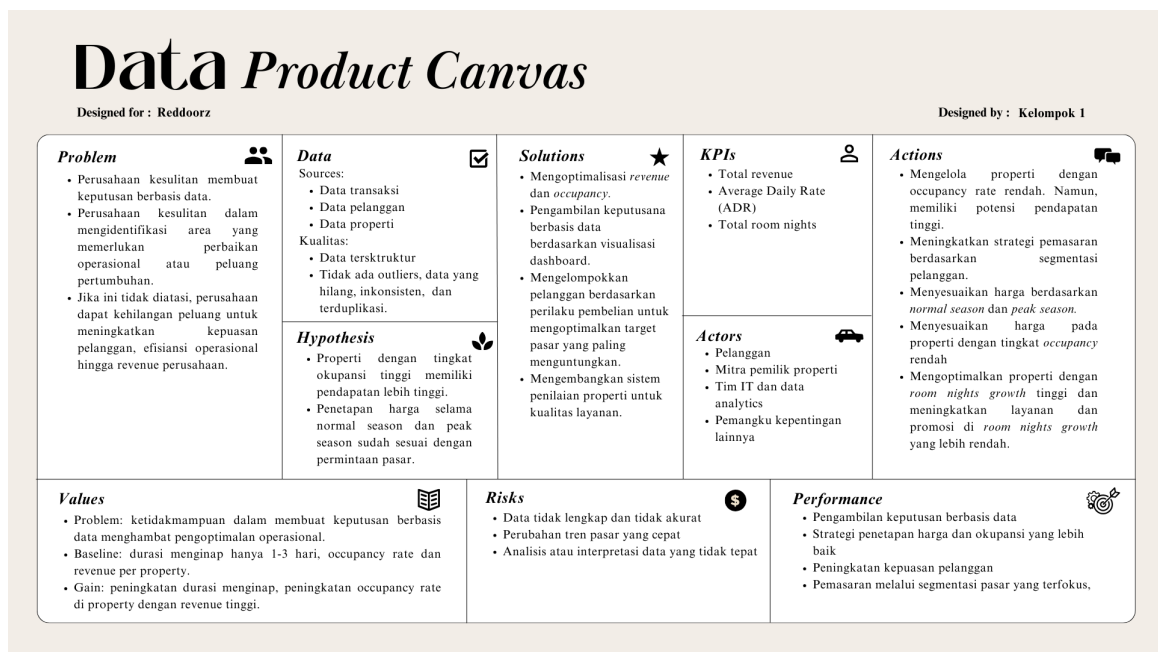
Hasil dan Pembahasan

A. Gambaran Umum Objek Penelitian

Data-data yang digunakan dalam analisis bisnis ini adalah data yang berisi informasi transaksi, pelanggan, dan properti. Dataset ini terdiri dari 100.000 baris dan 19 kolom, antara lain *series date, revenue, room/nights (RNs), occupancy, Average Daily Rate, user type, gender, travel purpose, age, city, district, inventory, date joined, dan brand type*. Secara keseluruhan, kualitas data cukup baik dengan tidak ditemukan data yang hilang, tidak konsisten atau format tipe data yang tidak sesuai, serta tidak ditemukan adanya nilai ekstrim (*outliers*).

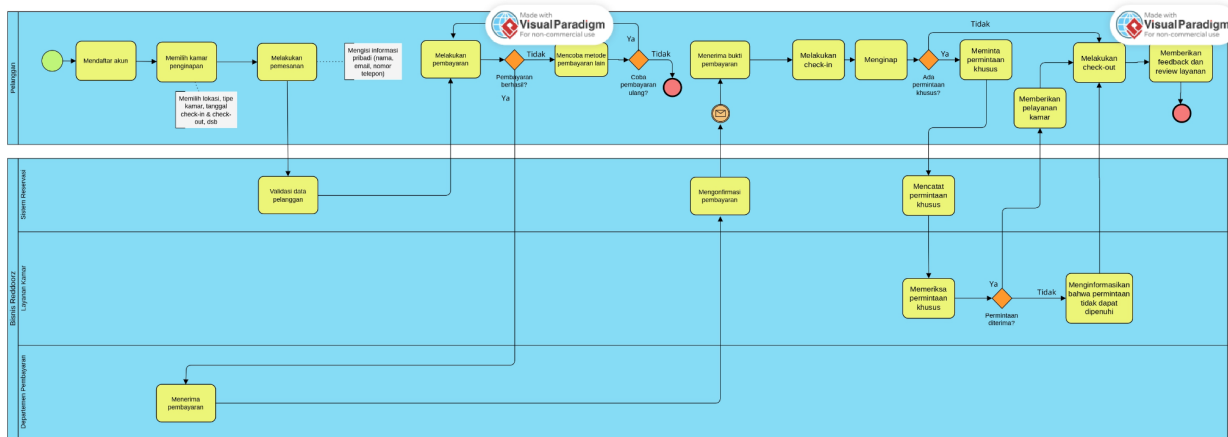
B. Temuan Hasil Penelitian

1. Product Canvas

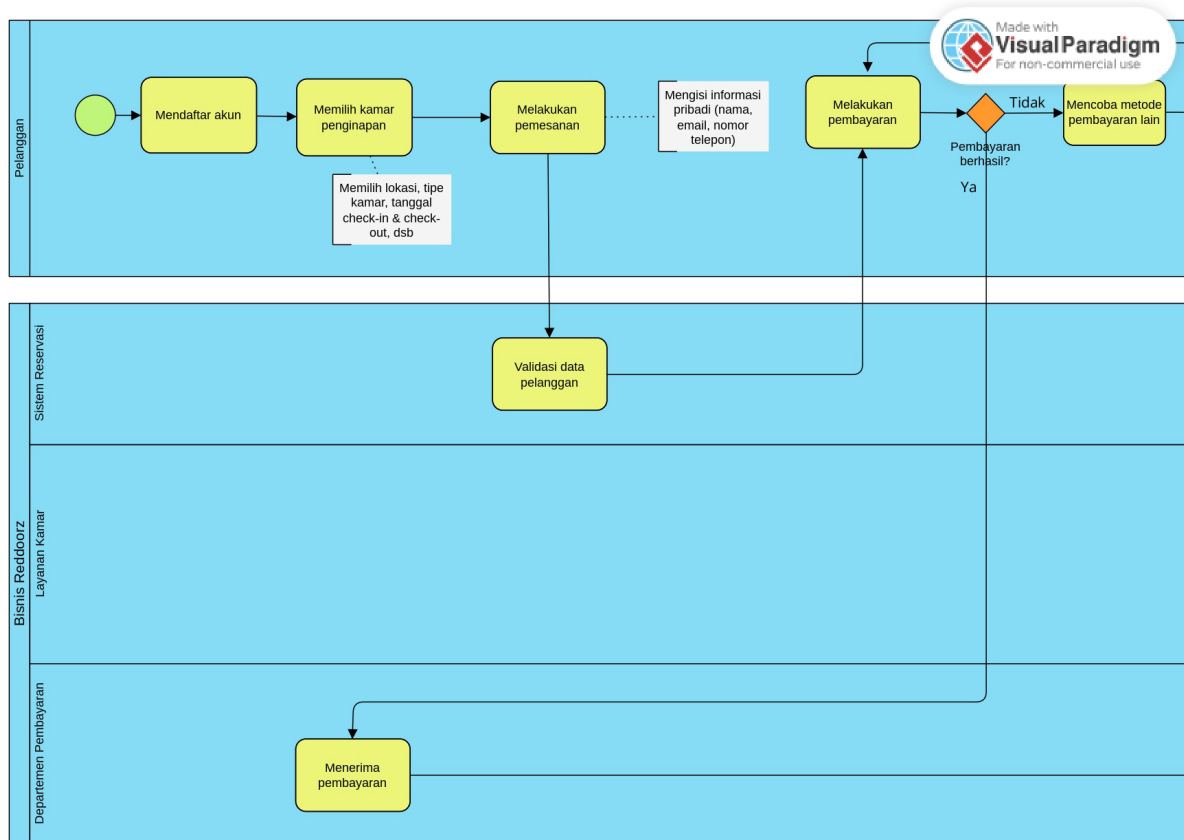


Gambar 2. Product canvas

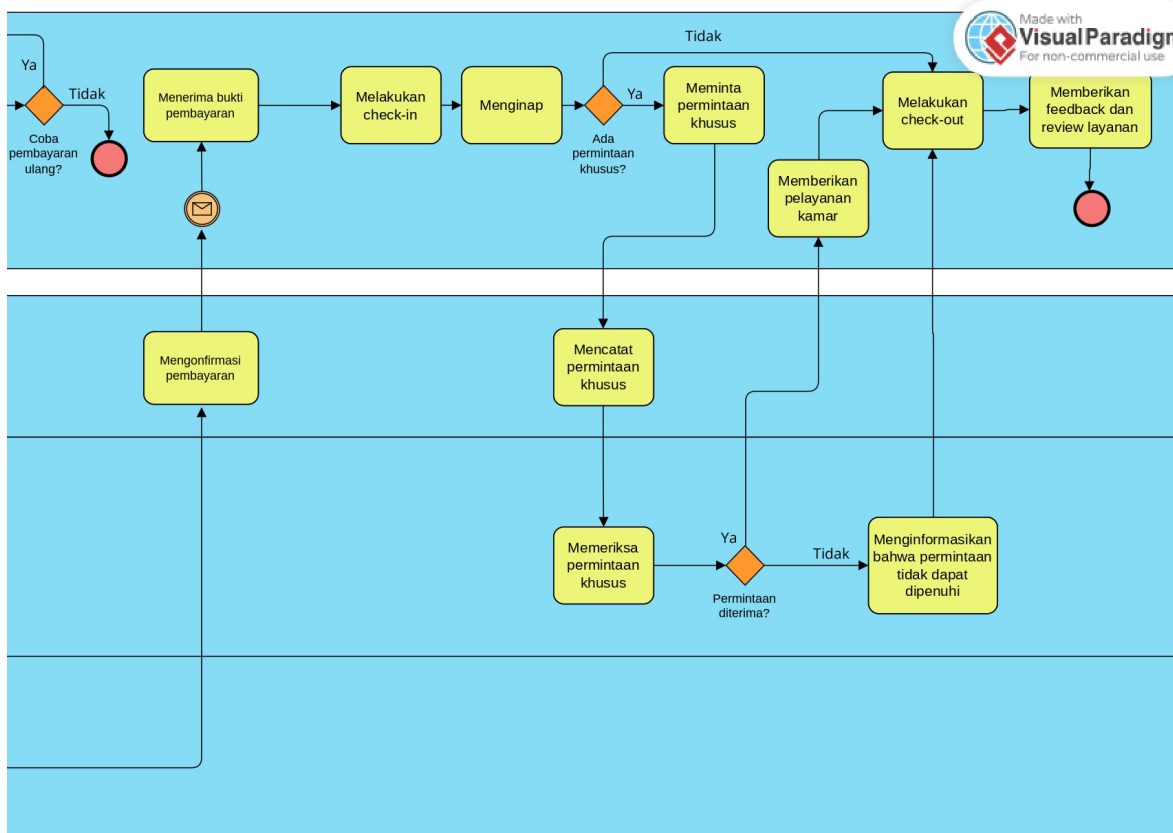
2. Business Modeling (BPMN)



Gambar 3. BPMN diagram keseluruhan



Gambar 4. BPMN diagram bagian pertama



Gambar 5. BPMN diagram bagian kedua

Alur proses bisnis dimulai dari pelanggan membuka website RedDoorz dan *login* akun atau mendaftar akun jika belum ada. Pelanggan memilih kamar yang tersedia beserta tipe kamar, tanggal *check-in* dan *check-out*, serta informasi lainnya. Kemudian, pelanggan melakukan pemesanan dengan mengisi informasi pribadi berupa nama, *e-mail*, dan nomor telepon. Sistem akan memvalidasi pemesanan tersebut lalu pelanggan akan melakukan pembayaran. Jika pembayaran gagal, pelanggan dapat mencoba pembayaran ulang. Jika pembayaran berhasil, sistem akan mengkonfirmasi pembayaran tersebut dengan mengirimkan notifikasi *e-mail* pembayaran berhasil. Setelah pelanggan menerima bukti pembayaran, pelanggan tinggal menunggu hingga waktu *check-in* tiba dan menginap. Pelanggan dapat meminta permintaan khusus selama menginap. Jika ada permintaan khusus, bagian layanan kamar akan memeriksa permintaan khusus tersebut untuk diperiksa apakah permintaan dapat diterima atau ditolak. Bagian layanan kamar akan memberikan layanan kamar jika permintaan dapat dipenuhi dan menginformasikan permintaan tidak dapat dipenuhi jika permintaan ditolak. Setelah masa

menginap berakhir, pelanggan melakukan proses *check-out* dan memberikan *feedback* dan *review* terkait layanan Reddoorz.

3. Python Programming

a. Data Preprocessing

Data preprocessing adalah sebuah tahapan yang dilakukan untuk mentransformasi data dari yang awalnya kotor menjadi bersih dalam artian data tersebut sudah bersih dari *missing value*, ketidakkonsistenan data, kesalahan tipe data, duplikasi data, serta bebas dari *outliers* atau pencilan. Secara lengkapnya langkah-langkah yang dilakukan dalam proses *data preprocessing* adalah sebagai berikut:

Kode Program 1. Mengimpor library dan memuat dataset

```
1. # Mengimport Library
2. import pandas as pd
3. import numpy as np
4. import seaborn as sns
5. import matplotlib.pyplot as plt

6. # Mengimport Dataset yang diperlukan
7. file_dataset = '/content/drive/MyDrive/Data Set - Reddoorz/Online
   Budget Hotel Dataset.xlsx'
8. df_bookings = pd.read_excel(file_dataset, sheet_name= 'Bookings
   Table')
9. df_user = pd.read_excel(file_dataset, sheet_name= 'User Table')
10. df_property = pd.read_excel(file_dataset, sheet_name= 'Property
    Table')
```

Kode di atas adalah sebuah rangkaian awal *preprocessing data* yang dimulai dari mengimpor *library* serta me-load dataset. Di mana terdapat tiga dataset yang dimuat, yaitu dataset *bookings*, *user*, dan *property*.

Kode Program 2. Penggabungan dataset

```
1. # Menggabungkan Dataset
2. df = df_bookings.merge(df_user, how='left',
   on='USER_ID').merge(df_property, how='left', on='PROPERTY_CODE')
```

Setelah dataset berhasil dimuat, ketiga dataset tersebut digabungkan dengan fungsi *merge()*, tujuan dari penggabungan tersebut adalah agar lebih efisien ketika dalam melakukan analisis data atau EDA. Ketiga dataset tersebut digabungkan menjadi satu dataframe bernama *df*. Tahapan selanjutnya adalah melakukan penyesuaian gaya huruf pada tiap kolom.

Kode Program 3. Mengatur nama tiap kolom menjadi *lower*

```
1. # Mengubah nama-nama kolom menjadi lower
2. df.columns = df.columns.str.lower()
```

Kode di atas bertujuan untuk mengubah nama tiap kolom yang awalnya *uppercase* menjadi *lowercase*. Hal tersebut bertujuan untuk mempermudah pemanggilan kolom.

Kode Program 4. Memeriksa dan menghapus *missing value*, duplikat, dan memeriksa *outliers*

```
1. # Mengecek dan menghapus duplikasi data
2. print(f'Jumlah Duplikasi Data : {df.duplicated().sum()}\n')
3. df = df.drop_duplicates()

4. # Mengecek Missing Value
5. print(f'Jumlah Missing Value : \n{df.isna().sum()}')
6.
7. # Mengecek Outliers Pada Data
8. df_to_plot = df.select_dtypes(include=('int64','float64'))
9. df_to_plot = df_to_plot.drop(columns=['booking_id','user_id'])
10. df_to_plot.plot(subplots=True, layout=(2,3), kind='box',
    figsize=(10,8), grid=True)
11. plt.subplots_adjust()
```

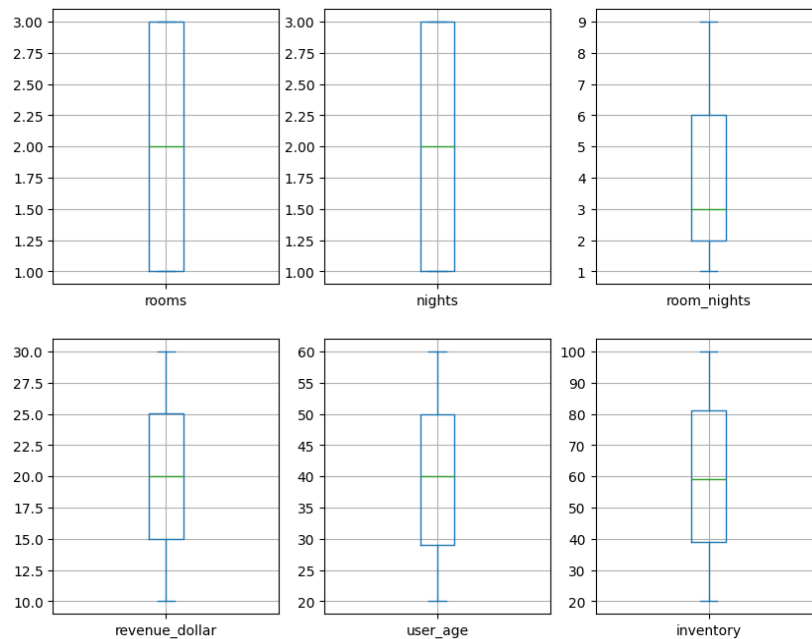
Kode di atas bertujuan untuk mengecek apakah terdapat *missing value* dan duplikasi data dalam dataset serta mengecek apakah terdapat *outliers* dalam kolom yang bertipe *numeric*. Berikut adalah *output* yang dihasilkan:

```
Jumlah Duplikasi Data : 0

Jumlah Missing Value :
booking_id      0
user_id         0
property_code   0
booking_date    0
check_in_date   0
check_out_date  0
rooms           0
nights          0
room_nights     0
revenue_dollar  0
user_type       0
user_gender     0
travel_purpose    0
user_age        0
city            0
district        0
cohort_date     0
brand_type      0
inventory       0
dtype: int64
```

Gambar 6. Pengecekan duplikasi dan *missing value*

Setelah dilakukan pengecekan didapatkan hasil bahwa di dalam dataset tidak terdapat *missing value* dan duplikasi data. Itu artinya dataset yang sudah digabung di awal tidak memiliki nilai NaN dan begitu pun dengan duplikasi data. Langkah selanjutnya adalah pengecekan *outliers*.



Gambar 7. Pengecekan *outliers*

Jika dilihat dari beberapa *boxplot* di atas terlihat bahwa untuk kolom atau *feature* yang bertipe data *numeric* tidak memiliki *outliers* atau pencilan. Semua variabel menunjukkan distribusi data yang baik tanpa *outlier* yang signifikan. Variabel *rooms*, *nights*, dan *room nights* memiliki rentang kecil dengan median yang seimbang, sedangkan *revenue dollar*, *user age*, dan *inventory* menunjukkan rentang lebih luas dengan distribusi yang simetris. Secara keseluruhan, data memiliki kualitas yang baik dan siap untuk analisis lebih lanjut. Dengan demikian, dataset sudah siap digunakan untuk tahapan selanjutnya, yaitu *exploratory data analysis*.

b. Exploratory Data Analysis

Exploratory Data Analysis adalah tahapan yang paling penting dalam rangkaian analisis data. Sebelum melakukan proses EDA, perlu diketahui terlebih dahulu pertanyaan yang akan dijawab. Berikut adalah pertanyaan serta proses analisis data menggunakan Python *programming*.

- Berapa lama pelanggan menghabiskan waktu menginap dan berapa lama waktu mereka menunggu *check-in*.

Kode Program 5. Membuat *feature* baru dan visualisasi

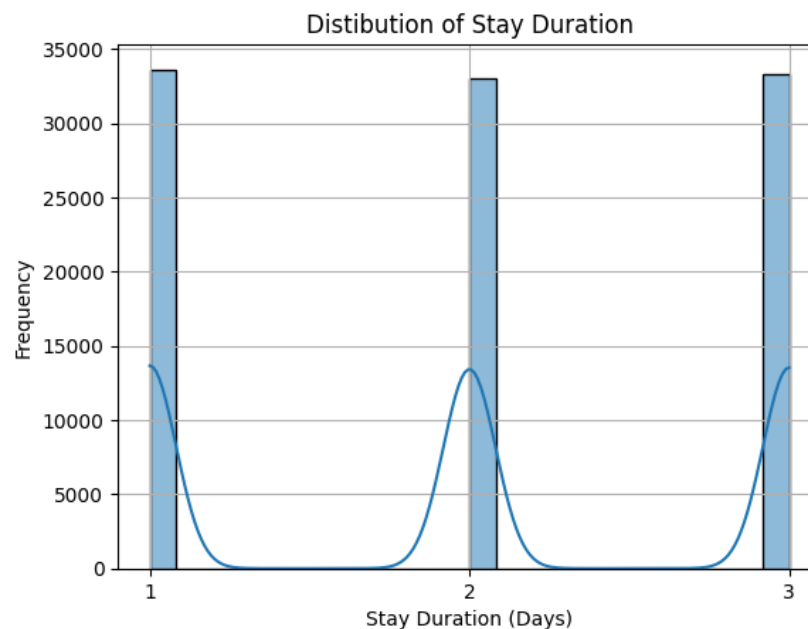
```
1. # Menambahkan kolom baru (Stay duration dan Lead Time)
2. df['stay_duration']=(df['check_out_date']-
   df['check_in_date']).dt.days
3. df['lead_time'] = (df['check_in_date'] - df['booking_date']).dt.days

4. # Visualisasi Distribusi data Stay Duration dan Lead Time
5. sns.histplot(data=df, x='stay_duration', kde=True)
6. plt.title('Distribution of Stay Duration')
7. plt.xlabel('Stay Duration (Days)')
8. plt.ylabel('Frequency')
9. plt.xticks([1,2,3])
10. plt.grid()
11. plt.show()

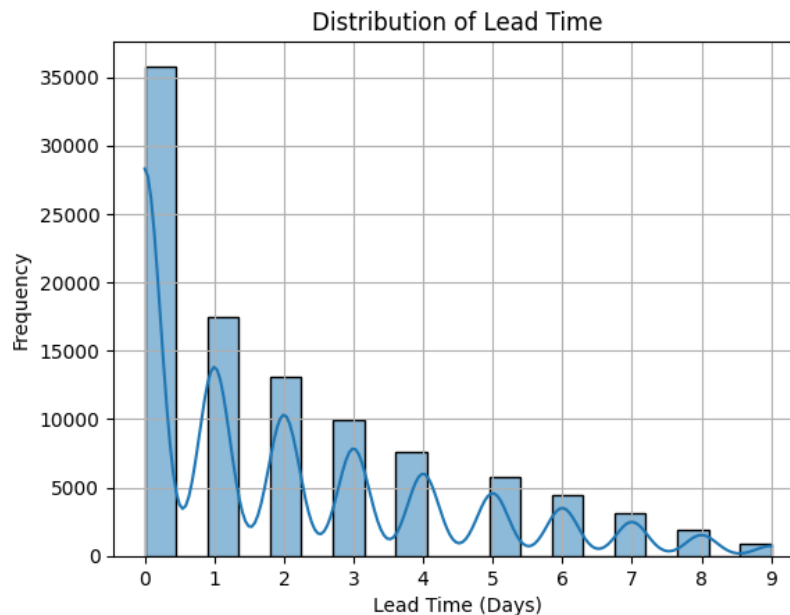
12. print('\n')

13. sns.histplot(data=df,bins=20, x='lead_time', kde=True)
14. plt.title('Distribution of Lead Time')
15. plt.xlabel('Lead Time (Days)')
16. plt.ylabel('Frequency')
17. plt.xticks([0,1,2,3,4,5,6,7,8,9])
18. plt.grid()
19. plt.show()
```

Dalam kode program di atas terdapat *syntax* untuk membuat kolom baru bernama *stay duration* dan *lead time* serta terdapat *syntax* untuk membuat visualisasi histogram. Berikut adalah *output* yang dihasilkan dari kode di atas.



Gambar 8. Distribusi dari kolom *stay duration*



Gambar 9. Distribusi dari kolom *lead time*

Dalam grafik di atas bisa dilihat bahwa untuk sebaran waktu menginap atau lamanya pelanggan menginap adalah sebanyak 1, 2, dan 3 hari dengan jumlah pemesanan yang lumayan merata di atas 30.000 pemesanan. Sedangkan untuk lamanya waktu tunggu pelanggan dari waktu *booking* ke *check-in* berjarak antara 0 hari sampai dengan 9 hari, tetapi kebanyakan para *customer* melakukan *booking* serta *check-in* dilakukan di hari yang sama, artinya mereka tidak memiliki waktu tunggu. Distribusi pemesanan yang tidak memiliki waktu tunggu atau *lead time*-nya 0 adalah sebanyak 35.000 lebih pemesanan atau 35% dari total jumlah pemesanan. Untuk jumlah pemesanan dengan waktu tunggu terlama yaitu 9 hari dengan jumlah di bawah 5.000 pemesanan dengan persentasenya sebesar 0,91%.

- **Menghitung pendapatan per malam serta pendapatan dari *booking***

Kode Program 6. Membuat *feature* baru dan visualisasi

```
1. # Menambahkan Kolom baru (Revenue per Night dan Total Booking Revenue)
2. df['revenue_per_night'] = df['revenue_dollar'] / df['room_nights']
3. df['total_booking_revenue'] =
   df.groupby('user_id')['revenue_dollar'].transform('sum')

4. # Menampilkan 10 Customer dengan Total Booking Revenue Tertinggi
5. top_10_customers =
   df[['user_id', 'total_booking_revenue']].drop_duplicates().sort_values(
       by='total_booking_revenue',
       ascending=False).head(10).reset_index(drop=True)
6. print(top_10_customers, '\n')
```

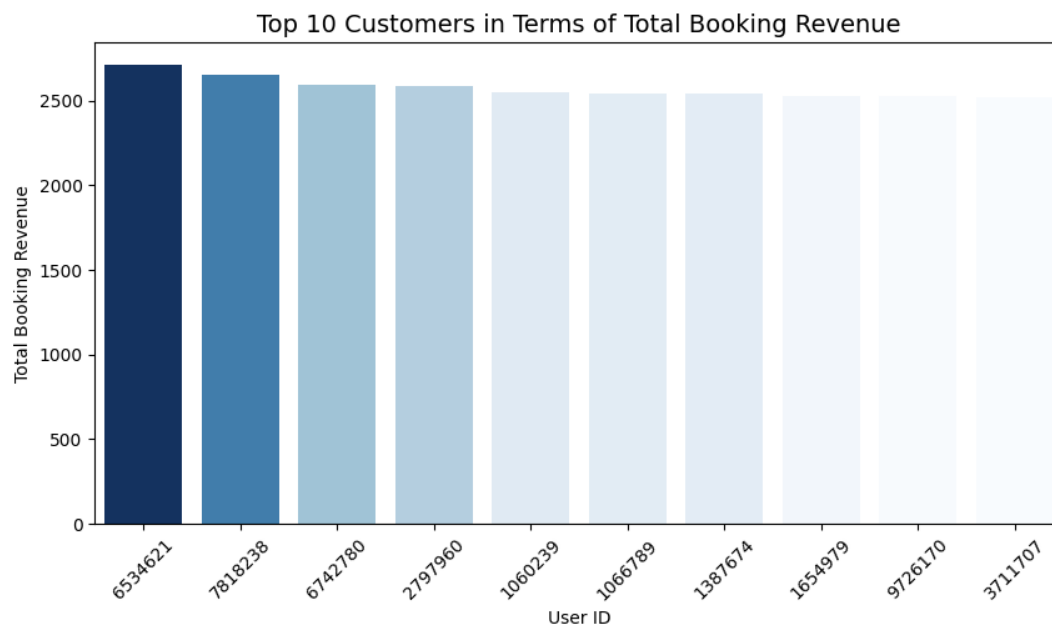


```

7. # Visualisasi 10 Customer dengan total booking revenue tertinggi
8. plt.figure(figsize=(10,5))
9. sns.barplot(x='user_id',y='total_booking_revenue',
    data=top_10_customers,
    palette='Blues',order=top_10_customers['user_id'],hue='total_booking
    _revenue',dodge=False)
10. plt.title('Top 10 Customers in Terms of Total Booking Revenue',
    fontsize=14, fontweight=10)
11. plt.xlabel('User ID')
12. plt.xticks(rotation=45)
13. plt.ylabel('Total Booking Revenue')
14. plt.legend([], [], frameon=False)
15. plt.show()

```

Dalam kode program di atas terdapat *syntax* untuk membuat kolom baru bernama *revenue per night* dan *total booking* serta terdapat *syntax* untuk membuat visualisasi *bar chart*. Berikut adalah *output* yang dihasilkan dari kode di atas:



Gambar 10. Top 10 customers terhadap revenue total booking

```

user_id      6534621
user_gender  Female
user_age     58
total_booking_revenue  2708.86
city         Jakarta
stay_duration  2
travel_purpose Business
Name: 0, dtype: object

```

Gambar 11. Profil customer tertinggi

Di atas merupakan total pendapatan booking tertinggi dari 10 pelanggan. Pelanggan dengan User ID 6534621 merupakan yang paling tinggi dibandingkan dengan *user* yang lain dalam hal kontribusi terhadap *revenue booking* di mana *user* tersebut ber-gender perempuan yang berasal dari kota Jakarta dengan durasi menginap selama 2 hari, ada pun tujuan menginap dari *user* tersebut adalah kepentingan bisnis. *Customer* tersebut melakukan pemesanan sebanyak 133 kali pemesanan.

- **Menghitung Total Revenue per Property dan Occupancy Rate**

Kode Program 7. Membuat *feature* baru dan visualisasi

```
1. # Membuat Kolom Baru Days Active
2. df['days_active'] = (df['check_in_date']-df['cohort_date']).dt.days

3. # Menambahkan kolom baru Occupancy Rate
4. df['occupancy_rate'] = df['room_nights'] / (df['inventory'] *
df['days_active'])

5. # Melakukan Pengkondisian Apabila Hari Aktif nol atau negatif untuk
menghindari kesalahan pembagian
6. df.loc[df['days_active'] <= 0, 'occupancy_rate'] = 0

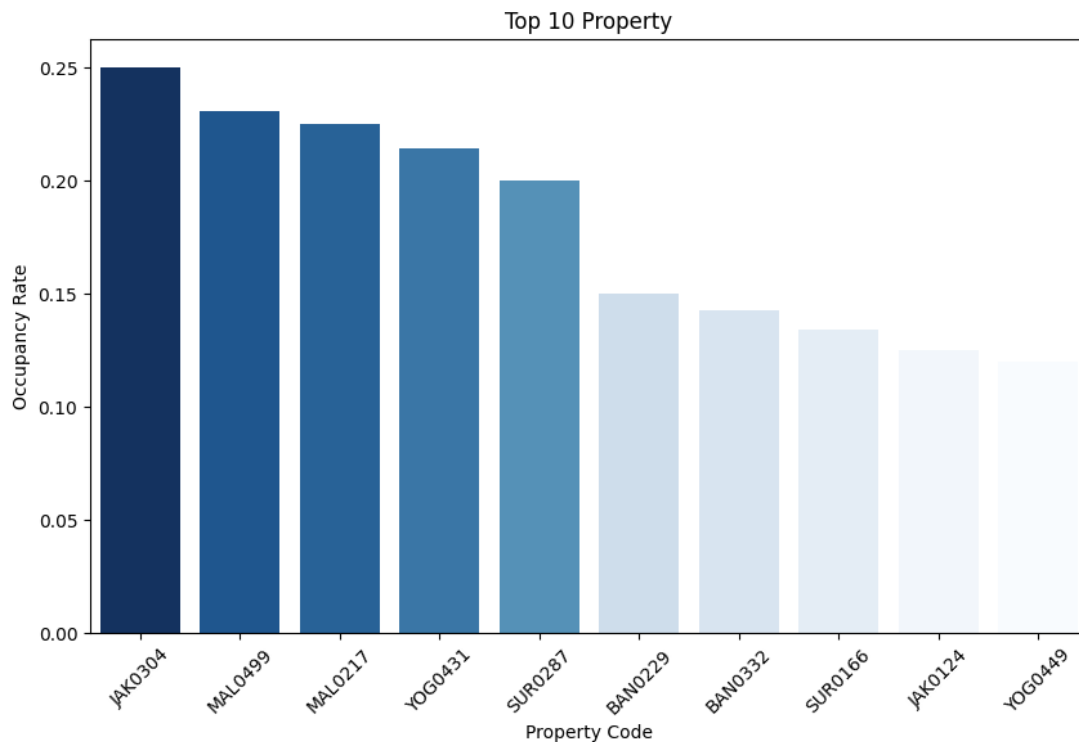
7. # Menambahkan Kolom Baru Total Revenue Per Property
8. total_revenue_per_property =
df.groupby('property_code')['total_booking_revenue'].sum()
9. df =
df.merge(total_revenue_per_property.rename('total_revenue_per_property'),
on='property_code')

10. top_10_properties = df[['property_code',
'occupancy_rate']].drop_duplicates().sort_values(by='occupancy_rate',
ascending=False).head(10).reset_index(drop=True)
11. print(top_10_properties, '\n')

12. plt.figure(figsize=(10,6))
13. sns.barplot(x='property_code', y='occupancy_rate', data=top_10_properties,
palette='Blues', hue='occupancy_rate')
14. plt.title('Top 10 Property')
15. plt.xlabel('Property Code')
16. plt.ylabel('Occupancy Rate')
17. plt.xticks(rotation=45)
18. plt.legend([], [], frameon=False)
19. plt.show()

20. top_property = df[['property_code',
'occupancy_rate', 'city', 'district', 'brand_type', 'inventory']].drop_duplic
ates().sort_values(by='occupancy_rate',
ascending=False).head(10).reset_index(drop=True)
21. print(top_property.loc[0])
```

Dalam kode program di atas terdapat *syntax* untuk membuat kolom baru bernama *days active*, *occupancy rate*, dan *total revenue per property* serta terdapat *syntax* untuk kemudian membuat visualisasi *bar chart*. Berikut adalah *output* yang dihasilkan dari kode di atas:



Gambar 12. Top 10 property terhadap occupancy rate

```
property_code    JAK0304
occupancy_rate   0.25
city             Jakarta
district         South
brand_type       Koolkost
inventory        36
Name: 0, dtype: object
```

Gambar 13. Profil properti tertinggi

Kita bisa melihat bahwa untuk properti dengan *rate* tertinggi ada pada properti dengan kode JAK0304, dengan *rate* sebesar 0.250000. Properti tersebut berada di kota Jakarta dengan *brand type*-nya adalah Koolkost. Ada pun jumlah *inventory* dari properti tersebut sebanyak 36 dengan jumlah pemesanan sebanyak 209 pesanan.

Kode Program 8. Menghitung korelasi *occupancy rate* dengan *total revenue per property*

```
1. # Menghitung Korelasi Antara Occupancy Rate dengan Total Revenue
   Per Property
2. correlation = df[['occupancy_rate',
                    'total_revenue_per_property']].corr().iloc[0, 1]
```

1.0521750205043423e-05

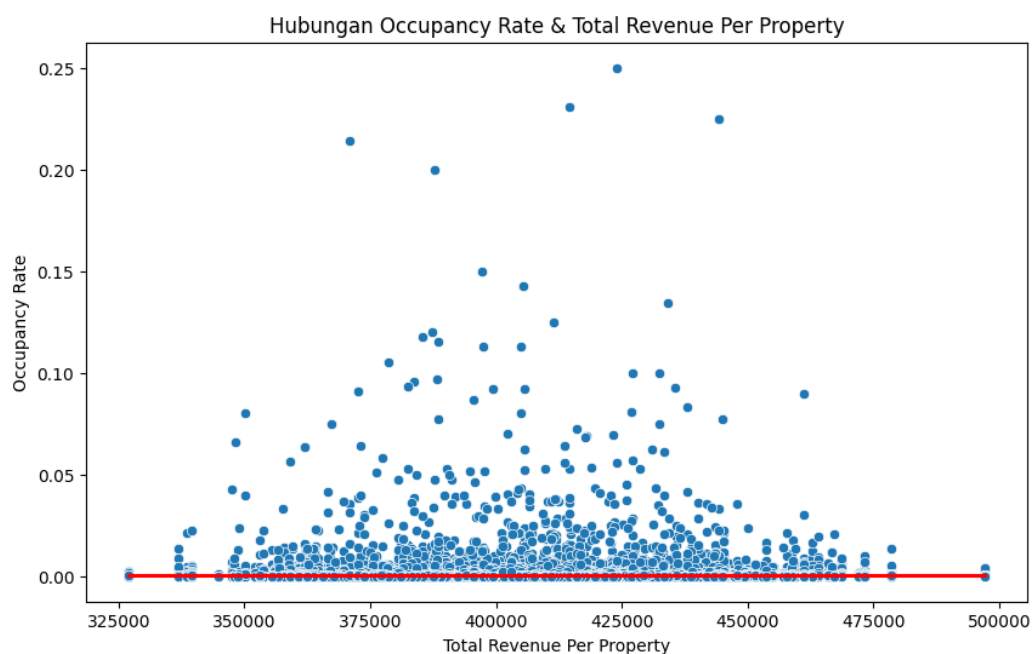
Gambar 14. Angka korelasi *occupancy rate* dan *total revenue per property*

Setelah menghitung korelasi antara *occupancy rate* dengan *total revenue per property* didapatkan hasil sebesar 1.0521750205043423e-05. Angka tersebut menunjukkan bahwa tidak ada korelasi atau hubungan yang linear antara dua variabel tersebut karena angka sebenarnya dari hasil tersebut adalah 0.00001.

Kode Program 9. Membuat *scatter plot* antara *occupancy rate* dengan *total revenue per property*

```
1. plt.figure(figsize=(10,6))
2. sns.scatterplot(y='occupancy_rate',x='total_revenue_per_property',
  data=df)
3. sns.scatterplot(y='occupancy_rate',x='total_revenue_per_property',
  data=df)
4. plt.title('Hubungan Occupancy Rate & Total Revenue Per Property')
5. plt.xlabel('Total Revenue Per Property')
6. plt.ylabel('Occupancy Rate')
7. plt.show()
```

Kode di atas berisi *syntax* untuk membuat sebuah *scatter plot* yang berfungsi untuk mengetahui hubungan antara *occupancy rate* dengan *total revenue per property*. Berikut adalah visualisasinya:



Gambar 15. Hubungan *Occupancy rate* & *Total Revenue per property*

Begitu pun jika dilihat dari *scatter plot* di atas menunjukkan tidak ada hubungan yang linear antara *occupancy rate* dan *total revenue per property*.

4. Advanced Data Analytics SQL

- Menghitung total RNs, *revenue*, dan ADR berdasarkan kota untuk setiap bulannya di tahun 2024.

Kode Program 1. *Menghitung total RNs, revenue, and ADR*

```
3. SELECT
4.     month_name,
5.     city,
6.     sum(room_nights) as total_room_nights,
7.     sum(revenue_dollar) as total_revenue,
8.     sum((revenue_dollar / room_nights)) as adr
9. FROM reddoorz r
10. WHERE year = 2024
11. GROUP by city, month_name
12. ORDER BY 5 Desc
```

Output dari SQL *query* di atas adalah sebagai berikut:

	A-Z month_name	A-Z city	123 total_room_nights	123 total_revenue	123 adr
1	April	Yogyakarta	1,562	8,040.33	3,153.7591666667
2	August	Yogyakarta	1,583	7,992.58	3,028.1619444444
3	February	Yogyakarta	1,432	7,337.42	2,900.2819444444
4	March	Yogyakarta	1,547	7,527.28	2,864.2461111111
5	May	Yogyakarta	1,494	7,333.26	2,802.565
6	April	Bandung	1,523	7,551.99	2,772.4191666667
7	May	Bandung	1,505	7,406.45	2,756.8922222222
8	January	Bandung	1,500	7,555.65	2,742.05
9	August	Bandung	1,452	7,385.79	2,710.3063888889
10	March	Surabaya	1,398	7,042.46	2,699.215
11	January	Jakarta	1,452	7,031.38	2,693.9588888889
12	January	Yogyakarta	1,537	7,408.56	2,693.7413888889
13	April	Surabaya	1,333	6,958.56	2,679.4480555556
14	July	Jakarta	1,391	6,972.43	2,672.4469444444
15	September	Jakarta	1,491	7,313.62	2,668.615
16	May	Jakarta	1,574	7,605.04	2,627.2994444444
17	June	Yogyakarta	1,468	7,231.78	2,624.1125
18	July	Bandung	1,491	7,296.94	2,613.7361111111

Gambar 16. *Total RN, revenue, dan ADR*

Insight yang dapat kita ambil dari hasil *query* di atas adalah bahwa Yogyakarta merupakan kota dengan total *room nights* terbanyak pada bulan Agustus dengan total 1.583. Total *revenue* dan ADR tertinggi ada pada bulan April dengan total *revenue* 8,040.33 dan total ADR 3,153.75 di kota yang sama, yaitu Yogyakarta.

- Mengidentifikasi top 10 *properties* berdasarkan RNs *growth* dengan membandingkan antara tahun 2023 dan tahun 2024.

Kode Program 2. Top 10 *properties* berdasarkan RNs *growth*

```
1. SELECT
2.     property_code,
3.     SUM(CASE WHEN YEAR = 2023 THEN room_nights ELSE 0 END) AS
   room_nights_2023,
4.     SUM(CASE WHEN YEAR = 2024 THEN room_nights ELSE 0 END) AS
   room_nights_2024,
5.     SUM(CASE WHEN YEAR = 2024 THEN room_nights ELSE 0 END) -
6.     SUM(CASE WHEN YEAR = 2023 THEN room_nights ELSE 0 END) AS
   room_nights_growth,
7.     ((SUM(CASE WHEN year = 2024 THEN room_nights ELSE 0 END) -
8.     SUM(CASE WHEN year = 2023 THEN room_nights ELSE 0 END)) * 100) /
9.     SUM(CASE WHEN year = 2023 THEN room_nights ELSE 0 END) AS
   growth_percentage
9. FROM reddoorz r
10. GROUP BY property_code
11. ORDER BY 5 DESC
12. LIMIT 10
```

Output dari SQL query di atas adalah sebagai berikut:

A-Z property_code	123 room_nights_2023	123 room_nights_2024	123 room_nights_growth	123 growth_percentage
YOG0342	127	217	90	70
JAK0311	109	164	55	50
BAN0251	132	194	62	46
YOG0297	89	129	40	44
JAK0160	105	151	46	43
JAK0197	112	159	47	41
YOG0219	111	156	45	40
JAK0183	139	195	56	40
BAN0306	117	164	47	40
SUR0174	140	193	53	37

Gambar 17. Top 10 pertumbuhan *room night*

Bisa dilihat bahwa untuk properti dengan *growth* tertinggi ada pada *property code* 'YOG0342' dengan pertumbuhan sebesar 70% dari *room nights* tahun sebelumnya.

- Menentukan *customer segmentation* berdasarkan perilaku pembelian untuk mengidentifikasi target pasar yang lebih terfokus.

Kode Program 3. *Customer segmentation* berdasarkan perilaku pembelian

```
1. WITH cte AS (SELECT
2.     user_id,
3.     city,
4.     user_gender,
5.     user_age,
6.     SUM(revenue_dollar) AS total_revenue,
7.     sum(room_nights) AS total_room_nights
8. FROM reddoorz r
9. GROUP BY 1
10. ORDER BY total_revenue desc
11. )
```

```

12.
13. SELECT
14. user_id,
15. city,
16. user_gender,
17. user_age,
18. CASE
19.     WHEN total_revenue > 2146 THEN 'Top'
20.     WHEN total_revenue BETWEEN 1848 AND 2146 THEN 'Middle'
21.     ELSE 'Lower'
22. END AS segmentation
23. FROM cte
24.

```

Output dari SQL query di atas adalah sebagai berikut:

123 user_id	A-Z city	A-Z user_gender	123 user_age	A-Z segmentation
6,534,621	Yogyakarta	Female	58	Top
7,818,238	Jakarta	Female	47	Top
6,742,780	Jakarta	Male	20	Top
2,797,960	Jakarta	Male	29	Top
1,060,239	Surabaya	Male	52	Top
1,066,789	Yogyakarta	Female	58	Top
1,387,674	Bandung	Male	47	Top
1,654,979	Surabaya	Male	53	Top
9,726,170	Jakarta	Male	40	Top
3,711,707	Surabaya	Female	40	Top
7,599,624	Surabaya	Female	41	Top
10,411,891	Surabaya	Female	54	Top
12,527,751	Bandung	Male	25	Top
11,016,762	Surabaya	Female	48	Top
3,641,580	Bandung	Male	23	Top
12,078,740	Surabaya	Female	49	Top
8,188,992	Malang	Female	52	Top
11,266,912	Surabaya	Female	50	Top

Gambar 18. Segmentasi pelanggan

Insight yang dapat diambil dari hasil *query* di atas adalah bahwa kota dengan segmentasi Top terbanyak adalah kota Jakarta dan Malang dengan gender *female* dan usia rata-ratanya adalah 35 tahun hingga 38 tahun. Kemudian Surabaya dengan segmentasi Lower terbanyak dengan didominasi oleh gender *male* dan rata-rata usianya, yaitu 41 tahun. Untuk segmentasi Middle terbanyak berada di kota Jakarta yang didominasi oleh gender *female* dengan usia rata-rata 41 tahun.

- Membandingkan *normal season vs. peak season pricing* di tahun 2023 dan 2024 untuk menentukan apakah strategi penetapan harga perusahaan sudah sesuai dengan permintaan berdasarkan musim

Kode Program 4. *Normal season vs. peak season pricing*

```

1. with cte_2 AS (SELECT
2.     YEAR,
3.     month_name,
4.     CASE
5.         WHEN month_name IN ('January','March','July') THEN
6.             'Peak Season'
7.         Else 'Normal Season'
8.     END AS season,
9.     revenue_dollar
10. FROM reddoorz r
11. WHERE year in (2023,2024)
12. )
13. SELECT
14.     YEAR,
15.     month_name,
16.     season,
17.     sum(revenue_dollar)
18. FROM cte_2
19. GROUP BY YEAR, month_name, season
20. ORDER BY revenue_dollar DESC

```

Output dari SQL query di atas adalah sebagai berikut:

	123 YEAR	A-Z month_name	A-Z season	123 sum(revenue_dollar)
1	2023	August	Normal Season	36,850.41
2	2023	January	Peak Season	35,923.64
3	2023	July	Peak Season	35,752.72
4	2023	December	Normal Season	35,751.18
5	2024	April	Normal Season	35,723.35
6	2023	November	Normal Season	35,703.78
7	2023	March	Peak Season	35,528.81
8	2024	March	Peak Season	35,383.49
9	2023	September	Normal Season	35,282.81
10	2024	August	Normal Season	35,264.87
11	2024	May	Normal Season	35,215.24
12	2024	January	Peak Season	35,043.36
13	2023	June	Normal Season	34,750.1
14	2023	October	Normal Season	34,594.08
15	2024	July	Peak Season	34,550.09
16	2024	February	Normal Season	34,411.57
17	2023	April	Normal Season	33,749.85
18	2023	May	Normal Season	33,693.65
19	2024	September	Normal Season	33,418.77
20	2023	February	Normal Season	32,885.44
21	2024	June	Normal Season	32,238.73

Gambar 19. Tren musiman tahun 2023 dan 2024

Sebagian besar pendapatan sudah *aligned* dengan *seasonal demand*. Namun, terdapat beberapa pendapatan yang bisa ditinjau ulang. Contohnya, permintaan musiman di bulan Januari 2024 dan Juli 2024, ini belum dimanfaatkan dengan optimal di mana *revenue* bulan-bulan tersebut lebih rendah dibandingkan bulan lain yang masuk kategori *peak season*, meskipun perbedaan pendapatannya tidak terlalu signifikan.

- **Menganalisis kualitas properti yang diakuisisi menggunakan penilaian berbasis logika.**

Kode Program 5. Menganalisis kualitas properti yang diakuisisi menggunakan penilaian berbasis logika.

```
1. SELECT
2.     property_code ,
3.     brand_type,
4.     CASE
5.         WHEN total_revenue_per_property > 426784 THEN 'A'
6.         WHEN total_revenue_per_property BETWEEN 386557 and
7.             426784 THEN 'B'
8.         WHEN total_revenue_per_property < 386557 THEN 'C'
9.     END AS Grade
10. FROM reddoorz r
```

Output dari SQL *query* di atas adalah sebagai berikut:

A-z property_code	A-z brand_type	A-z Grade
MAL0084	Koolkost	A
SUR0320	RedDoorz	C
SUR0381	RedPartner	C
BAN0082	Koolkost	A
BAN0482	Koolkost	B
YOG0230	RedPartner	B
SUR0193	RedDoorz	C
YOG0290	RedPartner	B
MAL0108	RedPartner	B
SUR0338	RedDoorz	A
YOG0245	RedDoorz	A
MAL0356	RedDoorz	B
BAN0291	Koolkost	B
SUR0120	Koolkost	B
MAL0117	RedPartner	B
BAN0140	RedPartner	B
MAL0281	RedDoorz	A
MAL0044	Koolkost	A

Gambar 20. Kualitas properti berdasarkan *total revenue*

Gambar di atas merupakan penggalan *output* dari *query* yang telah dibuat. *Insight* yang didapat adalah *brand type* dengan Grade A dan B terbanyak dipegang oleh Reddoorz dengan

jumlah Grade A sebanyak 8.935 dan Grade B sebanyak 17.856. Untuk Grade C yang paling banyak dipegang oleh RedPartner dengan jumlah sebanyak 9.738. Dengan catatan, grade A adalah *brand type* dengan *total revenue per property*-nya lebih dari 426.784, Grade B antara 386.557 sampai 426.784, dan Grade C kurang dari 386.557.

- **SQL Views: Membuat *regular reporting* untuk menyoroti *key metrics* selama enam bulan terakhir di tingkat kota.**

Kode Program 6. *SQL Views*

```
1. CREATE VIEW season AS SELECT
2.     city,
3.     month_name,
4.     sum(room_nights) AS total_room_nights,
5.     sum(revenue_dollar / room_nights) as adr,
6.     sum(revenue_dollar) AS total_Revenue
7. FROM reddoorz r
8. WHERE YEAR = 2024 AND month_name IN
9.     ('April', 'May', 'June', 'July', 'August', 'September')
10. GROUP BY city, month_name
11. ORDER BY 2 DESC
```

Output dari SQL *query* di atas sebagai berikut:

A-Z city	A-Z month_name	123 total_room_nights	123 adr	123 total_Revenue
Bandung	September	1,363	2,513.1294444444	6,819.71
Jakarta	September	1,491	2,668.615	7,313.62
Malang	September	1,220	2,456.985	6,394.9
Surabaya	September	1,139	2,025.8913888889	5,535.77
Yogyakarta	September	1,513	2,608.2822222222	7,354.77
Bandung	May	1,505	2,756.8922222222	7,406.45
Jakarta	May	1,574	2,627.2994444444	7,605.04
Malang	May	1,384	2,329.0458333333	6,481.75
Surabaya	May	1,251	2,485.1713888889	6,388.74
Yogyakarta	May	1,494	2,802.565	7,333.26
Bandung	June	1,387	2,426.7780555556	6,732.3
Jakarta	June	1,248	2,402.3236111111	6,244.38
Malang	June	1,287	2,231.1394444444	6,130.39
Surabaya	June	1,147	2,370.3561111111	5,899.88
Yogyakarta	June	1,468	2,624.1125	7,231.78
Bandung	July	1,491	2,613.7361111111	7,296.94
Jakarta	July	1,391	2,672.4469444444	6,972.43
Malang	July	1,466	2,478.0102777778	6,809.99
Surabaya	July	1,335	2,363.4458333333	6,358.56
Yogyakarta	July	1,498	2,589.1869444444	7,112.17
Bandung	August	1,452	2,710.3063888889	7,385.79
Jakarta	August	1,273	2,361.2747222222	6,380.81
Malang	August	1,345	2,521.3083333333	6,753.5
Surabaya	August	1,344	2,512.7652777778	6,752.19
Yogyakarta	August	1,583	3,028.1619444444	7,992.58
Bandung	April	1,523	2,772.4191666667	7,551.99
Jakarta	April	1,412	2,357.1077777778	6,636.83
Malang	April	1,371	2,291.3163888889	6,535.64
Surabaya	April	1,333	2,679.4480555556	6,958.56
Yogyakarta	April	1,562	3,153.7591666667	8,040.33

Gambar 21. Output SQL Views

Berdasarkan penggalan *output* di atas, kota Yogyakarta memiliki rata-rata total pendapatan tertinggi, khususnya pada bulan April dan Agustus dengan pendapatan di atas 7.500. Yogyakarta, Bandung, dan Jakarta menunjukkan performa kuat dengan nilai ADR (*Average Daily Rate*) rata-rata di atas 2.500, menandakan harga kamar yang lebih tinggi dibandingkan kota lain. Sementara itu, Surabaya dan Malang menjadi kota dengan total malam kamar, ADR, dan pendapatan yang lebih rendah di antara kota lainnya, dengan rata-rata ADR Malang di kisaran 2.400 dan rata-rata ADR Surabaya 2.300 per bulan, rata-rata pendapatan Surabaya di kisaran 6.300 dan rata-rata pendapatan Malang di kisaran 6.500 serta rata-rata *room night* Surabaya di kisaran 1.200 dan rata-rata *room night* Malang di kisaran 1.300.

5. Tableau for Data Visualization



Gambar 22. Dashboard visualisasi

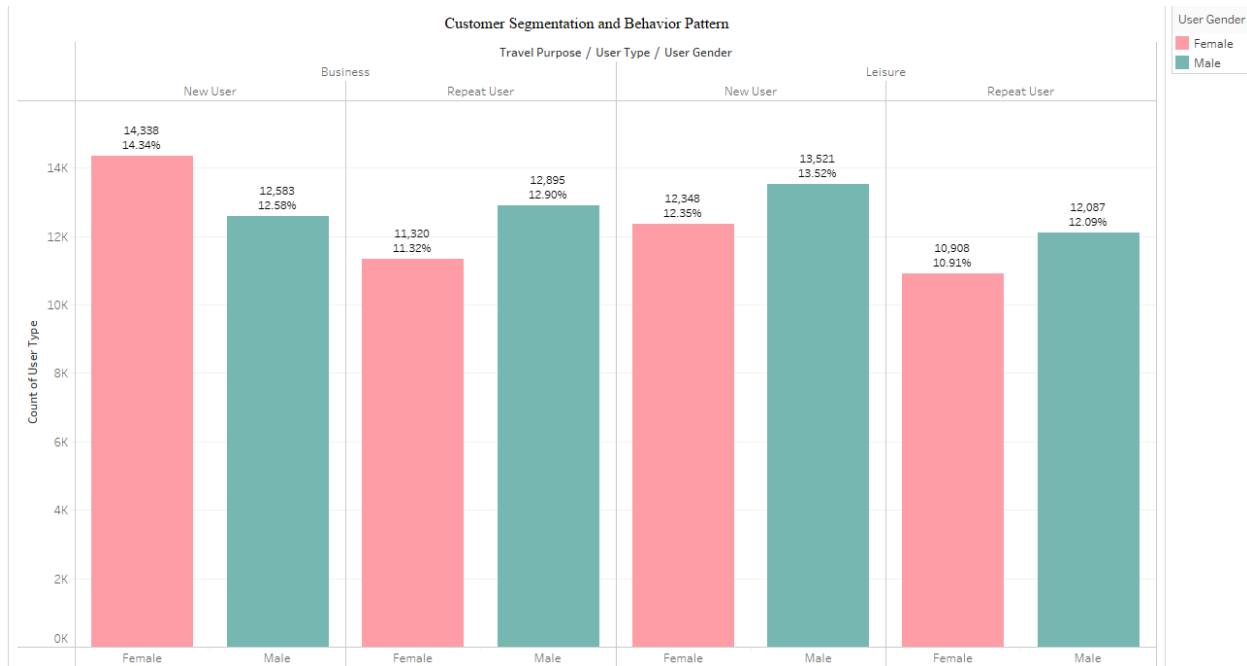
a. Revenue trends by city and brand type



Gambar 23. Visualisasi tren pendapatan

- Koolkost
 - Surabaya memiliki pendapatan terendah dibandingkan kota lain.
 - Bandung menghasilkan pendapatan tertinggi untuk brand ini.
 - Pendapatan Koolkost menunjukkan tren stabil di seluruh kota.
- RedDoorz
 - Bandung memiliki pendapatan terendah.
 - Surabaya mencatat pendapatan tertinggi dengan perbedaan signifikan dibandingkan kota lain.
- RedPartner
 - Surabaya menunjukkan pendapatan terendah.
 - Yogyakarta menjadi kota dengan pendapatan tertinggi untuk RedPartner.

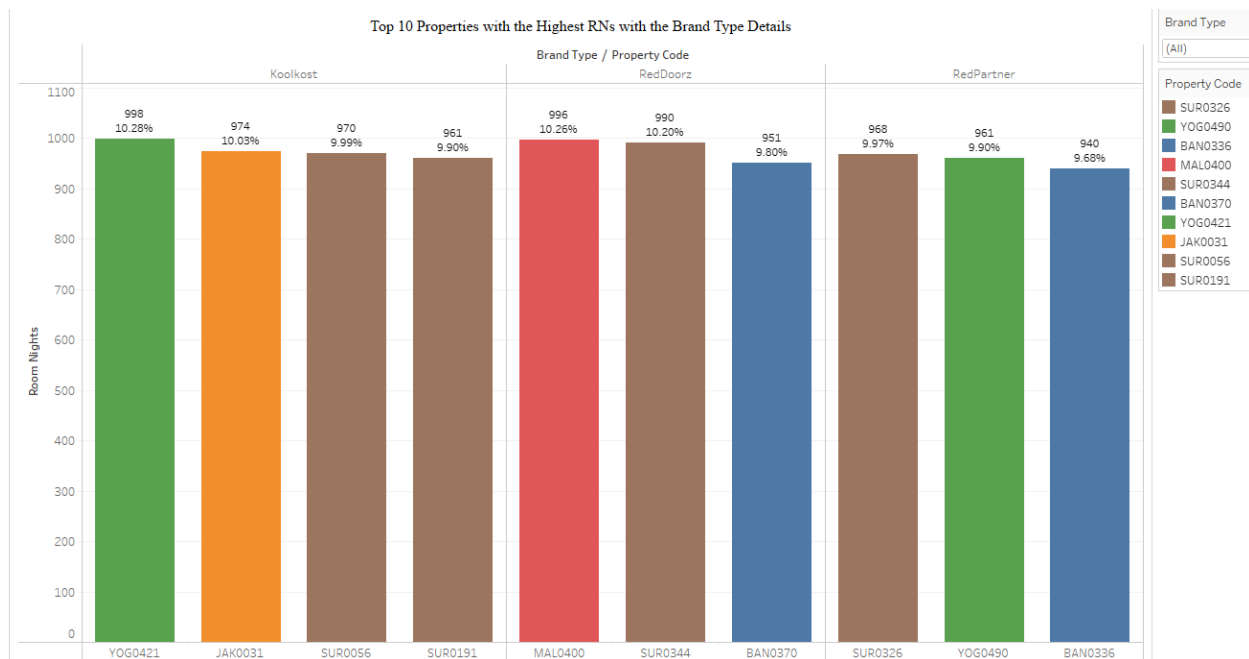
b. Customer segmentation and behavior pattern



Gambar 24. Visualisasi segmentasi pelanggan (*travel purpose, user type & gender*)

- New Users
 - Untuk perjalanan bisnis, pengguna wanita (14,338) lebih banyak dibandingkan pria (12,583).
 - Untuk perjalanan leisure, pengguna pria (13,521) mendominasi dibandingkan wanita (12,348).
- Repeat Users
 - Untuk perjalanan bisnis, pengguna pria (12,895) lebih tinggi dibandingkan wanita (11,320).
 - Untuk perjalanan leisure, pengguna pria (12,087) juga lebih tinggi dibandingkan wanita (10,908).

c. The top 10 properties with the highest RNs with the brand type details



Gambar 25. Visualisasi top 10 properties berdasarkan room nights tertinggi

- Jakarta: Pendapatan cenderung stabil dan tidak terlalu ekstrem.
- Bandung dan Yogyakarta: Kedua kota ini memiliki performa pendapatan relatif tinggi untuk semua jenis brand.
- Surabaya: Memerlukan strategi tambahan untuk meningkatkan pendapatan, terutama untuk brand Koolkost dan RedPartner.
- Malang: Memiliki pendapatan yang tinggi untuk kategori brand RedDoorz, untuk kategori brand lain perlu ditingkatkan lagi *campaign*-nya.

C. Pembahasan

Penting untuk mengidentifikasi area-area memerlukan peningkatan dan menyusun strategi khusus sebagai upaya dalam meningkatkan kinerja bisnis. Berdasarkan analisis yang telah dilakukan, beberapa area utama yang perlu ditingkatkan, yaitu *sales*, *pricing*, dan *property acquisition*. Strategi perbaikan di setiap area ini disusun dengan memperhatikan data tren pasar dan preferensi pelanggan. Selain itu, pengembangan kampanye pemasaran yang lebih terfokus dan tepat sasaran sangat penting untuk mendorong peningkatan pendapatan dan efisiensi operasional perusahaan. Berikut ini adalah rincian strategi yang dapat diimplementasikan untuk mengoptimalkan kinerja bisnis tersebut:

1. Area yang Perlu Ditingkatkan

- Sales
 - Bandung menghasilkan pendapatan tertinggi untuk hampir semua jenis brand, terutama RedPartner. Namun, Surabaya memiliki pendapatan rendah, terutama untuk Koolkost dan RedPartner.
 - Jakarta menunjukkan pendapatan stagnan sehingga memiliki peluang untuk meningkatkan performa di semua jenis brand.
 - New Users mendominasi pendapatan, terutama untuk perjalanan bisnis. Namun, Repeat Users untuk perjalanan leisure, terutama perempuan, perlu dioptimalkan.
- Pricing
 - Properti RedPartner memiliki jumlah Room Nights lebih rendah dibanding Koolkost dan RedDoorz. Strategi harga yang lebih kompetitif atau penambahan layanan bernilai tambah diperlukan untuk menarik lebih banyak pemesanan.
 - Diskon khusus atau harga paket dapat meningkatkan loyalitas Repeat Users pria di kategori leisure.
- Property Acquisition
 - Properti Koolkost di Jakarta dan Bandung menunjukkan performa konsisten dengan room nights dan pendapatan tinggi. Ekspansi ke kota seperti Bandung, Jakarta, dan Yogyakarta layak diprioritaskan.

- RedPartner sebaiknya mengakuisisi properti baru di Surabaya dan Malang untuk meningkatkan performa, karena pendapatan brand ini sangat rendah di kedua kota tersebut.

2. Rekomendasi untuk Target Marketing Campaign

- Berdasarkan Kota
 - Surabaya dan Malang
 - Fokuskan pada peningkatan kesadaran merek melalui media sosial, promosi lokal, atau kerja sama dengan platform e-booking.
 - Gunakan penawaran promosi harga spesial atau diskon untuk Koolkost dan RedPartner.
 - Mempromosikan konsep staycation bagi eksekutif muda dan keluarga.
 - Jakarta dan Bandung
 - Tingkatkan kehadiran digital dengan mempromosikan properti terbaik melalui ulasan positif dan lokasi strategis.
 - Lakukan upselling atau cross-selling dengan layanan tambahan (paket liburan, transportasi, atau makanan) untuk meningkatkan pendapatan.
 - Investasikan kampanye untuk menonjolkan properti Koolkost dan RedPartner di kawasan bisnis dan leisure utama.
- Berdasarkan User
 - New Users - Leisure

Jalankan promo khusus untuk pemesanan pertama, misalnya di Bandung, dengan fokus pada liburan akhir pekan dan staycation.
 - Repeat Users - Leisure

Buat program loyalitas atau member untuk pengguna bisnis, khususnya perempuan, agar lebih sering melakukan pemesanan.
- Berdasarkan Brand
 - Koolkost

Tonjolkan sebagai akomodasi terjangkau dan nyaman bagi wisatawan hemat, khususnya di Jakarta dan Bandung.
 - RedDoorz

Tingkatkan citra brand dengan memasarkan layanan premium seperti Wi-Fi, coworking space, dan lokasi strategis, terutama di Surabaya.

- RedPartner

Gunakan influencer media sosial untuk menarik demografi muda yang mencari akomodasi stylish dan terjangkau, terutama di Yogyakarta yang merupakan pusat mahasiswa.

BAB IV

Kesimpulan dan Saran

A. Simpulan

Penelitian ini bertujuan untuk menganalisis performa bisnis RedDoorz melalui pendekatan berbasis data. Berdasarkan analisis data yang telah dilakukan, beberapa temuan dapat dirangkum sebagai berikut:

1. Pola perilaku pelanggan

Mayoritas pelanggan cenderung memesan akomodasi dengan durasi yang singkat, yaitu 1-3 hari dan sebagian besar pelanggan melakukan pemesanan dan *check-in* di hari yang sama (35% dari total pemesanan).

2. Kinerja properti

Beberapa properti unggul dalam total pendapatan. Namun, diperlukan strategi untuk meningkatkan pendapatan di properti dengan performa rendah.

3. Pendapatan

Tren pendapatan perusahaan dipengaruhi oleh strategi harga selama musim puncak dan normal. Kota tertentu juga memberikan kontribusi pendapatan terbesar seperti Yogyakarta, Bandung, Jakarta yang menunjukkan bahwa kota-kota ini memiliki potensi pasar yang tinggi. Sedangkan, Surabaya dan Malang menyumbang pendapatan yang lebih rendah sehingga diperlukan strategi khusus untuk peningkatan.

B. Saran

Saran tambahan untuk bisnis perusahaan adalah melakukan diversifikasi strategi pemasaran berdasarkan musim dan segmentasi. Meskipun pemasaran sudah disesuaikan dengan musim (*normal* dan *peak season*), pemasaran bisa lebih dipersonalisasi lagi. Misalnya selama *normal season*, targetkan promosi pada pelanggan dengan tujuan perjalanan bisnis dan pada *peak season* lebih banyak targetkan pada pelanggan dengan tujuan liburan. Selain itu, gunakan juga analitik prediktif untuk memprediksi tren permintaan di masa depan agar membantu perusahaan menyesuaikan strategi lebih tepat waktu.

Lampiran

A. Online Diagram

- Product Canvas:
<https://drive.google.com/file/d/1NjuEO9JrN1jWvkW0NA0fZ5Hrr8v-X9ey/view?usp=sharing>
- Flowchart Pengerjaan Case Study:
<https://drive.google.com/file/d/1FVMEOuvCPIGcMSodwtUldqcWHpbS7rRR/view?usp=sharing>
- Business Flow:
https://drive.google.com/file/d/1G_0pLElLk1EURc-d-Ad77fCNHwIINNlQh/view?usp=sharing

B. Python Code

<https://colab.research.google.com/drive/1p31i7CB6NAiXR1fVTmIM4yuplO6dr0al?usp=sharing>

C. Tableau Dashboard

https://public.tableau.com/views/TABLEAUVIZ-PBLKELOMPOK1CS-17update/Dashboard?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

D. Notulen

https://docs.google.com/document/d/1aAUlkEmoAthTiB_bOMHzFc4YfQxSpSebzL54D-8qJDk/edit?usp=sharing

E. Recording

<https://youtu.be/4Gngj7LtrU4>