

Investigate Business Hotel using Data Visualization



Created by:

Suny Guinesya Ardiansyah

sannyguinesya@gmail.com

<https://www.linkedin.com/in/suny-guinesya-ardiansyah-b839761ba/>

Suny adalah seorang *Data Analyst* dan *Business Intelligence Enthusiast* yang mahir dalam SQL, pemrograman Python, dan alat visualisasi data. Dengan pengalaman yang mendalam dalam PostgreSQL, Suny mampu mengelola database dan mengembangkan query yang efisien. Keahlian dalam Python memungkinkannya untuk menganalisis data dan membangun model dengan menggunakan Google Colab dan Jupyter Notebook. Suny juga terampil dalam mengoperasikan Google Looker Studio, Tableau, dan Microsoft Power BI, menciptakan dashboard interaktif dan laporan yang informatif yang mendukung keputusan bisnis yang cerdas.

“Sangat penting bagi suatu perusahaan untuk selalu menganalisa performa bisnisnya. Pada kesempatan kali ini, kita akan lebih mendalami bisnis dalam bidang perhotelan. Fokus yang kita tuju adalah untuk mengetahui bagaimana perilaku pelanggan kita dalam melakukan pemesanan hotel, dan hubungannya terhadap tingkat pembatalan pemesanan hotel. Hasil dari insight yang kita temukan akan kita sajikan dalam bentuk data visualisasi agar lebih mudah dipahami dan bersifat lebih persuasif.”

Optimalisasi Data untuk Analisis: Sebuah Pendekatan Sistematis Menggunakan Google Colab

Langkah 1: Persiapan Data

- **Menggunakan Google Colab:** Alat utama untuk analisis data.
- **Upload Data ke Google Drive:** Data diunggah dan disimpan di Google Drive.
- **Import Data ke Google Colab:** Data yang disimpan diimpor ke Google Colab untuk analisis.

Langkah 2 : Pemeriksaan Data

- **Cek Data dan Isinya:** Memastikan data telah terimport dengan benar.
- **Cek Kesesuaian Tipe Data:** Verifikasi tipe data sesuai dengan kebutuhan analisis.
- **Cek Nilai Null:** Identifikasi dan dokumentasikan nilai yang hilang atau null.

Langkah 3 : Penanganan Data

- **Mengatasi Nilai Null :** Menghapus kolom, menghapus, dan mengganti nilai Null sesuai kondisi.
- **Mengatasi Nilai yang Tidak Sesuai:** Koreksi atau penghapusan nilai data yang tidak sesuai.
- **Menghapus Kolom yang Tidak Digunakan:** Penghapusan kolom yang tidak relevan untuk analisis lebih lanjut.

Penanganan Nilai Null

Presentase Nilai Null

```
[12] print(round(100*(df.isnull().sum()/len(df.index)),2))
```

```
hotel 0.00
is_canceled 0.00
lead_time 0.00
arrival_date_year 0.00
arrival_date_month 0.00
arrival_date_week_number 0.00
arrival_date_day_of_month 0.00
stays_in_weekend_nights 0.00
stays_in_weekdays_nights 0.00
adults 0.00
children 0.00
babies 0.00
meal 0.00
city 0.41
market_segment 0.00
distribution_channel 0.00
is_repeated_guest 0.00
previous_cancellations 0.00
previous_bookings_not_canceled 0.00
booking_changes 0.00
deposit_type 0.00
agent 13.69
company 94.31
days_in_waiting_list 0.00
customer_type 0.00
adr 0.00
required_car_parking_spaces 0.00
total_of_special_requests 0.00
reservation_status 0.00
dtype: float64
```

Penanganan Nilai Null

```
[15] ## Mengganti nilai null pada agent dengan median karena data memiliki distribusi skewed
df['agent'].fillna(df['agent'].quantile(0.5), inplace=True)
```

```
[16] ## Drop kolom company karena data tidak bisa digunakan (memiliki nilai null sebanyak 94%)
df=df.drop(['company'],axis=1)
```

```
[17] ## Menghapus baris yang berisi nilai Null pada kolom city karena nilai null hanya 0.4%
df = df.dropna(axis = 0)
```

Baris yang memiliki nilai Null pada kolom city di drop/hapus karena presentase data Null hanya 0.41%. Sedangkan kolom company di drop/hapus karena hamper 100% tidak memiliki Data. Untuk kolom agent, karena data bersifat skewed, nilai yang kosong diganti oleh nilai median.

Mengganti Value Count yang tidak sesuai

Kolom Meal

```
meal
Breakfast    91863
Dinner       14434
No Meal      10638
Undefined     1165
Full Board    798
Name: count, dtype: int64
```



Penanganan kolom Meal

```
[20] # Mengganti nilai Undefined dengan Break Fast (mengganti dengan nilai yang sering muncul)
      df.replace('Undefined', 'Breakfast', inplace=True)
```

Ada 4 jenis tipe pemberian makanan, ada tipe yang tidak terdefinisi (Undefined). Tipe ini akan menyebabkan kebingungan dalam proses analisis. Maka nilai Undefined pada kolom meal di ganti oleh nilai yang paling banyak muncul yaitu Breakfast.

Membuang Data Tidak Penting

```
print('Presentase Nilai 0 pada kolom adulst sebesar', round(100*(zero_adults/len(df)),2), '%')
print('Presentase Nilai 0 pada kolom children sebesar', round(100*(zero_children/len(df)),2), '%')
print('Presentase Nilai 0 pada kolom babies sebesar', round(100*(zero_babies/len(df)), 2), '%')

Presentase Nilai 0 pada kolom adulst sebesar 0.33 %
Presentase Nilai 0 pada kolom children sebesar 92.78 %
Presentase Nilai 0 pada kolom babies sebesar 99.23 %
```

Penanganan

```
# Hapus kolom children dan babies karena data tidak bisa digunakan
df=df.drop(['children', 'babies'],axis=1)
```

Hampir dari 100% nilai dari kolom children dan babies adalah 0, hal ini menyebabkan bahwa kolom ini tidak bisa digunakan. Maka kedua kolom tersebut di drop(di hapus)

Monthly Hotel Booking Analysis Based on Hotel Type

Penginapan dengan tipe “City Hotel” selalu lebih banyak diminati daripada “Resort Hotel”

Berdasarkan data dari tahun 2017 – 2019

Tahun 2017

Month	Amount Reservation	
	City Hotel	Resort Hotel
Jan	1234	1086
Feb	1654	1246
Sep	1398	1365
Oct	2476	1397
Nov	3529	1555
Dec	3372	1547

Tahun 2018

Month	Amount Reservation	
	City Hotel	Resort Hotel
Jan	3120	1312
Feb	2475	1353
Mar	1364	851
Apr	2371	1486
May	3046	1742
Jun	3561	1843
Jul	3676	1790
Aug	3923	1357
Sep	3131	1421
Oct	3378	1676
Nov	3871	1512
Dec	4219	1957

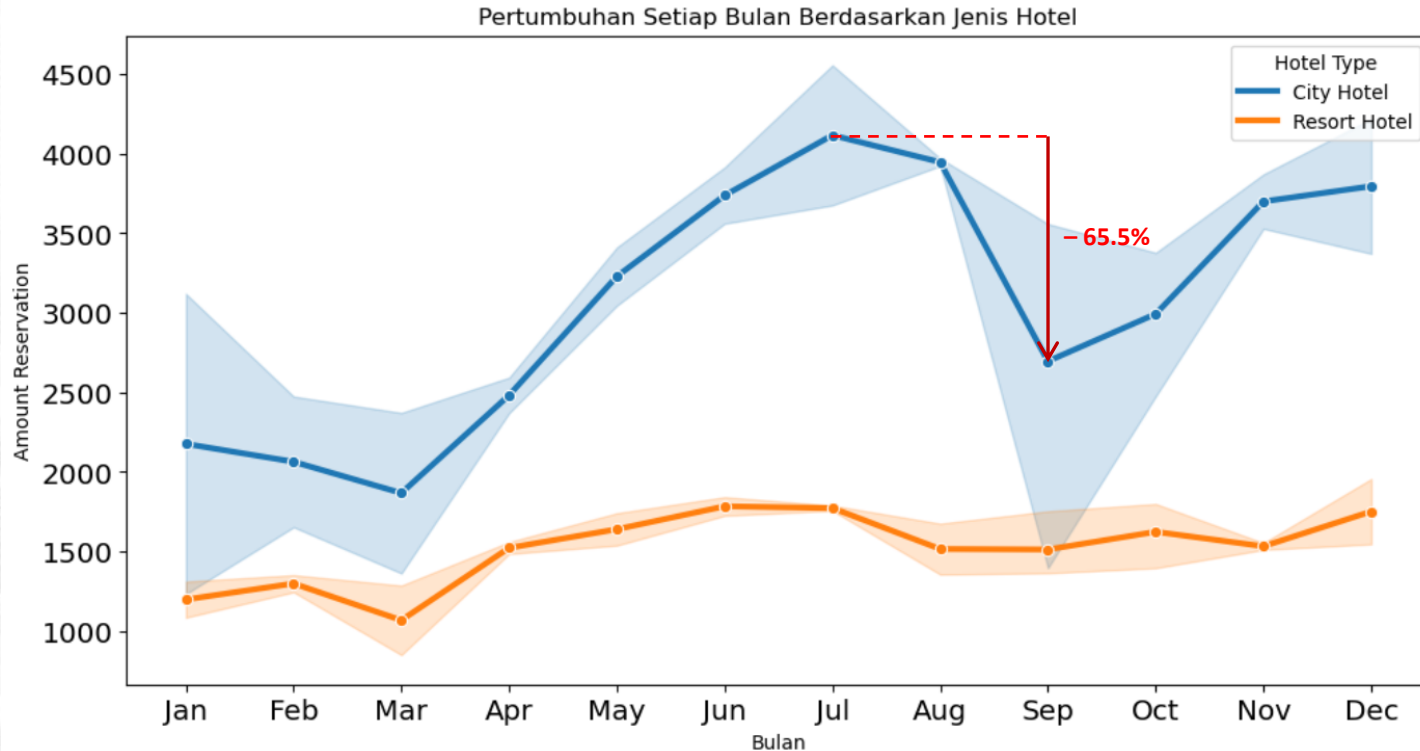
Tahun 2019

Month	Amount Reservation	
	City Hotel	Resort Hotel
Mar	2372	1287
Apr	2594	1561
May	3412	1539
Jun	3915	1726
Jul	4556	1757
Aug	3971	1676
Sep	3559	1754
Oct	3125	1800

Monthly Hotel Booking Analysis Based on Hotel Type

Adanya penurunan yang cukup signifikan pada bulan Agustus – September pada tipe City Hotel

Berdasarkan data dari tahun 2017 – 2019



Hasil Analisis

Berdasarkan data dari tahun 2017 – 2019

Berdasarkan hasil analisis, di dapat bahwa :

- Penginapan yang bertipe City Hotel tidak stabil. Sedangkan Resort Hotel memiliki pertumbuhan yang konstan (naik turunnya cenderung konstan).
- Walaupun tidak stabil, City Hotel tetap menjadi tipe yang paling diminati dibandingkan dengan Resort Hotel.
- Kenaikan yang paling terlihat pada kedua tipe jenis hotel ada di bulan Maret sampai dengan Juni. Hal ini bisa disebabkan karena adanya cuti bersama Idul Fitri dan libur semester anak Sekolah (Sekolah dasar dan Menengah)
- **Uniknya jumlah Reservation untuk City Hotel di bulan Agustus – September mengalami penurunan yang signifikan mencapai 65.5% dari bulan juli. Penyebabnya bisa karena momen di hari kemerdekaan Indonesia yang jatuh pada 17 Agustus juga pada akhir bulan Juli – September di setiap sekolah seluruh Indonesia waktu efektif belajar.**
- Kenaikan terjadi pada akhir tahun yaitu pada November dan Desember. Penyebabnya bisa jadi karena adanya cuti bersama yang cukup Panjang yaitu pada peringatan Natal dan Tahun Baru. Juga sekolah mengalami libur pergantian semester.

Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

City Hotel: Hampir selalu lebih dari 50% pelanggan melakukan pembatalan reservasi

Berdasarkan data dari tahun 2017 – 2019

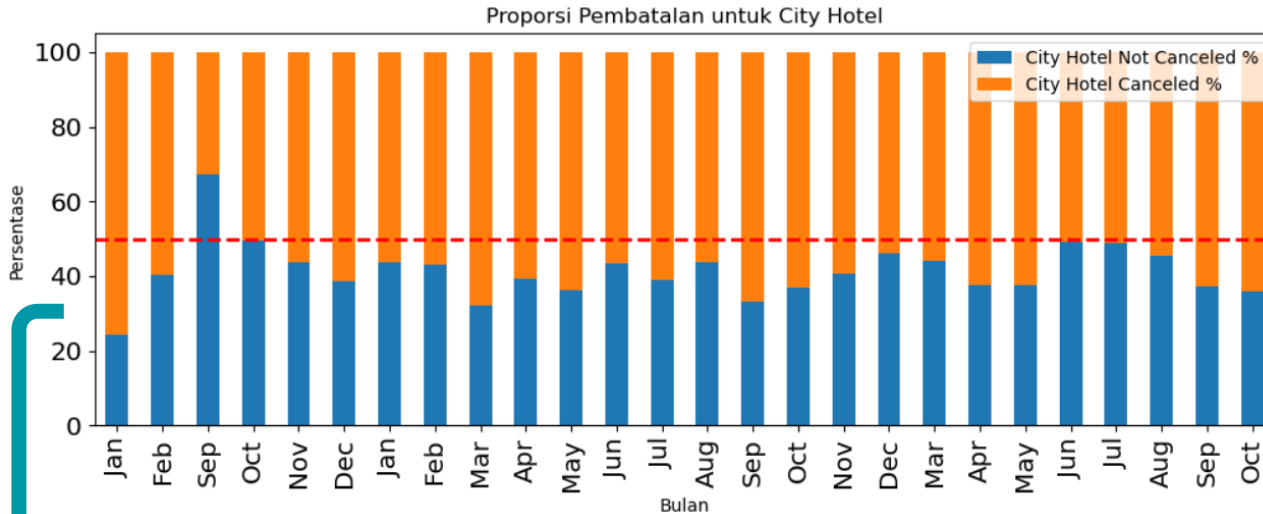
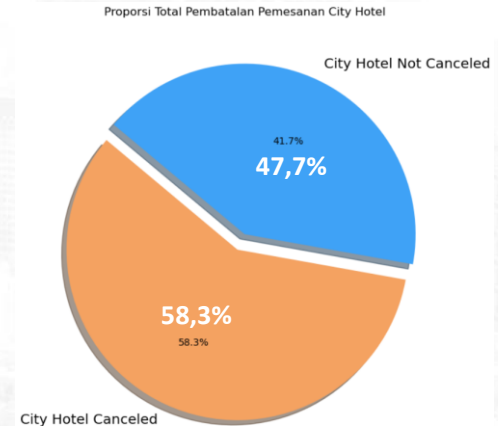


Diagram di atas merupakan diagram untuk melihat presentase pelanggan yang melakukan Check Out dan juga yang melakukan Canceled. Terlihat bahwa hanya ada 1 bulan yang presentase pembatalan kurang dari 40% yaitu pada Oct 2017.



Jika pelanggan yang melakukan reservasi ditotalkan, presentase pelanggan yang membatalkan reservasi mencapai angka 52,3%.

Resort Hotel: Lebih dari 60% setiap bulannya pelanggan melakukan pembatalan reservasi

Berdasarkan data dari tahun 2017 – 2019

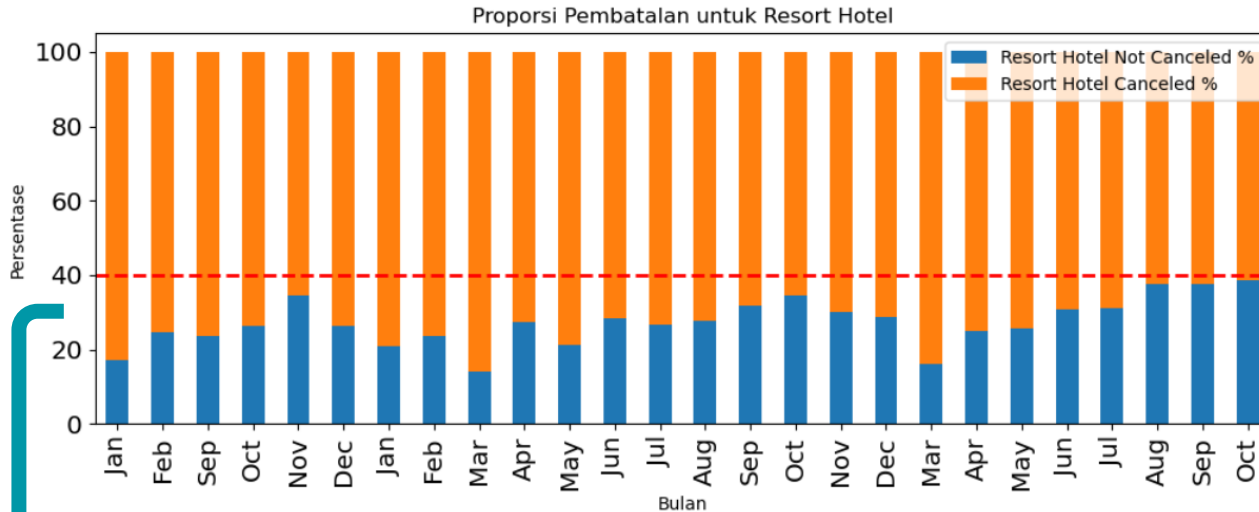
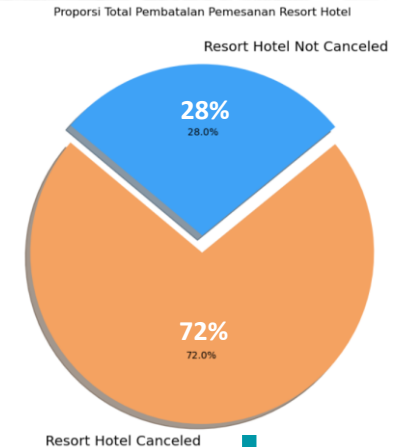


Diagram di atas merupakan diagram untuk melihat presentase pelanggan yang melakukan Check Out dan juga yang melakukan Canceled. Di semua bulan tidak ada yang melebihi 40% total reservasi yang melanjutkan untuk menginap.

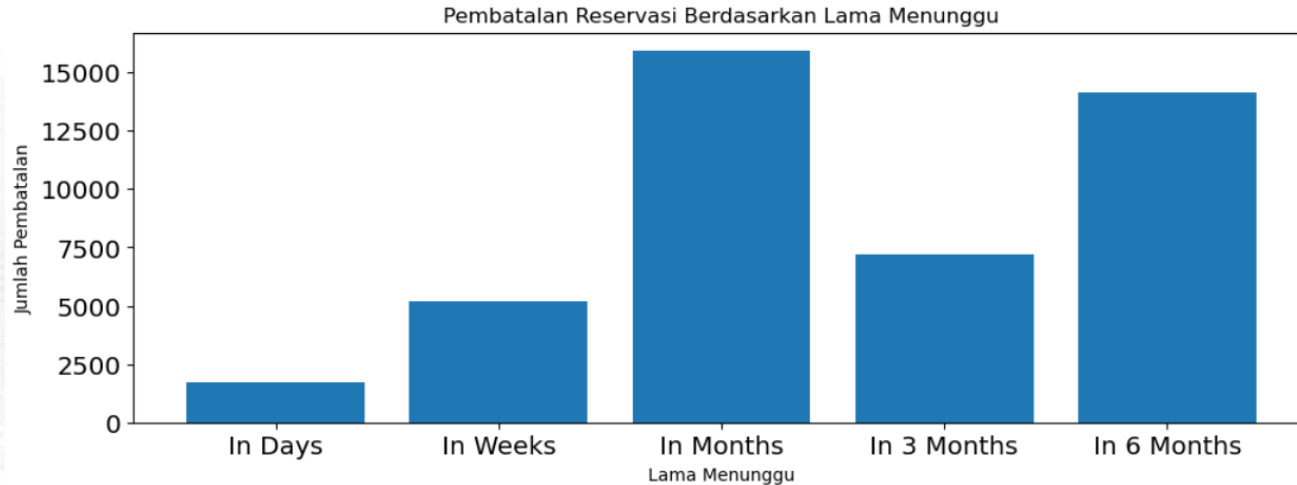


Jika pelanggan yang melakukan reservasi di totalkan, presentase pelanggan yang membatalkan reservasi mencapai angka 72%.

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Pelanggan Hotel Lebih Sering Menginap dengan Waktu Tunggu Singkat

Berdasarkan data dari tahun 2017 – 2019



Keterangan :

- In Days** = Kurang dari 1 Minggu (7 Hari)
- In Weeks** = Kurang dari 1 Bulan (30 Hari)
- In Months** = Kurang dari 3 Bulan (120 Hari)
- In 3 Months** = Kurang dari 6 Bulan (180 Hari)
- In 6 Months** = Lebih dari 6 Bulan (180 Hari)

Dalam rentang waktu tunggu selama 1 sampai 3 bulan, pelanggan banyak yang melakukan pembatalan. Ini menjadi yang paling banyak. Di ikuti oleh pelanggan yang menunggu lebih dari 6 bulan.

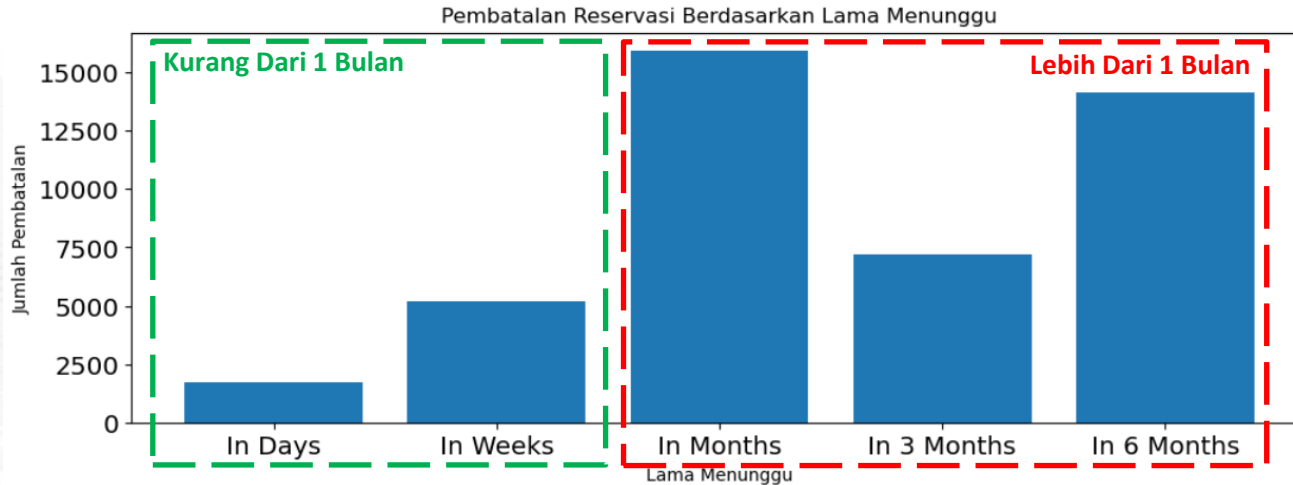


Sebaliknya, pelanggan dengan waktu tunggu singkat cenderung lebih sedikit melakukan pembatalan.

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Pelanggan Hotel Lebih Sering Membatalkan Jika Menunggu Terlalu Lama

Berdasarkan data dari tahun 2017 – 2019



Keterangan :

- In Days = Kurang dari 1 Minggu (7 Hari)
- In Weeks = Kurang dari 1 Bulan (30 Hari)
- In Months = Kurang dari 3 Bulan (120 Hari)
- In 3 Months = Kurang dari 6 Bulan (180 Hari)
- In 6 Months = Lebih dari 6 Bulan (180 Hari)

Jika di bagi dua segmentasi, pelanggan yang banyak melakukan pembatalan reservasi adalah pelanggan yang menunggu lebih dari 1 Bulan.



Sebaliknya, pelanggan yang menunggu kurang dari 1 bulan tidak banyak yang melakukan pembatalan reservasi.

Link Proses Pengerjaan Menggunakan Bahasa Pemrograman Python

Dari dataset `hotel_bookings_data.csv`

Semua File Python hasil pengerjaan ada pada Link Github

<https://github.com/sunyardiansyah/Investigate-Hotel-Business-using-Data-Visualization/tree/main>

Mini Project

Certificate of Completion

has been presented to

Suny Guinesya Ardiansyah

For successfully completing the Investigate Hotel Business using Data Visualization at Bukalapak. Some of the skills learned include Analytical Thinking, Business understanding, and Python Data Visualization

CEO Rakamin Academy

A handwritten signature in black ink, appearing to read "Andika".

Andika Deni Prasetya

Data Scientist,
Bukalapak

A handwritten signature in black ink, appearing to read "Figarri".

Figarri Keisha

Credential: 257412IAPSGIB672024