

Investigate Business Hotel using Data Visualization



Created by:

Haolia

haoliaaul@gmail.com

<https://www.linkedin.com/in/haolia/>

“Geophysical engineering graduates who diligently want to fulfil a role where intellectual, integrity, and curiosity are highly valued. Motivated, able to research, design, implement new features and learn various software. Skill handling problems with unique ways to develop innovative solutions. Proficient using Python, SQL, Tableau and other statistical tools for data multi purposes. Looking for opportunities in data analyst, data science, data engineer and Business Intelligence. ”

“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.”

Data Preprocessing

Exploring Data

- Dengan keyword `df.info()`, diperoleh informasi jumlah feature, tipe data.
- Terdapat missing data pada beberapa feature, serta kemungkinan adanya data yang duplikat.

Handling Duplicated Data

- Dilakukan drop row data yang duplikat

Handling Missing Value

- Menghitung persentase missing value, drop feature yang memiliki missing value lebih dari 90 % yaitu feature `company`.
- Mengganti missing value dengan :
- Feature categorical diisi dengan simple imputer dengan data yang sering muncul (`most_frequent`)
- Feature Numerical diisi dengan mean value
- Mengganti value yang tidak sesuai pada feature meal dengan membuat label baru

Cek Uniq and high cardinality

- Pada tahap ini dilakukan pengecekan fitur yang memiliki nilai unik yang sangat tinggi (high cardinality) dan fitur yang hanya memiliki satu nilai unik saja.
- Fitur `deposit_type` sangat didominasi oleh salah satu nilai saja yaitu nilai (no deposit = 98%) dan dibuang pada tahap ini.