



# Final Project



## Deskripsi

Sejak 2008, AirBnB hadir untuk mempertemukan penyedia tempat tinggal (*host*) dan traveler yang ingin merasakan pengalaman bepergian yang berbeda dari biasanya. AirBnB memungkinkan traveler tinggal di rumah orang lokal dengan harga yang terjangkau. Data yang tersedia disini meliputi aktivitas listing dan berbagai metriknya di Kota New York pada tahun 2019.



## Konten

Informasi yang terkandung meliputi semua data terkait host, ketersediaan lokasi, dan metrik penting untuk melakukan prediksi serta menarik kesimpulan.

# Tugas Kelompok A

---

1. Jelaskan apa gambaran besar yang anda dapatkan dari data tersebut.
2. Tunjukkan *properties* dari data pada setiap kolom, seperti *type*, *missing value*, dan lain-lain.
3. Tunjukkan jumlah data, mean, median, dan standar deviasi.
4. Berapa harga minimal dan maksimal?

# Tugas Kelompok A

5. Selesaikan masalah *missing value* dengan cara berikut:
  - a) Drop kolom-kolom yang mengandung *missing value* yang tidak dalam diisi dengan nilai yang logis, seperti nama, host\_id, dan lain-lain.
  - b) Isi *missing value* pada kolom ***last\_review*** dengan data yang paling banyak muncul.
  - c) Isi *missing value* pada kolom ***reviews\_per\_month*** dengan mean.
6. Buat kategori berdasarkan Price: *Cheap*, *Medium*, dan *Expensive*
7. Filter semua data yang memiliki ***last\_review*** di > 2019
8. Visualisasikan ***Price*** untuk membandingkan harga kamar yang tersedia

## Deskripsi

Pernahkah kamu bertanya kapan waktu terbaik untuk memesan kamar hotel?  
Atau berapa lama waktu tinggal untuk mendapatkan harga dengan *rate* terbaik?  
Bagaimana jika kamu ingin memprediksi kapan hotel mendapat pesanan unik?



## Konten

Data mengandung informasi pemesanan hotel dan resort. Informasi yang terkandung meliputi kapan waktu pemesanan, durasi *stay*, jumlah orang dewasa, anak, dan/atau bayi, jumlah tempat parkir yang tersedia, dan lain-lain.

Semua informasi personal terkait pemesan telah dihapus dari data.

## Tugas Kelompok B

1. Jelaskan apa gambaran besar yang anda dapatkan dari data tersebut.
2. Tunjukkan *properties* dari data pada setiap kolom, seperti *type*, *missing value*, dan lain-lain.
3. Tunjukkan jumlah data, mean, median, dan standar deviasi.
4. Berapa nilai minimal dan maksimal untuk kolom yang mengandung data numerik?



## Tugas Kelompok B

5. Selesaikan masalah *missing value* dengan cara berikut:
  - a) Drop kolom-kolom yang mengandung *missing value* lebih dari 50% keseluruhan data.
  - b) Isi *missing value* pada kolom **country** dan **children** dengan data yang paling banyak muncul.
  - c) Isi *missing value* pada kolom **agent** dengan mean.
6. Buat kategori berdasarkan *lead time*: *Short*, *Medium*, dan *Long*
7. Filter semua data yang memiliki **reservation\_status\_date** di > 2017
8. Visualisasikan **arrival\_date\_month** untuk melihat bulan dimana pengunjung terbanyak



**Indonesia AI**  
AI for Everyone, AI for Indonesia

# Terima Kasih!

[Indonesia AI | AI for Everyone, AI for Indonesia](#)

[contact@aiforindonesia.org](mailto:contact@aiforindonesia.org)