



**Kelompok: Mandalika**

**Stage: 1**

**Mentor: Mas Hafizh Adi Prasetya**

**Pukul/ Tanggal: 19.00 /26 Maret 2022**

**Pembagian tugas di stage ini:**

Nama: Arni Cici Suryani- Fakhry Husnul(Descriptive Analysis)

Nama: Riska Diah - Lise Listianti(Univariate Analysis)

Nama: Ardilla Safitri- Jonse Kennedy (Multivariate Analysis)

**Poin pembahasan:**

1. Stage 0 & 1
2. Feedback (Boxplot, Corr Heatmap, Category vs Target, Business Insight.)
3. Q&A
4. Target untuk stage 2

**Hasil Diskusi:**

1. Stage 0: Business metrics untuk supporting harus yang bisa diukur, diberikan saran oleh mentor misal gunakan gmv.
2. Feedback:
  - a. Pada bagian boxplot perhatikan outlier, untuk kedepannya dalam data pre-processing perhatikan outlier tersebut secara statistik dan kontekstual,
  - b. Pada bagian corr heatmap mungkin ada korelasi non-linear dengan menggunakan random forest dan perlu di cek kembali,
  - c. Pada bagian cat vs target karena untuk membandingkan kiri kanan maka pengaturan common\_norm perlu di set False,
  - d. Pada bagian business insight harus lebih spesifik dan mungkin bisa menggunakan kombinasi beberapa kolom.
3. Q&A: Outlier utama di fokuskan pada outlier terjauh dibandingkan banyaknya jumlah outlier.
4. Target untuk stage 2: Melakukan Data Pre-processing, Cleansing, Feature Engineering. Catatan: Pembersihan data dilakukan sebelum dan setelah data di split, over dan under sampling datanya perlu di train.

**Tindak Lanjut:**

1. Stage 0: Menghapus business metrics supporting yang tidak bisa diukur (Satisfaction Score dan Complain), menambahkan business metrics supporting yang berkaitan dengan revenue.
2. Feedback:
  - a. Pada bagian outlier sudah didiskusikan dan akan menjadi fokus diskusi pada saat pre-processing apakah outlier digunakan atau dihapus
  - b. Pada bagian corr heatmap saat ini masih menggunakan korelasi linear dan kedepannya akan mencoba menggunakan korelasi non-linear.
  - c. Pada bagian cat vs target telah diaplikasikan common\_norm=False.
  - d. Pada bagian business insight sudah lebih spesifik dan hasil kombinasi beberapa kolom
3. Q&A: Outlier terfokus pada data terjauh.
4. Target untuk stage 2: Mempersiapkan untuk stage 2.