**Notulen stage 2**

**Mentor:** Kak Salsa  
**Anggota:** Mufti, Syahdilla, Isma, Johannes

### **Pembahasan:**

1. **Penggunaan K-Means**
   * Digunakan untuk memudahkan analisis data.
2. **Model yang Dicoba**
   * Minimal mencoba 3 model:
     + Linear
     + XGBoost
     + Random Forest
   * Status percobaan model:
     + Awalnya direncanakan dua model.
     + Fokus beralih ke clustering karena kendala data yang sangat besar.
     + Data yang masuk ke listing dan kalender mencapai 900 ribu data.
     + Fokus pada clustering untuk memberikan insight.
3. **EDA dan Clustering**
   * Clustering digunakan untuk prediksi.
   * Kategori hasil clustering dapat memberikan insight untuk model prediksi harga.
4. **Arah Pengembangan Clustering**
   * Melanjutkan analisis clustering untuk prediksi.
   * Disarankan agar hasil prediksi dapat dideploy ke dalam dashboard Streamlit.
   * User dapat memasukkan data untuk mengetahui masuk ke cluster mana.
   * Muncul rekomendasi berdasarkan cluster tersebut.
5. **Implementasi dalam Web App**
   * Menyediakan **Top 10** hasil clustering dalam web app Streamlit.
   * Web app Streamlit akan digunakan untuk visualisasi hasil clustering dan prediksi harga.