



Kelompok: e-numpy

Stage: 1

Mentor: Muhammad Mirza Fahmi

Pukul/ Tanggal: 20.00, 27 Maret 2022

Pembagian tugas di stage ini:

Nama: Muhammad Chosasih Mahendra, Samuel Akwila, Ilham Taufiqurrohman, Rama Putra

Data Exploration, EDA, Visualization

Nama: Sonia Epifany, Putri Maylita

Business Insight, Visualization

Poin pembahasan:

1. Mengecek kesesuaian tipe data dan pengelompokan data (categorical dan numerical)
2. EDA
3. Menggali insight dari EDA

Hasil Diskusi:

- | | |
|---|---|
| <ol style="list-style-type: none">1. Menyesuaikan kembali pengelompokan data categorical2. Mengganti countplot categorical dari vertical menjadi horizontal3. Coba visualisasikan menggunakan catplot4. Cek korelasi ke variable target dan korelasi antar fitur | <ol style="list-style-type: none">5. Job Involvement cek datanya apakah informasinya hilang atau memang seperti itu6. Dilihat kenapa Job Level 1 lebih banyak attritionnya dan cari korelasinya sama variable lain7. Coba cek korelasi DistanceFromHome dengan variable lain, misal Age dkk |
|---|---|

Tindak Lanjut:

- | | |
|--|--|
| <ol style="list-style-type: none">1. Menyesuaikan lagi pengelompokan data categorical2. Mengubah plot di bagian categorical countplot menjadi horizontal, karena xlabel gak kelihatan.3. Menggali lagi korelasi antar variable | <ol style="list-style-type: none">4. Menggunakan catplot di bagian multivariate categorical5. Perlu hilangkan outlier nanti di stage 26. Perlu melihat efek over/undersampling kepada variable target di stage 3 |
|--|--|