

Homework 3 – DataFrame

Data Science Batch 31 Digital Skola

Gunakan sheet “**Paid-Transaction**” dan “**Seller**” di berkas excel “Homework-DataFrame-Dataset.xlsx” sebagai sumber data untuk diimpor ke DataFrame, lalu jalankan instruksi dan jawablah pertanyaan di bawah ini.

1. Buatlah kolom baru bernama “Full Name” dan sisipkan kolom tersebut setelah kolom “Last Name” yang berisikan nama lengkap dari setiap isian record dengan ketentuan sebagai berikut:
 - Nama lengkap menggabungkan “First Name” dan “Last Name”.
 - Nama lengkap menggunakan “title case”, yaitu untuk setiap kata, huruf awalnya harus huruf kapital.
2. Tambahkan sebuah kolom setelah kolom “Seller Discount” dan beri nama kolom tersebut sebagai “GMV”. GMV adalah Gross Merchandise Value yang nilainya dihitung berdasarkan “Transaction Amount” dikurangi “Seller Discount” ditambah dengan “Delivery Fee”.
3. Kelompokkan setiap Meta Category menjadi 2 group sesuai tabel di bawah ini. Buatlah DataFrame baru yang menyimpan pivot table dari DataFrame “Paid-Transaction” di atas untuk menampilkan jumlah total nilai GMV **per-bulan** untuk setiap grup.

| Group 1 | Group 2 |
|------------------------|------------------|
| Home/ Food | Gadget/ Komputer |
| Sports/ Hobi/ Otomotif | Elektronik |
| Fashion | Service/ Mokado |
| Beauty/ Health | Babies/ Kids |

4. Buatlah python statement untuk mencari Seller mana yang memiliki nilai GMV paling tinggi di bulan Agustus 2017. (Hint: gunakan juga sheet “Seller”)
5. Buatlah python statement untuk mencari Seller mana yang memiliki banyaknya transaksi (count) paling banyak di bulan September 2017 khusus untuk Meta Category “Fashion”. (Hint: gunakan juga sheet “Seller”)

Kumpulkan dalam bentuk Google Colab Notebook, dan jangan lupa aksesnya dijadikan **public**. Dilarang keras melakukan tindakan plagiarisme.

Happy Programming :)