

## HOMEWORK

### RAKAMIN TRIAL CLASS

Setelah mengikuti 4 sesi dari program Rakamin Trial Class, tentunya kamu pasti semakin memahami bagaimana pengolahan data menggunakan python. Kini saatnya untuk kamu mengasah kemampuan dengan mengerjakan mini case dibawah ini. Kamu dapat menggunakan jupyter notebook atau google colab.

#### Mengukur Performa Penjualan Ritel Online

Bayangkan kamu sedang bekerja di sebuah perusahaan ritel online, dan pelanggannya sudah tersebar dari berbagai negara. Kamu diberikan data yang berisi semua transaksi yang terjadi di tahun 2009-2011. Saat ini kamu diminta untuk menganalisis bagaimana performa penjualan dalam kurun waktu 3 tahun terakhir.

Link download dataset:

[https://drive.google.com/drive/folders/1UbsUuQJgkF-7ilhhNL2tOnpzU-WS\\_Fu?usp=sharing](https://drive.google.com/drive/folders/1UbsUuQJgkF-7ilhhNL2tOnpzU-WS_Fu?usp=sharing)

Berikut adalah penjelasan untuk masing-masing kolom.

#### Attributes

- **Invoice** : Nomor invoice 6 digit yang ditetapkan secara unik untuk setiap transaksi. Jika kode ini dimulai dengan huruf 'C', itu menunjukkan pembatalan.
- **StockCode** : Kode produk (barang). Angka 5 digit yang ditetapkan secara unik untuk setiap produk yang berbeda.
- **Description** : Nama produk.
- **Quantity** : Jumlah kuantitas setiap produk per transaksi.
- **InvoiceDate** : Tanggal dan waktu invoice, yakni hari dan waktu saat transaksi dibuat.
- **UnitPrice** : Harga satuan atau harga produk per unit dalam sterling (£).
- **CustomerID** : Nomor 5 digit yang ditetapkan secara unik untuk setiap pelanggan.
- **Country** : Nama negara tempat tinggal pelanggan.

#### Section 1: Menganalisis Rata-Rata Pendapatan Per Tahun

Lakukan langkah-langkah dibawah ini:

1. Buat kolom baru dengan nama Year yang berisi nilai tahun dari Invoice Date

- Ubah tipe data kolom InvoiceDate menjadi tipe 'datetime'
  - Gunakan function dari library pandas untuk mendapatkan tahun dari kolom InvoiceDate  
`DatetimeIndex(data['InvoiceDate']).year`
2. Buat filtering data dengan ketentuan di bawah ini dan simpan dalam variabel baru, misalnya sales
    - Quantity minimal 1 (tidak boleh 0 dan minus)
    - Kolom Invoice tidak mengandung huruf 'C' karena hal tersebut menandakan pelanggan tidak menyelesaikan belanjanya atau melakukan pembatalan.
  3. Buat kolom baru bernama Revenue dengan nilai Quantity dikali dengan Price
  4. Hitung rata-rata Revenue per tahun. Akan lebih baik jika kamu juga membuat visualisasinya (grafik).
  5. Tuliskan interpretasi-mu dari output nomor 4

## Section 2: Menganalisis Transaksi Pelanggan Per Tahun

Lakukan langkah-langkah di bawah ini:

1. Lakukan filtering menggunakan data sales (data yang sudah di filter pada section 1) dengan ketentuan CustomerID tidak boleh kosong atau null. Kemudian simpan dalam variabel finished.
2. Lakukan filtering data untuk mengelompokkan pelanggan yang membatalkan belanjanya, dengan cara mendeteksi kolom Invoice mengandung huruf 'C'. Kemudian simpan dalam variabel baru bernama cancel.
3. Hitung jumlah transaksi yang berhasil (dari variabel finished) dan jumlah transaksi yang dibatalkan (dari variabel cancel) untuk setiap tahunnya. Akan lebih baik jika kamu juga membuat visualisasinya (grafik).
4. Hitung cancellation rate untuk setiap tahunnya.  
Cancellation rate adalah persentase pelanggan yang melakukan pembatalan order yang telah dilakukan. Formulasnya adalah jumlah customer yang cancel dibagi jumlah seluruh customer kemudian dikali 100%.
5. Bandingkan hasil output nomor 3 dan 4 untuk setiap tahunnya, dan tuliskan interpretasi-mu.

Selamat mengerjakan. Good luck! :)

Ketentuan:

- Mini case ini bersifat opsional, boleh dikerjakan atau tidak.
- Download template pengerjaan mini case [disini](#).
- Bagi yang ingin mengerjakan, batas pengumpulan pada tanggal 10 Oktober 2021 pukul 23:59.
- Kumpulkan dalam bentuk upload file dengan format .ipynb
- Namakan file tersebut dengan nama masing-masing.