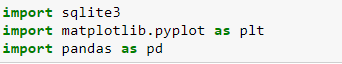
**TUGAS 5**

**“Analisis Data Pembelian di Bulan Februari – Juni”**

Analisis ini menggunakan data *Chinook* untuk menampilkan pembelian items yang dilakukan oleh beberapa custober di bulan Februari – Juni tahun 2009. Data yang akan ditampilkan adalah nama customer dan total pembelian yang dilakukan oleh setiap customer di rentan bulan Februari – Juni tahun 2009. Berikut penjelasan detail dari analisis ini :

1. **Code program**

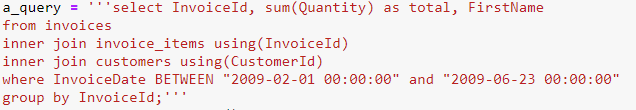
Pemrosesan data dan visualisasi yang dilakukan menggunakan bahasa pemrograman python. Langkah pertama yang dilakukan untuk mengolah data adalah menyiapkan *library* yang dibutuhkan dalam analisis ini.



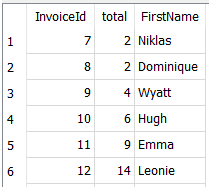
Selanjutnya melakukan koneksi terhadap basisdata yang akan digunakan dengan menggunakan code sebagai berikut.



Menuliskan *query* untuk menampilkan tabel-tabel yang akan digunakan untuk proses analisis data pembelian. Berikut  *query* yang digunakan.



Dalam *query* tersebut dilakukan fungsi untuk menampilkan tabel *InvoiceId,*  jumlah *Quantity* dan nama *customer.* Jika *query* tersebut dijalankan dengan aplikasi *DB Browser* akan menunjukan tampilan seperti berikut :



Data yang dibutuhkan sudah tersedia, kemudian proses yang akan dilakukan selanjutnya adalah menjalankan *query* yang telah disimpan ke dalam atribut *a\_querry* dan melakukan pengaplikasian perubahan ke dalam tabel. Berikut code yang dituliskan.



Setelah menyimpan perubahan dari *query,* kemudian proses yang dilakukan adalah memasuka data perubahan itu kedalam format dataframe untuk dilakukan visualisasi data. Data tersebut kemudian disimpan dalam format dataframe dengan nama “df”. Berikut code yang dituliskan.

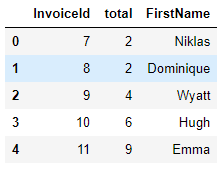


Tahap perakhir yang dilakukan adalah menutup koneksi kursor dan basisdata. Berikut code penulisan.

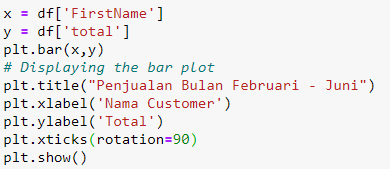


1. **Visualisasi**

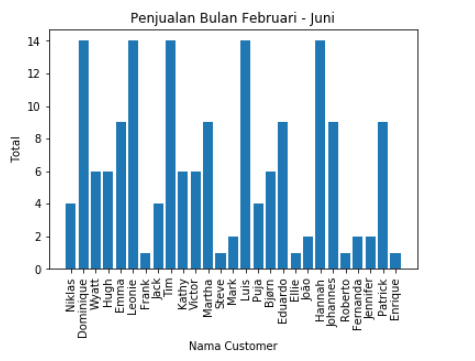
Visualisasi yang dilakukan pada data ini menggunakan visualisasi *bar plot.* Data yang digunakan untuk visualisasi seperti yang sudah dijelaskan diatas sebelumnya. Dalam python tampilan data yang digunakan adalah :



Setelah data yang dibutuhkan benar-benar siap, maka dapat dilakukan tahap visualisasi data dengan menggunakan *bar plot*. Berikut code penulisan.



Dari penjelasan data diatas bahwa variable x digunakan untuk menampung nilai dari tabel *FirstName,* kemudian variable y digunakan untuk menampung nilai dari total *Quantity*. Jika dijalankan akan menghasilkan tampilan visualisasi sebegai berikut.



1. **Analisis hasil data**

Analisis yang didapat dari pengolahan data tersebut adalah:

* Diketahui bahwa nilai pembelian tertinggi yang terjadi di bulan Februrari sampai Juni pada tahun 2009 berada di angka penbelian sebanyak 14 dengan nama customer yaitu Dominique, Leonie, Tim, Luis, dan Hannah. Jika ditotal banyaknya customer dengan nilai pembelian tertinggi sebanyak 5 orang.
* Sedangkan untuk data terendah dalam pembelian items berapada pada angka pembelian sebanyak 1. Data customer tersebut adalah Frank, Steve, Ellie, Roberto, dan Enrique. Jika ditotal banyakkan customer dengan nilai pembelian terrendah sebanyak 5.

Nama ; Mila Putri Kartika Dewi