

## Tugas Day 5 Week 1

### 1. Melakukan analisis terhadap Quantity yang terjual untuk masing-masing Genre musik dari data tahun 2009-2013

- a. Proses melakukan koneksi ke database, melakukan query untuk mengambil data-data yang diperlukan, serta mengubahnya menjadi DataFrame

```
2. #Sales Trend 2009-2013
3. #using google colab
4.
5. import numpy as np
6. import pandas as pd
7. import sqlite3
8.
9. connection = sqlite3.connect('/content/chinook.db')
10.
11.     crud_query = 'select g.name genre,
12.                  sum(i.Quantity) Quantity
13.                  from genres g inner join tracks t on g.GenreId = t.GenreId
14.                  inner join invoice_items i on t.TrackId = i.TrackId
15.                  inner join invoices j on i.InvoiceId = j.InvoiceId
16.                  group by genre'
17.                  order by Quantity
18.
19.     cursor = connection.cursor()
20.     cursor.execute(crud_query)
21.
22.     rows = cursor.fetchall()
23.     my_df = pd.DataFrame(rows)
24.     my_df
25.     my_df1 = my_df.rename(columns={0 : 'Genre',
26.                                     1: 'Quantity (Unit)' })
27.
28.     my_df1
29.
```

b. Hasil DataFrame yang diperoleh

	Genre	Quantity (Unit)
0	Alternative	14
1	Alternative & Punk	244
2	Blues	61
3	Bossa Nova	15
4	Classical	41

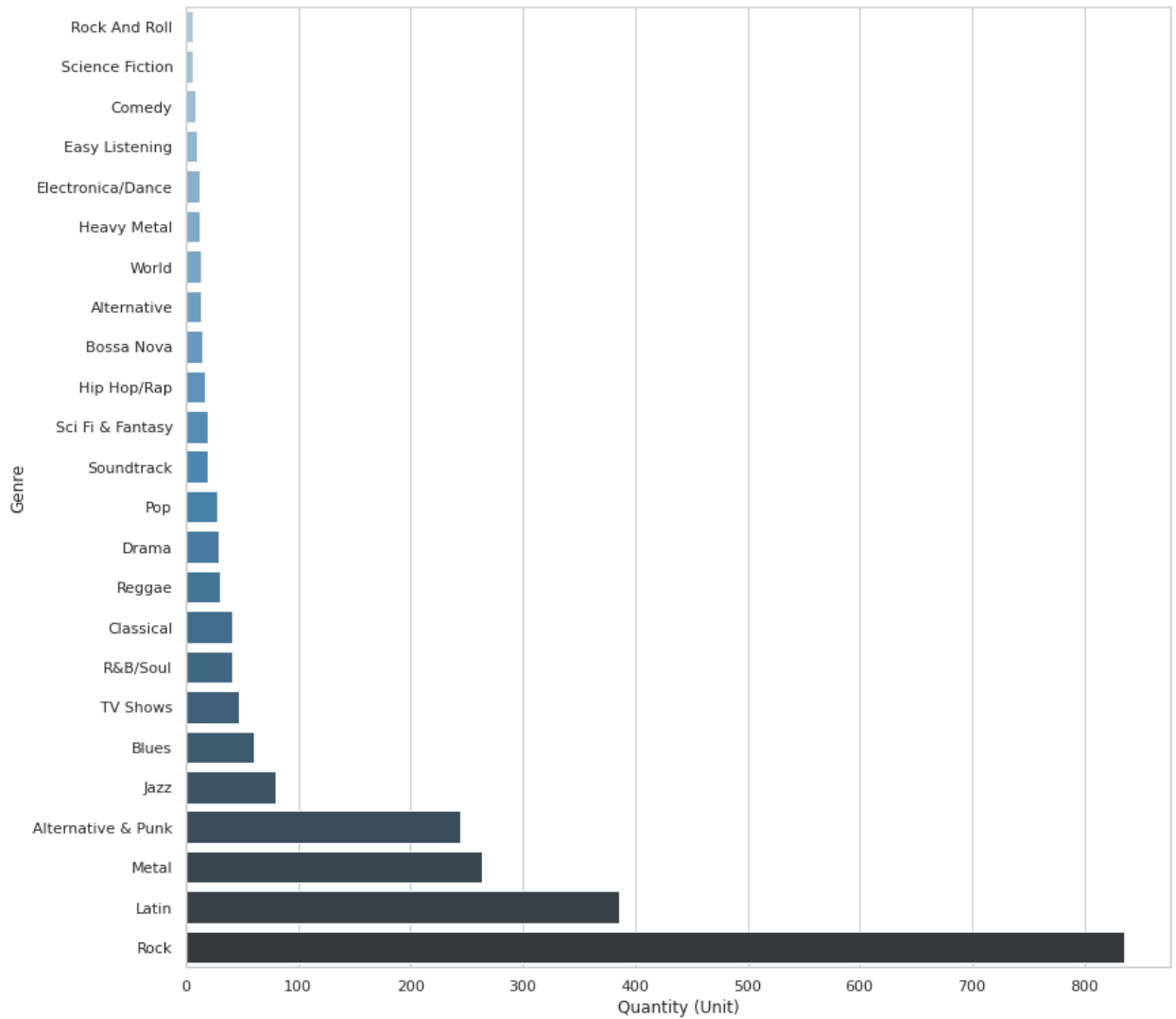
c. Melakukan visualisasi ke dalam bentuk barchart dengan menggunakan seaborn

```
import seaborn as sns
sns.set({'figure.figsize':(13,13)})

sns.set_theme(style="whitegrid")

ax = sns.barplot(x="Quantity (Unit)", y="Genre", data=my_dfl, palette="Blues_d")
```

d. Hasil visualisasi



#### Analisis :

Dari hasil visualisasi tersebut dapat dilihat bahwa Genre yang paling banyak terjual selama tahun 2009-2013 dari database Chinook adalah Genre Rock dan diikuti dengan Latin, sedangkan yang paling sedikit adalah Rock and Roll dan science fiction.

## 2. Melakukan analisis terhadap Quantity yang terjual untuk masing-masing Billing Country dari data tahun 2009-2013

- a. Proses melakukan koneksi ke database, melakukan query untuk mengambil data-data yang diperlukan, serta mengubahnya menjadi DataFrame

```
30.
31.     import numpy as np
32.     import pandas as pd
33.     import sqlite3
34.
35.     connection = sqlite3.connect('/content/chinook.db')
36.
37.     crud_query = 'select j.BillingCountry country,
38.                   sum(i.Quantity) Quantity from genres g
39.                   inner join tracks t on g.GenreId = t.GenreId
40.                   inner join invoice_items i on t.TrackId = i.TrackId
41.                   inner join invoices j on i.InvoiceId = j.InvoiceId
42.                   group by country
43.                   order by Quantity'
44.
45.     cursor = connection.cursor()
46.     cursor.execute(crud_query)
47.
48.     rows = cursor.fetchall()
49.     my_df = pd.DataFrame(rows)
50.     my_df
51.     my_df1 = my_df.rename(columns={0 : 'Country',
52.                                     1: 'Quantity (Unit)' })
53.
54.     my_df1.head()
55.
```

- b. Hasil DataFrame yang diperoleh

	Country	Quantity (Unit)
0	Argentina	38
1	Australia	38
2	Austria	38
3	Belgium	38
4	Chile	38

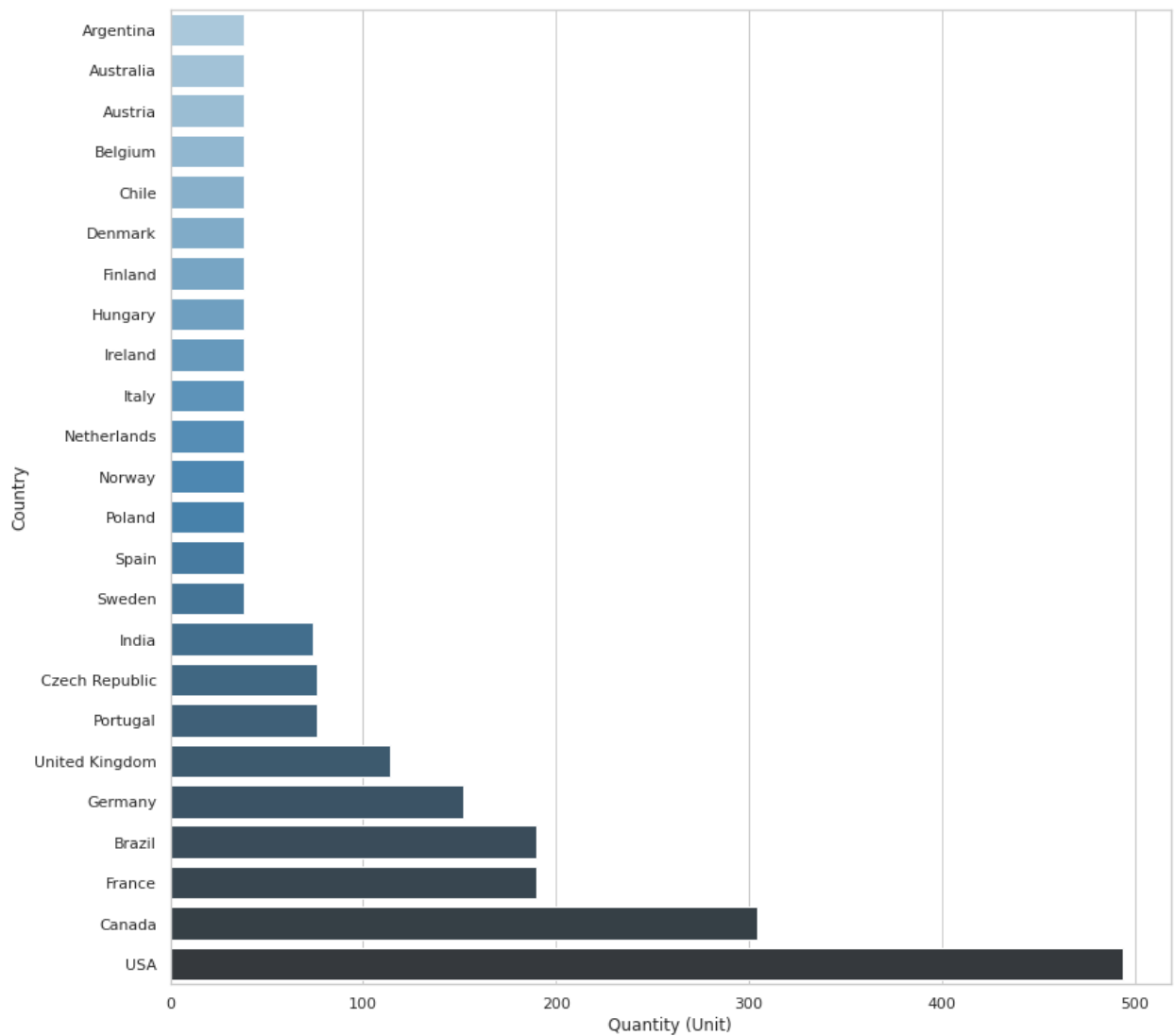
- c. Melakukan visualisasi ke dalam bentuk barchart dengan menggunakan seaborn

```
import seaborn as sns
sns.set({'figure.figsize':(13,13)})

sns.set_theme(style="whitegrid")

ax = sns.barplot(x="Quantity (Unit)", y="Country", data=my_df1,
palette="Blues_d")
```

- e. Hasil visualisasi



Analisis :

Dari hasil visualisasi tersebut dapat dilihat bahwa pembeli paling banyak berdomisili di Amerika Serikat dan Canada, sedangkan pembeli dari Negara lainnya, jumlahnya relative sama.