



Soal Try Out - Data Analytics & Visualization

EDA (Exploratory Data Analysis) World Happiness Report 😊

Disediakan dataset laporan tingkat kebahagiaan di berbagai negara di dunia, unduh di sini: [World Happiness](#). Gunakan hanya dataset **2019.csv**, lalu buatlah sebuah file *notebook* (.ipynb) dan selesaikanlah beberapa soal berikut:

Soal: Buatlah sebuah file *notebook* (**EDA_Happy.ipynb**) dan ikuti panduan *Exploratory Data Analysis* berikut:

1. Untuk *feature/column* **Generosity & Perceptions of corruption**, hitung dan beri penjelasan (*insight*) terkait:
 - Titik Pusat Data (*Central Tendency*)
 - Persebaran Data (*Dispersion*)
 - Distribusi Data (*Distribution*)
 2. Carilah data *outlier* dan beri penjelasan (*insight*) pada *feature/column* **Social support** ! Tambahkan plot untuk menampilkan adanya data *outlier*!
 3. Tampilkan distribusi (sebaran data) dalam bentuk plot yang tepat, untuk *feature/column* **Freedom to make life choices**. Jelaskan *insight* apa yang Anda dapatkan di setiap *plot* tersebut!
 4. Di antara **GDP per capita**, **Social support**, & **Healthy life expectancy**, manakah *feature/column* yang memiliki nilai korelasi tertinggi pada *feature/column* **Score**? Jelaskan metode korelasi apa yang Anda gunakan, serta apa *insight* untuk pemerintah dalam merumuskan kebijakan?
 5. Hitung berapa negara yang **Score** kebahagiaan warganya di atas rata-rata! Dari negara-negara yang **Score**-nya di atas rata-rata tersebut:
 - Tampilkan 5 negara yang **GDP per capita**-nya tertinggi dalam bentuk Barplot!
 - Tampilkan 5 negara yang **Perceptions of corruption** terendah dalam bentuk Barplot!
 - Tampilkan 5 negara yang **Generosity** terendah dalam bentuk Barplot!
 6. Lakukan Uji Hipotesis untuk menguji apakah ada perbedaan **Score** antara negara yang **GDP per capita**-nya di atas rata-rata dengan yang di bawah rata-rata!
- ✓ Di setiap visualisasi, beri penjelasan (*insight*) yang bermanfaat untuk pembangunan suatu negara agar dapat membahagiakan warganya 😊.

Soal 2 - PyMySQL - Distributor Die Cast 🚗 Database

Anda adalah distributor ternama di bidang jual-beli *die cast* (miniatur moda transportasi) yang menjual beragam produk dari vendor-vendor berkualitas, dengan 7 kantor cabang tersebar di beberapa negara dunia. Disediakan database yang menyimpan berbagai data terkait usaha Anda, mulai dari daftar customer, karyawan, produk hingga transaksi yang dilakukan.

Import database **retrowheels** ke local MySQL server Anda. Tata cara import database:

Import retrowheels.sql to your local MySQL

1. Download & place this *retrowheels.sql* on your local MySQL **bin** directory, for example: `C:\Program Files\MySQL\MySQL Server 8.0\bin`.
2. Open terminal/cmd as an administrator, then go to MySQL **bin** directory.

```
$ cd C:\Program Files\MySQL\MySQL Server 8.0\bin
$ mysql -u username -p
```

3. Create database *retrowheels*, and exit.

```
mysql> create database retrowheels;
mysql> exit
```

4. Import *retrowheels.sql* to *retrowheels* database.

```
$ mysql -u username -p -D retrowheels < retrowheels.sql
$ mysql -u username -p

mysql> show databases;
mysql> use retrowheels
```

5. Enjoy!

Soal : Buatlah file *jupyter (.ipynb)* dan buat *connection* Python dengan server MySQL Anda, lalu gunakan database *retrowheels*. **Tuliskan langkah-langkah/query MySQL dan tampilkan hasilnya berupa DataFrame.** Anda **dilarang** membuat database baru, merubah struktur table, membuat view atau segala bentuk tindakan yang mengubah struktur database.

1. Rumuskan *single query* untuk menampilkan jumlah **total customer** yang Anda layani, beserta jumlah **total kota & negara asal** customer-customer Anda. Contoh output yang diharapkan:

```
+-----+-----+-----+
| Customers | Cities | Countries |
+-----+-----+-----+
```

```
|      122 |      95 |      27 |
+-----+-----+-----+
```

Anda memiliki **122 customer** dari **95 kota** yang tersebar di **27 negara**!

2. Rumuskan *single query* untuk menampilkan *resources* yang Anda miliki, mulai dari jumlah **karyawan**, jumlah **kantor & lokasi negaranya**, jumlah **barang yang dijual**, total **stok barang** & jumlah **vendor** yang menjadi partner Anda. Contoh output yang diharapkan:

```
+-----+-----+-----+-----+-----+-----+
| Employees | Offices | Country | Products | StockProducts | Vendors |
+-----+-----+-----+-----+-----+-----+
|      23 |      7 |      5 |      110 |      555131 |      13 |
+-----+-----+-----+-----+-----+-----+
```

Anda memiliki **23 karyawan** di **7 kantor** yang berada di **5 negara**, dengan **110 model die cast** (total stok mencapai **555131 item**) yang didistribusikan dari **13 vendor partner**.

3. Dari soal sebelumnya tercatat Anda memiliki **110 model die cast** dengan total stok **555131 item**. Jika dikategorikan, produk yang Anda jual terbagi menjadi **7 product line die cast**, yakni model mobil klasik, mobil *vintage*, sepeda motor, pesawat terbang, kapal laut, kereta api serta truk & bus. Rumuskan *single query* yang dapat menampilkan **harga produk terendah & tertinggi** untuk masing-masing kategori. Contoh output yang diharapkan:

```
+-----+-----+-----+
| productLine | minPrice | maxPrice |
+-----+-----+-----+
| Classic Cars |      15.91 |      103.42 |
| Motorcycles |      24.14 |      91.02 |
| Planes |      29.34 |      77.27 |
| Ships |      33.30 |      82.34 |
| Trains |      26.72 |      67.56 |
| Trucks and Buses |      24.92 |      84.76 |
| Vintage Cars |      20.61 |      86.70 |
+-----+-----+-----+
```

4. Rumuskan *single query* yang dapat menampilkan daftar 10 customer paling *royal* (paling banyak mendatangkan uang bagi kita), yang total nominal transaksinya paling tinggi. Data yang ditampilkan adalah nama customer, kota & negara asal, beserta total uang yang dihabiskan di produk kita. Contoh output yang diharapkan:

```
+-----+-----+-----+-----+
| customerName | city | country | total |
+-----+-----+-----+-----+
| Euro+ Shopping Channel | Madrid | Spain | 715738.98 |
| Mini Gifts Distributors Ltd. | San Rafael | USA | 584188.24 |
```

Australian Collectors, Co.	Melbourne	Australia	180585.07	
Muscle Machine Inc	NYC	USA	177913.95	
Dragon Souvenirs, Ltd.	Singapore	Singapore	156251.03	
Down Under Souvenirs, Inc	Auckland	New Zealand	154622.08	
AV Stores, Co.	Manchester	UK	148410.09	
Anna's Decorations, Ltd	North Sydney	Australia	137034.22	
Corporate Gift Ideas Co.	San Francisco	USA	132340.78	
Saveley & Henriot, Co.	Lyon	France	130305.35	
+-----+-----+-----+-----+				

5. Pada **2003-06-05**, terdapat pembayaran masuk sebesar **US\$ 14571.44**. Tampilkan data seputar transaksi tersebut, mencakup **nama customer** yang melakukan pembayaran, **nama produk** yang dibeli, **jumlah tiap produk** yang dibeli dan **harga satuannya**. Pastikan total harga yang dibeli sesuai dengan data pembayaran masuk. Output yang diharapkan:

```

+-----+-----+-----+-----+
-----+
| customername      | productname                | quantityOrdered |
priceEach |
+-----+-----+-----+-----+
-----+
| Atelier graphique | 1965 Aston Martin DB5      | 26 |
120.71 |
| Atelier graphique | 1999 Indy 500 Monte Carlo SS | 46 |
114.84 |
| Atelier graphique | 1948 Porsche Type 356 Roadster | 34 |
117.26 |
| Atelier graphique | 1966 Shelby Cobra 427 S/C    | 50 |
43.27 |
+-----+-----+-----+-----+
-----+

```

Jika Anda cek jumlah total dari **quantityOrdered** dikali **priceEach** dari tabel di atas, hasilnya **14571.44**. Dan transaksi pembayaran tersebut terjadi tepat pada tanggal **2003-06-05**.

Good luck & Happy Coding