

Capstone Project - Module 2

Data Source :

<https://www.mysqltutorial.org/wp-content/uploads/2018/03/mysqlsampledatabase.zip>

General Question

1. Bagaimana konteks bisnis berdasarkan dengan data yang telah diberikan? Silahkan buat asumsi atau penamaan jika diperlukan.
2. Ada berapa banyak tabel yang disediakan oleh database yang tersedia? Jabarkan setiap tabelnya. Jika terdapat ERD, silahkan ditampilkan.

SQL

1. Apakah tabel customers, orders, ordersdetail, dan products dapat digabungkan menjadi 1 tabel? Jika memungkinkan, tampilkan tabel yang memuat informasi transaksi yang melibatkan keempat tabel tersebut.
2. Berapa banyak jenis produk yang tersedia? Berapa banyak jumlah produk untuk setiap jenis produknya? Serta berapa total quantity dan total harga beli untuk setiap jenis produknya?

Data Manipulation

1. Apakah terdapat anomaly berupa missing values, data duplicate, kesalahan data formatting, dan/atau inconsistency typing? Jika ada tunjukkan serta lakukan penanganan pada anomaly tersebut.
2. Apakah terdapat tipe data yang berupa datetime pada data? Jika iya, apakah tipe data yang berupa datetime tersebut dapat dicari tahu selisihnya? Silahkan tampilkan hasilnya, berikan insight yang sesuai (jika memungkinkan).
3. Ada berapa banyak transaksi untuk setiap status pengiriman yang ada? Lalu, bagaimana rata-rata dari buyPrice untuk setiap status pengiriman tersebut? Jabarkan dan berikan insight.
4. profitEach adalah sebuah variabel baru yang merupakan hasil pengurangan antara priceEach dan buyPrice. Apakah terdapat outlier pada variable baru yang bernama profitEach tersebut? Jika ada, tunjukkan di data keberapa. Lalu, bagaimana outlier ini akan memberikan pengaruh terhadap keuntungan setiap penjualan barang? Coba jelaskan.

Data Visualisation & Statistics

1. Apakah terdapat pelanggan yang loyal? Jika ada, bagaimana treatment dan strategi yang bisa dimanfaatkan dengan melihat adanya pelanggan setia tersebut?
2. Negara mana yang menjadi pemesan kendaraan terbanyak? Tampilkan dan jabarkan.
3. Bagaimana persentase status pengiriman? Apakah pengiriman yang bermasalah jauh lebih banyak dibandingkan dengan pengiriman yang berhasil?
4. Jenis product apa yang paling laku dan paling banyak terjual?
5. Pada tahun berapa keuntungan terbesar perusahaan didapatkan? Serta pada bulan apa saja penjualan mencapai titik tertingginya?
6. Terdapat 7 product yang tersedia, apakah terdapat perbedaan nilai profit antar setiap product tersebut? Buktikan.
7. Apakah terdapat hubungan antara 'quantityOrdered', 'quantityInStock', 'buyPrice', 'profitEach'? Jika iya, variable mana saja yang paling mempengaruhi satu sama lain? Gambarkan bentuk hubungannya.