**Capstone Project - Module 2**

**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.

**SQL**

1. Apakah tabel customers, orders, ordersdetail, employeeID, dan products dapat digabungkan menjadi 1 tabel? Jika memungkinkan, tampilkan tabel yang memuat informasi yang melibatkan kelima tabel tersebut.

2. Berapa jumlah produk yang terjual? Berapa jumlah customer yang melakukan pemesanan?

3. Berapa total pemasukan pendapatan untuk setiap jenis produknya? Serta berapa minimal gaji dan maksimal gaji karyawan di perusahaan tersebut?

**Data Manipulation**

1. Apakah terdapat anomaly berupa missing values, data duplicate, kesalahan data formatting, dan outlier? Jika ada tunjukan 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 tentang rentang tahun karyawan bekerja di perusahaan tersebut? Silahkan tampilkan hasilnya.

3. Apakah terdapat outlier pada setiap feature tersebut? Jika ada, tunjukan di data keberapa. Lalu, bagaimana outlier ini akan memberikan pengaruh terhadap data tersebut? Coba jelaskan.

**Data Visualisation & Statistics**

1. Pada tahun berapa keuntungan terbesar perusahaan didapatkan?
2. Berapa jumlah penghasilan tiap karyawan terhadap penjualan produk? Berikan penjelasan.
3. Berapa jumlah customer yang order dan berapa jumlah produk yang terjual? Serta apa yang bisa dilihat dari visualisasi tersebut?
4. Terdapat 4 departemen didata karyawan, apakah tinggi gaji yang dikeluarkan untuk karyawan sesuai dengan hasil penjualan produk yang didapatkan oleh karyawan tersebut? Buktikan.
5. Ada feature baru yang namanya total\_pemasukan. Apakah feature tersebut terdistribusi normal? Jelaskan.
6. Apakah terdapat hubungan antara 'salary', ‘unitprice’, ' quantity', ' total\_pemasukan', ‘quantity\_product’, ‘jumlah\_pemesanan’? Jika iya, variable mana saja yang paling mempengaruhi satu sama lain? Gambarkan bentuk hubungannya.