

# Capstone Project – Modul 2

## NORTHWIND

### General Question

1. Bagaimana konteks bisnis berdasarkan dengan data yang telah diberikan?  
Silakan buat asumsi dan penamaan jika diperlukan
2. Ada berapa banyak table yang disediakan oleh database yang tersedia?  
Jabarkan setiap tabelnya.

### SQL

1. Apakah table supplier, products, orderdetails, orders dan categories dapat menjadi satu table? Jika memungkinkan tampilkan table yang memuat informasi tersebut!
2. Berapa banyak supplier yang tersedia serta total revenue untuk setiap order?
3. Berapa banyak jumlah order yang tersedia serta kuantitas tiap order?

### Data Manipulation

1. Apakah terdapat anomaly berupa missing values, data duplicate, serta kesalahan data formatting? Jika ada tunjukkan dan lakukan penanganan pada anomaly tersebut.
2. Apakah terdapat tipe data berupa datetime? Jika iya apakah tipe data tersebut dapat dicari selisihnya? Silakan tampilkan hasilnya, berikan insight yang sesuai jika memungkinkan
3. Ada berapa transaksi yang ada? Buat variable baru yang merupakan hasil perkalian antara unit price dan quantity pada order?
4. Restunit adalah variable baru yang merupakan pengurangan antara unit in stock dan unit in order. Apakah terdapat outlier atau data yang tidak konsisten? Tunjukkan!

### Data Visualization

1. Berapakah proporsi pengiriman antara jalur pengiriman satu sama lain?
2. Negara supplier manakah yang memiliki revenue paling tinggi dibandingkan negara supplier yang lain?
3. Apa nama perusahaan supplier yang memiliki jumlah pendapatan tertinggi dibandingkan dengan yang lain? Bentuk dalam Treemap!
4. Berapakah jumlah pendapatan berdasarkan kategori produk? Bentuk dalam bar plot horizontal!
5. Gambarkan supplier dan kategori dengan revenue tertinggi. Apakah perusahaan yang mendiversifikasi produk supplynya memiliki pendapatan tertinggi dibandingkan perusahaan yang tidak mendiversifikasi produk? Jelaskan!
6. Berapakah jumlah pesanan antar negara? Negara manakah yang memesan produk paling banyak?

7. Seperti apakah perbandingan unit in stock dan unit in order tiap bulan? Bentuk dalam lineplot!
8. Seperti apakah perbandingan unit in stock dan unit in order tiap tahun? Bentuk dalam lineplot!
9. Gambarkan boxplot dari order per kategori! Apakah terdapat nilai outlier dan bagaimanakah persebaran datanya?

### **Statistik**

1. Apakah kategori makanan memiliki dampak signifikan pada Revenue? Jabarkan hipotesisnya menggunakan aplikasi statistik.
2. Apakah kategori makanan memiliki dampak signifikan pada kuantitas pesanan? Jabarkan hipotesisnya menggunakan aplikasi statistik.