

Northwind Employee Data Analysis

For Capstone Project Module 2

Data Source:

https://drive.google.com/drive/folders/1fTHrwh_gcLsOFKXHnUzUGEu_APxLoD9i

General Question

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

SQL

1. Apakah beberapa tabel khususnya tabel employee, employee territories, territories, region, dan order dapat digabungkan menjadi 1 tabel? Jika bisa, tampilkan tabel yang memuat informasi transaksi yang melibatkan kelima tabel tersebut.
2. Berapa banyak title dan jumlah employee perusahaan Northwind? Bagaimana sebaran rentang usia tiap employee perusahaan Northwind? Serta bagaimana sebaran negara (country) dari employee perusahaan Northwind?

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 dataset yang digunakan dapat ditambahkan feature baru yang dapat memberikan insight lain dari analisis yang dibuat? Seperti contohnya membuat feature usia dengan memanfaatkan tipe data datetime dari selisih antara tanggal lahir employee dengan waktu employee sekarang bekerja. Jika iya, buat dan jelaskan (Jika memungkinkan).
3. Ada berapa banyak order transaction yang dibuat oleh employee perusahaan Northwind? Lalu kategorikan berdasarkan usia dari employee. Jabarkan dan berikan insight.
4. Bagaimana salary atau gaji yang diterima tiap employee perusahaan Northwind berdasarkan durasi kerja employee. Jabarkan dan berikan insightnya.
5. Apakah terdapat outlier pada variable dalam dataset yang digunakan? Jika ada, tunjukkan di data keberapa atau bagian mana. Lalu bagaimana penanganan outlier ini, apabila outlier dianggap tidak memberikan pengaruh yang signifikan cukup tampilkan saja. Jika sebaliknya, lakukan penanganan sesuai interpretasi analisa data yang dibuat (Jika memungkinkan).

Data Manipulation

1. Gender mana yang menjadi jumlah terbanyak dari employee perusahaan Northwind? Tampilkan dan jabarkan.
2. Apa kategori umur dari employee yang mendominasi pelayanan order dari customer dari perusahaan Northwind? Apa insight yang dapat diambil, jabarkan dan jelaskan.
3. Bagaimana distribusi employee perusahaan Northwind berdasarkan territory tempat employee bekerja?
4. Bagaimana presentase employee perusahaan Northwind berdasarkan region tempat employee bekerja?
5. Terdapat beberapa rentang salary yang diterima oleh tiap employee perusahaan Northwind dan juga lama waktu bekerja employee terhadap perusahaan, apakah terdapat perbedaan nilai salary antar setiap employee antar setiap lama waktu bekerja tersebut? Buktikan.
6. Apakah terdapat hubungan antara 'Salary', 'DurationWorkingTime', dan 'Age'? Jika iya, variable mana saja yang paling mempengaruhi satu sama lain? Gambarkan bentuk hubungannya.