

## Supplerende opgave – Northwind Foods

### Baggrund

Northwind Foods er en fiktiv international fødevarevirksomhed, der sælger et bredt udvalg af fødevareprodukter til både private kunder og virksomheder. Virksomheden har en omfattende database, der indeholder vigtige oplysninger om deres produkter, salgshistorik, og kunder. Databasen indeholder data fra flere år, og ledelsen ønsker at få indsigt i forskellige nøgleområder for at optimere forretningsstrategierne. Der vil være fejl i data sættet, så det er ikke altid sikkert at det forventede resultat stemmer overens med hvad du finder. Skriv de fejl du kan finde ned

**Din Rolle:** Du arbejder som dataanalytiker for Northwind Foods. Din leder har givet dig en liste over 10 nøgleinformationer, som virksomhedens topchefer ønsker at vide mere om. Det er din opgave at udtrække, analysere og præsentere disse data på en måde, der kan hjælpe ledelsen med at træffe strategiske beslutninger.

### Del 1: Dataoprettelse

Du skal oprette data til opgaverne ved at afvikle SQL-fil, **northwind.sql**, i MySQL Workbench.

Det er en eksempelsdatabase fra Microsoft over en fiktiv virksomhed, **Northwind**, som sælger forskellige produkter.

### Del 2: Grundlæggende dataudtræk

Du skal oprette følgende 10 dataudtræk, altså følgende SQL-sætninger, en for hver opgave, men alle SQL-sætninger skal være samlet i en SQL-fil – Gem SQL filen som: **uge3\_opgave4.sql**

1. Tegn et **ER-diagram** over Northwind databasen. I bestemmer selv hvilket program i brugere til dette, det kan også "*bare*" være på papir og et billede af dette. Afleveres som en PDF-fil.
2. List alle Produkter fra tabellen Products. Sorter faldende efter feltet UnitPrice
3. Find alle kunder fra UK og Spain.  
Brug tabellen Customers  
*Resultat: 12 Rows*
4. Find alle Produkter hvor vi har mere end 100 på lager (UnitsInStock) og prisen (UnitPrice) er større end eller lig med 25.  
Brug tabellen Products.  
*Resultat: 10 Rows*
5. Find alle lande der er sendt en ordre til, vis dem kun én gang  
Brug tabellen Orders  
*Resultat: 21 Rows*

6. Find alle ordrer der har en OrdreDate i 10 måned af 1996.  
Brug tabellen Orders.  
*Resultat: 26 Rows*
7. Find alle ordrer hvor ShipRegion er blank, ShipCountry = Germany, Freight er større end eller lig med 100, EmployeeID = 1 og OrdreDate er fra 1996.  
Brug tabellen Orders  
*Resultat: 2 Rows*
8. Find de ordre der ikke er leveret til tiden, ShippedDate er større end RequiredDate  
Brug tabellen Orders.  
*Resultat: 37 Rows*
9. Find alle Ordre (OrderDate) fra 1997 i månederne januar, februar, marts og april fra Canada (ShipCountry)  
Brug tabellen Orders.  
*Resultat: 8 Rows*
10. Find de ordrer hvor EmployeeID er lig med 2, 5 eller 8. ShipRegion ikke er '' og ShipVia enten er 1 eller 3. Skal sorteres først efter EmployeeID derefter ShipVia begge ascending  
Brug tabellen Orders.  
*Resultat: 57 Rows*
11. Opgave 3 Find de Employees hvor der ikke er en værdi (' ') i Region eller ReportsTo er "blank" (NULL)  
Desuden skal de være født i 1960 eller tidligere  
Brug tabellen Employees.  
*Resultat: 3 Rows*

### Del 3: Analyse og rapportering med Python og SQL

I denne del af opgaven skal du udvikle et Python-script, der integrerer SQL til at udtrække data, bruger Pandas til dataanalyse og anvender matplotlib til at visualisere resultaterne. Formålet er at skabe et workflow, hvor du effektivt kan trække data fra en database, bearbejde og analysere det i Python, og præsentere resultaterne i form af visualiseringer.

Det er også **Northwind databasen** I skal bruge i denne opgave.

#### Opgavebeskrivelse:

- Brug Python og SQLite3-pakken til at forbinde til Northwind databasen.
- Brug SQL til at hente data fra de forskellige tabeller.
- Brug Pandas til at indlæse dataene og udføre en analyse for at finde salget for hvert land.
- Brug matplotlib til at lave et søjlediagram, der viser salget for hvert land.
- Analyser data og opret 2-3 yderligere relevante diagrammer ud fra data

**Aflevering**

- ER-diagram afleveres som en PDF-fil.
- I skal aflevere en SQL-fil, med de 10 SQL-sætninger
- Link til github-projekt med kode