

Teoretiska Frågor:

1- Relationell databas:

En relationell databas innebär en logisk koppling eller relation mellan två och flera tabeller i databasen. Vi kan samla information från olika tabeller och organisera dessa data som tabeller enligt våra behov.

2- CRUD:

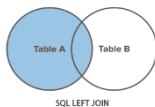
CRUD hänvisar till de fyra funktionerna:

- Skapa (*Create*): För att skapa eller infoga information.
- Läsa (*Read*): För att läsa information.
- Uppdatera (*Update*): För att uppdatera information.
- Radera (*Delete*): För att radera information.

De fyra funktionerna beskriver operationer för hantering av data.

3- Left Join:

Left Join returnerar information från den vänstra tabellen (A) och matchande information från den högra tabellen (B). Se visualisering:



Dock returnerar en Inner Join information som har matchande värden i båda tabellerna. Se visualisering:



4- Indexering:

Vi använder indexering i bakgrunden av databasen för att påskynda hämtning av data och öka effektiviteten för att hitta data.

5- Vy:

En vy är en virtuell tabell som innehåller fält (rader och kolumner) som en verklig tabell. Dessa fält kommer från en eller flera verkliga tabeller i databasen. Vi kan presentera vyn som om data kom från en enda tabell.

6- Lagrad procedur i SQL:

Vi använder lagrad procedur för att gruppera en eller flera transaktioner i logiska enheter, och den lagras som ett objekt. Sedan anropas den för att utföra genom att använda (EXEC).

Deskriptiv sammanfattning

AdventureWorks2022 databas är designad för Microsoft SQL Server, har en komplex och omfattande struktur som representerar olika aspekter av en typisk affärsorganisation. Dess struktur kan sammanfattas enligt följande:

1. Schema Organisation

Databasen är uppdelad i flera scheman som vardera representerar ett eget affärsområde, såsom personal, försäljning, produktion och inköp. Detta modulära tillvägagångssätt hjälper till att organisera data och operationer logiskt.

2. Tabelltyper

Business Entity Tables: Lagra information om enheter som kunder, anställda och leverantörer.

Product Tables: Inkludera information om produkter, kategorier och lager.

Sales and Order Tables: Registrera information om försäljningsorder, orderdetaljer och fakturor.

Human Resources Tables: Innehåller personaluppgifter, befattningsbenämningar och avdelningsdata.

Purchasing Tables: Dokumentera leverantörsinformation och inköpsorder.

3. Nyckeltabeller och deras funktioner

Person: Lagrar individuell personlig information som namn och kontaktuppgifter.

Product: Innehåller information om produkter inklusive namn, produktnummer och säkerhetslagernivå.

SalesOrderHeader and SalesOrderDetail: Används för att lagra rubrikinformation för försäljningsorder respektive enskilda orderrader.

Employee: Innehåller personalrelaterad data som anställningsdatum och befattning.

4. Relationer mellan tabeller

Databasen använder i stor utsträckning primära och främmande nycklar för att upprätta relationer.

Till exempel länkar SalesOrderDetail till Product genom en främmande nyckel för att indikera which product is in each sales order.

Affärsmodell representation

Person: This schema includes tables related to individuals, such as customers and employees. Key tables include Person, EmailAddresses, BusinessEntityContact, representing pe Detta schema innehåller tabeller relaterade till individer, såsom kunder och anställda. Nyckeltabeller inkluderar Person, EmailAddresses, BusinessEntityContact, som representerar personlig information, kontaktinformation och relationer mellan människor och företag.rsonal details, contact information, and relationships between people and businesses.

Product: Det här schemat innehåller information om de produkter företaget säljer, inklusive tabeller som produkt, produktbeskrivning och produktmodell. Dessa tabeller lagrar information om produkter, deras beskrivningar och deras kategoriseringar.

Sales: Detta schema hanterar försäljningsinformation. Den innehåller tabeller som SalesOrderHeader och SalesOrderDetail, som registrerar försäljningstransaktioner, och Customer, som sparar kundinformation.

Purchasing: Det här schemat är relaterat till inköp av råvaror eller varor, med tabeller som Vendor, PurchaseOrderHeader och PurchaseOrderDetail, som beskriver inköpstransaktioner och information om leverantörer.

HumanResources: Det här schemat innehåller tabeller Employee, Department, and Shift, som ger information om anställda, deras avdelningar och arbetsskift.

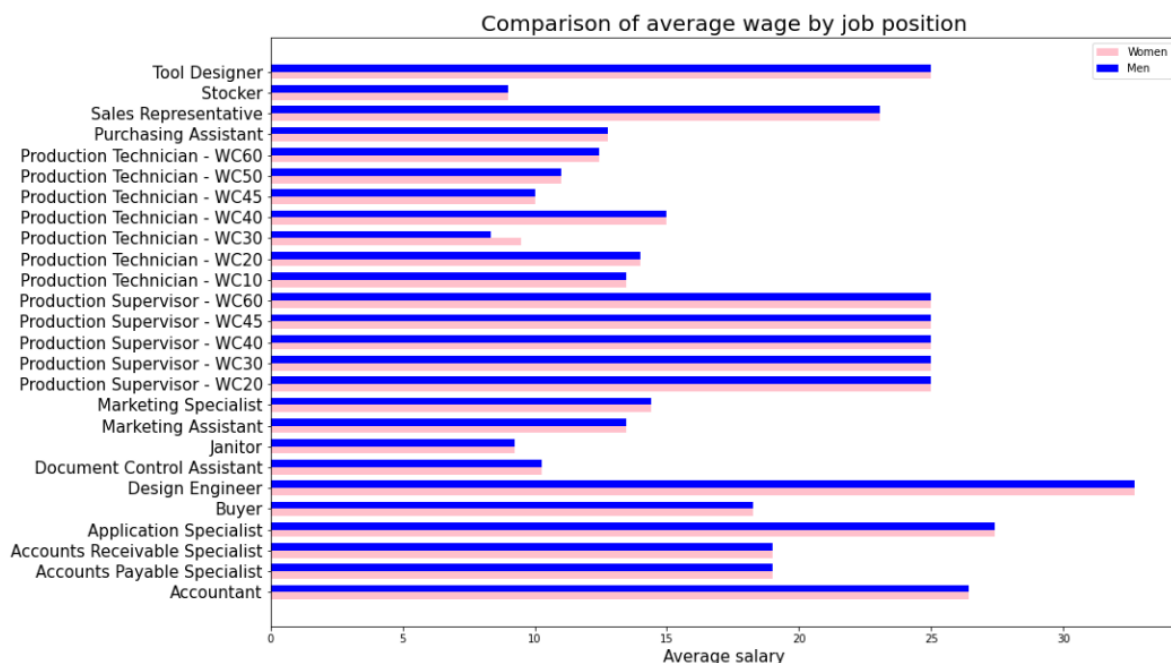
Production: Det här schemat täcker tillverkningsaspekten, inklusive tabeller som WorkOrder, ProductInventory och BillOfMaterials, som används för att hantera produktionsdetaljer, inventering och det material som krävs för produktion.

Statistics Analysis

AdventureWorks ansöker om offentliga bidrag men enligt Ankeborgslagen måste företaget för att de ska godkännas uppfylla vissa standarder relaterade till jämställdhet.

Vi vill analysera om AdventureWorks diskriminerar på grund av kön, konkret gällande lön. Resultatet av denna utredning kommer att vidarebefordras till den behöriga myndigheten för att fatta beslutet.

Efter att ha kontrollerat tabellerna på Human Resources department och efter lite databearbetning med pandas fick jag följande resultat.



Slutsatser och rekommendationer

Efter att ha samlat in och bearbetat uppgifterna rekommenderar jag följande:

- AdventureWorks ådrar sig inte någon uppenbar lönediskriminering baserad på kön.
- Den bäst betalda positionen (där både kvinnor och män arbetar) är designingenjör.
- De tre bäst betalda positionerna (där både kvinnor och män arbetar) visas i följande tabell.

JobTitle	Gender	Rate
Design Engineer	F	32.6923
	M	32.6923
Application Specialist	F	27.4038
	M	27.4038
Accountant	F	26.4423
	M	26.4423

Därför bör AdventureWorks få de bidrag de söker.

Reflektion på eget arbete

1. Utmaningar du haft under arbetet samt hur du hanterat dem:

- Jag fastnade när jag försökte ansluta till SQL Server i python. Jag hittade några kodexempel från min föreläsare, och det visade sig vara lättare än jag trodde.
- Att rita ett stapeldiagram med två datauppsättningar för samma kategori var också utmanande men med mycket googlande föll det på plats.
- Den absolut svåraste delen var att hitta vad jag skulle analysera eftersom jag inte ville upprepa samma analys som i exemplet (sales). Jag blev inspirerad av Pink Programmeringssamtal som har forskats om löneskillnader. Det verkar vara ett trendämne i branschen nuförtiden.

2. Vilket betyg du anser att du skall ha och varför:

- Jag tycker att jag har trots överraskningar klarat av att hantera SQL och hittat lösningar när problem har uppstått samt sökt hjälp om jag efter all förmåga inte klarat av att lösa det, därför tycker jag att jag anses få ett väldigt betyg.

3. Tips du hade "gett till dig själv" i början av kursen nu när du slutfört den:

- Följ lärarens instruktioner för att undvika överraskningar.
- Konsultera när det är möjligt flera informationskällor.