# Teoretiska frågor

1. Beskriv kort hur en relationsdatabas fungerar.

Det är en databas som sammanhänger olika tabeller med överlappande information. Att databasen spara upp information i flera olika tabeller i stället för en stor tabell. Denna struktur effektivisera arbetsflödet och underlätta hur du kan arbeta med information som är sammanhängande med varandra.

1. Vad menas med ”CRUD” flödet? Create, Read, Update and Delete.

Det är hur ett arbetsflöde inom databas operationer ska se ut. Flödet lyder att du ska skapa, läsa, uppdatera och radera data/tabeller/information. Inom SQL kan detta se ut som att du INSERT, SELECT, UPDATE, DELETE records or tables.

1. Beskriv kort vad en ”left join” och ”inner join” är. Varför använder man det?

Left join – väljer du left join så väljer du med hela vänster tabellen och den högra tabellen matcha med den vänstra tabellen.

Inner join – väljer du inner join så kommer båda tabellerna matchas med varandra för att hitta de gemensamma värden från båda.

Man använder det när det behövs att kombinera tabellerna för att hitta förhållanden med en tabell till en annan tabell. Detta för att får en tydlig bild och en helhet om olika tabellers information för att se bland annat om något inte stämmer.

1. Beskriv kort vad indexering i SQL innebär.

Indexering – fungera som referenser till data i tabeller, detta för att kunna effektivt hitta fram data eller tillhörande tabell i en väldigt stor databas (exempelvis). Indexering gör det lättare för användaren att snabbt hitta upp vad de vill leta fram eller arbeta med utan att behöva går igenom hela databasen.

1. Beskriv kort vad en vy i SQL är.

Vy – en virtuell visualisering av tabeller. Om koden eller tabellerna är komplicerade att läsa av, blir det enklare om man ser dem utifrån en vy. I en vy kan du redigera vad du vill se och även gömma känsliga data om det behövs.

1. Beskriv kort vad en lagrad procedur i SQL är.

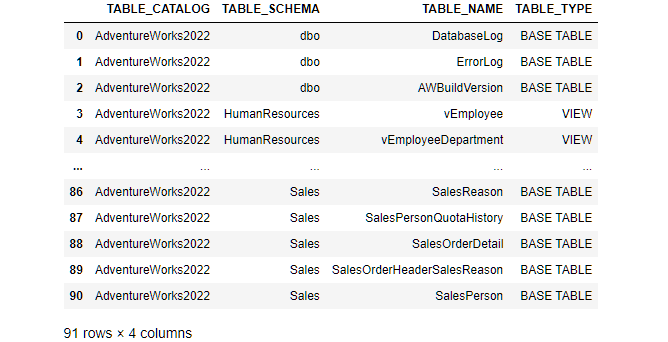
Lagrad procedur – är en förberedd kod som kan användas om och om igen i stället för att skriva om din query. Med andra ord det är en förberedd kod som lagra din query kommando för återanvändning i stället för att skriva om hela query igen.

# 2. Programmeringsuppgift och Rapport

**1. Deskriptiv sammanfattning**

Adventureworks2022 är en databas över ett företag som säljer cyklar. I databasen finns bland annat innehåll som personregister, HR historik, lönehistorik, försäljnings historik, butiker, cykel komponenter och mer annat som går att utforska med hjälp av databas hanteringssystem som SQL.

Här kommer några exempel data från Adventureworks2022 databasen som går att tar fram och utforska.

  
*Figur 1. Alla scheman som finns i Adventureworks2022 databasen.*

En bild som visar text, skärmbild, Teckensnitt, nummer

Automatiskt genererad beskrivning En bild som visar text, skärmbild, nummer, Teckensnitt

Automatiskt genererad beskrivning  
*Figur 2. Antal produkter och beställningar. Figur 3. Kundernas genomsnitt inköphistorik.*

**2. Statistisk analys**

Efter har bekantat mig med företagets databas utfördes det fyra analyser för att får fram hur situationen ser ut just nu i företaget. De fyra analyser ligger i Jupyter Notebook filen i [Github](https://github.com/Mogenavokado/SQL_kunskapskontroll.git), med namnet ”PythonSQL”. [Github](https://github.com/Mogenavokado/SQL_kunskapskontroll.git) länken kommer ni hittar längst ned på detta papper, under rubriken ”Källor”.

Den analysen som kommer presenteras här är analysen över företagets försäljningsshistorik per månad, se figur nedan. Mer hitta ni i Jupyter Notebook filen.

En bild som visar text, skärmbild, Teckensnitt, nummer

Automatiskt genererad beskrivning  
*Figur 4. försäljningsshistorik per månad.*

En bild som visar text, skärmbild, Graf, nummer

Automatiskt genererad beskrivning

*Figur 5. Histogram över totala order per månad. (Röd markering – medelvärdet)   
(skuggat område – konfidensintervall)*

**3. Slutsatser och rekommendationer**

Analysen gav fynd till hur det ser ut för företagets försäljningar genom åren. Genom att veta vårt konfidensintervall kan vi se vilka perioder försäljningen har gått bättre och perioder det har gått sämre. Med den insikten kan vi diskutera fram strategier på t.ex. rabatterade försäljningar, erbjudanden eller hur/vad företaget ska göra för att öka försäljningen under perioderna det har gått sämre på.

**4. Executive summary**

Introduktion

Som ny dataanalytiker anställd på ett företag måste individen vara effektivt med att lära in sig företagets databas så man kan fort bidra till företagets utveckling. Det är därmed en viktig aspekt att vara flexibel och vara bekväm med att hantera databaser och programmeringsspråk.

Deskriptiv sammanfattning

Analysering av företagets databas i Adventureworks2022 skapar helhets bild på företagets nuvarande situation samt lär man sig att vara bekant med nuvarande databas.

Statistisk analys

En djupare analys utförs för att leta fram förbättringar och rekommendationer som potentiellt kan hjälpa företagets framtid.

Slutsats

Diskussion och rekommenderade åtgärder som bidra till företagets utveckling.

**5. Datum för muntlig presentation**

2024-01-02 kl. 10:40

# Reflektion på eget arbete

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

Det svåraste var nog att hålla sig till en idé, eftersom uppgiften gav mycket frihet var det svårt att hålla sig på en riktning.

Det var också utmaningsfullt att sortera ut rätt tabell och data för att får det jag vill få fram i resultatet.

Min lösning var att lära mig att vara bekväm med CTE då det har hjälpt mig väldigt mycket och lägga upp vyer i SQL för att får en bild på hur min plan ser ut samt om det går att utföra.

1. Vilket betyg du anser att du skall ha och varför.

I början tyckte jag SQL var lite mäktigt att hantera då det var så stora databaser som jag fick se från lektionerna men efter att har insett att det finns kommandon som underlätta arbetsflödet så gick det smidigare att välja fram data.

Jag förstår och kan hantera grundliga operationer i SQL så jag är nöjd att ha uppnått den delen. Men det betyder inte att alla data jag tar fram och sedan koda genom python alltid stämmer. Eftersom jag nu har skapat en vana att dubbelkolla så inser jag ibland att något inte stämmer men löser det under dagen dock tar det lite längre tid.

Med all sagt ovan anser jag mig som ett G då grundliga operationer i SQL, CTE-metoden och bland annat python till SQL har blivit en del av min tillväxt.

1. Tips du hade ”gett till dig själv” i början av kursen nu när du slutfört den.

Något jag hade velat öva mer på att konvertera olika format samt lära mig att bli bättre på grafritning, om jag fick hoppa tillbaka i tiden.

**Källor**

<https://github.com/Mogenavokado/SQL_kunskapskontroll.git>