

# 1. Teoretiska frågor

1. Beskriv kort hur en relationsdatabas fungerar.
2. Vad menas med "CRUD" flödet?
3. Beskriv kort vad en "left join" och "inner join" är. Varför använder man det?
4. Beskriv kort vad indexering i SQL innebär.
5. Beskriv kort vad en vy i SQL är.
6. Beskriv kort vad en lagrad procedur i SQL är.

## 1.1 Svar

1. En relationsdatabas är ett system för att samla och organiseras stora datamängder i olika tabeller med rader och kolumner. Relationen mellan dessa tabeller görs genom nycklar kallade primärnycklar och främmande nycklar. Dessa relationer gör det möjligt att få ut informationen ifrån olika tabeller och ställa avancerade sql-frågor kallade queries.
2. CRUD står för de 4 grundläggande operationer som används för att hantera datan i en databas. Dessa är: **Creat** som skapar nya dataposter. **Read** som läser och hämtar befintliga data utifrån vad som specificerats av sökkriterier en användare väljer. **Update** som uppdaterar befintliga dataposter. **Delete** som tar bort dataposter.
3. **Left join** returnerar alla rader från vänstertabell och matchande rader i högra. Finns det inga matcher i högra tabellen blir resultat "NULL" där. Detta används när man behöver alla rader från en tabell (den vänstra) och vill se vilka rader som matchar med den högra. **Inner join** hämtar däremot enbart radar där det finns en match i båda tabellerna, används när man bara vill ha raderna där det finns matcher i båda tabellerna. Dessa olika joins används för att kombinera data från flera olika tabeller.
4. En indexering i SQL skapar en datastruktur med uppslagstabeller som gör det möjligt att söka snabbare efter rader i dessa tabeller. Vilket är nödvändigt när man använder sig av mycket stora datamängder.
5. En vy är en virtuell tabell som skapas utifrån en sparad fråga på en eller flera tabeller i databasen. Genom att använda vyer kan man förenkla komplexa frågor och dölja den underliggande databasstrukturen, vilket gör informationen mer lättillgänglig. Vyer bidrar också till ökad säkerhet genom att begränsa åtkomsten till viss data.

6. En lagrad procedur i SQL är ett fördefinierat kommando som sparas i databasen och kan köras med ett enkelt anrop. Den består av en samling SQL-kommandon som gör det möjligt att automatisera och återanvända komplexa databashanteringsuppgifter. Detta gör det mycket effektivt att använda lagrade procedurer för återkommande uppgifter.

## 2. Rapport

### 2.1 Executive summary

Adventureworks2022 är en exempeldatabas på ett stort företag med 290 anställda med en medellön på 18,19/h. De har väldigt bra personalbehållning där majoriteten av de anställda har arbetat där över 15 år. Detta är dock något de måste vara medvetna om då många går i pension relativt samtidigt. Över 80% av deras intäkter kommer från cyklar som också är produkten med högst vinstmarginal. Men de borde fortsätta sälja alla produkter de gör i nuläget även fast det inte ger mycket intäkter. Medelpriset för cyklar är 1586,74 med en övre gräns på 1785 och en nedre gräns på 1388.48 med ett konfidensintervall på 95%.

### 2.2 Deskriptiv sammanfattning

Adventureworks2022 är en exempeldatabas som har skapats av Microsoft och är ett påhittat företag som driver en verksamhet som cykeltillverkare och även tillbehör till dessa. Det är ett stort företag med 290 anställda. Företaget behandlar sina anställda extremt bra med majoriteten av de anställda har stannat över 15 år på företaget. Genomsnittsåldern på de anställda när det började på företaget var 31 år. Det finns 14 säljare på företaget, stor variation finns det på säljarna gällande hur mycket de säljer, från 1,3Miljoner/år till över 4Miljoner/år. Majoriteten av intäkterna kommer från deras cyklar som är deras dyraste produkt som också har högst vinstmarginal. Efter det kommer kläder, tillbehör och till sist komponenter.

## 2. Analyser

### 2. Personal analys

Analysen visar att företaget har ganska hög medelålder även fast de anställer ganska ungt med en medelålder på 31 år. Majoriteten av personalen har stannat i företaget över 15 år vilket lett till en stor expertis inom sitt yrke. Det har dock lett till att inte så många nya har anställts vilken kan leda till om det fortsätter som det gör idag, att många slutar snabbt samtidigt och företaget måste anställa många nya samtidigt. Med många äldre som går kan det leda till att det finns få att lära upp de nya som börjar vilket kan leda till en kris i företaget. Men inser företaget detta kan de anställa nya folk i tid som lärs upp av de äldre mer erfarna innan de går i pension. Det är också viktigt då det inte anställs många nya i företaget att personalen regelbundet utbildas i den nya teknologin som utvecklas för att kunna prestera så bra som möjligt.

Det finns en väldigt stor variation gällande hur bra säljarna presterar, även fast de ska ha lika mycket erfarenhet då de varit nästan anställda lika många år på företaget. Här kan

företaget analysera vad de bättre säljarna gör för att sedan lära ut de som inte säljer lika bra.

Medellönen för alla i företaget ligger på 18,19/h. Med ingen skillnad gällande herrar och damers löner. Utan företaget har givit personalen lön utifrån deras arbetsposition vilket syns på tabellen av löner. Detta kan vara en anledning till att personalen stannar så länge på företaget, det är jämställt.

## Försäljningsanalys

Adventureworks2022 är en cykeltillverkare och det ser man också på deras försäljningssiffror, över 80% av deras försäljning kommer från deras försäljning av dessa. De säljer dock mycket fler klädesprodukter än antalet cyklar men de produkterna är mycket billigare än cyklarna, vilket leder till att deras försäljningsvärde blir betydligt mindre. Kollar man på vilka produkter som har högst vinstmarginal ligger cyklarna betydligt högre än de andra produkterna. Man skulle därför kunna säga till dem att helt sluta med de andra produkterna, enbart fokusera på deras cyklar om man bara kollar på vinstmarginal och intäkter. Men en kund vill ha service och enkelhet. Att kunna köpa tillbehör, produkter till sin cykel och cykelkläder i samma butik är väldigt smidigt för en kund. Det kan också locka kunder till sina butiker som bara tänkt köpa något till sin cykel, men väl där bestämmer sig att köpa en ny cykel istället. Ett köp som inte hade varit möjligt om de inte sålde andra produkter.

## Statistisk Analys

Här har jag valt att fokusera på priset av cyklar då det är den viktigaste produkten i deras sortiment. I python fick jag fram priset på 97 olika cyklar och räknade ut priset med ett konfidensintervall på 95%. Där jag fick följande svar:

Medelvärde	1586,74
Standardavvikelse	996.27
Urvalsstorlek	97
Nedre gräns	1388.48
Övre gräns	1785

Detta är en mindre prisskillnad än jag förväntade mig, då det finns stort variation på en cykels kvalité. Men om vi kollar på det dyraste produkterna så ser vi att deras dyraste cykel går för 3578, vilket är ganska mycket mer än 1785. Så varians finns det, och här har vi bara ett konfidensintervall på 95%. Detta tyder dock på att majoriteten av deras cyklar ligger i samma prisklass. Skulle rekommendera företaget att ha flera prisklasser så kunden har flera cyklar av olika kvalité att välja mellan. Om företaget lyckas göra dessa cyklar med bra vinstmarginal.

## 1. Utmaningar du haft under arbetet samt hur du hanterade dem.

Man måste tänka lite annorlunda när man skriver SQL-kod i python, inte som koden vi skrivit innan, men delar du upp koden i mindre steg blir det mer tydligt vad man ska göra. Sen tyckte jag det var lite svårt att få ut rätt information med alla olika joins så jag löste först hur jag skulle få ut informationen i MSSMS och sedan i python.

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

Gjort uppgiften på rätt sätt, men med tanke på hur sen jag är med den kan jag inte förvänta mig högre än g.

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

Verkligen förstå grunderna innan jag försökte mig på mer komplexa joins, så jag vet vad jag gör i stället för att prova mig fram. Ha en gameplan från början vilka analyser jag vill göra, inte bara göra massa olika queries för att få information som inte riktigt är nödvändig.