# Teoretiska frågor

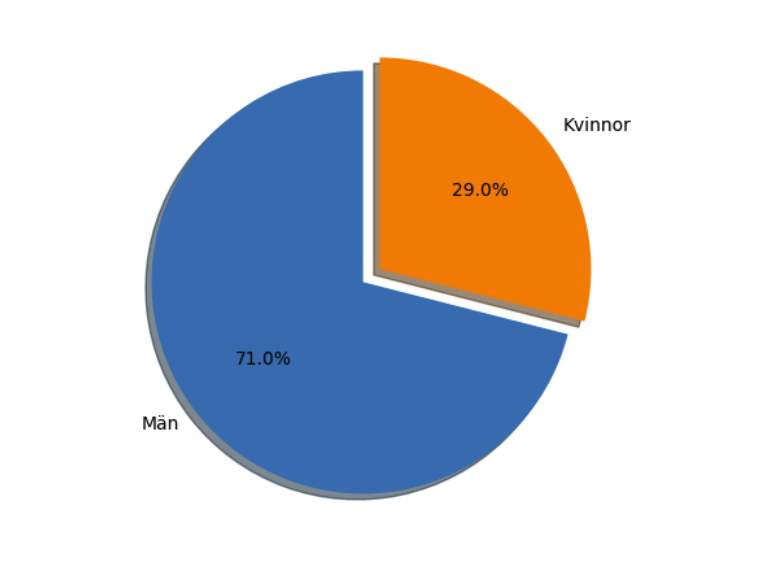
* Beskriv kort hur en relationsdatabas fungerar.  
  En relationsdatabas är ett databassystem som organiserar och lagrar data i tabeller eller relationer vilket gör det lätt att kunna söka och hämta data som man söker efter och är ett effektivt sätt att hantera data mellan olika system.
* Vad menas med ”CRUD” flödet?  
  CRUD står för. Create, Read, Update och Delete. Detta är de fyra grundläggande databasoperationer som utförs på data.
* Beskriv kort vad en ”left join” och ”inner join” är. Varför använder man det?  
  “left join” tar fram all data i den yttre tabellen och jämnför det med den innre tabellen.  
  “inner join” tar fram all data som finns i båda tabellerna.  
  Man använder detta för att jämnföra, kombinera och seperera olika tabeller.
* Beskriv kort vad indexering i SQL innebär.  
  Indexering används för att snabbt ta fram data och att användarna inte ser själva indexeringen
* Beskriv kort vad en vy i SQL är.  
  En vy i SQL är en virtuell tabell som är baserad på resultatet av en SQL sökning  
  Vyer ser ut som en vanlig tabell och innehåller Rader och kolumner.
* Beskriv kort vad en lagrad procedur i SQL är.  
  En lagrad procedur i SQL är en förberedd kod som du kan spara och använda om och om igen i olika förfrågningar och sökningar.

# Programmeringsuppgift och Rapport

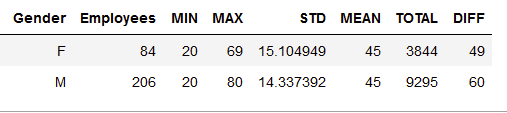
**1. Deskriptiv sammanfattning**

Jag har blivit ombedd att undersöka företagets anställdas sjukfrånvaro för att hitta och sortera något samband mellan dem anställda och deras sjukfrånvaro, få ut deras totala sjuktimmar, Högsta sjukfrånvaron och lägsta sjukfrånvaron, beräkna medelvärdet av dem anställdas sjukfrånvaro och bekräfta om det finns något samband mellan sjukfrånvaron och dess anställda. Samt att förslagsvis kunna hitta ett funktionellt sätt att kunna sänka sjukfrånvaron av de anställda på företaget AdventureWorks2022.

Utvald data.

Jag valde att analysera datan i Humanresources Employees tabellen där jag studerade och analyserade  
vad alla anställda hade mest gemensamt och den gemensama nämnaren var att de anställda är uppdelade  
i två kön.

Vi kan se att det finns en markant skillnad där arbetskraften  
är mestadels män. Vi kan med hjälp av datan fastställa   
det vi behöver för att kunna påbörja våran analys.

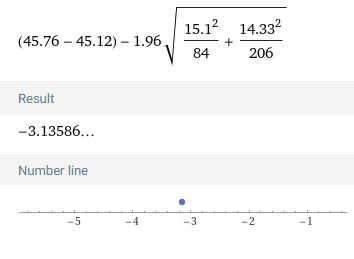
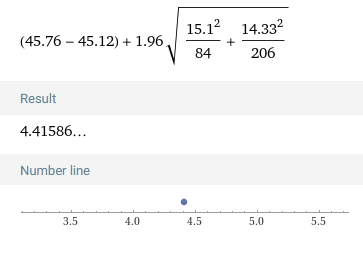


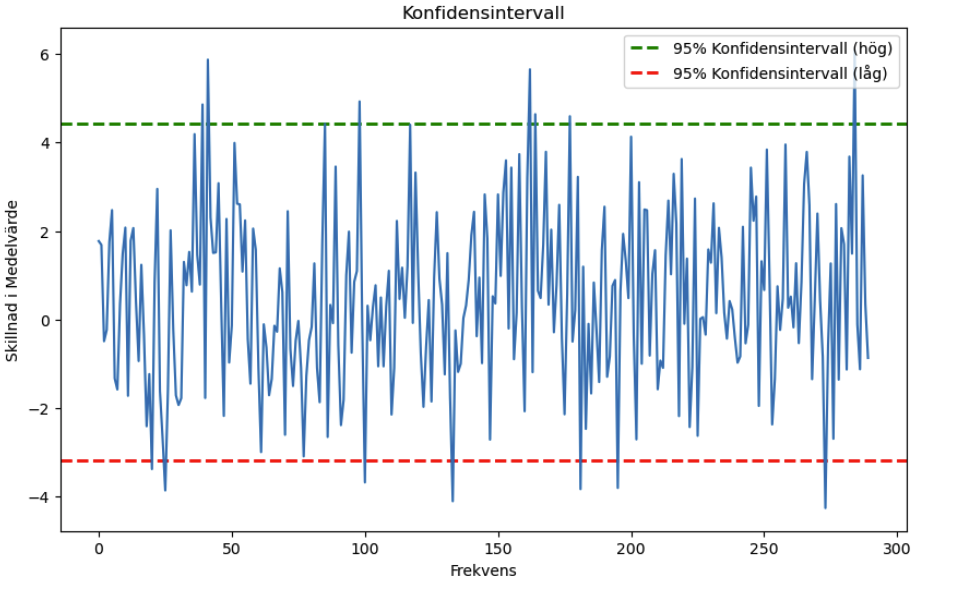
**2. Statistisk analys (konfidensintervall)**

Mellan året 2022 och 2023 har AdventureWorks totalt 290 anställda varav 84 är kvinnor   
och 206 är män, totalt har Adventureworks 13139 sjukfrånvarotimmar bland deras anställda.

Både de kvinnliga och manliga anställda har en minimumfrånvaro av 20 timmar och en maximumfrånvaro av 69 respektive 80 timmar, både de kvinnliga och manliga anställda har en medelfrånvaro på 45 timmar med en decimal skillnad på ungefär 0.6 timmar.

Med den tillgängliga data kan vi framställa konfidensintervallen för dessa två medelvärden





Här ser vi att trots den stora skillnaden mellan arbetskraftsantalet i grupperna så är kan vi se att Nollan finns med och därmed kan vi inte vara säkra på att det finns någon skillnad mellan könen.

3. Slutsatser och rekommendationer

Med denna datan kan vi sammanfatta att det finns en skillnad mellan arbetskraften men att vi inte säkert kan säga att den skillnaden har något samband med sjukfrånvaron.

Med mer data skulle man kunna analysera hurvida andra faktorer har något samband med detta resultat. Har inkomsten någon påverkan, Har shiftbestämelsen någon påverkan eller har åldern någon påvärkan?.

4. Executive summary

Mellan året 2022 och 2023 har AdventureWorks totalt 290 anställda varav 84 är kvinnor   
och 206 är män, totalt har Adventureworks 13139 sjukfrånvarotimmar bland deras anställda.

Där männen hade en minimisjuktid från 20 timmar och en maximumsjuktid på 80 timmar,  
Kvinnorna hade en minimisjuktid på 20 timmar och en Maximumsjuktid på 80 timmar.

Båda grupperna hade en medelsjuktid på 45 timmar med en decimalskillnad på 0.6 timmar  
och vi kunde med hjälp av konfidensintervall se en skillnad mellan grupperna men osäkert säga att det alltid är en skillnad.

Föreslaget att ta in mer data för att djupare kunna undersöka och fastställa ett effektivt sätt att säkna sjukfrånvaron.

5. Datum för muntlig presentation

2024-01-04. Kl 13:40 – 14:00

# Reflektion på eget arbete

* Utmaningar du haft under arbetet samt hur du hanterat dem.   
  Största utmaningen jag haft med detta arbetet var rapporten, Hur skulle jag sätta ord på mitt arbete?. Jag tacklade problemet med att bara skriva och sedan anpassa.
* Vilket betyg du anser att du skall ha och varför.   
  Jag anser att jag förtjänar ett godkänt för jag kunde haft med mer data i uträkningen men använde inte tiden väl.
* Tips du hade ”gett till dig själv” i början av kursen nu när du slutfört den.  
  Stressa inte, Ställ mer frågor, Labba, Labba, Labba.