

## 1. TEORETISK DEL

### 1.1. Relationsdatabas

En RDB är en databas som är uppbyggt av tabeller som relaterar sig till varandra. Varje tabell har unika rader/records och har specifik information om olika data, som tex. kunder, priser, adresser. Relationerna mellan tabellerna etableras genom primary och foreign keys, som knyter ihop unika värden i en tabell, med värden i andra tabeller. Fördelen med detta är att det blir flexibelt att ta fram specifika data och kombinera information från olika tabeller utan att skapa nya tabeller.

### 1.2. "CRUD" flödet?

CRUD står för Create, Read, Update, Delete och är processor som en RDBMS ska kunna utföra. Create: processen att skapa nya records i tabellerna. Read: att kunna läsa data i tabellerna. Update: uppdatera records. DELETE: ta bort records från tabellerna.

### 1.3. LEFT JOIN & INNER JOIN

LEFT JOIN mellan två tabeller (tabell A LEFT JOIN tabell B) betyder att vi tar alla rader från tabell A (den vänster tabellen) och lägger till rader som matcher från tabell B (den höger tabellen).

INNER JOIN mellan två tabeller betyder att man tar de rader som finns i både tabellerna. Det är som att ta A snitt B. Om inga rader matcher kommer resultatet av queryn vara noll rader.

Vi använder joins för att kombinera data från olika tabeller. I stället för att skapa oändlig många tabeller kan vi genom joins skapa temporära tabeller och kombinera data på ett flexibelt sätt.

### 1.4. Indexering

Indexering i tabellerna är till för att datorn kan hitta information snabbare. I en indexkolumn är varje värde unikt – ett exempel är tabellernas primary keys som skapar ett clustered index. Andra kolumner, som inte är primary key kolumner, kan också vara indexerade. Man gör typisk en indexering på kolumner som man ofta använder i joins eller andra query funktioner. För många indexes i en tabell gör det långsammare att lägga till nya rader då alla indexes sen ska uppdateras.

### 1.5. En VY

En vy i SQL är en tabell som kan tas fram med SQL-queryn CREATE VIEW och som blir sparad i Object Explorer under Views. Man kan göra en vy om det är information som användas ofta.

### 1.6. En lagrad procedur

En lagrad procedur är en SQL-query som är sparad, och fungerar som en funktion eller metod som kan anropas. Proceduren kan ha en eller flera variabler som kan definieras när man anropar proceduren. Det är praktiskt om man vill ta fram specifik information från en tabell eller få ett output som är bestämt av några fastsatta variabler.

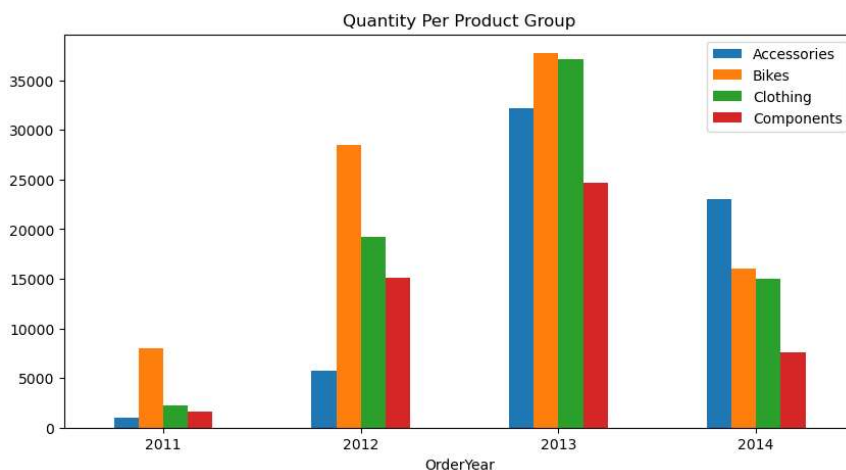
## 2. PRAKTISK DEL

### 2.1 Deskriptiv beskrivning

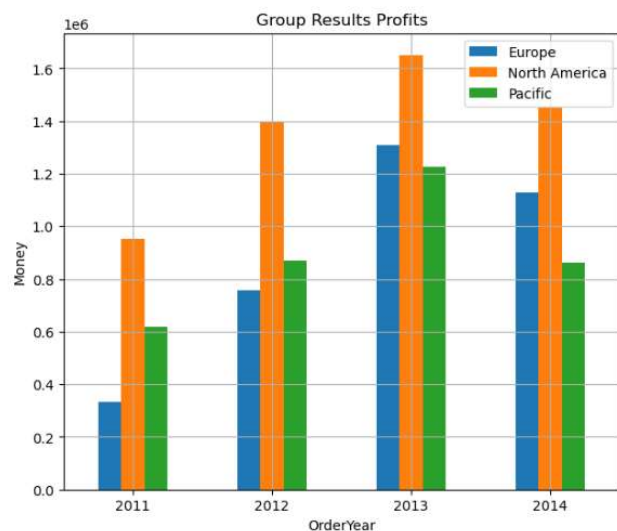
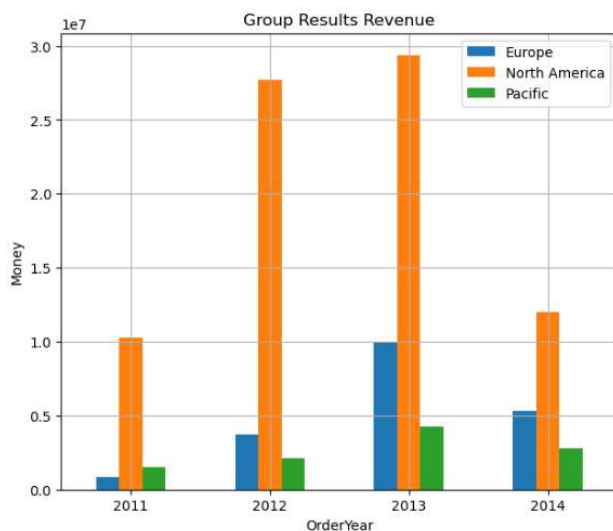
Databasen för företaget AdventureWorks2022 är uppbyggt av 5 centrala schemas (Purchasing, HumanResources, Person, Sales, Production), som tillsammans har 69 tabeller och 23 views, där flest tabeller och views finns i Production och Sales.

I 2014 hade AdventureWorks 290 anställda fördelad över 16 avdelningar, uppdelad på 6 grupper. 62% av arbetsstyrkan var anställda inom produktion. Företaget har växt sedan 2006, med det högste antal nyanställningar i 2009 där 148 personer blev anställd - härav 104 inom Manufacturing.

I 2014 hade de 504 produkter där 266 av dessa var uppdelade på 4 produktkategorier: Bikes, Accessories, Components och Clothing. De resterande 238 hade ingen ProductSubCategoryID och var inte registrerade på några SalesOrders, så jag antar att de är reservdelar till andra produkter. På försäljningen, Quantity per produktkategori från 2011 till 2014, ser vi att försäljningen av cyklar har varit den centrala del av företaget, men att de andre kategorier har ökad över åren.



De säljer till USA, Canada, Frankrike, Tyskland, Australien och Storbritannien. Nedan syns fördelningen av Revenue och SalesProfit per världsdel över 4 år, där vi ser att Nordamerika är deras största marknad.



## 2.2. Statistisk analys

### Metod och syfte med analysen

I min analys vill jag undersöka effekten av två rabattkampanjer. Syftet med detta är att ge insikter som kan användas till framtida kampanjer.

Jag vill fokusera på två kampanjer där Resellers, via SalesPersons, fick rabatt på 3 produkter (hjälm). Effekten av dessa två kampanjer vill jag jämföra med försäljningen av de samma produkter, utan discount, såld via SalesPersons. Jag vill använda konfidensintervall för att få kunskap om normalvärden för försäljningen. De centrale frågorna är:

- Finns det en statistisk signifikant skillnad på hur kampanjerna presterade och den vanliga försäljningen per månad. Både kampanjerna varade en månad.
- Finns det en statistisk signifikant skillnad per SalesOrder, både kampanjerna emellan och i förhållande till den vanliga försäljningen.

### Antagande

I rapporten har jag utgått från att alla kolumnerna med money är räknad i US-dollar. Alla konfidensintervall är räknad med konfidensgraden 95%.

### Kampanjerna

Nedan ses information om de två kampanjer som jag har kallad A och B. Rabatten var på produkterna: 707, 708 och 711.

Campaign	SpecialOfferID	Description	DiscountPct	Category	StartDate	EndDate
A	8	Sport Helmet Discount-2002	0.10	Reseller	30-05-2012	29-06-2012
B	11	Sport Helmet Discount-2003	0.15	Reseller	30-05-2013	29-06-2013

Nedan syns hur kampanjerna presterade. Vi ser att UnitPriceDiscount – det pris som produkterna blev såld för inkl. discount – är långt under det pris som är satt i Production.Product tabellen (ListPrice). Marginalen per produkt är väldigt lite, vilket resulterar i en låg SalesProfit på både kampanjerna. Kostnaden per produkt är tagit från StandardCost per produkt från Production.Product tabellen.

Campaign	Orders	UnitPriceDiscount	ListPrice	StandardCost	Quantity	Revenue	SalesProfit
A	98	15,1399	34,99	13,8782	492	7448,83	620,75
B	84	13,3837	34,99	13,0863	680	9100,9	202,22

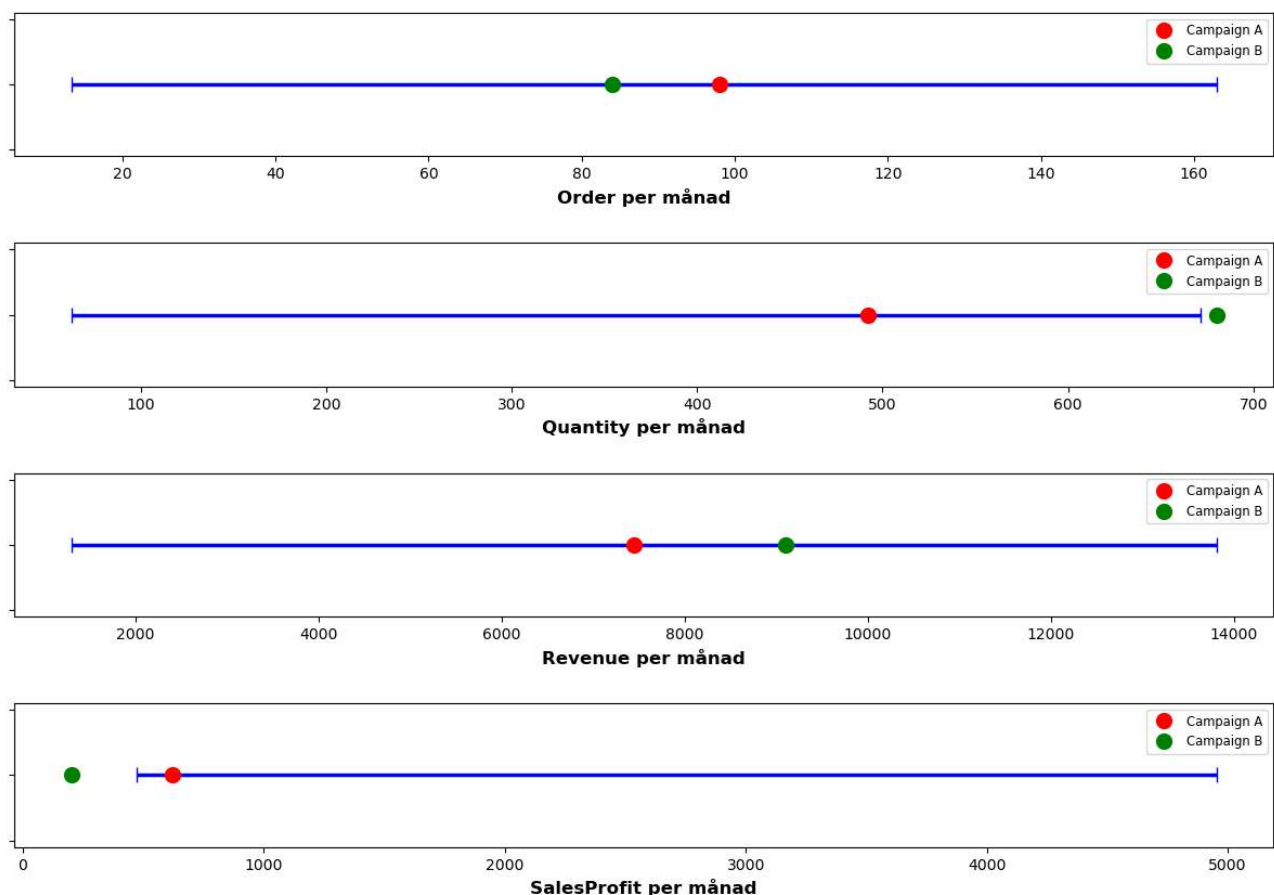
Kollar vi på den vanliga försäljningen var UnitPrice (UnitPrice är priset per produkt på ordena) på dessa produkter 20,1865 fram till den 30.06.2014. Efter det datumet var UnitPrice 20,994. Dvs. UnitPriceDiscount i kampanjerna är inte bara 10% och 15% under, men hela 25% och 34% under det lägsta UnitPrice som produkterna vanligtvis blev såld för. Varför man har vald att ha ett lägre UnitPrice än vanligt och ha discount på detta kan jag inte besvara i denna rapport. Utan konstatera att informationen i databasen om kampanjens rabattsats kan ifrågasättas.

## Försäljningen av hjälmar per månad

Jag har tagit fram försäljningen av de tre hjälmar utan discount, såld via SalesPersons i åren 2011 till 2014 och grupperad per månad. De parametrar som jag har valt att kolla på är Orders, Quantity, Revenue och SalesProfit. SalesProfit är uträknad utifrån LineTotal (ex. moms och frakt) minus standardkostnaden för produkterna på orderna. Enligt Shapiro Wilks testen på de fyra parametrarna kommer de från en population som är normalfördelad, då testerna alla gav p värden över 0.05.

## Konfidensintervall för försäljningen per månad

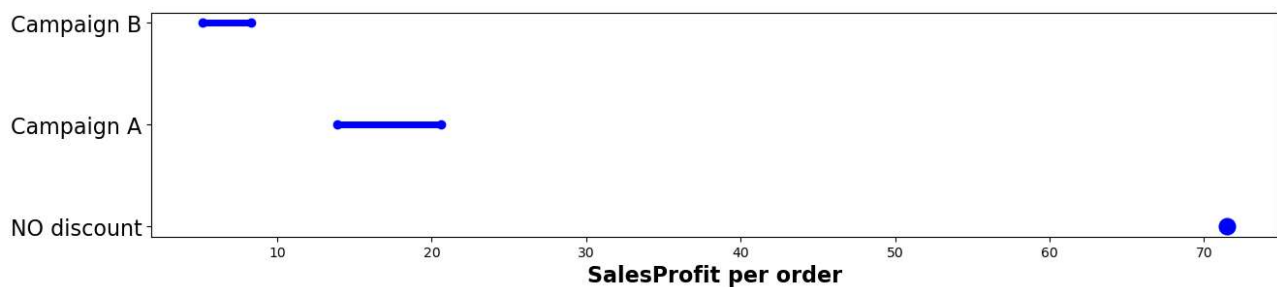
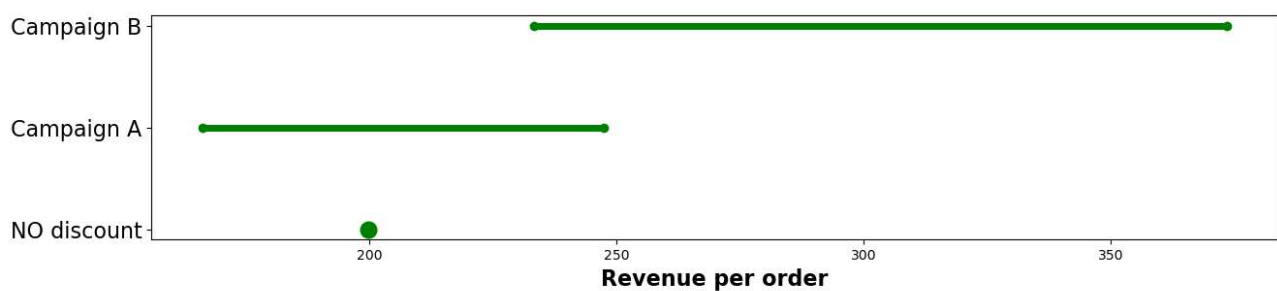
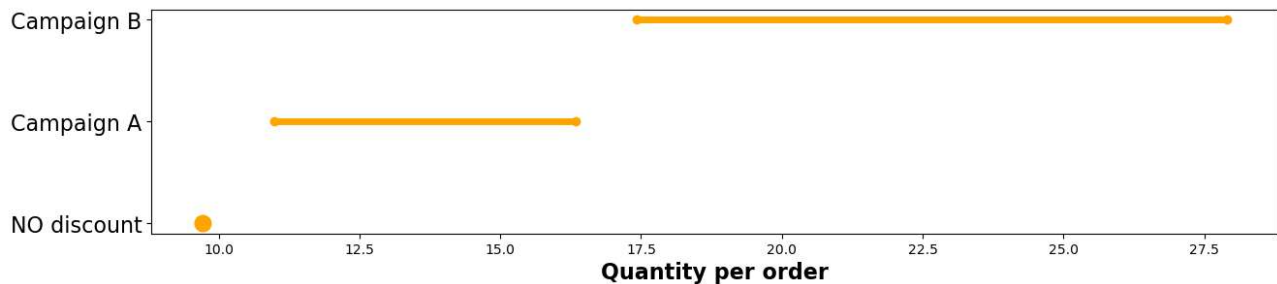
Nedan syns konfidensintervallen för försäljningen per månad på parametrarna Order, Revenue, Quantity och SalesProfit. Jag har också lagt till de enskilda kampanjerna som punkter. Då kampanjerna enbart varade en månad kan jag inte göra ett konfidensintervall per månad för kampanjerna som går att jämföra med den vanliga försäljningen per månad på de valde parametrarna. Jag har i stället lagt in resultaten för kampanjerna som prickar på konfidensintervallen för att få en uppfattning om hur de presterade i förhållande till en vanlig månad.



Kampanj B ligger utanför KI för Quantity och SalesProfit. Kampanjen resulterade alltså i en större kvantitet av produkter blev såld än en vanlig månad, men detta gav inte ökad intjäning, tvärtom. Pga. den låga marginal per produkt var intjäningen under en vanlig månads intjäning. Kampanj A ligger inom områden för det normala och vi kan inte se någon signifikant skillnad på några av de valde parametrarna.

### Konfidensintervall per SalesOrder

Nästa fråga är om kampanjerna har haft en effekt på de enskilda SalesOrders. Jag har tagit fram data för både kampanjerna och för den vanliga försäljningen per SalesOrder. Enligt Shapiro Wilks testen kommer data från kampanjerna från en normalfördelad population, medan det inte är fallet för den vanliga försäljningen per SalesOrder. Jag har därför enbart gjort konfidensintervall på de två kampanjerna och lagt till medelvärden för försäljningen utan discount per SalesOrder som punkter.



Vi kan se att både kampanjerna skiljer sig från den vanliga försäljningens medelvärde förutom på Revenue, där det enbart är kampanj B som ligger över. Vi ser också att de två kampanjernas KI går över varandra på Revenue och det finns då ingen signifikant skillnad på Revenue mellan kampanjerna.

Det är dock en signifikant skillnad mellan kampanjerna när vi kollar på Quantity och på SalesProfit. Kampanj B har både såld signifikant flera produkter och haft en signifikant lägre intjäning per SalesOrder än kampanj A. De extra order har alltså inte bidragit till en ökad intjäning per SalesOrder. Både kampanjerna har högre kvantitet och lägre SalesProfit än en vanlig SalesOrder med samma produkter.

## 2.3 Slutsatser och rekommendationer

De frågor jag ställde i början var följande.

- Finns det en statistisk signifikant skillnad på hur kampanjerna presterade och den vanliga försäljningen per månad.
- Finns det en statistisk signifikant skillnad per SalesOrder, både mellan kampanjerna och i förhållande till den vanliga försäljningen.

I den första fråga kunde vi konstatera att kampanj A låg inom värden för en vanlig månads försäljning på Orders, Quantity, Revenue och SalesProfit. Däremot låg kampanj B över på Quantity och under på SalesProfit. Kampanj B resulterade i flera sålda produkter och lägre intjäning än en vanlig månad.

I analysen av parametrarna Quantity, Revenue och SalesProfit per SalesOrder, så vi likaväl att kampanj B stack ut i förhållande till Quantity och SalesProfit. Inte bara när vi jämförde med medelvärden för den vanliga försäljningen, men också i förhållande till kampanj A. Kampanj A var också över medelvärden på Quantity och under medelvärden för SalesProfit när vi jämförde med en vanlig SalesOrder.

Utifrån detta kan vi konkludera att de två kampanjerna har haft en positiv effekt på antalet produkter som blev såld, men en negativ effekt på intjäningen per SalesOrder. Detta reflekteras också när vi kollar på resultaten per månad. Vi kan också konkludera att rabattsatsen hade en betydelse, då de två kampanjerna hade en signifikant skillnad på Quantity och SalesProfit per SalesOrder. Ju högre rabattsats, ju större Quantity, men ju lägre SalesProfit.

Konklusionerna ovan baseras enbart på analys av två kampanjer och har därför sina begränsningar. Det behövs mera data, om vi ska kunna säga att denna rapports konklusioner kan appliceras på kampanjer generellt, eller om resultaten är specifika för dessa två kampanjer. Om företaget vill använda rabatter i försäljningen, hade jag rekommenderat att köra flera kampanjer med olika rabattsatser och till olika segment inom kunddatabasen. På detta sätt kan vi också lära mera om vilken effekt olika kampanjer kan ha.

Rabatter kan vara ett bra verktyg i försäljningen, men det behövs riktlinjer. Vem ska rabatten gå till, hur länge ska de vara, hur liten blir marginalen på produkterna, vad blir konsekvensen för intjäningen och vad kan vi förvänta oss av kampanjerna. Om det finns en strategi för användning av rabatter i kampanjer, kan vi sätta rättvisande KPIer för kampanjerna, samt få en bättre uppföljning och bättre förståelse för deras effekt.

## 2.4 Executive summary

De två kampanjerna hade en positiv effekt på kvantiteten av produkter per SalesOrder. Vi ser också att rabattsatsen har betydelse, då kampanj B med den högre rabattsatsen presterade signifikant bättre än kampanj A på kvantitet per SalesOrder. Det var också enbart kampanj B som låg utanför konfidensintervallen för Quantity per månad.

På trots att den ökade kvantiteten hade både kampanjerna en negativ effekt på SalesProfit per SalesOrder och också i förhållande till den vanliga försäljningen per månad. Att ha rabatt kan vara ett effektivt verktyg om man vill sälja flera produkter, men man ska i sin tur acceptera att det kan skada intjäningen.

För framtida kampanjer är det därför viktigt att sätta rätt KPIer, så uppföljningen på kampanjerna blir rättvisande och att företaget har realistiska förväntningarna till utfallen.

### **3. REDOGÖRELSE**

#### **3.1. Utmaningar**

Det har varit utmanande att få en överblick över databasen och vilken information som kunde vara intressant att lyfta fram i en rapport och göra konfidensintervall på. Jag gjorde många queries på mycket av data innan jag fick begränsa mig till det som skulle vara med i rapporten.

#### **3.2. Betyg**

Jag har lärt mig väldigt mycket och jag tycker det är kul att göra queries och gräva i data. Jag tycker att jag har gjort det som krävs för att bli godkänd.

#### **3.3. Tips.**

Den första tiden var det svårt att begränsa sig och inte snöa in sig på att få en tabell till att lyckas framför att jobba mera målriktad. Jag hade behövt flera pauser och reflektioner över vad detektivarbetet skulle leda fram till.

#### **3.3. Presentation**

Den muntliga presentationen var onsdag den 3. Januari 2024

-