Kunskapskontroll SQL

# 1. Teoretiska frågor

1.**Relationsdatabas:**

En databas som lagrar information i tabeller med rader och kolumner. Tabellerna kan sedan länkas ihop med hjälp av nycklar vilket gör det enkelt att få fram information från olika tabeller.

2.**CRUD:**

Står för Create, Read, Update och Delete. Detta vill man kan kunna göra i en relationsdatabas.

3.**Left join**:

Tar alla värden från den vänstra tabellen och alla matchande värden från den högra tabellen, om det saknas värden i högra tabellen får man tillbaka såkallade NULL värden.

**Inner join:**

Tar endast värden som har matchande värden i både vänster och höger tabell. Detta gör att rader som saknar matchande värden inte kommer med och man filtrerar på så sätt bort NULL värdena. Detta är den vanligaste typen av JOIN

4. **Indexering:**

Indexering är en slags separat tabell man kan skapa för att skynda på processen av att hitta information i en tabell, detta saktar dock ner andra processer som t. ex uppdatering av en databas.

5.**Vy** är en sparad Select query som man kan tillkalla för att få fram viss information från en databas, detta används oftast när man vill begränsa vad vissa användare ska få ta del av och därför skapar man en vy med endast den tillåtna informationen.

6.**Lagrad Procedur** är en kod som sparas i databasen under ett givet namn. Detta gör att användare kan tillkalla proceduren via namnet och få ut resultatet av proceduren utan att skriva om hela queryn.

# 2. Programmeringsuppgift och Rapport

**1. Deskriptiv sammanfattning**

AdventureWorks är ett företag som ägnar sig åt försäljning av cyklar och tillbehör.

Databasen är uppdelad i 5 olika scheman, varje schema innehåller i sin tur en del olika tabeller. I dessa tabeller kan vi hitta information om olika delar av företaget, t. ex inom schemat HumanResources hittar vi tabeller med information om anställda, deras löner med mera. Vi har även schemat Sales, där hittar vi tabeller med information om försäljning, säljare med mera. Utöver dessa 2 har vi även schemana Persons, Production och Purchases.

AdventureWorks har 296 anställda, det finns flera olika avdelningar men den överlägset största i form av antal anställda är ”Production” där 61% av de anställda arbetar. Utöver Production har vi även t. ex ”Finance” och ”Marketing”. På AdventureWorks är 72% av de anställda män och resterande kvinnor.

En bild som visar text, diagram, cirkel, skärmbild

Automatiskt genererad beskrivning

Databasen innehåller termen ”Persons”, detta är alla som associeras med företaget. Det vill säga kunder, anställda och kontakter hos återförsäljare. Nedan ser vi de 10 städer med flest personer inom AdventureWorks. London är högst på listan följt av Paris. AdventureWorks är alltså ett internationellt företag som finns runt om i världen.

En bild som visar text, skärmbild, nummer, linje

Automatiskt genererad beskrivning

Som tidigare nämnt arbetar AdventureWorks med försäljning av cyklar och tillbehör. Produkterna är som på de flesta företag indelade i olika kategorier, vi ser att flest produkter tillhör kategorin Components följt av bikes. Det finns även clothing och accessories kategorierna som innehåller lite mindre produkter än de först nämnda.

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

AdventureWorks omsätter tiotals miljoner dollar om året och omsättning har ökat stadigt mellan 2011–2013. 2014 ser vi en kraftig minskning från föregående år och detta beror på att vi saknar data för andra halvan av 2014. Detta gör att grafen visar en så stor minskning. Likaså ser vi en mycket stor ökning från 2011 till 2012 och det beror på att vi endast har data från och med 2011–05. Vi ser fortfarande en ökning mellan 2012–2013, alltså de två åren där vi har fullständig data.

En bild som visar text, diagram, linje, Graf

Automatiskt genererad beskrivning

Sammanfattningsvis kan vi konstatera att AdventureWorks är ett internationellt företag med tiotalsmiljoner dollar i omsättning per år. I databasen har vi data för åren 2011–05 till 2014–06 och datan är uppdelad i tydliga scheman vilket gör det enkelt att leta upp relevant information och navigera i databasen. Databasen innehåller mycket information om olika delar av företaget som t. ex försäljning, anställningsdatum, priser för produkter och mycket mer.

**2. Statistisk Analys**

Jag har fått i uppgift att undersöka sambandet mellan försäljning och kampanjer, för att ta reda på om tidigare kampanjer har varit lyckade och om framtida kampanjer kan förbättras.

Graferna nedan visar den totala försäljningen per månad samt försäljningen av vissa utvalda produkter per månad. Tabellen visar alla föregående kampanjer.

Om vi kollar på perioden för den första kampanjen 2012-04-13 till 2012-05-29 ser vi en ökning i försäljning. Vid en närmare titt kan vi se att ökningen började innan kampanjen startades och fortsätter öka efter kampanjen fram till 2012–06 innan den börjar minska. Kampanjen gäller för ”Mountain 100” inom Subcategoryn ”Mountain Bikes”, i nedersta grafen framgår det ingen ökning i försäljning av ”Mountain Bikes”.

Om vi kollar på nästa 2 kampanjer som är under period 2012–06 till 2012–07 och den andra fram till 2012-08 ser vi en minskning i försäljning. Om vi kollar närmare på ”helmets” och ”road bikes” som kampanjerna gäller ser vi att ”road bikes” har minskat i försäljning under denna period medan ”helmets” har hållit sig på samma nivå.

Det är alltså svårt att se något tydligt samband här och intuitivt kan vi säga att kampanjer inte verkar ha någon positiv påverkan på försäljning. För att få tydligare svar får vi undersöka närmare.

En bild som visar text, diagram, linje, Graf

Automatiskt genererad beskrivning

En bild som visar text, skärmbild, Teckensnitt, nummer

Automatiskt genererad beskrivning

Vi börjar med att jämföra medelvärdena för att få en överblick av skillnaderna.

Medelvärdet för försäljningar under en dag är ungefär 97 728 dollar.

Om vi kollar på medelvärdet för en dag under kampanj 7 ”Mountain-100 Clearance Sale” får vi ett medelvärde på ungefär 43 151 dollar, alltså långt under.

Detta ger oss inget svar på frågan direkt då det kan finnas ett flertal faktorer som påverkar vårt resultat. För att få tydligare svar får vi ställa upp en hypotesprövning.

Vi sätter signifikansnivån på 95%.

Vi jämför stickprov 1 som är försäljning under kampanjer och stickprov 2 som är försäljning utanför kampanjer.

H0: μ1 = μ2

H1: μ1> μ2

En bild som visar text, Teckensnitt, vit, kalligrafi

Automatiskt genererad beskrivning

P> 0.05 och därmed kan vi inte förkasta H0. Det är inte statiskt säkerställt att det finns en signifikant skillnad mellan försäljning under en kampanj och försäljning utan en kampanj.

Resultatet ger oss svar på om kampanjer i helhet har någon påverkan på försäljning. Detta utesluter inte möjligheten att en viss kampanj har varit lyckad. Så för att kolla lite närmare kan vi kolla specifikt på kampanj 7 och ställa upp en likadan hypotesprövning för denna.

H0: μ1 = μ2

H1: μ1> μ2

En bild som visar text, Teckensnitt, vit, typografi

Automatiskt genererad beskrivning

Även här är P> 0.05 och vi kan därmed inte förkasta nollhypotesen, detta innebär att det inte är statistiskt säkerställt att det finns någon signifikant skillnad mellan försäljning under kampanj 7 jämfört med övrig försäljning.

**3.Data och Resultat/Slutsatser och Rekommendationer**

Enligt min analys av datan kan jag konstatera att med 95 % säkerhet finns det inget samband mellan försäljning och kampanjer. Jag har kollat på kampanjer som helhet men även på specifika kampanjer och resultatet är likadant. Slutsatsen blir därmed att föregående kampanjer inte har gett något resultat i form av ökad försäljning. Jag rekommenderar en överblick och omkonstruktion av marketingavdelningens arbetssätt. Nästa logiska steg skulle vara att lansera kampanjer baserat på data, t. ex försöka hitta trender i försäljning och lansera kampanjer baserat på detta. Efter det kan vi analysera resultaten och optimera kommande kampanjer för att få ännu bättre resultat. Undersöka olika målgruppers reaktioner på olika kampanjer är också lämpligt, då får vi mer information om vilka målgrupper man ska försöka nå med olika kampanjer. Detta ger oss möjlighet att använda olika marknadsföringssätt för olika kampanjer så att vi kan nå rätt målgrupp.

**4.Executive Summary**

**Översikt:**

Analys av kampanjers påverkan på försäljning och hur det kan förbättras.

**Problem:**

Min analys av kampanjer och försäljning har visat att kampanjer inte har varit effektiva och det går inte att säkerställa att kampanjer ökar försäljning. Detta gör att i nuläget spenderas resurser och tid på kampanjer som inte ger något resultat. Detta innebär att vi förlorar pengar på misslyckade kampanjer dels för själva kostnaden av att lansera en kampanj, dels för att vi går miste om försäljningar från lyckade kampanjer.

**Lösningar:**

Det krävs en överblick och genomgång av marketingavdelningen och dess arbetssätt. Deras nuvarande strategier fungerar inte och det behövs ett nytt sätt att lansera kampanjer. Min rekommendation är att lansera kampanjer baserat på data. Detta kan göras på olika sätt men ett exempel är att lansera kampanjer baserat på trender som kan finnas i försäljningen. Efter det kan vi anpassa kommande kampanjer efter resultaten och förbättra framtida kampanjer.

**Nästa steg:**

1. Överblick av Marketing avdelning.
2. Lansera kampanjer baserat på data.
3. Anpassa och optimera kommande kampanjer.

**Muntlig presentation 03/01/2024**

# 3.Reflektion på eget arbete

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

Under arbetet tyckte jag det svåraste var att komma igång med själva analysen. Jag försökte tänka analytiskt och logiskt för att komma fram till något som jag tyckte var nödvändigt för företaget att ta reda på, samt något som man kan förbättra. När jag bestämde mig för vad jag vill analysera gjorde jag ännu fler queries för att få fram mer information och arbetade därifrån. Hade jag fått göra om det hade jag utfört ännu fler queries vid utforskningsdelen så att man har en ännu större grund att arbeta ifrån.

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

Jag anser att jag ska ha betyget VG då jag har fördjupat mig i analysen, efter att jag har fått fram ett svar har jag tittat närmare på resultatet och inte bara nöjt mig med svaret. Jag har under min analys haft i åtanke att det ofta finns flera faktorer som påverkar resultaten man får. Eftersom analysen inte ska vara för lång har jag inte fördjupat mig i alla faktorer.

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

Att utforskningen av databasen är lika viktig om inte viktigare än resterande delar. Här kan man få en god överblick över datan och det blir lättare att se samband och potentiella fördjupningar man kan göra. Utför många queries så att du kan se så mycket information som möjligt och jobba utifrån det. Jag skulle även sagt att det är viktigt att fortsätta arbeta, att inte fastna för länge på en detalj utan gå vidare och kom tillbaka senare. Tänka analytiskt är också viktigt och det hjälper en oftast att komma vidare om man känner att man har fastnat med något.