

Kunskapskontroll SQL

Camilla Månsson

2024-01-05

Teoretiska Frågor

1. Beskriv kort hur en relationsdatabas fungerar.

En relationsdatabas är en databas som har data organiserad i tabeller som har en koppling/relation mellan sig. Varje tabell består av rader och kolumner.

2. Vad menas med "CRUD" flödet?

"CRUD" flödet står för Create, Read, Update, Delete. I detta flöde skapar man tabeller i create, hämtar data i read, uppdaterar data i update och raderar data i delete.

3. Beskriv kort vad en "left join" och "inner join" är. Varför använder man det?

En join används i SQL för att kombinera data från olika tabeller som har en gemensam kolumn. En left join returnerar alla rader från den vänstra tabellen och matchande rader från den högra tabellen. Om det finns rader i den högra tabellen som inte finns i den vänstra kommer det skapas NULL-värden för den högra tabellen. En inner join däremot returnerar bara de rader som har matchande värden i båda tabellerna efter det angivna villkoret. Inner join används då man endast är intresserad av att få fram de matchande raderna mellan tabellerna. Left join används då man vill behålla alla rader i den vänstra tabellen oavsett om det finns en matchning i den högra eller ej.

4. Beskriv kort vad indexering i SQL innebär.

Indexering i SQL är ett sätt att effektivt hämta data från en tabell. Det går att skapa index för att påskynda sökningar men det är viktigt att inte skapa index på alla kolumner då det gör processen långsammare. Varje gång en uppdatering görs måste indexet byggas om.

5. Beskriv kort vad en vy i SQL är.

En vy är en tabell som har skapats av en viss query med ett visst önskat innehåll. Vyer kan tex användas för att transformera data och göra den mer lättöverskådlig. Det kan också användas i syfte att bara göra viss data synlig för vissa personer.

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

En lagrad procedur är en SQL-kod som kan lagras i databasen och sedan anropas och köras. Det betyder att koden kan återanvändas och på så sätt göra det enklare och snabbare att använda än om koden måste skrivas om varje gång.

Programmeringsuppgift och Rapport

Metod

För att närma mig den angivna uppgiften började jag med en grundlig undersökning av databasen Adventure Works 2022. Jag utforskade de olika tabellerna och dess relation genom att använda mig av olika SQL-frågor. På så sätt skapade jag mig en överblick av dess struktur och innehåll. Därefter valde jag att fördjupa mig i specifika delar av datan. Med hjälp av mer avancerade SQL-frågor plockade jag ut relevant information som sedermera låg till grund för en djupare analys. För att tydliggöra mina insikter genom visualisering tog jag hjälp av Python.

Syfte

Syftet med mitt arbete var att få en bred översyn över databasen Adventure Works 2022, vad den innehåller och vad den handlar om. Detta för att sedan kunna göra en mer riktad analys.

Mål

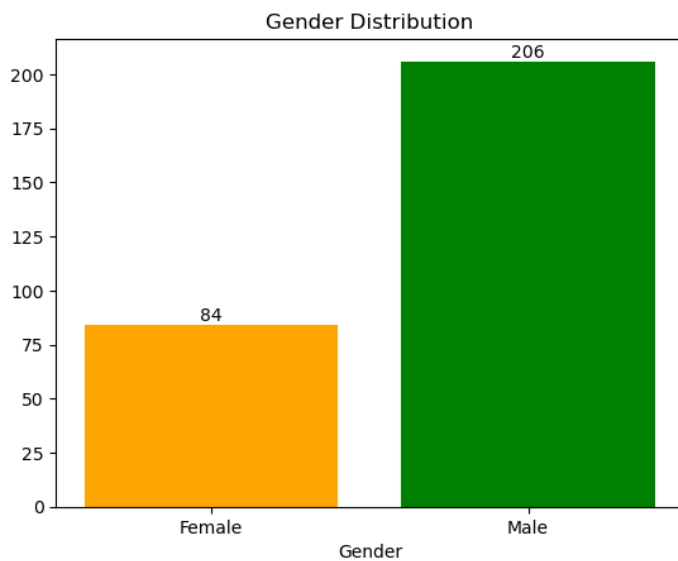
Målet med analysen var att identifiera samband, trender och/eller mönster som kunde vara av intresse. Genom att presentera informationen i form av diagram och grafer ville jag underlätta för andra att dra egna slutsatser av den sammanställda datan.

1. Deskriptiv sammanfattning

Genom en bred översyn av Adventure Works databas går det snabbt att konstatera att företaget tillverkar och säljer cyklar. De säljer även reservdelar och olika tillbehör. Databasen har många scheman som innehåller en stor mängd information om bla deras produkter, produktion, försäljning, anställda och kunder. De säljer sina produkter via internet till kunder i flera länder världen över men har även fysiska butiker.

1.1 Personal

Adventure Works har 290 anställda med en klar majoritet av män (se bild nedan). Den äldsta som jobbar på företaget är 73 år gammal medan den yngsta är 33. Medelåldern är 46 år för kvinnorna och 45 bland männen. Företaget har 16 olika avdelningar, produktion, sälj och ekonomi för att nämna några. I de schema som rör personal går det tex även att hitta deras titel, lön, schema, personuppgifter och sjuk- samt semestertimmar. Den som tjänar mest på företaget är en man vid namn Ken Sánchez som har titeln CEO att jämföra med de som har lägst lön som har titlarna Stocker och Janitor.



1.2 Produkter

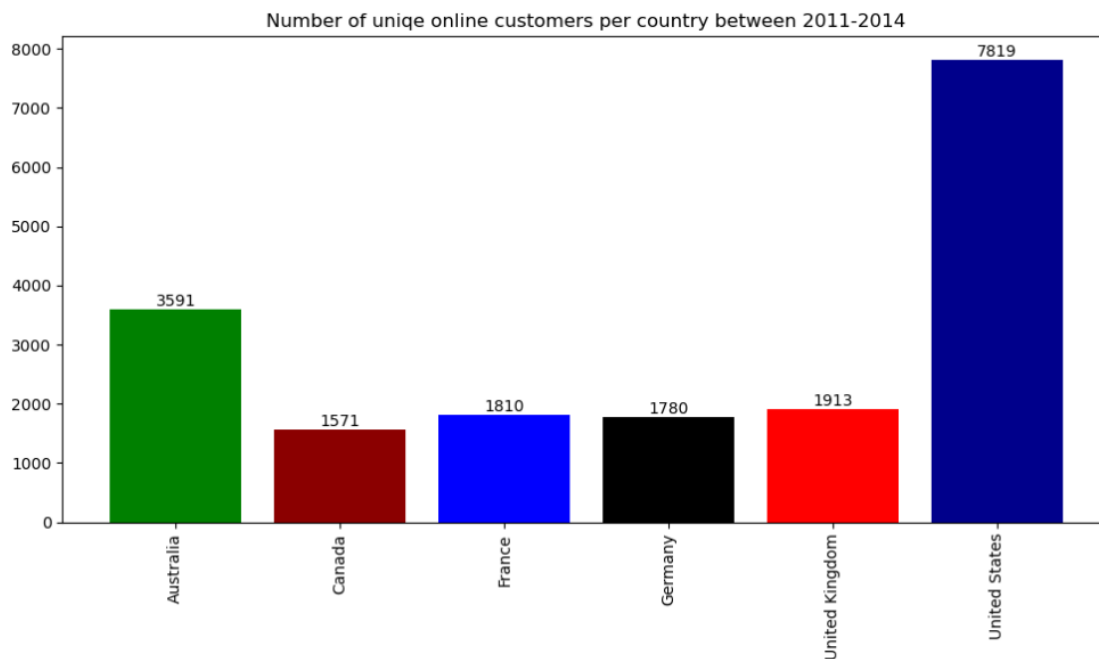
Adventure Works har hela 505 olika produkter i sitt sortiment. Förutom cyklar säljer de även tex hjälmar, kläder och diverse utrustning för cykling. Kunder kan efter en beställning betygsätta produkten och lägga till en kommentar. En av de produkter som fått bäst betyg är en gul cykel som heter "Road-550-W". Dock är det hitintills inte många av produkterna som blivit kommenterade eller betygsatta.

ProductID	Product Name	Product Description	Rating by customers
798	Road-550-W Yellow, 40	Same technology as all of our Road series bike...	5.0 The f
709	Mountain Bike Socks, M	Combination of natural and synthetic fibers st...	5.0
937	HL Mountain Pedal	Stainless steel; designed to shed mud easily.	4.0
937	HL Mountain Pedal	Stainless steel; designed to shed mud easily.	2.0 N

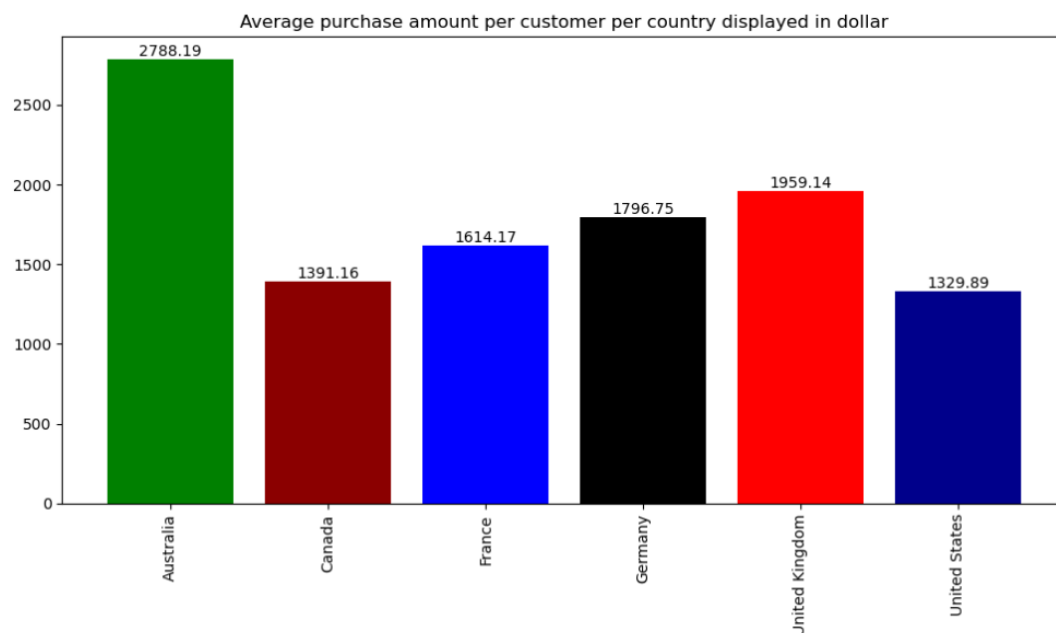
I de schema som rör produkterna går det även att hitta tex priser, kostnad för produkten, kategorisering av produkt och reklamation/tillverkningsfel.

1.3 Kunder

Företagets online-kunder finns i flera länder men flest i USA. Just nu finns det kunder i tre olika europeiska länder, varav störst antal i Storbritannien.



Ordern med det högsta beloppet är beställd av en kund i Frankrike medan den lägsta är gjord av en kund i USA. Men det är kunderna i Australien som har det högsta medelvärdet av köpsumma per kund.

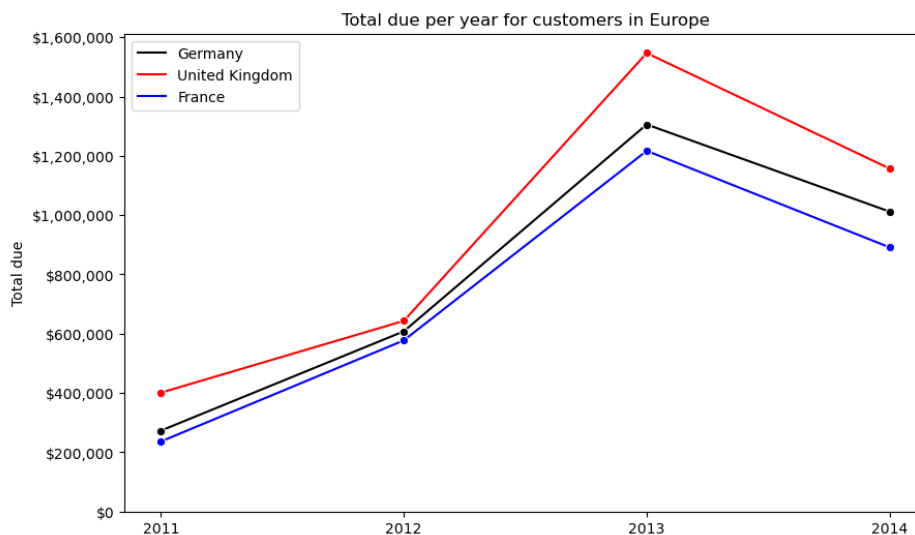


I databasen finns givetvis även kontaktuppgifter och användar-/inloggningsuppgifter för respektive kund.

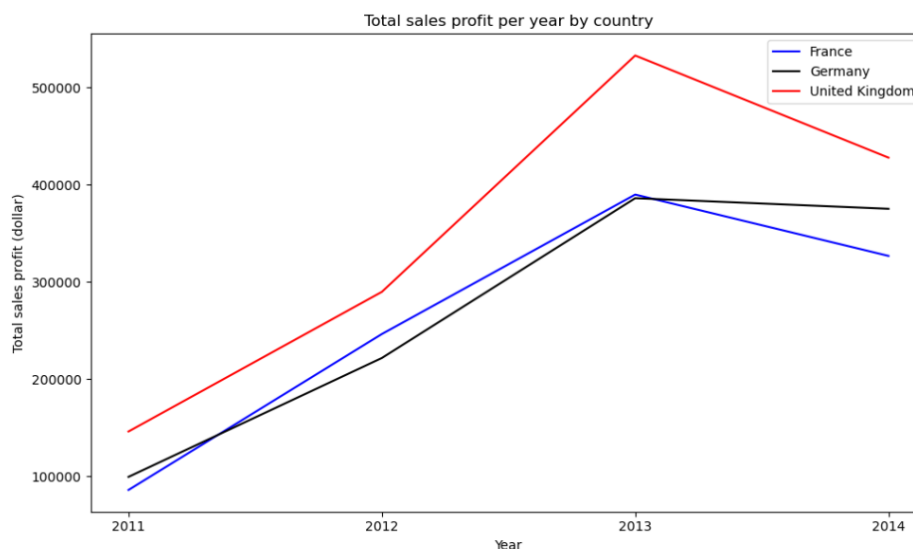
2. Statistisk analys

I min statistiska analys har jag valt att titta närmre på länderna i Europa där Adventure Works i dagsläget har online-kunder i Tyskland, Storbritannien och Frankrike. Vid granskning av antalet ordrar som lagts är det enkelt att konstatera att de ökat stadigt mellan år 2011-2013 i samtliga

länder. De har också minskat i alla länder år 2014 men ligger inte särskilt långt från antalet ordrar året innan. I Tyskland lades 81 beställningar 2011 att jämföra med 1142 stycken år 2013. Året efter minskade antalet med 130 stycken. Utvecklingen för den totala köpsumman per land följer den för antalet ordrar, vilket visserligen inte är så konstigt. Men det hade lika gärna kunnat vara så att kunder gjort fler ordrar men för lägre belopp. För att exemplifiera så är den totala köpsumman i Tyskland 2011 strax över 270 000 dollar medan den stigit till över 1 300 000 dollar 2013 för att sedan gå ner till strax över 1 000 000 dollar 2014. Samma förlopp kan man se både i Storbritannien och Frankrike.



Tittar vi vidare på försäljningsvinsten är utvecklingskurvan i de tre länderna ungefär som den för den totala köpsumman. Men med skillnaden att den mer planas ut mot 2014 och knappt går ner.



Att försäljningen och förtjänsten så framgångsrikt gått uppåt från 2011-2013 beror troligtvis på flertalet faktorer. Men framgångssagan beror säkerligen främst på Adventure Works priser i jämförelse med konkurrenter. Det är vad som anges som säljorsak (SalesReasonID) på överlägset flest produkter. Att det 2014 gått lite neråt hoppas vi inte beror på att marknaden är mättad utan på att det rent generellt varit ett tufft ekonomiskt år. Vi får undersöka detta närmre.

Jag valde att beräkna ett konfidensintervall med konfidensgraden 95 procent gällande medelvärdet på den totala köpsumman år 2011-2014 för kunder i Frankrike, Tyskland och Storbritannien. Sedan tidigare hade jag beräknat medelvärdet på köpbeloppet per kund och fått resultat enligt nedan.

	Country	Number of Customers	Average Purchase Amount per Customer
2	France	1810	1614.1655
3	Germany	1780	1796.7501
4	United Kingdom	1913	1959.1437

När jag sedan beräknar konfidensintervallet får jag resultatet, efter avrundning, att det ligger mellan 1595 dollar och 1985 dollar. Detta skulle kunna ge oss vägledning inför framtida marknadsföringsstrategier om vad som kan förväntas vara det genomsnittliga köpbeloppet per kund med konfidensgraden 95%. Detta kanske kan vara till nytta vid kommande affärsbeslut.

Intressant att veta är också vilka produkter som säljer bäst till antal. I alla tre länderna kan vi se att det är samma produkter som är populärast. En speciell keps, vattenflaska och hjälm är det som säljer mest.

3. Slutsatser och rekommendationer

Av den statistiska analysen att dömma går försäljningen framåt i de Europeiska länder som Adventure Works för närvarande finns tillgänglig för. Med det konstaterat är det lägligt att börja marknadsföra sig i fler Europeiska länder. Förslagsvis i första hand i de länder där cykelsporten är stor. Kanske Italien? Eller varför inte i Nederländerna somt sägs vara det enda landet med fler cyklar än invånare.

Hjälm är en av de produkter som säljer bäst i Europa liksom kepsar och vattenflaskor. Det är inte produkter som ger bäst intäkt men kan vara bra produkter att börja slå sig in på marknaden med. Troligtvis är de enklare att sälja pga bla priset. Med rätt marknadsföring och bra erbjudande kan försäljningen av cyklar, som ger bättre förtjänst, sätta fart. Givetvis är den statistiska analysen inget facit till en prognos för framtiden men ger ändå en fingervisning inför framtida marknadsföringsstrategier och kommande affärsbeslut.

Jag vill också uppmärksamma om att företaget bör sikta mot en något jämnare könsfördelning bland de anställda.

4. Executive summary

Genom en granskning av Adventure Works 2022-databasen framträder en tydlig bild av företagets framsteg och potential. Med över 500 produkter och en stadigt ökande försäljningstrend i Tyskland, Storbritannien och Frankrike, finns det en klar möjlighet till tillväxt och expansiv strategi i Europa.

Datan visar på intressanta mönster i kundbeteende och produktpopularitet, vilket öppnar för möjligheter till strategiska försäljningsinitiativ. Genom att utforska dessa mönster går det att optimera erbjudande och marknadsposition ytterligare.

Genom att fortsätta analysen, fokusera på kundengagemang samt försäljningsstrategier kan företaget fortsätta att driva sin framgångsrika expansion och samtidigt förstärka närvaron i Europa.

5. Tidpunkt för min muntliga presentation var 2024-01-03 klockan 13:40-14:00.

Reflektion på eget arbete

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

Jag har alltid tyckt att det är väldigt svårt med uppgifter där man kan välja helt fritt. Så min allra största utmaning i denna uppgift har helt klart varit vad jag ska välja och vad ska jag fokusera på. En annan del av uppgiften som var väldigt utmanande var att beräkna konfidensintervallet. Själva uträkningen i sig var kanske inga större konstigheter. Mitt problem låg i att förstå varför jag skulle göra det, på vad skulle jag räkna och hur jag skulle tolka det? Jag tyckte inte riktigt den delen passade in i uppgiften.

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

Jag tycker absolut att jag förtjänar godkänt i kursen. Jag har lagt ner mycket tid och möda och känner att jag har lärt mig mycket nytt. Men framför allt för att jag känner att jag obehindrat kan få fram information i en databas genom SQL-kod. Jag anser att jag har en god förmåga att tolka och analysera datan på ett tillfredställande sätt

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

Eftersom att jag hade problem med att komma igång med själva uppgiften ger jag mig själv tipset att köra igång direkt och testa, testa, testa - det kommer leda till något. Försök att begränsa dig tidigt för en djupare analys. Påbörja skrivandet av rapporten i ett tidigt skede parallellt med kodandet och avsett lite mer tid till den.