Teoretiska frågor:

1. Beskriv kort hur en relationsdatabas fungerar.

En relationsdatabas består av länkade tabeller med Primarykey relationer(Unika värden som är desamma I båda tabellerna) relationen mellan tabellerna är en-många –en-en vanligtvis. Tabellerna består av Kolumner och Rader även kallade Fält och Tupler. En viss rad går ofta också under benämningen “record”. Relationsdatabaser är idag den vanligaste databasarkitekturen inom de flesta organisationer.

2. Vad menas med ”CRUD” flödet?

Create, read,use, delete är datats “livscykel”. Det är en terminologi som delar upp behoven olika användare kan ha av en databas. Begreppet kommer från James Martin’s bok “Managing the database environment” som ko mut 1983.

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

En left Join mellan tabeller innebär att alla värden I den vänstra tabellen ingår och de matchande I den högra ingår också. En inner join innebär att bara matchande värden I BÅDA tabeller ingår. I ett Venn Diagram motsvaras en inner join av Snittet. Ett Venn Diagram av en left join innehåller hela den vänstra mängden + snittet. Joins är vanliga operationer för att binda ihop data från flera tabeller.

4. Beskriv kort vad indexering i SQL innebär.

Indexering innebär att varje position I tabellen har ett index och är därmed ett snabbt sätt att läsa data med hjälp av denna position oavsett vad det är för data på positionen. Det är ofta till hjälp för att snabba upp läsning av stora datamängder.

5. Beskriv kort vad en vy i SQL är.

En vy är en bild av datat som man kan spara. Det förändrar inte databasen. Mark hade en bra metafor till detta att det är som att titta ut I ett fönster. Där det du ser är din vy men databasen är allt därute. En vy är data sett ifrån ett visst hall kan man säga.

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

En lagrad procedur är en subrutin som är lagrad I själva databasen. Exempel på när det är användbart är t.ex vid dataverifiering och åtkomstkontroll. Det kan även spara nätkapacitet när denna kod är sparad I själva databasen och inte behöver hämtas.

1. Deskriptiv sammanfattning

Adventure Works Cycles är ett multinationellt företag som tillverkar och säljer cyklar till en global marknad. Huvudkontor I Bothell Washington med 290 anställda och regionala säljare globalt. Försäljningen uppgår till ca 110 MUSD över åren 2011-2014 varav 95MUSD (se nedan) är cyklar. Övrig försäljning består av reservdelar, tillbehör och kläder. Cyklar och cykelkomponenter tillverkas till viss del I Mexico för att monteras I Bothell. Kläder och övriga tillbehör såsom vattenflaskor köps in. Här ses försäljningskategorierna nedan. De egentillverkade försäljningskategorierna star för nästan hela försäljningen.

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **USD Pris** | **Kvantitet** | **USD Försäljning** |
| Bikes | 1,255 | 90268 | 94,651,173 |
| Components | 251 | 49044 | 11,802,593 |
| Clothing | 32 | 73670 | 2,120,543 |
| Accessories | 20 | 61932 | 1,272,073 |

Här ser vi tydligt att de egentillverkade produktkategorierna står för >96% av försäljningen.

A blue circle with orange and green text

Description automatically generated

För att skapa en nulägesrapport (2014) så är det centralt att titta på senaste historik för att dra slutsatser om vilka produkter som går bra och vilka som går mindre bra. Detta är viktig information bade för försäljningsavdelningen och tillverkande enheter I Mexico och Bothell.

För att neutralisera effekten av Säsong och kalendereffekter generellt så väljer jag att jämföra Rullande 12 Månader Juli2013-Juni2014 mot Juli2012-Juni2103. Detta ger en top-down bild av förändringen I försäljning. Försäljningsdiffen delas upp I Förändring Pris och förändring Kvantitet. Denna tabells styrka är att den visar vilken kategori som är mest relevant. Tabellen visar aggregat I form av t.ex snittkvantiteter som kan vara viktig input vid tillverkning. Då det är grova kategorier så måste en djupare analys göras för att få kännedom om specifika produkter. En tydlig slutsats är att någonstans inom Cykelkategorin finns den största förbättringspotentialen. Även gällande Accessoarer kan det vara läge att ta höjd för större inköp. Det kan även finnas gömda “vinnare”/”förlorare” med en något mer granular kategorisering av kategorierna det blir nästa steg.

Diff Juli2013-Juni2014 mot Juli2012-Juni2103 (R12 YoY)

Pris(USD) Kvantitet Försäljning(USD)

|  |  |  |  |
| --- | --- | --- | --- |
| **Bikes** | 102 | 5112 | 9,362,528 |
| **Components** | -28 | -323 | -431,941 |
| **Clothing** | 3 | 9019 | 292,335 |
| **Accessories** | 0 | 38326 | 811,901 |

Efter att det är tydligt att Cyklar är den stora kategorin även trendmässigt I försäljningsförbättring (Inte bara I absoluta tal) så går jag vidare med förändringsfaktorer för en något mer granulär produktindelning mellan tidigare perioder Juli2013-Juni2014 mot Juli2012-Juni2103.

Högsta Förändringsfaktorer Produkter

Pris(USD) Kvantitet Försäljning(USD)

|  |
| --- |
| Produkter |
| **Tires and Tubes** | 1.0 | 37.7 | 51.0 |
| **Bike Stands** | 1.0 | 40.5 | 40.5 |
| **Fenders** | 1.0 | 39.8 | 39.8 |
| **Bottles and Cages** | 1.1 | 11.7 | 17.5 |
| **Touring Bikes** | 1.4 | 5.5 | 7.5 |
| **Cleaners** | 1.3 | 4.7 | 5.8 |
| **Hydration Packs** | 1.3 | 4.5 | 5.3 |
|  |  |  |  |

Touringcyklar är den cykelkategori som uppvisar den största förbättringen med en nästan sjudubbling av försäljningen trots en genomsnittlig prisökning på 40% (Faktor 1.4). Här kan finnas potential för en ännu större försäljningsvolym med en tillfällig sänkning av priset. Accessoarer såsom cykelställ och däck/slangar uppvisar mycket stora volymökningar här finns möjligheter att eventuellt höja priset. Touringcyklarna är den viktigaste trenden att följa och maximera då det är en stor förändring I en viktig kategori.Nedan ses Förändringsfaktorerna för Touringcyklar.

Förändringsfaktorer Touringcyklar Juli2013-Juni2014 mot Juli2012-Juni2103

Touringcyklar Pris(USD) Kvantitet Försäljning(USD)

|  |  |  |  |
| --- | --- | --- | --- |
| **Touring-3000 Blue, 62** | 1.7 | 14.6 | 25.5 |
| **Touring-1000 Yellow, 54** | 1.8 | 9.4 | 19.7 |
| **Touring-2000 Blue, 50** | 1.1 | 11.9 | 12.9 |
| **Touring-1000 Yellow, 50** | 1.6 | 6.6 | 12.6 |
| **Touring-1000 Yellow, 46** | 1.6 | 5.6 | 10.4 |
| **Touring-3000 Yellow, 54** | 1.5 | 6.1 | 9.8 |
| **Touring-3000 Blue, 58** | 1.3 | 6.0 | 9.1 |
| **Touring-3000 Yellow, 50** | 1.4 | 5.7 | 8.8 |

Ett flertal Gula Touringmodeller uppvisar stora kvantitetsförbättringar trots prisökningar på 40-80%. (Även en blå modell men det är tydligare för flera gula). Denna goda trend för Gula cyklar är något som förmodligen går att kapitalisera på I nuläget. Ett sätt är att I en bred marknadsföringskampanj (med fokus på gula touringcyklar) erbjuda ett något lägre pris under en period. Ett alternativ är att lämna med en populär accessoar gratis för att kunna höja försäljningskvantiteten under denna period. Tillverkningsorganisationen behöver ta höjd för en högre volym under perioden.

2. Statistisk analys

En total försäljningsserie över åren 2011 till 2014 skapas där dagstotalerna (ca 1100st) är data att beräkna läges och spridningsmått på.

Beräkning av normalfördelat konfidensintervall för dagsförsäljning i USD är för de 1100 dagstotalerna med 95% konfidensgrad:

83097------139785

Detta intervall bygger på hela samplet (1100st) Hur högt mål måste vi sätta för att kunna säga med hög sannolikhet att vi förflyttat µ vid en framtida två månaders kampanj ? Endast två månader I framtida sampel gör att intervallet blir bredare. En sampling med n=60 på datat ger det normalfördelade konfidensintervallet:

17844-----279908

Här ser vi att en kampanjperiod på två månader kräver ett högt mål för att kunna säga att vi förbättrat oss. Konfidensintervallet blir mycket brett för n=60. 300000 testas som hypotes på dagens data och det visar sig att µ=300000 kan förkastas idag enligt resonemang nedan:

På dagens data med samplestorlek 60 (två månaders kampanj) o t-distribution (lite bredare än normalfördelning vid 59 degrees of freedom) så är sannolikheten bara 4% att dagens distribution täcker den fixa men okända µ. Alltså förkastas H0: µ=300000 idag.

Om 300000 uppstår som dagsmedel I vår 2 månaders kampanj med Gula Touringcyklar så kan vi alltså anta att vi förbättrat oss mot dagens distribution. 300000 som framtida dagsförsäljningsmedel visar att det inte är sannolikt att det är slumpförändringar som ger USD 300000/dag. Det är fortfarande viktigt att komma ihåg att på dagens data är det 4% sannolikhet att USD 300000/dag uppstår. Så vi kan med hög sannolikhet säga(96%) att vi förbättrats om det sker.

3. Slutsatser och rekommendationer

Rekommendationen är fokus på Touringcyklar då de uppvisar hög förbättring R12 YoY (Juli2013-Juni2014 mot Juli2012-Juni2103). Flera Touringmodeller (särskilt Gula) har gått att höja priset på med 40% och uppvisar ändå mycket stora volymökningar(på vissa modeller över 10dubbling) Dessa cyklar bör användas till en “super sale” med något lägre priser under en tvåmånaders period. Kampanjen bör kunna ge mycket högre volymer pga cyklarnas höga populariet för tillfället. Ett alternativ

4. Executive summary

Försäljningen för Adventure Works är inne I god trend vissa specifika produkter sticker ut mycket tydligt. Kampanj för att kapitalisera på detta föreslås.

* Touringcyklar är inne i en mycket stark trend där försäljningsvolym femdubblats under senaste rullande helår mot föregående rullande helår
* Detta har skett trots prishöjningar på I genomsnitt 40% för våra touringcyklar.
* Detta ger ett bra tillfälle att genomföra en kampanj med attraktiva priser för Touringcyklar
* Rekommendationen är att genomföra en två månaders marknadsföringskampanj med fokus på gula Touringcyklar.

5. Muntlig presentation

Muntlig presentation gjordes den 4/1 kl 9:40.

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

En utmaning var att få en överblick på databasen. Jag studerade varje Schema individuellt och fick med tiden en bra bild av var kunde göra en logiskt sammanhängande undersökning.

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

Jag tycker att jag är värd VG för att jag har ett mycket tydligt “pärlband” av delar I mitt resonemang. Alla delar hänger ihop. I ett verkligt scenario så är det så klart inte en heltäckande beskrivning av AdventureWorks,men det har inte heller varit mitt mål och behöver inte vara ett mål som jag förstått uppgiften. Jag har heller inte sett ett behov av mycket visualiseringar då jag tycker mitt resonemang framgår ändå. Kort smalt och koncist. Att visa att jag förstått vad jag tagit fram är viktigast och det tycker jag att jag gjort.

I början gjorde jag en mängd “spretiga” queries men sedan landade jag I att försöka göra ett fåtal skalbara Queries som har med det jag vill undersöka. Det är ett medvetet programmeringsval och jag tror det är att föredra I all programmering. (Att helst jobba I en skalbar “ram” för olika funktioner).

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

Studera databasen I sig själv några dagar I början hade varit väl investerad tid. Jag gjorde det parallellt med queries I början.

Att direkt tänka I att skapa ett fåtal skalbara Queries för de frågeställningar jag har var något som växte fram. Idag hade jag gjort det direkt och inte testat olika “spretiga queries” I början.