# Teoretiska frågor

1. Beskriv kort hur en relationsdatabas fungerar.

En relationsdatabas fungerar genom att använda nycklar. Till exempel innehåller en tabell en främmande nyckel som refererar till primärnyckeln i en annan tabell. Då kan man koppla de två tabellarna med join functionen ” tabell1.PK\_column = tabell2.FK\_column”

1. Vadmenas med ”CRUD” flödet?

”CRUD” är kortname av SQL functioner.

”C” betyder ”create” med functionen “CREATE DATABASE database-name” eller ”CREATE TABLE table-name ”

“R” betyder “Read” med functionen

”SELECT \*

FROM table-name”

“U” bytyder “Update” med functionen

“UPDATE table-name

SET column1 = value1, column2 = value2

WHERE \*\*\*\*\*”

“D” betyder “Delete” med functionen “DROP DATABASE databasename” eller ”DROP TABLE tablename”

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

”left join” menar att alla rader från den vänstra tabellen kommer att vara krav. Men rader som matchar med den vänstra tabellen kan bara kvar från den högra tabellen. Om man vill kolla alla rader från den vänstra tabellen oavsett om de har matchningar från den högra tabellen kan man använda ”left join”.

”inner join” menar att bara de rader som har matchat i båda vänstra och högra tabellerna kunna vara kvar i det filter resultatet. Om man är bara intresserad av data som är gemensam i båda västra och högra tabeller kan man använda ”inner join”.

1. Beskriv kort vad indexering i SQL innebär.

Ett index är ett annat sätt att visa data med en specifik dataordning som skiljer sig från den ordning i vilken data ligger fysiskt på databas. Funktionen av indexering är att man kan sortera om data och skapa en ny strukturerad information. Då får man snabbt hitta värden från den rätta raden.

1. Beskriv kort vad en vy i SQL är.

En vy i SQL är en virtuell tabell som inte har något fysiskt lagrat data utan bara bygger på resultatet av en SELECT frågor. En vy som skapas genom att definiera en fråga kan användas på samma sätt som en vanlig tabell. Men de vyn skapas inte fysiskt. Man kan skapa en vy från en eller flera tabellerna från den databasen. Vyn visar alltid den uppdaterande data. När man använder en vy behöver man inte veta den grundläggande förhållande mellan tabellerna. Då ökar det säkerheten på grund av att man bara får använda data som visas av vyn.

Exempel av en vy

CREATE VIEW namn

AS

SELECT \*\*\*\*\*

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

En lagrad procedur i SQL är en grupp av SQL kod som man kan spara och använda direkt. Så den koden kan återanvändas om och om igen. Om en query användas ofta kan man spara den som en lagrad procedur. Då kan man köra den koden direkt. En lagrade procedur inkluderar processdeklaration och exekveringsprocessdel vilken kan innehålla SQL frågor, logik, variabler, join och andra programmeringskonstruktioner. Den lagrad proceduren är lättare att underhålla och uppdatera. Man använder lagrade procedurer för att organisera och sortera komplexa databasoperationer.

Exempel av en lagrad procedur

CREATE PROCEDURE namn-av-procedur

Variable type

AS

SELECT \*\*\*\*\*

================= Svar till teorifrågor ska ta max 1 sida (men detta är flexibel=================

=================Teorifrågorna ingår inte i den 2-5 sidor och räknas separat==================

==========================Börja rapporten på en ny sida==========================

# 2. Programmeringsuppgift och Rapport

1. Deskriptiv sammanfattning

I den här uppgiften vill jag analysera och diskutera den totala försäljningen, försäljningskvantiteten, produkt med olika färger, region, kund, säljare, månade etc, så att ledare kan använda dessa data att göra strategi, beslut och skapa värde. Jag delade rapport i tre delarna. Den första delen handlar om hur kraft av färg är i marknadsföring. Jag kommer att undersöka vilken specifika färg som kund tycker om mest och distribution av färger på olika produkter. Genom att analysera den data kunde vi analysera förutse konsumenternas köpbeteende, vilket kan instruera ledare att designa flera nya produkter i olika specifika färger. Å andra sidan diskuterar jag förhållande mellan mängd av kund och fördelning av total försäljning på olika region baserade data för att identifiera potentiella värdefulla kunder och fatta konsumkraft i olika region och bidra till att ordna säljare på olika region. Till sist, sammanfattar jag slutförandet av försäljning uppgifter i augusti, vilket ger ett starkt datastöd för att göra försäljning schema och mål för näst år. Därför delade jag upp innehåll i tre delar för att göra en tydlig analys av olika data.

1. Statistisk analys (konfidensintervall)

Del 1. Kraft av färg i marknadsföring

I marknadsföringskriget mellan Coca-Cola och Pepsi-Cola visar det hur avgörande av färg, alltså kallar man kriget mellan röd och blå. Så i den här delen ska vi, på den ena sidan, analysera och förstå vilken specifika färg som kund tycker om mest. På den andra sidan ska vi analysera behöver av någon specifik färg hos konsument så att designer kunna kåppla efterfrågan, färg och produkt ihop, vilket kan förhöja tävling av produkten och slutligen ökar produktförsäljningen. Figur 1 handlar om variationen i totalt antal beställningar och totalt priset för olika färger. Man kan se att svart är det mest populära och den totala försäljningen är också störst. När man kollar på Figur 3 och Tabell 1 vet man att svart färg distribueras huvudsakligen i hjälm, cykel, shorts och handske. Jag tycker att man kan expandera svarta produkter till fler områden, till exempel kläder och cykeltillbehör. Överraskande är nog att gul och röd är den andra och tredje mest populära färgen. Det finns två typ cyklar med röd färg. Och gul handske är ganska populär eftersom gult är en mycket iögonfallande färg som ger människor en öppen och ljus psykologisk känsla. Total antal beställningar av silver och blå följer sedan med men det totalt priset är tvärtom på grund av att genomsnittligt enhetspris av silvers produkter är betydligt högre än de av blå produkter, som visas i figur 2(a). Så trots att totalt antal beställningar för blå och silverprodukter är liknande, är den totala försäljningen är olika. Från figur 3 kan vi se att blå hjälmar är mer populära än röda och svarta hjälmar. På samma sätt som svarta och röda cyklar kan vi utveckla blå cyklar.

|  |  |
| --- | --- |
|  |  |
| (a) | (b) |

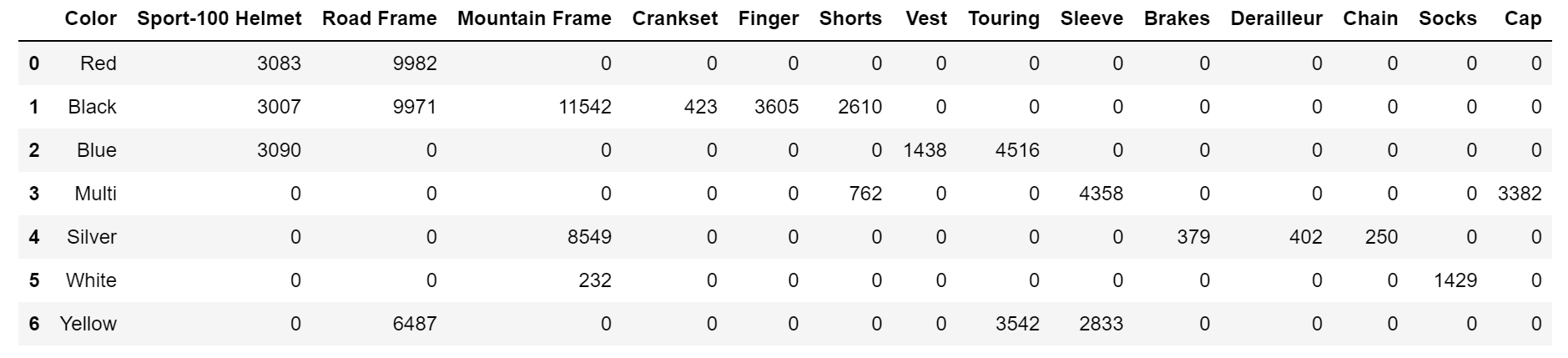
Figur 1 (a) Variationen i totalt antal beställningar för olika färger, (b) Variationen i totalt priset för olika färger.

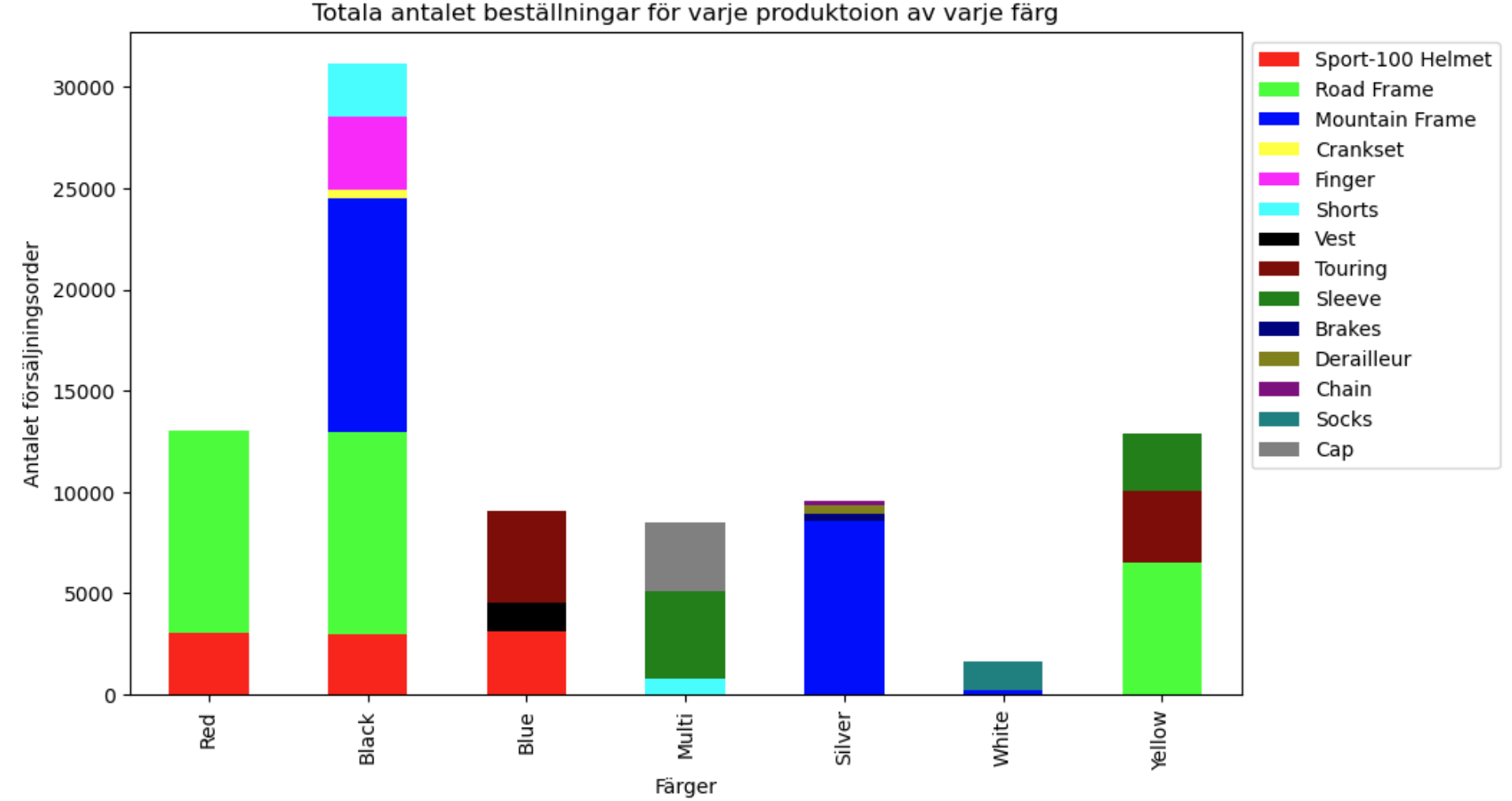
Figur 2 handlar om konfidensintervallet för genomsnittligt enhetspris, genomsnittlig kvantitet för varje order och genomsnittligt pris för varje order av varje olika färg, respektive. Man kan se från figur 2(a) att silver, röda och gula varor är dyrast men multi och vit är billigast. Däremot kan man se från figur 2(b) att genomsnittlig försäljning kvantitet av vit och multi är högst. När man tittar på figur 3 får man veta att kläder, keps och strumpor är huvudsakligen vita och flerfärgade och har den största försäljning kvantitet. Då kan man utveckla några nya produkter inom sportkläder med vit och multi färger.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (a) | (b) | (c) |

Figur 2 Beräknade konfidensintervallet för (a) genomsnittligt enhetspris, (b) genomsnittlig kvantitet för varje order och (c) genomsnittligt pris för varje order av varje olika färg, respektive.

Tabell 1 Det totala antalet försäljningsorder beroende på olika färger med fördelning i olika produkter.





Figur 3 Det totala antalet försäljningsorderna beroende på olika färger med fördelning i olika produkter.

Del 2 Förhållande mellan mäng av kund och säljare

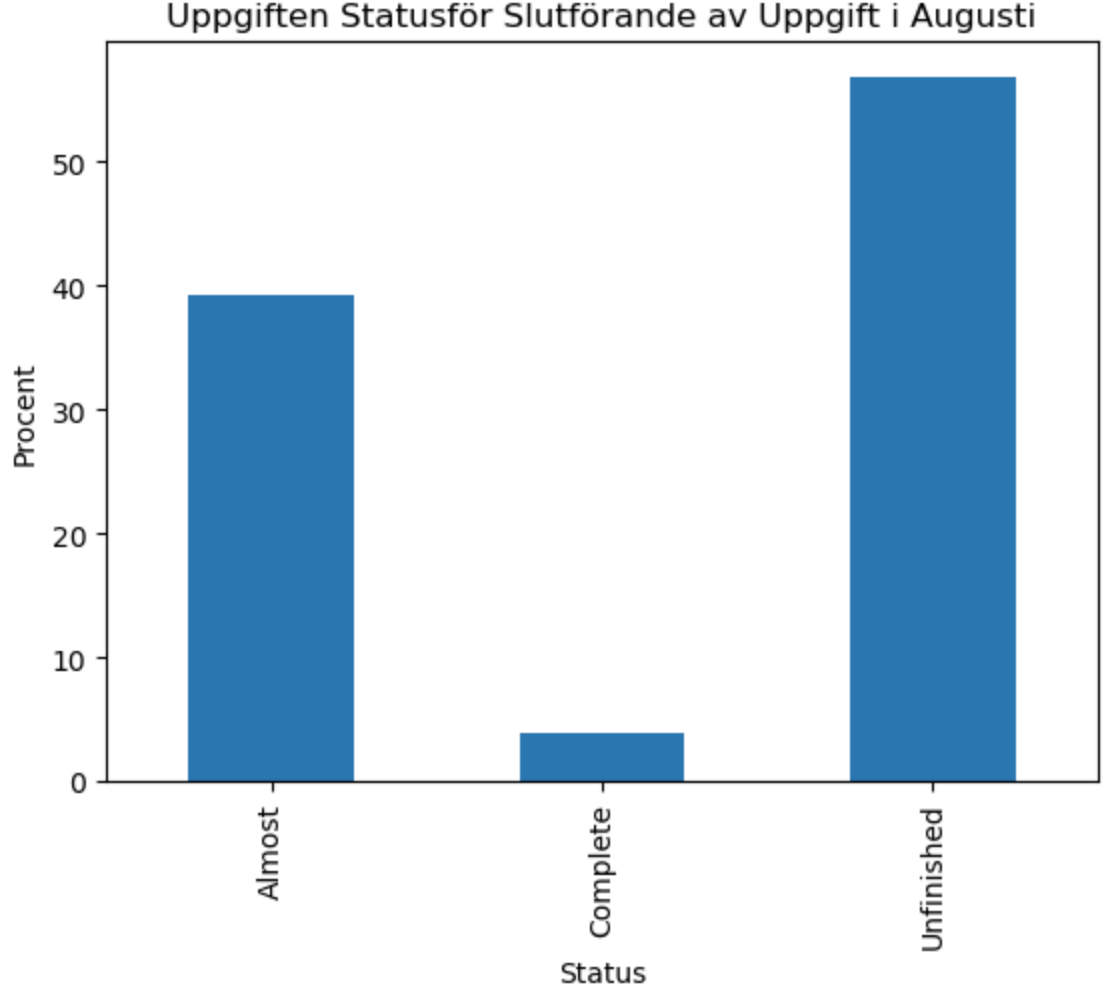
I den här delen ska vi analysera förhållande mellan mängd av kund och fördelning av total försäljning på olika region så att fatta konsumkraft i olika region. Från figur 2(a) hänvisar vi att antalet kunder i sydvästra USA är störst och följs av Australien och nordvästra USA. Men total försäljning i Australien är svag i de tre områdena. Dessutom anger det att köpkraften i nordvästra USA är mest stark. Trots att antalet kund i Kanada är mindre än i Australien är den totala försäljningen i Kanada är tydligt högre än i Australien. Det betyder att chefer kan ordna mer säljare till Kanada på grund av att köpkraft är starkare i Kanada. Jämförande med andra delar av regionen från figur 2 kan vi konstatera att köpkraften i Centrala, nordöstra och sydöstra delarna av USA är liknande den i Frankrike, Tyskland och Storbritannien. Men antalet av kund i store i Frankrike, Tyskland och Storbritannien. Då antalet av säljare kan distribuera lite grand mer i Frankrike, Tyskland och Storbritannien för att förbättra, attrahera och betjäna de potentiella målkunderna.

|  |  |
| --- | --- |
|  |  |
| (a) | (b) |

Figur 4 (a) Fördelning av kunder i olika regioner, (b) Fördelning av total årlig försäljning i olika regioner.

Del 3 Försäljningsmålsplan

I den här delen ska vi analysera slutförd uppgift status av försäljningsuppgifter i augusti, vilken kan ge datastöd för ledare att formulera nästa årsplaner och försäljning. Från figur 5 kan vi se att status av den slutförd uppgifter är jättedålig. Endast mindre än 4% av produkterna överskred det målsantalet av försäljning. Försäljningsstatus av resten av produkterna var långt ifrån idealisk mål. Om vi minska försäljningsmängd med 3 st för varje produkt skull det ha cirka 43% av produkterna nått det målsantalet av försäljning. Då kanske det är ett bra sätt att minska målsantalet av försäljning med 3 till 5 st för varje produkt näst år. På motsvarande sätt kan verkstadstillverkningen även göra mindre justeringar för att minska produktlager och sparar kostnader.



Figur 5 Slutförande av försäljningsuppgifter i augusti.

3. Slutsatser och rekommendationer

Enlig den analyserande data i delen 1 slutsatser jag att svart är det mest populära färg och att man kan försätta expandera svarta produkter till fler områden, till exempel kläder och cykeltillbehör. Gul och röd färg är de andra och tredje populära färg. Dessutom kan man utveckla blå cyklar på grund av att blå hjälmar är mer populära än röda och svarta hjälmar. Designer kan skapa några nya sportkläder med vit och multi färger. I delen 2 uppfattar vi att köpkraften i nordvästra USA är mest stark, följande med sydvästra och Kanada. Då chefen kan ordna mer säljare till Kanada. Å andra sidan, kan antalet av säljare vara större i Frankrike, Tyskland och Storbritannien för att förbättra och locka till sig de potentiella målkunderna. I delen 3 får man veta att slutförd försäljningsuppgift är väldigt dålig. Så nästa årsplaner och försäljning måst vara minska lite.

4. Exekutive summary

I den här rapporten analyserar vi hur färg påverkar totalt antal beställningar och totalt priset av olika varor. Dessutom använder vi statistiska metoder beräkna konfidensintervall för genomsnittligt enhetspris, genomsnittlig kvantitet av varje olika färg och genomsnittligappriset för varje order av varje olika färg. Vi diskuterar också förhållande mellan mängd av kund och fördelning av total försäljning på olika region och försäljning uppgifter i augusti baserade data. Vi får några intressanta och viktiga slutsatser.

5. Datum för muntlig presentation

Jag ska göra muntlig presentation den fjärde januari 2024.

==Rapporten ska vara 2-5 sidor. Det går bra att skriva mer, speciellt om man använder grafer== =================MEN skriv koncist och fundera på vad du vill lyfta fram================

============================Börja reflektionen på en ny sida============================

=========================Längden av reflektionen är ungefär 1 sida=======================

# Reflektion på eget arbete

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

Jag upptäckte att skriva kod och knacka på tangentbordet är jätteviktig än att läsa boken och lyssna video när man studera ett programmeringsspråk och teknik. Till exempel jag har läst boken <Sams teach yourself SQL in 10 minutes, 4th edition> av författare Forta. B innan SQL kursen började. I den boken fattade jag många metoder om hur hanterar och analyserar data med SQL kod. Jag kände många query för att sortera olika parameter och logisk bakom den. Men när jag själv skrev query kod kan jag inte komma ihåg att använda några speciell parameter, till exempel ”FETCH”, ”LIMIT”, ”BETWEEN .. AND ..” på grund av att jag inte tränar mycket använda de. Därför tycker jag att praktik är viktigare. Medan jag gjorde den kunskapskontrollen skrev jag några rader av SQL kod varje dag från nätet för att inte glömde de. Jag fick många användbar SQL och Python kunskap och färdigheter på nätet. Å andra sidan har jag lagt märke till att språk är viktigt som jag inte har förmåga att berättar och förklarar bra mina data. Jag planerar att studera om svenska under sommarsemester. Jag lärde mig många saker bland annat SQL och Python när jag gjorde kunskapskontroll. Det är jättebra upplevelser och i framtiden ska jag göra några liten uppgift själv från nätet för att träna.

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

Enligt min egen inlärningssituation, ansträngningsnivå, kunskaper som jag fick tycker jag att jag kan få VG. Den främsta anledningen är att jag har medvetenhet och förmåga att lära mig självständigt och jag pluggar varje dag. Jag var färdig i tid alla lärandeinnehållet som läraren berättade. Dessutom läst jag en SQL bok självständigt i förväg som heter ”Sams Teach Yourself SQL in 10 Minutes”. För det andra var jag inte frånvarande någon kurs. Och jag var modig att ställa några dumafrågor på svenska. Det tredje och viktigaste är att jag inte bara trånade kunskapen jag har lärt mig men också lärt mig mycket nya kunskap när jag gjorde kunskapskontrollen. Denna kunskapskontroll hade en stor inverkan på mig. Först och främst insåg jag att lärande inte handlar om att lära sig kunskapen i läroböcker och i den tiden. Praktisk är tvärtom viktigare. En annan känsla är att det in är viktigt när man ska lära sig vilken kunskapspunkt. det viktigaste är att kontinuerlig inlärning. Att inse den tidigt är mycket viktigt för mig att söka jobb i framtiden. Den fjärde anledningen är att den här kunskapskontrollen är en öppen testning. Varje student ska analysera samma data och gräva värdefull information enligt sina egna intressen och idéer. Så alla uppdrag är unikt och speciellt. Till exempel en av mina klasskamrater tog reda på längden på semester och sjuk för nya och gamla anställda. Detta var mycket värdefullt och intressant att veta personalernas situation och SQL-koden hon skrev var kort, tydlig och snyggt. Men jag skulle vilja upptäcka den totala försäljningen av produkten, såsom totalt antal beställningar och total priset av produkter med olika färger, etc., så att man kan veta vilka färger och produkter som säljer bäst och kan fokusera på försäljning eller någon förbättrad försäljningsarrangeman. Dessutom är det också intressant och värdefulla att förstå försäljningsområdet, relationen mellan försäljningsstatus och kundsdistribution för varje område, och vilka kunder som är högvärdiga kunder eller har hög köpkraft samt identifiera potentiella kunder. Så jag tycker att jag kan få VG för min rapport.

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

I början av SQL kursen har jag redan förstått grundläggande begrepp, syntax, nyckelkod och relationslogik, så jag kan snabbt fattade vad läraren berättade om. Det viktigaste är att använda verkliga data såsom ”AdventureWorks2022” och ”WideWorldimporters” för att träna. Jag upplevde förstå, inte förstå och återförstå tre processer på grund av att jag inte så bekant kod och data. När jag gjorde rapport repeterade jag SQL och Python kunskaper varje dag vilket ha hjälpt mig att stärka minnet av syntax och koncept. Dessutom insåg jag att jag kan minna bättre när jag diskuterar, pratar och ställa några dumafrågor i kursdiskussioner. Och mina klasskamraters perspektiv och erfarenheter är speciella och värdefulla för mig att inlära med.