Del 1 – Teoretiska frågor

1. Relationsdatabas lagrar och organiserar datapunkter med definierade relationer för snabb åtkomst. Det är data som innehåller information om varje föremål och står som typ för diverse kategorier i rader och kolumner som i sin tur är staplade i tabeller.
2. CRUD står för create, read, update och delete. Det ska möjliggöra att man kan skapa data, ha tillgång till data, uppdatera och radera data.
3. När man vill att varje rad från t.ex. första raden i tabellen oavsett om det matchar med raden från andra tabellen så kan man använda’’left join’’. En’’inner join’’ däremot är lite striktare än ’’left join’’ med anledning av att den används när man vill använda endast rader som matchar en ON villkor. Skulle inga rader matcha ON villkoren resulterar det i att vi inte kommer få några resultat.
4. Indexering i SQL underlättar att hitta information som är lagrad i en databas. Detta är optimalt om man har en stor databas och genom indexering kommer datorn att leta i varje rad tills informationen man söker hittas.
5. En vy i SQL gör det lättare att kunna titta på data. Det refererar till namngivna kolumner och rader med data.
6. En lagrad procedur i SQL innebär ett avsnitt av förbered SQL kod som går att spara och kan även återvändas. Till exempel om man har en query som man skriver flera gånger om så kan man spara det som lagrad procedur och sedan ropa koden för att det ska köras. Det går även att spara parametrar som lagrad procedur så att den kan agera utifrån tidigare parameter-värden.

Del 2 – Rapport

2.1 Inledning

Denna rapport kommer att fokusera på att studera databasen AdventureWorks2022 som har upprättats i Python genom att integrera det med SQL. Det primära fokuset är att studera query Production.WordOrder och dess relationen mellan två queries ’’scrapped quantity’’ och ’’stocked quantity’’. Det har även studerats skillnaden mellan ’’scrapped’’ och ’’stocked quantities’’. Anledningen till att just ’’scrapped’’ och ’’stocked quantity’’ valdes är för att studera sambandet mellan produktionen av varorna och hur mycket som faktiskt slängas och vilka som finns i lager för försäljning samt vad produktionsansvariga kan förbättra. Det har inte studerats vad för varor som har producerats utan samtliga varor har varit en faktor i analysen. Det har heller inte studerats orsaken till att dessa varor har slängts.

Tidsperioden maj 2014 valdes mot bakgrund av att det skulle vara lika många varor som förekommer i både ’’scrapped’’(29 rader) och ’’stocked’’ (30 rader) för att kunna analysera datan. En left join har använts i analysen av datan för att studera en eventuell matchning i raderna som har ’’scrapped’’ och ’’stocked’’ samt skillnaden mellan dessa. Det är tre Work Order ID som matchar båda raderna och har resulterat i följande:

* 60 scrapped med en stocked quanity om 3101 under perioden 4 maj 2014 till 20 maj 2014,
* 109 scrapped med en stocked quantity om 4278 under perioden 4 maj 2014 till 14 maj 2014
* 42 scrapped med en stocked quantity om 1539 under perioden 4 maj 2014 till 20 maj 2014.

Ovanstående siffor och data har analyserats med hjälp av beräkning av medelvärdet, variansen, standardavvikelse och konfidensintervall så väl som normalfördelning för att veta om varorna som slängs utgör en risk för företaget.

Det som har framkommit under analysen är att ’’stocked’’ har betydligt högre medelvärde än medelvärdet för ’’scrapped’’ tyder på att produktionen har mindre varor att slänga. Jämförelsen mellan ’’stocked’’ och ’’scrapped’’ tyder även på att produktionen har en effektiv process för att minska felaktiga produktioner som i sin tur behöver slängas. Detta kommer att behandlas djupgående i nästa kapitel ’’analys av resultatet’’

2.2 Analys av resultatet

Som det har tidigare nämnts i inledningen är syftet med rapporten att studera relationen mellan ’’stocked’’ och ’’scrapped quantity’’ samt skillnaden mellan dem. Detta är för att veta om produktionen har ett effektivt arbetssätt som minskar varor som slängs i förhållande till vad som finns i lager. I detta avsnitt kommer fokuset vara på att analysera datan och resultatet och därefter kommer slutsats och rekommendationer att presenteras i ett annat avsnitt.

2.2.1 Stocked quantity

Som tidigare nämnt har endast 30 rader av ’’stocked quantity’’ för maj 2014 använts för att kunna få en rättvis statistisk analys av ’’stocked’’, ’’scrapped’’ och skillnaden. Analysen av datan kan eventuellt medföra felaktigt resultat då ’’scrapped’’ endast har 29 rader. Utifrån beräkningar så utgör det ingen större skillnad i analysen för att primära fokuset är att förstå sambandet mellan dessa queries men även räkna om produktionen av varorna påverkas av antalet varor som slängs så väl som om företaget arbetar utifrån ett effektivt arbetssätt och kostnadseffektivt.

Det som har framkommit med hjälp av en left join är tre Work Order IDs som matchar med ’’scrapped quantity’’. Det har vidare visat att under 4 maj 2014 till 20 maj 2014 fanns 3101 varor i lagret varav 60 fick slängas. Under perioden 4 maj 2014 till 14 maj 2014 fanns det 4278 varor i lagret varav 109 fick slängas. Under 4 maj 2014 till 20 maj 2014 fanns 1539 varor i lager varav 42 fick slängas. Denna period gäller även för ’’scrapped’’ och även skillnaden.

Det har gjorts en beräkning på samtliga Work Order ID i ’’stocked’’ för att veta vad medelvärdet, variansen och standardavvikelsen. Även om antalet varor i lager är relativt höga så förväntas ett högt medelvärde så väl som standardavvikelsen efter att variansen är uträknat. Medelvärdet för ’’stocked’’ är 2 979, variansen 1 270 034 och standardavvikelsen är 1 126. Variansen och standardavvikelsen visar hur mycket de avviker från medelvärdet och det indikerar på att det finns en variation i produktionens resultat. I och med att variansen är hög talar detta för att det eventuellt kan studeras vidare huruvida produktionens process innefattar olika typer av produkter med diverse komplexitet och då kan man förvänta sig en hög variation som 1 270 034. Därtill visar standardavvikelsen att den avviker från medelvärdet med 1126 vilket kan också bero på variationer i resultatet.

2.2.2 Scrapped quantity

I ’’scrapped quantity’’ har 29 rader använts för att kunna studera hur många varor slängs efter produktion. Det har därtill använts, precis som i ’’stocked’’, en left join som i sin tur resulterade i tre Work Order IDs som kunde analyseras. Det som har framkommit är att 42 varor fick slängas av 1539 under perioden 4 maj 2014 till 20 maj 2014, 109 varor fick slängas av 4278 under perioden 4 maj 2014 till 14 maj 2014 och slutligen under perioden 4 maj 2014 till 20 maj 2014 fick 60 varor slängas av 3101.

Likt tidigare kapitel har samma metod använts för att beräkna medelvärdet, variansen och standardavvikelsen. I detta fall är antalet varor som har fått slängas lågt och därmed förväntas ett lägre medelvärde, varians och standardavvikelse. Medelvärdet för ’’scrapped’’ är 70, variansen är 801 och standardavvikelsen är 28. Detta innebär att variansen har en måttlig spridning och att standardavvikelsen indikerar på att det avviker med ungefär 28 från medelvärdet på 70. Detta visar sammantaget att variansen ger en mer konsekvent nivå av ’’scrapped’’ kan också vara ett resultat på att medelvärdet är låg.

2.2.3 Skillnaden mellan stocked och scrapped quantity

En jämförelse mellan ’’stocked’’ och ’’scrapped quanities’’ har gjorts med anledning av att det är intressant att veta skillnaden mellan varandra och om de påverkar varandra i större utsträckning. Båda queries är kopplade till varandra men frågan i denna rapport är att ta reda på om det har en stor påverkan på produktionen när man slänger produkterna.

Det har presenterats antalet varor som finns i lager (stocked) och de som har fått slängas (scrapped) i tidigare kapitel. Skillnaden mellan dessa queries är, 3 041, 4 169 och 1497 varor från tre samma Work Order IDs som tidigare. Samma metoder har använts för att beräkna medelvärdet, variansen och standardavvikelsen. Medelvärdet är 2902, variansen är 1 199 544 och standardavvikelsen är 1095. Medelvärdet innebär att skillnaden visar att fler varor är lagrade än vad som har slängts. Det kan dock även betyda att det finns en obalans i produktionen med att lagra för mycket som är mer vanligt än att slänga. Det som talar för varför variansen och standardavvikelsen är relativt höga kan beror på att vissa Work Order IDs lagrar för många produkter för att det är mer hållbart medans andra har en mer balanserad och effektiv produktion.

2.2.4 Konfidensintervall av stocked, scrapped quantities och dess skillnad

Urvalet för denna rapport har varit att studera tre Work Order IDs som matchar med varandra dock utgör det utrymme för en viss osäkerhet, det vill säga ett urvalsfel. Konfidensintervallet underlättar för att ta reda på felets storlek. Det har gjorts tre olika konfidensintervall utifrån queries så som ’’stocked’’, ’’scrapped’’ och skillnaden mellan dessa två. Detta för att ta reda på var medelvärdet av dessa queries ligger vid en normalfördelning.

För ’’stocked quantity’’ är konfidensintervallet avrundat till mellan -449.03 och 6408.36. För ’’scrapped quantity’’ är konfidensintervallet avrundat till mellan -15.80 och 156.47. Slutligen är konfidensintervallet för skillnaden mellan ’’stocked’’ och ’’scrapped’’ avrundat till -429.86 och 6234.52. Det innebär att det är 95 procent säkert att det sanna medelvärdet av ’’stocked’’, ’’scrapped quantity’’ och skillanden mellan dessa två queries faller inom respektive intervall.

Detta resultat möjliggör för ’’stocked quantity’’ att ansvariga inom produktionen kan fortsätta ha sin lagernivå inom ett visst omfång eftersom det tillhandahåller en nivå av säkerhet om var det sanna medelvärdet kan vara. För ’’scrapped quantity’’ möjliggör detta resultat för ansvariga som har som mål att minska på antalet produkter som slängs inom ett visst omfång. En kombination av att hålla nere antalet produkter som slängs och lagrar för mycket, är detta konfidensintervall väldigt lovande för produktionen.

2.3 Slutsats och rekommendationer

Det har i tidigare kapitel presenterats resultatet av datan för Production.WorkOrder med ’’scrapped’’ och ’’stocked quantities’’. Resultatet visar att företagets produktion av ’’stocked’’ har en betydande variation med en hög standardavvikelse. För att kunna åtgärda detta bör produktionsansvariga undersöka vilka faktorer som bidrar till en sådan variation och bedöma om det är i enlighet med produktionsmålen. Å andra sidan visar ’’scrapped’’ en måttlig nivå av variation och för att fortsätta minska antalet bör produktionsansvariga se över andra strategier för att kunna upprätthålla en jämn nivå av produktionseffektivitet. Produktionsansvariga kan med hjälp av konfidensintervallet ha en god planering över produktionen med fortsatt minskning av ’’scrapped’’.

Under arbetsgång av koden och analysen har det framkommit några frågor som skulle vara intressant att studera vidare för att kunna ha en effektiv produktion med mindre spill. För att kunna göra det krävs det vidare analyser av samtliga Work Order IDs som lagrar för mycket för ett hållbart sätt och de som har en balanserad och effektiv produktion. Detta för att veta vilka som har ett effektivt arbetssätt och är kostnadseffektivt. Det hade även varit intressant att studera vidare specifika Work Order IDs som har höga quantities för att förstå vilka dessa bidragande faktorer är avseende lagring och slängning av produkter.

2.4 Executive summary

Analysen av produktionen visar värdefulla insikter om effektivitet och resultatet av produktionsprocessen. Skillnaden ’’stocked’’ och ’’scrapped’’ visar ett genomsnittligt överskott på 2902 i ’’stocked’’. Med konfidensintervallet, som har tidigare presenterats för skillnader, finns det en möjlighet att optimera produktionsprocessen och minimera lagra för mycket.

Visualiseringarna i datan visar en tydlig representation av specifika Work Order IDs, vilket underlättar att identifiera mönster och ha en god översikt. Produktionsansvariga uppmuntras därmed att utforska andra Work Order IDs som har en hög quantity för att kunna förstå orsakerna men även potentiella förbättringsområden. Produktionen behöver se över att man har en mer balanserad, enhetligt och effektivt arbetssätt för samtliga Work Order IDs för att ha en hållbar produktion.

Sammanfattningsvis, belyser denna rapport vikten av att ta itu med olika variationer i ’’stocked quantity’’, minimera skrot och optimera produktionsprocessen. Produktionsansvariga har värdefulla verktyg som konfidensintervallen som är baserad på normalfördelningen och med dessa verktyg kan ansvariga med god säkerhet ha strategisk planering och resursallokering i produktionsmiljön.

2.5 Redogörelse

Mina utmaningar har varit att bekanta mig med SQL och hur det används samt tillföra det till Python. Det var ganska omfattande med filmerna som man skulle se på den korta tiden inför nästa lektion. Jag tyckte till slut att jag löste det genom att googla och kolla på flera videos. Mitt tips till mig själv skulle nog vara lugn och stressa inte för att det är inte funkar.

Jag skulle ge mig själv ett VG eftersom jag har reflekterat över mina val och analyserat resultatet av min kod. Jag har även förklarat varför det var intressant men också vilka fler studier man kan göra baserat på vad jag har hittat.

Jag gjorde den muntliga presentationen den 5 januari 2024 genom att spela in mig själv.