# Teoretiska frågor

**1. Beskriv kort hur en relationsdatabas fungerar.**

En databas där tabeller och data är relaterade till varandra.

**2. Vad menas med ”CRUD” flödet?**

Create

Read

Update

Delete

De olika metoderna vi använder för att skapa och hantera relationsdatabaser

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

Left Join är där vi kombinerar två tabeller där vi bevarar all data från den vänstra tabellen men bara det som har matchande data i högra tabellen.

En Inner Join är där vi kombinerar två tabeller men bevarar bara dem raderna som har matchande data i båda tabellerna.

**4. Beskriv kort vad indexering i SQL innebär.**

Indexering hjälper oss att kolla upp data i databasen snabbare. Den vanligaste typen av index i sql skapas automatiskt när vi bestämmer vad som är primary key i en tabell.

**5. Beskriv kort vad en vy i SQL är.**

En vy är en virtuell tabell baserad på "riktiga" tabeller. Det är alltså en sparad SELECT call som kallar på tabeller utan att lagra datan på nytt.

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

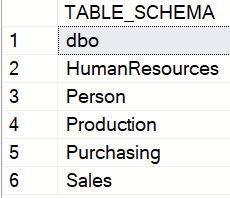
En lagrad procedur är en sparad kod som går att kalla på flera gånger utan att behöva skriva hela koden igen.

# Rapport

AdventureWorks är en amerikansk återförsäljare av mountainbikes och tillhörande utrustning.

Ken Sánchez har varit VD sedan 2009 med Laura Norman som ekonomichef.

AdventureWorks databas innehåller 72 tabeller indelat i 6 scheman.

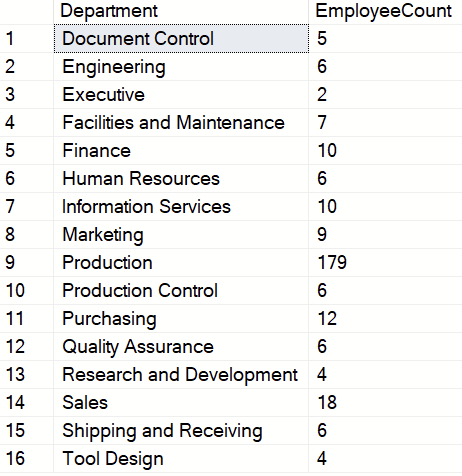


Databasen har säljdata från 2011-05-31 till 2014-06-30 så allt i rapporten är baserat på dessa 3 år.

Under dessa år hade dem 290 anställda spridda över 6 länder.

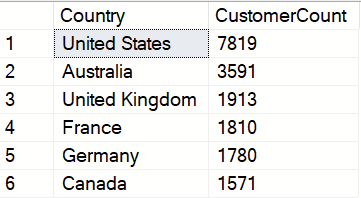


Företaget har 16 olika avdelningar där största avdelningen är Produktion



Dem har en snittlön på $18,19 / timme men $17,67 / timme om vi inte räknar med Executive avdelningen.

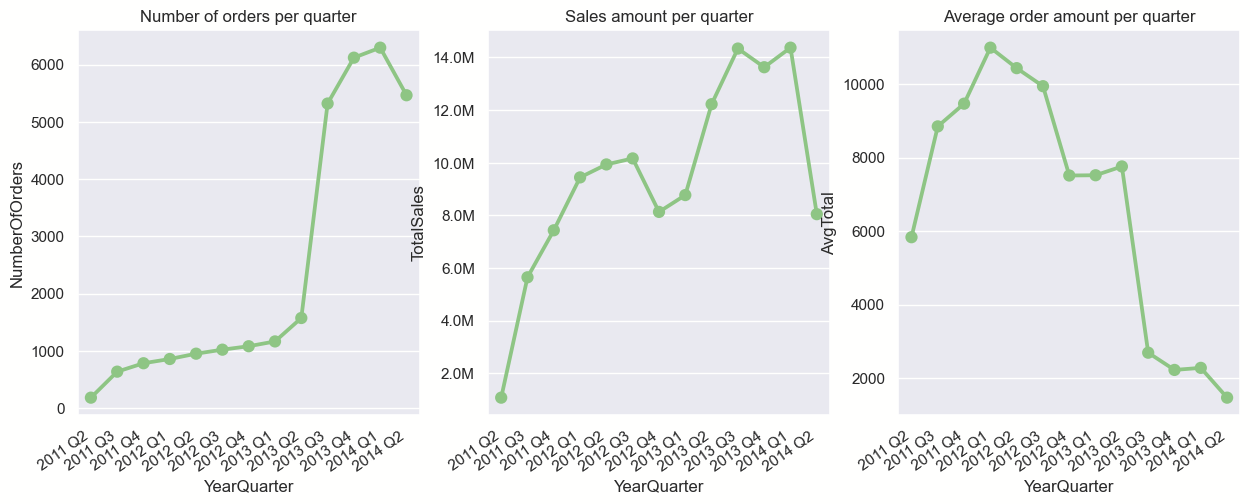
Dem hade 18 484 registrerade kunder spridda över 6 länder med den största marknaden i USA men deras internationella kundbas är större totalt.



Mellan 2011-05-31 och 2014-06-30 omsatte dem årligen i snitt $41M och har haft i snitt en ökning på omsättningen med $2,3M under den perioden.

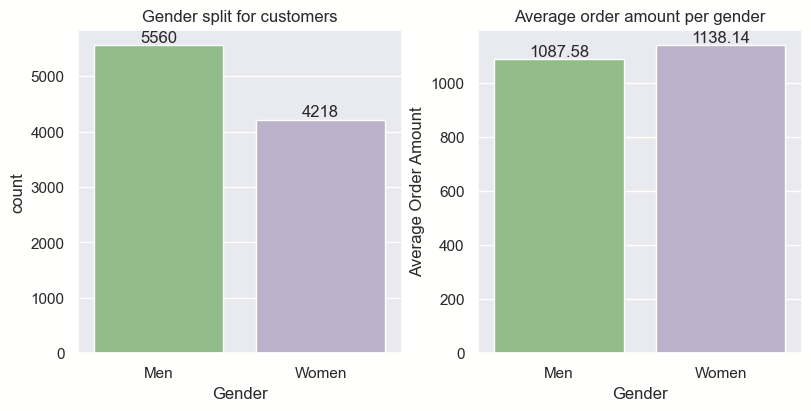
Det omfattar i snitt 10 500 ordrar per år med en årlig ökning på 1800 ordrar i snitt.

Antalet ordrar och försälning per år i snitt visar dock inte så tydligt hur det faktiskt ser ut därför tittar vi istället på varje kvartal för att ge oss en bättre bild.



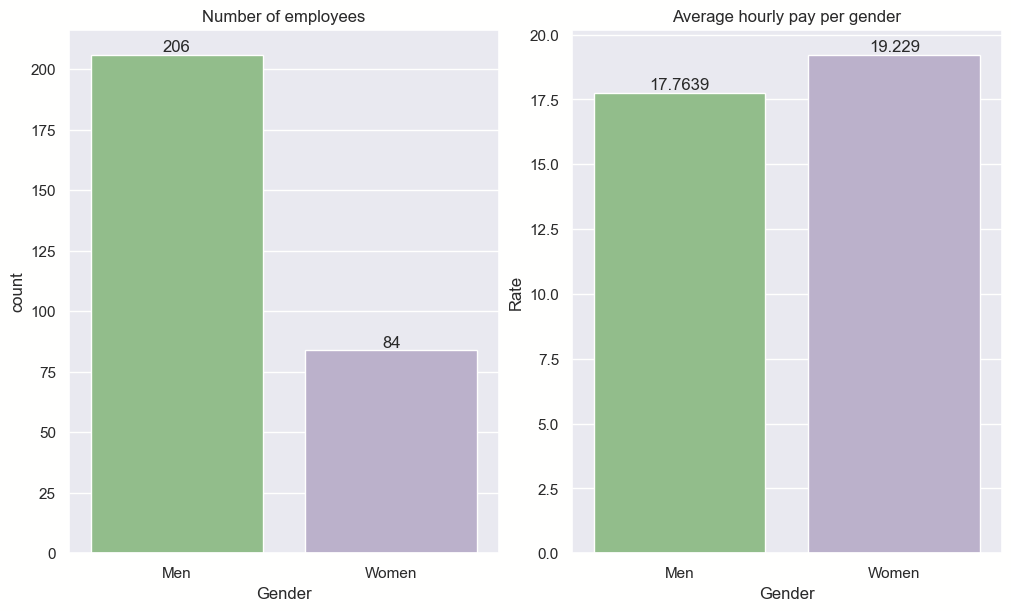
Vi ser här att antal ordrar ökade markant från Q2 2013 till Q3 2013 men totala försäljningen för samma kvartal ökade inte alls lika mycket. I procent så ökade antalet ordrar med 238% medans totala försäljningen bara ökade med 17,3%. I den tredje grafen ser vi att snittsumman för en order minskade markant över samma period och vi kan från detta dra slutsatsen att trots att ordrarna ökat så minskade värdet per order. Därför ser vi inte så stor ökning i totala försäljningen.

För att veta mer om kunderna så kan vi kolla på könsfördelningen och hur mycket respektive beställer för.

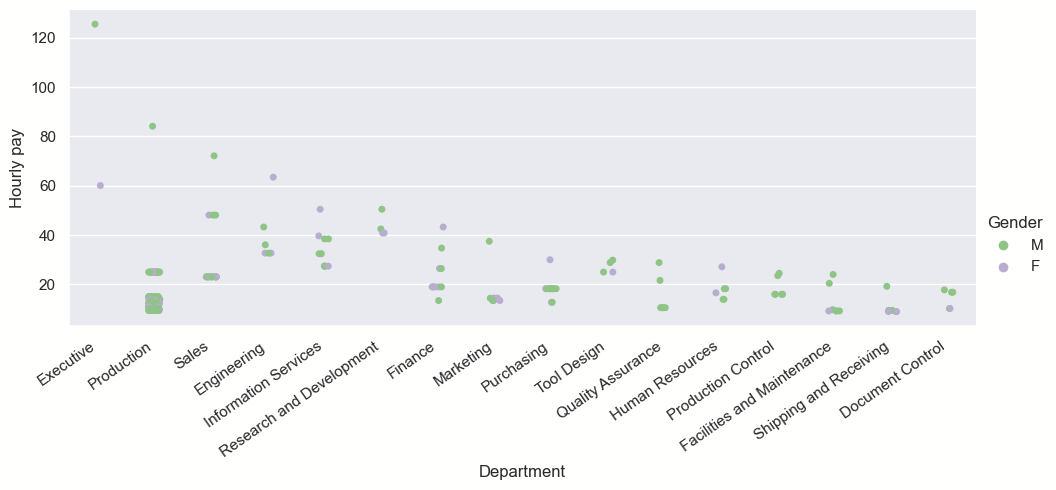


Här ser vi att det är fler män än kvinnor men vi ser också att det inte är alla registrerade kunder utan bara ungefär hälften. Dessa är bara dem som uppgett kön men vi kan använda detta för att ge en generell bild över alla kunder. Vi kan alltså anta att könsfördelningen för alla kunder ser liknande ut. Männen gör upp ca 56,9% av alla kunder men deras genomsnittsorder är lägre än kvinnornas så vi ser att männen gör fler små ordrar jämfört med kvinnorna. Med detta kan vi säga att med 95% säkerhet så kommer snittordern för alla manliga kunder att vara mellan $1057 och $1118.

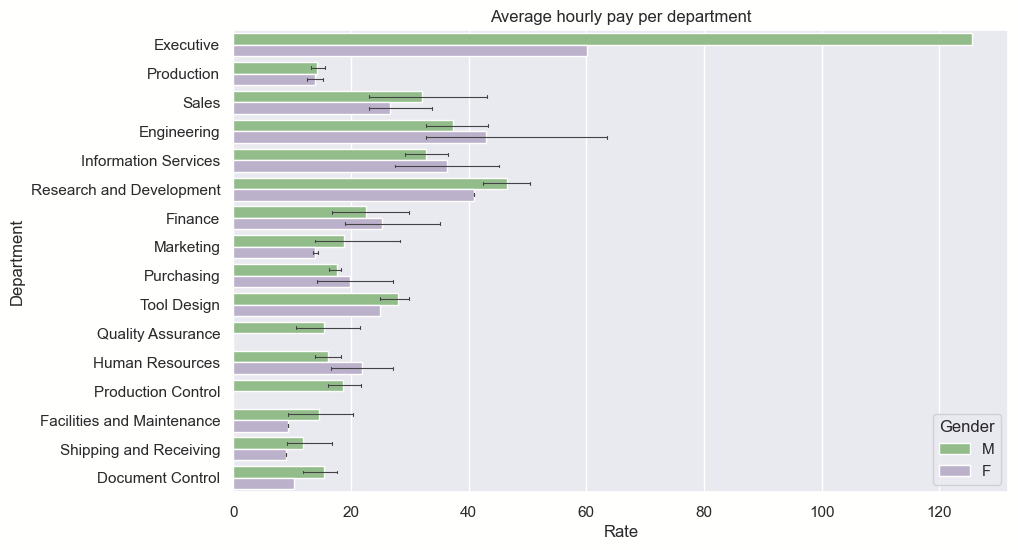
Denna fördelning av män och kvinnor är också väldigt lik könsfördelningen och lönefördelningen av dem anställda. Men jag tror dock inte att det finns något samband mellan dem.



Det är intressant att se att trots att VDn är en man så har kvinnorna högre lön i snitt, alltså att kvinnorna har högre positioner generellt sett än männen. Vi kan också kolla på hur fördelningen ser ut över dem olika avdelningarna.



Som vi såg innan så är det flest anställda i produktionen men har också en av de lägsta snittlönerna som vi kan se tydligare i grafen nedan.



I grafen så har vi med konfidensintervall men dem är inte helt korrekta då grafsystemet som vi använder inte kan räkna ut konfidensintervall helt korrekt på så små antal men det kan ändå ge oss en uppfattning om spridningen av lön i varje avdelning.

Sammanfattningsvis så såg vi att företaget har ökat antalet ordrar över åren men minskat i medelvärdet på varje order. Utifrån datan kan vi inte fastslå varför det är så men vi kan se att det är billigare varor så som kläder som köps mest. Företaget hade 18 484 registrerade kunder med totalt 31 465 ordrar vilket är i snitt 1,7 ordrar per kund. Det betyder att nästan varje kund kommer tillbaka för att beställa mer.

Muntlig presentation gjordes 2024-01-03

1. **Utmaningar jag haft under arbetet.**

Jag tyckte det var svårt att ha en bra struktur på rapporten. Jag har försökt få allt att hänga ihop men tycker fortfarande att det är ganska spritt. Jag hade även velat gå djupare på något men tiden sprang ifrån mig.

1. **Vilket betyg jag anser att jag skall ha.**

Jag anser att jag uppnått VG nivå i aspekterna att jag har reflekterat över vilken data jag tagit fram och dragit slutsatser utifrån den data. Jag tycker dock att jag varit lite för ytlig och spridd i min analys.

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

Jag hade sagt till mig själv att gå djupare på ett område tidigare så att jag hade haft mer tid att analysera och dra slutsatser.