

1. Teoretiska frågor

1. *Beskriv kort hur en relationsdatabas fungerar.*

I en relationsdatabas är datan organiserad i tabeller bestående av rader och kolumner, där varje kolumn har en viss datatyp. Tabellerna kan länkas samman på olika sätt genom unika identifierare, s.k. "nycklar".

2. *Vad menas med "CRUD" flödet?*

"CRUD" står för de fyra grundläggande funktionerna i en databashanterare; Create, Read, Update och Delete. Med Create (INSERT i SQL) skapar vi tabeller eller lägger till rader. Med Read (SELECT i SQL) söker vi upp den data vi vill åt. Med Update (UPDATE i SQL) uppdaterar vi viss data och med Delete (DELETE i SQL) tar vi bort viss data.

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

En **"left join"** är en join mellan två tabeller som har ett visst gemensamt index/nyckel. Från den vänstra tabellen väljs alla rader och från den högra väljs endast de rader som har det gemensamma värde som valts.

En **"inner join"** är en join mellan två tabeller som har ett visst gemensamt index/nyckel. Endast de rader som har detta värde i båda tabellerna väljs ut.

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

Indexering i SQL gör att sökningar och sammanslagning (joins) av tabeller går smidigare. Samtidigt gör indexering att det går långsammare att lägga till, ändra eller ta bort en rad, eftersom indexet då måste byggas om.

5. *Beskriv kort vad en vy i SQL är.*

En vy (view) i SQL är ett sparad resultat av den tabell som skapats utifrån en viss frågeställning (query). En vy kan sedan användas som en tabell för att till exempel göra ytterligare frågeställningar eller "join" med en annan tabell.

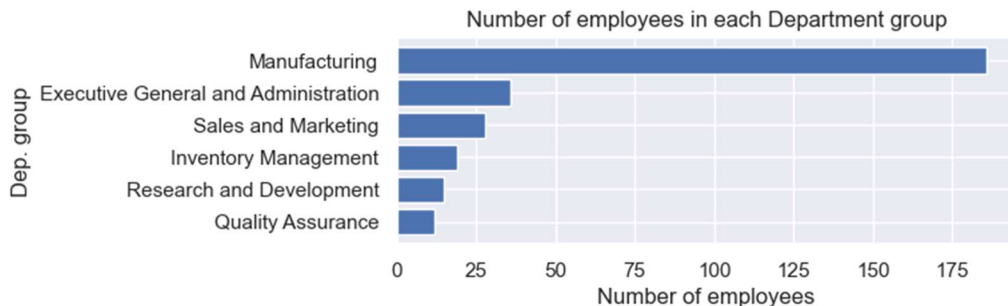
6. *Beskriv kort vad en lagrad procedur i SQL är.*

En lagrad procedur (stored procedure) i SQL är en typ av funktion som ger ett resultat utifrån valda värden på angivna parametrar. En lagrad procedur kan vara särskilt användbar vid upprepade komplexa frågeställningar.

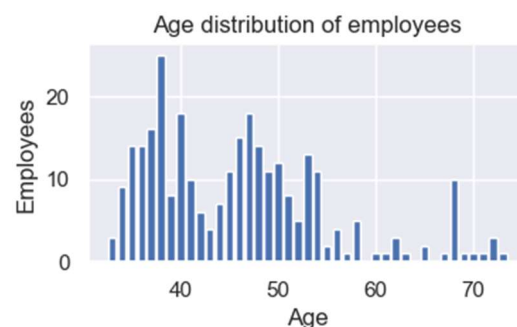
2. Programmeringsuppgift och Rapport

Deskriptiv sammanfattning

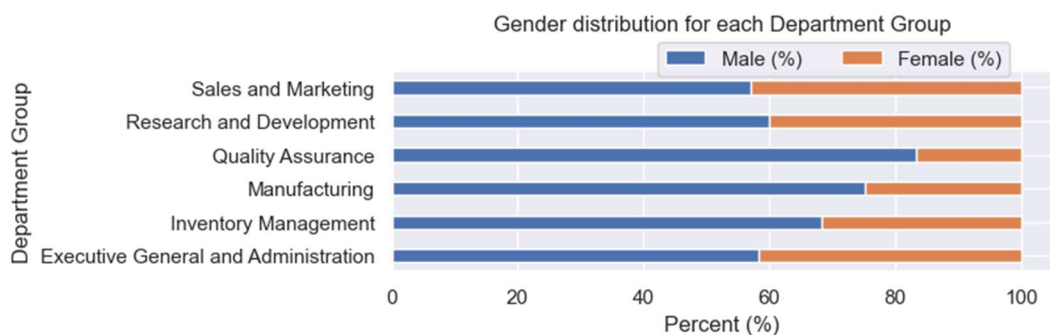
Företaget har totalt 290 anställda, fördelade på de olika avdelningsgrupperna enligt bilden nedan. Tillverkningsavdelningen har betydligt fler anställda än övriga avdelningar.



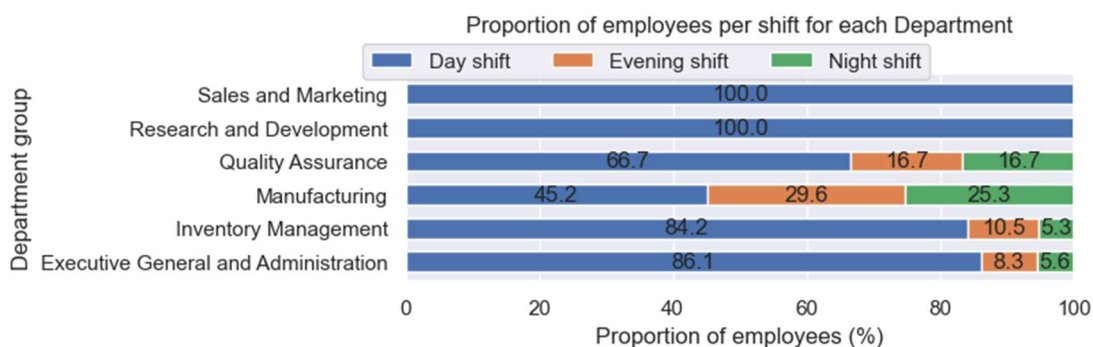
De anställdas ålder varierar mellan 33 – 73 år, enligt fördelningen i bilden till höger. De flesta av de anställda är i åldrarna 35–40 år samt 45–50 år.



De flesta anställda i företaget är män. De mest blandade avdelningarna avseende kön är Sales and Marketing och Executive General and Administration.



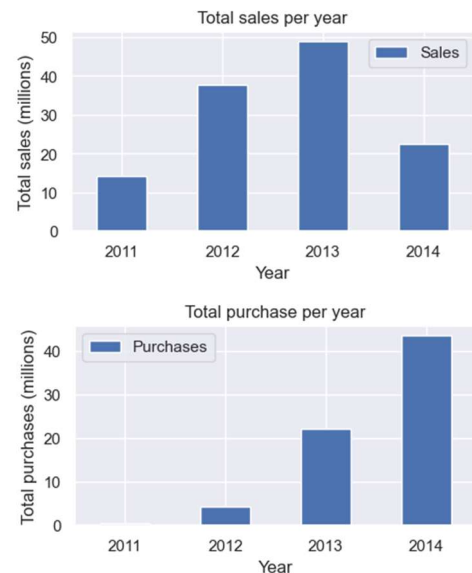
Flera av företagets avdelningar har skiftarbete, där de flesta arbetar under dagskiftet och en mindre andel under kvällsskiftet respektive nattsiftet.



Databasen innehåller bland annat försäljningsdata från 2011-05-31 till 2014-06-30. Data för inköpen sträcker sig under en lite längre period, från 2011-04-16 till 2014-09-22. Beloppen i databasen antas vara i dollar.

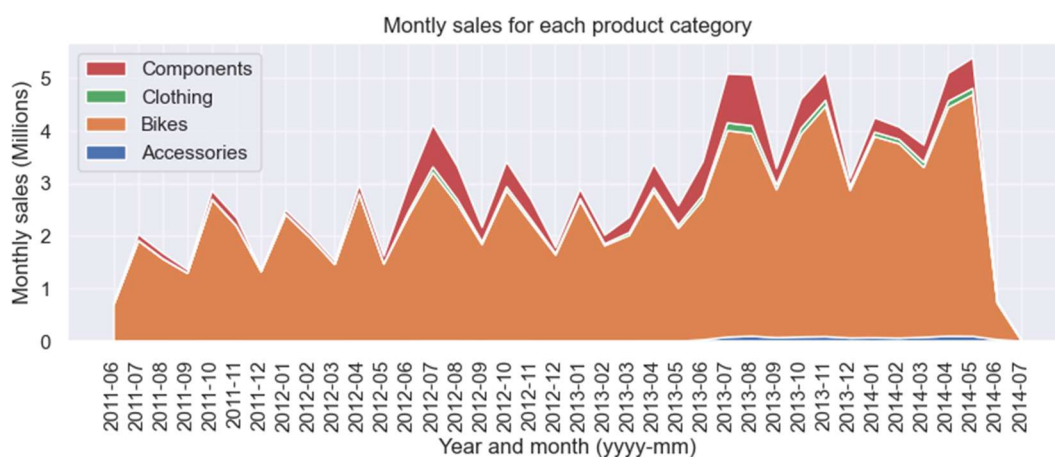
Den totala försäljningen har ökat under åren, men ökningen avtog något mellan åren 2012 och 2013. För året 2014 finns bara data första halvåret vilket gör att det året inte är fullt jämförbart räknat per år.

Inköpen har ökat mer och mer under åren och det har gjorts inköp under 2014 för dubbelt så stort belopp som under 2013, trots att det inte finns data för hela år 2014.

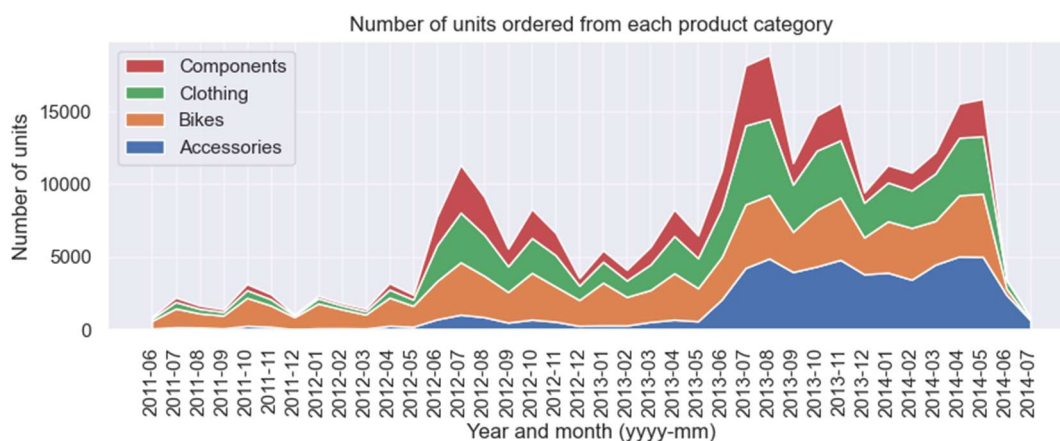


Statistisk analys

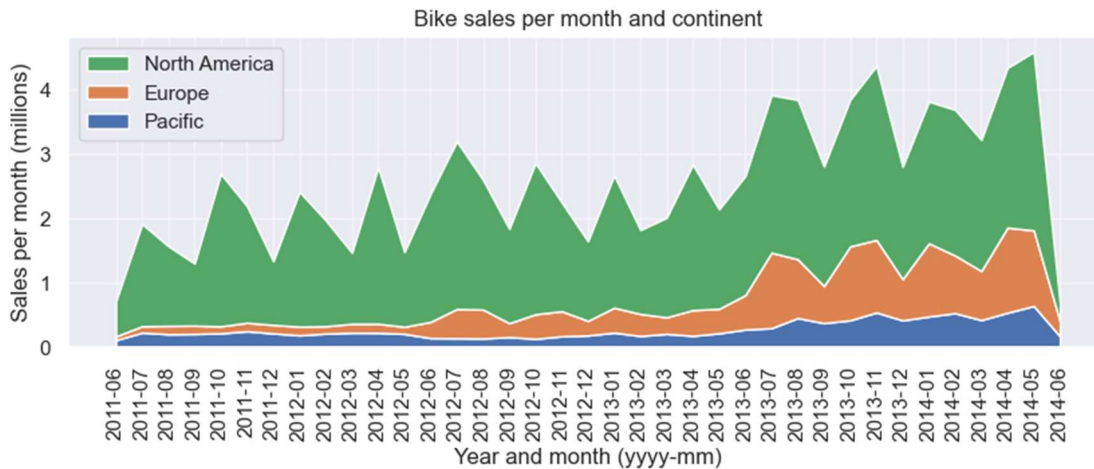
En vidare analys görs av försäljningsdata. Inom produktförsäljningen studeras försäljningsordrar. Syftet är att hitta eventuella mönster eller samband, för att kunna ge rekommendationer som kan optimera försäljningen ytterligare. Sett till försäljning mellan de olika produktkategorierna, är det kategorin "Bikes" som säljer för klart störst belopp och övriga kategorier ser relativt obetydliga ut i jämförelse.



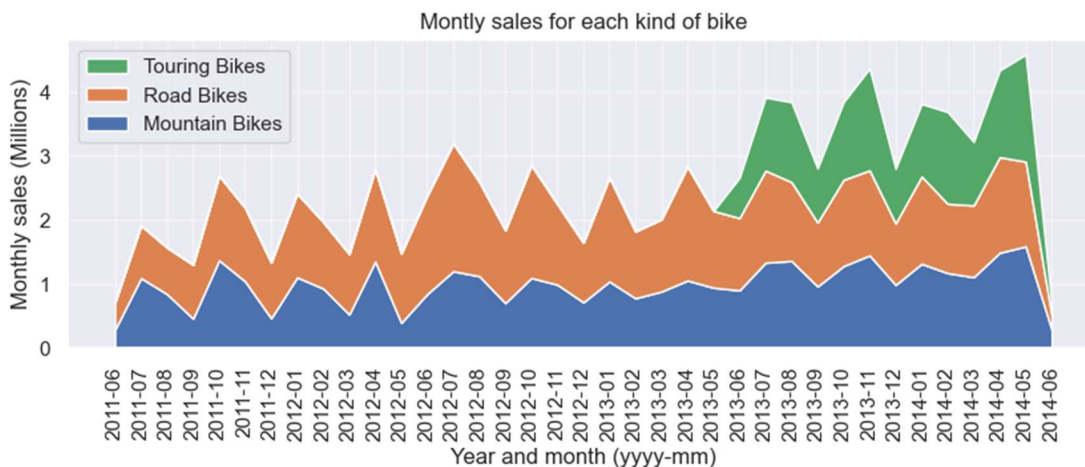
Men ser vi i stället till antal sålda enheter blir det genast tydligt att det även säljs många andra produkter, men till lägre pris än för cyklarna.



Försäljningen av cyklar sker mest inom Nordamerika men för övriga representerade kontinenter ökar försäljningen, särskilt från sommaren 2013 och framåt. De övriga kontinenterna utgör en större andel av den totala försäljningen under det sista året än de gjorde under det första året.



Företaget säljer tre typer av cyklar. I grafen nedan kan utläsas att försäljningen av "Mountain Bikes" och "Road Bikes" har varit relativt konstant under åren och när "Touring Bikes" började säljas inför sommaren 2013 så har den totala försäljningen ökat ytterligare.

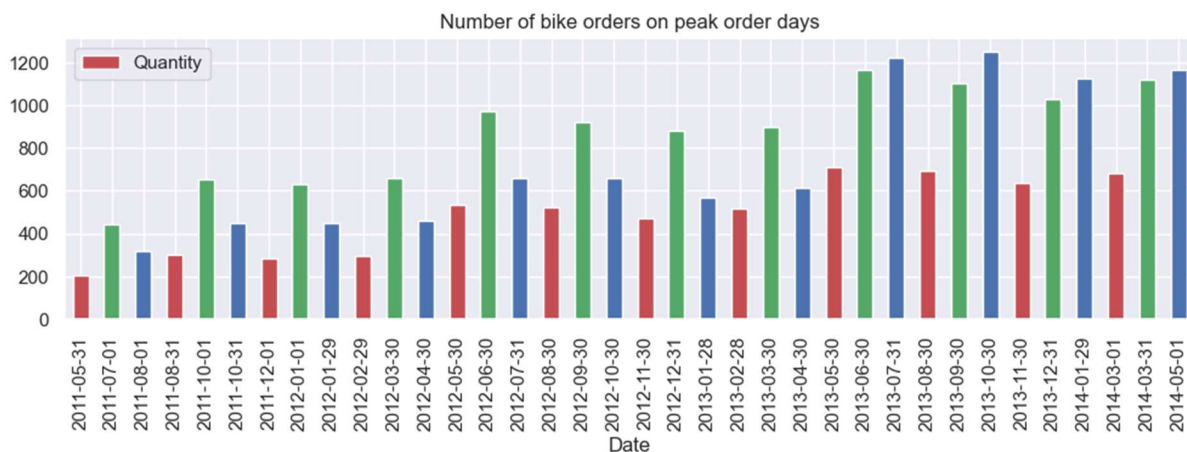


Genom att analysera antalet cyklar per orderdag, framträder ett intressant mönster enligt grafen nedan.



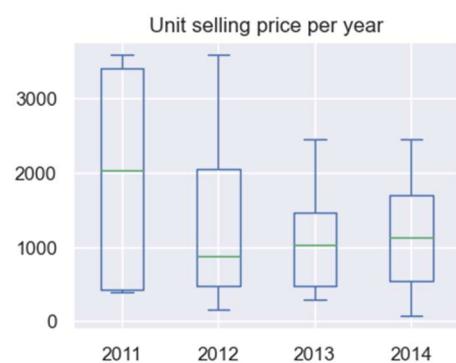
Det visar sig att det görs flest cykelköp ett visst datum i månaden, kring månadsskiftet, antingen i slutet av månaden eller i början av nästa månad. Efter en jämförelse av vilka kunder som handlar

dessa datum, visar det sig att de kunderna handlar cyklistiskt var tredje månad. Dessa kunder består av tre grupper (representerade här av varsin färg) som varvas var tredje månad, enligt grafen nedan.

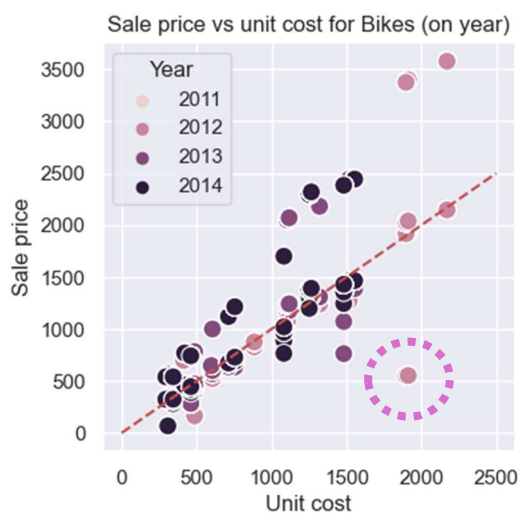
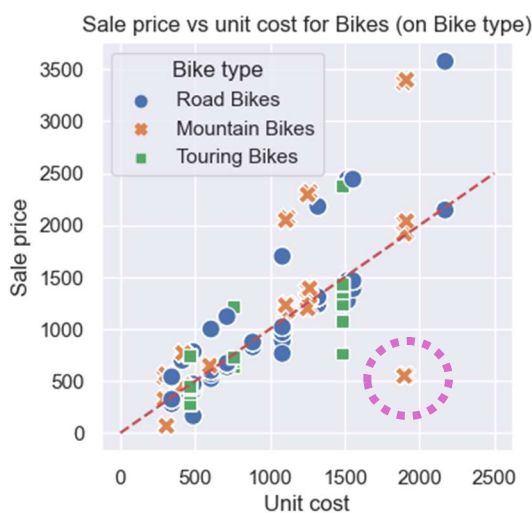


Sett till försäljningspriset för cyklar så varierar det mycket över tid och olika mycket för olika år.

Year	count	mean	std	min	25%	50%	75%	max
2011	3826.0	1784.35	1258.73	397.37	419.46	2024.99	3393.74	3578.27
2012	10776.0	1263.95	922.84	164.43	469.79	874.79	2039.99	3578.27
2013	16485.0	1122.32	697.00	282.14	469.79	1020.59	1466.01	2443.35
2014	8944.0	1246.42	743.64	67.80	539.99	1120.49	1700.99	2443.35



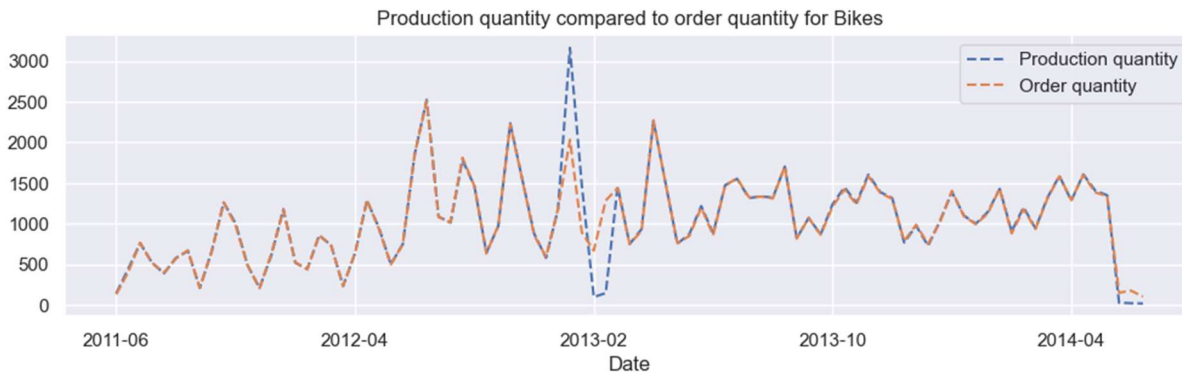
Genom att göra en "scatter plot" av försäljningspriset jämfört med tillverkningspriset kan vi se spridningen av enheter som sålt med vinst eller med förlust. Den röda streckade linjen visar "break-even", där enheter ovanför linjen sålts med vinst och enheter under har sålts med förlust. Många enheter säljs kring "break-even" och flera säljs med bra vinst, men det finns även enheter som säljs med förlust. Produkter som sticker ut särskilt har ett tillverkningspris på ca 1900 men ett försäljningspris på bara ca 500 (se nedan inringade punkter).



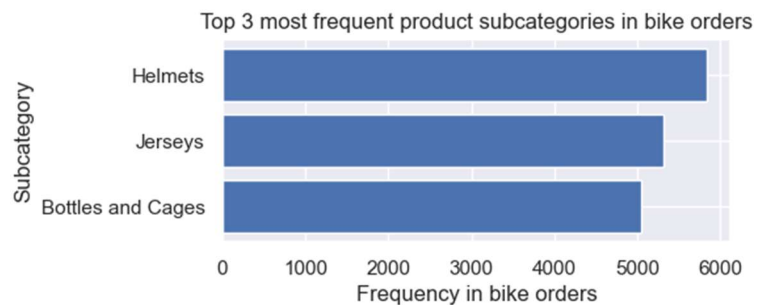
De cyklar som sålts till mycket lägre pris visar sig ha sålts 2012-04-30, och är de sista enheterna innan den cykelmodellen slutar säljas.

Order date	Sale price	Unit cost	Bike model	Color	Quantity
2012-04-30	548.435875	1898.0944	Mountain-100	Black	69
2012-04-30	552.498375	1912.1544	Mountain-100	Silver	68

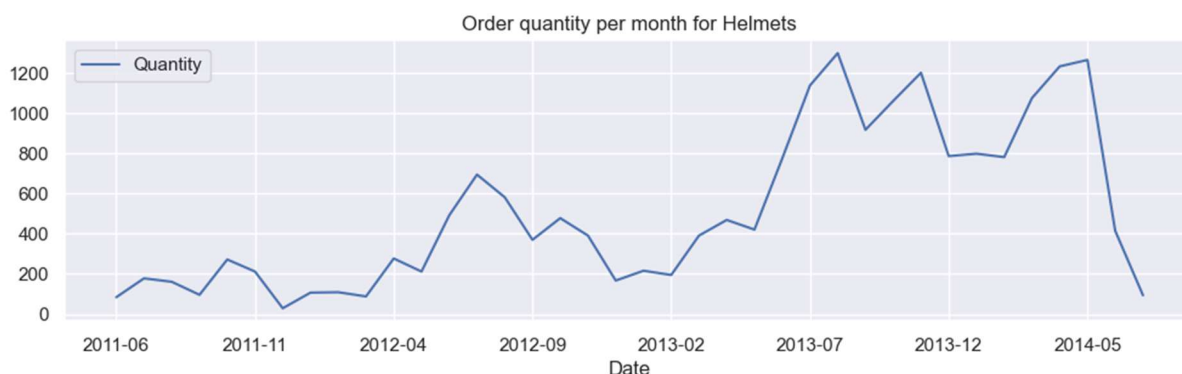
En teori kan vara att det producerats för många cyklar och att företaget sedan försökt bli av med dem. Genom att studera antalet tillverkade cyklar jämfört med antalet beställda cyklar ser vi att det inte är fallet. I stället syns en överproduktion av cyklar vid årsskiftet 2012/2013, vilket kompenseras med att tillverkningen minskar under några månader tills antalet är i våg igen.



Genom analyser av vilka produkter som är vanligast att köpa i samma order som cykelköp kan vi se köpmönster som kan vara intressanta gällande framtida marknadsföring. Den produktgrupp som återkommer oftast i samma order som cykelköp är hjälmar.



Antalet hjälmar som säljs per månad ökar markant inför sommaren 2013, vilket kan ha sin förklaring i att "Touring bikes" började säljas då och ökade på försäljningen.



I genomsnitt innefattar en cykelorder hjälm/-ar i ca 14,59 % av fallen, räknat på alla cykelordrar under hela försäljningsperioden. Konfidensintervallet för andelen hjälmköp i samband med cykelköp, med konfidensgraden 95 %, blir 0,1418 till 0,1499 (avrundat till fyra decimaler). Sett till alla cykelordrar som gjorts innebär det att intervallet, med 95 % sannolikhet, täcker det faktiska värdet.

Slutsatser och rekommendationer

Cyklar ger klart störst inkomst för företaget men det säljs även många andra produkter. Många av orderarna görs kring månadsskiftet och utgör en slags "peak days". Framtida marknadsföringsinsatser skulle kunna riktas efter dessa resultat. Exempelvis kan synlighetskampanjer sättas in i andra halvan av månaden för att speciellt rikta sig till de aktuella kundgrupperna.

De vanligaste produktgrupperna i samma order som cykelköp är hjälmar, tröjor och flaskor. Genom att marknadsföra dessa produkter ihop med cyklar kanske fler lockas av att köpa tillbehör när de köper cykel. Det bör även analyseras vidare om storlek, modell eller kulörer på produkterna kan optimeras för att förbättra försäljningen.

Företaget skulle kunna skicka ut enkäter för att få in mer data kring kundnöjdhet eller andra åsikter från olika kundgrupper (såsom ålder, kön eller socioekonomisk bakgrund). Sådan data hade gjort det möjligt att hitta fler mönster och korrelationer, så att företagets försäljning och produktsortiment kan optimeras ytterligare.

Företaget som arbetsplats och kreativ miljö har också potential att utvecklas. Exempelvis är de flesta anställda i företaget män. Om fler kvinnor anställs i strategiska produktutvecklings- och inköpsroller skulle det kunna bidra till bredare insikter, vilket i sin tur skulle kunna bidra till merförsäljning. Detta område är däremot inte analyserat i denna rapport.

Executive summary

Cyklar säljer för störst belopp men övriga produktkategorier säljer också många enheter, om än till lägre priser.

Cykelförsäljningen i länder utanför Nordamerika har det senaste året ökat mycket och utgör vid försäljningsperiodens slut en större andel av den totala försäljningen än de gjorde vid periodens början.

Försäljningen av "Touring Bikes", som startade inför sommaren 2013, har bidragit starkt till den ökade försäljningen det sista året.

Det finns grupper av kunder som gör cykelköp i månadsskiftet, cykliskt var tredje månad. Det finns tre grupper vars köpmönster varvas över tre månader så att varje månadsskifte finns representerat.

Analyserna visar på att vissa cyklar har sålts till betydligt lägre pris än tillverkningskostnaden. Det kan finnas olika orsaker till ett lågt försäljningspris och i detta fall berodde det troligtvis på en utgående cykelmodell.

De produktgrupper som är vanligast förekommande i samma order som cykelköp är hjälmar, tröjor och flaskor.

Försäljningen av produktgruppen hjälmar följer i stora drag cykelförsäljningen.

Datum för muntlig presentation

2024-01-02

3. Reflektion på eget arbete

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

Mycket tid gick åt i början för att hitta saker att analysera vidare. Att göra snabb-plottning i Python hjälpte mycket för att inte arbeta vidare med något som inte möjliggjorde vidare analyser.

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

VG. Jag hoppas att mitt arbete lever upp till kraven på djupare analyser som gäller för betyget VG.

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

Börja tidigt med att lägga in i rapporten och skriva analyser. Då är det lättare att få en överblick och lättare att hitta vidare till nästa analys.