Prosjektoppgave

Kandidatnummer 66 og 73

07 June, 2022

- $\bullet \ \, \text{Koder er insperert av tidligere arbeid, forelesnings koder i SOK_1005, https://towardsdatascience.} \\ \text{com/business-analysis-in-r-part-1-advanced-filtering-and-conditional-calculations-c6b3ef7eb7a2}$
- Laster inn datasett og pakker

```
library(tidyverse)
## -- Attaching packages -----
                                        ----- tidyverse 1.3.1 --
## v ggplot2 3.3.6
                      v purrr
                                0.3.4
## v tibble 3.1.7
                      v dplyr
                              1.0.9
## v tidyr 1.2.0
                     v stringr 1.4.0
## v readr
           2.1.2
                      v forcats 0.5.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()
                    masks stats::lag()
library(lubridate)
##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:base':
##
##
      date, intersect, setdiff, union
WEEKLY_WEATHER <- read.csv("WEEKLY_WEATHER.csv")</pre>
WEEKLY_SALES_10STORES <- read.csv("WEEKLY_SALES_10STORES.csv")</pre>
county_crime <- read.csv("county_crime.csv")</pre>
county_demographic <- read.csv("county_demographic.csv")</pre>
county_employment <- read.csv("county_employment.csv")</pre>
AppWichStoreAttributes<- read.csv("AppWichStoreAttributes.csv")
```

- I Oppgave 1 har vi startet med å se på hvilke kolonner som samsvarer med de forskjellige datasettene til oppgaven. Enkelte kolonner har samme data, men forskjellige navn på kolonnen. Dette endrer vi slik at de kan slås sammen til ett større datasett.
- Dette gjør vi med Store_Num, Date, Week og Store_County sånn at de passer med de andre datasettene. Videre endrer vi på Annual_Rent_Estimate der den ene variabelen har ett komma som gjør at kolonnen blir til character. For å endre på dette bruker vi gsub som erstatter det vi ber den om for hele kolonnen. Her erstatter vi da kommaet med ingenting og gjør den om til numeric samtidig. Weather_Date endrer vi som sagt til Date slik at den samsvarer med Date i WEEKLY_SALES_10STORES. Parallelt gjør vi om begge fra character til Date. Dette gir oss mulighet til å kunne slå sammen alle datasettene til ett stort datasett.
- Når vi skal begynne å slå sammen datasettene finnes det en del metoder å gjøre dette på. merge og left_join er to metoder som vi mente fungerte best, men valgte å benytte bare left_join. Dette bestemte vi ettersom at dplyr:join funksjonene er raskere enn merge og informerer om hva de slår sammen med, samtidig var det en kode som var allsidig å jobbe med. Vi kjører koden og setter datasettene sammen ett datasett ved hjelp av County_Name. Ett datasett ved hjelp av Weather_Station og Date. Tilslutt slår vi sammen disse to datasetten med County_Name Dette gir et stort datasett på 79459 observasjoner som tilsvarer det største datasettet WEEKLY_SALES_10STORES og vi får 72 variabler. Dette datasettet har vi gitt navnet Data.

```
AppWichStoreAttributes <- AppWichStoreAttributes %>%
  rename(Store_num = Store_Num,
         Weather_Station = Store_Weather_Station) %>%
  mutate(Annual_Rent_Estimate = as.numeric(gsub(",", "", Annual_Rent_Estimate)))
WEEKLY_WEATHER <- WEEKLY_WEATHER %>%
  mutate(Weather Date=dmy(Weather Date)) %>%
  rename(Week = Weather_Week,
         Date = Weather Date)
WEEKLY_SALES_10STORES <- WEEKLY_SALES_10STORES %>%
  mutate(Date=mdy(Date))
DF1 <- left_join(county_crime, county_demographic, by = 'County_Name') %%
  left_join(county_employment)
## Joining, by = "County_Name"
DF2 <- left join(AppWichStoreAttributes, WEEKLY SALES 10STORES, by = 'Store num') %>%
  rename(County_Name = Store_County) %>%
  left_join(WEEKLY_WEATHER)
## Joining, by = c("Weather_Station", "Date")
Data <- left_join(DF2, DF1)
## Joining, by = "County_Name"
```

• I oppgave to skal vi benytte datasettet til en ukentlig salgsrapport. En ukentlig salgsrapport skal samle informasjon om salg- og inntekts fremgang for bedriften. Vi tar datasettet og lager to nye variabler. En for ukentlig og en for månedlig estimert leiekostnad for å ha til videre analyse. Vi vil ta for oss ukentlig omsetning, kostnader og profitt for utsalgsstedet.

```
Data = Data %>%
  mutate(Annual_Rent_Estimate = as.numeric(Annual_Rent_Estimate),
          "Weekly rent" = Annual_Rent_Estimate/52) %>%
  mutate(Annual_Rent_Estimate = as.numeric(Annual_Rent_Estimate),
          "Monthly rent" = Annual_Rent_Estimate/12)
```

• Vi har laget en funskjon slik at en kan velge hvilket utsalgssted de vil analysere og hvilken uke de vil ha. Da får vi nye datasett for hver uke som vi vil analysere.

```
Utsalgssted <- function(data, x, y){
    {{data}} %>%
    filter(Store_Name == {{x}}, Week == {{y}})
}

PCS_17 <- Utsalgssted(Data, "Power City StripMall", 30)

PCS_18 <- Utsalgssted(Data, "Power City StripMall", 31)</pre>
```

- Utsalgstedet vi har valgt å analysere er Power City StripMall. Vi skal se på uke 31 og sammenligne med uken før, uke 30.
- Vi lager en funskjon som summerer alle variabler sammen og får totalte verdier for gitt uke. Dette gjør vi for å ha en generell funksjon som kan brukes til andre analyser.

```
TotalSUM <- function(a){
    return(x <- (a) %>%
        summarise('Total units sold' = sum(Sold), 'Total revenue' = sum(Sales),
        'Total expenses' = sum(Cost),'Total Cost %' = sum(Cost_Percent),
        'Total profit' = sum(Profit),'Total Margin' = sum(Margin),
        'Weekly average rent' = `Weekly rent`[1],
        'Week' = round(`Week`[1])))
}
```

• Og vi får nye datasett som inneholder summen av beløpene vi vil analysere

Table 1: Totale resultater på ukebasis

Total units sold	Total revenue	Total expenses	Total Cost %	Total profit	Total Margin	Weekly average rent	Week
$4571 \\ 4257$	$13309.47 \\ 12438.63$	$4120.88 \\ 3767.93$	35.06256 33.64654	9188.59 8670.70	0.9996 0.9997	$1038.462 \\ 1038.462$	30 31

- Ut fra tabell 1 kan vi se på at det har vært en endring på -314 enheter solgt på en uke som er en prosentvis endring på -6.8694%.
- Totale inntekter har hatt en endring på -870.84 dollar som utgjør en prosentvis endring på -6.543%.
- Totale kostnader eksludert leiekostnader hadde en endring på -352.95 dollar som utgjør en prosentvis endring på -8.5649%.
- Total kostnad i prosent utgjor en marginal endring på -1.416029
- Total profit hadde en enring på -517.89 dollar som utgjør en prosentiv endring på -5.6362%.

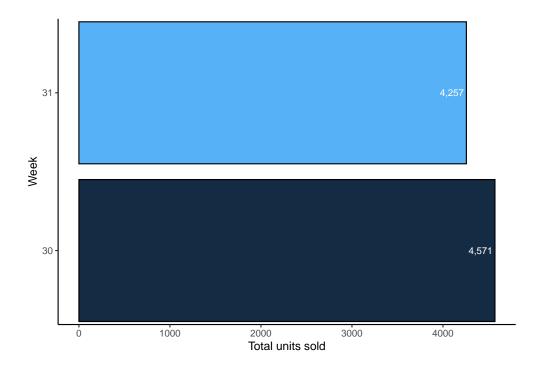


Figure 1: Sammenligning mellom totale enheter solgt per uke

• I Figur 1 viser vi grafisk hvor mye som selges totalt for Power City StripMall. Vi ser at butikken har en nedgang i salg for denne uken sammenlignet med forrige uke, som kan være påvirket av forskjellige faktorer. Det er viktig å se på hvor mye som selges hver uke sammenlignet med tidligere uker om det kommer en stor nedgang i salg må kanskje butikken gjøre endringer for å opprettholde salget. For Power City StripMall vil ikke denne nedgangen være bekymringsverdig da det er en liten nedgang fra forrige uke, men om trenden fortsetter må de gå dypere inn å se hva som forårsaker dette.

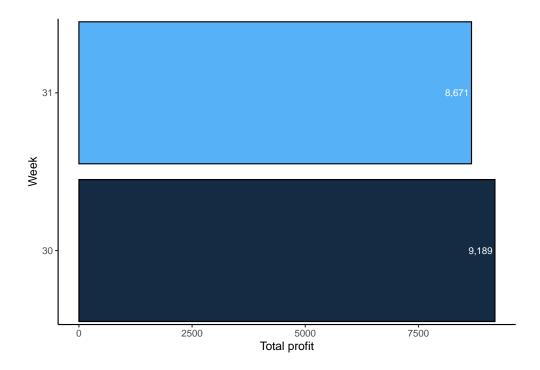


Figure 2: Sammenligning mellom total profit per uke

• I Figur 2 ser vi grafisk på forskjellen mellom uke 30 og 31 av total profitt. Fra figur 1 ser vi at salget har gått ned og derfor er det ikke overraskende at Power City StripMall har en lavere profitt denne uken heller. Det er viktig å presisere at vi ser her på Netto profitt, da kostnader som leie og lønnskostnader ikke er trekt fra.

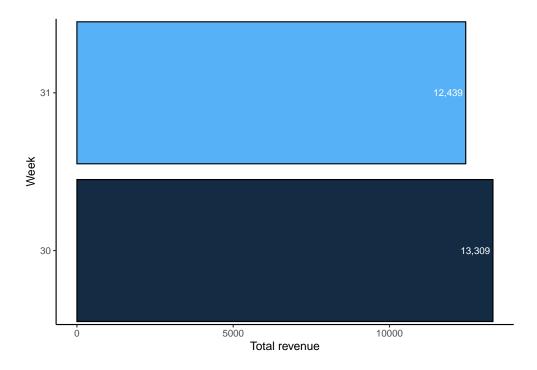


Figure 3: Sammenligning mellom total inntekter per uke

• Også i Figur 3 ser vi på den totale inntekten til butikken. Her også ser vi at det korrelerer med salget. Vi tar med alle disse grafene for å vise at salget gjenspeiler seg i gjennom alle ledd, og det er av salg en bedrift tjener profitt. Derfor er det viktig å passe på at salget opprettholdes og om det går ned kan enkle tiltak som rabatt kampanjer tiltrekke flere folk. Om det ikke fungerer må konsernledelsen gå dypere inn i butikken og se på andre faktorer som kostnader og gjøre en evaluering.

Table 2: Månedlige resultater i Juli per butikk

	Store Location	County	Monthly revenue	Monthly expenses	Monthly profit	Monthly rent	Monthly result after rent	Percent profit of franchise profit
Lake City StripMall	Strip Mall	Lake County	56690.83	17371.15	39319.68	3450.000	35869.68	10.955%
Littletown StripMall	Strip Mall	Appleton County	55113.25	16773.83	38339.42	3000.000	35339.42	10.682%
North Town BigBox	Big Box	Farm County	43750.93	13360.82	30390.11	1725.000	28665.11	8.467%
North Town StripMall	Strip Mall	Farm County	67797.28	21393.62	46403.66	3150.000	43253.66	12.928%
Power City FreeStand	Free Standing	Power County	87798.45	26805.79	60992.66	5250.000	55742.66	16.993%
Power City StripMall	Strip Mall	Power County	65937.23	20260.39	45676.84	4500.000	41176.84	12.726%
Rail City BigBox	Big Box	Rail County	35742.63	10993.31	24749.32	1458.333	23290.99	6.895%
River City StripMall	Strip Mall	River County	38301.23	12394.58	25906.65	3750.000	22156.65	7.218%
University Town BigBox	Big Box	River County	30613.16	9239.37	21373.79	1250.000	20123.79	5.955%
West Power StripMall	Strip Mall	Power County	37163.66	11384.04	25779.62	4050.000	21729.62	7.182%

```
MånedligData <- Månedsdata(Data,7) %>%
    arrange(.$`Store Name`)

## 'summarise()' has grouped output by 'Store_Name', 'Store_Location'. You can
## override using the '.groups' argument.

rownames(MånedligData) <- c(MånedligData$`Store Name`)

knitr::kable((MånedligData %>% select(-c(`Store Name`))),
```

caption = "Månedlige resultater i Juli per butikk",

kableExtra::kable_styling(font_size = 10, latex_options = "scale_down")

align = "lccccccccccc") %>%

- I tabell 2 har vi sortert ut resultatene for Juli måned for hver enkelt butikk i franchisen. Her ser vi at det er Power city FreeStand som har størst månedlig inntekt, kostnad og profitt. Allikevel er det Littletown StripMall som har høyest samlet profittmargin av butikkene i franchisen. Dette skyldes av at kostnadene er lavere i forhold til salgs inntektene til denne butikken og de genererer mer profitt i forhold til de andre butikkene med tanke på forholdet mellom salg og kostnad.
- For Franchisen sin del er det interessant å se på at Big Box butikkene genererer vesentlig mindre profitt i prosent av franchisens totale profitt for juli måned. Sammenligner vi River City StripMall og University Town BigBox ligger disse i samme County, samtidig som University Town BigBox bare har 3 andre konkurrenter nær seg i forhold til River City StripMall som har 38. Allikevel har den nesten 1,5% mindre i andel profitt til franchisen. Det vi ser er at alle Big Box butikkene ligger ikke nær en skole, som gjør at de mister kunder som familier med skolebarn, og at det kan ligge litt i utkanten, som gjør det vanskeligere å besøke. Alle Big Box butikkene har under 9% som en franchiseleder burde ta hensyn til, om det finnes bedre muligheter.
- Det å se på hvor mye hver butikk bidrar med profitt til franchisen hjelper konsernledelsen med å se hvor og hvilken butikk som trenger endringer og hvilken som går veldig bra. I tabellen kan vi se at begge butikkene i River county gir relativt lav profitt for franchisen samlet sett i forhold til de andre county's som har to butikker. Siden begge butikkene har lav profitt andel og University Town BigBox har den laveste profitten av alle butikkene, vil vi anbefale konsernledelsen sette inn tiltak for å øke profitten, men siden den går i profitt vil det ikke være nødvendig å fjerne den.

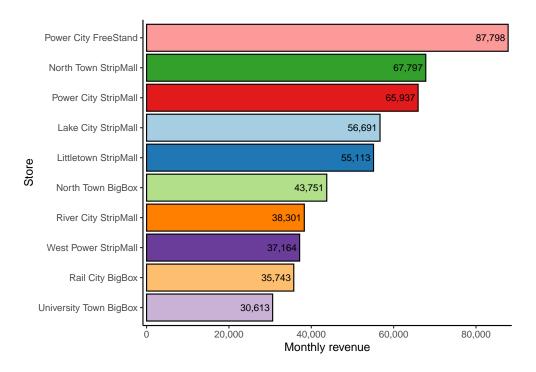


Figure 4: Totale inntekter i Juli måned for hver butikk

• I figur 4 ser vi på den månedlige inntekten til hver enkelt butikk i franchisen i stigene rekkefølge. Som Nevnt tidligere ser vi at Power city FreeStand har en mye høyere inntekt en de andre butikkene og at det er mer jevnt nedover på hvor mye hver enkelt tjener. Her får konsernledelsen et innblikk i hvor mye en måned tjener for hver enkelt butikk og hvilke butikker som har størst inntjening.

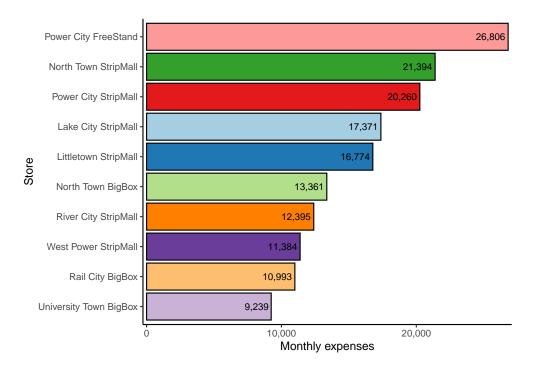


Figure 5: Total kostnad i Juli måned for hver butikk

• Vi mener det også er viktig å ha kontroll på hvor mye hver enkelt butikk har i kostnader for måneden. Vi ser at at figur 4 og figur 5 korrelerer med at de butikkene som tjener mest er også de med størst kostnader. dette er fordi de som tjener mest her mest salg og dermed trenger mer varer, og kostnader i dette tilfellet er kostnader for produktene de selger. Om man ser nøye på grafene kan vi se at det er noen av butikkene som er litt mer kostnadseffektive en andre i forhold til avstanden på inntekt og dermed kan generere litt mer profitt.

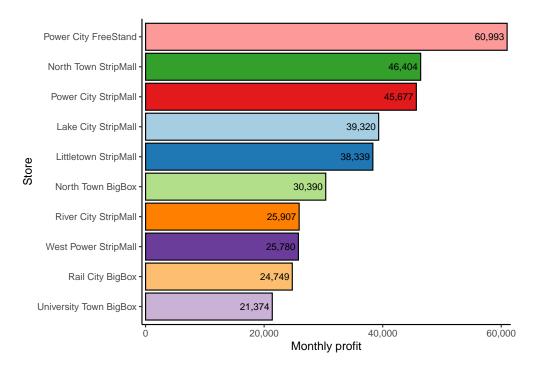


Figure 6: Total profit i Juli måned for hver butikk

• Og som nevnt med profitt ovenfor ser vi her i figur 6 at sånn som North Tower StripMall og West Power StripMall kommer nærmere den butikken med mer inntekt i profitt, da dette skyldes lavere kostnader. For konsernledelsen er det interessant om de vil se nærmere på hvorfor disse butikkene klarer å være mer kostnadsbesparende enn de andre butikkene.

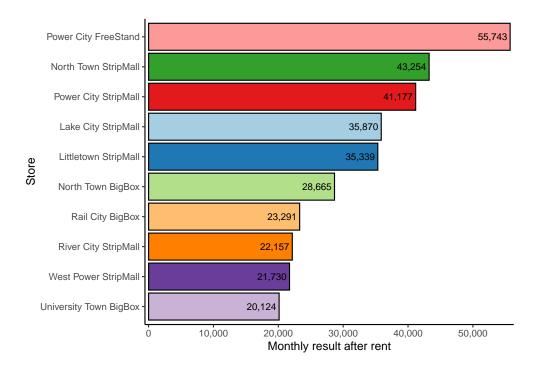


Figure 7: Total profit etter leie i Juli måned for hver butikk

• Ved å trekke fra leiekostnadene for hver butikk av profitten kan man få en slags brutto profitt, det skal nevnes at vi ikke har med lønnskostnader som gjør at det ikke blir helt nøyaktig bruttoprofitt. Her ser vi at det er noen av butikkene som har så høye leiekostnader at rekkefølgen på hvem som har meest profitt og hvem som har mest profitt etter leiekostnader endres. West Power har så høye leiekostnader at de havner nest nederst mens Rail City klatrer to plasser. for en franchise er det viktig å ha lave kostnader og leiekostnader er ofte en av de største kostnadene. Derfor mener vi det er viktig å se på hvor mye denne kostnaden tar av total profitt av en hel måned. og vi ser at de butikkene med høy profitt har også relativt høye leiekostnader som kan komme av at de har en god lokasjon de må betale mer for, men i gjengjeld får de intjening som gir høy profitt av dette også.

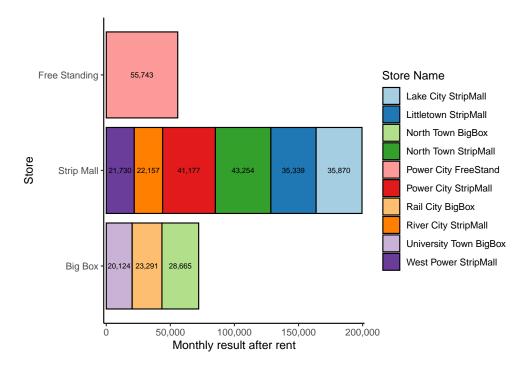


Figure 8: Månedlig salg per type butikk i Juli måned

• Her kan vi se hvor mye de forskjellige butikkene har i profitt etter leiekostnad for hvilken type butikk de er. Som nevnt tidligere er Big Box en type butikk som gir minst profitt generelt sett, men vet er to StripMall's som har på lik linje profitt som Big Box etter leiekostnad. Dette er da henholdsvis West Power StripMall som ikke ligger nær noe skole og går gantagelig glipp av mange kunder på dette aspektet, også er det River City StripMall som har veldig få innbyggere og masse konkurranse rundt seg. FreeStand er den som tjener mest, men etter leiekostnader synker profitten masse ettersom leien der er høy.

```
årTall<- function(data){</pre>
  {{data}} %>%
  group_by(Store_Name, Store_Location, County_Name, Annual_Rent_Estimate,
           County_Total_Census_Pop)%>%
  summarise("Revenue" = sum(Sales), "Profit" = sum(Profit),
            "Expenses" = sum(Cost))%>%
  select("Store Name" = Store_Name, "Store Location" = Store_Location,
         "County" = County_Name,
         "Revenue", "Expenses", "Profit",
         "Annual Rent" = Annual_Rent_Estimate,
         County_Total_Census_Pop)%>%
  as.data.frame() %>%
   mutate(
           "Result after rent" = Profit - `Annual Rent`,
           "Dollar profit per 1000 Citizen" =
             (`Result after rent`*1000)/County_Total_Census_Pop,
           'Percent profit of franchise profit' =
           scales::percent((`Profit`/sum(`Profit`)))) %>%
    select(-c(County_Total_Census_Pop))
}
årStatistikk <- function(data){</pre>
  {{data}} %>%
  group_by(Store_Name,
           County_Total_Crime_Rate, Store_Near_School,County_Total_Census_Pop,
           Store_Competition_Otherfood, Store_Competition_Fastfood,
           Store_Drive_Through, County_Unemployment_Rate, Store_Traveller_Clients)%%
    summarise() %>%
  select("Store Name" = Store Name,
         "County Population" = County_Total_Census_Pop,
         "Unemployment Rate" = County Unemployment Rate,
         "Total Crime Rate" = County_Total_Crime_Rate,
         "Drive Through?" = Store_Drive_Through,
         "Near School?" = Store_Near_School,
         "Travelling Clients?" = Store_Traveller_Clients,
         Store_Competition_Otherfood, Store_Competition_Fastfood) %>%
  as.data.frame() %>%
   mutate("Total competitors" = Store_Competition_Otherfood + Store_Competition_Fastfood) %%
        select(-c(Store_Competition_Otherfood, Store_Competition_Fastfood))
årligTall <- årTall(Data) %>%
  arrange(.$`Store Name`)
## 'summarise()' has grouped output by 'Store_Name', 'Store_Location',
## 'County_Name', 'Annual_Rent_Estimate'. You can override using the '.groups'
## argument.
```

```
årligStat <- årStatistikk(Data) %>%
 arrange(.$`Store Name`)
## 'summarise()' has grouped output by 'Store_Name', 'County_Total_Crime_Rate',
## 'Store_Near_School', 'County_Total_Census_Pop', 'Store_Competition_Otherfood',
## 'Store_Competition_Fastfood', 'Store_Drive_Through',
## 'County_Unemployment_Rate'. You can override using the '.groups' argument.
Weather <- left_join(AppWichStoreAttributes, WEEKLY_WEATHER, by = 'Weather_Station')%>%
  rename(County_Name = Store_County) %>%
  select("Store Name" = Store_Name, "Bad Weather Days" = Weather_Bad_Weather_days) %>%
  group_by(`Store Name`) %>%
  summarise("Bad Weather Days" = sum(`Bad Weather Days`))
årligStat <- left_join(årligStat, Weather)</pre>
## Joining, by = "Store Name"
rownames(arligTall) <- c(arligTall$`Store Name`)</pre>
rownames(arligStat) <- c(arligStat$`Store Name`)</pre>
årligTall <- årligTall %>%
 select(-c(`Store Name`))
årligStat <- årligStat %>%
 select(-c(`Store Name`))
```

Table 3: Årlig resultat

	Store Location	County	Revenue	Expenses	Profit	Annual Rent	Result after rent	Dollar profit per 1000 Citizen	Percent profit of franchise profit
Lake City StripMall	Strip Mall	Lake County	543373.6	169291.94	374081.6	41400	332681.6	3627.025	11.11%
Littletown StripMall	Strip Mall	Appleton County	522831.1	160356.71	362474.4	36000	326474.4	1321.882	10.76%
North Town BigBox	Big Box	Farm County	414275.0	131007.88	283267.1	20700	262567.1	3417.952	8.41%
North Town StripMall	Strip Mall	Farm County	628040.1	199575.96	428464.1	37800	390664.1	5085.448	12.72%
Power City FreeStand	Free Standing	Power County	870859.6	276774.03	594085.5	63000	531085.5	2911.685	17.64%
Power City StripMall	Strip Mall	Power County	581628.1	186421.95	395206.2	54000	341206.2	1870.668	11.73%
Rail City BigBox	Big Box	Rail County	352152.2	110225.89	241926.4	17500	224426.4	2614.321	7.18%
River City StripMall	Strip Mall	River County	336361.2	107229.24	229131.9	45000	184131.9	3099.655	6.80%
University Town BigBox	Big Box	River County	302451.7	92232.74	210219.0	15000	195219.0	3286.293	6.24%
West Power StripMall	Strip Mall	Power County	365501.8	116420.07	249081.7	48600	200481.7	1099.144	7.40%

Table 4: Statistikk for året

	Ct Dl-ti	II D-+-	Total Crime Rate	D.: Tl	M C-L19	T11:	T-4-1	Bad Weather Days
	County Population	Unemployment Rate		Drive Through?		Travelling Clients?	Total competitors	Bad Weather Days
Lake City StripMall	91723	9.5	86.3	Yes	Yes	Yes	11	118
Littletown StripMall	246977	10.6	108.8	No	No	Yes	25	106
North Town BigBox	76820	9.1	NA	No	No	No	24	122
North Town StripMall	76820	9.1	NA	No	Yes	Yes	22	122
Power City FreeStand	182398	9.3	68.8	Yes	Yes	No	35	115
Power City StripMall	182398	9.3	68.8	No	No	No	11	115
Rail City BigBox	85845	9.7	63.3	No	No	No	30	106
River City StripMall	59404	7.5	49.5	No	Yes	No	38	146
University Town BigBox	59404	7.5	49.5	No	No	No	3	146
West Power StripMall	182398	9.3	68.8	No	No	No	10	115

- Når vi ser på dataene for å se etter nytt utsalgssted for konsernledelsen har vi satt sammen to tabeller, tabell 3 og tabell 4, som kan gi oss innblikk i hvilke steder som samlet sett vil være positivt for konsernet å åpne ny butikk på. Det første vi så på var hvilket county som hadde høyest dollar profit per 1000 innbygger og i tabellen ser vi at Farm county ligger på topp og tredjeplass med to butikker, samtidig generere North Town StripMall nest høyest profitt av den totale profitten til franchisen. Videre har vi sett på arbeidsledighets raten for alle stedene, og spesielt til Farm county. Der kommer River county opp på topp, men Farm county har nest lavest arbeidsledighets rate. Dette ser vi på som positivt med tanke på at det er et county som har innbyggere med inntekt og dette er viktig om man ser på et nytt sted å starte opp en butikk. Hadde det vært høy arbeidsledighets rate, er det mindre av befolkningen som er villig til å bruke penger på franchisekjeden og det vil påvirke resultatet. Samtidig er en lav arbeidsledighet et tegn på at det er få arbeidsuføre, og at det er få som ikke aktivt søker jobb. Derfor vil ikke det å kunne ansette nye arbeidere og bidra til flere arbeidsplasser være en bekymring.
- Ved å se nærmere på farm county kan vi evaluere litt nærmere hvilken type butikk de burde starte opp og om det er noen elementer som vi hjelpe få profitten til butikken og franchisen. Fra figur 7 i oppgave tre ser vi at majoriteten til butikkene i franchisen er StripMalls og at de har gjennomsnittlig høyere profitt etter leiekostnader en det BigBox butikkene har. Siden det bare er en FreeStand butikk i hele kjeden kan dette indikere at det ikke er like enkelt å finne tilgjengelige butikker av dette formatet, men siden den i Power City har så høy profitt vil vi anbefale konsernledelsen å se etter muligheten for en Freestand i Farm county. Om det ikke er tilgjengelig vil et StripMall være det beste alternativet. Om butikken burde ha en Drive-through ser vi at de to butikkene Power city og Lake City som har det har ganske høy profitt og profitt per 1000 innbygger. Og trekker vi med antall dager med dårlig vær så har Farm county en tredjedel av året med dårlig vær, da kan en drive-through hjelpe med å øke salget siden man slipper å gå ut av bilen for å få mat. Til slutt burde de legge butikken nær en skole, siden det viser positiv korrelasjon på de butikkene som er nærme skoler og profitt. Dette gir

en kundegruppe som inneholder flere barnefamilier, samt at fra dataene ser vi at Farm county har reisende kunder som gir en økning i kundegruppen til de innbyggerne som allerede er der.

• For å summere opp vil vi anbefale konsernledelsen å starte opp en ny butikk i Farm county og se etter et lokale med drive-through som er nær skole iform av enten Freestand eller StripMall.