

Prosjektoppgave

Marcus Edvardsen og Yves Sebazungu

2022-06-07

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.6      v dplyr  1.0.8
## 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(dplyr)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union

library(plotly)

##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##   last_plot

## The following object is masked from 'package:stats':
##
##   filter

## The following object is masked from 'package:graphics':
##
##   layout

library(lubridate)
library(data.table)

##
## Attaching package: 'data.table'

## The following objects are masked from 'package:lubridate':
##
```

```
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
##      yday, year

## The following objects are masked from 'package:dplyr':
##
##      between, first, last

## The following object is masked from 'package:purrr':
##
##      transpose

library(ggplot2)
library(tibble)
```

Oppgave 1

```
appwich<- read.csv("AppWichStoreAttributes.csv")
countycrime <- read.csv("county_crime.csv")
demographic<- read.csv("county_demographic.csv")
weeklysales <- read.csv("WEEKLY_SALES_10STORES.csv")
weeklyweather<- read.csv("WEEKLY_WEATHER.csv")
countyemployment <- read.csv("county_employment.csv")

appwich = appwich %>%
  rename(Store_num =Store_Num,
         Weather_Station = Store_Weather_Station)

weeklyweather <- weeklyweather %>%
  mutate(Weather_Date=dmy(Weather_Date)) %>%
  rename(Week = Weather_Week,
         Date = Weather_Date) %>%
  mutate(Week =strftime(.$Date, format = "%U")) %>%
  replace(is.na(.), 0)

weeklysales <- weeklysales %>%
  mutate(Date=mdy(Date))
```

Merger tabellene Bruker merge på df2

```
df1 <- inner_join(countycrime, demographic, countyemployment, by = 'County_Name') %>%
  replace(is.na(.), 0)
df2 <- merge(appwich, weeklysales, by = 'Store_num') %>%
  rename(County_Name = Store_County) %>%
  left_join(weeklyweather)
```

```
## Joining, by = c("Weather_Station", "Date")
```

```
data_clean <- left_join(df2,df1)
```

```
## Joining, by = "County_Name"
```

Oppgave 2

Til oppgave 2 har vi valgt å foreta en videre analyse av Lake city Stripmall som er et kjøpesenter lokalisert i byen Lake City. Grunnen til at vi valgte denne butikken er at vi ser et vidt spekter av faktorer som påvirker senteret. (Eksempler på dette er Drive Through, flerkulturelt samfunn, prisklasser)

```
lake_city_week <- data_clean %>%
  filter(Store_num == "16") %>%
  filter(Month == "12") %>%
  filter(Date >="2012-12-02", Date <="2012-12-09")
```

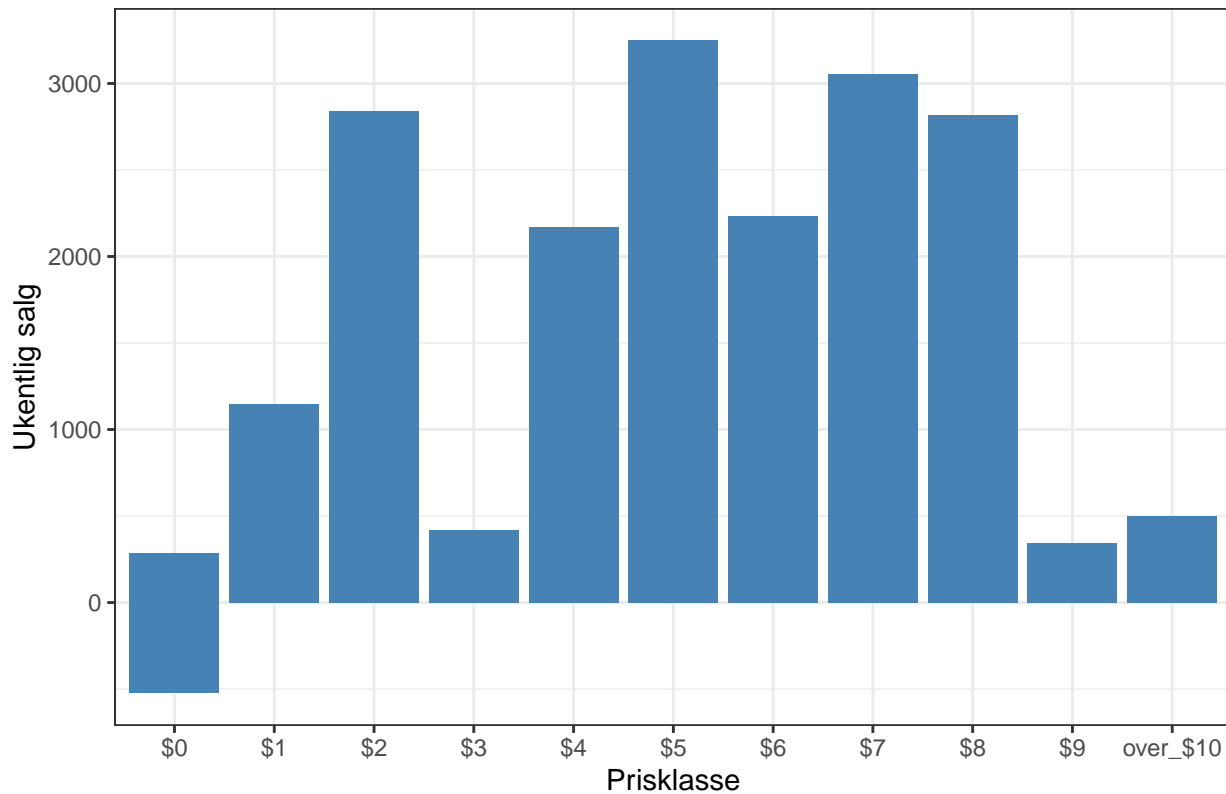
```
sg <- function(data_clean){
  return(x <- (data_clean) %>%
    group_by(price_group = ifelse(Price <= 0.49, "$0",
                                   ifelse(Price <= 1.49, "$1",
                                           ifelse(Price > 1.49 & Price <= 2.49, "$2",
                                                  ifelse(Price > 2.49 & Price <= 3.49,
                                                         ifelse(Price > 3.49 & Price <= 4.49,
                                                                ifelse(Price > 4.49 & Price <= 5.49,
                                                                     ifelse(Price > 5.49, "$6", ""))
                                                                )
                                                         )
                                                  )
                                                         )
                                                  )
                                   )
    )
    summarise(Sold, Price, Sales, Profit, Margin, Store_Name) %>%
    arrange(.$price_group))
}

lake_city <- sg(lake_city_week)
```

Grafen under er en graf som tar for seg de forskjellige prisklassene og sammenligner opp mot total solgte varer, allerede her får vi en indikasjon på hvordan konsernledelsen satser, men vi må huske på at kostnader ikke er med i betraktning. Dette betyr at selv om prisklassen 5 dollar er den som selger mest er det ikke dermed gitt at denne klassen gir best profit.

3

Ukentlig salg – Lake City Stripmall



Hva er de mest lønnsome produktene?

```
lake_city_week %>%
  group_by(Description) %>%
  summarise(Profit = sum(Profit)) %>%
  arrange(desc(Profit))
```

```
## # A tibble: 173 x 2
##   Description      Profit
##   <chr>          <dbl>
## 1 REGULAR SAVORY TURKEY    705.
## 2 REGULAR CHICKEN BACON RANCH 550.
## 3 CHIPS                   539.
## 4 MINI SAVORY TURKEY      531.
## 5 REGULAR STEAK & CHEESE   527.
## 6 REGULAR HONEY MUSTARD HAM 457.
## 7 REGULAR CHICKEN TERIYAKI 437.
## 8 REGULAR BMT             426.
## 9 MINI COLD CUT MUSHROOM   424.
## 10 REGULAR ROASTED CHICKEN 407.
## # ... with 163 more rows
```

Hva er de minst lønnsome produktene?

```
lake_city_week %>%
  group_by(Description) %>%
  summarise(Profit = sum(Profit)) %>%
  arrange(Profit)
```

```
## # A tibble: 173 x 2
##   Description      Profit
##   <chr>          <dbl>
## 1 VAL MEAL 1 MINI MTBALL CHIP -243.
## 2 FREE MINI SUB -188.
## 3 VAL MEAL 3 MINI TURKEY CHIP -37.4
## 4 BOGO MINI SUB -29.5
## 5 MISCELLANEOUS COUPON -12
## 6 FREE REGULAR -11.8
## 7 VAL MEAL #1 MINI DOUBLE CHIP -0.9
## 8 FREE COOKIE -0.5
## 9 TRANSFER FROM -0.01
## 10 MINI ITALIAN WHT BREAD 0
## # ... with 163 more rows
```

Hva selger vi mest av?

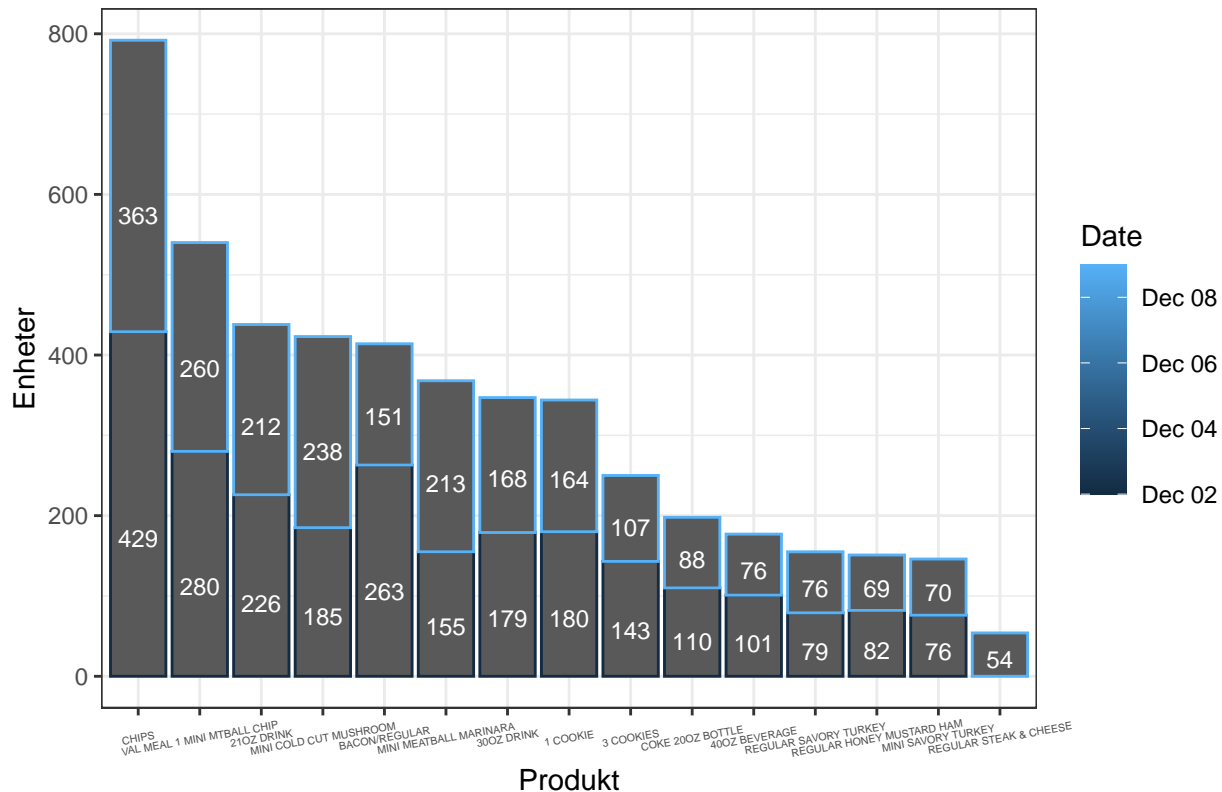
```
lake_city_week %>%
  group_by(Description) %>%
  summarise(Sold = sum(Sold)) %>%
  arrange(desc(Sold))
```

```
## # A tibble: 173 x 2
##   Description      Sold
##   <chr>          <int>
## 1 CHIPS 792
## 2 VAL MEAL 1 MINI MTBALL CHIP 540
## 3 21OZ DRINK 438
## 4 MINI COLD CUT MUSHROOM 423
## 5 BACON/REGULAR 414
## 6 MINI MEATBALL MARINARA 368
## 7 30OZ DRINK 347
## 8 1 COOKIE 344
## 9 3 COOKIES 250
## 10 COKE 20OZ BOTTLE 198
## # ... with 163 more rows
```

Graf over mest solgte produkter:

```
lake_city_week %>%
  filter(Sold >= 50) %>%
  ggplot(aes(x = reorder(Description, -Sold), y = Sold, colour = Date)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = Sold, position = position_stack(vjust = 0.4), colour = "White", size=3)) +
  labs(title = "Mest solgte varer for uken 02-09, 2012") +
  xlab("Produkt") +
  ylab("Enheter") +
  theme_bw() +
  theme(axis.text.x = element_text(vjust= 0.6, size = 4, angle = 10))
```

Mest solgte varer for uken 02–09,2012



Av grafen over kommer det tydelig frem at varene som er solgt mest av er potetgull og drikke varer. Dette er det de selger mest av og en justering av pris her vil derfor påvirke budsjette deres stort. Man kan eventuelt da sammenligne prisene hos andre butikker for potetgull og drikke, og eventuelt satse på om de vil undergå de andre butikk kjedene.

Oppgave 3

En månedlig rapport vil konsernledelsen kunne se det totale bildet av virksomhetene og derfor kunne foreta strategiske vurderinger basert på resultatene. Dette indikerer også at salg og fortjeneste totalt og per butikk vil være sentrale faktorer for konsersledelsen. Sammenligning av data på varer tvers av butikker vil også være relevant.

Starter med å filtrere både for Lake city stripmall og for alle utsalgsstedene for valgte måned.

```
lake_city_mnd <- data_clean %>%
  filter(Store_num == "16") %>%
  filter(Month == "12") %>%
  filter(Year == "2012")

mnd_data <- data_clean %>%
  filter(Year == "2012", Month == "12")
```

Utsalgsstedene og deres profitt for desember 2012

```
mnd_data %>%
  group_by(Store_num) %>%
  summarise(Profit = sum(Profit)) %>%
  arrange(desc(Profit))
```

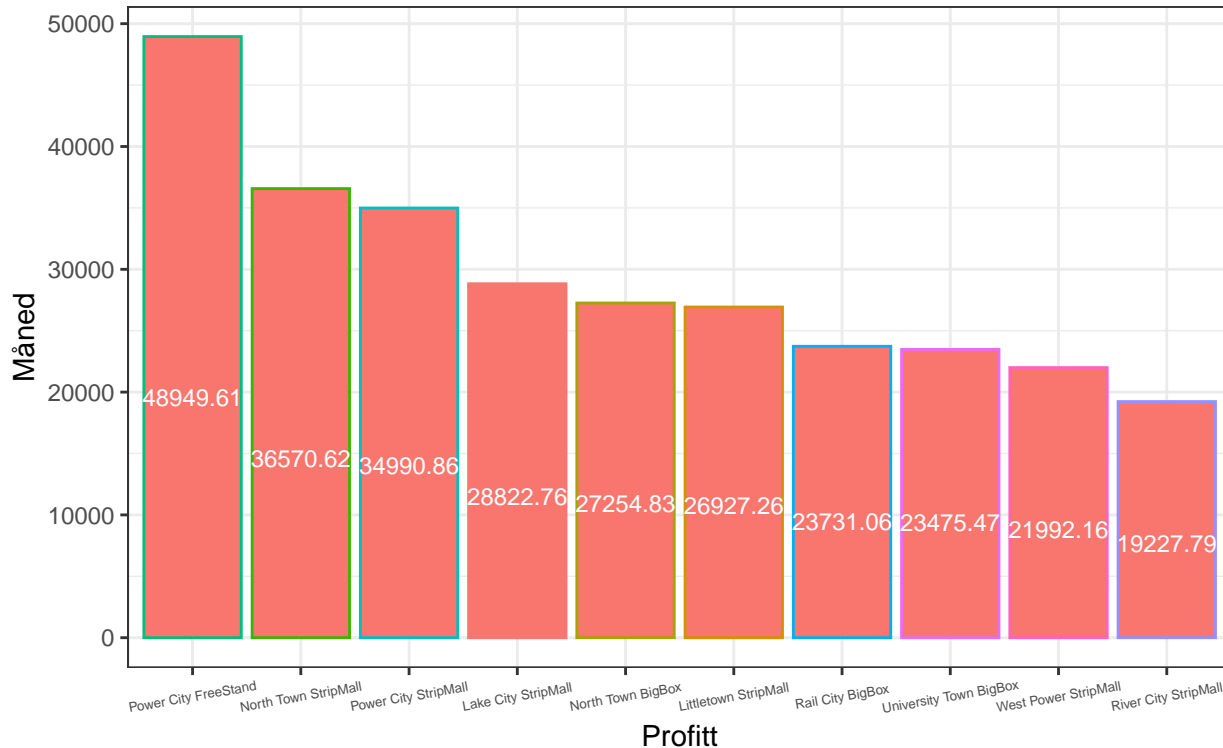
```
## # A tibble: 10 x 2
##   Store_num Profit
##   <int>   <dbl>
## 1         2 48950.
## 2        19 36571.
## 3         7 34991.
## 4        16 28823.
## 5        23 27255.
## 6         9 26927.
## 7        14 23731.
## 8        24 23475.
## 9         5 21992.
## 10       11 19228.
```

```
mnd_data %>%
  group_by(Month, Store_Name) %>%
  summarise(Mnd_profitt = sum(Profit)) %>%
  ggplot(aes(x = reorder(Store_Name, -Mnd_profitt), y = Mnd_profitt, col= Store_Name)) +
  geom_col(aes(fill= "red"), show.legend = FALSE) +
  geom_text(aes(label = Mnd_profitt), position = position_stack(vjust = 0.4), colour = "White", size=3)+
  labs(title = "Profitt for alle utsalgsstedene desember 2012",
        subtitle = "I USD") +
  xlab("Profitt") +
  ylab("Måned") +
  theme_bw() +
  theme(axis.text.x = element_text(vjust= 0.6, size = 5, angle = 10))
```

```
## `summarise()` has grouped output by 'Month'. You can override using the
## `.groups` argument.
```

Profitt for alle utsalgsstedene desember 2012

I USD



Av grafen ser vi at lake City stripmall verken gjør det gjennomsnittlig i forhold til de andre kjedene. Om tilfellet er slik at konsernledelsen nå skulle ønsket å åpne enda en restaurant ville vi heller bedt han om å bytte lokasjon og heller etterligne powercity freestand. Dette da måneds proffitten er 20 000 \$ mer.

Hvilke produkter er mest lønnsom?

```
lake_city_mnd %>%
  group_by(Description) %>%
  summarise(Profit = sum(Profit)) %>%
  arrange(desc(Profit))
```

```
## # A tibble: 208 x 2
##   Description      Profit
##   <chr>           <dbl>
## 1 REGULAR SAVORY TURKEY 1611.
## 2 REGULAR CHICKEN BACON RANCH 1286.
## 3 CHIPS                1235.
## 4 REGULAR CHICKEN TERIYAKI 1206.
## 5 REGULAR HONEY MUSTARD HAM 1194.
## 6 REGULAR STEAK & CHEESE 1154.
## 7 MINI SAVORY TURKEY 1054.
## 8 REGULAR SPICY ITALIAN 945.
## 9 MINI COLD CUT MUSHROOM 920.
## 10 REGULAR ROASTED CHICKEN 908.
## # ... with 198 more rows
```

Hvilke produkter er minst lønnsome?

```
lake_city_mnd %>%
  group_by(Description) %>%
```



```
summarise(Profit = sum(Profit)) %>%
arrange(Profit)
```

```
## # A tibble: 208 x 2
##   Description      Profit
##   <chr>           <dbl>
## 1 VAL MEAL 1 MINI MTBALL CHIP -559.
## 2 FREE MINI SUB             -281
## 3 VAL MEAL 3 MINI TURKEY CHIP  -78.4
## 4 BOGO MINI SUB             -54.5
## 5 FREE REGULAR              -44.8
## 6 MISCELLANEOUS COUPON        -40
## 7 VAL MEAL #1 MINI DOUBLE CHIP  -6.3
## 8 FREE COOKIE               -1.5
## 9 TRANSFER FROM             -0.12
## 10 MINI ITALIAN WHT BREAD         0
## # ... with 198 more rows
```

Hva selger vi mest av?

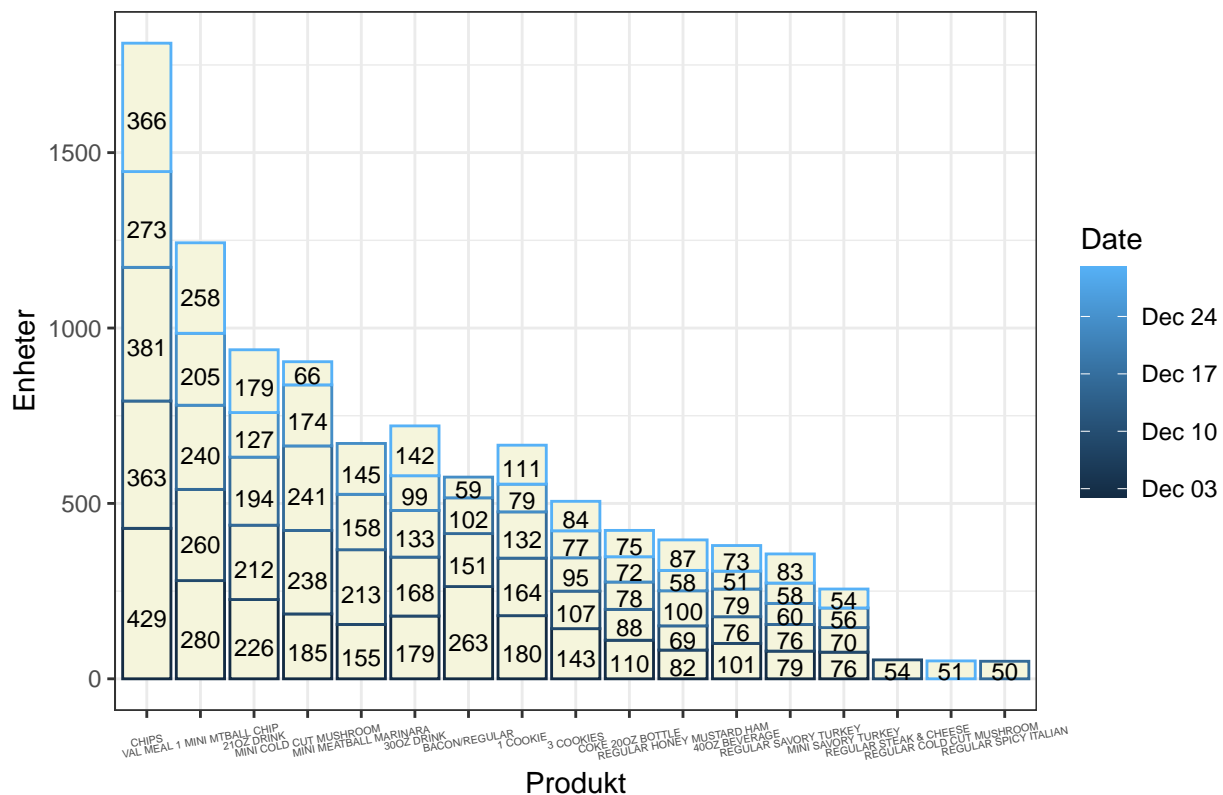
```
lake_city_mnd %>%
  group_by(Description) %>%
  summarise(Sold = sum(Sold)) %>%
  arrange(desc(Sold))
```

```
## # A tibble: 208 x 2
##   Description      Sold
##   <chr>           <int>
## 1 CHIPS             1812
## 2 VAL MEAL 1 MINI MTBALL CHIP  1243
## 3 21OZ DRINK         938
## 4 MINI COLD CUT MUSHROOM       904
## 5 30OZ DRINK         721
## 6 MINI MEATBALL MARINARA       702
## 7 1 COOKIE           666
## 8 BACON/REGULAR       621
## 9 3 COOKIES           506
## 10 COKE 20OZ BOTTLE       423
## # ... with 198 more rows
```

Graf over mest solgte produkter denne måneden:

```
lake_city_mnd %>%
  filter(Sold >= 50) %>%
  ggplot(aes(x = reorder(Description,-Sold), y = Sold, colour = Date )) +
  geom_bar(stat = "identity", fill="beige") +
  geom_text(aes(label = Sold),position = position_stack(vjust = 0.4), colour = "Black", size=3)+
  labs(title = "Mest solgte varer for uken 02-09,2012") +
  xlab("Produkt") +
  ylab("Enheter") +
  theme_bw() +
  theme(axis.text.x = element_text(vjust= 0.6, size = 4, angle = 10))
```

Mest solgte varer for uken 02–09,2012



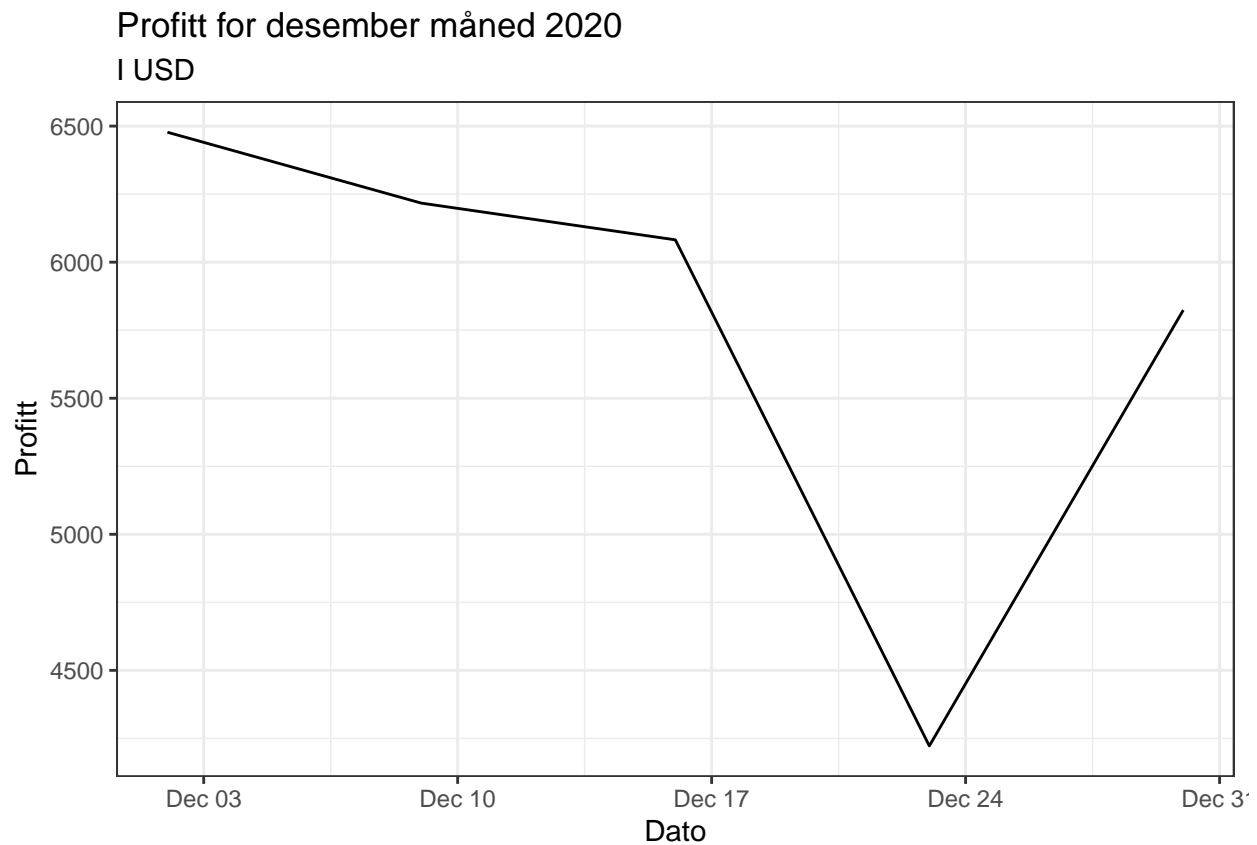
Her kan vi igjen se at det er “Chips” som dominerer salget ut måneden.

Graf for profitt gjennom måneden:

```
plotmnd <- lake_city_mnd

plotmnd <- aggregate(plotmnd$Profit, by=list(Dato=plotmnd$Date), FUN=sum)
plotmnd <- rename(plotmnd, Profitt = x)

plotmnd %>%
  ggplot(aes(x = Dato, y = Profitt)) +
  geom_line() +
  labs(title = "Profitt for desember måned 2020",
        subtitle = "I USD") +
  theme_bw()
```



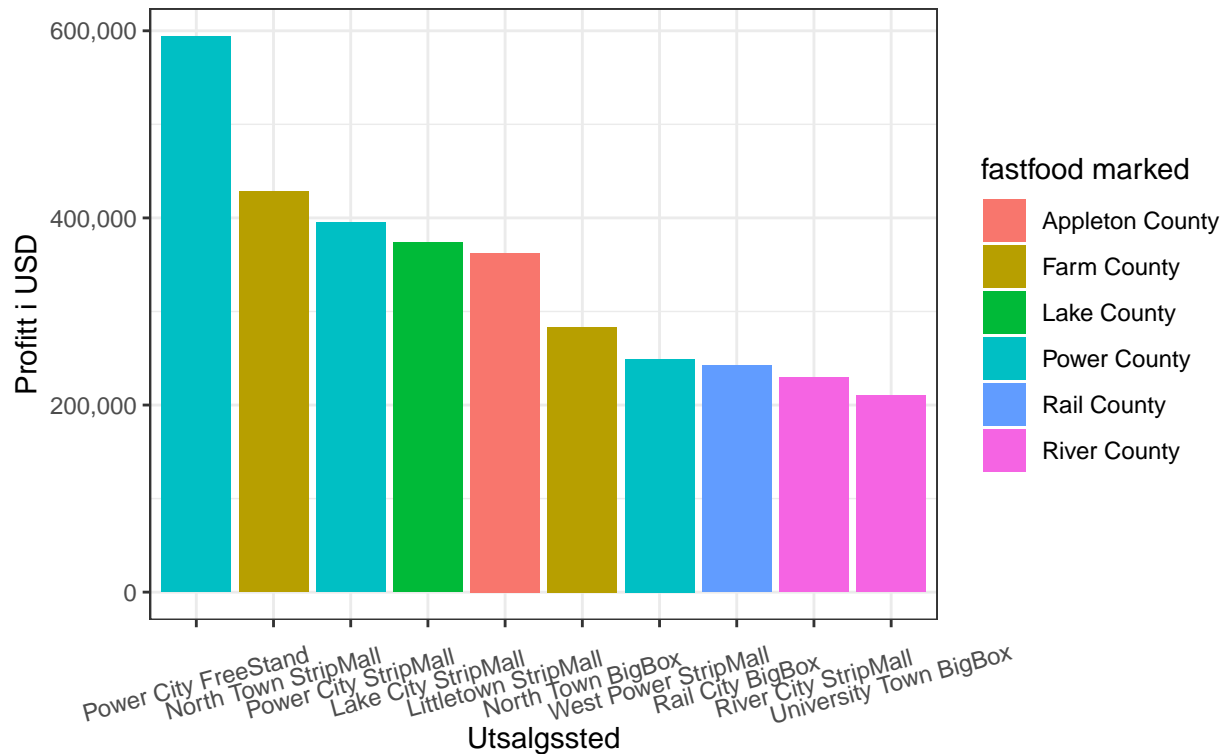
Oppgave 4

```
data_clean %>%
  select(Store_Name, County_Name, Profit) %>%
  group_by(Store_Name, County_Name) %>%
  summarise(Profit = sum(Profit)) %>%
  ggplot(aes(x=reorder(Store_Name,-Profit) , y =Profit))+
  geom_col(aes(fill = County_Name))+
  scale_y_continuous(labels = scales::comma)+
  labs(title = "Har konkurransen en effekt på lønnsomheten?",
        subtitle = 'Farger etter hvilket county de ligger i.',
        fill = 'fastfood marked' )+
  ylab("Profitt i USD")+
  xlab("Utsalgssted")+
  theme_bw()+
  theme(axis.text.x = element_text(angle=15, vjust = 0.4))
```

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

Har konkurransen en effekt på lønnsomheten?

Farger etter hvilket county de ligger i.

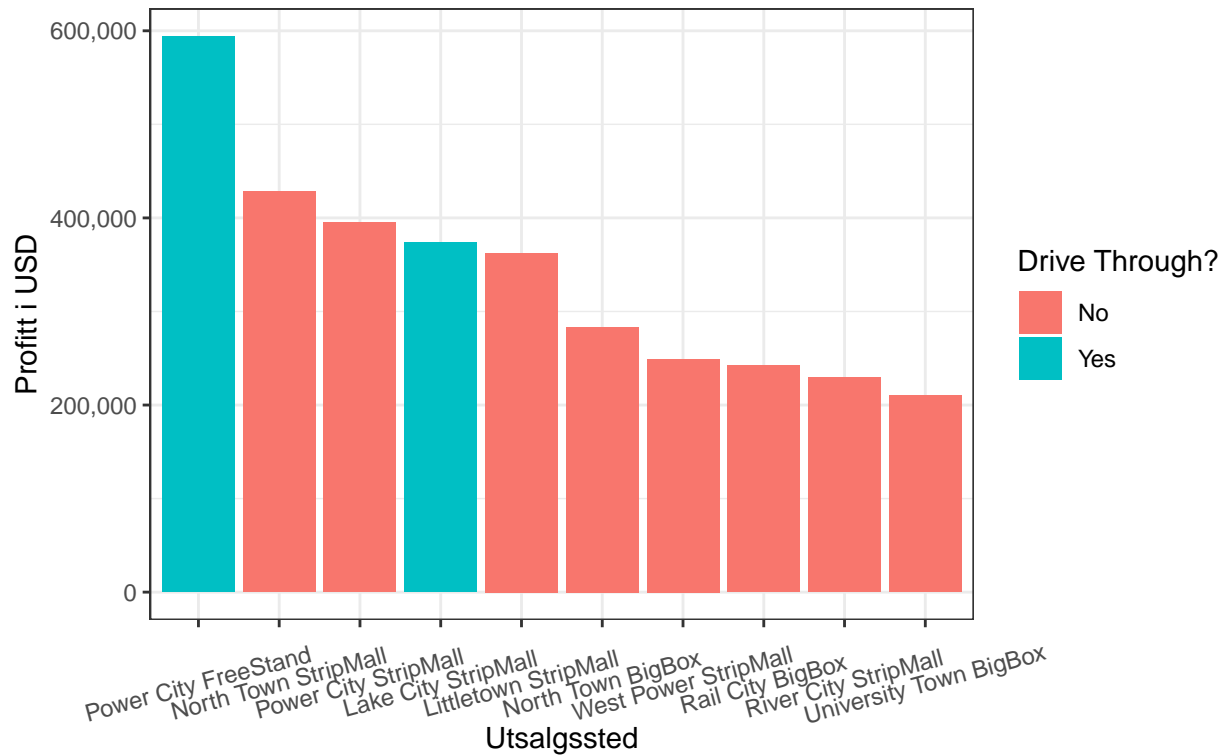


```
data_clean %>%
  select(Store_Name, Profit, Store_Drive_Through) %>%
  group_by(Store_Name, Store_Drive_Through) %>%
  summarise(Profit = sum(Profit)) %>%
  ggplot(aes(x=reorder(Store_Name,-Profit) , y =Profit))+
  geom_col(aes(fill= Store_Drive_Through )) +
  labs(title = "Har drive-through noe å si på lønnsomheten?",
        subtitle = "Burde det nye utsalgsstedet ha drive-through?",
        fill = 'Drive Through?' )+
  ylab("Profitt i USD")+
  xlab("Utsalgssted")+
  scale_y_continuous(labels = scales::comma)+
  theme_bw()+
  theme(axis.text.x = element_text(angle=15, vjust = 0.4))
```

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

Har drive-trough noe å si på lønnsomheten?

Burde det nye utsalgsstedet ha drive-trough?



Av dataanalysen gjort i denne oppgaven indikerer at det vil være bedre å etablere butikker utenfor et kjøpesenter enn inne, dette fordi at av grafen ser man at konkurransen påvirker salget i noe grad negativt. Om man tillegg klarer å opprettholde en Drive-In løsning, så ser det ut til å gi best resultat.