Assignment STAT702

Product name: BIC Round Stic Xtra Life Ballpoint Pen, Medium Point (1.0mm), Red, 12-Count Sales sku_id: 219884 Reviews asin: B00006IE7J

#1 Analysis of Sales Data

1(a) For the product (sku_id) which has been assigned to your group (see page 6), compute the total monthly sales from January 2011 – September 2013. Present your results in an appropriate plot and write 2 – 3 sentences describing your results.

Hint: This will require some "wrangling" of the variable week. To do this, format week as a date and then use the appropriate lubridate function to extract the month.

Marking Criteria

- Total monthly sales have been correctly computed and are displayed in an appropriate plot.
- Description of results/plot is correct and provides useful insights.
- Plot is constructed using ggplot2 and has appropriate titles, labels, scales etc.

```
# Load libraries
library(tidyverse)
```

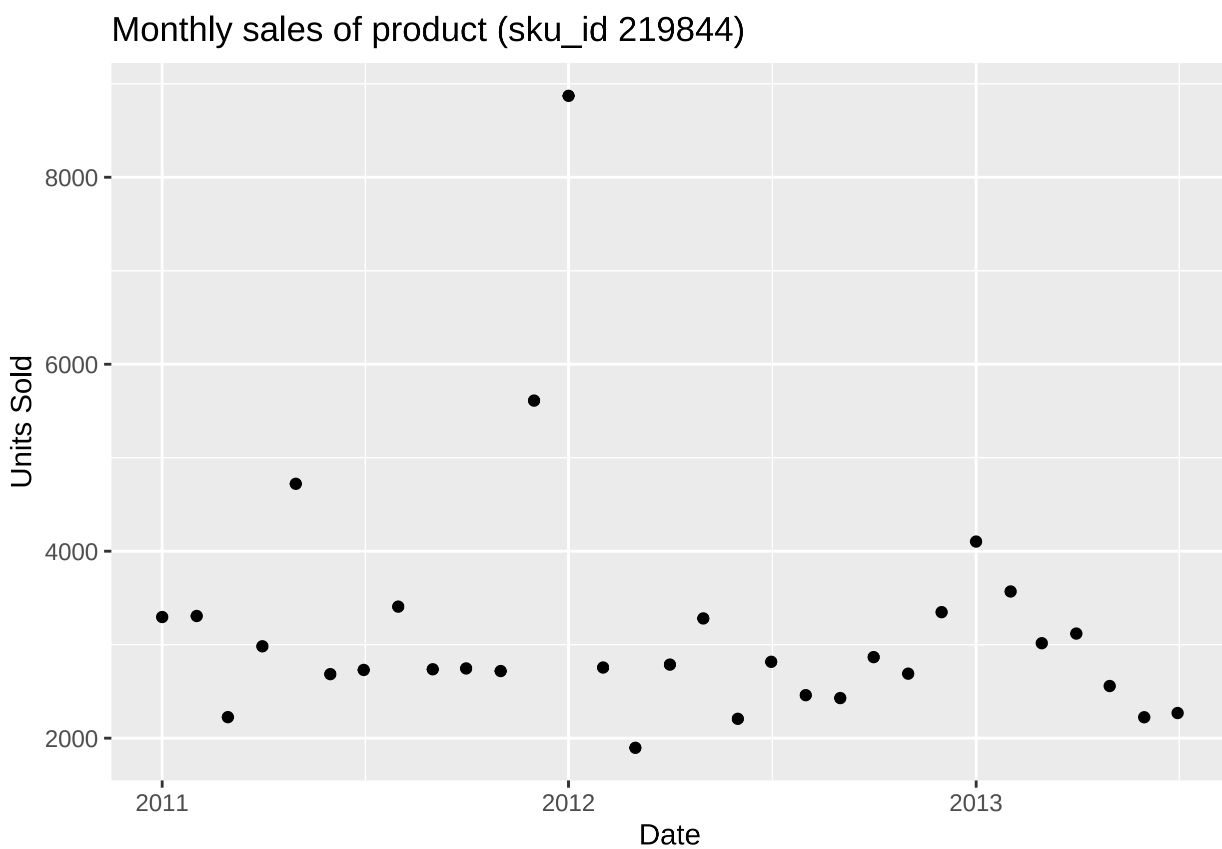
```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.6     v dplyr   1.0.4
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     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
```

```r
# Read in data and convert to tibbles
reviews_data <- read.csv("reviews_data.csv")
reviews_data <- as_tibble(reviews_data)

sales_data <- read.csv("sales_data.csv")
sales_data <- as_tibble(sales_data)
```

```r
# Create a summary table, grouped by month and year (single column with m/y)

sales_data %>%
        filter(sku_id == 219844) %>%
        mutate(week = as_date(week, format = "%d/%m/%y"),
               date = format(week, "%m/%y"),
               date = my(date)) %>%
        group_by(date) %>%
        summarise(total_units_sold = sum(units_sold))-> sales_summary


# Scatterplot of monthly sales

ggplot(sales_summary) +
        geom_point(aes(x = date, y = total_units_sold)) +
        xlab("Date") +
        ylab("Units Sold") +
        ggtitle("Monthly sales of product (sku_id 219844)")
```



```r
# # Print the summary, displaying month and year in separate columns
#
# sales_summary %>%
#         mutate(month = month(date, label = TRUE), year = (year(date))) %>%
#         select(month, year, total_units_sold) %>% view()
```

```r
# Compute basic stats
summary(sales_summary$total_units_sold)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1897    2622    2787    3175    3302    8871
```

Mean monthly sales are 3175. Outlier of 8871 sales in Jan 2012, minimum observation is 1897 in March 2012. 50% of observations lie between the values 2622 (first quartile) and 3302 (third quartile). No seasonal variation or trend identified.

1(b) The GM Sales wants to know which stores are performing well and which are not, in terms of product sales. For the product (sku_id) which has been assigned to your group, use appropriate summary statistics and plots to investigate sales performance across the stores and write 2 – 3 paragraphs summarising your findings.

Hint: You will need to decide what it means for a store to be "performing well" and how you will evaluate this using the data.

Marking criteria

- Sales performance is clearly defined.
- Written summary includes relevant and appropriate summary statistics and plots.
- Plot/s are constructed using ggplot2 and have appropriate titles, labels, scales etc.
- Descriptions of results and plots are correct and provides useful insights.

```r
# Summarise by year
sales_data %>%
  filter(sku_id == 219844) %>%
  mutate(week = as_date(week, format = "%d/%m/%y"),
         date = format(week, "%m/%y"),
         date = my(date),
         year = year(date),
         quarter = quarter(date),
         store_id = as.character(store_id))%>%
  group_by(year, store_id, date) %>%
  summarise(total_units_sold = sum(units_sold)) -> sales_summary_2
```

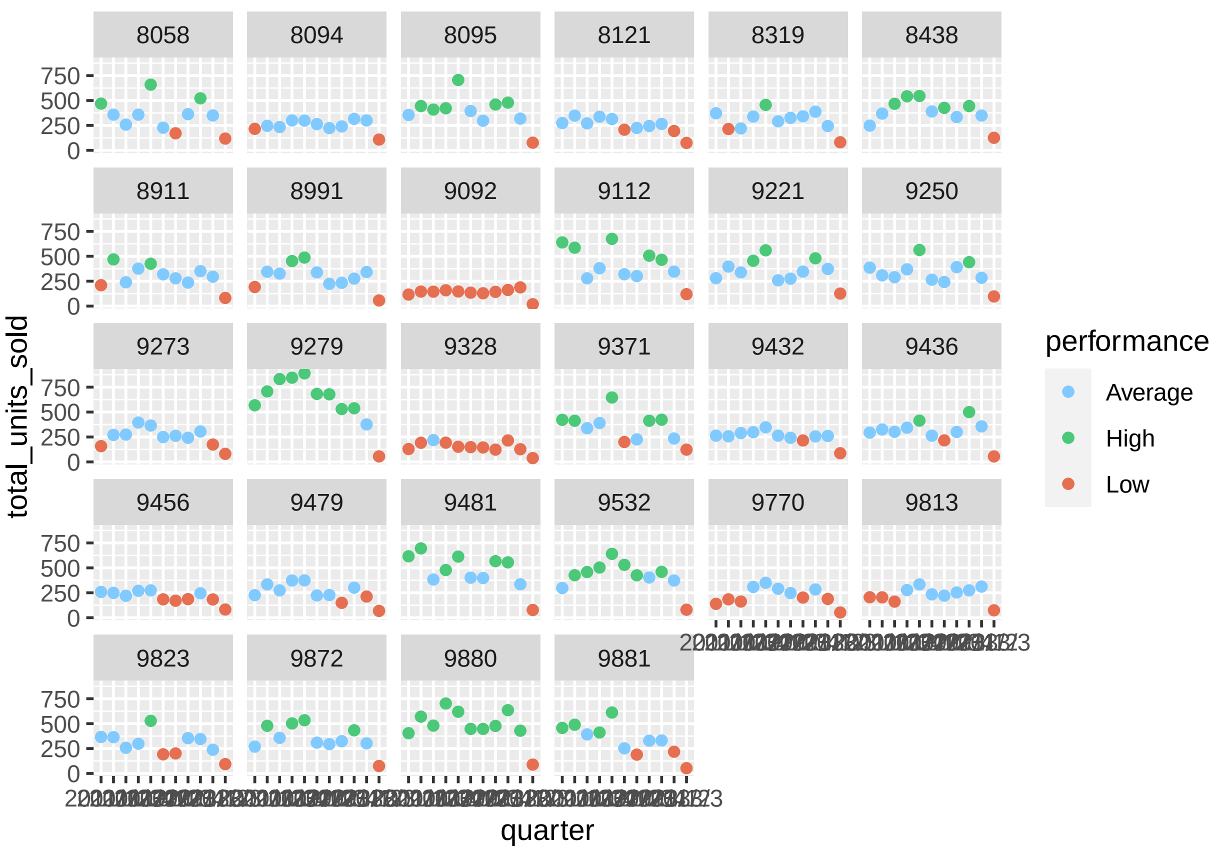
```
## `summarise()` has grouped output by 'year', 'store_id'. You can override using the `.groups` argumen
```

```r
# Summarise by quarter

sales_summary_2 %>%
  mutate(quarter = quarter(date)) %>%
  unite("quarter", year, quarter, remove = FALSE, sep = "/") %>%
  group_by(quarter, store_id) %>%
  summarise(total_units_sold = sum(total_units_sold)) -> sales_summary_q
```

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

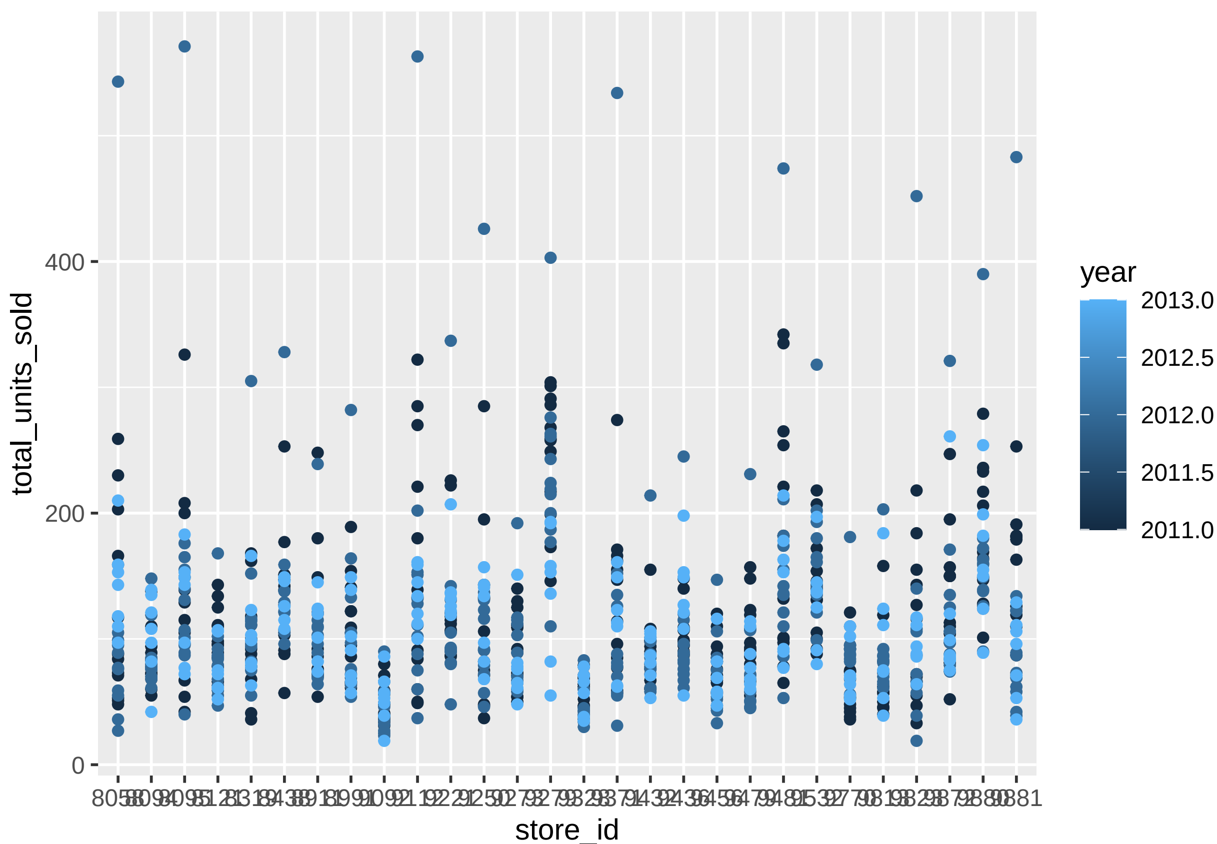
```r
# Calculate quarterly quantiles
quartiles <- summary(sales_summary_q$total_units_sold)
q1 <- quartiles[2]
q3 <- quartiles[5]
```

```r
# Quarterly plots for each store with colour based on performance
# low is less than q1
# high is greater than q3

sales_summary_q %>%
  mutate(performance = ifelse(total_units_sold < q1, "Low", (ifelse(total_units_sold > q3, "High", "Ave
  ggplot()+
  geom_point(aes(x = quarter, y = total_units_sold, colour = performance)) +
  facet_wrap(~ store_id, ncol = 6) +
  scale_colour_manual(values = c("#80CAFF", "#4BC979", "#E76F51"))
```

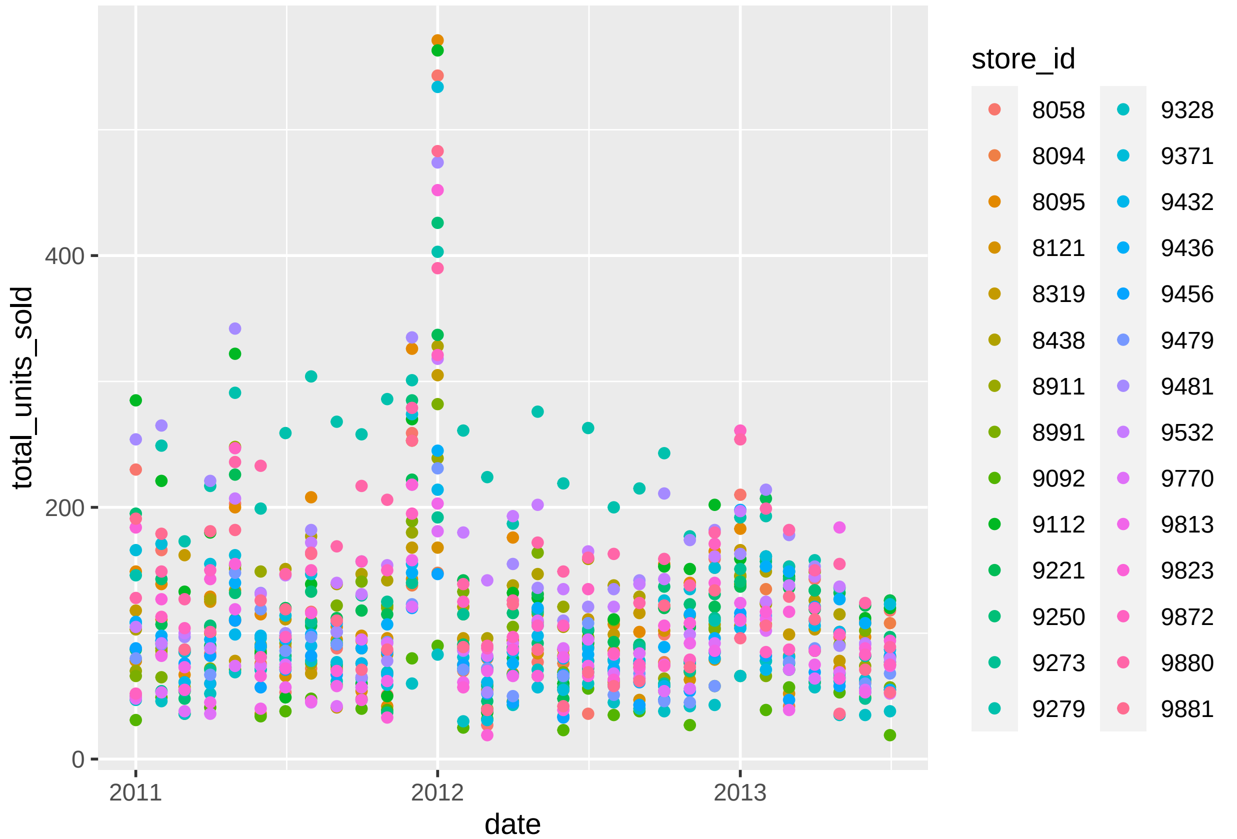


```
# Quarterly axis is not good
```

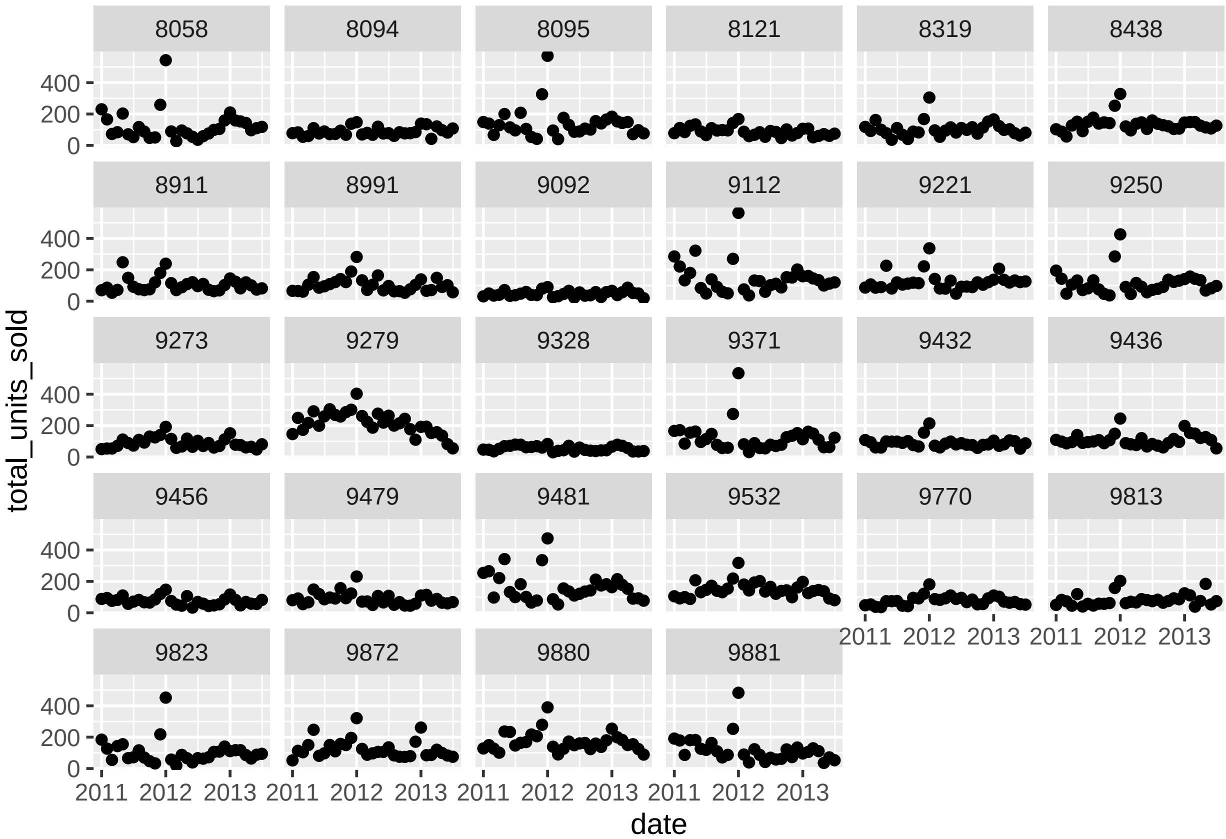
```
# Plots not used

# summarise by year
sales_summary_2 %>%
        ggplot() +
        geom_point(aes(x = store_id, y = total_units_sold, colour = year))
```



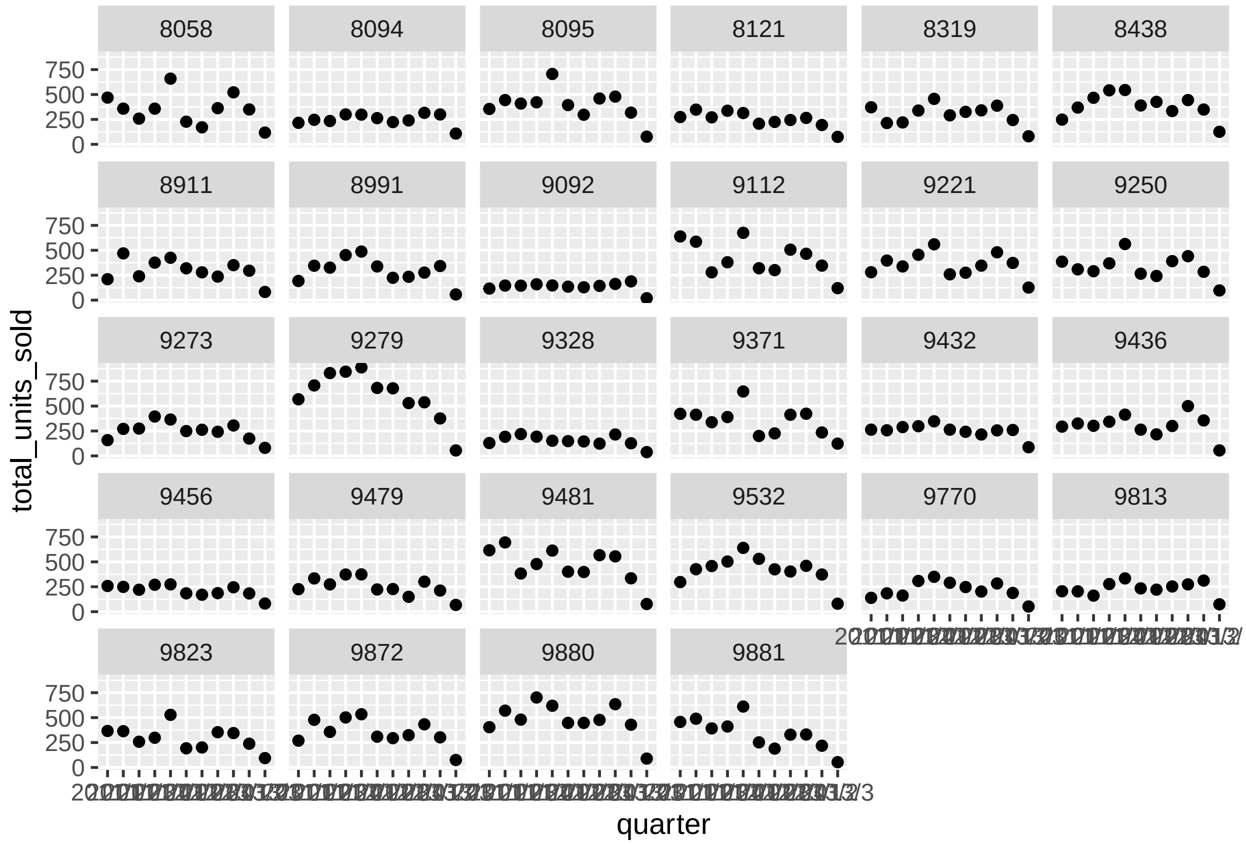

```
sales_summary_2 %>%
        ggplot()+
        geom_point(aes(x = date, y = total_units_sold, colour = store_id))
```



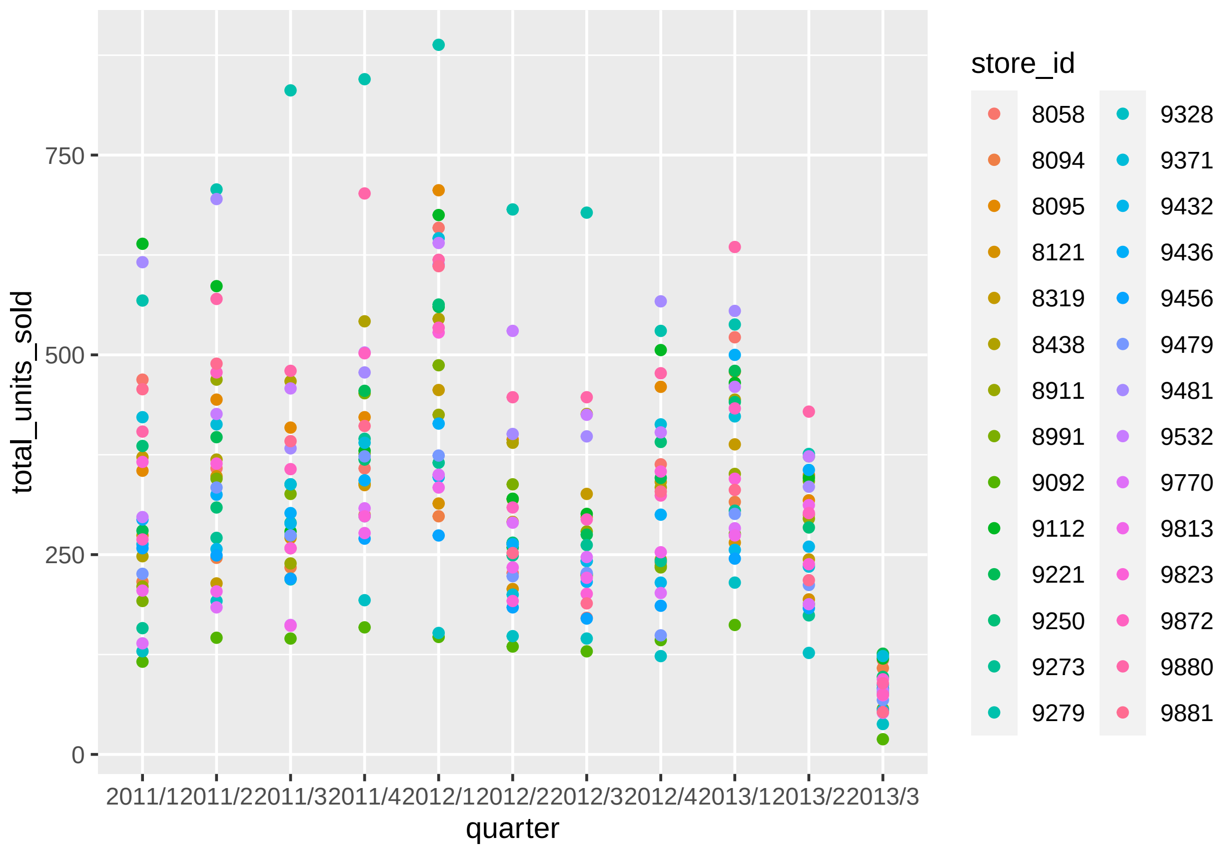

```
sales_summary_2 %>%
        ggplot()+
        geom_point(aes(x = date, y = total_units_sold)) + facet_wrap(~ store_id, ncol = 6)
```




```
# summarise by quarter

sales_summary_q %>%
        ggplot()+
        geom_point(aes(x = quarter, y = total_units_sold)) + facet_wrap(~ store_id, ncol = 6)
```




```
sales_summary_q %>%
        ggplot()+
        geom_point(aes(x = quarter, y = total_units_sold, colour = store_id))
```



Question 2

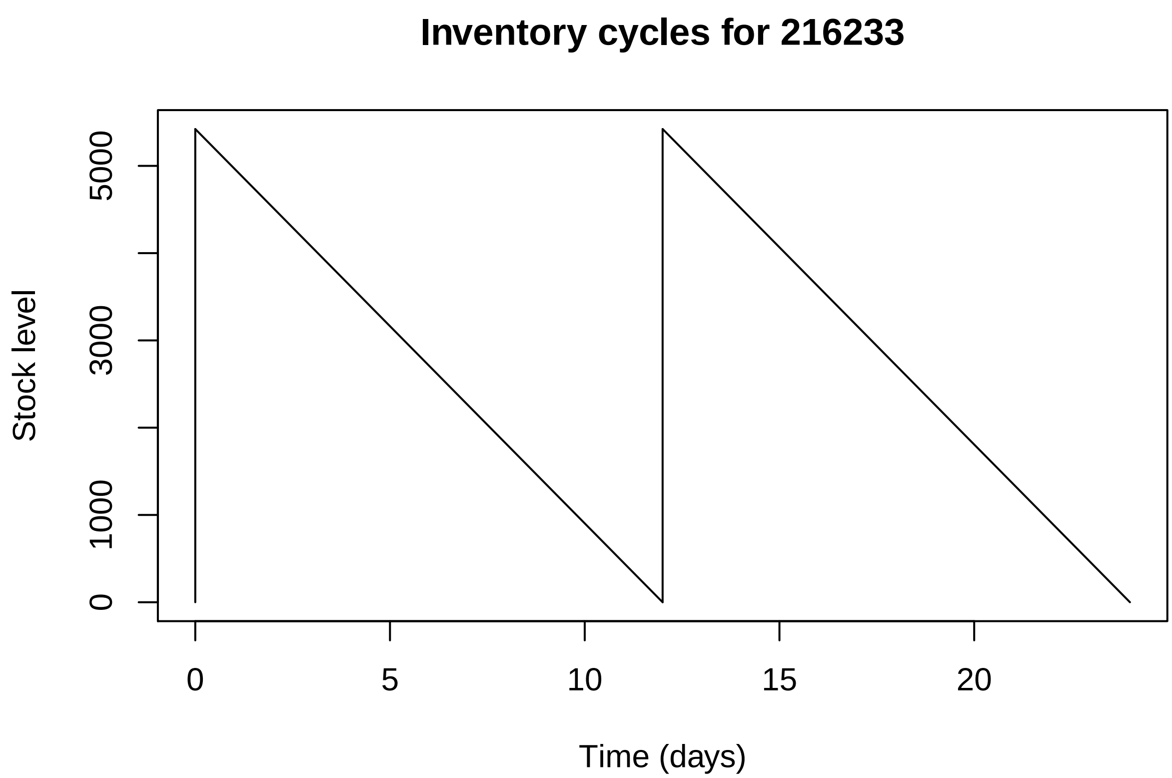
(a) The Operations Manager is interested in studying an EOQ model for product 216233, based on sales in 2012. The setup and holding costs are known to be 130 per order and 1.50 per unit per year, respectively.

i) Determine the best order quantity in such a way that the costs are minimised. Write 1 – 2 paragraphs summarising your findings.

Marking criteria • Number of orders during a year, number of days between orders, and the total annual inventory cost are correctly computed and included in the findings. • The paragraphs clearly explain your findings. • Assumptions of the EOQ model are clearly stated

```
# Annual demand in 2012 for product 216233
sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week)) %>%
  filter(sku_id == "216233", year == 2012) %>%
  select(units_sold) %>%
  sum()-> A

k <- 130 # set up costs per order
h <- 1.5 # holding costs per unit
```

```r
Q <- sqrt(2*k*A / h) # optimum order quantity
Q <- round(Q, 0)


t <- sqrt(2*k/(A*h)) # inventory cycle
t <- round(365*t, 0)

Tc <- k*A/Q + h*Q/2 # annual inventory cost


plot(0, xlim = c(0, 2*t), ylim = c(0, Q), type = "n",
     xlab = "Time (days)", ylab = "Stock level",
     main = "Inventory cycles for 216233")
segments(x0 = c(0,t), y0 = Q, x1 = c(t,2*t), y1 = 0) # diagonals
segments(x0 = c(0,t), y0 = 0, x1 = c(0,t), y1 = Q)  # vertical
```



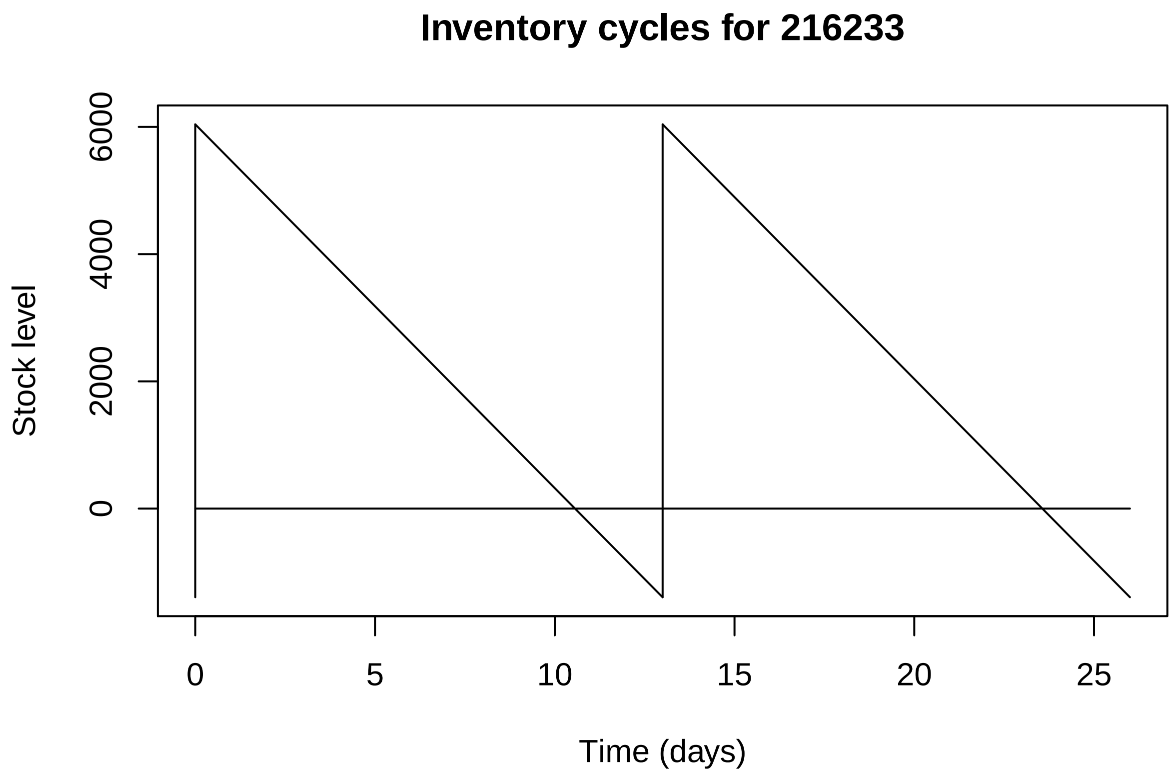
Optimum order quantity is 5422 Inventory cycle is 12 Annual inventory cost is 8132.6803762

Assumptions of EOQ model: Known and constant demand No lead time - orders arrive instantaneously No back orders

ii) The Operations Manager is also interested in studying a model in which backorders are permitted. According to its estimates, the cost of backorders is approximately 5% of the total price (price per unit). Determine the best order quantity in the sense that inventory costs are minimised. Write 1 – 2 paragraphs summarising your findings and plot the first two inventory cycles.

Marking criteria • The optimum order quantity, maximum level of stock, optimum time between orders, proportion of time the company have to take backorders, and total annual inventory cost are correctly computed and included in your answer. • The paragraphs clearly explain your findings. • Assumptions of the model are clearly stated. • The first two inventory cycles are correctly plotted

```
# Annual demand in 2012 for product 216233
sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week)) %>%
  filter(sku_id == "216233", year == 2012) %>%
  select(units_sold) %>%
  sum()-> A

k <- 130 # set up costs per order
h <- 1.5 # holding costs per unit

sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week)) %>%
  filter(sku_id == "216233", year == 2012) %>%
  select(total_price)-> x # create vector of prices for product 216233

p <- mean(x$total_price)*0.05 # back order cost per unit (0.05 of total price)

Q <- sqrt(2*k*A/h) * sqrt((p+h)/p)
Q <- round(Q, 0) # Optimum ordering quantity

S <- sqrt(2*k*A/h) * sqrt(p/(p+h))
S <- round(S, 0) # Optimum maximum inventory level

t <- round(365*Q/A, 0) # Optimum time between orders

t1 <- round(365*S/A, 0) # Time supplier is in stock
pt1 <- 100*(t-t1)/t # Proportion of time supplier taking back orders

Tc <- k*A/Q + h/2 * S^2/Q + p/2 * (Q-S)^2 /Q # Total inventory cost

bo <- Q*(pt1/100) # back orders before reordering

# Plotting
plot(0, xlim = c(0, 2*t), ylim = c(-bo, Q), type = "n",
     xlab = "Time (days)", ylab = "Stock level",
     main = "Inventory cycles for 216233")
segments(x0 = c(0,t), y0 = Q, x1 = c(t,2*t), y1 = -bo) # diagonals
segments(x0 = c(0,t), y0 = -bo, x1 = c(0,t), y1 = Q)  # vertical
segments(x0 = c(0,t), y0 = 0, x1 = c(t,2*t), y1 = 0) # horizontal
```



Optimum ordering quantity 6040 Optimum maximum inventory level 4867 Optimum time between orders 13 Proportion of time taking back orders 23.0769231 Total inventory cost 7300.4389305

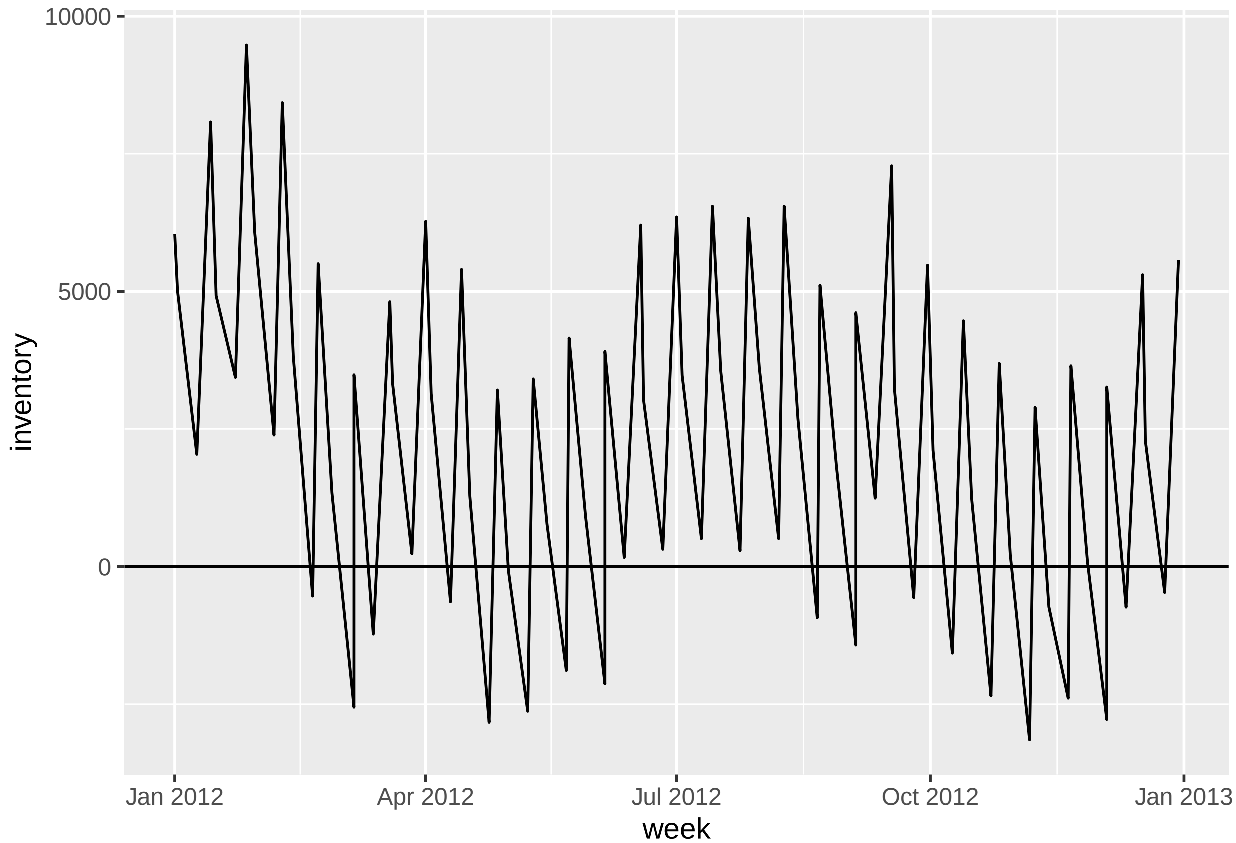
Assumptions of model: known and constant demand no lead time orders arrive instantaneously back orders allowed demand can be backordered when no stock

iii) Plot the inventory cycles associated with the model in part ii and compare with the observed inventory levels in 2012, assuming actual demand during 2012, and the order frequency and order quantity from the model. Write 2 – 3 sentences describing your plot.

Marking criteria • The inventory levels from the model and data are correctly plotted. • Accurate and insightful comments are made about the plot. • Note: This is a bonus question. The maximum mark that could be awarded for this project is 100

```
# Calculate weekly sales in 2012 for product 216233
sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week))%>%
  filter(sku_id == "216233", year == 2012) %>%
  group_by(week) %>%
  summarise(quantity = sum(units_sold)) -> weekly_sales

# create data frame with reorder quantities and dates
seq(from = 0, to = 365, by = t) %>%
  as_tibble() %>%
  mutate(week = as_date("2012-01-01") + value, quantity = Q) %>%
```

```
  select(week, quantity) -> inventory

# combine data frames and create a running total
weekly_sales %>%
  mutate(quantity = quantity * (-1)) %>%
  rbind(inventory) %>%
  arrange(week) %>%
  mutate(inventory = cumsum(quantity),
         type = (ifelse(quantity > 0, "inventory", "sales"))) -> inventory2

ggplot(inventory2) +
  geom_line(aes(x = week, y = inventory)) +
  geom_hline(yintercept = 0)
```



The Operations Manager is considering the option of a multi-period inventory model. The company, as a policy, is not willing to tolerate more than 5% chance of a stock-out. The Operations Manager has estimated that the annual holding cost is 6.50 per unit and the ordering cost is 20.50 per order.

i. Calculate a multi-period inventory model for product 216425, based on the 2012 sales data. Create plot/s of the weekly average demand of this product. Use the costs stated in part (b)above. Write a paragraph explaining the results of your model and the plot/s.

Hint: Use the weekly demand to estimate the demand during a one-week lead time.

Marking criteria • The optimal order quantity, safety stock, expected annual cost, orders per years are correctly computed and included in your answer. • The paragraph clearly explains your findings. • The assumption of normality for the demand during a one-week lead time is discussed. • The weekly average demand of this product is correctly plotted and discussed

```
sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week))%>%
  filter(sku_id == "216425", year == 2012) %>%
  group_by(week) %>%
  summarise(average = mean(units_sold)) -> weekly_sales2

sales_data %>%
   mutate(week = as_date(week, format = "%d/%m/%y"),
         year = year(week))%>%
  filter(sku_id == "216425", year == 2012) %>%
  group_by(week) %>%
  summarise(units_sold)
```

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

```
## # A tibble: 3,432 x 2
## # Groups:   week [52]
##    week       units_sold
##    <date>          <int>
##  1 2012-01-02         15
##  2 2012-01-02         50
##  3 2012-01-02          9
##  4 2012-01-02         46
##  5 2012-01-02         20
##  6 2012-01-02         39
##  7 2012-01-02          6
##  8 2012-01-02          2
##  9 2012-01-02         14
## 10 2012-01-02         13
## # ... with 3,422 more rows
```