Assignment STAT702

Group 5: Hitarth Asrani and Genevieve Connell

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 – July 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.**

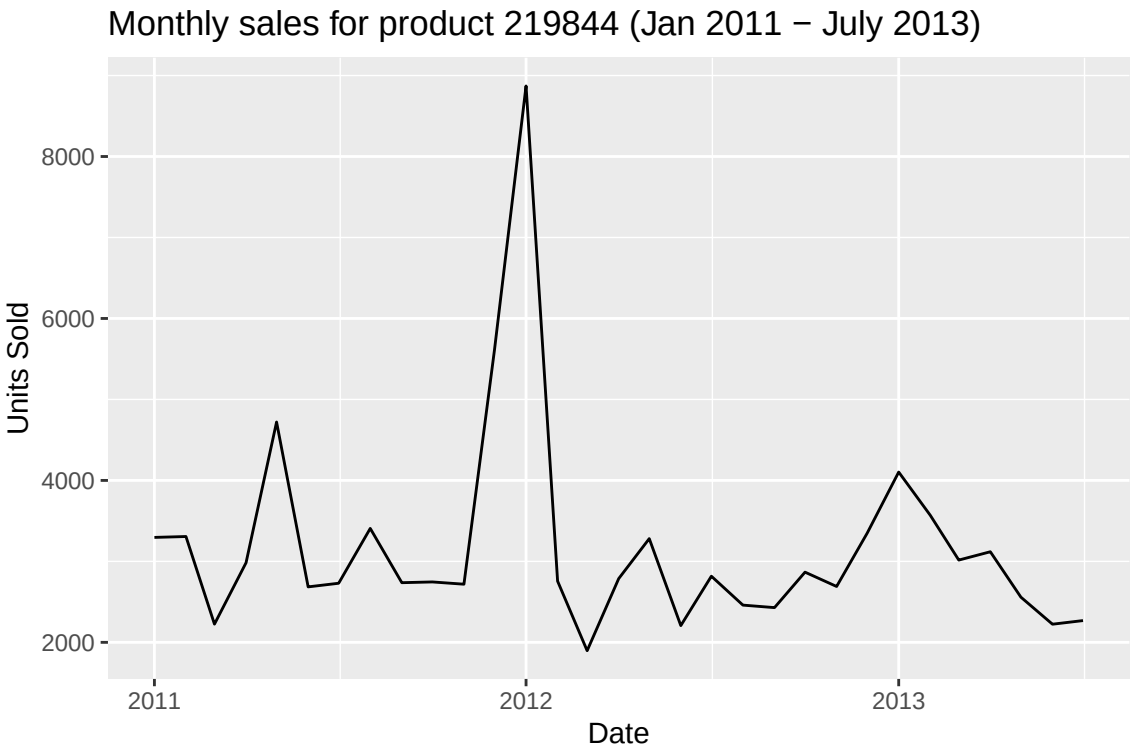
Answer

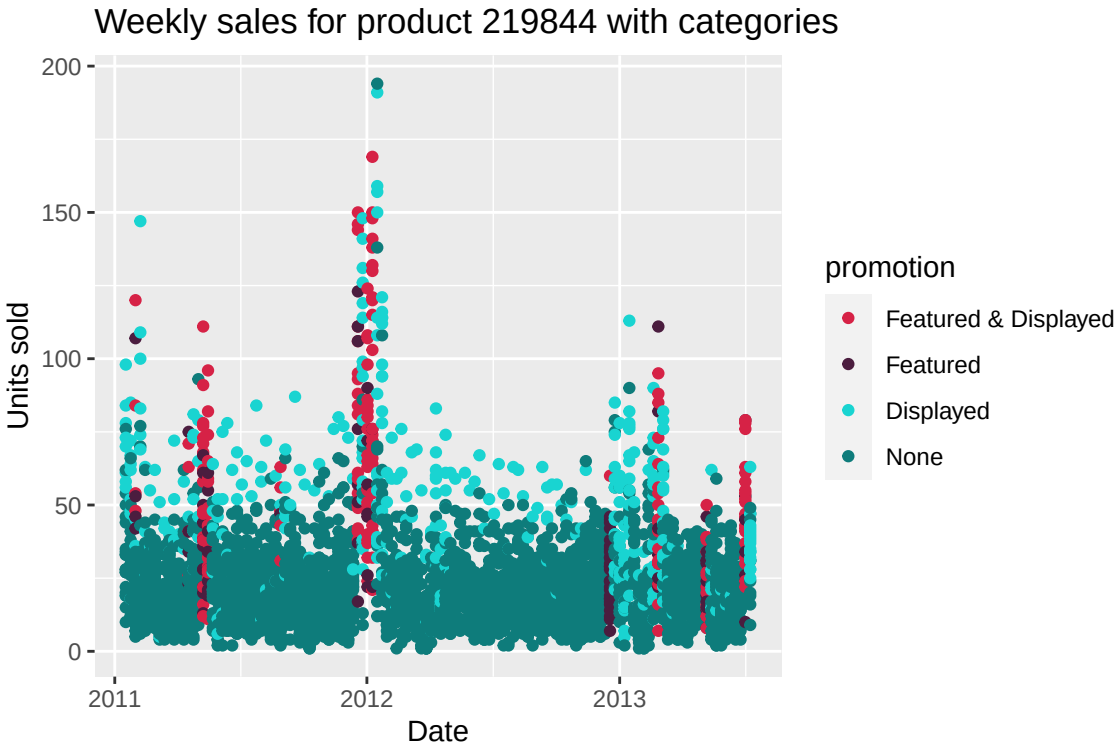
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1897    2622    2787    3175    3302    8871
```

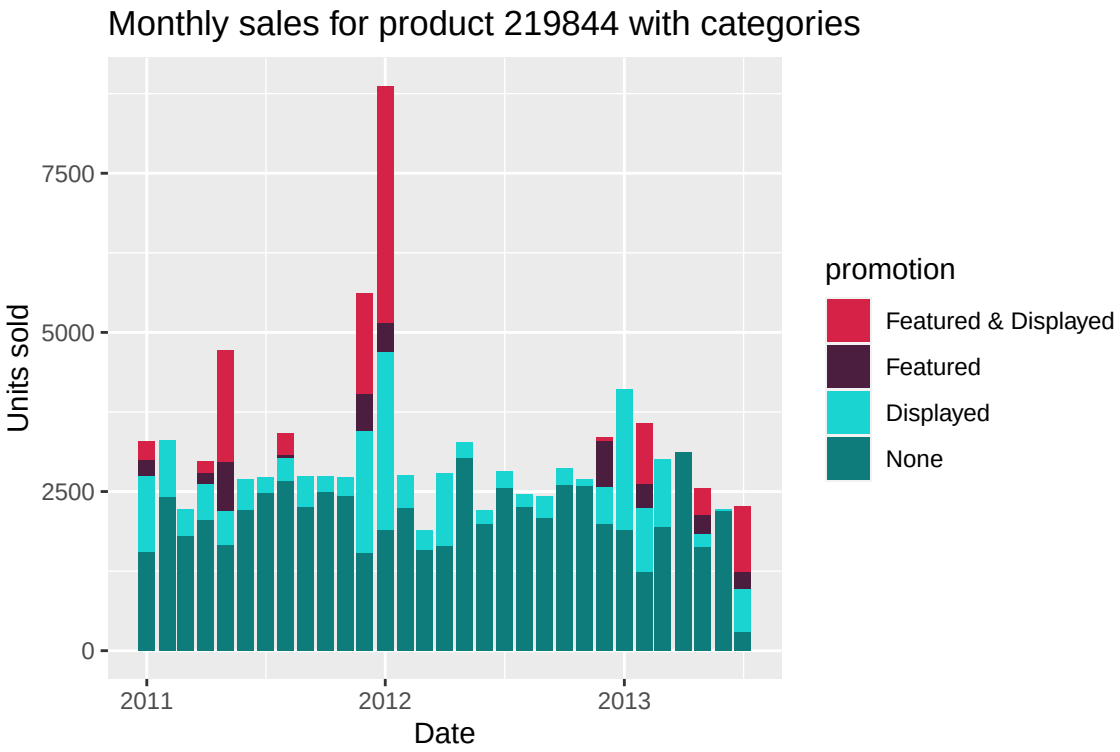
From Jan 2011 - July 2013 98434 units of sku 219844 were sold with a mean monthly sale of 3175.2903226. The interquartile range is2621.5 -3301.5. This means 50% of monthly sales lie within this range.

Monthly sales are plotted below, no trend or seasonal pattern is evident in this plot. There are three months with significantly more sales, May 2011, December 2011 and Januarary 2012. The most significant outlier was on when 8871 units were sold.

In these months a large proportion of sales were made at stores where product 219844 was featured and/or displayed. In the scatterplot there is a pattern of high weekly store sales where products are featured and/or displayed.


Monthly sales for product 219844 (Jan 2011 − July 2013)
Units Sold
8000
6000
4000
2000
2011
2012
2013
Date


Weekly sales for product 219844 with categories
Units sold
200
150
100
50
0
2011
2012
2013
Date
promotion
Featured & Displayed
Featured
Displayed
None



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.

Answer

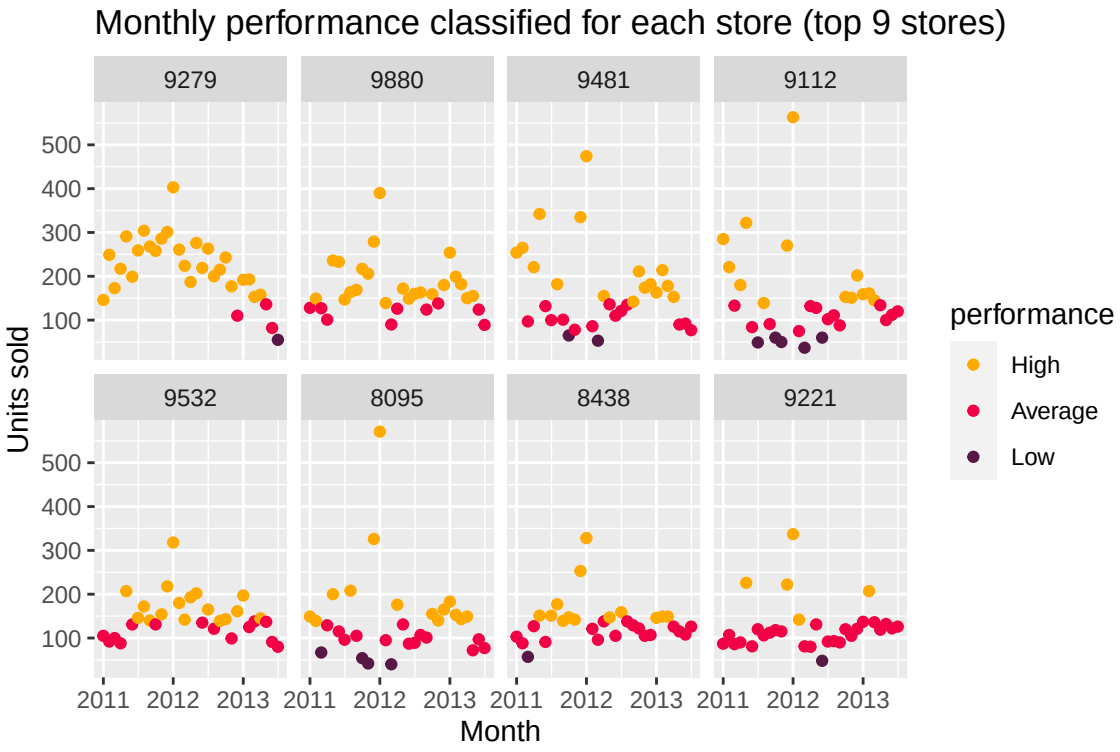
All stores have been ranked based on their total sales from Jan 2011 to July 2013. The store with the highest total units sold is ranked '1', this is store 9279 with 6698 units. Stores ranked 1-9 are performing highly. Stores 10 - 19 are performing adequately. Stores 20 - 29 are performing poorly.

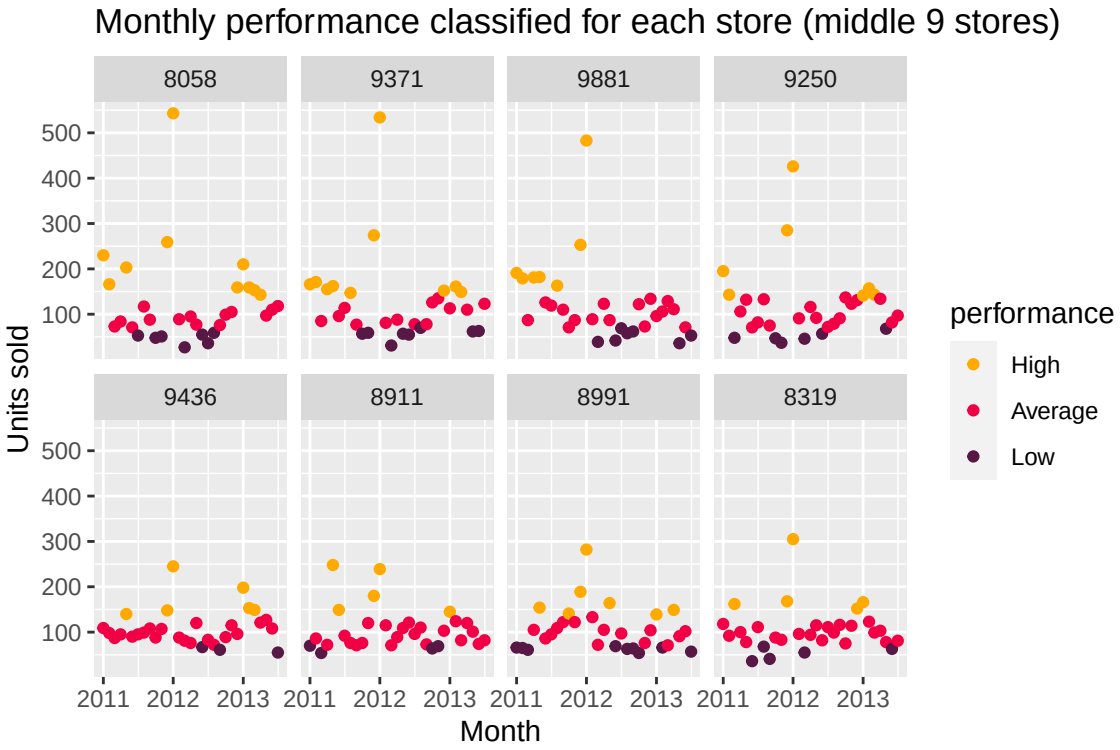
Rank	Store ID	Total sales	Rank	Store ID	Total sales	Rank	Store ID	Total sales
1	9279	6698	10	8058	3853	19	9432	2777

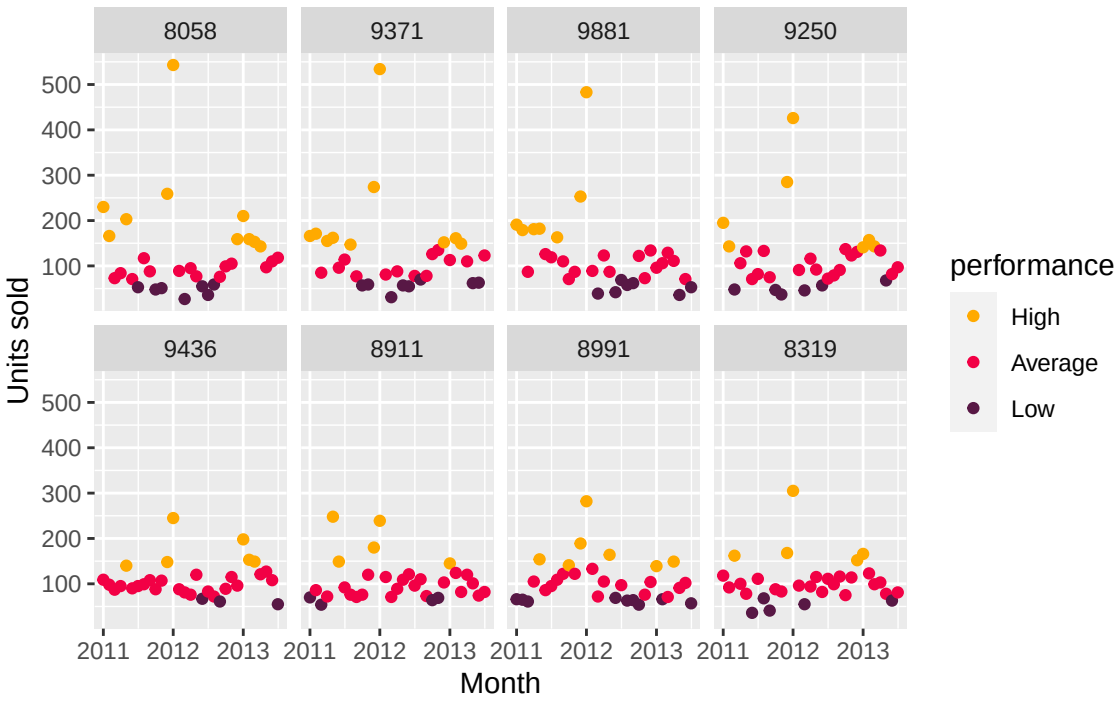
Below is a bar chart showing the total units sold with stores ordered by rank. Sales for 2013 are much lower than 2011 and 2012 as only half the years data is included.

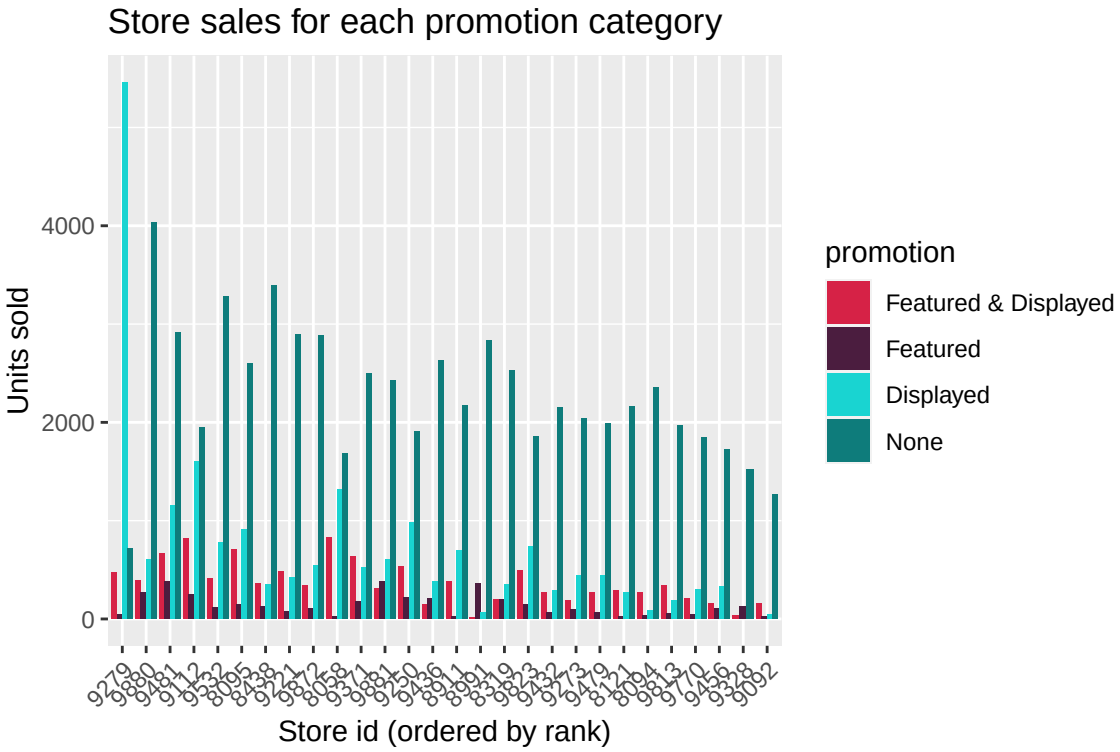
|3 |9880 |5118|{r 1b store totals} # Bar chart of total yearly sales for stores (ordered by rank) sales_219844 %>% mutate(year = as.character(year)) %>% group_by(store_id_ranked, year) %>% summarise(total_units = sum(units_sold)) %>% ggplot() + geom_bar(aes(x = store_id_ranked, y = total_units, fill = year), stat = "identity") + theme(axis.text.x = element_text(angle = 45, hjust = 1)) + scale_fill_manual(values = c("#5BC0BE", "#3A506B", "#0B132B")) + ggtitle("Total units sold by each store") + xlab("Store id (ranked)") + ylab("Total units sold")

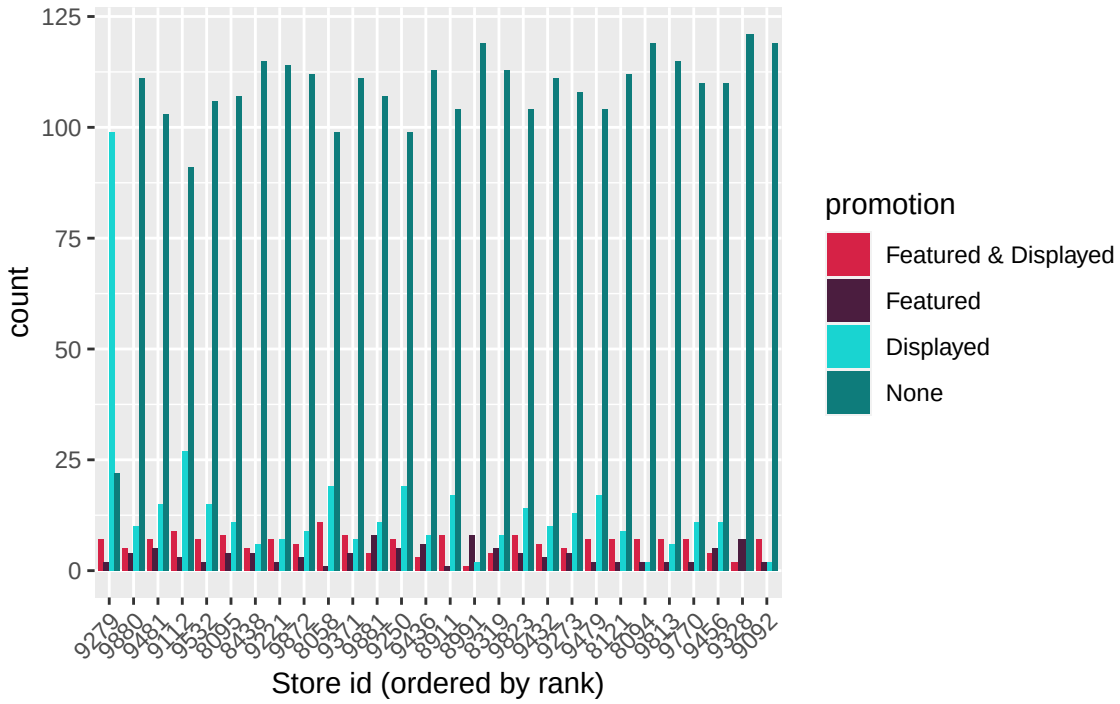
Monthly performance has been classified for each store. An interquartile range for monthly sales has been calculated with the first quartile at 71 and the third quartile at 138. Monthly sales greater than the third quartile are classified as 'High' performing. Sales

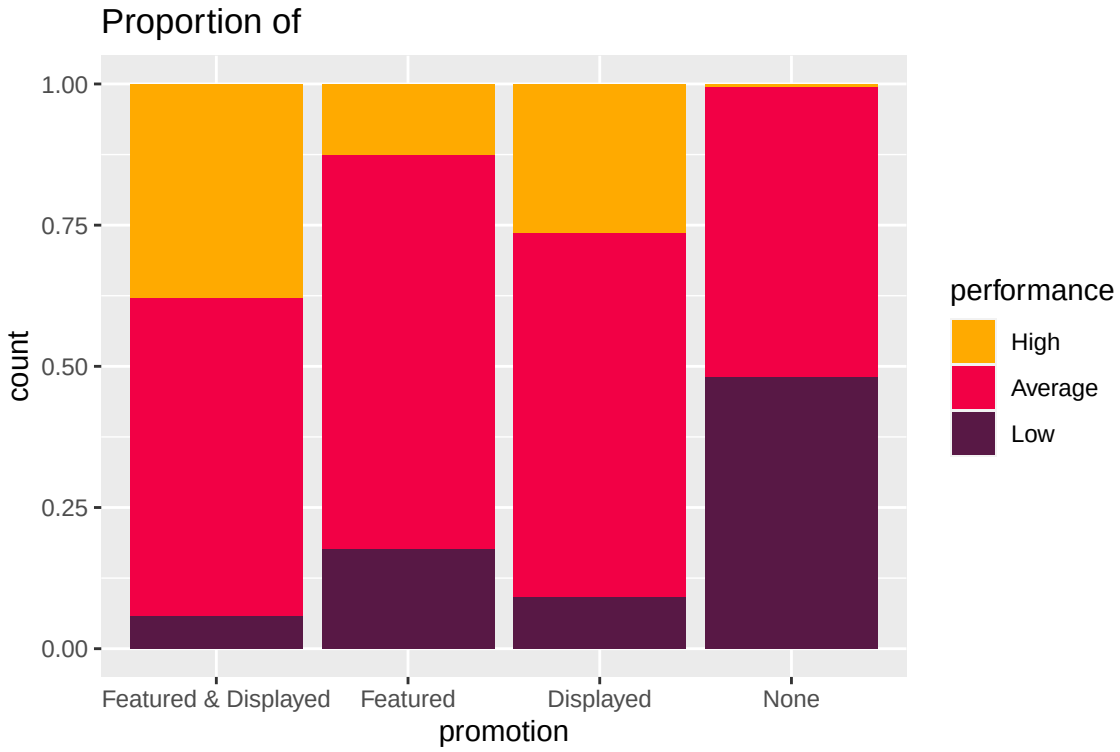


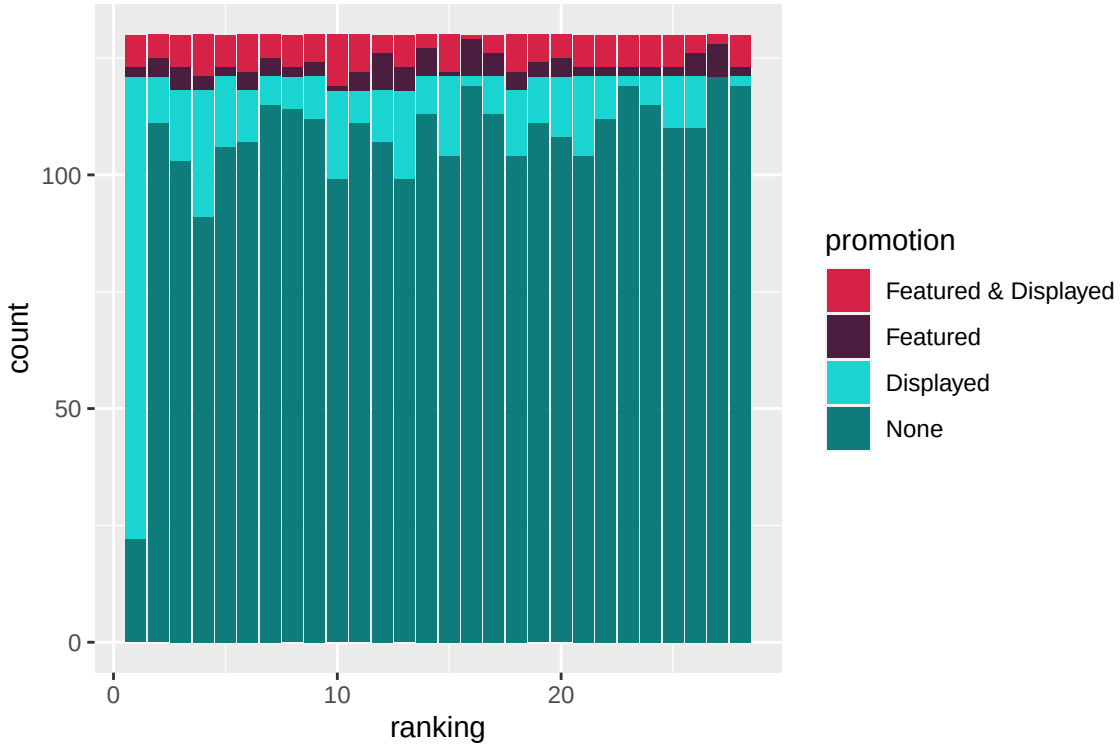

Monthly performance classified for each store (middle 9 stores)
8058
9371
9881
9250
9436
8911
8991
8319
Units sold
500
400
300
200
100
2011
2012
2013
Month
performance
High
Average
Low


Monthly performance classified for each store (bottom 10 stores)
8058
9371
9881
9250
9436
8911
8991
8319
Units sold
500
400
300
200
100
2011
2012
2013
Month
performance
High
Average
Low


Store sales for each promotion category
Units sold
4000
2000
0
9279
9880
9481
9112
9532
8095
8438
9221
9872
8058
9371
9881
9250
9436
8911
8991
8319
9823
9432
9273
9479
8121
8094
9813
9770
9456
9328
9092
Store id (ordered by rank)
promotion
Featured & Displayed
Featured
Displayed
None


count
125
100
75
50
25
0
9279
9880
9481
9112
9532
8095
8438
9221
9872
8058
9371
9881
9250
9436
8911
8991
8319
9823
9432
9273
9479
8121
8094
9813
9770
9456
9328
9092
Store id (ordered by rank)
promotion
Featured & Displayed
Featured
Displayed
None


Proportion of
1.00
0.75
0.50
0.25
0.00
count
Featured & Displayed
Featured
Displayed
None
promotion
performance
High
Average
Low


100
50
0
count
0
10
20
ranking
promotion
Featured & Displayed
Featured
Displayed
None

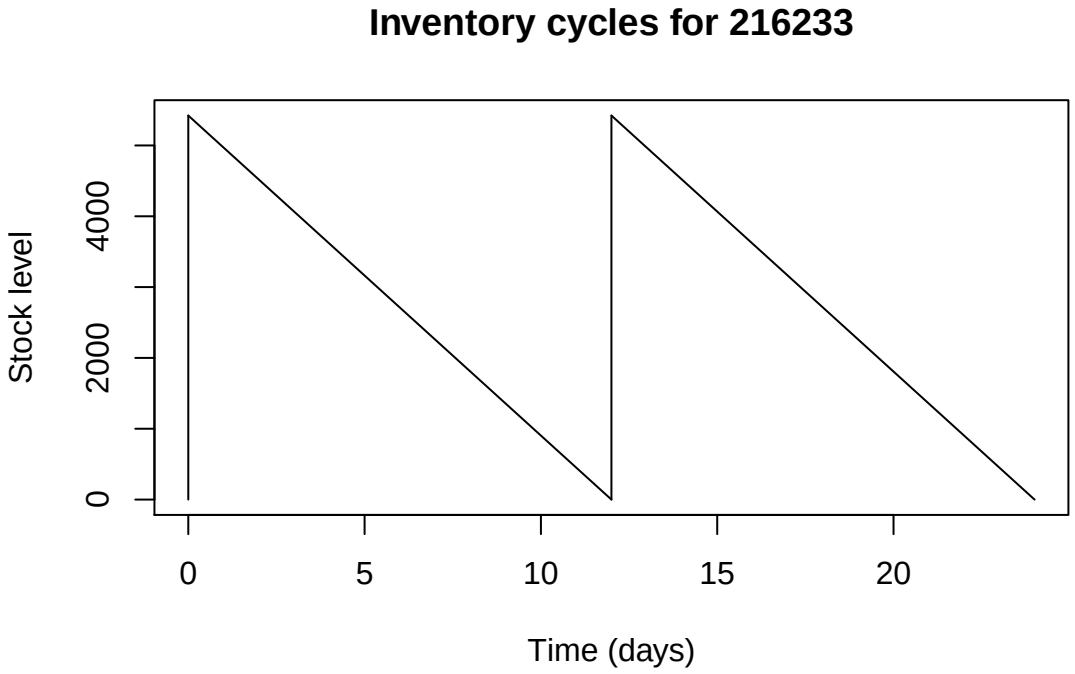
Question 2

(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



Optimum order quantity is 5422

Inventory cycle is 12

Annual inventory cost is 8132.6803762

Assumptions of EOQ model:

Known and constant demand

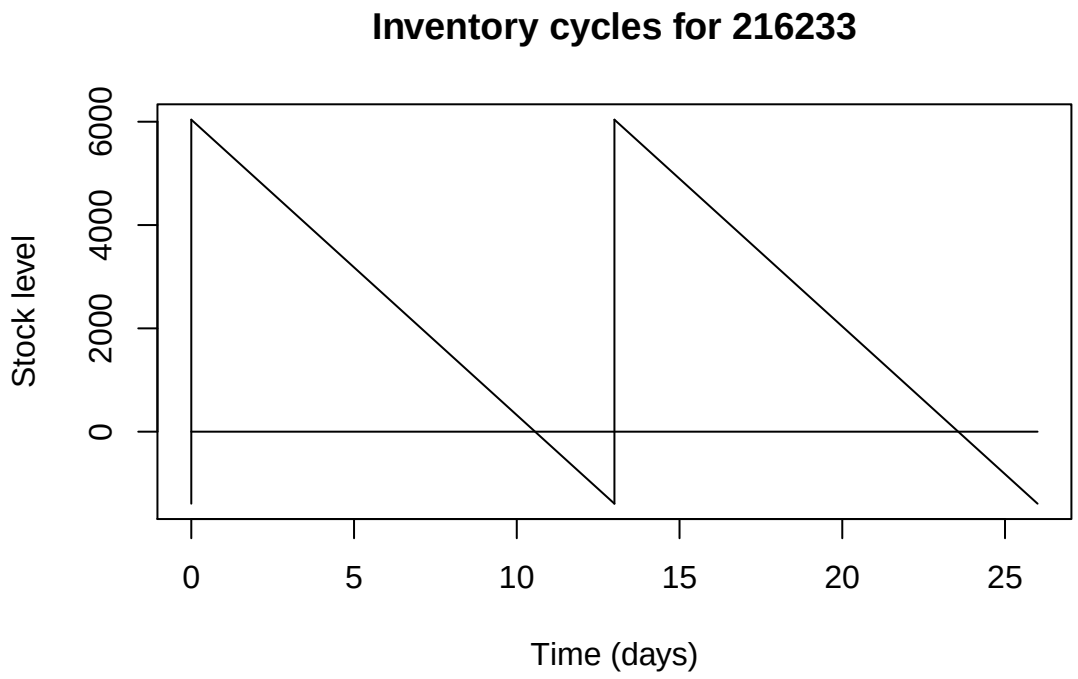
No lead time - orders arrive instantaneously

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



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

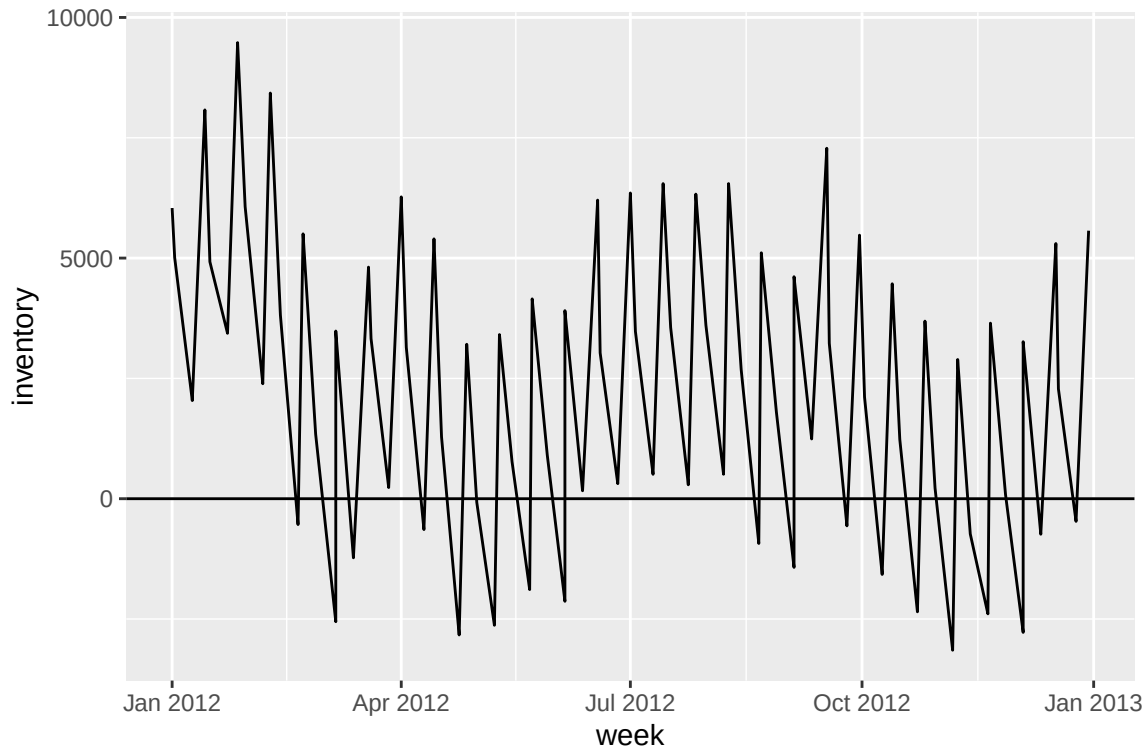
orders arrive instantaneously

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

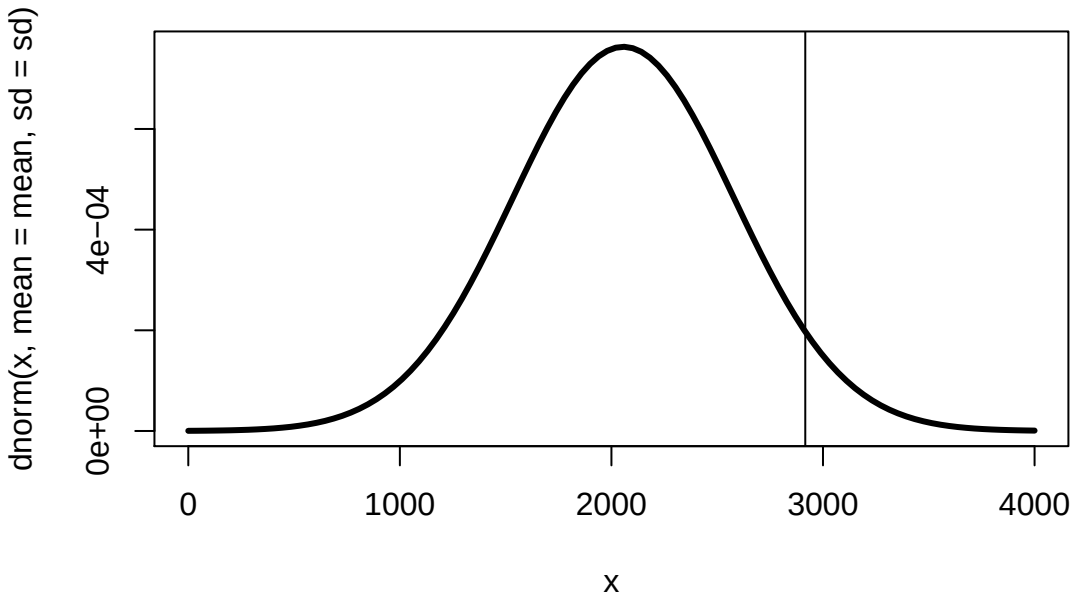


The Operations Manager is considering the option of a multi-period inventory model. The company, as a policy, is not willing to tolerate more than 5% chance of a stock-out. The Operations Manager has estimated that the annual holding cost is 6.50 per unit and the ordering cost is 20.50 per order.

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




```
## integer(0)
```

Demand during a one week lead time has been estimated based off 2012 weekly demand. The mean weekly demand from 2012 and standard deviation have been applied to a normal distribution to estimate demand for this multi-inventory model.

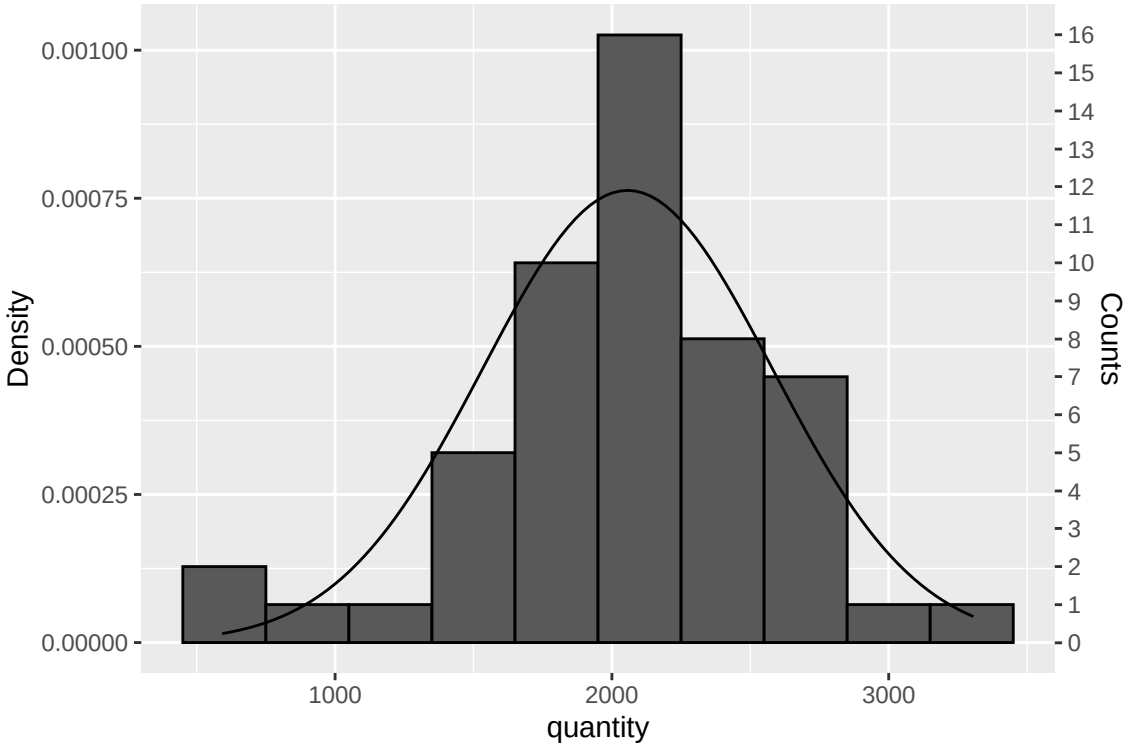
As the plot below shows, the actual demand for 2012 does not perfectly follow a normal distribution. It is difficult to determine the distribution of weekly demand based on only one year of observations.It is recommended that more data is used to for a more accurate model of demand distribution.

Based on 2012, where mean weekly demand was 2056.5192308 with standard deviation of 522.6820535, the expected annual demand is estimated to be 1.06939×10^5.

Given this annual demand and the costs of holding and reordering stock, the recommended multi-inventory model is to order 821 units whenever the order quantity reaches the reorder point of 2916.2547022 units. Approximately 130 orders will be placed per year. Safety stock is 821.

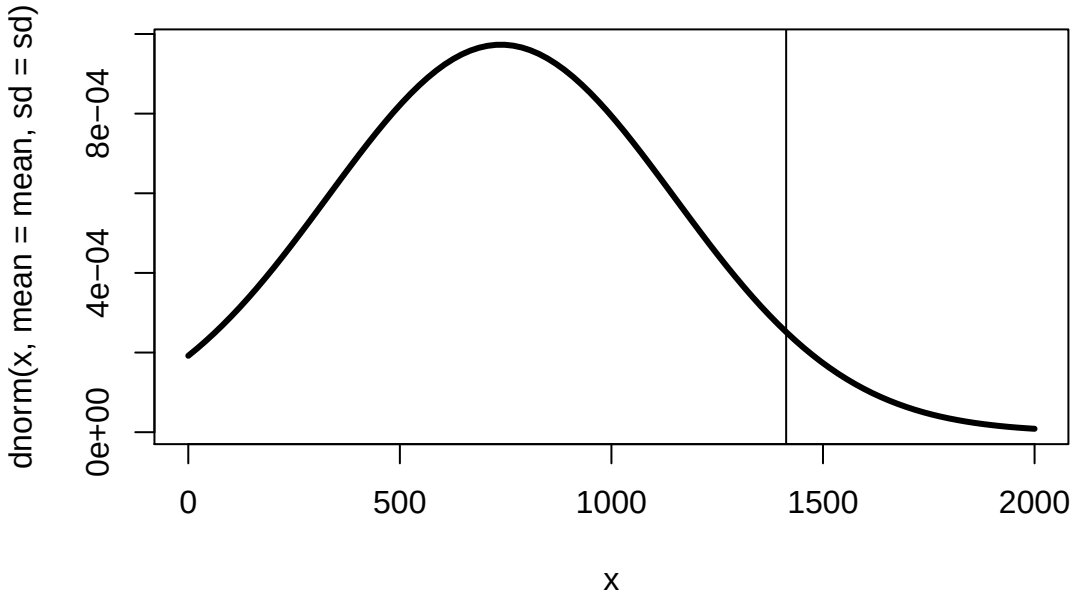
This approach ensures roughly 95% of the time the 2916.2547022 units will be able to satisfy demand during the lead time.

The expected annual costs are 1.0926749×10^4 per year. If demand was certain the annual costs would only be 1.0926749×10^4, the additional cost of holding safety stock is the cost of uncertain demand.


0.00100
0.00075
0.00050
0.00025
0.00000
Density
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
Counts
1000
2000
3000
quantity

2.b.ii. Investigate the use of a multi-period inventory model for the product which has been assigned to your group, 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.

Discuss the assumptions of the model and suggest a solution, in case of finding any problems. Write a paragraph explaining the results of your findings and the plot.




```
## integer(0)
```

Demand during a one week lead time has been estimated based off 2012 weekly demand. The mean weekly demand from 2012 and standard deviation have been applied to a normal distribution to estimate demand for this multi-inventory model.

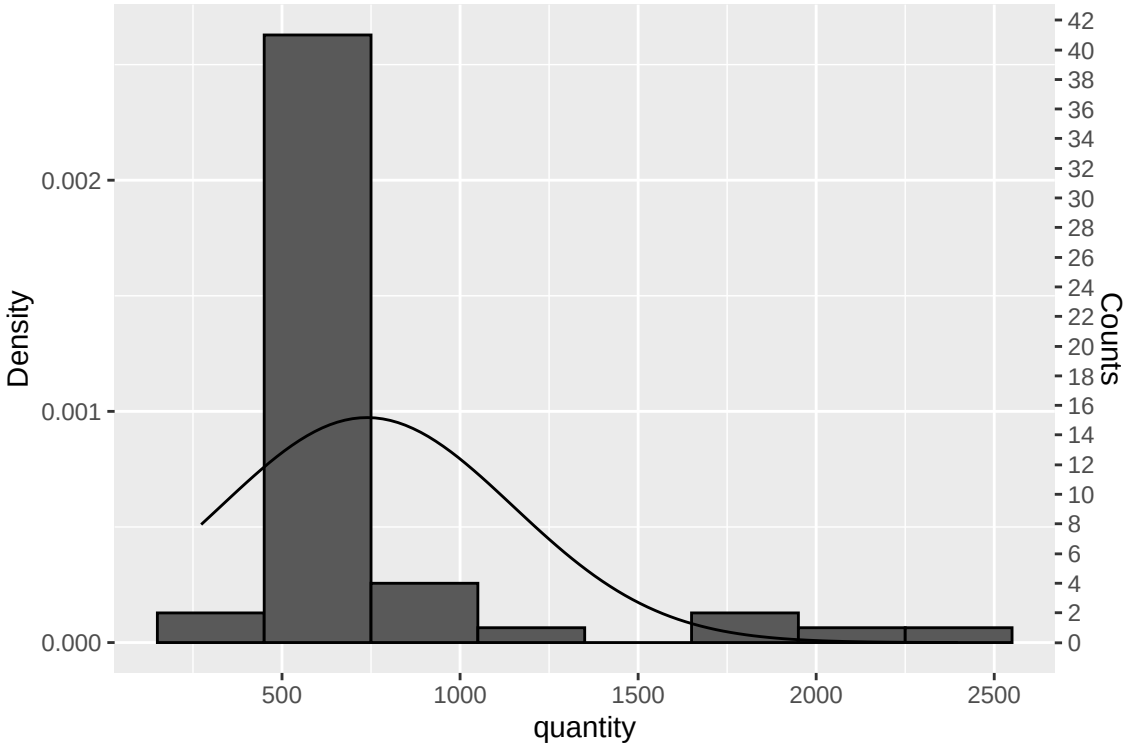
As the plot below shows the observed demand does not follow normal distribution with mean 738.6538462 409.9454715. It is recommended that a more suitable model is used to estimate deand.

Based on the flawed normal model, where mean weekly demand was 738.6538462 with standard deviation of 409.9454715, the expected annual demand is estimated to be 3.841×10^4.

Given this annual demand and the costs of holding and reordering stock, the recommended multi-inventory model is to order 492 units whenever the order quantity reaches the reorder point of 1412.9541418 units. Approximately 78 orders will be placed per year. Safety stock is 492.

This approach ensures roughly 95% of the time the 1412.9541418 units will be able to satisfy demand during the lead time.

The expected annual costs are 7582.3685882 per year. If demand was certain the annual costs would only be 7582.3685882, the additional cost of holding safety stock is the cost of uncertain demand.

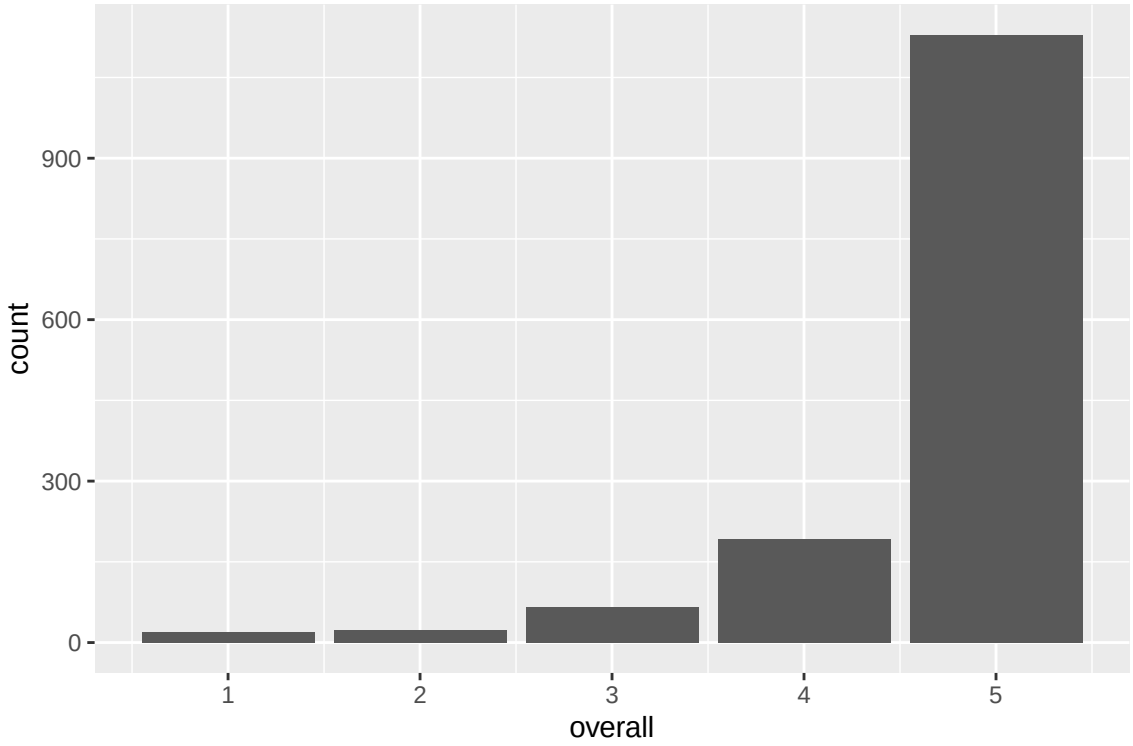


Question 3

Question 3a

```
## # A tibble: 6 x 14
##   title       brand main_cat price asin  document.id overall verified reviewTime
##   <chr>       <chr> <chr>    <chr> <chr>       <int>   <int> <lgl>    <chr>
## 1 BIC Round ~ BIC   Office ~ $2.18 B000~       48735       4 TRUE     05 1, 2018
## 2 BIC Round ~ BIC   Office ~ $2.18 B000~       48762       3 TRUE     04 24, 20~
## 3 BIC Round ~ BIC   Office ~ $2.18 B000~       48763       5 TRUE     04 22, 20~
## 4 BIC Round ~ BIC   Office ~ $2.18 B000~       48774       5 TRUE     04 21, 20~
## 5 BIC Round ~ BIC   Office ~ $2.18 B000~       48775       5 TRUE     04 16, 20~
## 6 BIC Round ~ BIC   Office ~ $2.18 B000~       48776       4 TRUE     04 10, 20~
## # ... with 5 more variables: reviewerID <chr>, reviewerName <chr>,
## #   reviewText <chr>, summary <chr>, unixReviewTime <int>

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   5.000   5.000   4.669   5.000   5.000
```


900
600
300
0
count
1
2
3
4
5
overall