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.

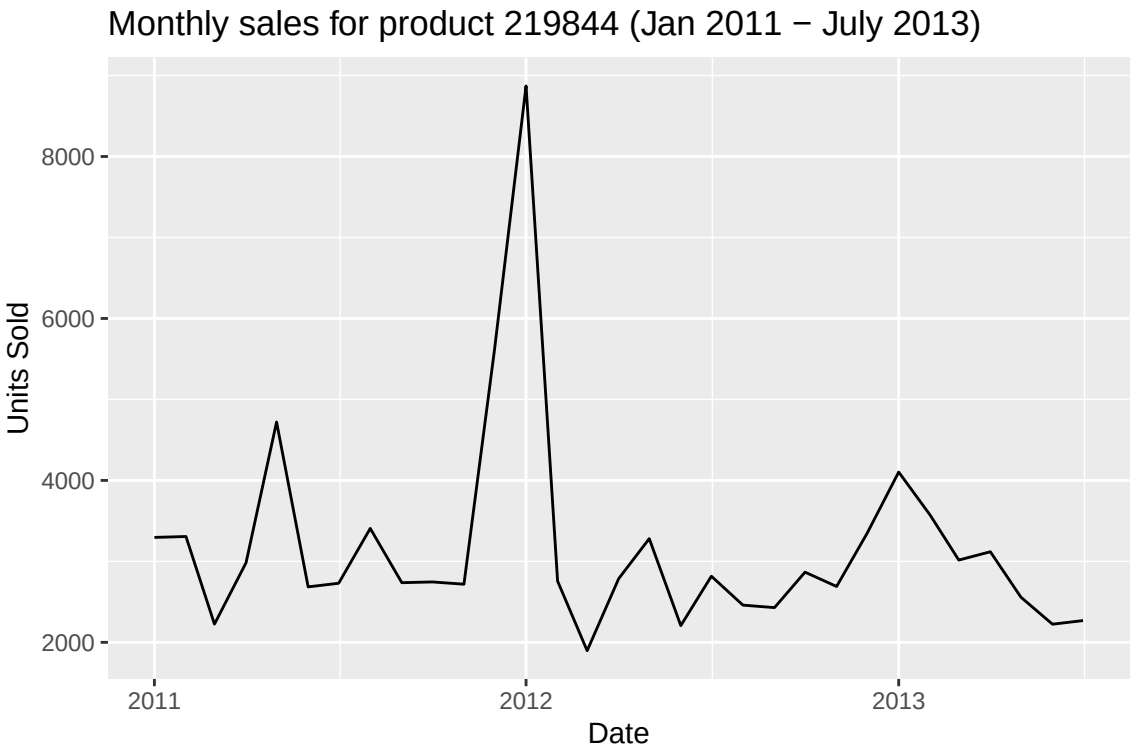
Marking Criteria

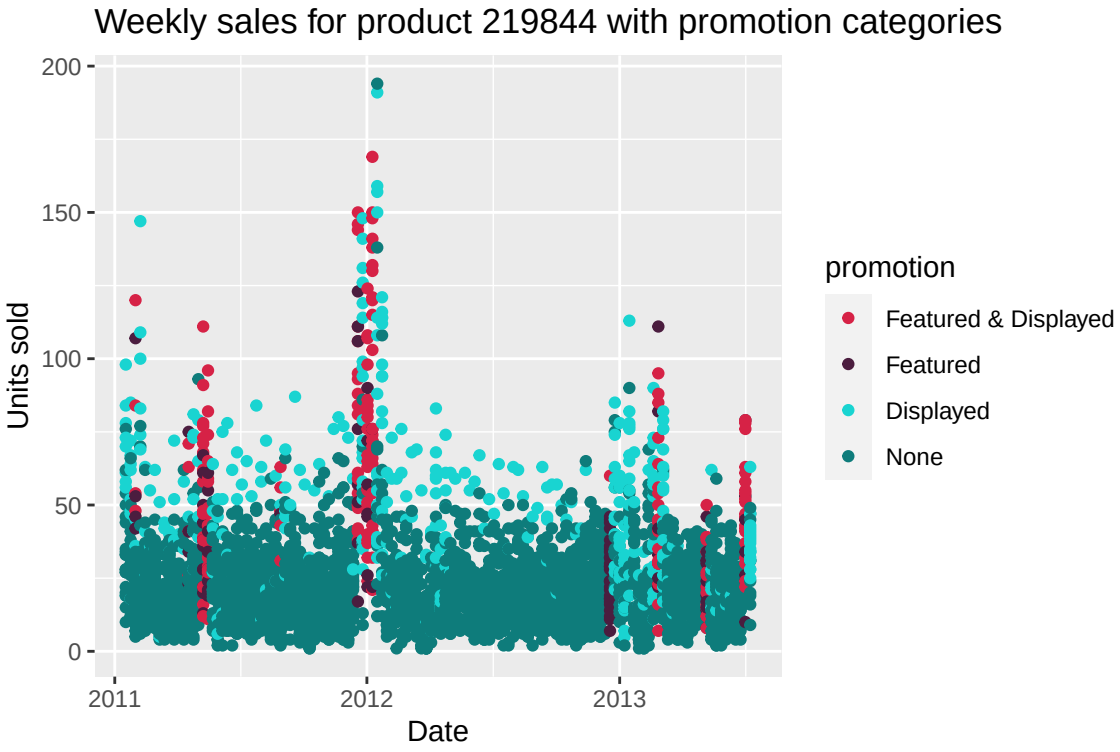
- Total monthly sales have been correctly computed and are displayed in an appropriate plot.
- Description of results/plot is correct and provides useful insights.
- Plot is constructed using ggplot2 and has appropriate titles, labels, scales etc.**

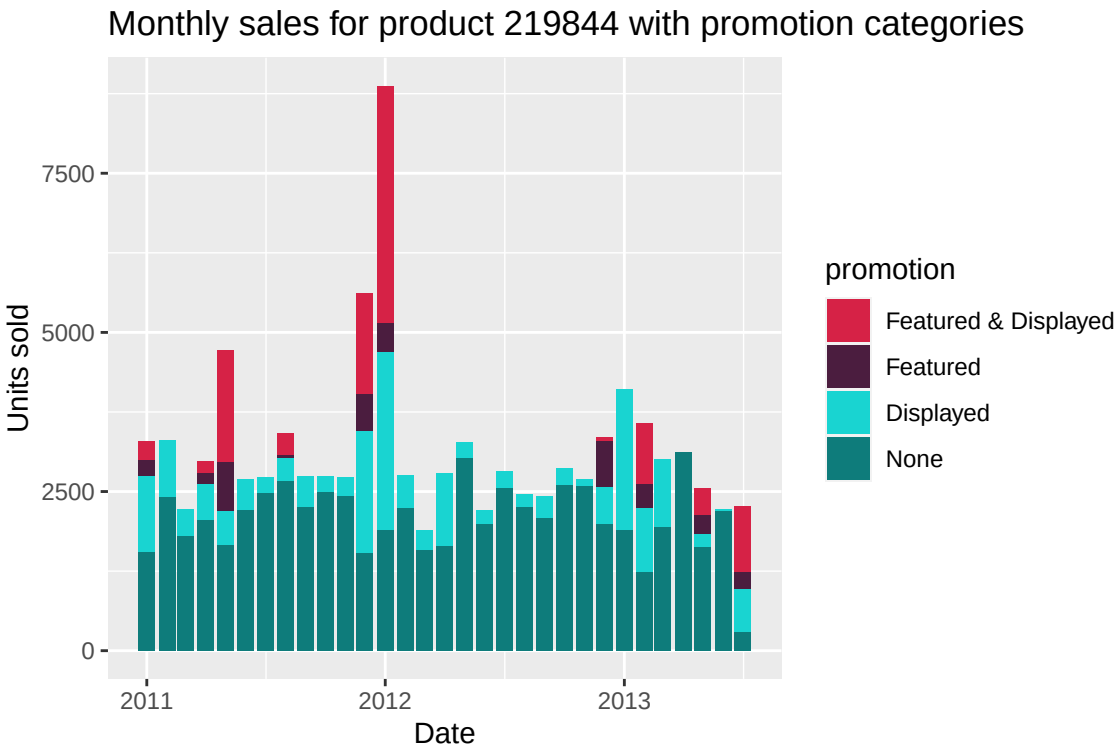
Answer

From Jan 2011 - July 2013, 98434 units of sku 219844 were sold with a mean monthly sale of 3175.3 and an interquartile range of 2621.5 - 3301.5.

Monthly sales are plotted below, no trend or seasonal pattern is evident in this plot. There are three months with significantly high sales, May 2011, December 2011 and Januarary 2012. The most significant outlier was in January 2012 when 8871 units were sold. As shown in the plots below these high monthly sales correspond with a high proportion of stores featuring and/or displaying the product.


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 promotion 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.

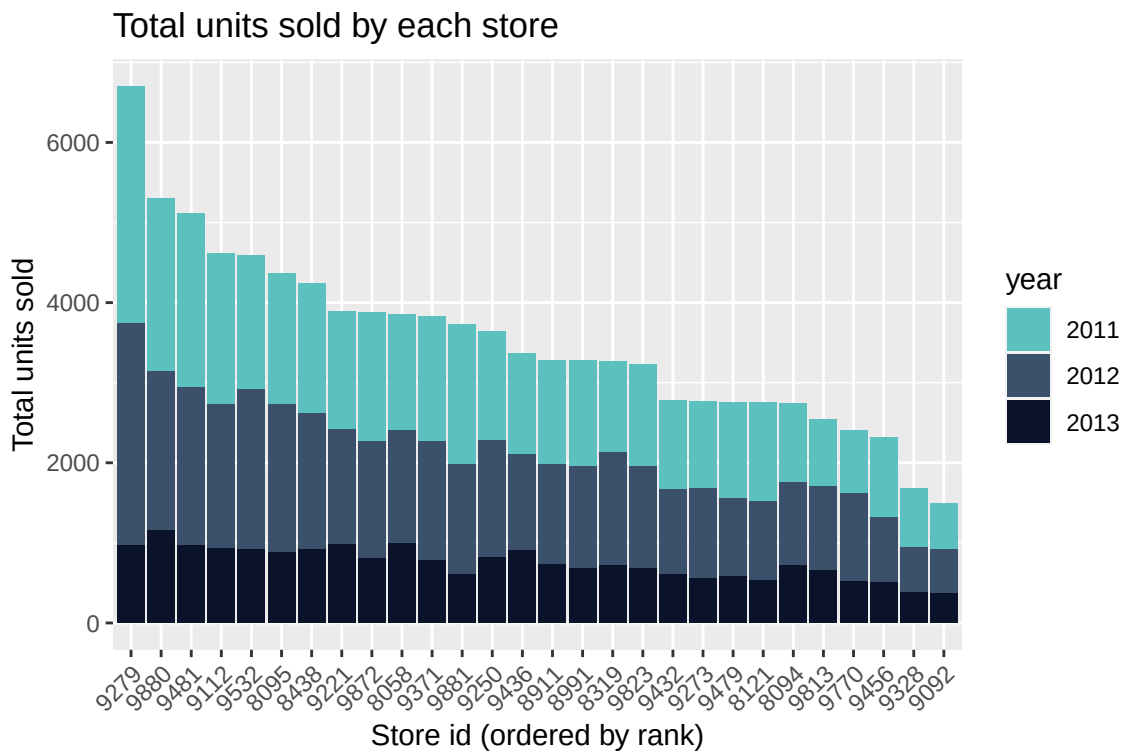
Marking criteria

- 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.

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.



For a more detailed look at store performance monthly sales have been plotted for each store and categorised as 'High', 'Average' or 'Low' performing. These categories are based on the interquartile range for monthly sales. This range is 71 - 138 and captures 50% of all monthly sales. Monthly sales within the range are classified as 'Average', sales greater than 138 are classified as 'High' performing and those below 71 are classified as 'Low' performing.

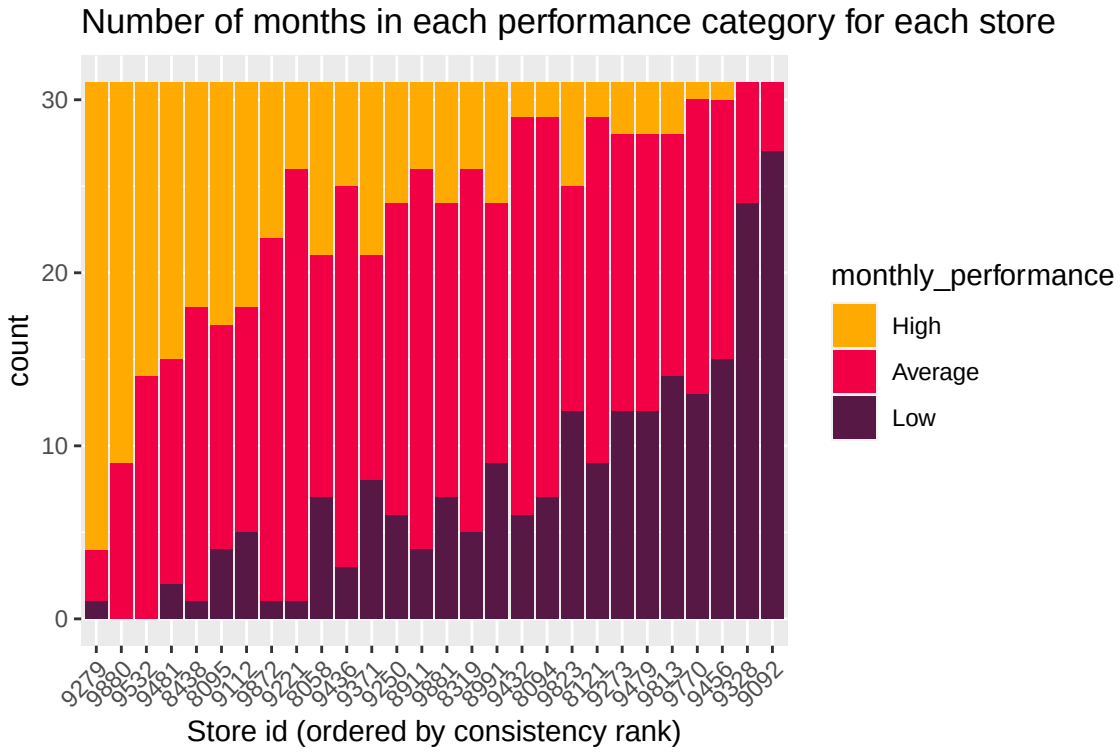
An interesting feature of these plots is that some stores have more consistent performance than others. For example store 9112 had 5 low performing months but was ranked at 4, much higher than store 9872 which had only 1 low performing month but was ranked 9. Despite having higher overall sales store 9112 was less reliably successful than store 9872.

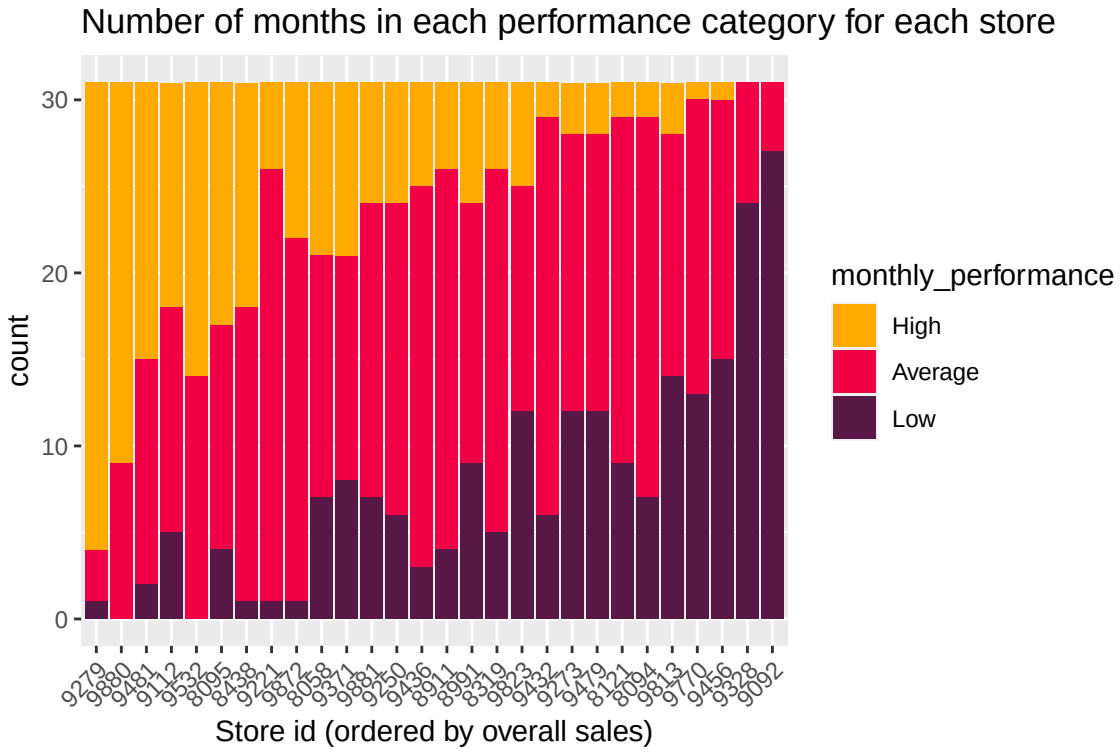
Stores with high occurrences of low sales are not reliable and are a source of risk. As an alternative to ranking stores based on their overall sales, stores have been given a ranking based on their consistency. For

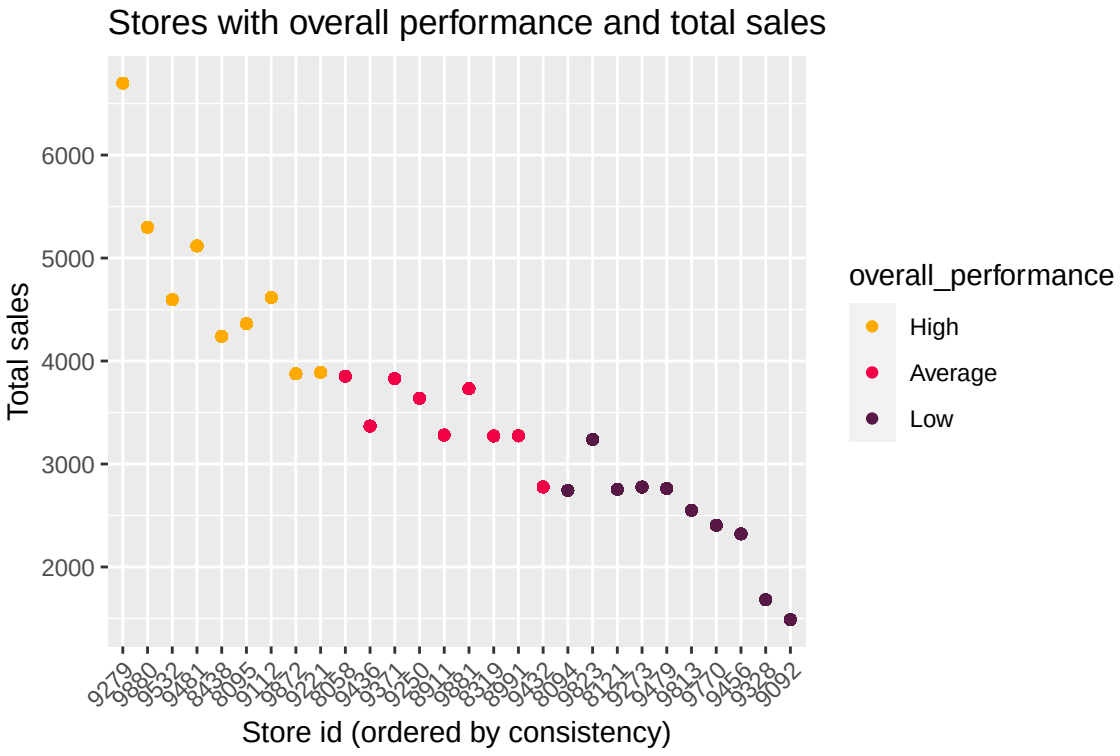
every high performing month stores are given a score of 3, for every average month a score of 2 and for every low performing month a score of 1.

Below is a plot showing stores ordered according to the new consistency rank along with the number of High, Average and Low performing months they achieved. Another plot is included for comparison where stores are ranked according to overall sales. Under the new consistency ranking scheme stores with unreliable performance such as 9112 drop in rank whilst stores with reliable performance such as 9532 are ranked more highly.

Overall consistency ranking seems most appropriate for sales performance as it rewards stores with reliable monthly sales. Consistency ranking has been used to split stores into three groups of overall performance. The top third of stores (consistency rank 1-9) are evaluated as 'High' performing. The middle third of stores (rank 10-18) are evaluated as having 'Average' performance and stores in the bottom third (rank 19-28) are evaluated as 'Low' performing.

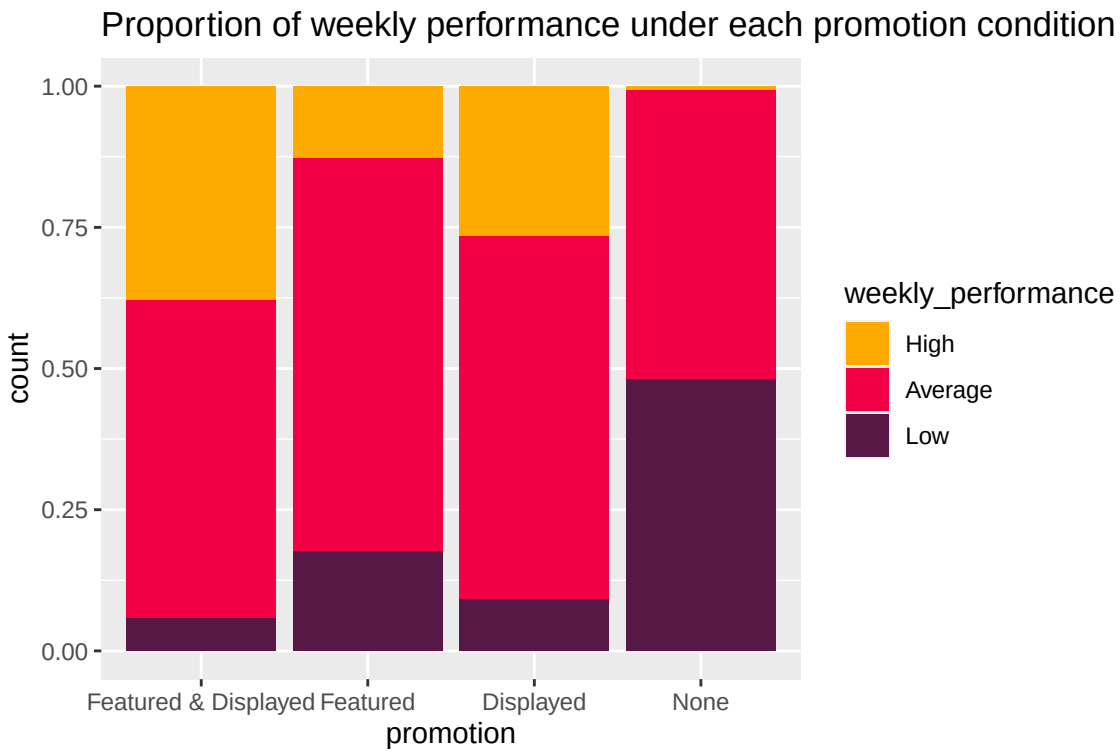






Below is a plot showing that when a product is featured and/or displayed weekly sales are more likely to be high. Weekly performance has been categorised into 'High', 'Average' and 'Low' based on the interquartile range of weekly sales. When the product is featured and/or displayed a higher proportion of weekly sales have been 'High' or 'Average' compared to no promotion.

Low performing stores could boost their sales by increasing the number of weeks they promote the product and unreliable stores could ensure high sales by holding regular promotions.



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

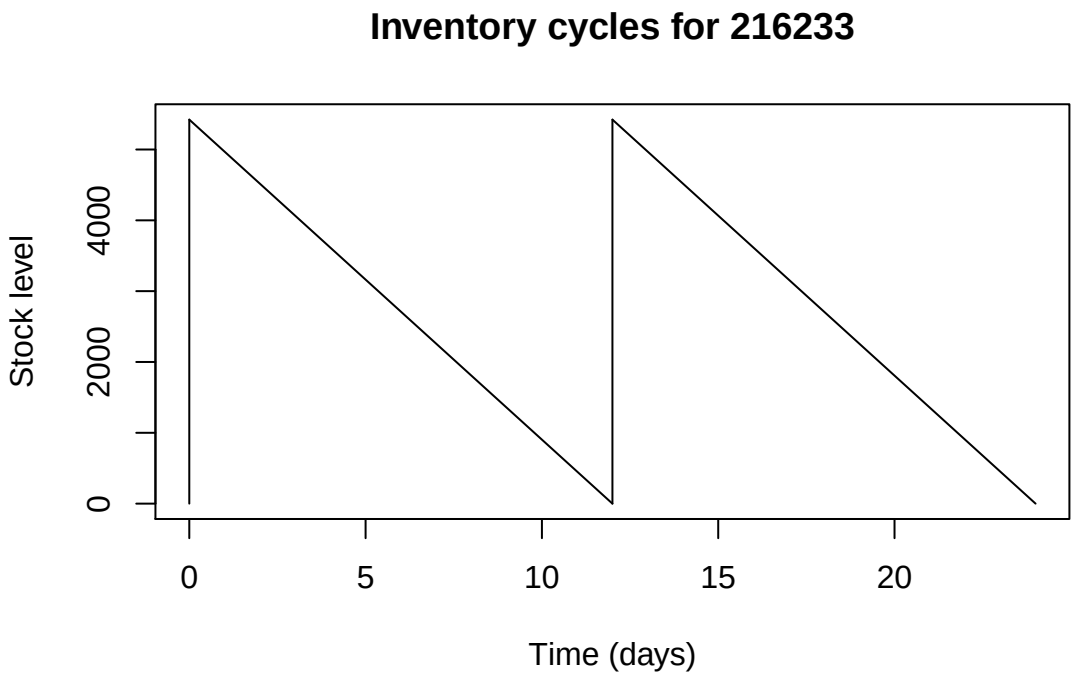
- Number of orders during a year, number of days between orders, and the total annual inventory cost are correctly computed and included in the findings.
- The paragraphs clearly explain your findings.
- Assumptions of the EOQ model are clearly stated

The Economic Order Quantity (EOQ) model is used to find the best order quantity so that total costs are minimised. Key assumptions to this model are that demand is constant and known, there is no lead time, orders arrive instantaneously and back orders are not allowed. Another assumption is that stock levels are under continuous review.

Demand for product 216233 has been estimated based on the annual demand from 2012, this is 169591. The optimum order quantity is calculated based on this demand and the annual order and holding costs. The following EOQ formula is used to determine optimal order quantity where k is order cost and h is holding cost.

$$Q^* = \sqrt{\frac{2kA}{h}}$$

The optimal order quantity is calculated to be 5422 with an inventory cycle of 12 days. This means that every 12 days 5422 units are ordered, resulting in 31 annual orders. This model results in the smallest possible annual inventory cost of 8132.6803762 with an annual order cost of 4066.1803762 and an annual holding cost of 4066.5. This EOQ model is plotted for two cycles below.



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.

- 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

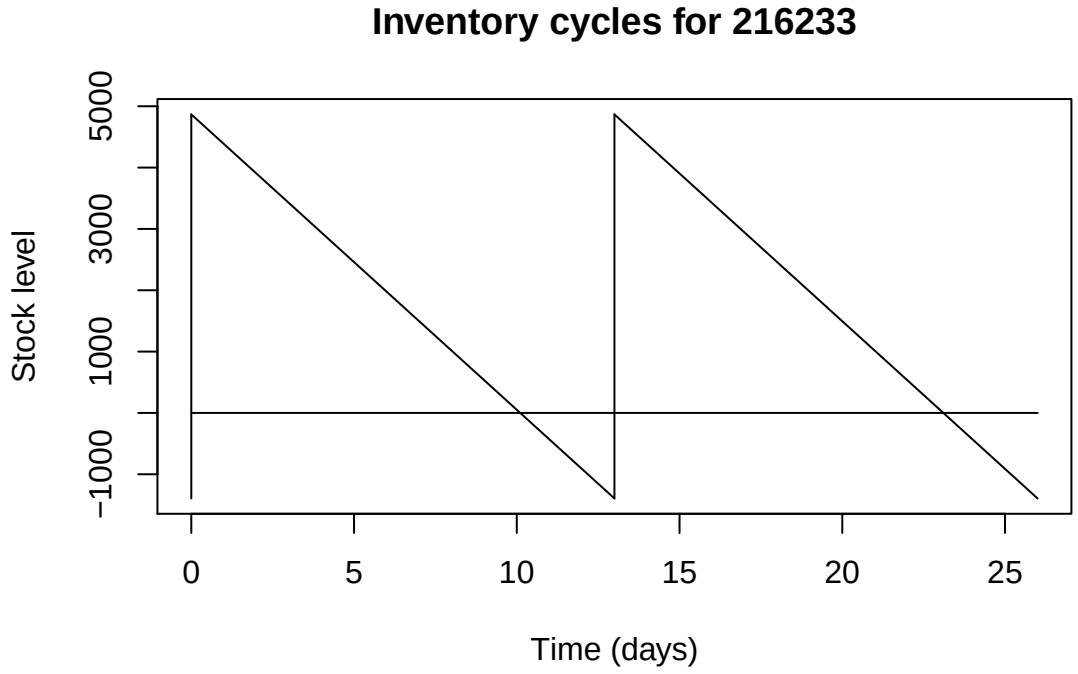
The Optimum Backorder Model is used to find the best order quantity so that total costs are minimised when backorders are allowed. Similar to the EOQ model assumptions are made that demand is constant and known, there is no lead time, orders arrive instantaneously and stock levels are under continuous review.

Annual demand is estimated as 169591 using 2012 data. Backorders cost is approximately 5% of the total price per unit. In the 2012 sales data total price for product 216233 varies from 78.4 to 134.7. For evaluating the backorder cost the mean total price 78.4 has been used, resulting in a backorder cost (p) of 6.22 per unit. The optimum quantity (Q^*) and optimum maximum inventory level (S^*) are calculated using the following forumlas.

$$Q^* = \sqrt{\frac{2kA}{h}}\sqrt{\frac{p+h}{p}}$$

$$S^* = \sqrt{\frac{2kA}{h}}\sqrt{\frac{p}{p+h}}$$

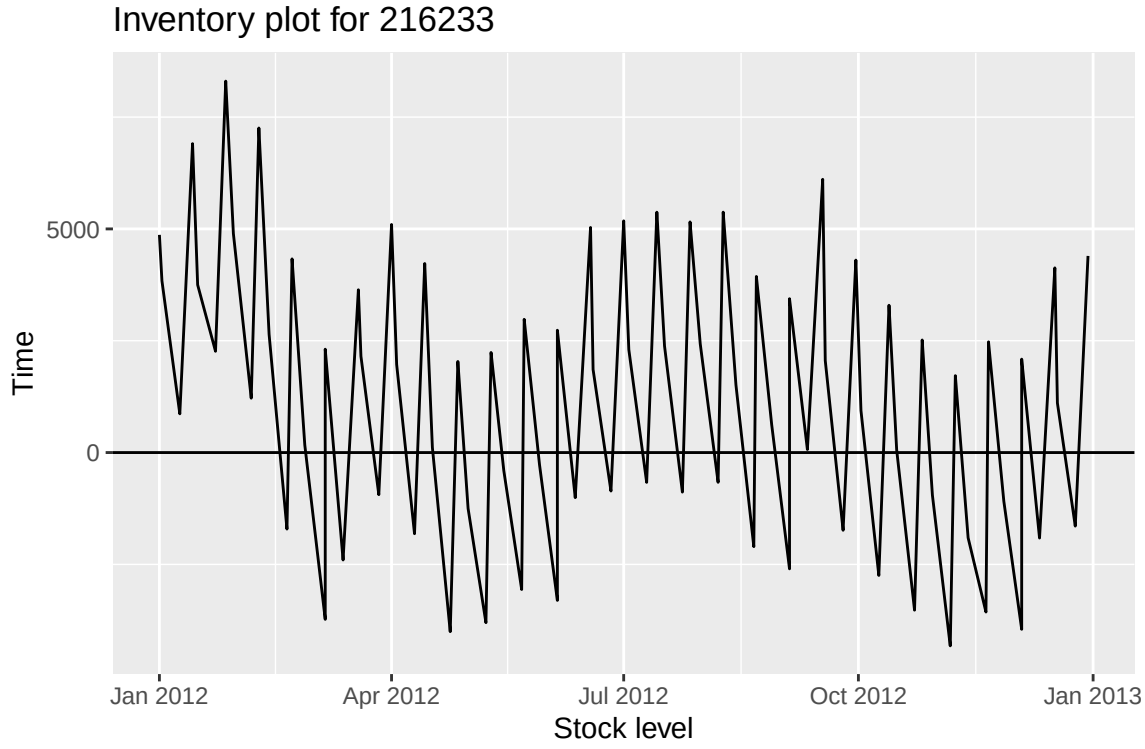
Optimal order quantity is calculated to be 6040 with an inventory cycle of 13 days. This means that every 13 days 6040 units are ordered, resulting in 28 annual orders. The optimum inventory level is 4867 and the proportion of time taking back orders is 23%. This model results in the smallest possible annual inventory cost of 7300.44 with an annual order cost of 3650.14 and an annual holding cost of 2941.35 and annual backorder cost of 709. This EOQ model is plotted for two cycles below.



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.

- 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 plot above shows the weekly demand from the 2012 sales data plotted with the optimum order frequency, 13 days and optimum order quantity 6040 from the Optimum Backorder Model. Stock starts at the optimum inventory level 4867, decreases with every weekly sale quantity from 2012 data and increases with inventory added at the optimum frequency and quantity.

Actual demand is not constant and as a result the quantity and frequency found with the back order model is not always appropriate. The model performs best when demand is constant, for example in July.

2b

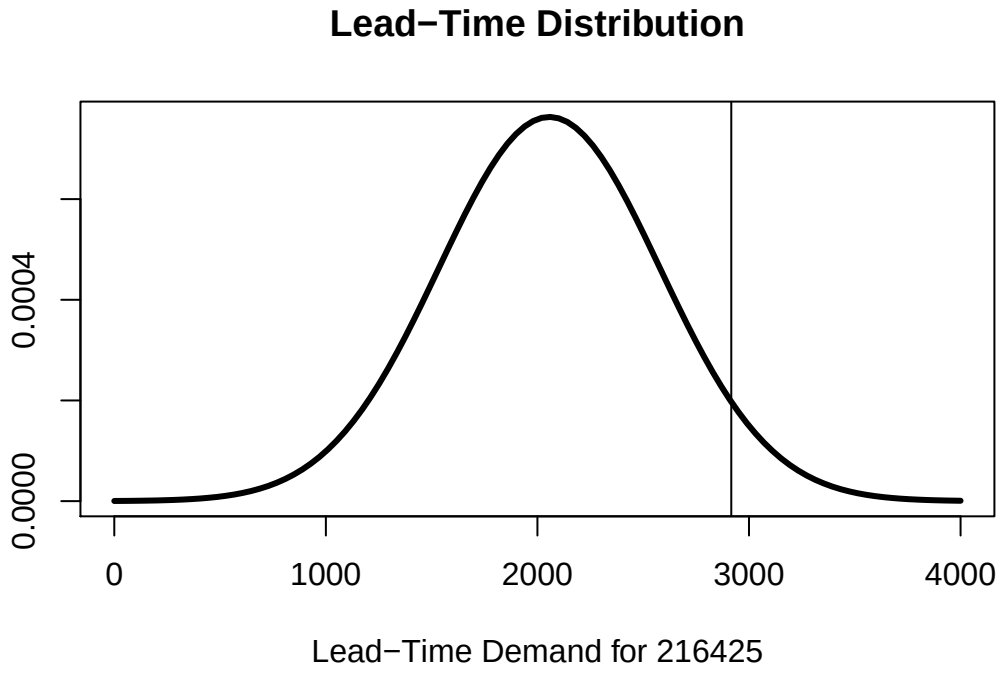
The Operations Manager is considering the option of a multi-period inventory model. The company, as a policy, is not willing to tolerate more than 5% chance of a stock-out. The Operations Manager has estimated that the annual holding cost is 6.50 per unit and the ordering cost is 20.50 per order.

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.

- 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.

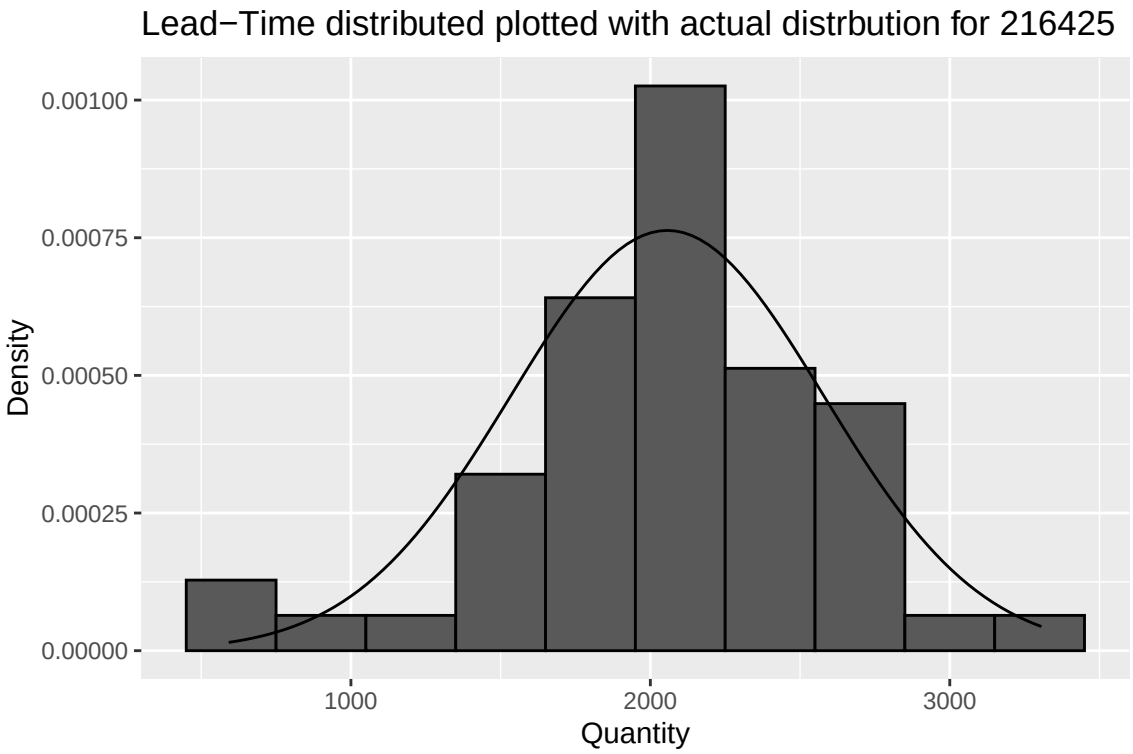
- The assumption of normality for the demand during a one-week lead time is discussed.
- The weekly average demand of this product is correctly plotted and discussed




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

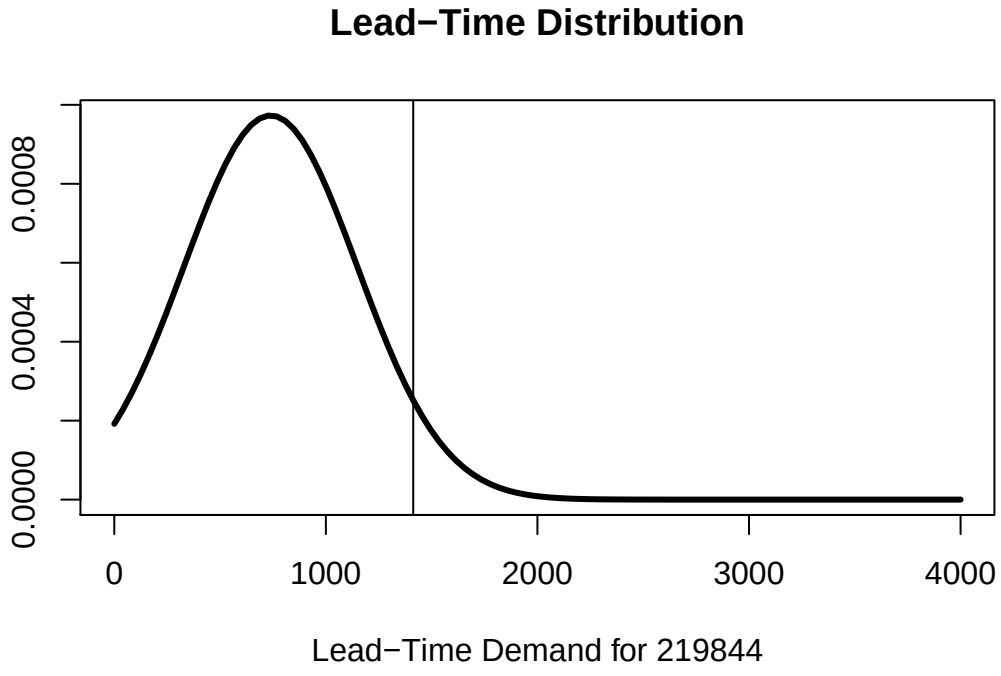
For this multi-inventory model demand during a one week lead time has been estimated using the mean and standard deviation of observed data in 2012. Demand has been estimated as a normal distribution with a mean of 2057 and standard deviation of 523. As the plot below shows, whilst the actual demand for 2012 does not perfectly follow this distribution it is a adequate approximation.

The expected annual demand is estimated to be 106939. 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 units. Approximately 130 orders will be placed per year and safety stock is 821. This approach ensures roughly 95% of the time stock will be sufficient for weekly demand. The expected annual costs are 10926.7492 per year. If demand was certain the annual costs would only be 5338.4686358 so the additional cost of holding safety stock is 5588.2805642.


Lead−Time distributed plotted with actual distrbution for 216425
0.00100
0.00075
0.00050
0.00025
0.00000
Density
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.

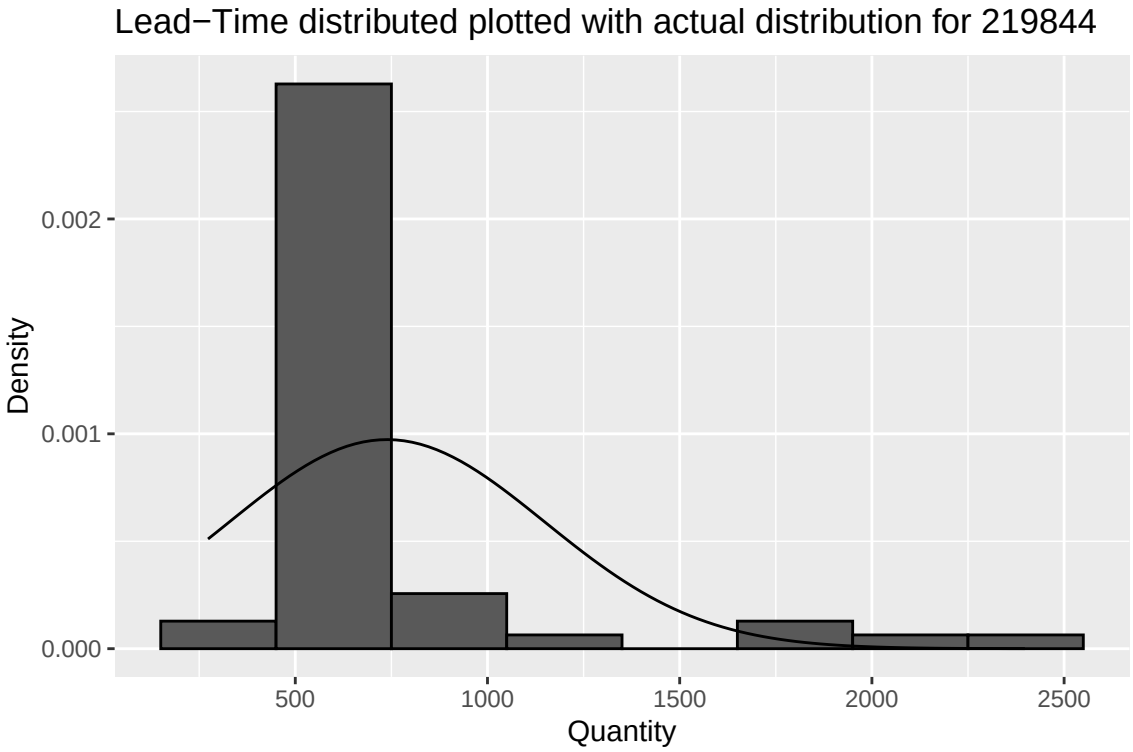
Discuss the assumptions of the model and suggest a solution, in case of finding any problems. Write a paragraph explaining the results of your findings and the plot.




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

For this multi-inventory model demand during a one week lead time has been estimated using the mean and standard deviation of observed data in 2012. Demand has been estimated as a normal distribution with a mean of 739 and standard deviation of 410. As the plot below shows this distribution is a very poor fit for the observed data. It is recommended that a more accurate distribution is used for estimating the Lead-Time distribution and reorder point. It is likely that the reorder point calculated with this normal distribution is unnecessarily high.

The expected annual demand is estimated to be 38410. 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 1413 units. Approximately 78 orders will be placed per year and safety stock is 492. This approach ensures roughly 95% of the time stock will be sufficient for weekly demand. The expected annual costs are 7582.3685882 per year. If demand was certain the annual costs would only be 3199.4166667 so the additional cost of holding safety stock is 4382.9519216.

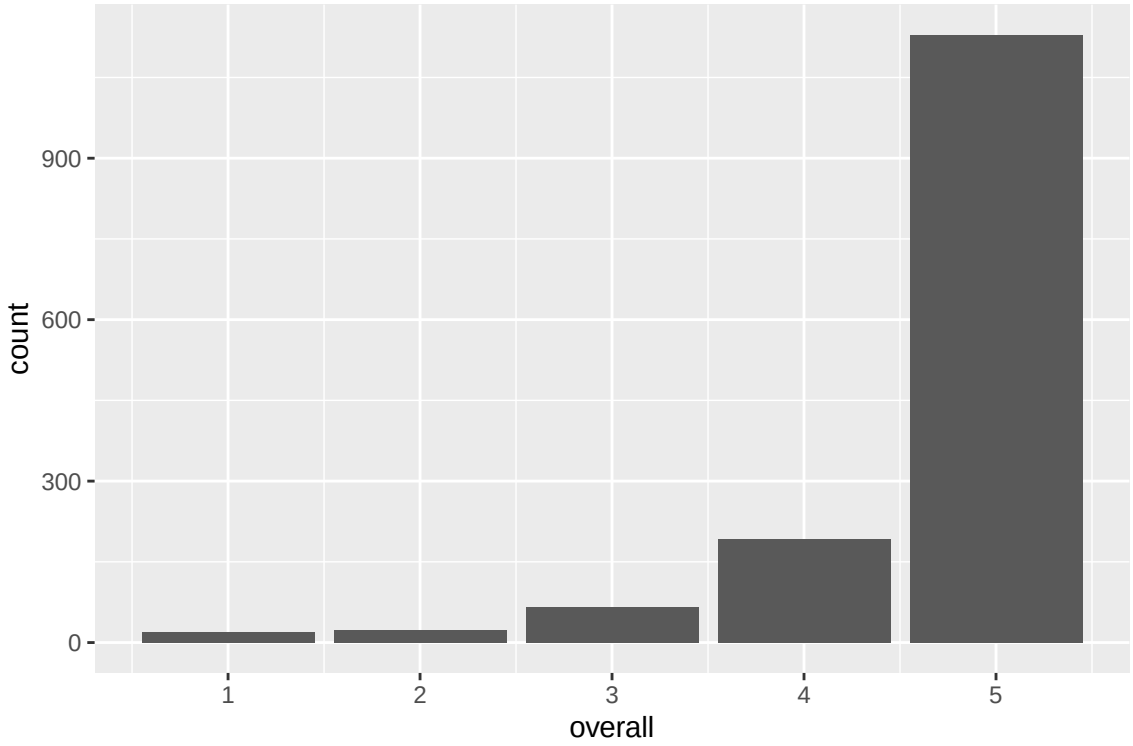


Question 3

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