Forecasting Supply-Chain Needs: A Regression-Based Approach

Introduction

Background

An FMCG company entered the instant noodles market two years ago. It has faced challenges with supply-demand mismatches in its warehouses across the country. In high-demand areas, supply is inadequate, leading to lost sales. In low-demand regions there is extra supply resulting in increased inventory costs. Due to this the company now needs an optimized supply chain model.

Research Question

Using historical data, how can the weight of product shipments being shipped to the warehouse each time be optimized in a way where inventory costs are minimized while regional demands are met?

Hypotheses

In this analysis there are several hypotheses to be tested. The analysis will check for a linear relationship between the independent variables and product weight. We will assess the impact of storage issues, analyzing if they have a significant effect on product weight or not. The model will also evaluate the influence of government checks and location on supply chain performance, checking if these factors improve the model. Additionally, the effect of electric supply on the model's accuracy determining if it improves the model will be tested. Finally, the analysis will test whether interaction terms between different factors add significant variation and improve the model's predictive power.

Description of Data

The initial dataset has 25,000 rows and 24 variables. Some of the important variables that are important to our model are

Location_type: Indicates whether the warehouse is in rural or urban area
WH_capacity_size: The storage capacity of the warehouse
zone: The geographical zone of the warehouse
WH_regional_zone: The specific regional zone under each larger zone
num_refill_req_l3m: Number of times the warehouse required refilling in the past 3 months
transport_issue_l1y: Any transportation issues like accidents or goods being stolen reported in the last year
retail_shop_num: Number of retail shops selling the product in the warehouse's area
wh_owner_type: Whether the warehouse is owned by the company or rented
distributor_num: Number of distributors working between the warehouse and retail shops
flood_impacted: Whether the warehouse is in a flood-impacted area
electric_supply: Whether the warehouse has electric backup like a generator
dist_from_hub: Distance from the warehouse to the production hub in kilometers
workers_num: Number of workers in the warehouse
storage_issue_reported_l3m: Number of storage issues (like rats or fungus due to moisture) reported in the last 3 months

govt_check_l3m: Number of times government officers have checked the warehouse in the last 3 months
product_wg_ton: The amount of product shipped from the warehouse(in tons) in the last 3 months

Exploratory Data Analysis (EDA)

Data Cleaning

- Removed null values and filtered rows with missing critical data, such as government certification statuses.
- Renamed columns such as product_wg_ton to product_weight and storage_issue_reported_l3m to storage_issues for better clarity and removed non-informative attributes like Ware_house_ID and WH_Manager_ID.

Variable Exploration

- The dataset includes sixteen numerical variables (e.g. distance_hub, workers_num, storage_issues, product_weight) and six categorical variables (e.g. location, capacity, zone, reg_zone, warehouse_owner, govt_cert).
- Pair plots, bar plots and histograms revealed significant variability in the distribution of product_weight. The interaction between product_weight and categorical variables like zone and location was explored.
- Correlation between select numerical variables and product_weight revealed that there was positive correlation in two cases suggesting possible transformations for modeling.

Regression Analysis

Final Regression Model

A multiple linear regression model with quadratic and interaction terms was developed with `product_weight` as the dependent variable and the following predictors:
w_est_year, govt_cert, storage_issues, transport_issue, temperature_regulation.

Multiple Linear Regression with interactions and Polynomial Terms (MLR with I&Q)

» $E(Y_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i1}^2 + \beta_6 X_{i2}^2 + \beta_7 (X_{i3} * X_{i4}) + \varepsilon_i$

`

```
updated_both_model_mc2 <- lm(product_weight ~ w_est_year + govt_cert + storage_issues + transport_issue +
                                temperature_regulation, data =train_df)
```

The model provides an Adjusted R-squared of 89.76%, which indicates that it explains 89.76% of the variance in the data, demonstrating a strong relationship between the independent variables and the product weight. Additionally, the BIC (Bayesian Information Criterion) for this model is -3980.228.

Rationale for Model Selection

- The inclusion of interaction and quadratic terms allows the model to capture more complex relationships between the predictors and the dependent variable.
- With the improved Adjusted R-squared of 89.76%, the model outperforms the simple linear regression (81.1%) and the MLR without interaction and quadratic terms (88.81%). The model's ability to capture relationships, which improved its predictive accuracy.

Interpretations and Interesting Findings from the Coefficients:

The final model had a F-statistic value of 2139 with p-value lesser than the significance level, which suggests the overall model is statistically significant.

Significant Predictors:

W_est_year i.e. Warehouse Established year shows a clear negative trend in product weight over time, also significant negative coefficients starts from 2007. Also, by 2022, the product weight decreased by about 0.85 units compared to the baseline year.

Government Certification (govt_cert): Compared to baseline levels certification B and B+ show significant negative effects on produce weights and certification C also shows a negative effect.

Storage and Transport (Storage and Transport Issues): Out of all, storage issues column have a strong positive effect with a weightage of 0.305 and more storage issues correlate with increased product weight, whereas higher transport issue levels correspond to lower product weights.

Temperature Regulation: Presence of temperature regulation slightly increases with product weight.

Confidence Intervals of Coefficients: Almost all of the coefficients have narrow confidence intervals, indicating precise estimates.

Model Assumptions and Fit analysis:

Observed vs. Predicted Values Plot:

The plot of observed vs. predicted values shows a clear linear relationship, with the data points clustered around the regression line and suggests that the model is a good fit for the data and capturing the underlying trend

QQ Plot of residuals: The respective plot shows that the residual generally follows a linear pattern, indicating the assumption of normality for the residuals is reasonably met but there are some outliers in the end of the line, overall residuals appears to be normally distributed.

Residuals vs Leverage Plot: The residuals vs. leverage plot shows a clear pattern, with residuals distributed evenly around the horizontal line. This suggests that the model is not overly influenced by any individual points and the leverage is not a significant issue.

Lack of fit-test: The lack of fit test F-statistic of 10.6 with a p-value of 0 indicates that model does not fit the data entirely but suggests that there may be some non-linearity in the data that the model is not capturing

Overall, model appears to be well fitted, with a good R-square value and residuals that are generally well-behaved. However there may be some non-linear patterns that the current model is not fully capturing.

To summarize,

- Linearity: The relationship between the predictors and the dependent variable is assumed to be linear.
- Independence of Errors: We assume that the residuals are independent. This assumption was checked using residual analysis which showed no signs of autocorrelation.
- Homoscedasticity: We also assume constant variance of the residuals across levels of the independent variables.
- Normality of Residuals: Shapiro-Wilk test and Q-Q plots's results indicated that the residuals are approximately normally distributed.
- Multicollinearity: To ensure that multicollinearity does not distort the regression estimates, the Variance Inflation Factors (VIFs) for the predictors were calculated and the predictors with VIF>2 were removed.

Model Fit Metrics:

Adjusted R-squared of 89.76% explains a substantial portion of the variance in the dependent variable, reflecting excellent model performance.
Residual standard error (RSE) of 0.194 aligns with low prediction errors, and residual plots indicate no significant patterns, supporting the assumptions of homoscedasticity and normality.
The extremely high F-statistic (2139) and a low p-value < 2.2e-16 confirm the overall significance of the model.
BIC = -3980.228 which means that the model strikes an optimal balance between complexity and fit.

Discussion

The regression-based approach employed in this project demonstrates significant potential in optimizing supply chain performance for the FMCG instant noodles company. The findings indicate strong correlations, such as between product weight and storage issues, which offer actionable insights into warehouse management. The integration of interaction terms and quadratic variables in the MLR models enhanced the model's explanatory power, yielding an adjusted R-squared of 88.85%. The systematic data cleaning, transformation, and encoding processes ensured the model leveraged reliable and meaningful features.

The selected model—MLR with interaction and quadratic terms—proved to be the most effective, achieving robust predictive capabilities while maintaining interpretability. This demonstrates the importance of considering complex relationships between predictors in supply chain contexts. Moreover, the use of validation techniques, such as k-cross validation, ensured the model's generalizability to unseen data.

These findings provide the supply chain and logistics teams with a strong foundation for demand prediction and inventory optimization. The practical implications include reducing excess stock, minimizing shortages, and ultimately improving customer satisfaction and cost efficiency.

Limitations

1. Data Limitations:

    - The dataset initially contained null values and non-meaningful column names, which may have led to the loss of potentially valuable information during the cleaning process.

    - The data may not fully represent all operational conditions, such as seasonal demand variations or abrupt market changes.

2. Model Assumptions:

    - The regression models assume linearity, independence of errors, and constant variance, which might not perfectly align with real-world supply chain complexities.
    - Despite steps to reduce multicollinearity, some residual dependencies between predictors may still exist, potentially impacting coefficient interpretability.

3. External Factors:

    - Factors like sudden transportation disruptions, geopolitical events, or shifts in consumer preferences were not explicitly accounted for, limiting the model's adaptability to such scenarios.

4. Feature Limitations:

    - Although interaction and quadratic terms improved performance, they increased model complexity, which may challenge interpretability for non-technical stakeholders.
    - The reliance on VIF for multicollinearity detection may not fully capture all issues, especially with a high-dimensional dataset.

5. Future Validation:

    - The model requires regular re-validation with updated data to maintain its relevance and accuracy in dynamically evolving market conditions.
    - Incorporating additional data, such as real-time logistics or demand forecasting metrics, could further enhance predictive power.

Conclusion:
This report presents an analysis aimed at optimizing the supply chain performance of an FMCG company in the instant noodles market, addressing supply-demand mismatches in warehouses. Using a historical dataset of 25,000 rows and 24 variables, a multiple linear regression (MLR) model with interaction and quadratic terms was developed. Key predictors included storage issues, government checks, and transportation factors. Data cleaning and transformation ensured meaningful insights, while exploratory analysis highlighted critical relationships, such as the impact of warehouse capacity and location on product shipment weight.

The final model demonstrated robust performance, achieving an Adjusted R-squared of 89.76% and explaining 88.85% of the variation in product weight, with a consistent MSPR/MSE ratio of 1.022. Narrow confidence intervals reflected reliable and stable estimates of key coefficients, while precise prediction intervals indicated strong robustness in model predictions. Residual diagnostics confirmed the validity of key regression assumptions, and the low BIC score highlighted the model's balance between complexity and fit. Significant findings included the effects of storage issues and transportation challenges on product weight, providing actionable strategies to optimize inventory management by balancing supply and demand, reducing excess stock, and minimizing shortages.

Despite its strong performance, the analysis has limitations, including potential biases from data cleaning, unaccounted seasonal variations, and reliance on linear assumptions that may not fully capture real-world complexities. Future improvements could involve incorporating additional data sources and regularly re-validating the model to maintain accuracy. Overall, this study offers a solid foundation for enhancing inventory management, improving customer satisfaction, and achieving operational efficiency in a competitive market.

Additional Work:

In addition to the multiple linear regression (MLR) model, alternative modeling approaches were explored to optimize the supply chain performance for the FMCG company. A General linear regression model was initially tested to capture potential non-linear relationships and interactions between variables without requiring explicit transformations. While the decision tree model provided some interpretability in identifying key decision paths, it underperformed compared to the MLR model, with an Adjusted R-squared of 82.4%, and tended to overfit the data due to the high dimensionality of the predictors.

Ultimately, while these models offered alternative perspectives and predictive capabilities, the MLR model with interaction and quadratic terms was selected for its superior interpretability, simplicity, and comparable performance, making it more aligned with the project's goals of delivering actionable insights for supply chain optimization.

Additionally, there was a plan to include a contour plot, but in the end we ultimately scrapped this plan because the plot was difficult to interpret and other plots could do a better job as conveying the information.

Appendix:

Code Output:

Initial data structure

```
> str(df)
'data.frame':	25000 obs. of  24 variables:
 $ Ware_house_ID               : chr  "WH_100000" "WH_100001" "WH_100002" "WH_100003" ...
 $ WH_Manager_ID               : chr  "EID_50000" "EID_50001" "EID_50002" "EID_50003" ...
 $ Location_type               : chr  "Urban" "Rural" "Rural" "Rural" ...
 $ WH_capacity_size            : chr  "Small" "Large" "Mid" "Mid" ...
 $ zone                        : chr  "West" "North" "South" "North" ...
 $ WH_regional_zone            : chr  "Zone 6" "Zone 5" "Zone 2" "Zone 3" ...
 $ num_refill_req_l3m          : int  3 0 1 7 3 8 8 1 8 4 ...
 $ transport_issue_l1y         : int  1 0 0 4 1 0 0 0 1 3 ...
 $ Competitor_in_mkt           : int  2 4 4 2 2 2 4 4 4 3 ...
 $ retail_shop_num             : int  4651 6217 4306 6000 4740 5053 4449 7183 5381 3869 ...
 $ wh_owner_type               : chr  "Rented" "Company Owned" "Company Owned" "Rented" ...
 $ distributor_num             : int  24 47 64 50 42 37 38 45 42 35 ...
 $ flood_impacted              : int  0 0 0 0 1 0 0 0 0 0 ...
 $ flood_proof                 : int  1 0 0 0 0 0 0 0 0 0 ...
 $ electric_supply             : int  1 1 0 0 1 1 1 0 1 0 ...
 $ dist_from_hub               : int  91 210 161 103 112 152 77 241 124 78 ...
 $ workers_num                 : num  29 31 37 21 25 35 27 23 22 43 ...
 $ wh_est_year                 : num  NA NA NA NA 2009 ...
 $ storage_issue_reported_l3m  : int  13 4 17 10 18 17 32 19 15 7 ...
 $ temp_reg_mach               : int  0 0 0 1 0 1 0 0 1 0 ...
 $ approved_wh_govt_certificate: chr  "A" "A" "A" "A+" ...
 $ wh_breakdown_l3m            : int  5 3 6 3 6 3 3 6 5 6 ...
 $ govt_check_l3m              : int  15 17 22 27 24 3 6 24 2 2 ...
 $ product_wg_ton              : int  17115 5074 23137 22115 24071 32134 30142 24093 18082 7130 ...
> 
```

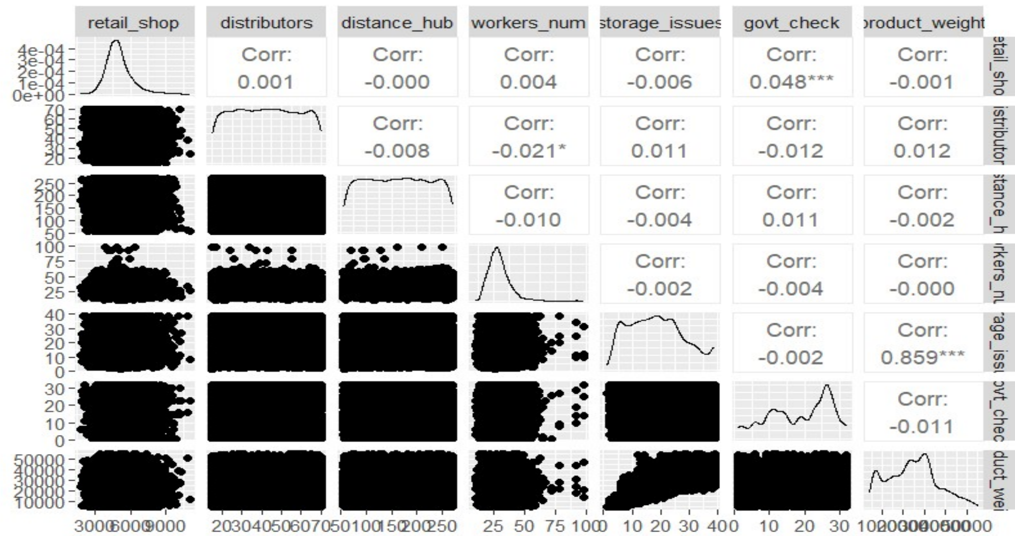
Before and after removing null values

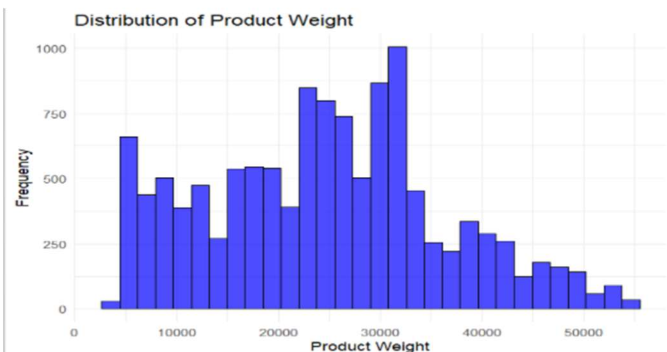
```
> colSums(is.na(df))
           Ware_house_ID                WH_Manager_ID                Location_type 
                       0                            0                            0 
        WH_capacity_size                         zone             WH_regional_zone 
                       0                            0                            0 
      num_refill_req_l3m          transport_issue_l1y            Competitor_in_mkt 
                       0                            0                            0 
         retail_shop_num                wh_owner_type              distributor_num 
                       0                            0                            0 
          flood_impacted                  flood_proof              electric_supply 
                       0                            0                            0 
           dist_from_hub                  workers_num                  wh_est_year 
                       0                          990                        11881 
storage_issue_reported_l3m                temp_reg_mach approved_wh_govt_certificate 
                       0                            0                            0 
       wh_breakdown_l3m               govt_check_l3m               product_wg_ton 
                       0                            0                            0 
> 
> colSums(is.na(df))
               Location_type             WH_capacity_size                         zone 
                           0                            0                            0 
            WH_regional_zone           num_refill_req_l3m          transport_issue_l1y 
                           0                            0                            0 
           Competitor_in_mkt              retail_shop_num                wh_owner_type 
                           0                            0                            0 
             distributor_num               flood_impacted                  flood_proof 
                           0                            0                            0 
             electric_supply                dist_from_hub                  workers_num 
                           0                            0                            0 
                 wh_est_year   storage_issue_reported_l3m                temp_reg_mach 
                           0                            0                            0 
approved_wh_govt_certificate             wh_breakdown_l3m               govt_check_l3m 
                           0                            0                            0 
              product_wg_ton 
                           0 
> 
```

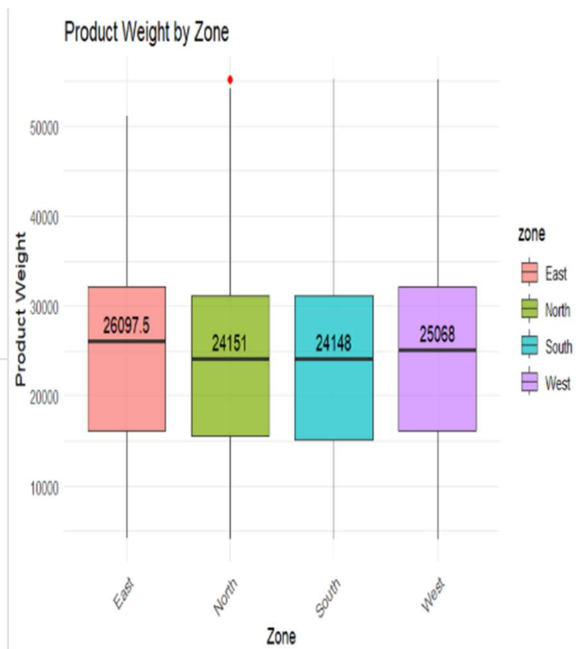
Dropping the ID columns and changing the column names

```
> str(df)
'data.frame':	12127 obs. of  22 variables:
 $ location              : chr  "Rural" "Rural" "Rural" "Rural" ...
 $ capacity              : chr  "Large" "Small" "Large" "Small" ...
 $ zone                  : chr  "North" "West" "West" "South" ...
 $ reg_zone              : chr  "Zone 5" "Zone 1" "Zone 6" "Zone 6" ...
 $ refill                : int  3 8 8 8 7 7 4 6 4 8 ...
 $ transport_issue       : int  1 0 0 1 1 0 0 1 1 0 ...
 $ competitor            : int  2 2 4 4 3 5 3 2 4 2 ...
 $ retail_shop           : int  4740 5053 4449 5381 4623 4627 5012 6858 4598 5678 ...
 $ warehouse_Owner       : chr  "Company Owned" "Rented" "Company Owned" "Rented" ...
 $ distributors          : int  42 37 38 42 31 40 48 26 58 31 ...
 $ flood_impacted        : int  1 0 0 0 0 0 0 0 0 0 ...
 $ flood_proof           : int  0 0 0 0 0 0 0 0 0 0 ...
 $ electric_supply       : int  1 1 1 1 1 0 0 1 1 1 ...
 $ distance_hub          : int  112 152 77 124 150 225 95 242 159 65 ...
 $ workers_num           : num  25 35 27 22 37 16 28 36 22 41 ...
 $ w_est_year            : num  2009 2009 2010 2013 1999 ...
 $ storage_issues        : int  18 17 32 15 17 11 4 22 36 11 ...
 $ temperature_regulation: int  0 1 0 1 0 0 0 1 1 0 ...
 $ govt_cert             : chr  "C" "A+" "B" "A+" ...
 $ warehouse_breakdown   : int  6 3 3 5 4 2 1 5 5 4 ...
 $ govt_check            : int  24 3 6 2 6 28 1 11 27 1 ...
 $ product_weight        : int  24071 32134 30142 18082 21125 14115 5124 30063 38082 24062 ...
> 
> cat("Number of numerical variables:", num_numerical, "\n")
Number of numerical variables: 16 
> cat("Number of categorical variables:", num_categorical, "\n")
Number of categorical variables: 6 
> 
```

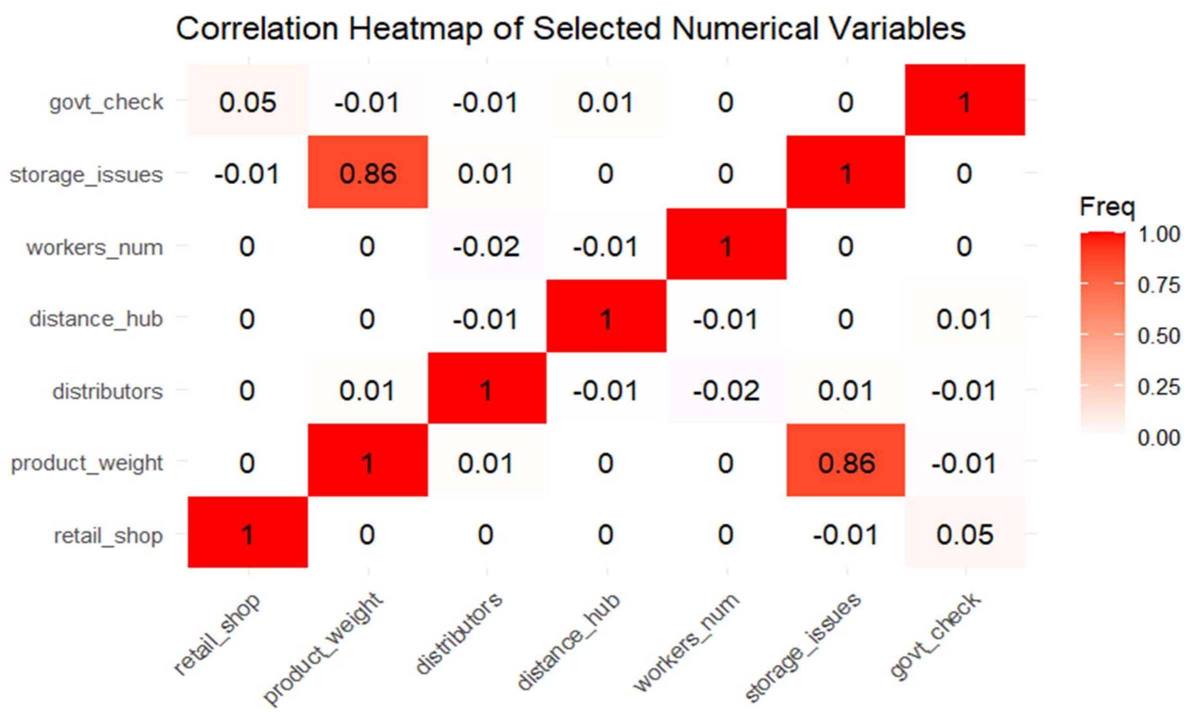
EDA:


retail_shop
distributors
distance_hub
workers_num
storage_issues
govt_check
product_weight
Corr: 0.001
Corr: -0.000
Corr: 0.004
Corr: -0.006
Corr: 0.048***
Corr: -0.001
Corr: -0.008
Corr: -0.021*
Corr: 0.011
Corr: -0.012
Corr: 0.012
Corr: -0.010
Corr: -0.004
Corr: 0.011
Corr: -0.002
Corr: -0.002
Corr: -0.004
Corr: -0.000
Corr: -0.002
Corr: 0.859***
Corr: -0.011


Distribution of Product Weight
Frequency
Product Weight


Product Weight by Zone
Product Weight
26097.5
24151
24148
25068
zone
East
North
South
West
Zone


Average Product Weight by Warehouse Location
24352.95
26736.55
Average Product Weight
location
Rural
Urban
Warehouse Location



Final Model:

Summary:

```
Residual standard error: 0.194 on 9664 degrees of freedom
Multiple R-squared:  0.8885,	Adjusted R-squared:  0.8881
F-statistic:  2139 on 36 and 9664 DF,  p-value: < 2.2e-16
```

Anova Table:

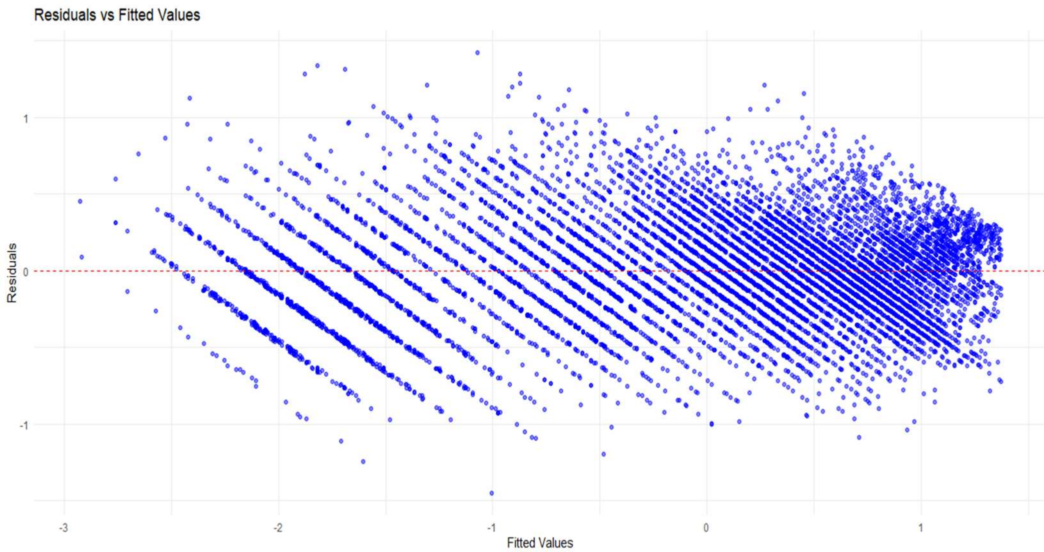
```
> anova(updated_both_model_mc2)
Analysis of Variance Table

Response: product_weight
                         Df  Sum Sq Mean Sq  F value    Pr(>F)    
w_est_year               26 2546.34  97.936 2603.524 < 2.2e-16 ***
govt_cert                 4   34.39   8.598  228.571 < 2.2e-16 ***
storage_issues            1  305.26 305.255 8114.867 < 2.2e-16 ***
transport_issue           4    9.24   2.311   61.431 < 2.2e-16 ***
temperature_regulation    1    0.84   0.837   22.250 2.428e-06 ***
Residuals              9664  363.53   0.038                       
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

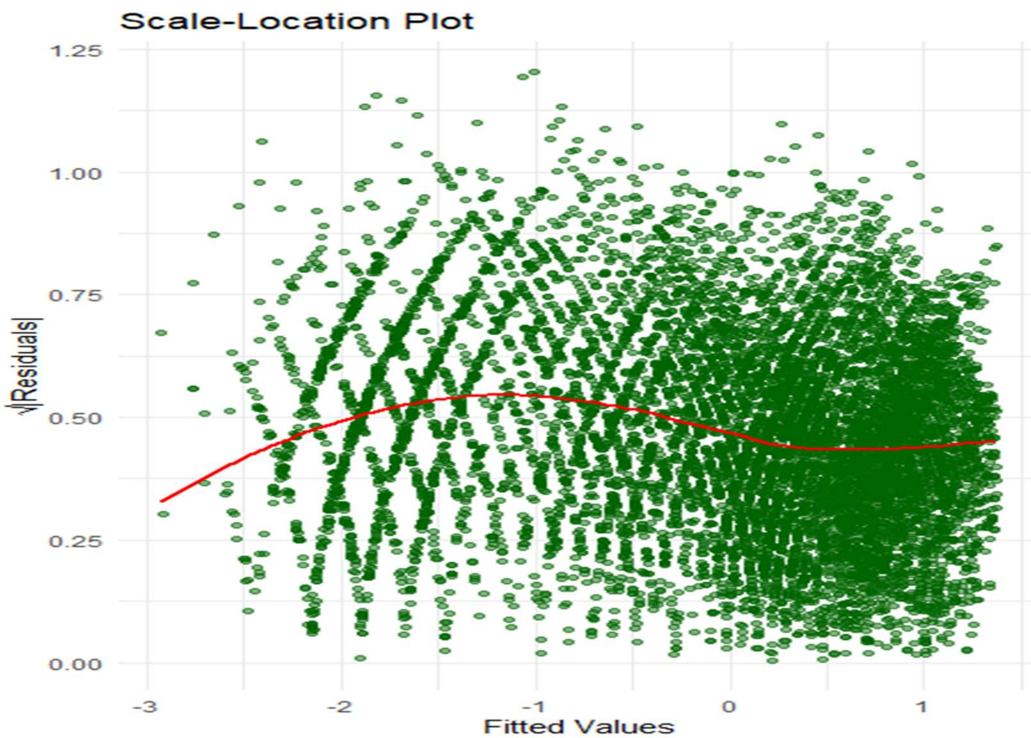
Multicollinearity:

```
> vif(updated_both_model_mc2)
                           GVIF Df GVIF^(1/(2*Df))
w_est_year             3.270324 26        1.023048
govt_cert              1.555713  4        1.056796
storage_issues         3.102658  1        1.761436
transport_issue        1.026658  4        1.003294
temperature_regulation 1.456053  1        1.206670
```

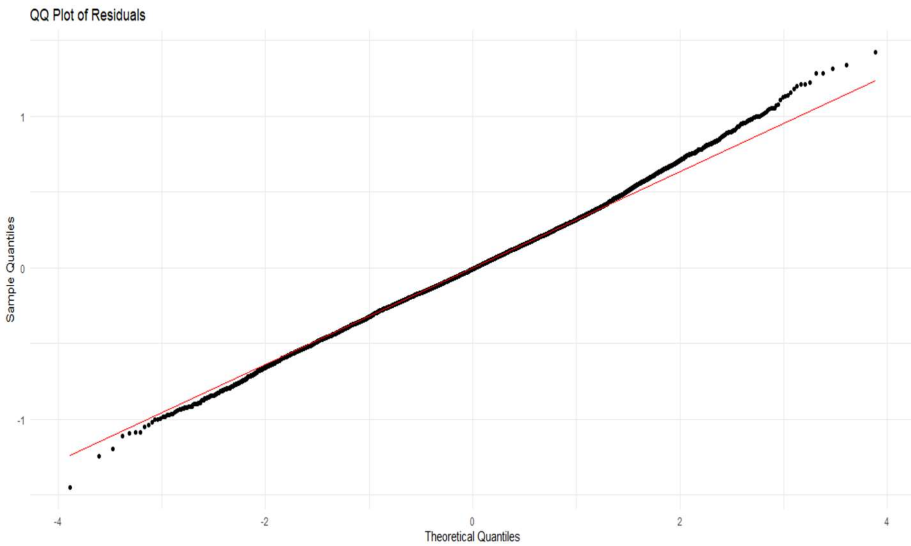
Residual Plots:



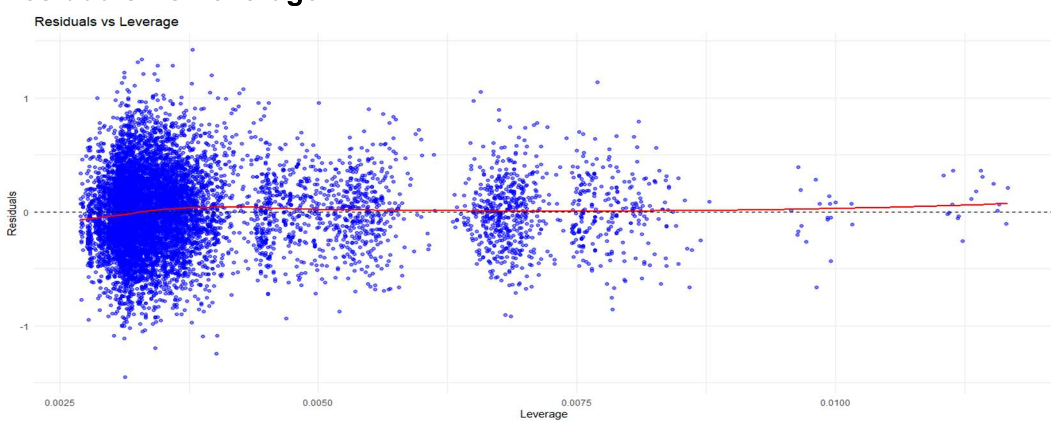
Scale-Location:

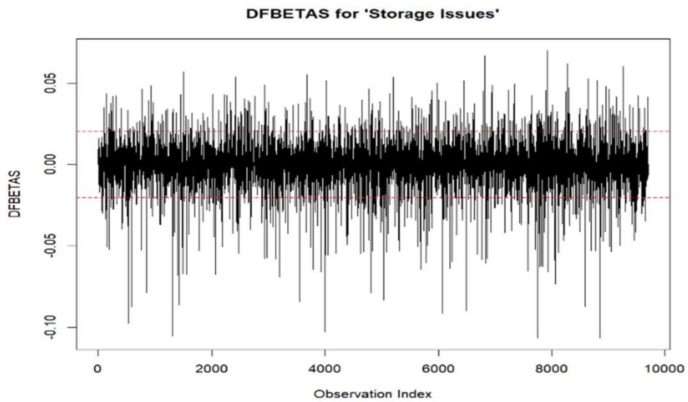


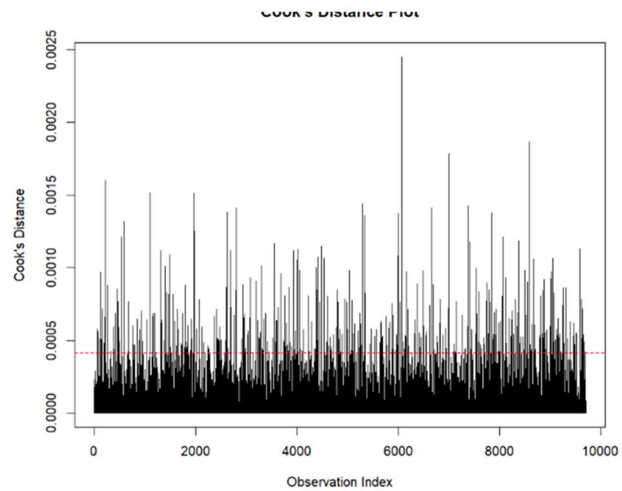
Q-Q Plot:

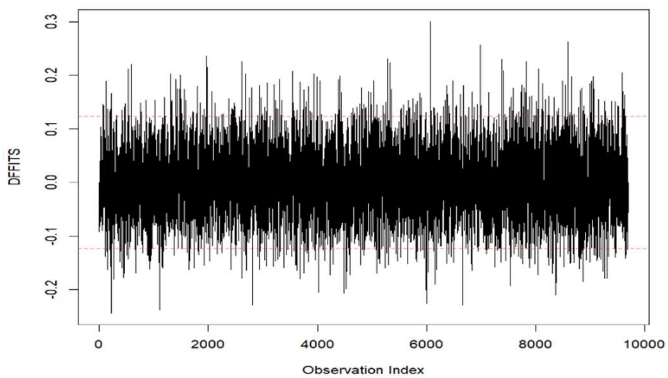


Residuals Vs Leverage:










```
Linear Regression

12127 samples
    5 predictor

No pre-processing
Resampling: Cross-Validated (5 fold)
Summary of sample sizes: 9701, 9700, 9702, 9702, 9703
Resampling results:

  RMSE       Rsquared   MAE
  0.1947053  0.8874002  0.1530196
```

The MSPR value and R-square value of test data is 0.03829644 and 0.8868954

The ratio of MSPR and MSE is: 1.021966

Inference in Regression Analysis:

```
> print(conf_intervals)
                                     Coefficient     Estimate  Lower_95_CI   Upper_95_CI
(Intercept)                          (Intercept) 10.257407831 10.224011222 10.2908044400
w_est_year1997                    w_est_year1997 -0.004355300 -0.044139953  0.0354293534
w_est_year1998                    w_est_year1998  0.008010816 -0.029095884  0.0451175159
w_est_year1999                    w_est_year1999  0.012617894 -0.024596558  0.0498323460
w_est_year2000                    w_est_year2000  0.016378655 -0.020238453  0.0529957629
w_est_year2001                    w_est_year2001  0.001709980 -0.035414332  0.0388342922
w_est_year2002                    w_est_year2002 -0.006838073 -0.043855797  0.0301796507
w_est_year2003                    w_est_year2003 -0.003757591 -0.041205335  0.0336901537
w_est_year2004                    w_est_year2004  0.020890102 -0.016255381  0.0580355851
w_est_year2005                    w_est_year2005  0.003440587 -0.034220861  0.0411020346
w_est_year2006                    w_est_year2006 -0.035805483 -0.072602676  0.0009917103
w_est_year2007                    w_est_year2007 -0.111135743 -0.148102476 -0.0741690097
w_est_year2008                    w_est_year2008 -0.162147227 -0.199687966 -0.1246064892
w_est_year2009                    w_est_year2009 -0.182804207 -0.220530478 -0.1450779363
w_est_year2010                    w_est_year2010 -0.177422926 -0.214791284 -0.1400545687
w_est_year2011                    w_est_year2011 -0.175330558 -0.212822288 -0.1378388274
w_est_year2012                    w_est_year2012 -0.241695438 -0.279178942 -0.2042119343
w_est_year2013                    w_est_year2013 -0.297396082 -0.335353959 -0.2594382051
w_est_year2014                    w_est_year2014 -0.342639799 -0.380660858 -0.3046187399
w_est_year2015                    w_est_year2015 -0.340294635 -0.378514590 -0.3020746807
w_est_year2016                    w_est_year2016 -0.414797171 -0.453286804 -0.3763075377
w_est_year2017                    w_est_year2017 -0.502793043 -0.541835121 -0.4637509648
w_est_year2018                    w_est_year2018 -0.608203669 -0.647995597 -0.5684117402
w_est_year2019                    w_est_year2019 -0.701686722 -0.741718000 -0.6616554440
w_est_year2020                    w_est_year2020 -0.752543095 -0.793526154 -0.7115600359
w_est_year2021                    w_est_year2021 -0.817342081 -0.861098411 -0.7735857518
w_est_year2022                    w_est_year2022 -0.849939096 -0.898736501 -0.8011416917
govt_certA+                          govt_certA+  0.011089449 -0.003573829  0.0257527279
govt_certB                            govt_certB -0.117898904 -0.130128683 -0.1056691257
govt_certB+                          govt_certB+ -0.105981699 -0.118264515 -0.0936988836
govt_certC                            govt_certC -0.036786122 -0.048966769 -0.0246054744
storage_issues                    storage_issues  0.305306369  0.298504182  0.3121085572
transport_issue1                transport_issue1 -0.044119312 -0.054515531 -0.0337230921
transport_issue2                transport_issue2 -0.074750500 -0.093833292 -0.0556677078
transport_issue3                transport_issue3 -0.124880136 -0.148633758 -0.1011265132
transport_issue4                transport_issue4 -0.097039647 -0.120795679 -0.0732836148
temperature_regulation1 temperature_regulation1  0.022787837  0.013317928  0.0322577454
```

```
The 95% confidence interval for E(X_h) is:
 8.7814 <= E(X_h) <= 8.7917
```

```
The 95% prediction interval for New Observation is:
 8.4054 <= Y_h_new <= 9.1677
```

```
The 95% prediction interval for New Observation is:
 8.7598 <= mean of m obsr <= 8.8133
```


```
The 95% confidence interval for E(X_h) is:
 8.7814 <= E(X_h) <= 8.7917
```