Garment Factory Productivity Model

Anil Raju

2023-07-26

I. Introduction : Data and Variables

The dataset utilized for this study is not owned by the author but rather borrowed from the Kaggle database for training and educational purposes. The dataset *"productivity-prediction-of-garment-employees.csv"* contains recorded data for a specific time period, encompassing various variables and a productivity measure, representing a garment factory. As with any factory, identifying the crucial variables and their interactions that influence productivity levels holds significant value. This notebook will conduct exploratory data analysis and linear regression analysis on the dataset to construct an interpretable productivity model (not a prediction or time series model). Additionally, we will address a few questions that arise after the exploratory analysis using hypothesis testing.

actual_productivity - measures the actual productivity of each team (in percentage, but represented as a range between 0 and 1). It's a unit less variable where value of 1 represents 100% productivity from the team. This is the **'Response Variable'** for our models.

The remaining variables are considered as predictor variables for the model.
Numerical Variables:

- targeted_productivity - targeted productivity set by the supervisor for each team for each day(in percentage, but recorded as a range between 0-1)
- smv - standard time allocated for a task(in minutes)
- wip - number of unfinished items for products(number)
- over_time - amount of overtime allocated to each team (in minutes)
- incentive - amount of financial incentive that enables or motivates the team(in BDT)
- idle_time - amount of time when the production was interrupted due to several reasons(in minutes)
- idle_men - number of workers who were idle due to production interruption(number)
- no_of_workers - number of workers in each team(number)

Categorical Variables:

- quarter - portion of the month; a month was divided into five quarters(5 levels)
- department - associated department(2 levels)
- day - days of the week(6 levels)
- team - team identifier (12 levels)
- no_of_style_change - number of changes in the style of a particular product(3 levels)

Since the data is unfamiliar I have no a-priori hypothesis or questions.

```
##
##  COUNT OF NA VALUES(TRUE) IN THE DATA FRAME

##     date          quarter        department         day
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:1197      FALSE:1197      FALSE:1197      FALSE:1197
##
##     team         targeted_productivity    smv             wip
##  Mode :logical   Mode :logical         Mode :logical   Mode :logical
##  FALSE:1197      FALSE:1197            FALSE:1197      FALSE:691
##                                                        TRUE :506
##  over_time       incentive       idle_time        idle_men
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:1197      FALSE:1197      FALSE:1197      FALSE:1197
##
##  no_of_style_change no_of_workers   actual_productivity
##  Mode :logical      Mode :logical   Mode :logical
##  FALSE:1197         FALSE:1197      FALSE:1197
##
```

We noticed that only the 'wip' variable contains NA (missing) values, accounting for approximately 42% of the variable's data. As we cannot ascertain whether the NA cells should be treated as having a value of 0, we have decided not to alter the existing values. Imputing the missing data is not feasible due to the high count/percentage of NA values, and knnimpute functions are converting them back to NA. Therefore, it is advisable to remove the 'wip' variable from the dataset.

As we are not conducting a time series analysis with the dataset, we will exclude the 'date' variable from our analysis.

```
##
##  CONTENTS OF THE DATA FRAME

## 'data.frame':    1197 obs. of  13 variables:
##  $ quarter              : chr  "Quarter1" "Quarter1" "Quarter1" "Quarter1" ...
##  $ department           : chr  "sweing" "finishing" "sweing" "sweing" ...
##  $ day                  : chr  "Thursday" "Thursday" "Thursday" "Thursday" ...
##  $ team                 : int  8 1 11 12 6 7 2 3 2 1 ...
##  $ targeted_productivity: num  0.8 0.75 0.8 0.8 0.8 0.8 0.75 0.75 0.75 0.75 ...
##  $ smv                  : num  26.16 3.94 11.41 11.41 25.9 ...
##  $ over_time            : int  7080 960 3660 3660 1920 6720 960 6900 6000 6900 ...
##  $ incentive            : int  98 0 50 50 50 38 0 45 34 45 ...
##  $ idle_time            : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ idle_men             : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ no_of_style_change   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ no_of_workers        : num  59 8 30.5 30.5 56 56 8 57.5 55 57.5 ...
##  $ actual_productivity  : num  0.941 0.886 0.801 0.801 0.8 ...
```

Notice the data type for the categorical variables, it doesn't have categorical data type in the data set, lets change that.

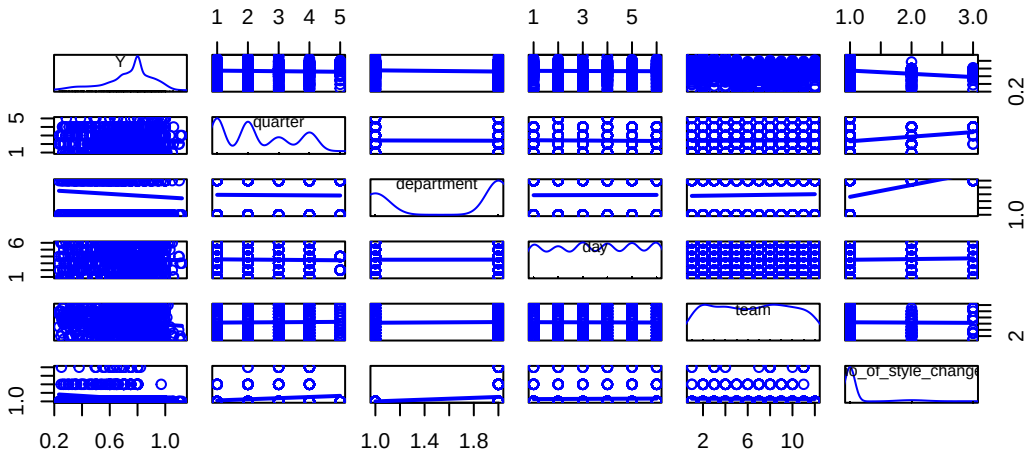
```
##
##  CONTENTS OF THE DATA FRAME

## 'data.frame':    1197 obs. of  13 variables:
##  $ quarter              : Factor w/ 5 levels "Quarter1","Quarter2",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ department           : Factor w/ 2 levels "finishing","sweing": 2 1 2 2 2 2 1 2 2 2 ...
##  $ day                  : Factor w/ 6 levels "Monday","Saturday",..: 4 4 4 4 4 4 4 4 4 4 ...
##  $ team                 : Factor w/ 12 levels "1","2","3","4",..: 8 1 11 12 6 7 2 3 2 1 ...
##  $ targeted_productivity: num  0.8 0.75 0.8 0.8 0.8 0.8 0.75 0.75 0.75 0.75 ...
##  $ smv                  : num  26.16 3.94 11.41 11.41 25.9 ...
##  $ over_time            : int  7080 960 3660 3660 1920 6720 960 6900 6000 6900 ...
##  $ incentive            : int  98 0 50 50 50 38 0 45 34 45 ...
##  $ idle_time            : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ idle_men             : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ no_of_style_change   : Factor w/ 3 levels "0","1","2": 1 1 1 1 1 1 1 1 1 1 ...
##  $ no_of_workers        : num  59 8 30.5 30.5 56 56 8 57.5 55 57.5 ...
##  $ actual_productivity  : num  0.941 0.886 0.801 0.801 0.8 ...

##
##  TOP FEW ROWS OF THE DATA FRAME

##    quarter department      day team targeted_productivity   smv over_time
## 1 Quarter1     sweing Thursday    8                  0.80 26.16      7080
## 2 Quarter1  finishing Thursday    1                  0.75  3.94       960
## 3 Quarter1     sweing Thursday   11                  0.80 11.41      3660
## 4 Quarter1     sweing Thursday   12                  0.80 11.41      3660
## 5 Quarter1     sweing Thursday    6                  0.80 25.90      1920
## 6 Quarter1     sweing Thursday    7                  0.80 25.90      6720
##   incentive idle_time idle_men no_of_style_change no_of_workers
## 1        98         0        0                  0          59.0
## 2         0         0        0                  0           8.0
## 3        50         0        0                  0          30.5
## 4        50         0        0                  0          30.5
## 5        50         0        0                  0          56.0
## 6        38         0        0                  0          56.0
##   actual_productivity
## 1           0.9407254
## 2           0.8865000
## 3           0.8005705
## 4           0.8005705
## 5           0.8003819
## 6           0.8001250

##
##  SUMMARY STATISTICS FOR EACH COLUMN IN THE DATA FRAME

##       quarter        department          day            team
##  Quarter1:360   finishing:506   Monday   :199   2      :109
##  Quarter2:335   sweing   :691   Saturday :187   8      :109
##  Quarter3:210                   Sunday   :203   1      :105
##  Quarter4:248                   Thursday :199   4      :105
##  Quarter5: 44                   Tuesday  :201   9      :104
##                                 Wednesday:208   10     :100
```

```
##                                                (Other):565
##  targeted_productivity      smv          over_time       incentive
##  Min.   :0.0700        Min.   : 2.90   Min.   :    0   Min.   :   0.00
##  1st Qu.:0.7000        1st Qu.: 3.94   1st Qu.: 1440   1st Qu.:   0.00
##  Median :0.7500        Median :15.26   Median : 3960   Median :   0.00
##  Mean   :0.7296        Mean   :15.06   Mean   : 4567   Mean   :  38.21
##  3rd Qu.:0.8000        3rd Qu.:24.26   3rd Qu.: 6960   3rd Qu.:  50.00
##  Max.   :0.8000        Max.   :54.56   Max.   :25920   Max.   :3600.00
##
##    idle_time           idle_men       no_of_style_change no_of_workers
##  Min.   :  0.0000   Min.   : 0.0000   0:1050             Min.   : 2.00
##  1st Qu.:  0.0000   1st Qu.: 0.0000   1: 114             1st Qu.: 9.00
##  Median :  0.0000   Median : 0.0000   2:  33             Median :34.00
##  Mean   :  0.7302   Mean   : 0.3693                      Mean   :34.61
##  3rd Qu.:  0.0000   3rd Qu.: 0.0000                      3rd Qu.:57.00
##  Max.   :300.0000   Max.   :45.0000                      Max.   :89.00
##
##  actual_productivity
##  Min.   :0.2337
##  1st Qu.:0.6503
##  Median :0.7733
##  Mean   :0.7351
##  3rd Qu.:0.8503
##  Max.   :1.1204
##
```

Note:
over_time,incentive, idle_time and idle_men has a very high maximum value, the mean and median are also very different - check the variables for skewness.
actual_productivity data ranges from 0.2 to 1.1

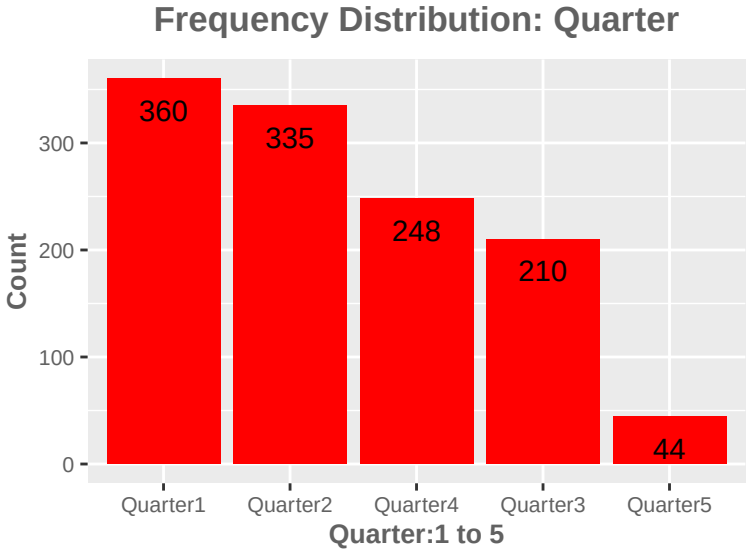
II. Explore Categorical Variables

```
##
##  SCATTER PLOT MATRIX OF RESPONSE AND THE FIVE CATEGORICAL VARIABLES
```



Based on the observations from the Scatter Plot Matrix, it is evident that for most variables, there is no noticeable trend between those variables and the response variable, except for the variable "no_of_style_change." This particular variable exhibits a negative correlation with the response variable.

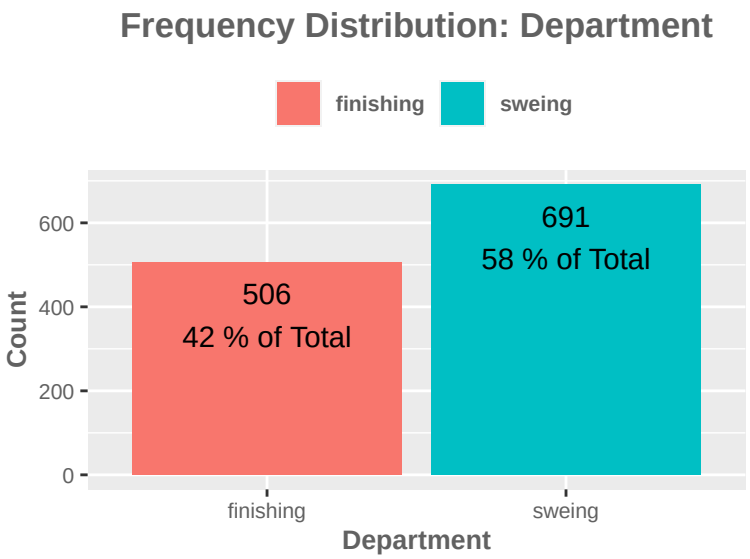
1. Quarter



A general trend observed in the above plot is that there are fewer observations recorded towards the end of the month. Additionally, there seems to be a notable difference between the number of data points in Quarter 4 of a month compared

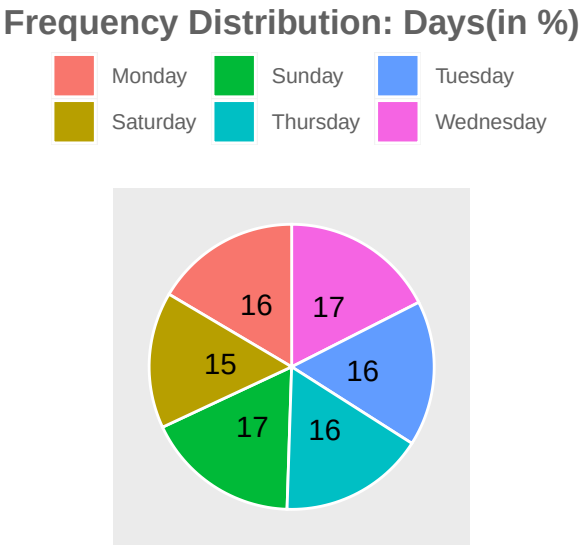
to Quarter 3. Quarter 4 appears to have more data points than Quarter 3. This observation could be significant and may require further investigation to understand the reasons behind this pattern.

2. Department



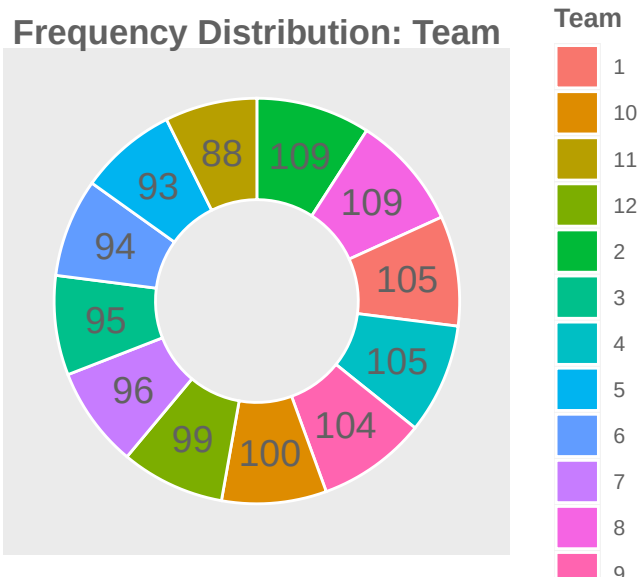
It is noted that the dataset shows an almost equal number of observations within each department.

3. Day



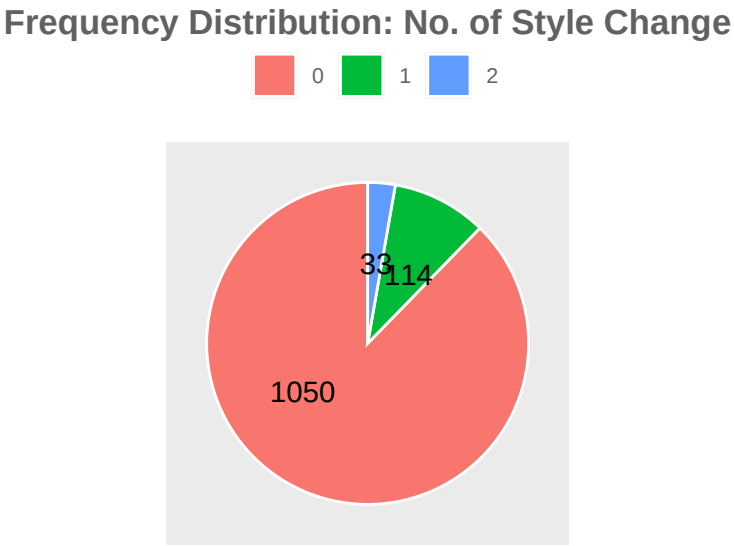
It has been observed that the dataset contains a very similar percentage of data from each day, except for Saturdays. This could be due to multiple reasons.

4. Team



The dataset reveals that the number of data points collected from each team is different and appears to be randomly distributed. Team 8 and team 2 have the highest number of data points, on the other hand, team 11 and team 5 have the least data points. This discrepancy in the amount of data from each team might be a crucial aspect to consider while analyzing and interpreting the results, as it could potentially influence the outcomes of our analysis or models.

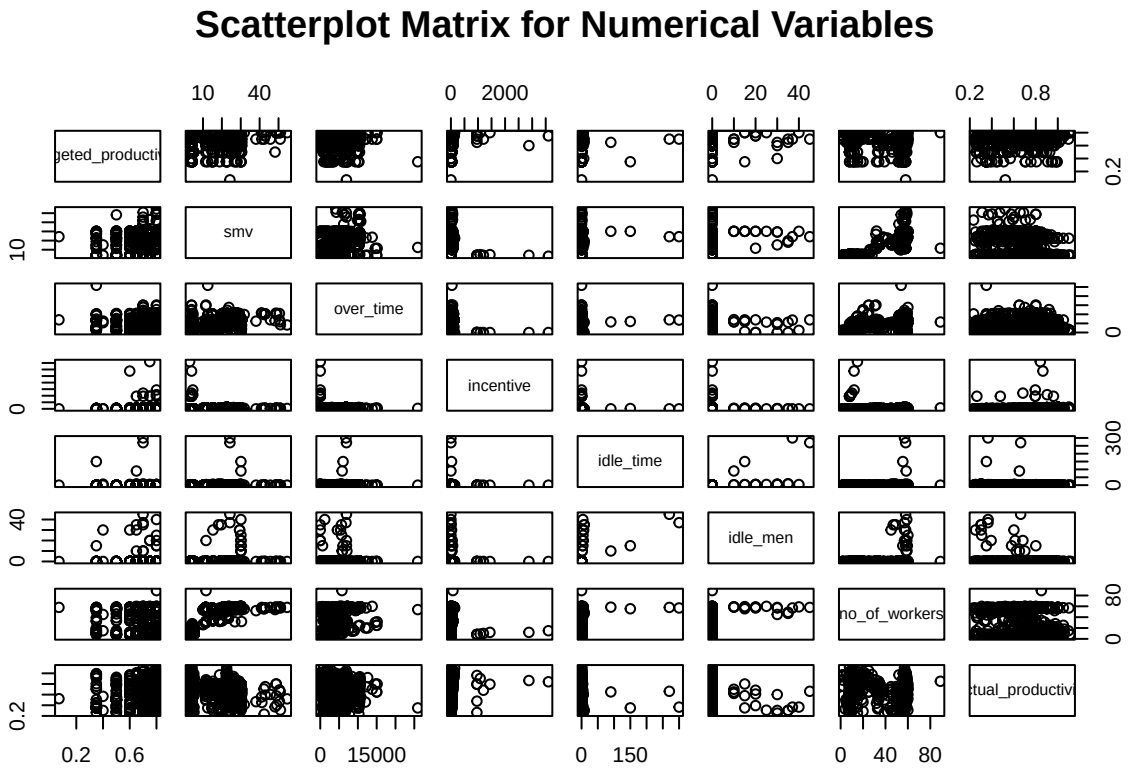
5. Number of Style Change



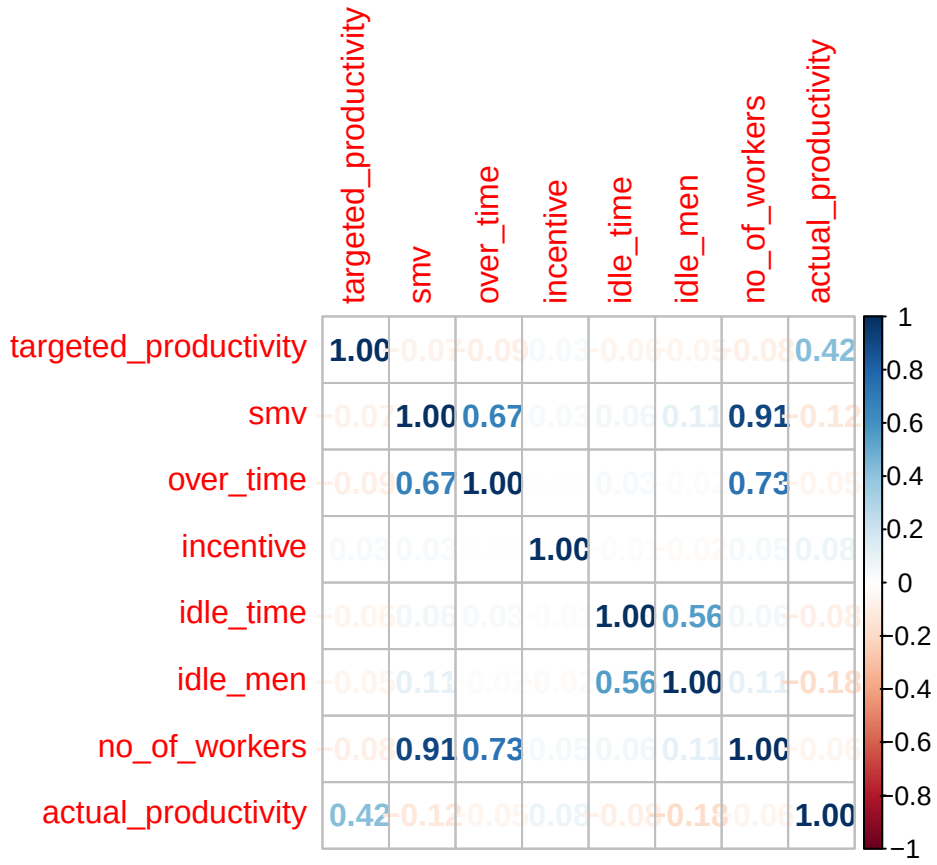
It is evident that there is a substantial difference in the data distribution among the various levels in the variable. The presence of a significant imbalance in the distribution of data across different levels can have implications on the analysis and modeling process.

III. Explore Numerical Variables

```
##
##  SCATTER PLOT MATRIX OF YOUR RESPONSE AND THE NUMERICAL VARIABLES
```




```
##
##  CORRELATION PLOT MATRIX OF YOUR RESPONSE AND THE NUMERICAL VARIABLES
```



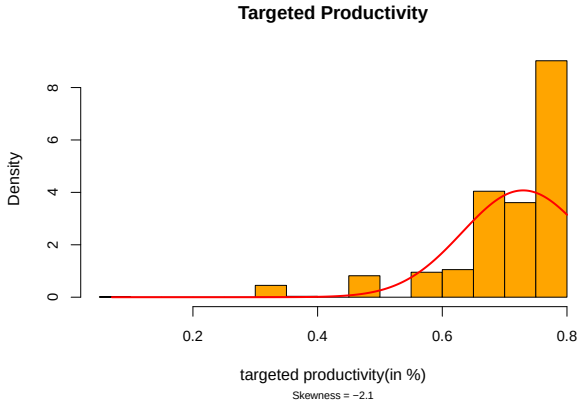
From the scatter plot, there doesn't appear to be any evident correlation among the variables or with the predictor variable. However, the correlation plot reveals certain correlations between specific variables. Notably, there seems to be a correlation between "over_time," "smv," and "no_of_workers." Additionally, there appears to be a correlation between "idle_men" and "idle_time".

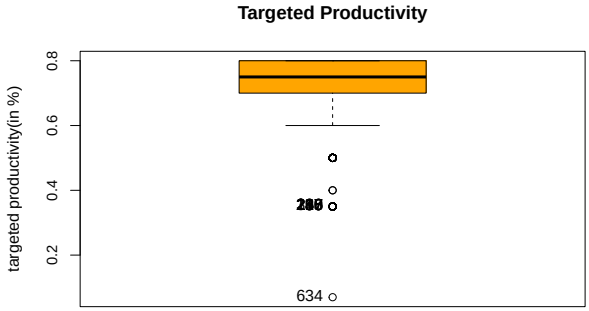
```
##
##  SKEWNESS VALUES FOR NUMERICAL VARIABLES

## targeted_productivity                  smv            over_time
##            -2.1414622            0.4054279            0.6724433
##             incentive            idle_time             idle_men
##            15.7709512           20.5196701            9.8427251
##         no_of_workers   actual_productivity
##            -0.1115997           -0.8064795
```

The above skewness values for the data set variables indicate that we have a lot of skewed variables in our dataset.

1. Targeted Productivity

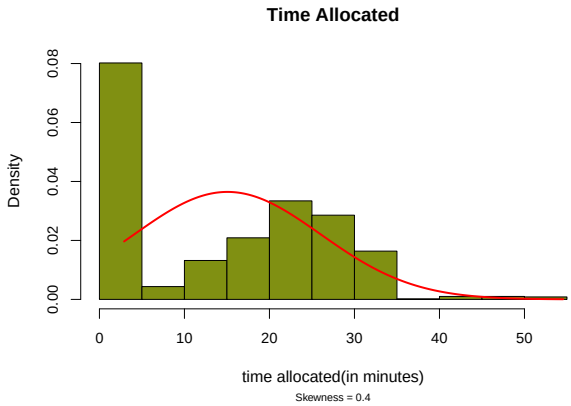
```
##  [1] 634 147 195 215 217 238 240 256 260 337
```

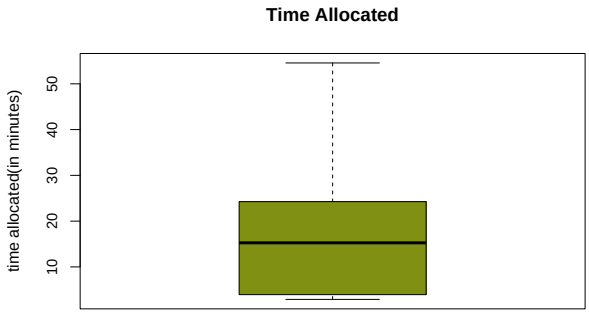




Skewness: -2.1
The skewness value indicates highly negative skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:
634, 147, 195, 215, 217, 238, 240, 256, 260, 337

2. SMV

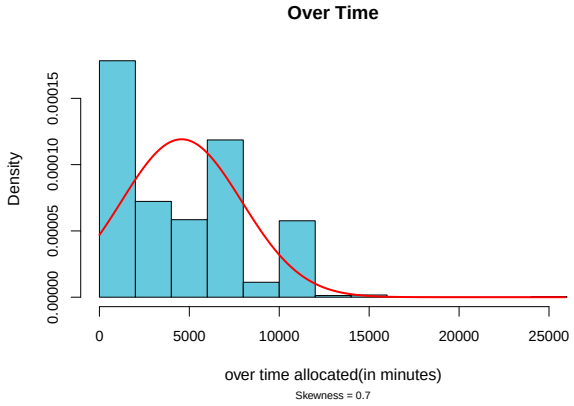


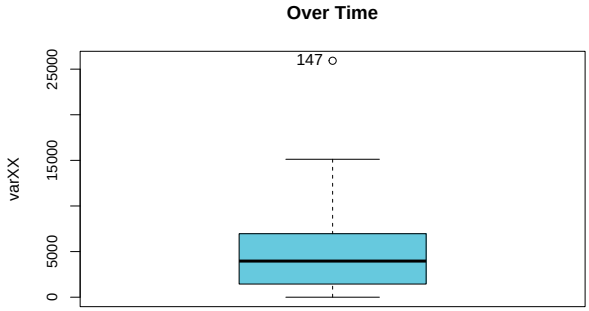


Skewness: 0.4
The skewness value indicates slightly positive skewed variable confirmed by the histogram.

3. Over Time

```
## [1] 147
```

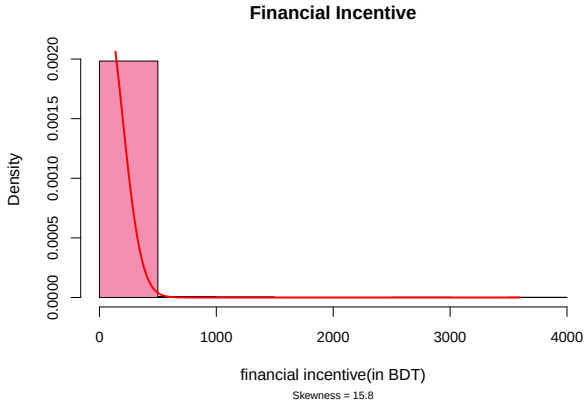


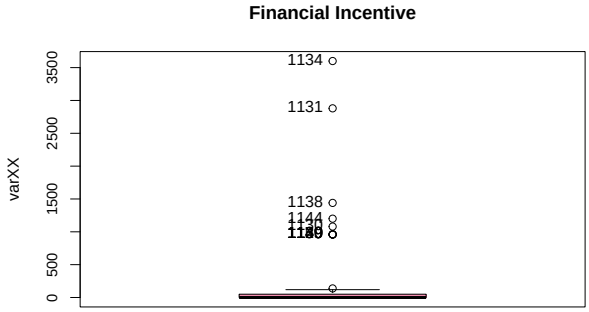


Skewness: 0.7
The skewness value indicates slightly positive skewed variable confirmed by the histogram.
We do have one outlier in the variable, mentioned below:
147

4. Incentive

```
##  [1] 1134 1131 1138 1144 1130 1129 1139 1140 1149 1150
```

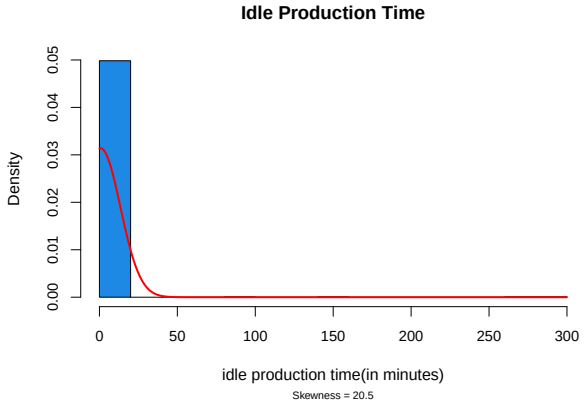


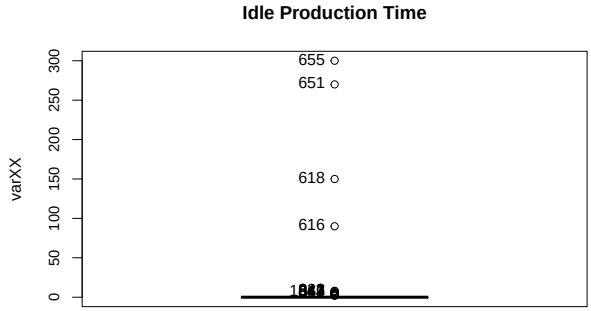


Skewness: 15.8
The skewness value indicates highly positive skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:
1134, 1131, 1138, 1144, 1130, 1129, 1139, 1140, 1149, 1150

5. Idle Time

```
##  [1]  655  651  618  616  823  842 1047  819  861  844
```

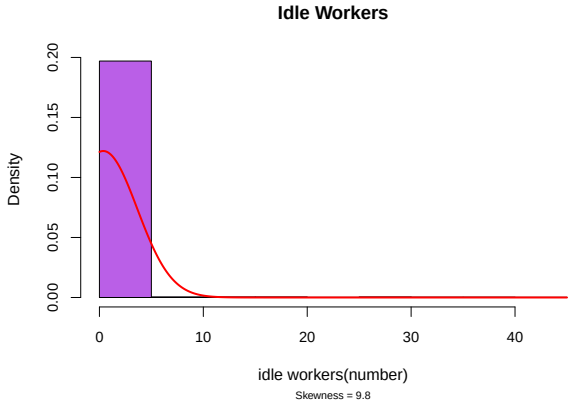


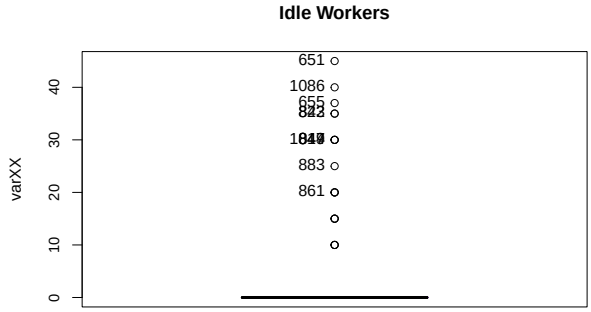


Skewness: 20.5
The skewness value indicates highly positive skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:
655, 651, 618, 616, 823, 842, 1047, 819, 861, 844

6. Idle Men

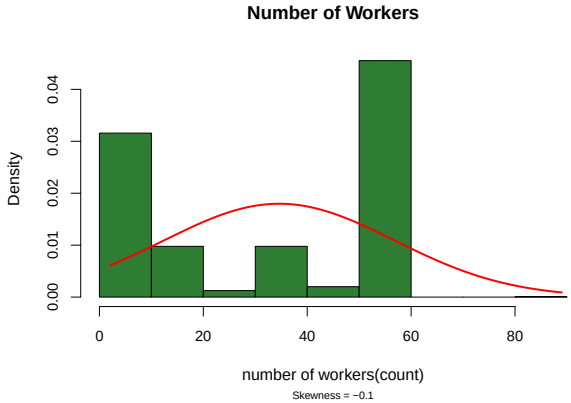
```
##  [1]  651 1086  655  823  842  819  844 1047  883  861
```

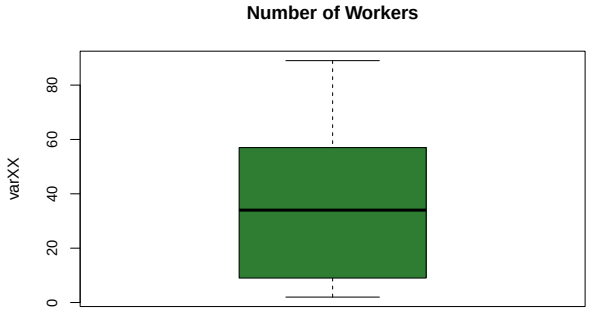




Skewness: 9.8
The skewness value indicates highly positive skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:
651, 1086, 655, 823, 842, 819, 844, 1047, 883, 861

7. Number of Workers

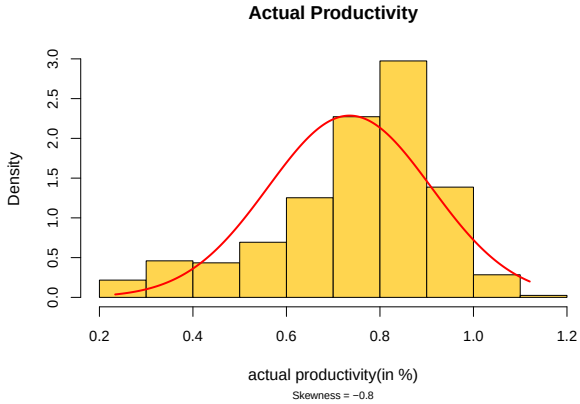


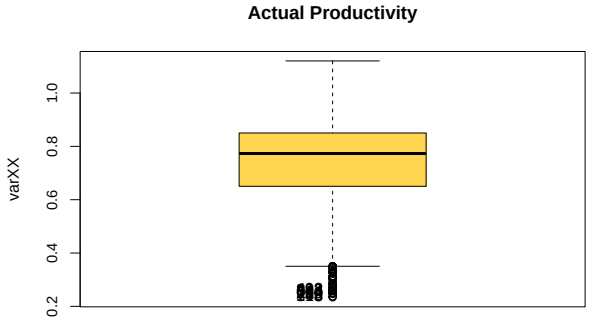


Skewness: -0.1
The skewness value indicates slightly negative skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:

8. Actual Productivity

```
##  [1] 148 318 766 216 259 864 844 843 803 498
```





Skewness: -0.8
The skewness value indicates slightly negative skewed variable confirmed by the histogram.
We do have few outliers in the variable, list mentioned below:
148, 318, 766, 216, 259, 864, 844, 843, 803, 498

IV. Develop Models

Apart from the theoretical understanding, we will be using the adjusted R-squared (R^2adj), p-value as the criterion for model selection, to assess the goodness of fit of different models. The adjusted R-squared accounts for the number of predictors in the model, preventing overfitting by penalizing complex models with additional variables.

Additionally, we will set a p-value threshold of 0.05 for the hypothesis tests. If the p-value obtained from a test is less than 0.05, we reject the null hypothesis in favor of the alternative hypothesis as this suggests that the observed effect or relationship is statistically significant.

Model 1: Numerical Variables

First Numerical Model:

```
##
## Call:
## lm(formula = Y ~ smv + over_time + incentive + idle_time + idle_men +
##     no_of_workers, data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55099 -0.08602  0.04076  0.11681  0.37924
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.462e-01  9.237e-03  80.786  < 2e-16 ***
## smv           -6.320e-03  1.089e-03  -5.801 8.42e-09 ***
## over_time     -2.864e-06  2.180e-06  -1.314    0.189
## incentive      7.380e-05  3.060e-05   2.412    0.016 *
## idle_time      4.540e-04  4.646e-04   0.977    0.329
## idle_men      -1.048e-02  1.834e-03  -5.714 1.40e-08 ***
## no_of_workers  2.827e-03  5.869e-04   4.817 1.64e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1688 on 1190 degrees of freedom
## Multiple R-squared:  0.0687, Adjusted R-squared:  0.064
## F-statistic: 14.63 on 6 and 1190 DF,  p-value: 3.576e-16
```

Performing an ANOVA (Analysis of Variance) test is a valuable way to determine whether removing variables from a model would lead to a significant reduction in the unexplained variability in the response variable.

ANOVA for Over Time:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_time + idle_men + no_of_workers
## Model 2: Y ~ smv + over_time + incentive + idle_time + idle_men + no_of_workers
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1   1191 33.961
## 2   1190 33.912  1  0.049167 1.7253 0.1893
```

p- value: 0.1892647

ANOVA for Idle Time:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers
## Model 2: Y ~ smv + incentive + idle_time + idle_men + no_of_workers
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1   1192 33.983
## 2   1191 33.961  1  0.022054 0.7734 0.3793
```

p- value: 0.3793418

Since the p-values from both ANOVA tests resulted in p-value greater than 0.05, hence we updated the model by removing the two variables.

Updated Numerical Model:

```
##
## Call:
## lm(formula = Y ~ smv + incentive + idle_men + no_of_workers,
##     data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.54873 -0.08950  0.04126  0.11828  0.38123
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.441e-01  9.095e-03  81.813  < 2e-16 ***
## smv           -6.351e-03  1.089e-03  -5.830 7.13e-09 ***
## incentive      7.639e-05  3.054e-05   2.501   0.0125 *
## idle_men      -9.198e-03  1.503e-03  -6.120 1.27e-09 ***
## no_of_workers  2.518e-03  5.375e-04   4.684 3.13e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1688 on 1192 degrees of freedom
## Multiple R-squared:  0.06674,    Adjusted R-squared:  0.06361
## F-statistic: 21.31 on 4 and 1192 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = Y ~ smv + incentive + idle_men + no_of_workers,
##     data = garwork)
##
## Coefficients:
##   (Intercept)            smv      incentive       idle_men  no_of_workers
##     7.441e-01     -6.351e-03      7.639e-05     -9.198e-03      2.518e-03
```

R^2adj = 0.0636
R^2adj degrades by a tiny fraction with the optimized numerical model.

Model 2: Add Categorical Variables

We could use ANOVA tests in a similar way to add categorical variables into the model.

ANOVA for Quarter:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers
## Model 2: Y ~ smv + incentive + idle_men + no_of_workers + quarter
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1192 33.983
## 2   1188 33.211  4   0.77256 6.9089 1.695e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p- value: 1.6950357×10^{-5}
Hence include the variable in the model.

ANOVA for Department:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter
## Model 2: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1   1188 33.211
## 2   1187 32.964  1   0.24705 8.8961 0.002916 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p- value: 0.0029161
Hence include the variable in the model.

ANOVA for Day:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department
## Model 2: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     day
##   Res.Df    RSS Df Sum of Sq     F Pr(>F)
## 1   1187 32.964
## 2   1182 32.849  5   0.11492 0.827 0.5304
```

p- value: 0.5304025
Hence **don't** include the variable in the model.

ANOVA for Team:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department
## Model 2: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     team
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1187 32.964
## 2   1176 30.123 11    2.8411 10.084 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p- value: $9.5826884 \times 10^{-18}$
Hence include the variable in the model.

ANOVA for Number of Style Change:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     team
## Model 2: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1176 30.122
## 2   1174 29.168  2   0.95421 19.203 6.219e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p- value: 6.2193907×10^{-9}
Hence include the variable in the model.

Base Numerical + Categorical Model:

```
##
## Call:
## lm(formula = Y ~ smv + incentive + idle_men + no_of_workers +
##     quarter + department + team + no_of_style_change, data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53297 -0.06754  0.02632  0.10862  0.35445
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          8.010e-01  2.034e-02  39.386  < 2e-16 ***
## smv                 -6.122e-03  1.112e-03  -5.506 4.51e-08 ***
## incentive            5.257e-05  2.884e-05   1.823 0.068553 .
## idle_men            -7.165e-03  1.431e-03  -5.005 6.43e-07 ***
## no_of_workers        4.704e-03  9.048e-04   5.199 2.37e-07 ***
## quarterQuarter2      6.441e-04  1.223e-02   0.053 0.958011
```

```
## quarterQuarter3     -2.481e-02  1.391e-02  -1.783 0.074764 .
## quarterQuarter4     -2.943e-02  1.353e-02  -2.175 0.029858 *
## quarterQuarter5      7.097e-02  2.522e-02   2.815 0.004964 **
## departmentsweing    -8.336e-02  3.643e-02  -2.288 0.022310 *
## team2               -3.601e-02  2.171e-02  -1.659 0.097428 .
## team3               -3.338e-03  2.246e-02  -0.149 0.881911
## team4               -3.646e-02  2.194e-02  -1.662 0.096819 .
## team5               -1.052e-01  2.255e-02  -4.664 3.46e-06 ***
## team6               -9.393e-02  2.517e-02  -3.732 0.000199 ***
## team7               -1.139e-01  2.255e-02  -5.051 5.10e-07 ***
## team8               -1.156e-01  2.178e-02  -5.310 1.31e-07 ***
## team9               -7.494e-02  2.188e-02  -3.424 0.000637 ***
## team10              -8.816e-02  2.215e-02  -3.980 7.32e-05 ***
## team11              -1.590e-01  2.391e-02  -6.651 4.45e-11 ***
## team12              -1.401e-02  2.488e-02  -0.563 0.573316
## no_of_style_change1 -1.056e-01  1.720e-02  -6.139 1.13e-09 ***
## no_of_style_change2 -4.803e-02  3.020e-02  -1.590 0.112016
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1576 on 1174 degrees of freedom
## Multiple R-squared:  0.199,  Adjusted R-squared:  0.184
## F-statistic: 13.26 on 22 and 1174 DF,  p-value: < 2.2e-16


##
##  TOP 10 POSITIVE COEFFICIENTS IN THE MODEL:

##     (Intercept) quarterQuarter5   no_of_workers quarterQuarter2       incentive
##    8.009793e-01    7.097245e-02    4.703913e-03    6.441162e-04    5.256976e-05
##           team3             smv        idle_men          team12 quarterQuarter3
##   -3.337659e-03   -6.121694e-03   -7.164945e-03   -1.401331e-02   -2.480619e-02
```

R^2adj = 0.184

R^2adj has improved in comparison to our optimized Numerical Model.

Model 3: Add Interaction Terms

Evaluate whether the 2-way interaction terms are necessary in our model using ANOVA test:

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change
## Model 2: Y ~ (smv + incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change)^2
##   Res.Df    RSS  Df Sum of Sq      F    Pr(>F)
## 1   1174 29.168
## 2   1024 17.205 150    11.963 4.7469 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p- value: $1.9570171 \times 10^{-51}$
The p-value from the ANOVA test is way less than 0.05, we can confirm that we need the 2 way interactive terms to explain our model.

Complete 2-way Model:

```
##
##  TOP 10 POSITIVE COEFFICIENTS IN THE MODEL

##           departmentsweing                (Intercept)
##                  1.9229154                  0.8798695
## team11:no_of_style_change1  team9:no_of_style_change1
##                  0.5685957                  0.5381277
##  team3:no_of_style_change1  team6:no_of_style_change1
##                  0.4817562                  0.4793441
##  team2:no_of_style_change1  team4:no_of_style_change1
##                  0.4765656                  0.4725530
##  team7:no_of_style_change1  team8:no_of_style_change1
##                  0.4461004                  0.4242623
```

Repeating the ANOVA test as shown below, we can remove non-significant interaction terms one by one and build our final model.

```
## Analysis of Variance Table
##
## Model 1: Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change + smv:incentive + smv:idle_men +
##     smv:no_of_workers + smv:quarter + smv:team + smv:no_of_style_change +
##     incentive:idle_men + incentive:no_of_workers + incentive:quarter +
##     incentive:department + incentive:team + incentive:no_of_style_change +
##     idle_men:no_of_workers + idle_men:quarter + idle_men:department +
##     idle_men:team + idle_men:no_of_style_change + no_of_workers:quarter +
##     no_of_workers:department + no_of_workers:team + no_of_workers:no_of_style_change +
##     quarter:department + quarter:team + quarter:no_of_style_change +
```

```
##     department:team + department:no_of_style_change + team:no_of_style_change
## Model 2: Y ~ (smv + incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change)^2
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1   1025 17.320
## 2   1024 17.205  1   0.11483 6.8343 0.009074 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction Term Model:

```
##
## Call:
## lm(formula = Y ~ smv + incentive + idle_men + no_of_workers +
##     quarter + department + team + no_of_style_change + smv:no_of_workers +
##     smv:department + incentive:quarter + incentive:department +
##     incentive:team + idle_men:department + no_of_workers:department +
##     no_of_workers:team + quarter:department + quarter:team +
##     quarter:no_of_style_change + department:team + department:no_of_style_change +
##     team:no_of_style_change, data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.56147 -0.03855  0.00724  0.06753  0.37958
##
## Coefficients: (12 not defined because of singularities)
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          8.496e-01  7.783e-02  10.917  < 2e-16 ***
## smv                                  8.901e-03  1.349e-02   0.660 0.509655
## incentive                           -8.804e-04  5.761e-04  -1.528 0.126810
## idle_men                            -3.696e-03  1.323e-03  -2.793 0.005315 **
## no_of_workers                       -8.409e-03  5.277e-03  -1.594 0.111316
## quarterQuarter2                      3.926e-02  3.449e-02   1.138 0.255358
## quarterQuarter3                     -1.389e-02  4.622e-02  -0.301 0.763746
## quarterQuarter4                     -2.650e-02  4.179e-02  -0.634 0.526101
## quarterQuarter5                      1.397e-01  7.270e-02   1.921 0.055024 .
## departmentsweing                     1.484e+00  3.560e-01   4.169 3.31e-05 ***
## team2                               -2.377e-01  7.197e-02  -3.303 0.000987 ***
## team3                               -1.974e-03  8.963e-02  -0.022 0.982429
## team4                               -2.598e-01  7.935e-02  -3.274 0.001094 **
## team5                               -3.581e-01  7.439e-02  -4.814 1.70e-06 ***
## team6                               -4.596e-01  7.473e-02  -6.150 1.09e-09 ***
## team7                               -3.394e-01  8.578e-02  -3.956 8.12e-05 ***
## team8                               -3.346e-01  8.239e-02  -4.061 5.25e-05 ***
## team9                               -2.469e-01  7.965e-02  -3.099 0.001993 **
## team10                              -3.713e-01  7.575e-02  -4.902 1.09e-06 ***
## team11                              -3.568e-01  7.570e-02  -4.713 2.76e-06 ***
## team12                              -2.798e-01  8.324e-02  -3.362 0.000802 ***
## no_of_style_change1                 -6.775e-01  1.593e-01  -4.253 2.29e-05 ***
## no_of_style_change2                 -2.103e-01  1.028e-01  -2.045 0.041110 *
## smv:no_of_workers                    3.534e-04  1.706e-04   2.071 0.038566 *
## smv:departmentsweing                -3.136e-02  1.566e-02  -2.002 0.045508 *
## incentive:quarterQuarter2            8.012e-04  5.640e-04   1.421 0.155701
## incentive:quarterQuarter3            5.158e-04  6.906e-04   0.747 0.455280
```

```
## incentive:quarterQuarter4             2.504e-04  6.900e-04   0.363 0.716730
## incentive:quarterQuarter5             3.318e-03  1.161e-03   2.859 0.004332 **
## incentive:departmentsweing            4.581e-03  4.394e-04  10.427  < 2e-16 ***
## incentive:team2                       1.135e-05  1.746e-04   0.065 0.948150
## incentive:team3                       1.806e-05  1.642e-04   0.110 0.912404
## incentive:team4                       6.467e-05  1.933e-04   0.334 0.738077
## incentive:team5                       8.232e-05  1.441e-04   0.571 0.567820
## incentive:team6                      -2.065e-03  1.095e-03  -1.887 0.059450 .
## incentive:team7                       1.355e-03  9.008e-04   1.504 0.132908
## incentive:team8                      -3.050e-04  1.930e-04  -1.581 0.114248
## incentive:team9                       8.630e-05  1.418e-04   0.609 0.542938
## incentive:team10                     -8.477e-05  1.927e-04  -0.440 0.660159
## incentive:team11                      3.693e-04  1.947e-04   1.897 0.058095 .
## incentive:team12                      1.570e-04  1.839e-04   0.854 0.393562
## idle_men:departmentsweing                    NA         NA      NA       NA
## no_of_workers:departmentsweing       -2.127e-02  4.092e-03  -5.198 2.41e-07 ***
## no_of_workers:team2                   2.069e-02  6.105e-03   3.389 0.000726 ***
## no_of_workers:team3                   5.468e-03  7.164e-03   0.763 0.445481
## no_of_workers:team4                   1.933e-02  6.406e-03   3.018 0.002602 **
## no_of_workers:team5                   2.639e-02  6.470e-03   4.079 4.86e-05 ***
## no_of_workers:team6                   2.612e-02  6.423e-03   4.066 5.14e-05 ***
## no_of_workers:team7                   2.249e-02  7.472e-03   3.010 0.002677 **
## no_of_workers:team8                   2.523e-02  7.342e-03   3.436 0.000612 ***
## no_of_workers:team9                   2.187e-02  7.961e-03   2.747 0.006109 **
## no_of_workers:team10                  2.930e-02  6.771e-03   4.327 1.65e-05 ***
## no_of_workers:team11                  2.609e-02  6.479e-03   4.026 6.08e-05 ***
## no_of_workers:team12                  2.618e-02  7.667e-03   3.414 0.000664 ***
## quarterQuarter2:departmentsweing     -7.377e-02  3.604e-02  -2.047 0.040938 *
## quarterQuarter3:departmentsweing     -4.425e-02  4.055e-02  -1.091 0.275406
## quarterQuarter4:departmentsweing     -4.159e-03  4.352e-02  -0.096 0.923882
## quarterQuarter5:departmentsweing     -4.494e-01  8.449e-02  -5.318 1.28e-07 ***
## quarterQuarter2:team2                -5.473e-02  4.609e-02  -1.188 0.235271
## quarterQuarter3:team2                -1.394e-02  6.043e-02  -0.231 0.817682
## quarterQuarter4:team2                -6.568e-02  5.432e-02  -1.209 0.226885
## quarterQuarter5:team2                -7.387e-02  9.880e-02  -0.748 0.454806
## quarterQuarter2:team3                -3.676e-02  5.097e-02  -0.721 0.471013
## quarterQuarter3:team3                -9.441e-02  6.165e-02  -1.532 0.125934
## quarterQuarter4:team3                 1.242e-03  5.740e-02   0.022 0.982743
## quarterQuarter5:team3                -1.729e-01  9.933e-02  -1.741 0.081948 .
## quarterQuarter2:team4                 3.463e-02  4.739e-02   0.731 0.465063
## quarterQuarter3:team4                -1.223e-02  5.811e-02  -0.210 0.833397
## quarterQuarter4:team4                 2.897e-02  5.758e-02   0.503 0.614983
## quarterQuarter5:team4                 6.244e-02  9.738e-02   0.641 0.521551
## quarterQuarter2:team5                 4.364e-02  5.341e-02   0.817 0.414055
## quarterQuarter3:team5                 1.463e-01  6.267e-02   2.335 0.019738 *
## quarterQuarter4:team5                 4.036e-02  6.161e-02   0.655 0.512533
## quarterQuarter5:team5                 1.195e-01  1.069e-01   1.118 0.263990
## quarterQuarter2:team6                 2.128e-02  5.099e-02   0.417 0.676463
## quarterQuarter3:team6                 9.775e-02  5.967e-02   1.638 0.101679
## quarterQuarter4:team6                 1.102e-02  5.690e-02   0.194 0.846447
## quarterQuarter5:team6                 1.490e-01  1.035e-01   1.440 0.150057
## quarterQuarter2:team7                -2.527e-02  5.163e-02  -0.489 0.624698
## quarterQuarter3:team7                -1.361e-02  6.105e-02  -0.223 0.823676
## quarterQuarter4:team7                -6.038e-02  5.903e-02  -1.023 0.306608
```

```
## quarterQuarter5:team7                 1.696e-01  1.197e-01   1.416 0.156954
## quarterQuarter2:team8                -4.395e-02  4.915e-02  -0.894 0.371410
## quarterQuarter3:team8                 4.325e-03  5.940e-02   0.073 0.941972
## quarterQuarter4:team8                -7.143e-02  5.922e-02  -1.206 0.228025
## quarterQuarter5:team8                 5.997e-02  1.107e-01   0.542 0.588055
## quarterQuarter2:team9                -6.791e-02  4.673e-02  -1.453 0.146434
## quarterQuarter3:team9                -3.564e-02  5.913e-02  -0.603 0.546881
## quarterQuarter4:team9                -1.534e-01  5.853e-02  -2.622 0.008874 **
## quarterQuarter5:team9                 1.037e-01  1.053e-01   0.985 0.325023
## quarterQuarter2:team10               -9.214e-02  4.821e-02  -1.911 0.056232 .
## quarterQuarter3:team10               -2.639e-02  6.014e-02  -0.439 0.660897
## quarterQuarter4:team10                1.902e-02  5.699e-02   0.334 0.738684
## quarterQuarter5:team10                9.735e-03  9.763e-02   0.100 0.920588
## quarterQuarter2:team11               -2.121e-02  5.099e-02  -0.416 0.677513
## quarterQuarter3:team11               -1.143e-01  6.298e-02  -1.816 0.069721 .
## quarterQuarter4:team11               -7.135e-02  6.319e-02  -1.129 0.259069
## quarterQuarter5:team11                9.963e-02  1.003e-01   0.993 0.320938
## quarterQuarter2:team12               -5.566e-03  4.833e-02  -0.115 0.908323
## quarterQuarter3:team12                2.603e-02  5.899e-02   0.441 0.659122
## quarterQuarter4:team12               -2.266e-02  5.451e-02  -0.416 0.677674
## quarterQuarter5:team12                1.951e-02  9.956e-02   0.196 0.844694
## quarterQuarter2:no_of_style_change1   1.265e-01  5.595e-02   2.261 0.023975 *
## quarterQuarter3:no_of_style_change1   1.628e-01  6.253e-02   2.604 0.009351 **
## quarterQuarter4:no_of_style_change1   2.169e-01  6.193e-02   3.503 0.000479 ***
## quarterQuarter5:no_of_style_change1          NA         NA      NA       NA
## quarterQuarter2:no_of_style_change2  -1.971e-01  1.154e-01  -1.707 0.088033 .
## quarterQuarter3:no_of_style_change2  -2.521e-02  7.370e-02  -0.342 0.732416
## quarterQuarter4:no_of_style_change2          NA         NA      NA       NA
## quarterQuarter5:no_of_style_change2          NA         NA      NA       NA
## departmentsweing:team2               -9.194e-01  2.938e-01  -3.130 0.001797 **
## departmentsweing:team3               -2.881e-01  3.361e-01  -0.857 0.391519
## departmentsweing:team4               -8.787e-01  3.037e-01  -2.894 0.003886 **
## departmentsweing:team5               -1.218e+00  3.150e-01  -3.868 0.000116 ***
## departmentsweing:team6               -9.197e-01  3.002e-01  -3.063 0.002244 **
## departmentsweing:team7               -9.687e-01  3.588e-01  -2.700 0.007052 **
## departmentsweing:team8               -1.087e+00  3.522e-01  -3.086 0.002079 **
## departmentsweing:team9               -9.652e-01  3.882e-01  -2.487 0.013053 *
## departmentsweing:team10              -1.270e+00  3.229e-01  -3.933 8.93e-05 ***
## departmentsweing:team11              -1.192e+00  3.112e-01  -3.832 0.000135 ***
## departmentsweing:team12              -1.195e+00  3.173e-01  -3.764 0.000176 ***
## departmentsweing:no_of_style_change1         NA         NA      NA       NA
## departmentsweing:no_of_style_change2         NA         NA      NA       NA
## team2:no_of_style_change1             5.305e-01  1.476e-01   3.594 0.000341 ***
## team3:no_of_style_change1             5.502e-01  1.587e-01   3.467 0.000547 ***
## team4:no_of_style_change1             5.874e-01  1.539e-01   3.816 0.000143 ***
## team5:no_of_style_change1             4.386e-01  1.575e-01   2.784 0.005464 **
## team6:no_of_style_change1             5.091e-01  1.603e-01   3.175 0.001541 **
## team7:no_of_style_change1             5.580e-01  1.561e-01   3.576 0.000365 ***
## team8:no_of_style_change1             5.542e-01  1.609e-01   3.445 0.000594 ***
## team9:no_of_style_change1             5.121e-01  1.558e-01   3.287 0.001047 **
## team10:no_of_style_change1            4.071e-01  1.526e-01   2.669 0.007729 **
## team11:no_of_style_change1            5.633e-01  2.026e-01   2.780 0.005532 **
## team12:no_of_style_change1                   NA         NA      NA       NA
## team2:no_of_style_change2                    NA         NA      NA       NA
```

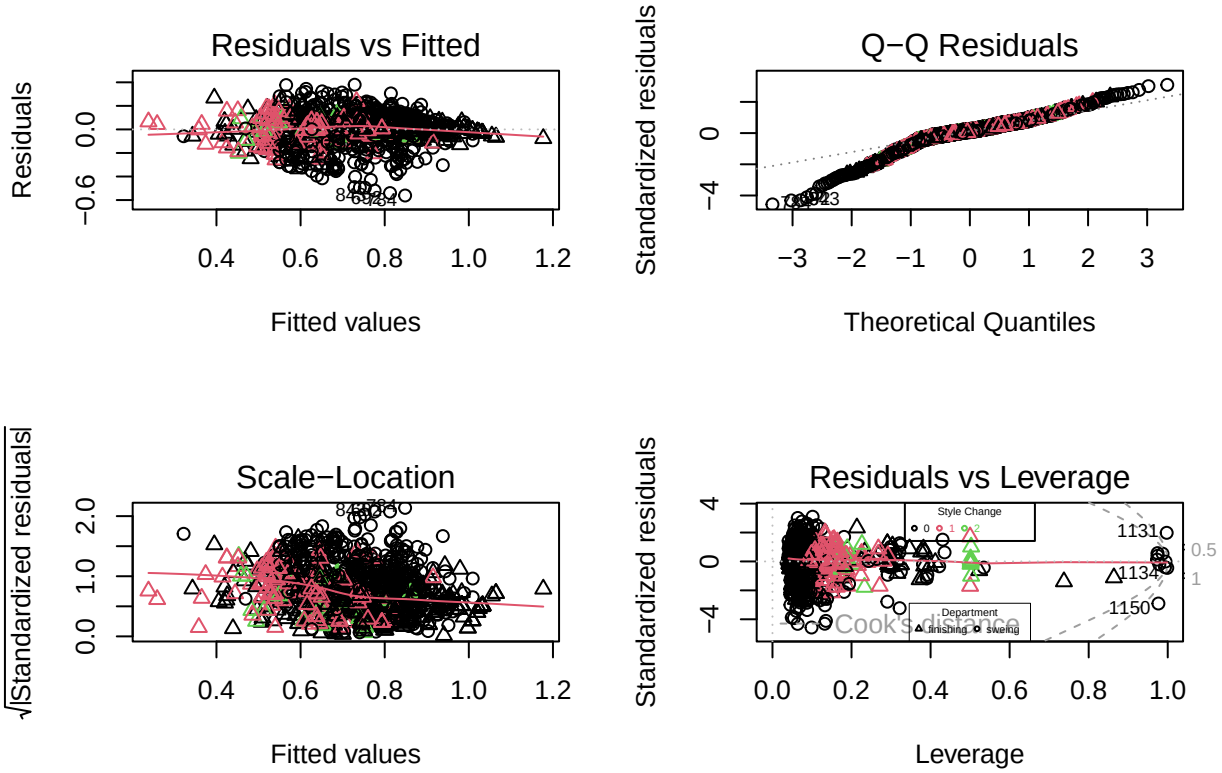
```
## team3:no_of_style_change2             2.186e-01  1.195e-01   1.828 0.067755 .
## team4:no_of_style_change2             5.942e-02  1.412e-01   0.421 0.673873
## team5:no_of_style_change2                    NA         NA      NA       NA
## team6:no_of_style_change2                    NA         NA      NA       NA
## team7:no_of_style_change2             2.830e-01  1.716e-01   1.649 0.099358 .
## team8:no_of_style_change2             2.624e-01  1.211e-01   2.167 0.030456 *
## team9:no_of_style_change2                    NA         NA      NA       NA
## team10:no_of_style_change2            3.380e-01  1.694e-01   1.995 0.046294 *
## team11:no_of_style_change2            4.047e-01  1.199e-01   3.377 0.000759 ***
## team12:no_of_style_change2                   NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1295 on 1064 degrees of freedom
## Multiple R-squared:  0.5103, Adjusted R-squared:  0.4496
## F-statistic:   8.4 on 132 and 1064 DF,  p-value: < 2.2e-16


##
##  TOP 10 POSITIVE COEFFICIENTS IN THE MODEL:

##           departmentsweing                (Intercept)
##                  1.4843349                  0.8496415
##  team4:no_of_style_change1 team11:no_of_style_change1
##                  0.5873998                  0.5633334
##  team7:no_of_style_change1  team8:no_of_style_change1
##                  0.5580347                  0.5541548
##  team3:no_of_style_change1  team2:no_of_style_change1
##                  0.5502489                  0.5304594
##  team9:no_of_style_change1  team6:no_of_style_change1
##                  0.5121158                  0.5091328
```

R^2adj = 0.4496

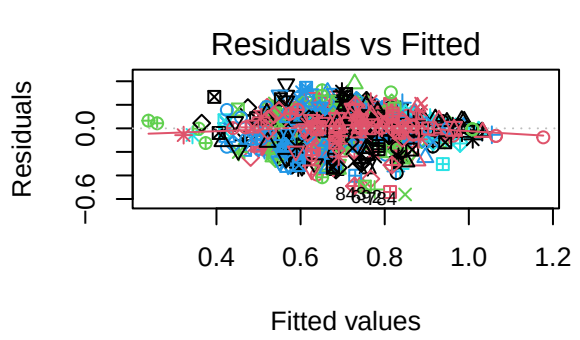
R^2adj has improved in comparison to our optimized Numerical+Categorical model.

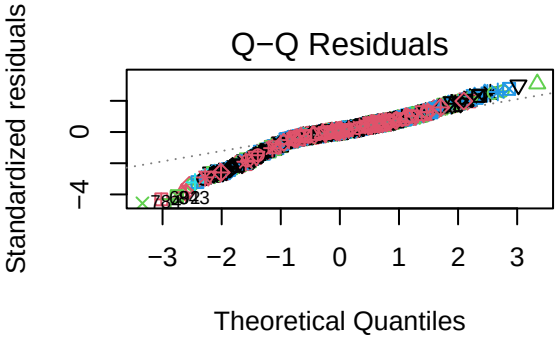
V. Check Assumptions

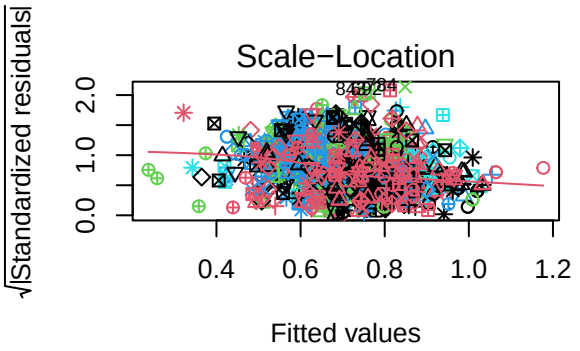
```
## STANDARDIZED RESIDUAL PLOT: STYLE CHANGE(COLOR) AND DEPARTMENT(SHAPE)
```

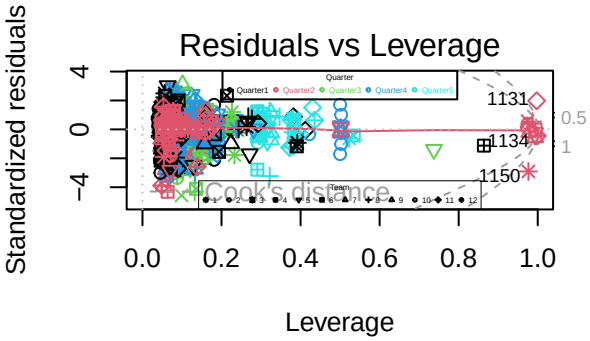



```
## STANDARDIZED RESIDUAL PLOT: QUARTER(COLOR) AND TEAM(SHAPE)
```

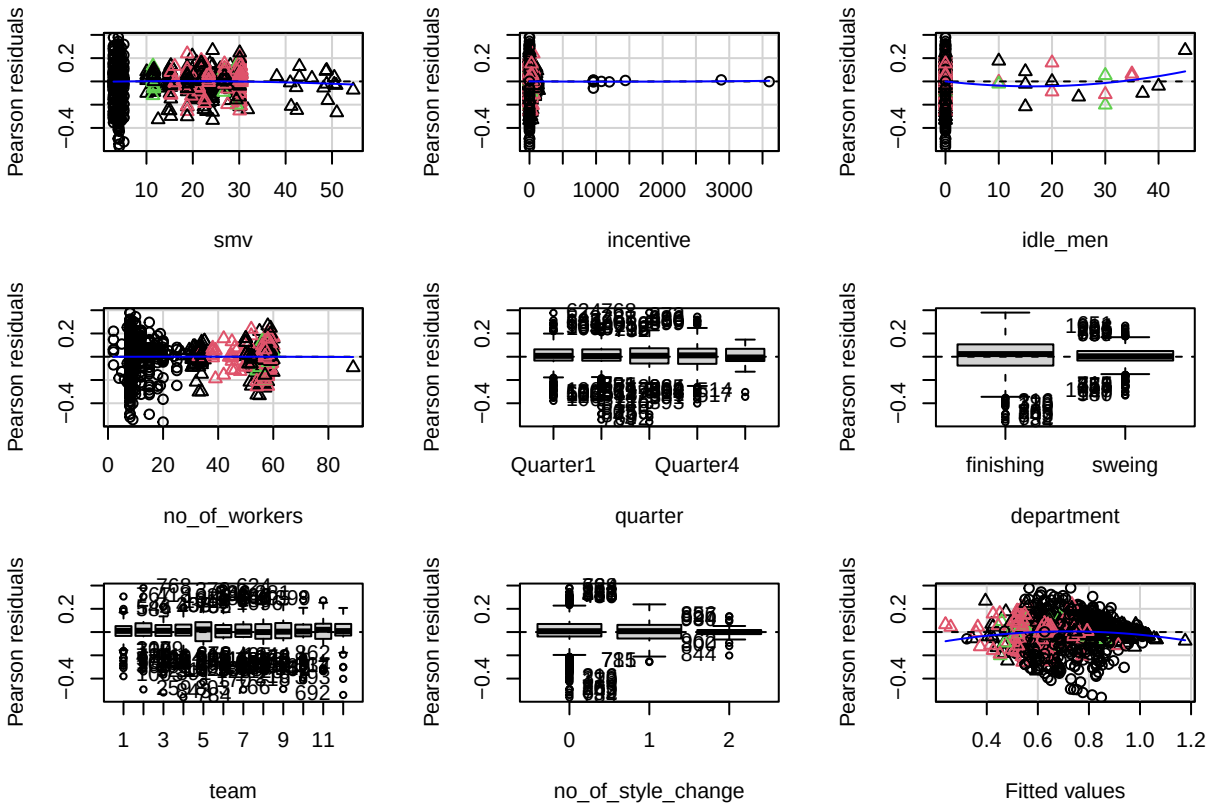




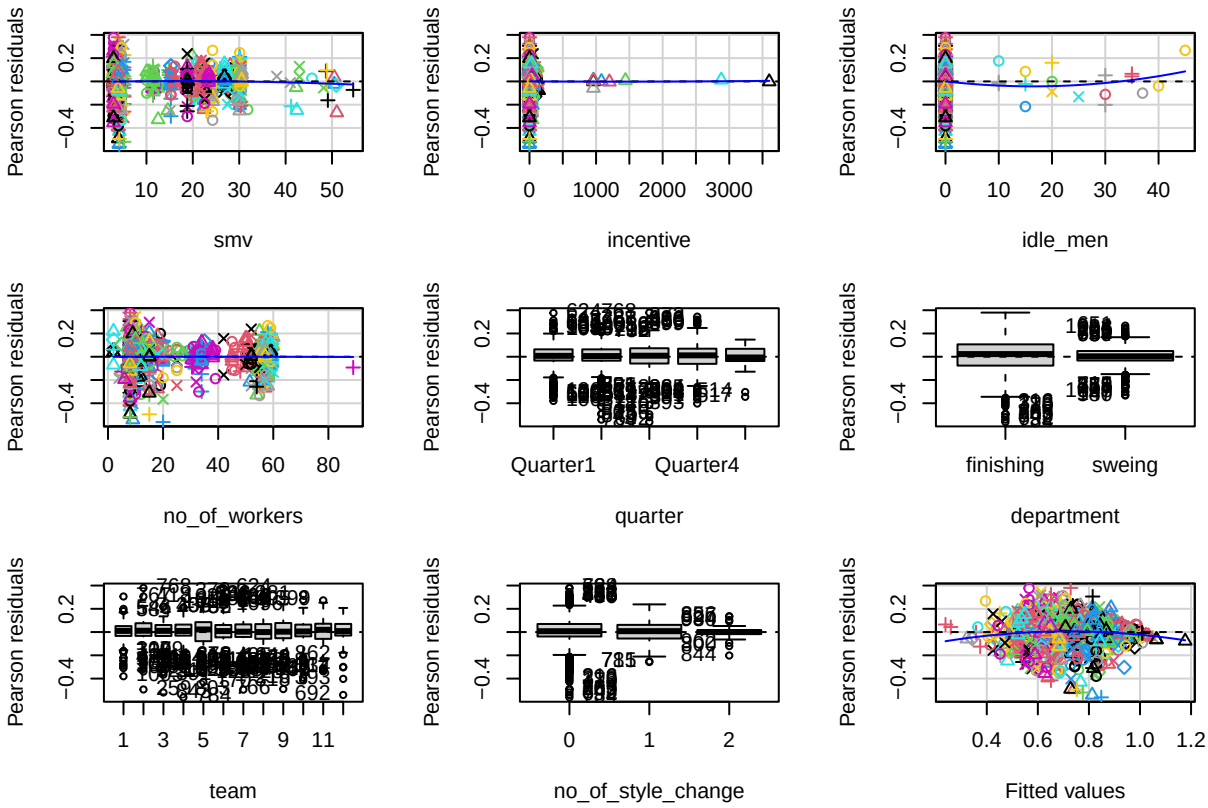





```
## NORMALITY TEST

##
##  Shapiro-Wilk normality test
##
## data:  rstandard(mod)
## W = 0.94563, p-value < 2.2e-16

## FITTED RESIDUAL PLOT: STYLE CHANGE(COLOR) AND DEPARTMENT(SHAPE)
```



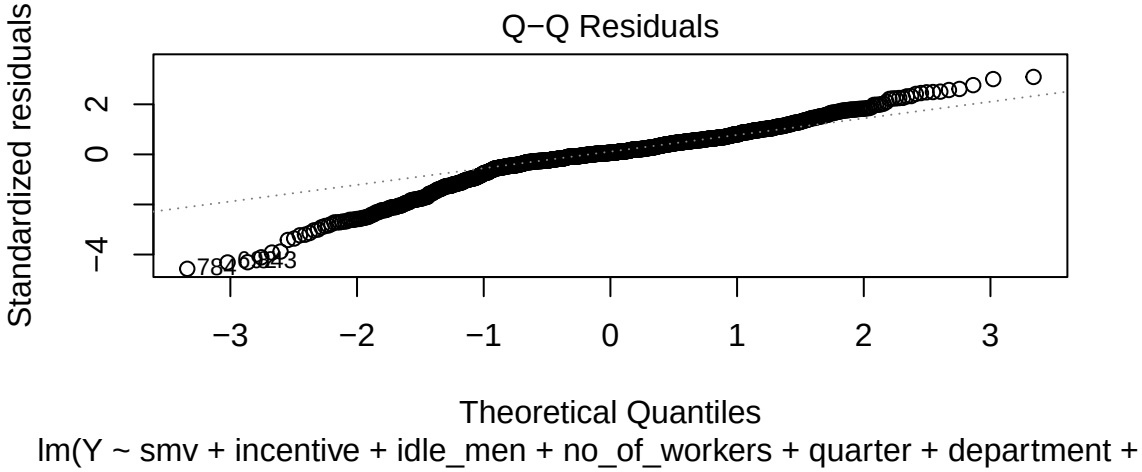

```
##                     Test stat Pr(>|Test stat|)
## smv                   -2.0657        0.0390997 *
## incentive              1.0380        0.2994841
## idle_men               1.3289        0.1841612
## no_of_workers         -0.3469        0.7287195
## quarter
## department
## team
## no_of_style_change
## Tukey test            -3.6064        0.0003105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## FITTED RESIDUAL PLOT: QUARTER(COLOR) AND TEAM(SHAPE)
```




```
##                     Test stat Pr(>|Test stat|)
## smv                   -2.0657        0.0390997 *
## incentive              1.0380        0.2994841
## idle_men               1.3289        0.1841612
## no_of_workers         -0.3469        0.7287195
## quarter
## department
## team
## no_of_style_change
## Tukey test            -3.6064        0.0003105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

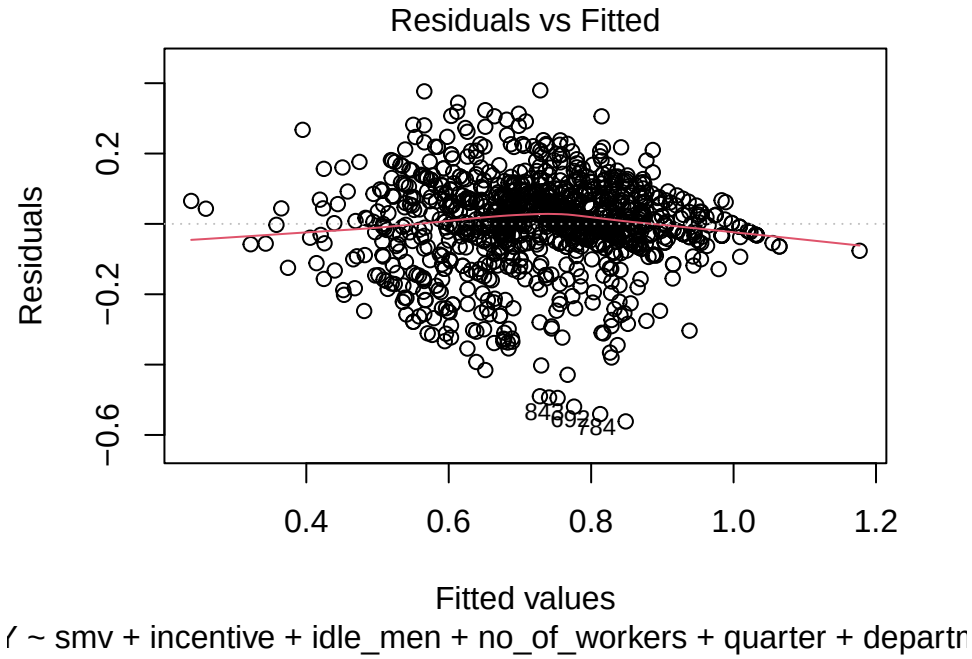
All of the above plots convey the message that our model assumptions are violated with our current model. Lets deep dive into each of them below:

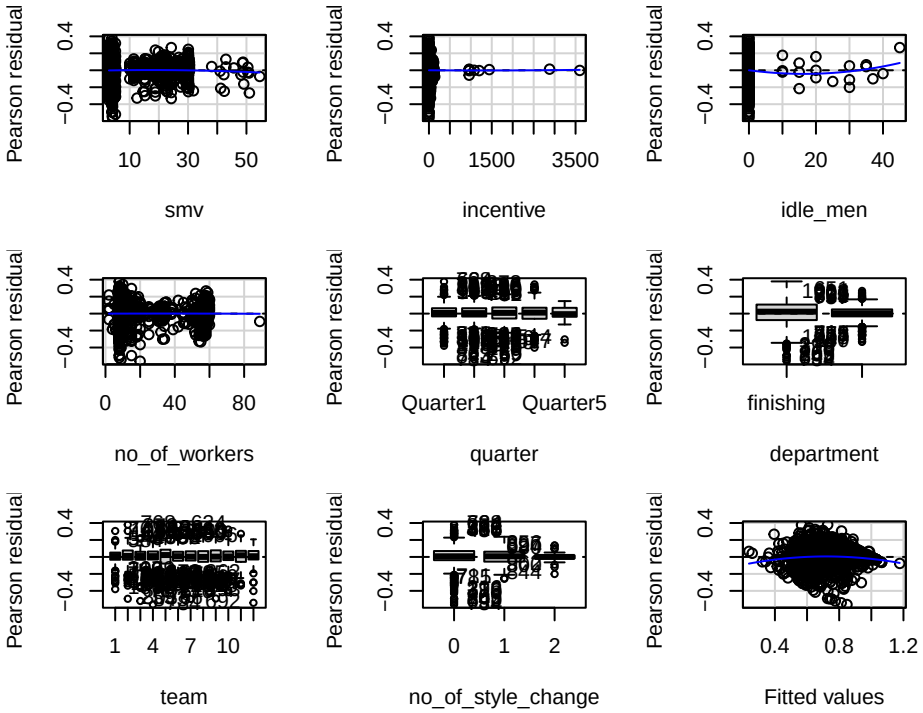
1. Normality:



The Q-Q plot shows that the points are highly biased on the negative axis - left skew. The Shapiro-Wilk test confirms with a very low p-value that there is a risk with the assumption of normality. Hence, the assumption of normal distribution is violated with the model. Therefore we will have to address this with proper transformations.

2. Linearity:






```
##                    Test stat Pr(>|Test stat|)
## smv                  -2.0657        0.0390997 *
## incentive             1.0380        0.2994841
## idle_men              1.3289        0.1841612
## no_of_workers        -0.3469        0.7287195
## quarter
## department
## team
## no_of_style_change
## Tukey test           -3.6064        0.0003105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The residual-fitted plot should show a straight line or equal distribution of points on top and bottom sides of the line to prove linearity.The plot from the given model shows a bias and hence rejects the assumption of linearity.Looking further, we do observe a quadratic trend in the residual plots for idle_men.Hence, we need these to be fixed with necessary transformations.

3.Homoscedasticity:

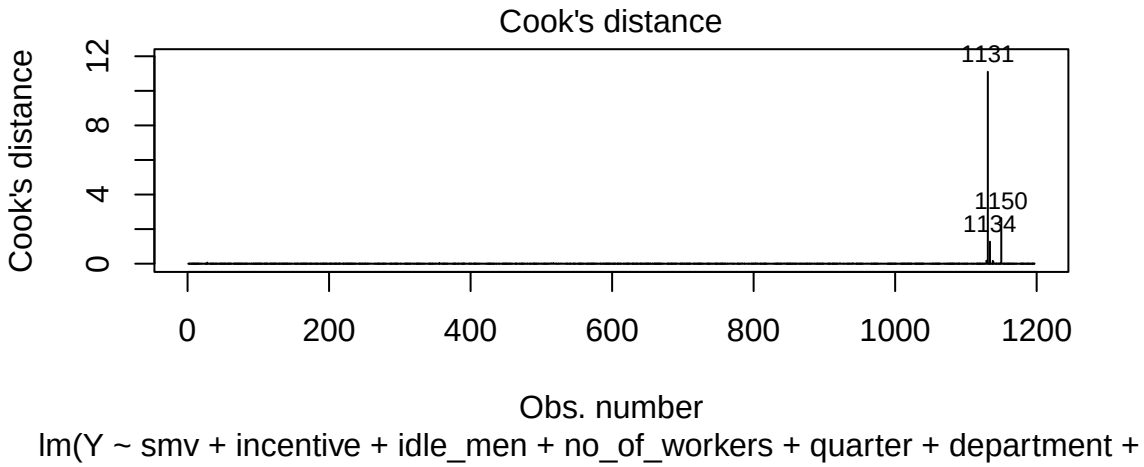
```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 54.74713, Df = 1, p = 1.3708e-13
```

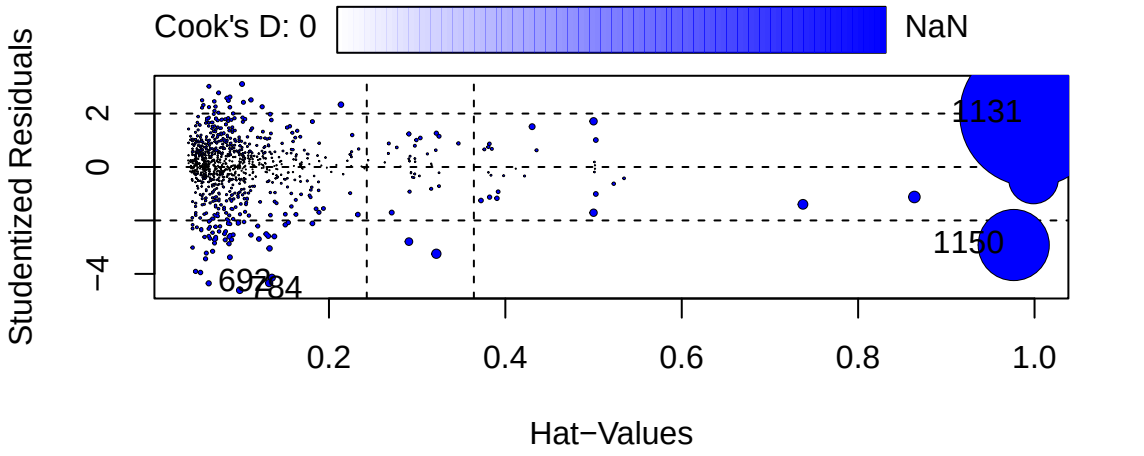
When the data follows homoscedasticity, we observe equal distribution of points in the residual-fitted and residual-predictor plots.We can also use the Breusch-Pagan test, where the p-value > 0.05 gives us enough proof to agree that the data follows the assumption of homoscedasticity.

The residual-fitted values plot show a quadratic trend. The Scale-location plot also can confirm with a strong trend that the assumption of homoscedasticity is invalid. From the above residual-fitted and residual plots for the predictors, we observe an uneven spread of the residuals. Also, the p-value from the Breusch-Pagan test is small thus confirming heteroscedasticity.We should make appropriate transformations to correct them.

Observing the residual vs fitted values curve, we can state that there is no strong linear relationship in our final model. From the individual residual plots, we come to the same conclusion that there is no variable with strong positive or negative linear relationship. We do have few outliers and hence check their leverage in the new model.

4.Influential Points:





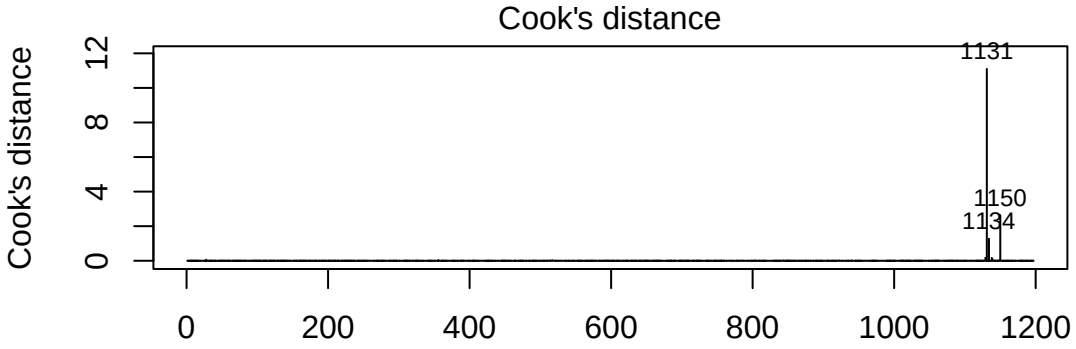

```
##         StudRes         Hat        CookD
## 692  -4.352590 0.06367490  0.009526248
```

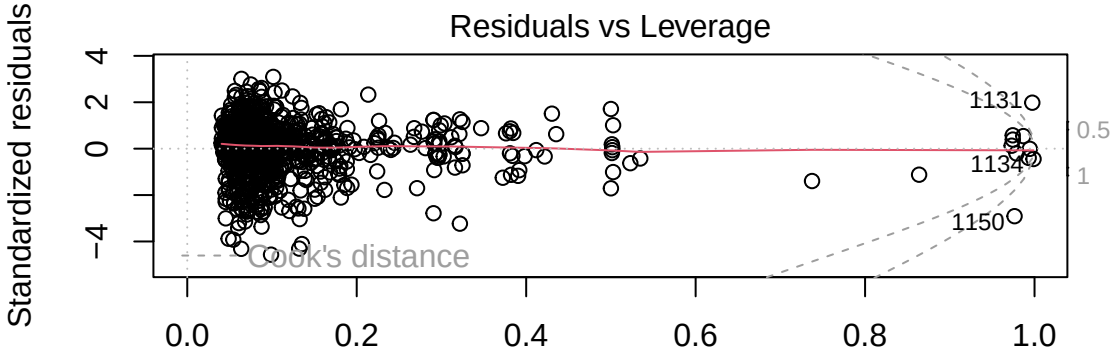
```
## 730        NaN 1.00000000          NaN
## 784  -4.612299 0.09889402  0.017225833
## 930        NaN 1.00000000          NaN
## 1131  1.983754 0.99734867 11.099672796
## 1150 -2.919838 0.97641332  2.634947247
```

Understanding the above residual-Leverage plot and by referring the rule of thumb for Cook's distance, we can state that we definitely have numerous high influential points. Among them, points 1131,113,1150,692,730,784,930 seems to be the most influential.

VI. Transformations

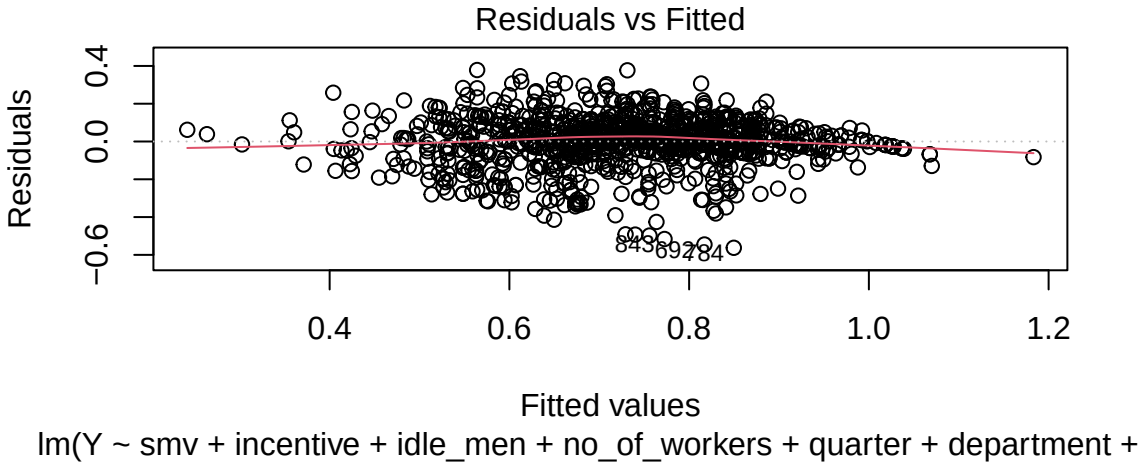
1. For Influential Points:

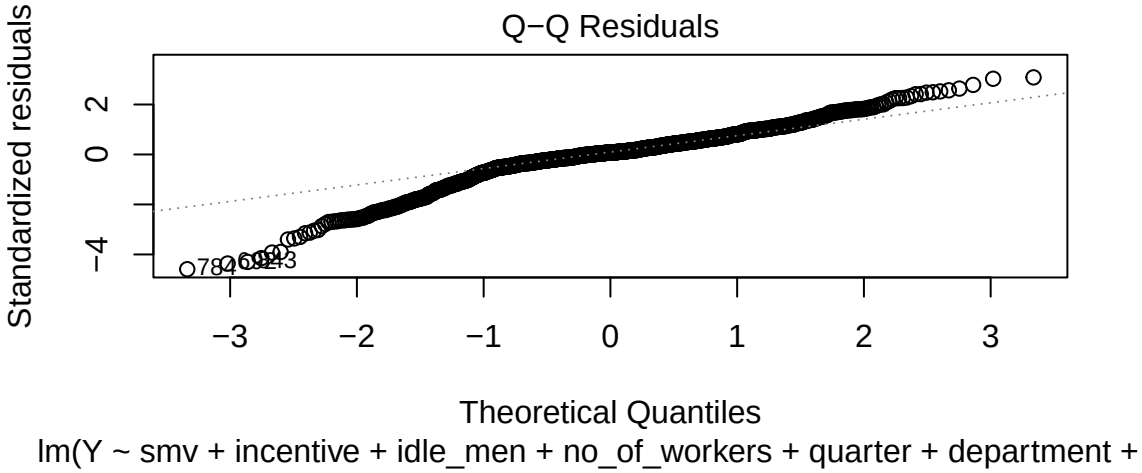


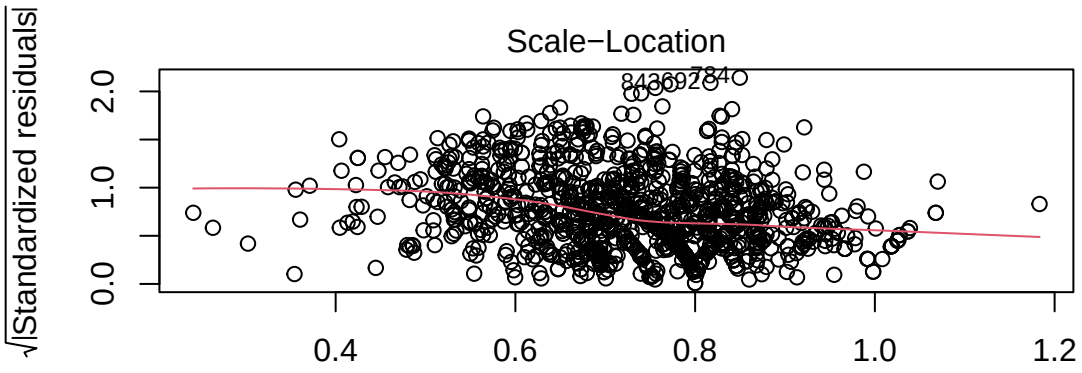


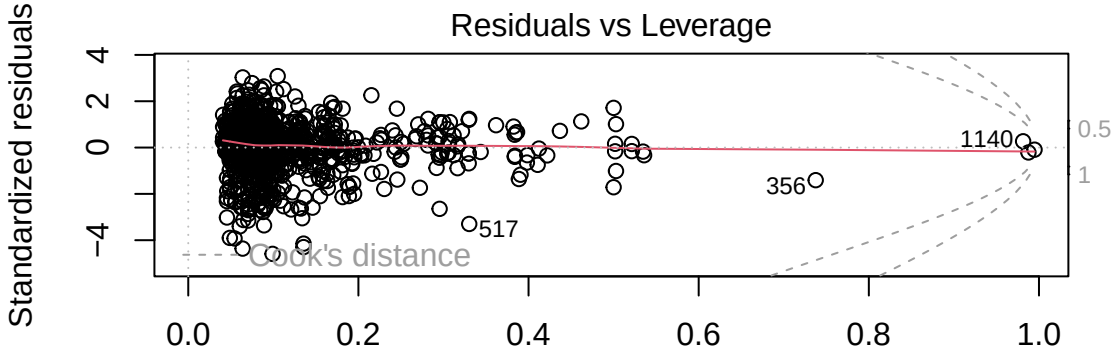
R^2adj before transformation = 0.4496

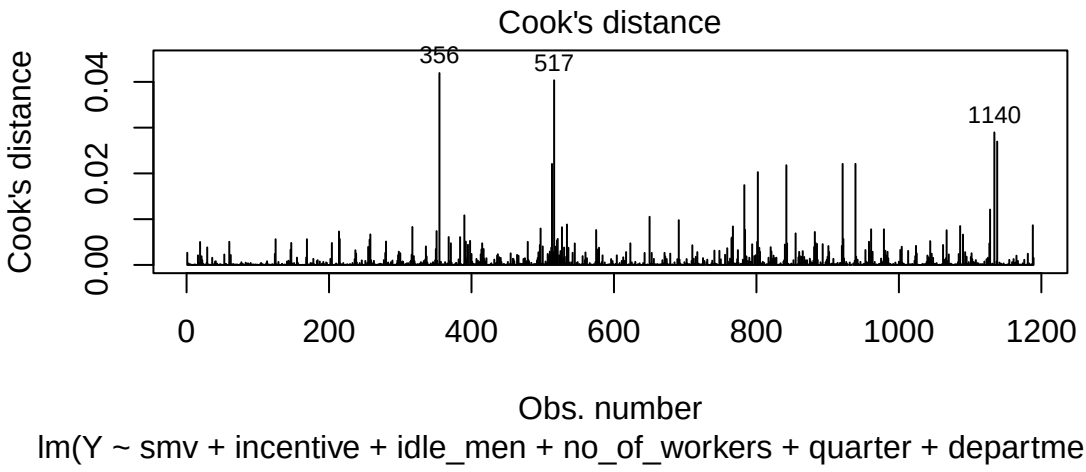
Based on the plots above, we can discern the data points with high leverage. To assess their impact on our model, we will proceed by removing the data points that exceed three times the mean value of Cook's distance as a critical threshold, using plots for evaluation.






Scale−Location
√|Standardized residuals|
0.0
1.0
2.0
0.4
0.6
0.8
1.0
1.2
Fitted values
lm(Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +

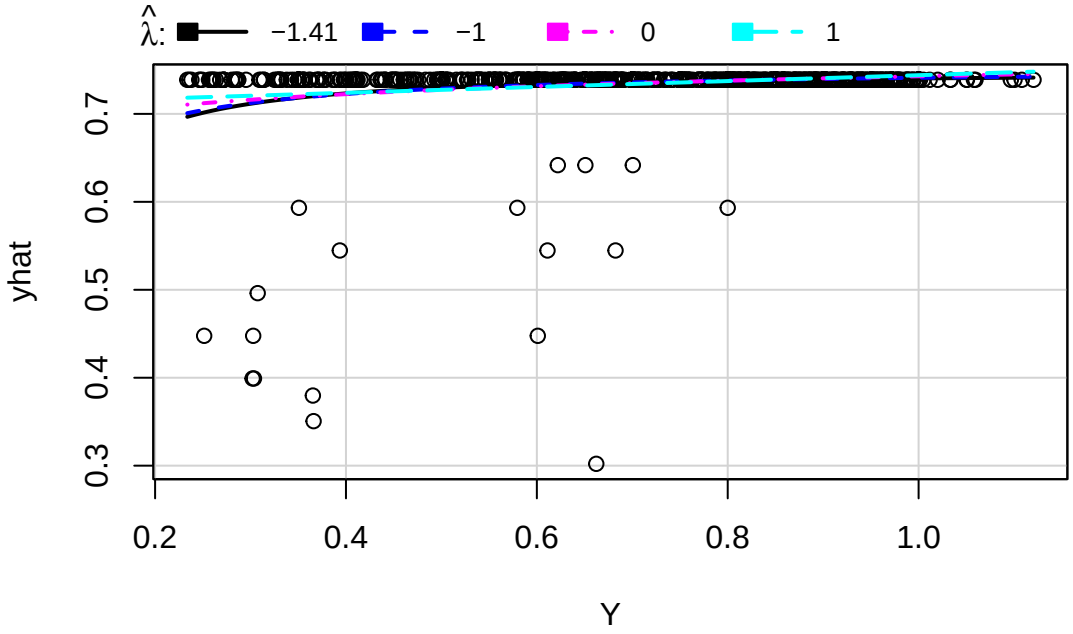

Residuals vs Leverage
Standardized residuals
−4
0
2
4
0.0
0.2
0.4
0.6
0.8
1.0
1140
356
517
Cook's distance
0.5
1
Leverage
lm(Y ~ smv + incentive + idle_men + no_of_workers + quarter + department +



Influential Data Point Row Numbers: **28, 1129, 1131, 1134, 1138, 1139, 1149, 1150**
R^2adj after transformation = 0.4503
After removing the highly influential data points, we have noted that there is no significant difference in the residual standard error, R^2, or R^2adj values. As a result, we can confidently proceed with the reduced number of observations.

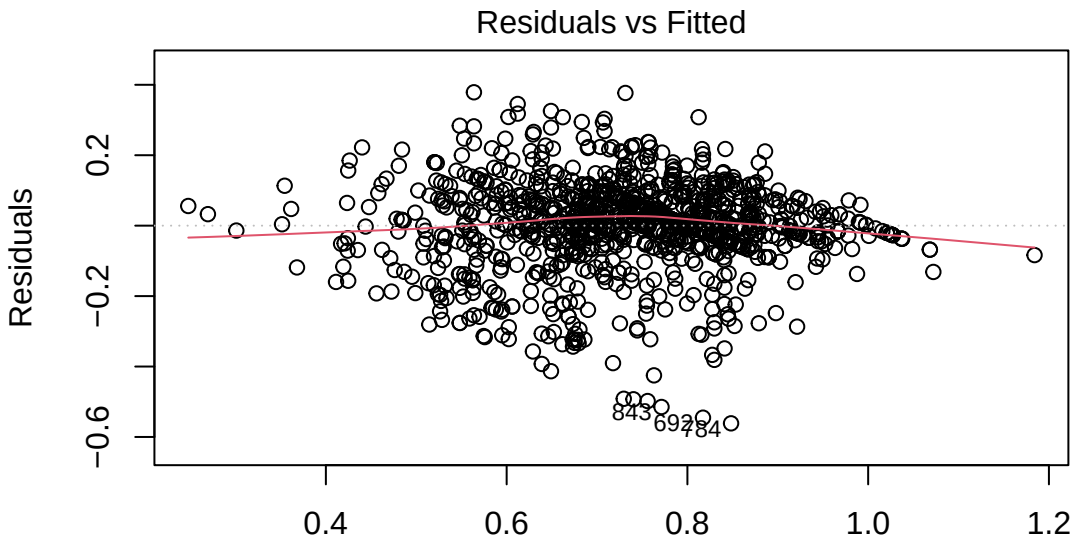
2. For Linearity:

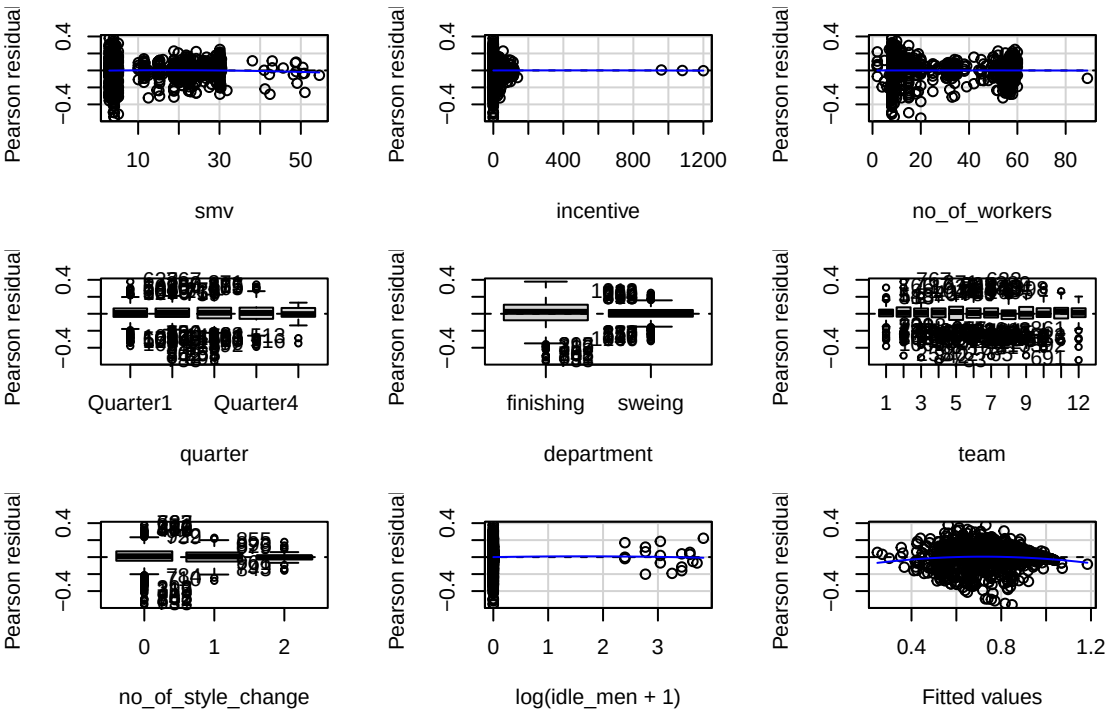
Since the residual plots for "idle_men" showed a quadratic trend, we will need to find a transformation to build a model that follows linearity assumption.




```
##      lambda      RSS
## 1 -1.410593 1.149083
## 2 -1.000000 1.149606
## 3  0.000000 1.154782
## 4  1.000000 1.163153
```

Based on the invResPlot analysis, we can determine that the appropriate lambda value for this transformation is 0, as "idle_men" contains values of 0, making an inverse transformation unsuitable. Therefore, I will introduce a log term for "idle_men + 1" in my model instead of using the variable itself.





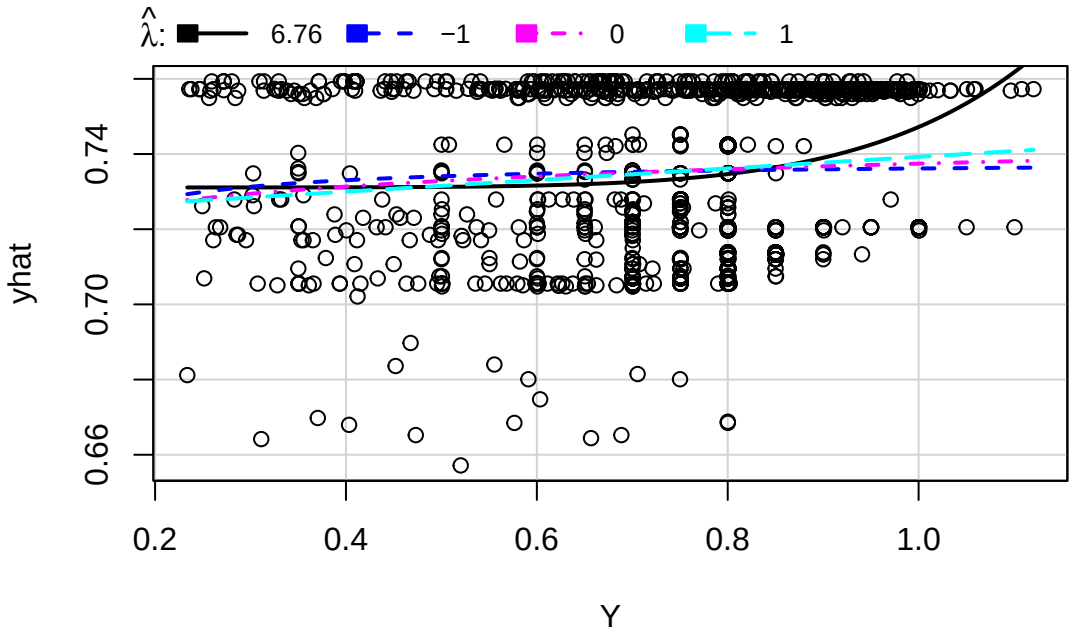

```
##                    Test stat Pr(>|Test stat|)
## smv                  -1.8508        0.0644811 .
## incentive            -0.9534        0.3406199
```

```
## no_of_workers        -0.2564        0.7976709
## quarter
## department
## team
## no_of_style_change
## log(idle_men + 1)    -0.1651        0.8688748
## Tukey test           -3.6390        0.0002737 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R^2adj after transformation = 0.4508

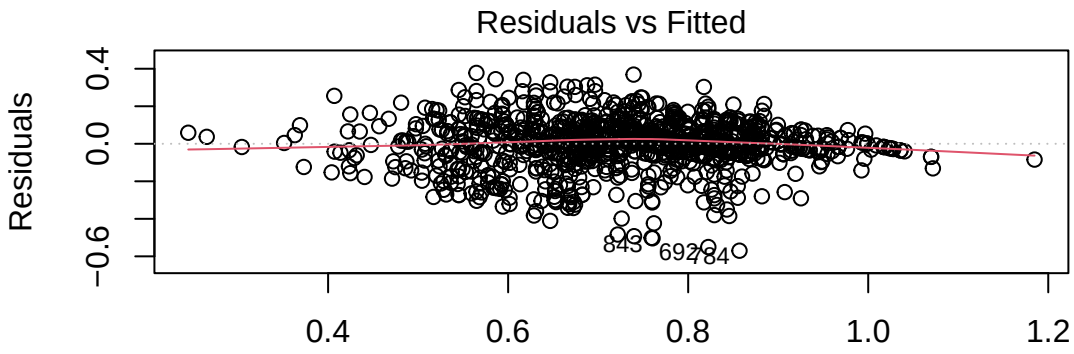
Checking the fitted residual and the residual vs log(idle_men+1) plot after the transformation, we can say that the linearity assumption is invalid for the model. The transformation on idle_men didn't result in improvement in the R^2adj levels and hence no transformation was included in the final model.

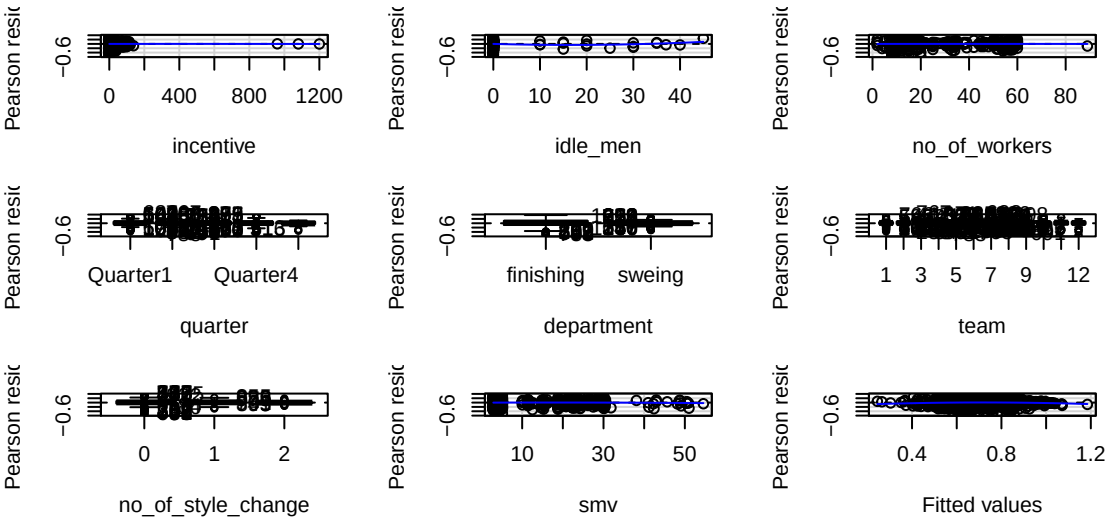
We also observe a slight quadratic trend in the residual vs smv plots:




```
##      lambda       RSS
## 1  6.759386 0.5318717
## 2 -1.000000 0.5544964
## 3  0.000000 0.5516594
## 4  1.000000 0.5474664
```

Based on the invResPlot, we can say that the appropriate lambda value is 6 for this transformation. Thus I will be introducing a smv^6 term instead of smv in my model.






```
##                    Test stat Pr(>|Test stat|)
## incentive            -0.9716        0.3314931
## idle_men              1.4854        0.1377435
## no_of_workers        -0.1318        0.8951324
## quarter
## department
## team
## no_of_style_change
## smv                  -1.9406        0.0525699 .
## Tukey test           -3.7010        0.0002147 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R^2adj after transformation = 0.4486

Checking the fitted residual and the residual vs smv[6] plot after the transformation, we can say that the linearity assumption is still invalid for the model.The transformation on smv didn't result in improvement in the R^2adj levels and hence no transformation was included.

3. For Linearity, Normality and Homoscedasticity:

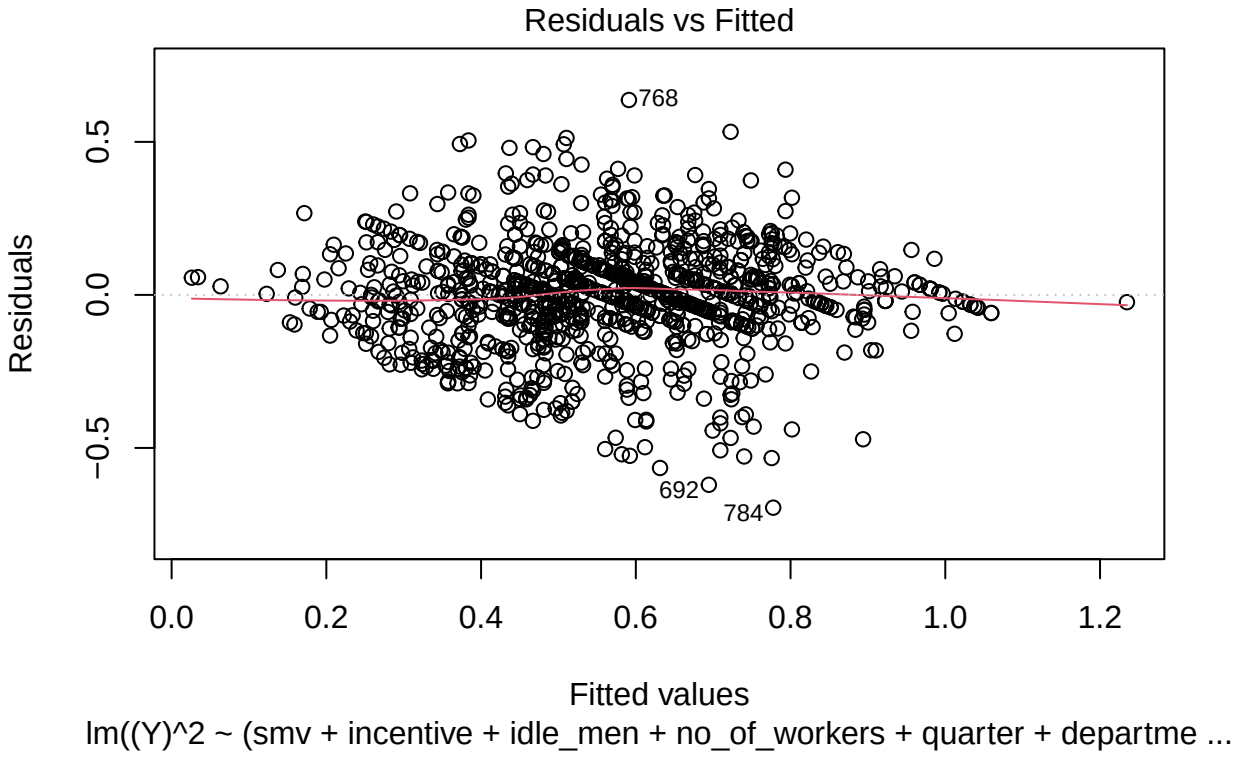
Evaluating a transformation on the response variable,
Response Variable Y^2:

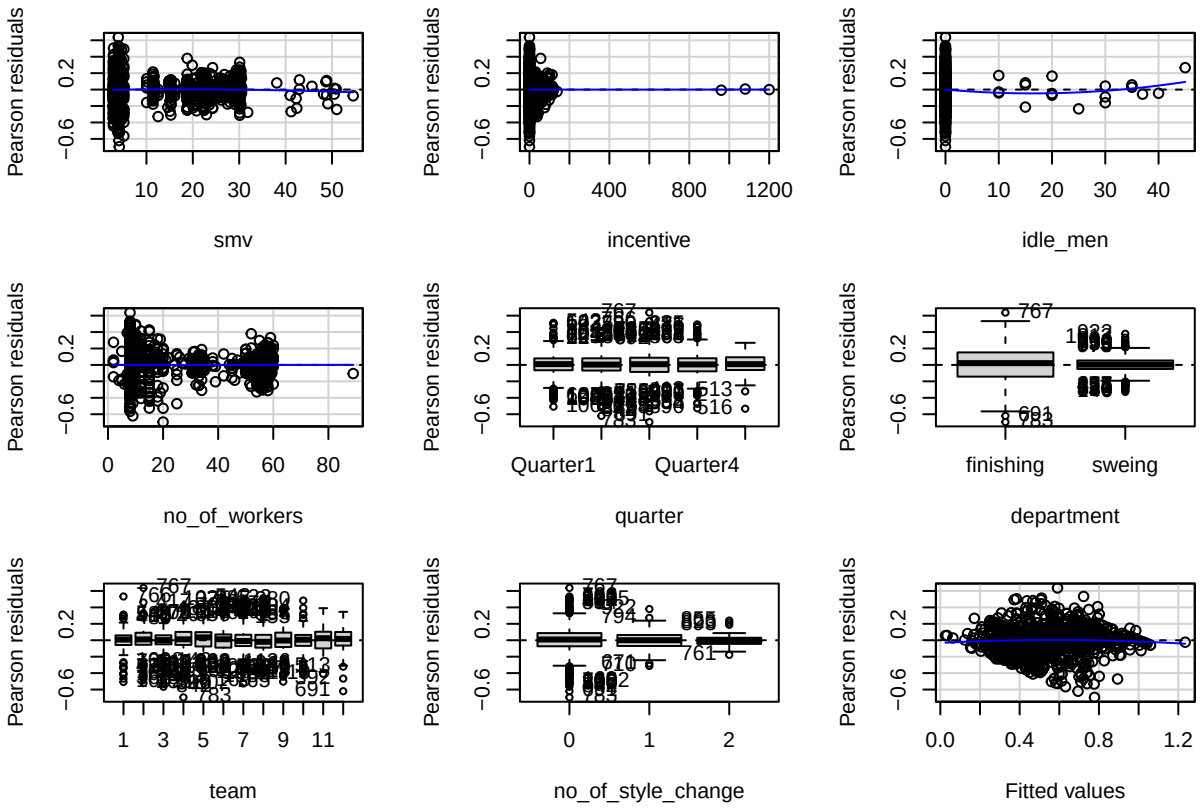
```
##
## Call:
## lm(formula = (Y)^2 ~ (smv + incentive + idle_men + no_of_workers +
##     quarter + department + team + no_of_style_change)^2 - smv:department -
##     incentive:no_of_workers - incentive:idle_men - incentive:no_of_style_change -
##     idle_men:no_of_workers - smv:incentive - smv:idle_men - smv:quarter -
##     smv:team - smv:no_of_style_change - quarter:department -
##     department:no_of_style_change - quarter:no_of_style_change -
##     no_of_workers:no_of_style_change - idle_men:no_of_style_change -
##     idle_men:team - idle_men:department - idle_men:quarter, data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69518 -0.07243  0.00683  0.08136  0.63685
##
## Coefficients: (6 not defined because of singularities)
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    8.605e-01  8.707e-02   9.883  < 2e-16 ***
## smv                           -1.651e-02  1.083e-02  -1.525 0.127577
## incentive                     -6.572e-04  7.289e-04  -0.902 0.367455
## idle_men                      -3.244e-03  1.722e-03  -1.885 0.059753 .
## no_of_workers                 -1.172e-02  6.941e-03  -1.689 0.091601 .
## quarterQuarter2                7.443e-02  4.695e-02   1.585 0.113195
## quarterQuarter3                1.585e-02  6.235e-02   0.254 0.799333
## quarterQuarter4               -6.237e-02  5.730e-02  -1.088 0.276638
## quarterQuarter5                3.614e-01  1.023e-01   3.535 0.000426 ***
## departmentsweing               1.550e+00  4.355e-01   3.559 0.000388 ***
## team2                         -3.949e-01  9.383e-02  -4.208 2.79e-05 ***
## team3                         -4.314e-03  1.176e-01  -0.037 0.970730
## team4                         -4.356e-01  1.043e-01  -4.175 3.22e-05 ***
## team5                         -4.746e-01  9.724e-02  -4.880 1.22e-06 ***
## team6                         -6.188e-01  9.746e-02  -6.349 3.21e-10 ***
## team7                         -5.194e-01  1.120e-01  -4.640 3.92e-06 ***
## team8                         -4.912e-01  1.083e-01  -4.536 6.38e-06 ***
## team9                         -3.520e-01  1.060e-01  -3.322 0.000924 ***
## team10                        -5.512e-01  9.967e-02  -5.530 4.02e-08 ***
## team11                        -4.759e-01  9.920e-02  -4.798 1.83e-06 ***
## team12                        -4.882e-01  1.731e-01  -2.821 0.004882 **
## no_of_style_change1           -5.082e-01  1.937e-01  -2.623 0.008837 **
## no_of_style_change2           -2.206e-01  1.367e-01  -1.614 0.106913
## smv:no_of_workers              2.406e-04  1.944e-04   1.238 0.216146
## incentive:quarterQuarter2      5.182e-04  7.095e-04   0.730 0.465349
```

```
## incentive:quarterQuarter3      -2.148e-04  8.465e-04  -0.254 0.799693
## incentive:quarterQuarter4      -2.920e-04  8.410e-04  -0.347 0.728510
## incentive:quarterQuarter5       4.798e-03  1.661e-03   2.889 0.003944 **
## incentive:departmentsweing      7.154e-03  7.821e-04   9.148  < 2e-16 ***
## incentive:team2                 5.368e-06  2.292e-04   0.023 0.981318
## incentive:team3                -7.884e-04  1.141e-03  -0.691 0.489606
## incentive:team4                -1.441e-03  1.199e-03  -1.202 0.229811
## incentive:team5                -3.349e-03  1.252e-03  -2.674 0.007604 **
## incentive:team6                -2.564e-03  1.474e-03  -1.740 0.082144 .
## incentive:team7                 6.150e-04  1.279e-03   0.481 0.630864
## incentive:team8                 5.929e-04  1.218e-03   0.487 0.626658
## incentive:team9                 1.683e-04  1.456e-03   0.116 0.907992
## incentive:team10               -1.076e-03  1.199e-03  -0.898 0.369542
## incentive:team11                2.022e-05  2.130e-03   0.009 0.992427
## incentive:team12                2.515e-04  2.419e-04   1.040 0.298803
## no_of_workers:quarterQuarter2  -1.427e-03  9.957e-04  -1.433 0.152165
## no_of_workers:quarterQuarter3  -3.203e-04  1.075e-03  -0.298 0.765796
## no_of_workers:quarterQuarter4   8.884e-04  1.193e-03   0.745 0.456481
## no_of_workers:quarterQuarter5  -1.391e-02  2.718e-03  -5.116 3.69e-07 ***
## no_of_workers:departmentsweing -2.467e-02  5.071e-03  -4.864 1.32e-06 ***
## no_of_workers:team2             3.419e-02  7.981e-03   4.284 2.00e-05 ***
## no_of_workers:team3             9.011e-03  9.390e-03   0.960 0.337447
## no_of_workers:team4             3.201e-02  8.410e-03   3.807 0.000149 ***
## no_of_workers:team5             3.393e-02  8.513e-03   3.986 7.19e-05 ***
## no_of_workers:team6             3.476e-02  8.380e-03   4.148 3.62e-05 ***
## no_of_workers:team7             3.480e-02  9.774e-03   3.560 0.000387 ***
## no_of_workers:team8             3.693e-02  9.667e-03   3.820 0.000141 ***
## no_of_workers:team9             3.014e-02  1.063e-02   2.835 0.004673 **
## no_of_workers:team10            4.312e-02  8.963e-03   4.811 1.72e-06 ***
## no_of_workers:team11            3.432e-02  8.445e-03   4.064 5.18e-05 ***
## no_of_workers:team12            4.940e-02  1.820e-02   2.714 0.006754 **
## quarterQuarter2:team2          -6.902e-02  6.006e-02  -1.149 0.250732
## quarterQuarter3:team2          -1.008e-02  7.956e-02  -0.127 0.899155
## quarterQuarter4:team2          -5.205e-02  7.095e-02  -0.734 0.463354
## quarterQuarter5:team2          -1.177e-01  1.314e-01  -0.896 0.370465
## quarterQuarter2:team3          -5.665e-02  6.744e-02  -0.840 0.401126
## quarterQuarter3:team3          -1.416e-01  8.127e-02  -1.743 0.081652 .
## quarterQuarter4:team3           2.872e-02  7.538e-02   0.381 0.703263
## quarterQuarter5:team3          -2.516e-01  1.329e-01  -1.894 0.058523 .
## quarterQuarter2:team4           6.470e-02  6.245e-02   1.036 0.300456
## quarterQuarter3:team4          -1.406e-02  7.622e-02  -0.184 0.853711
## quarterQuarter4:team4           1.097e-01  7.655e-02   1.434 0.152006
## quarterQuarter5:team4           1.286e-01  1.301e-01   0.989 0.323130
## quarterQuarter2:team5           4.577e-02  7.014e-02   0.653 0.514165
## quarterQuarter3:team5           1.604e-01  8.200e-02   1.956 0.050734 .
## quarterQuarter4:team5           1.276e-01  8.290e-02   1.540 0.123935
## quarterQuarter5:team5           1.995e-01  1.436e-01   1.390 0.164888
## quarterQuarter2:team6          -1.174e-02  6.465e-02  -0.182 0.855928
## quarterQuarter3:team6           7.650e-02  7.693e-02   0.994 0.320259
## quarterQuarter4:team6           2.898e-02  7.387e-02   0.392 0.694895
## quarterQuarter5:team6           3.663e-02  1.305e-01   0.281 0.779060
## quarterQuarter2:team7          -5.813e-02  6.726e-02  -0.864 0.387628
## quarterQuarter3:team7          -5.239e-02  7.983e-02  -0.656 0.511829
## quarterQuarter4:team7          -2.606e-02  7.687e-02  -0.339 0.734708
```

```
## quarterQuarter5:team7           2.658e-01  1.580e-01   1.683 0.092729 .
## quarterQuarter2:team8          -9.552e-02  6.454e-02  -1.480 0.139139
## quarterQuarter3:team8          -5.903e-02  7.740e-02  -0.763 0.445789
## quarterQuarter4:team8          -6.820e-02  7.641e-02  -0.892 0.372344
## quarterQuarter5:team8           1.233e-01  1.438e-01   0.857 0.391633
## quarterQuarter2:team9          -1.010e-01  6.129e-02  -1.647 0.099756 .
## quarterQuarter3:team9          -7.978e-02  7.755e-02  -1.029 0.303829
## quarterQuarter4:team9          -1.859e-01  7.653e-02  -2.430 0.015277 *
## quarterQuarter5:team9           1.008e-01  1.388e-01   0.726 0.467872
## quarterQuarter2:team10         -1.158e-01  6.337e-02  -1.828 0.067895 .
## quarterQuarter3:team10         -4.930e-02  7.892e-02  -0.625 0.532350
## quarterQuarter4:team10          5.477e-02  7.475e-02   0.733 0.463904
## quarterQuarter5:team10          7.951e-03  1.294e-01   0.061 0.951015
## quarterQuarter2:team11         -6.229e-02  6.742e-02  -0.924 0.355744
## quarterQuarter3:team11         -1.705e-01  8.155e-02  -2.091 0.036755 *
## quarterQuarter4:team11         -5.904e-02  8.154e-02  -0.724 0.469209
## quarterQuarter5:team11          5.000e-02  1.302e-01   0.384 0.701039
## quarterQuarter2:team12         -2.002e-02  6.330e-02  -0.316 0.751883
## quarterQuarter3:team12          1.153e-02  7.677e-02   0.150 0.880609
## quarterQuarter4:team12         -2.451e-02  7.243e-02  -0.338 0.735144
## quarterQuarter5:team12         -1.644e-01  1.299e-01  -1.266 0.205809
## departmentsweing:team2         -1.523e+00  3.843e-01  -3.964 7.88e-05 ***
## departmentsweing:team3         -4.313e-01  4.459e-01  -0.967 0.333629
## departmentsweing:team4         -1.369e+00  4.034e-01  -3.394 0.000715 ***
## departmentsweing:team5         -1.435e+00  4.229e-01  -3.393 0.000717 ***
## departmentsweing:team6         -1.218e+00  3.954e-01  -3.081 0.002116 **
## departmentsweing:team7         -1.420e+00  4.714e-01  -3.011 0.002665 **
## departmentsweing:team8         -1.573e+00  4.626e-01  -3.401 0.000696 ***
## departmentsweing:team9         -1.315e+00  5.124e-01  -2.566 0.010426 *
## departmentsweing:team10        -1.814e+00  4.254e-01  -4.264 2.19e-05 ***
## departmentsweing:team11        -1.531e+00  4.186e-01  -3.657 0.000268 ***
## departmentsweing:team12        -1.972e+00  5.534e-01  -3.563 0.000382 ***
## team2:no_of_style_change1       5.500e-01  1.901e-01   2.894 0.003883 **
## team3:no_of_style_change1       5.105e-01  2.055e-01   2.484 0.013139 *
## team4:no_of_style_change1       5.400e-01  2.019e-01   2.675 0.007591 **
## team5:no_of_style_change1       3.780e-01  2.053e-01   1.841 0.065886 .
## team6:no_of_style_change1       4.334e-01  2.066e-01   2.098 0.036137 *
## team7:no_of_style_change1       5.298e-01  2.045e-01   2.591 0.009694 **
## team8:no_of_style_change1       5.507e-01  2.096e-01   2.627 0.008738 **
## team9:no_of_style_change1       5.726e-01  2.081e-01   2.752 0.006020 **
## team10:no_of_style_change1      4.445e-01  2.095e-01   2.122 0.034036 *
## team11:no_of_style_change1      5.772e-01  2.769e-01   2.084 0.037359 *
## team12:no_of_style_change1             NA         NA      NA       NA
## team2:no_of_style_change2              NA         NA      NA       NA
## team3:no_of_style_change2       1.941e-01  1.622e-01   1.197 0.231643
## team4:no_of_style_change2      -4.306e-02  1.907e-01  -0.226 0.821433
## team5:no_of_style_change2              NA         NA      NA       NA
## team6:no_of_style_change2              NA         NA      NA       NA
## team7:no_of_style_change2       2.475e-01  2.261e-01   1.095 0.273896
## team8:no_of_style_change2       2.378e-01  1.559e-01   1.525 0.127594
## team9:no_of_style_change2              NA         NA      NA       NA
## team10:no_of_style_change2      3.634e-01  2.296e-01   1.583 0.113796
## team11:no_of_style_change2      4.100e-01  1.551e-01   2.643 0.008337 **
## team12:no_of_style_change2             NA         NA      NA       NA
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.17 on 1062 degrees of freedom
## Multiple R-squared:  0.5328, Adjusted R-squared:  0.4774
## F-statistic: 9.613 on 126 and 1062 DF,  p-value: < 2.2e-16
```

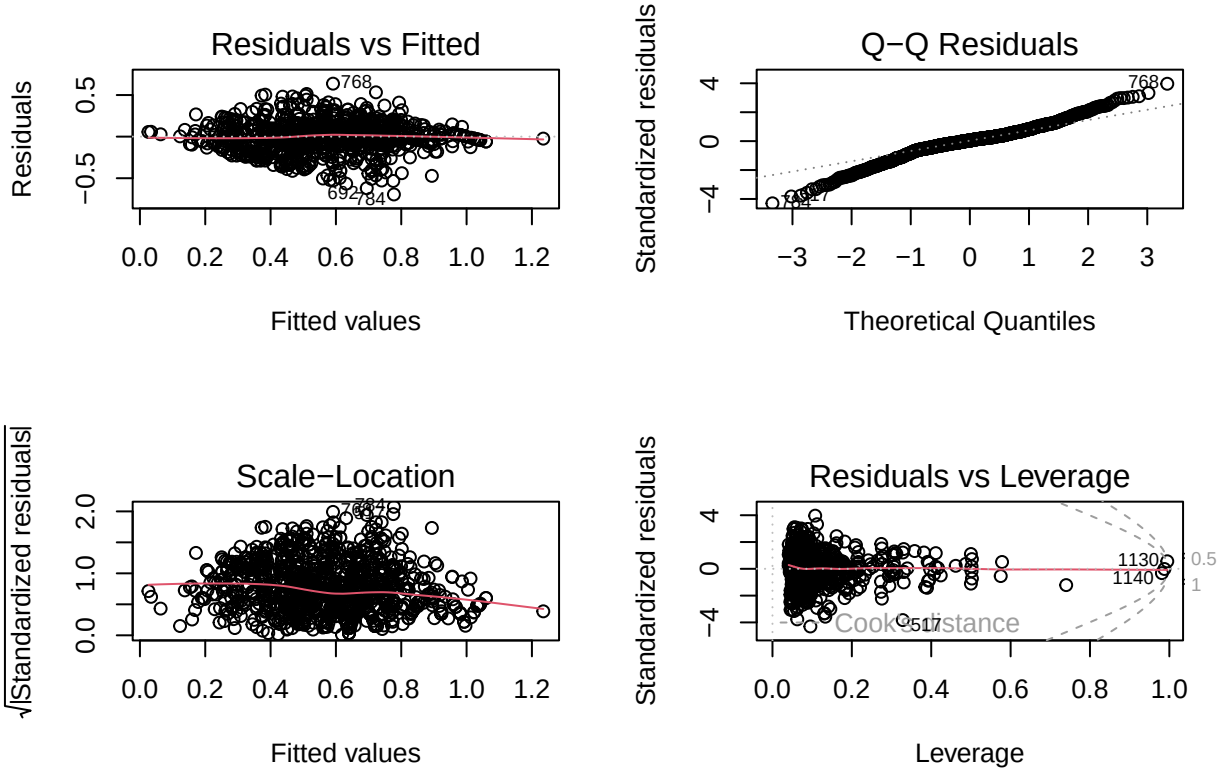





```
##                     Test stat Pr(>|Test stat|)
## smv                   -1.9216          0.05492 .
## incentive              0.0956          0.92382
## idle_men               1.0918          0.27515
## no_of_workers          0.0854          0.93198
## quarter
## department
## team
## no_of_style_change
## Tukey test            -1.5767          0.11486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R^2adj after transformation = 0.4774

The residual-fitted plot exhibits a straight line or equal distribution of points, providing evidence that the linearity assumption holds true in the model.




```
##
##  Shapiro-Wilk normality test
##
## data:  rstandard(mod)
## W = 0.97015, p-value = 6.448e-15
```

Both the plots and the p-value obtained from the Shapiro-Wilk tests confirm that the assumption of normality is still not met in the model

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 2.944745, Df = 1, p = 0.086157
```

The residual-fitted values plot indicates no discernible trend, while the Scale-location plot demonstrates a significant trend, rendering the assumption of homoscedasticity invalid. However, upon conducting the Breusch-Pagan test, we obtain a p-value greater than 0.05, confirming the presence of homoscedasticity in the model.

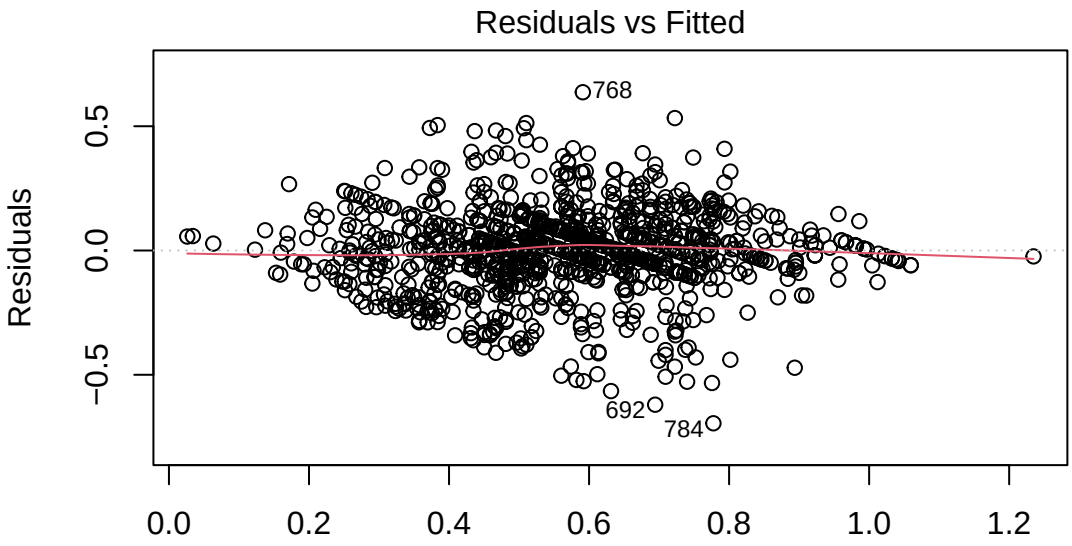
VII. Final Model

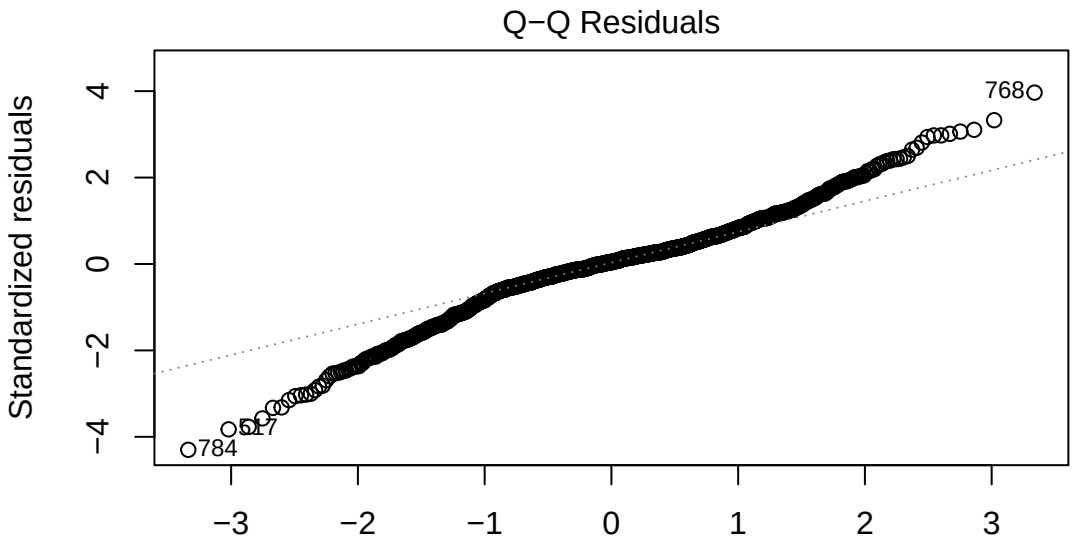
```
##
## Call:
## lm(formula = (Y)^2 ~ (smv + incentive + idle_men + no_of_workers +
##     quarter + department + team + no_of_style_change)^2 - smv:department -
##     incentive:no_of_workers - incentive:idle_men - incentive:no_of_style_change -
##     idle_men:no_of_workers - smv:incentive - smv:idle_men - smv:quarter -
##     smv:team - smv:no_of_style_change - quarter:department -
##     department:no_of_style_change - quarter:no_of_style_change -
##     no_of_workers:no_of_style_change - idle_men:no_of_style_change -
##     idle_men:team - idle_men:department - idle_men:quarter, data = garwork)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69518 -0.07243  0.00683  0.08136  0.63685
##
## Coefficients: (6 not defined because of singularities)
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    8.605e-01  8.707e-02   9.883  < 2e-16 ***
## smv                           -1.651e-02  1.083e-02  -1.525 0.127577
## incentive                     -6.572e-04  7.289e-04  -0.902 0.367455
## idle_men                      -3.244e-03  1.722e-03  -1.885 0.059753 .
## no_of_workers                 -1.172e-02  6.941e-03  -1.689 0.091601 .
## quarterQuarter2                7.443e-02  4.695e-02   1.585 0.113195
## quarterQuarter3                1.585e-02  6.235e-02   0.254 0.799333
## quarterQuarter4               -6.237e-02  5.730e-02  -1.088 0.276638
## quarterQuarter5                3.614e-01  1.023e-01   3.535 0.000426 ***
## departmentsweing               1.550e+00  4.355e-01   3.559 0.000388 ***
## team2                         -3.949e-01  9.383e-02  -4.208 2.79e-05 ***
## team3                         -4.314e-03  1.176e-01  -0.037 0.970730
## team4                         -4.356e-01  1.043e-01  -4.175 3.22e-05 ***
## team5                         -4.746e-01  9.724e-02  -4.880 1.22e-06 ***
## team6                         -6.188e-01  9.746e-02  -6.349 3.21e-10 ***
## team7                         -5.194e-01  1.120e-01  -4.640 3.92e-06 ***
## team8                         -4.912e-01  1.083e-01  -4.536 6.38e-06 ***
## team9                         -3.520e-01  1.060e-01  -3.322 0.000924 ***
## team10                        -5.512e-01  9.967e-02  -5.530 4.02e-08 ***
## team11                        -4.759e-01  9.920e-02  -4.798 1.83e-06 ***
## team12                        -4.882e-01  1.731e-01  -2.821 0.004882 **
## no_of_style_change1           -5.082e-01  1.937e-01  -2.623 0.008837 **
## no_of_style_change2           -2.206e-01  1.367e-01  -1.614 0.106913
## smv:no_of_workers              2.406e-04  1.944e-04   1.238 0.216146
## incentive:quarterQuarter2      5.182e-04  7.095e-04   0.730 0.465349
## incentive:quarterQuarter3     -2.148e-04  8.465e-04  -0.254 0.799693
## incentive:quarterQuarter4     -2.920e-04  8.410e-04  -0.347 0.728510
## incentive:quarterQuarter5      4.798e-03  1.661e-03   2.889 0.003944 **
## incentive:departmentsweing     7.154e-03  7.821e-04   9.148  < 2e-16 ***
## incentive:team2                5.368e-06  2.292e-04   0.023 0.981318
## incentive:team3               -7.884e-04  1.141e-03  -0.691 0.489606
## incentive:team4               -1.441e-03  1.199e-03  -1.202 0.229811
## incentive:team5               -3.349e-03  1.252e-03  -2.674 0.007604 **
## incentive:team6               -2.564e-03  1.474e-03  -1.740 0.082144 .
## incentive:team7                6.150e-04  1.279e-03   0.481 0.630864
```

```
## incentive:team8                 5.929e-04  1.218e-03   0.487 0.626658
## incentive:team9                 1.683e-04  1.456e-03   0.116 0.907992
## incentive:team10               -1.076e-03  1.199e-03  -0.898 0.369542
## incentive:team11                2.022e-05  2.130e-03   0.009 0.992427
## incentive:team12                2.515e-04  2.419e-04   1.040 0.298803
## no_of_workers:quarterQuarter2  -1.427e-03  9.957e-04  -1.433 0.152165
## no_of_workers:quarterQuarter3  -3.203e-04  1.075e-03  -0.298 0.765796
## no_of_workers:quarterQuarter4   8.884e-04  1.193e-03   0.745 0.456481
## no_of_workers:quarterQuarter5  -1.391e-02  2.718e-03  -5.116 3.69e-07 ***
## no_of_workers:departmentsweing -2.467e-02  5.071e-03  -4.864 1.32e-06 ***
## no_of_workers:team2             3.419e-02  7.981e-03   4.284 2.00e-05 ***
## no_of_workers:team3             9.011e-03  9.390e-03   0.960 0.337447
## no_of_workers:team4             3.201e-02  8.410e-03   3.807 0.000149 ***
## no_of_workers:team5             3.393e-02  8.513e-03   3.986 7.19e-05 ***
## no_of_workers:team6             3.476e-02  8.380e-03   4.148 3.62e-05 ***
## no_of_workers:team7             3.480e-02  9.774e-03   3.560 0.000387 ***
## no_of_workers:team8             3.693e-02  9.667e-03   3.820 0.000141 ***
## no_of_workers:team9             3.014e-02  1.063e-02   2.835 0.004673 **
## no_of_workers:team10            4.312e-02  8.963e-03   4.811 1.72e-06 ***
## no_of_workers:team11            3.432e-02  8.445e-03   4.064 5.18e-05 ***
## no_of_workers:team12            4.940e-02  1.820e-02   2.714 0.006754 **
## quarterQuarter2:team2          -6.902e-02  6.006e-02  -1.149 0.250732
## quarterQuarter3:team2          -1.008e-02  7.956e-02  -0.127 0.899155
## quarterQuarter4:team2          -5.205e-02  7.095e-02  -0.734 0.463354
## quarterQuarter5:team2          -1.177e-01  1.314e-01  -0.896 0.370465
## quarterQuarter2:team3          -5.665e-02  6.744e-02  -0.840 0.401126
## quarterQuarter3:team3          -1.416e-01  8.127e-02  -1.743 0.081652 .
## quarterQuarter4:team3           2.872e-02  7.538e-02   0.381 0.703263
## quarterQuarter5:team3          -2.516e-01  1.329e-01  -1.894 0.058523 .
## quarterQuarter2:team4           6.470e-02  6.245e-02   1.036 0.300456
## quarterQuarter3:team4          -1.406e-02  7.622e-02  -0.184 0.853711
## quarterQuarter4:team4           1.097e-01  7.655e-02   1.434 0.152006
## quarterQuarter5:team4           1.286e-01  1.301e-01   0.989 0.323130
## quarterQuarter2:team5           4.577e-02  7.014e-02   0.653 0.514165
## quarterQuarter3:team5           1.604e-01  8.200e-02   1.956 0.050734 .
## quarterQuarter4:team5           1.276e-01  8.290e-02   1.540 0.123935
## quarterQuarter5:team5           1.995e-01  1.436e-01   1.390 0.164888
## quarterQuarter2:team6          -1.174e-02  6.465e-02  -0.182 0.855928
## quarterQuarter3:team6           7.650e-02  7.693e-02   0.994 0.320259
## quarterQuarter4:team6           2.898e-02  7.387e-02   0.392 0.694895
## quarterQuarter5:team6           3.663e-02  1.305e-01   0.281 0.779060
## quarterQuarter2:team7          -5.813e-02  6.726e-02  -0.864 0.387628
## quarterQuarter3:team7          -5.239e-02  7.983e-02  -0.656 0.511829
## quarterQuarter4:team7          -2.606e-02  7.687e-02  -0.339 0.734708
## quarterQuarter5:team7           2.658e-01  1.580e-01   1.683 0.092729 .
## quarterQuarter2:team8          -9.552e-02  6.454e-02  -1.480 0.139139
## quarterQuarter3:team8          -5.903e-02  7.740e-02  -0.763 0.445789
## quarterQuarter4:team8          -6.820e-02  7.641e-02  -0.892 0.372344
## quarterQuarter5:team8           1.233e-01  1.438e-01   0.857 0.391633
## quarterQuarter2:team9          -1.010e-01  6.129e-02  -1.647 0.099756 .
## quarterQuarter3:team9          -7.978e-02  7.755e-02  -1.029 0.303829
## quarterQuarter4:team9          -1.859e-01  7.653e-02  -2.430 0.015277 *
## quarterQuarter5:team9           1.008e-01  1.388e-01   0.726 0.467872
## quarterQuarter2:team10         -1.158e-01  6.337e-02  -1.828 0.067895 .
```

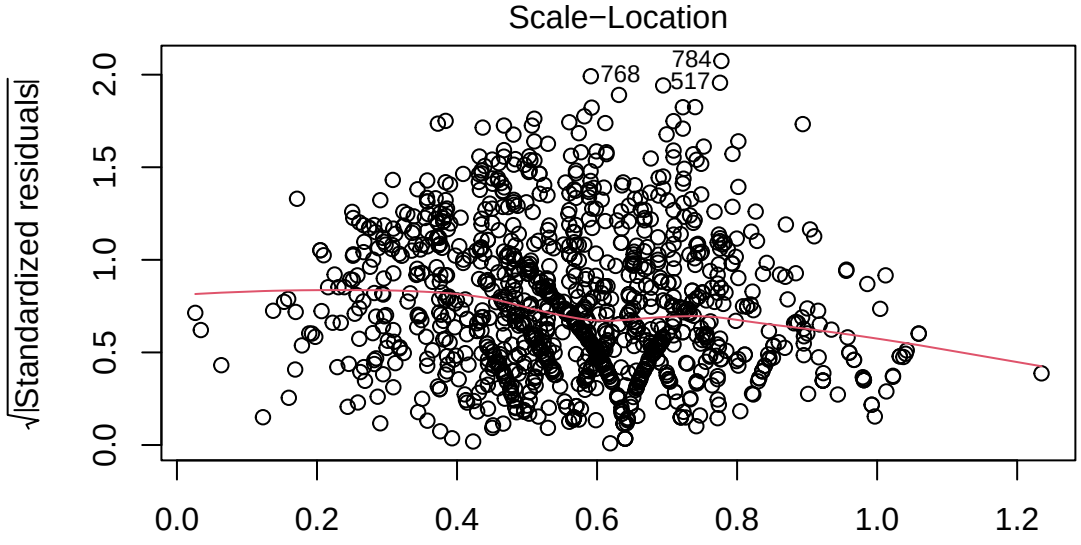
```
## quarterQuarter3:team10         -4.930e-02  7.892e-02  -0.625 0.532350
## quarterQuarter4:team10          5.477e-02  7.475e-02   0.733 0.463904
## quarterQuarter5:team10          7.951e-03  1.294e-01   0.061 0.951015
## quarterQuarter2:team11         -6.229e-02  6.742e-02  -0.924 0.355744
## quarterQuarter3:team11         -1.705e-01  8.155e-02  -2.091 0.036755 *
## quarterQuarter4:team11         -5.904e-02  8.154e-02  -0.724 0.469209
## quarterQuarter5:team11          5.000e-02  1.302e-01   0.384 0.701039
## quarterQuarter2:team12         -2.002e-02  6.330e-02  -0.316 0.751883
## quarterQuarter3:team12          1.153e-02  7.677e-02   0.150 0.880609
## quarterQuarter4:team12         -2.451e-02  7.243e-02  -0.338 0.735144
## quarterQuarter5:team12         -1.644e-01  1.299e-01  -1.266 0.205809
## departmentsweing:team2         -1.523e+00  3.843e-01  -3.964 7.88e-05 ***
## departmentsweing:team3         -4.313e-01  4.459e-01  -0.967 0.333629
## departmentsweing:team4         -1.369e+00  4.034e-01  -3.394 0.000715 ***
## departmentsweing:team5         -1.435e+00  4.229e-01  -3.393 0.000717 ***
## departmentsweing:team6         -1.218e+00  3.954e-01  -3.081 0.002116 **
## departmentsweing:team7         -1.420e+00  4.714e-01  -3.011 0.002665 **
## departmentsweing:team8         -1.573e+00  4.626e-01  -3.401 0.000696 ***
## departmentsweing:team9         -1.315e+00  5.124e-01  -2.566 0.010426 *
## departmentsweing:team10        -1.814e+00  4.254e-01  -4.264 2.19e-05 ***
## departmentsweing:team11        -1.531e+00  4.186e-01  -3.657 0.000268 ***
## departmentsweing:team12        -1.972e+00  5.534e-01  -3.563 0.000382 ***
## team2:no_of_style_change1       5.500e-01  1.901e-01   2.894 0.003883 **
## team3:no_of_style_change1       5.105e-01  2.055e-01   2.484 0.013139 *
## team4:no_of_style_change1       5.400e-01  2.019e-01   2.675 0.007591 **
## team5:no_of_style_change1       3.780e-01  2.053e-01   1.841 0.065886 .
## team6:no_of_style_change1       4.334e-01  2.066e-01   2.098 0.036137 *
## team7:no_of_style_change1       5.298e-01  2.045e-01   2.591 0.009694 **
## team8:no_of_style_change1       5.507e-01  2.096e-01   2.627 0.008738 **
## team9:no_of_style_change1       5.726e-01  2.081e-01   2.752 0.006020 **
## team10:no_of_style_change1      4.445e-01  2.095e-01   2.122 0.034036 *
## team11:no_of_style_change1      5.772e-01  2.769e-01   2.084 0.037359 *
## team12:no_of_style_change1             NA         NA      NA       NA
## team2:no_of_style_change2              NA         NA      NA       NA
## team3:no_of_style_change2       1.941e-01  1.622e-01   1.197 0.231643
## team4:no_of_style_change2      -4.306e-02  1.907e-01  -0.226 0.821433
## team5:no_of_style_change2              NA         NA      NA       NA
## team6:no_of_style_change2              NA         NA      NA       NA
## team7:no_of_style_change2       2.475e-01  2.261e-01   1.095 0.273896
## team8:no_of_style_change2       2.378e-01  1.559e-01   1.525 0.127594
## team9:no_of_style_change2              NA         NA      NA       NA
## team10:no_of_style_change2      3.634e-01  2.296e-01   1.583 0.113796
## team11:no_of_style_change2      4.100e-01  1.551e-01   2.643 0.008337 **
## team12:no_of_style_change2             NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.17 on 1062 degrees of freedom
## Multiple R-squared:  0.5328, Adjusted R-squared:  0.4774
## F-statistic: 9.613 on 126 and 1062 DF,  p-value: < 2.2e-16


##
##  TOP 10 POSITIVE COEFFICIENTS IN THE MODEL:
```

```
##           departmentsweing                (Intercept)
##                  1.5501180                  0.8604631
## team11:no_of_style_change1  team9:no_of_style_change1
##                  0.5771871                  0.5726113
##  team8:no_of_style_change1  team2:no_of_style_change1
##                  0.5506995                  0.5500422
##  team4:no_of_style_change1  team7:no_of_style_change1
##                  0.5399743                  0.5298029
##  team3:no_of_style_change1 team10:no_of_style_change1
##                  0.5105063                  0.4445395
```

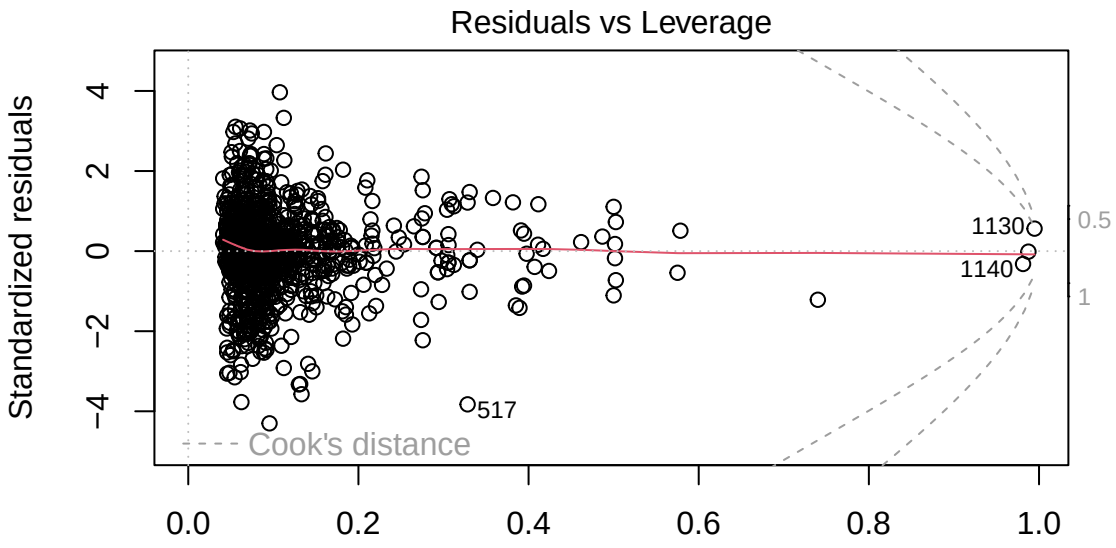




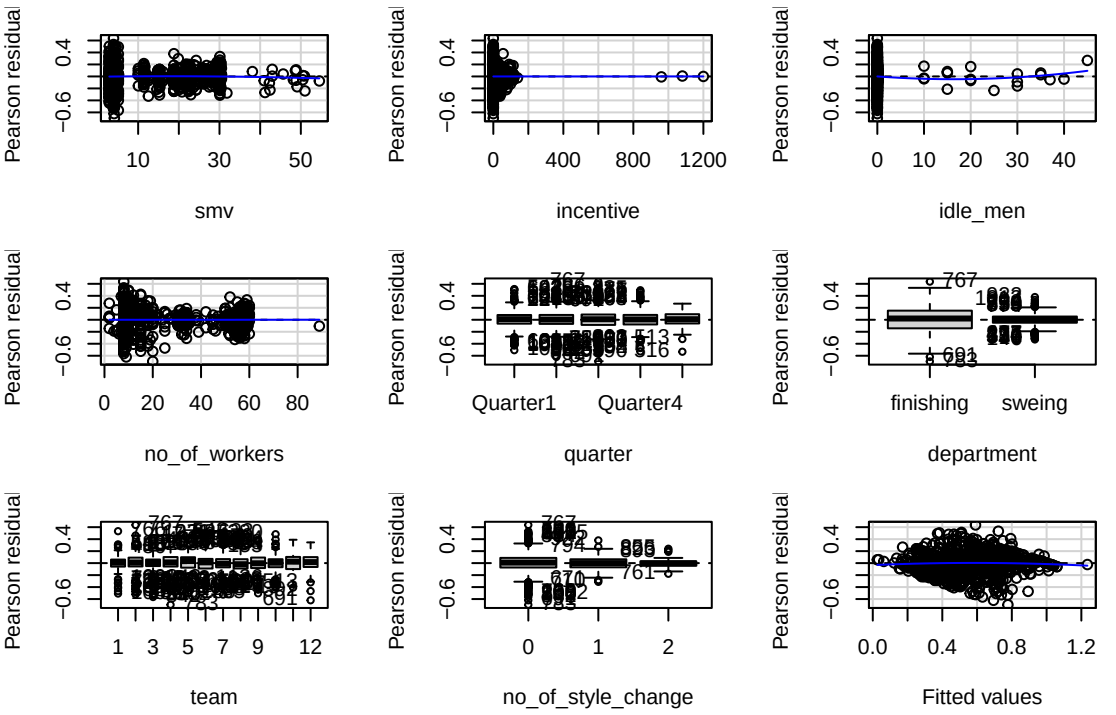
lm((Y)^2 ~ (smv + incentive + idle_men + no_of_workers + quarter + departme



lm((Y)^2 ~ (smv + incentive + idle_men + no_of_workers + quarter + departme



lm((Y)^2 ~ (smv + incentive + idle_men + no_of_workers + quarter + departme




```
##                    Test stat Pr(>|Test stat|)
## smv                  -1.9216          0.05492 .
## incentive             0.0956          0.92382
```

```
## idle_men              1.0918          0.27515
## no_of_workers         0.0854          0.93198
## quarter
## department
## team
## no_of_style_change
## Tukey test           -1.5767          0.11486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 2.944745, Df = 1, p = 0.086157
```

1. Hypothesis Test 1

Coefficient of smv = 0,i.e. the time allocated for each team doesn't affect the actual productivity.
This hypothesis will be helpful for allocating time for tasks and teams:

- If we find that the productivity is not related to the time allocated to the team we can conclude that the time provided to each team is optimum for maximum actual productivity.
- If we can't reject the null hypothesis, their is a high probability that the time provided needs to be optimized for achieving higher efficiency.

```
## Analysis of Variance Table
##
## Model 1: (Y)^2 ~ incentive + idle_men + no_of_workers + quarter + department +
##     team + no_of_style_change + smv:no_of_workers + incentive:quarter +
##     incentive:department + incentive:team + no_of_workers:quarter +
##     no_of_workers:department + no_of_workers:team + quarter:team +
##     department:team + team:no_of_style_change
## Model 2: (Y)^2 ~ (smv + incentive + idle_men + no_of_workers + quarter +
##     department + team + no_of_style_change)^2 - smv:department -
##     incentive:no_of_workers - incentive:idle_men - incentive:no_of_style_change -
##     idle_men:no_of_workers - smv:incentive - smv:idle_men - smv:quarter -
##     smv:team - smv:no_of_style_change - quarter:department -
##     department:no_of_style_change - quarter:no_of_style_change -
##     no_of_workers:no_of_style_change - idle_men:no_of_style_change -
##     idle_men:team - idle_men:department - idle_men:quarter
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1   1063 30.746
## 2   1062 30.679  1  0.067175 2.3254 0.1276
```

p-value = 0.1275774
We don't have enough proof to reject H0.Thus it is very likely that the standard time allocated for a task(in minutes) might not have impact on the actual productivity among the teams.

2. Hypothesis Test 2

Coefficient for “no_of_workers” = 0.05; i.e. number of workers affect the actual productivity by 5%.

This hypothesis will be helpful as we can find if we can increase actual productivity by 5% by increasing the number of workers.

```
## Analysis of Variance Table
##
## Model 1: (Y)^2 ~ smv + incentive + idle_men + quarter + department + team +
##     no_of_style_change + smv:no_of_workers + incentive:quarter +
##     incentive:department + incentive:team + no_of_workers:quarter +
##     no_of_workers:department + no_of_workers:team + quarter:team +
##     department:team + team:no_of_style_change + offset(I(0.05 *
##     no_of_workers))
## Model 2: (Y)^2 ~ (smv + incentive + idle_men + no_of_workers + quarter +
##     department + team + no_of_style_change)^2 - smv:department -
##     incentive:no_of_workers - incentive:idle_men - incentive:no_of_style_change -
##     idle_men:no_of_workers - smv:incentive - smv:idle_men - smv:quarter -
##     smv:team - smv:no_of_style_change - quarter:department -
##     department:no_of_style_change - quarter:no_of_style_change -
##     no_of_workers:no_of_style_change - idle_men:no_of_style_change -
##     idle_men:team - idle_men:department - idle_men:quarter
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1063 32.963
## 2   1062 30.679  1    2.2843 79.073 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p-value = $2.5217631 \times 10^{-18}$

We have enough proof to reject H0.
Thus it is very unlikely to increase productivity by 5% by increasing the number of workers.