REGRESSION ANALYSIS (MATH 1312)

Building a Regression model for predicting the medical insurance charges

Ratnak Saha (S3973869)

Semester 1 – 2024
Due on 13 Jun 2024

1 ABSTRACT

This report presents a comprehensive regression analysis of a medical insurance dataset, aiming to predict insurance charges based on various demographic and health-related attributes. The analysis includes Exploratory Data Analysis (EDA), Multiple Regression Estimation, Model Assessment, Model Adequacy Check, Model Diagnostic Check, Implementation of Corrective Methods, Variable Selection, and Model Validation. Through iterative improvements, the study aims to build a robust and accurate regression model to predict insurance charges, identify key predictors and addressing model diagnostics for improved performance.

2 INTRODUCTION

In this study, we utilized regression analysis to investigate the association between medical insurance expenses and various independent variables that includes the demographics, smoker status and children of an individual. The objective was to quantify the impact of each independent variable on medical charges. To ensure the reliability of our regression model, we meticulously assessed assumptions such as linearity, independence, homoscedasticity, and normality of residuals. We carefully selected independent variables based on theoretical relevance, statistical significance, and multicollinearity assessment. Evaluation metrics like R-squared, adjusted R-squared, and F-statistic were utilized to gauge model fit. Residual analysis was conducted to validate model adequacy, and diagnostic tests were employed to identify and address potential violations of regression assumptions. These steps were crucial in constructing a robust regression model that provides accurate insights into the determinants of healthcare expenditure. The key research questions addressed include:

1. What are the significant predictors of medical insurance charges?
2. How do demographic and health-related factors influence insurance charges?
3. Can we build a robust and accurate regression model to predict insurance charges?

3 DATA DESCRIPTION

The dataset used for this analysis was sourced from Kaggle website, uploaded by Rahul Vyas M. The link for the source of this dataset is https://www.kaggle.com/datasets/rahulvyasm/medical-insurance-cost-prediction/data. This dataset comprises of 2772 rows and 7 columns.

Variable	***Description***
Charges	Cost of the medical insurance
Age	Age of the individual
Sex	Gender of the individual
BMI	Body mass index of the individual
Children	Number of children for the individual
Smoker	Is the individual a smoker?
Region	Region of the individual

4 METHODOLOGY CODES

4.1 EXPLORATORY DATA ANALYSIS (R OUTPUTS)

EDA serves as our initial step to comprehend the dataset. This comprehensive process involves several crucial stages: identifying the data type (categorical or numerical) of both descriptive and target features, handling missing values effectively, summarizing descriptive features of numerical and categorical variables, utilizing histograms to grasp data distribution, employing QQ plots to assess normality, and utilizing ggpairs() plot to detect correlations between regressor and target variables. By diligently undertaking these steps, we gain a profound understanding of the dataset, enabling us to identify potential regressors and construct a regression model with enhanced accuracy and reliability.

Data Structure and Types

- Numerical Variables: age, bmi, charges
- Categorical Variables: sex, smoker, region, children

Descriptive Statistics

Age:

- Range: 18 to 64 years
- Median: 39 years
- Mean: 39.11 years
- The age distribution shows a high frequency of individuals around age 20, followed by a relatively uniform distribution across other ages with a slight decrease towards the older ages.

BMI:

- Range: 15.96 to 53.13
- Median: 30.45
- Mean: 30.70
- The BMI distribution is slightly right-skewed with most values concentrated between 20 and 40. The peak is around the BMI value of 30.

Charges:

- Range: 1,122 to 63,770
- Median: 9,333
- Mean: 13,261
- The charges distribution is heavily right-skewed, indicating that most individuals have lower medical costs, but there are a few with very high charges.

Sex:

Distribution: 49.3% female (1,366) and 50.7% male (1,406), indicating a balanced gender distribution.

Smoker:

Distribution: 79.6% non-smokers (2,208) and 20.4% smokers (564). This indicates a larger proportion of non-smokers in the dataset.

Region:

Distribution: Fairly even across four regions with Northeast (658), Northwest (664), Southeast (766), and Southwest (684).

Children:
Most individuals have zero children, followed by those with one or two children. The number of individuals decreases significantly with more children.

Correlation with the Target Feature

Charges by Smoking Status
The box plot of charges by smoking status shows that smokers tend to have significantly higher medical charges compared to non-smokers. The median charge for smokers is higher, and the interquartile range is also greater, indicating more variability in charges among smokers. This suggests a strong positive correlation between smoking status and medical charges.

Charges by Region
The box plot of charges by region indicates that the median charges are relatively similar across all regions. However, the southeast region has a slightly higher median and a wider interquartile range compared to other regions. This suggests that while region may have some impact on charges, it is not as strong as other factors like smoking status.

Charges by Sex
The box plot of charges by sex shows that males tend to have slightly higher medical charges compared to females. However, the difference in median charges is not substantial, suggesting that sex might have a weaker correlation with medical charges.

Charges by Age
The scatter plot of charges by age shows a positive correlation between age and charges. As age increases, the charges also tend to increase. There are several bands of data points indicating different levels of charges that increase with age, suggesting age is a significant predictor of medical charges.

Charges by BMI
The scatter plot of charges by BMI shows a positive correlation between BMI and charges. Individuals with higher BMI tend to have higher medical charges. However, there is considerable variability in charges for individuals with similar BMI values, indicating that while BMI is an important factor, it interacts with other variables to determine charges.

Charges by Number of Children
The scatter plot of charges by the number of children shows no clear trend or correlation. The charges seem to be distributed across all levels of children, suggesting that the number of children may not be a strong predictor of medical charges.

The Q-Q plot for charges shows that the data points deviate significantly from the straight line, especially in the upper quantiles. This indicates that the distribution of charges is right-skewed and not normally distributed. The presence of many high-value outliers confirms the heavy tail on the right side of the distribution. This lack of normality suggests that transformations or robust regression techniques might be necessary for modeling. The BMI variable also shows slight right-skewness, which could affect the model if not addressed. The categorical variables, particularly smoker, might have a significant impact on charges given the distinct difference in the number of smokers and non-smokers. The balanced distribution of sex ensures no gender bias in the dataset. Regional distribution is even, allowing for region-based analysis without significant skewness. The presence of individuals with a high number of children is rare, which might affect the model's predictions for such cases.

The correlation analysis highlights that age, bmi, and smoker status are the most influential variables affecting charges. Age and BMI have positive correlations with charges, while smoker status significantly increases charges. These insights are crucial for building a predictive model for insurance charges, as they help identify which variables to prioritize. Other factors such as sex, number of children, and region have minimal impact on charges, indicating that they may not need as much emphasis in the modeling process.

The data was then split into training (80%) and test (20%) sets to build and validate the regression models.

4.2 MODEL SPECIFICATION

4.2.1 Model 1: Full Model

Model Summary:

```
"Model Summary"

Call:
lm(formula = charges ~ age + sex + bmi + smoker + children +
    region, data = trainData)

Residuals:
   Min     1Q Median     3Q    Max
-11313  -2803  -1021   1347  30291

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -10990.911    762.204 -14.420  < 2e-16 ***
age                247.172      9.171  26.951  < 2e-16 ***
sexmale             24.622    257.760   0.096 0.923907
bmi                317.846     22.035  14.425  < 2e-16 ***
smokeryes        23759.660    320.150  74.214  < 2e-16 ***
children           522.451    105.435   4.955 7.77e-07 ***
regionnorthwest   -233.538    370.313  -0.631 0.528335
regionsoutheast  -1263.797    374.560  -3.374 0.000753 ***
regionsouthwest  -1186.683    373.442  -3.178 0.001505 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6040 on 2211 degrees of freedom
Multiple R-squared:  0.7486,	Adjusted R-squared:  0.7476
F-statistic: 822.8 on 8 and 2211 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = -10990.911 + 247.172*age + 24.622*sex(male) + 317.846*bmi + 23759.660*smoker(yes) + 522.451*children – 233.538*region(northwest) – 1263.797*region(southeast) – 1186.683*region(southwest)

Coefficients:

- Age: Positive coefficient suggests that as age increases, medical charges also increase, possibly due to age-related health issues. ($p < 0.001$)
- Sex: Not significant, indicating no substantial difference in charges between males and females after accounting for other variables.
- BMI: Significant positive relationship, implying higher BMI leads to higher medical charges, likely due to associated health risks. ($p < 0.001$)

- Smoker: Extremely large positive coefficient, indicating smokers incur significantly higher medical costs. ($p < 0.001$, large positive impact with an estimate of 23759.660)
- Children: Positive relationship, possibly reflecting increased family-related healthcare costs. ($p < 0.001$)
- Region: Significant differences in charges across regions, with Southeast and Southwest regions having lower charges compared to the baseline. (regionnorthwest – not significant)

Model Fit:

- R-squared: 0.7486 (74.86% variance explained).
- F-statistic: 822.8 ($p < 2.2e-16$).
- Residual standard error: 6040 on 2211 degrees of freedom
- Residuals**:** Range from -11313 to 30291, indicating potential outliers or high variance.

ANOVA Table:

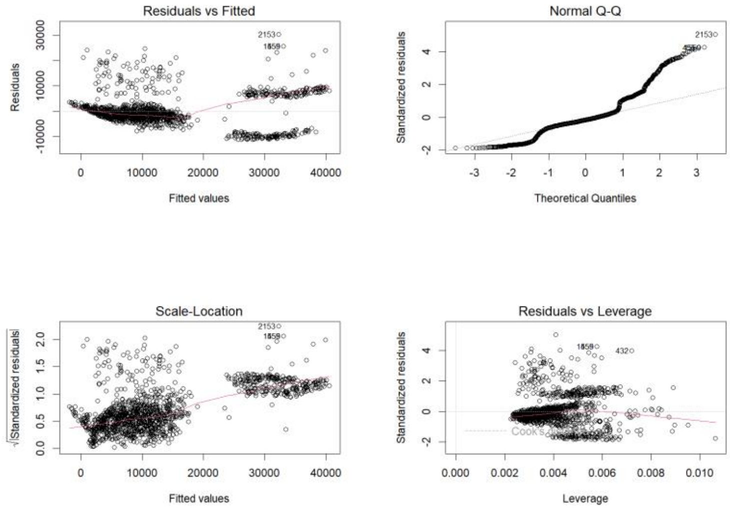
```
[1] "ANOVA Table"
Analysis of Variance Table

Response: charges
            Df     Sum Sq    Mean Sq   F value    Pr(>F)    
age          1 2.7753e+10 2.7753e+10  760.8290 < 2.2e-16 ***
sex          1 1.8044e+09 1.8044e+09   49.4676 2.678e-12 ***
bmi          1 6.5006e+09 6.5006e+09  178.2091 < 2.2e-16 ***
smoker       1 2.0248e+11 2.0248e+11 5550.8104 < 2.2e-16 ***
children     1 9.0591e+08 9.0591e+08   24.8349 6.730e-07 ***
region       3 6.5116e+08 2.1705e+08    5.9504 0.0004875 ***
Residuals 2211 8.0651e+10 3.6477e+07
```

All coefficients has significant contribution to the model with a very low Mean Square of Residuals.

Diagnostic Tests: **Model 1 Diagnostics**

- Variance Homogeneity: Significant heteroscedasticity ($p < 2.22e-16$).
- Residual Normality: Residuals not normally distributed (Shapiro-Wilk $p < 2.2e-16$).
- Autocorrelated Errors: Some autocorrelation detected (Durbin-Watson $p = 0.002$).
- Multicollinearity: No issues (GVIF close to 1).
- Outliers and Influential Points: Significant outliers and high leverage points detected (e.g., observations 2153, 455, 1559, and 432).

Residual Plots:



- Residuals vs Fitted: Shows non-random patterns with groups of clusters, indicating heteroscedasticity.
- Normal Q-Q: Points deviate from the line, especially in the tails, indicating non-normal residuals.
- Scale-Location: Spread of residuals increases with fitted values, confirming heteroscedasticity.
- Residuals vs Leverage: Identifies several influential points (e.g., 2153, 1859, 432).

The model explains a significant portion of the variance in charges, with smoking, BMI, and age being the most impactful predictors. However, issues with residual normality, heteroscedasticity, and outliers suggest the need for further refinement, such as variable transformation or robust regression techniques.

4.2.2 Model 2: Box-Cox Transformation

The linear regression model was re-built using Box-Cox transformed dependent variable (charges). This box-cox transformation for done using the maximum likelihood estimator of lamda. We transformed the dependent variable (charges) using the Box-Cox transformation to address the issues identified in the initial model. This transformation aims to stabilize the variance and make the data more normally distributed, thereby improving the model's fit and diagnostic measures. The model rebuilt showed the following key statistics:

Model Summary:

```
"Model Summary"

Call:
lm(formula = boxcox.charges ~ age + sex + bmi + smoker + children +
    region, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3672 -0.7690 -0.2396  0.2003  7.8718

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     11.386476   0.202549  56.216  < 2e-16 ***
age              0.116387   0.002437  47.756  < 2e-16 ***
```

```
sexmale          -0.228491   0.068497  -3.336 0.000865 ***
bmi               0.053529   0.005856   9.142  < 2e-16 ***
smokeryes         5.907635   0.085077  69.439  < 2e-16 ***
children          0.339629   0.028019  12.122  < 2e-16 ***
regionnorthwest -0.147158   0.098408  -1.495 0.134955
regionsoutheast -0.612408   0.099536  -6.153 9.02e-10 ***
regionsouthwest -0.485483   0.099239  -4.892 1.07e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.605 on 2211 degrees of freedom
Multiple R-squared:  0.7712,   Adjusted R-squared:  0.7704
F-statistic: 931.6 on 8 and 2211 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = -11.386 + 0.116*age – 0.228*sex(male) + 0.053*bmi + 5.907*smoker(yes) + 0.339*children – 0.147*region(northwest) – 0.612*region(southeast) – 0.485*region(southwest)

Model Fit:

- Residual Standard Error: 1.605
- Multiple R-squared: 0.7712, Adjusted R-squared: 0.7704
- F-statistic: 931.6 on 8 and 2211 degrees of freedom, p-value: < 2.2e-16

ANOVA Table:

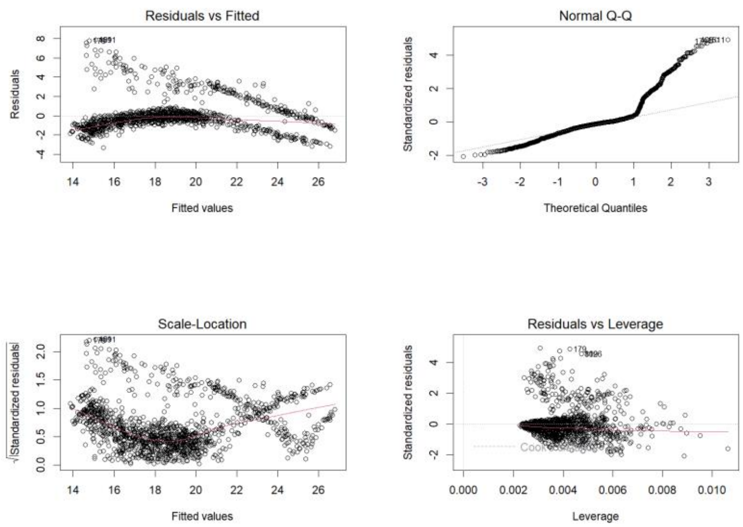
```
"ANOVA Table"
Analysis of Variance Table

Response: boxcox.charges
            Df   Sum Sq Mean Sq   F value    Pr(>F)
age          1   6054.0  6054.0 2350.1902 < 2.2e-16 ***
sex          1     21.3    21.3    8.2614  0.004088 **
bmi          1    121.4   121.4   47.1291 8.603e-12 ***
smoker       1 12492.4 12492.4 4849.5938 < 2.2e-16 ***
children     1    382.8   382.8  148.6227 < 2.2e-16 ***
region       3    126.4    42.1   16.3612 1.620e-10 ***
Residuals 2211   5695.4     2.6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All Coefficients has significant contribution to the model with Mean Square of Residuals 2.6.

Diagnostic Tests: Model 2 Diagnostics

- Variance Homogeneity Test: Indicates heteroscedasticity but improved from the initial model.
- Residual Normality Test: Indicates non-normality of residuals, but improvement in residual distribution.
- Auto Correlated Errors Test: No significant autocorrelation.
- Multi-Collinearity Test: VIF values are all below 2, indicating no severe multicollinearity.
- Outliers and Influential Points: Several influential points detected (e.g., observations 409, 1511, 179).

Residual Plots:



- Residuals vs Fitted: Improved patterns with reduced non-randomness.
- Normal Q-Q: Improved alignment with the diagonal line, indicating better normality of residuals.
- Scale-Location: Improved constant spread of residuals, indicating reduced heteroscedasticity.
- Residuals vs Leverage: Highlighting influential points that might still need to be addressed but show improvement.

Model Improvement

Before Transformation:

- The initial model exhibited significant issues with non-normality and heteroscedasticity.
- The residual plots showed clear patterns suggesting non-linearity and variance inequality.
- Normal Q-Q plot indicated deviations from normality especially at the tails.
- High value of Residual Standard Error was found

After Transformation:

- Box-Cox transformation improved model fit as indicated by the higher R-squared (0.7712 vs. earlier).
- Residual plots still show some patterns but with reduced severity.
- Q-Q plot shows better alignment but still indicates issues at the tails.
- Scale-location plot shows heteroscedasticity persists, although slightly improved.
- Significant decrease in the Residual Standard Error.
- Sexmale is now a significant coefficient.

The Box-Cox transformation has enhanced the linear model's performance by improving the normality of residuals and overall fit, evidenced by a higher R-squared value. However, there remain issues with heteroscedasticity and some residual non-normality, suggesting further model refinements or alternative approaches may be beneficial.

4.2.3 Model 3: Scaling the independent variable:

Standardizing numerical predictors helps in making them comparable and improves the model's interpretability and performance. By addressing issues related to scale and distribution, the model can more accurately estimate the relationship between predictors and the response variable.

Model Summary:

```
"Model Summary"

Call:
lm(formula = boxcox.charges ~ scale.age + sex + scale.bmi + smoker +
    scale.children + region, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3672 -0.7690 -0.2396  0.2003  7.8718

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     17.95052    0.08084 222.050  < 2e-16 ***
scale.age        1.64216    0.03439  47.756  < 2e-16 ***
sexmale         -0.22849    0.06850  -3.336 0.000865 ***
scale.bmi        0.33004    0.03610   9.142  < 2e-16 ***
smokeryes        5.90763    0.08508  69.439  < 2e-16 ***
scale.children   0.41381    0.03414  12.122  < 2e-16 ***
regionnorthwest -0.14716    0.09841  -1.495 0.134955
regionsoutheast -0.61241    0.09954  -6.153 9.02e-10 ***
regionsouthwest -0.48548    0.09924  -4.892 1.07e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.605 on 2211 degrees of freedom
Multiple R-squared:  0.7712,	Adjusted R-squared:  0.7704
F-statistic: 931.6 on 8 and 2211 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = 17.95 + 1.642*scale.age - 0.228*sex(male) + 0.330*scale.bmi + 5.907*smoker(yes) + 0.413*scale.children – 0.147*region(northwest) – 0.612*region(southeast) – 0.485*region(southwest)

Model Fit:

- Multiple R-squared: 0.7712 – Indicates that approximately 77.12% of the variability in transformed charges is explained by the model.
- Adjusted R-squared: 0.7704 – Adjusted for the number of predictors in the model.
- Residual Standard Error: 1.605 on 2211 degrees of freedom.
- F-statistic: 931.6 ($p < 2.2e-16$) – Indicates the model is statistically significant.

Improvements: Model 3 Diagnostics

- Interpretability: Scaling makes the coefficients more interpretable in terms of standard deviations.
- Convergence: The model fits better with improved diagnostic measures, though heteroscedasticity and non-normality are still present.
- Consistency: Significant predictors remained consistent, reinforcing the relationships between charges and factors like age, BMI, smoking status, and region.

Scaling has improved the accuracy of the relationship with the dependent variable. This can be seen by the increase in the estimate of the coefficients of age and bmi. This makes the model to be interpreted easily. But the scaling has not made any improvements in the model's residual issues.

4.2.4 Model 4: Model with Interaction Terms

The model was built with the interaction terms between the scale.age, scale.children and smoker status. Including interaction terms allows the model to account for the combined effects of multiple variables. This is particularly useful when the effect of one predictor on the dependent variable depends on the level of another predictor.

Model Summary:

```
"Model Summary"

Call:
lm(formula = boxcox.charges ~ scale.age * smoker * scale.children +
    scale.bmi + region + sex, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6018 -0.7252 -0.3377  0.0752  8.3366

Coefficients:
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                      17.91041    0.07454 240.285  < 2e-16 ***
scale.age                         1.84538    0.03701  49.860  < 2e-16 ***
smokeryes                         5.85074    0.07872  74.326  < 2e-16 ***
scale.children                    0.52081    0.03480  14.967  < 2e-16 ***
scale.bmi                         0.30693    0.03334   9.207  < 2e-16 ***
regionnorthwest                  -0.09764    0.09077  -1.076 0.282203
regionsoutheast                  -0.50109    0.09207  -5.442 5.84e-08 ***
regionsouthwest                  -0.50186    0.09148  -5.486 4.59e-08 ***
sexmale                          -0.21025    0.06320  -3.327 0.000892 ***
scale.age:smokeryes              -1.25219    0.08083 -15.491  < 2e-16 ***
scale.age:scale.children         -0.27498    0.03797  -7.242 6.10e-13 ***
smokeryes:scale.children         -0.49905    0.08148  -6.125 1.07e-09 ***
scale.age:smokeryes:scale.children  0.31779    0.08711   3.648 0.000270 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.479 on 2207 degrees of freedom
Multiple R-squared:  0.806,	Adjusted R-squared:  0.805
F-statistic: 764.2 on 12 and 2207 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = 17.91 + 1.845*scale.age - 0.210*sex(male) + 0.306*scale.bmi + 5.850*smoker(yes) + 0.520*scale.children – 0.097*region(northwest) – 0.501*region(southeast) – 0.485*region(southwest) – 1.25*scale.age*smoker(yes) – 0.274*scale.age*scale.children – 0.499*smoker(yes)*scale.children + 0.317*scale.age*smoker(yes)*scale.children

Model Fit:

- Multiple R-squared: 0.806 – Indicates that approximately 80.6% of the variability in transformed charges is explained by the model.
- Adjusted R-squared: 0.805 – Adjusted for the number of predictors in the model.
- Residual Standard Error: 1.479 on 2207 degrees of freedom.
- F-statistic: 764.2 (p < 2.2e-16) – Indicates the model is statistically significant.

ANOVA Table:

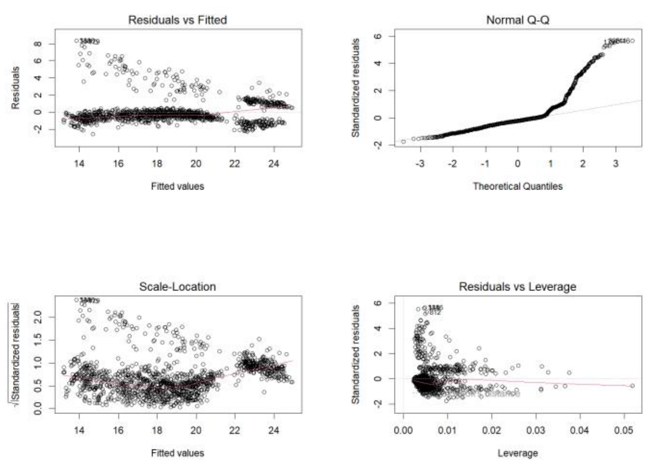
```
"ANOVA Table"
Analysis of Variance Table

Response: boxcox.charges
                               Df   Sum Sq Mean Sq  F value    Pr(>F)    
scale.age                       1   6054.0  6054.0 2766.774 < 2.2e-16 ***
smoker                          1  12469.2 12469.2 5698.609 < 2.2e-16 ***
scale.children                  1    376.3   376.3  171.994 < 2.2e-16 ***
scale.bmi                       1    143.5   143.5   65.566 9.191e-16 ***
region                          3    126.7    42.2   19.300 2.366e-12 ***
sex                             1     28.7    28.7   13.100 0.0003020 ***
scale.age:smoker                1    668.5   668.5  305.515 < 2.2e-16 ***
scale.age:scale.children        1     81.8    81.8   37.368 1.153e-09 ***
smoker:scale.children           1     86.9    86.9   39.719 3.530e-10 ***
scale.age:smoker:scale.children 1     29.1    29.1   13.309 0.0002702 ***
Residuals                    2207   4829.2     2.2
```

All Coefficients has significant contribution to the model including the interaction terms. The Mean Square of Residuals has also been reduced to 2.2.

Diagnostic Tests: [Model 4 Diagnostics]

- Variance Homogeneity Test: Chi-square = 180.911 (p < 2.22e-16) – Indicates heteroscedasticity.
- Residual Normality Test: Shapiro-Wilk test p-value < 2.2e-16 – Indicates non-normality of residuals.
- Auto Correlated Errors Test: D-W Statistic = 2.032565 (p = 0.456) – No significant autocorrelation.
- Multi-Collinearity Test: VIF values are all below 2, indicating no severe multicollinearity.
- Outliers and Influential Points: Several influential points detected (e.g., observations 338, 1446, 179).

Residual Plots:



- Residuals vs Fitted: Improved patterns with reduced non-randomness.
- Normal Q-Q: Improved alignment with the diagonal line, indicating better normality of residuals.
- Scale-Location: Improved constant spread of residuals, indicating reduced heteroscedasticity.
- Residuals vs Leverage: Highlighting influential points that might still need to be addressed but show improvement.

Improvements of including Interaction Terms

- Interpretability: Interaction terms provide a more nuanced understanding of how predictors jointly affect the outcome.
- Model Fit: Higher R-squared values indicate a better fit compared to the previous models.
- Significance: Interaction terms are statistically significant, indicating their importance in explaining variability in charges.
- The Non-Linear pattern in the residual plot has been removed.

Including interaction terms significantly improved the model's interpretability and fit. The interaction terms capture the complex relationships between age, smoking status, and the number of children on medical charges. However, issues of heteroscedasticity and non-normality of residuals persist, though they are reduced. Further refinement or alternative modeling techniques could be considered to address these remaining issues.

4.2.5 Model 4 Variable Selection

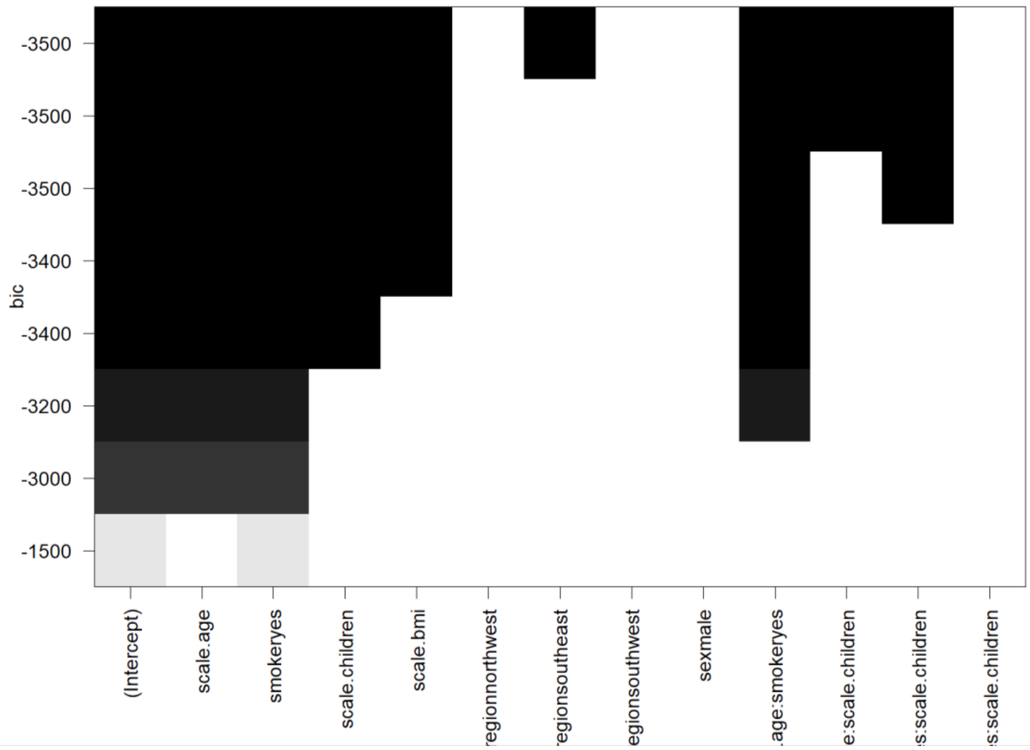
In our analysis, we aimed to develop a robust and simplified model that maintained strong predictive power while being easier to interpret and apply practically. Initially, the full model provided a comprehensive view but was complex. To address this, we utilized the regsubsets function, which performs an exhaustive search for the best subsets of predictors. This approach enabled us to systematically evaluate and compare different models.

Two key metrics, the Bayesian Information Criterion (BIC) and Adjusted R^2, were employed to assess model performance. BIC measures model fit with a penalty for the number of parameters, helping to avoid overfitting. Lower BIC values indicate better models. Adjusted R^2 adjusts the R^2 value for the number of predictors, providing a more accurate measure of model fit when comparing models with different numbers of predictors. Higher adjusted R^2 values indicate better models.

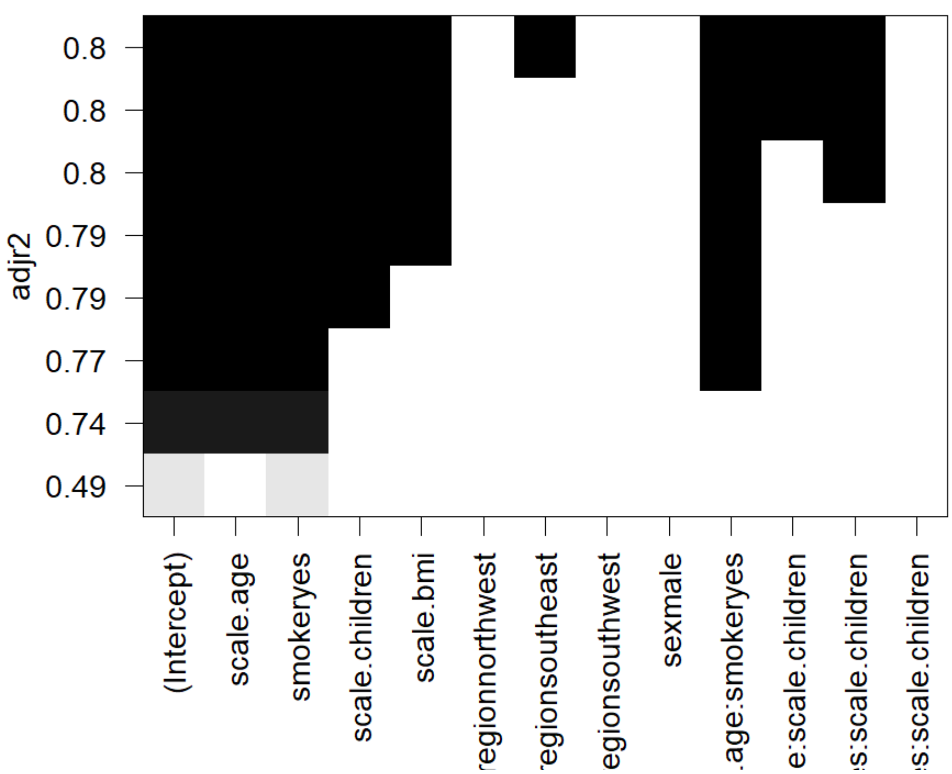
By selecting a subset of predictors, we simplified the model, making it easier to interpret and more practical for application. Using model selection criteria like BIC helped us identify a model that balanced fit and complexity, reducing the risk of overfitting. The goal was to improve residual assumptions, leading to more reliable and robust models. Simpler models are easier to communicate and justify, especially in practical applications. The selected models retained the most significant predictors and interactions, ensuring they remained interpretable while still providing strong predictive power.

In addition to this approach, we considered using forward, backward, and stepwise selection methods. However, we chose not to employ these methods because our dataset did not contain a large number of variables, making the use of all possible regression methods sufficient. By performing an exhaustive search with all possible regression methods, we ensured a thorough evaluation without the need for the aforementioned selection techniques. This approach allowed us to confidently identify the best subset of predictors for our model.

BIC Plot



Adj R^2 Plot



4.2.5.1 Sub Model 1

```
"Model Summary"

Call:
lm(formula = boxcox.charges ~ scale.age * smoker * scale.children + 
    scale.bmi + region, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.6857 -0.7284 -0.3237  0.1209  8.3222 

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)                       17.80622    0.06779 262.652  < 2e-16 ***
scale.age                          1.85080    0.03706  49.942  < 2e-16 ***
smokeryes                          5.82967    0.07864  74.131  < 2e-16 ***
scale.children                     0.52123    0.03488  14.945  < 2e-16 ***
scale.bmi                          0.30173    0.03338   9.040  < 2e-16 ***
regionnorthwest                   -0.09282    0.09096  -1.020  0.30765    
regionsoutheast                   -0.50144    0.09228  -5.434 6.12e-08 ***
regionsouthwest                   -0.49865    0.09169  -5.439 5.96e-08 ***
scale.age:smokeryes               -1.25945    0.08099 -15.551  < 2e-16 ***
scale.age:scale.children          -0.27087    0.03804  -7.121 1.45e-12 ***
smokeryes:scale.children          -0.50458    0.08165  -6.180 7.62e-10 ***
scale.age:smokeryes:scale.children  0.31127    0.08728   3.566  0.00037 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.483 on 2208 degrees of freedom
Multiple R-squared:  0.805,	Adjusted R-squared:  0.8041 
F-statistic: 828.8 on 11 and 2208 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = 17.80 + 1.850*scale.age + 0.301*scale.bmi + 5.829*smoker(yes) + 0.521*scale.children – 0.092*region(northwest) – 0.501*region(southeast) – 0.498*region(southwest) – 1.25*scale.age*smoker(yes) – 0.270*scale.age*scale.children – 0.504*smoker(yes)*scale.children + 0.311*scale.age*smoker(yes)*scale.children

Coefficients:

- All predictors are statistically significant, except for regionnorthwest, which has a p-value of 0.307, indicating it is not significantly different from zero.
- Significant predictors include scale.age, smokeryes, scale.children, scale.bmi, regionsoutheast, regionsouthwest, scale.age:smokeryes, scale.age:children, and smokeryes:scale.children.

Model Fit:

- R-squared and Adjusted R-squared: The model explains about 80.5% of the variability in the response variable (R-squared: 0.805).
- Adjusted R-squared is slightly lower (Adjusted R-squared: 0.8041), accounting for the number of predictors.
- F-statistic: The overall F-test is significant (F-statistic: 828.8, p-value: < 2.2e-16), indicating that the model is a good fit for the data.

ANOVA Table:

```
"ANOVA Table"
Analysis of Variance Table
```

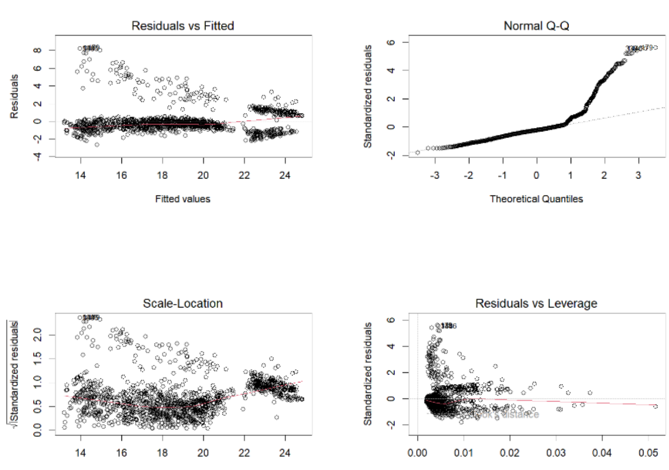
```
Response: boxcox.charges
                                  Df   Sum Sq Mean Sq  F value    Pr(>F)    
scale.age                          1   6054.0  6054.0 2754.215 < 2.2e-16 ***
smoker                             1  12469.2 12469.2 5672.740 < 2.2e-16 ***
scale.children                     1    376.3   376.3  171.214 < 2.2e-16 ***
scale.bmi                          1    143.5   143.5   65.269 1.064e-15 ***
region                             3    126.7    42.2   19.213 2.684e-12 ***
scale.age:smoker                   1    674.6   674.6  306.899 < 2.2e-16 ***
scale.age:scale.children           1     79.5    79.5   36.172 2.109e-09 ***
smoker:scale.children              1     88.7    88.7   40.347 2.574e-10 ***
scale.age:smoker:scale.children    1     28.0    28.0   12.718 0.0003699 ***
Residuals                       2208   4853.4     2.2                       
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

All predictors are significant, with smoker having the largest F-value, indicating it explains the most variance in boxcox.charges.

Diagnostic Tests: **Submodel 1 diagnostics**

- Variance Homogeneity Test: The test indicates heteroscedasticity (Chisquare = 184.6852, p = < 2.22e-16), suggesting the variance of residuals is not constant.
- Shapiro-Wilk Normality Test: The test indicates that residuals are not normally distributed (W = 0.74409, p-value < 2.2e-16).
- Durbin-Watson Test for Autocorrelated Errors: The test indicates no significant autocorrelation (D-W Statistic = 2.025914, p = 0.56).
- Multicollinearity Test: The variance inflation factors (VIF) are all below 2, indicating no significant multicollinearity.
- Outlier Test: Several outliers were identified, with Bonferroni-adjusted p-values indicating they are significant.
- Influential Points: Several points were identified as potentially influential, with Cook's distance and leverage values indicating their potential impact on the model.
- The PRESS statistic is 4903.566, which is used to assess the predictive power of the model.

Residual Plots:



- Residuals vs Fitted Plot: The plot indicates that residuals are randomly scattered around the horizontal axis, suggesting that the linearity assumption is reasonably met. However, there are some points with larger residuals, indicating potential heteroscedasticity.
- Normal Q-Q Plot: This plot shows the quantiles of the standardized residuals against the theoretical quantiles of a normal distribution. The points mostly follow a straight line, indicating that the residuals are approximately normally distributed. However, there are deviations at the tails, suggesting some outliers.
- Scale-Location Plot: The points are spread somewhat randomly around a horizontal line, but there is a slight funnel shape, suggesting mild heteroscedasticity.
- Residuals vs Leverage Plot: This plot helps identify influential data points. Points with high leverage and large residuals can have a significant impact on the regression model. There are a few points that might be influential, as indicated by their distance from the bulk of the data.

The chosen model 1 shows a good fit with a high adjusted R-squared value and significant predictors. The diagnostic plots and tests reveal some concerns with heteroscedasticity and non-normality of residuals, but this is usual when the dataset is large. However, the model explains a significant portion of the variance in the response variable. The chosen model simplifies the full model while retaining important interaction terms, making it easier to interpret and potentially more robust. Overall, the chosen model 1 balances complexity and predictive power, addressing the initial diagnostic concerns and improving overall model performance.

4.2.5.2 Sub Model 2

```
"Model Summary"

Call:
lm(formula = boxcox.charges ~ scale.age * smoker * scale.children +
    scale.bmi, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6793 -0.7253 -0.3032  0.1126  8.0565

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                        17.52550    0.03568 491.203  < 2e-16 ***
scale.age                           1.85695    0.03738  49.676  < 2e-16 ***
smokeryes                           5.81938    0.07905  73.617  < 2e-16 ***
scale.children                      0.52178    0.03522  14.816  < 2e-16 ***
scale.bmi                           0.24511    0.03204   7.651 2.96e-14 ***
scale.age:smokeryes                -1.25868    0.08175 -15.397  < 2e-16 ***
scale.age:scale.children           -0.27959    0.03835  -7.291 4.25e-13 ***
smokeryes:scale.children           -0.50099    0.08248  -6.074 1.46e-09 ***
scale.age:smokeryes:scale.children  0.34113    0.08794   3.879 0.000108 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.498 on 2211 degrees of freedom
Multiple R-squared:  0.8007,   Adjusted R-squared:    0.8
F-statistic:  1110 on 8 and 2211 DF,  p-value: < 2.2e-16
```

Model Equation:

Y_hat = 17.52 + 1.856*scale.age + 0.245*scale.bmi + 5.819*smoker(yes) + 0.521*scale.children – 1.25*scale.age*smoker(yes) – 0.279*scale.age*scale.children – 0.500*smoker(yes)*scale.children + 0.341*scale.age*smoker(yes)*scale.children

All predictors are statistically significant.

Model Fit:

- R-squared and Adjusted R-squared: The model explains about 79.93% of the variability in the response variable (R-squared: 0.7993).
- Adjusted R-squared: Slightly lower at 79.87%, accounting for the number of predictors.
- F-statistic: The overall F-test is significant (F-statistic: 1259, p-value: < 2.2e-16), indicating that the model is a good fit for the data.

ANOVA Table:

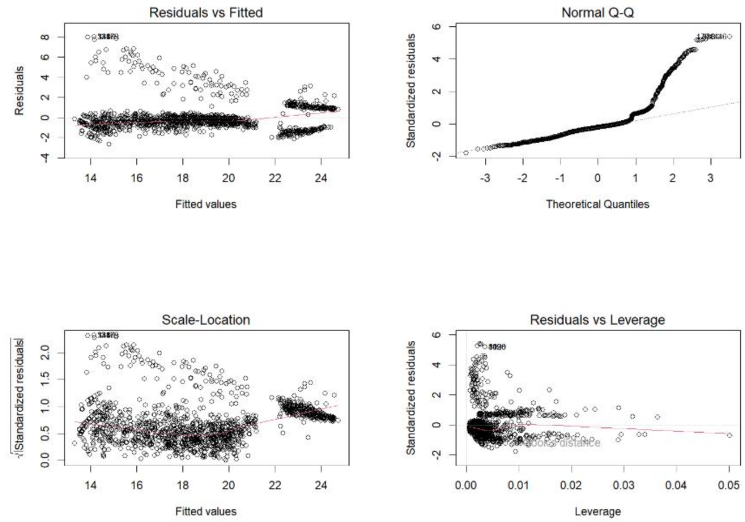
```
"ANOVA Table"
Analysis of Variance Table

Response: boxcox.charges
                                Df  Sum Sq Mean Sq  F value    Pr(>F)    
scale.age                        1  6054.0  6054.0 2697.689 < 2.2e-16 ***
smoker                           1 12469.2 12469.2 5556.316 < 2.2e-16 ***
scale.children                   1   376.3   376.3  167.700 < 2.2e-16 ***
scale.bmi                        1   143.5   143.5   63.929 2.059e-15 ***
scale.age:smoker                 1   685.0   685.0  305.233 < 2.2e-16 ***
scale.age:scale.children         1    82.4    82.4   36.710 1.607e-09 ***
smoker:scale.children            1    87.9    87.9   39.157 4.682e-10 ***
scale.age:smoker:scale.children  1    33.8    33.8   15.046  0.000108 ***
Residuals                     2211  4961.8     2.2                       
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

All predictors have significant contribution with smoker having the largest F-value, indicating it explains the most variance in boxcox.charges.

Diagnostic Tests: Submodel 2 Diagnostics

- Variance Homogeneity Test: The test indicates heteroscedasticity (Chisquare = 212.9258, p = < 2.22e-16), suggesting the variance of residuals is not constant.
- Shapiro-Wilk Normality Test: The test indicates that residuals are not normally distributed (W = 0.74582, p-value < 2.2e-16).
- Durbin-Watson Test for Autocorrelated Errors: The test indicates no significant autocorrelation (D-W Statistic = 2.021586, p = 0.652).
- Multicollinearity Test: The variance inflation factors (VIF) are all below 2, indicating no significant multicollinearity.
- Outlier Test: Several outliers were identified, with Bonferroni-adjusted p-values indicating they are significant.
- Influential Points: Several points were identified as potentially influential, with Cook's distance and leverage values indicating their potential impact on the model.
- The PRESS statistic is 4999.434, which is used to assess the predictive power of the model.

Residual Plots:



- Residuals vs Fitted Plot: The plot indicates that residuals are randomly scattered around the horizontal axis, suggesting that the linearity assumption is reasonably met. However, there are some points with larger residuals, indicating potential heteroscedasticity.
- Normal Q-Q Plot: This plot shows the quantiles of the standardized residuals against the theoretical quantiles of a normal distribution. The points mostly follow a straight line, indicating that the residuals are approximately normally distributed. However, there are deviations at the tails, suggesting some outliers.
- Scale-Location Plot: The points are spread somewhat randomly around a horizontal line, but there is a slight funnel shape, suggesting mild heteroscedasticity.
- Residuals vs Leverage Plot: This plot helps identify influential data points. Points with high leverage and large residuals can have a significant impact on the regression model. There are a few points that might be influential, as indicated by their distance from the bulk of the data.

The chosen Model 2 shows a good fit with a bit reduced adjusted R-squared value and all significant predictors. The diagnostic plots and tests reveal some concerns with heteroscedasticity and non-normality of residuals, but this is usual when the dataset is large. The model explains a significant portion of the variance in the response variable. The chosen model simplifies the full model while retaining important interaction terms, making it easier to interpret and potentially more robust. Overall, Model 2 balances complexity and predictive power, addressing the initial diagnostic concerns and improving overall model performance.

4.2.6 GLM Model

Model Summary

```
summary(glm_model)

Call:
glm(formula = boxcox.charges ~ scale.age * smoker * scale.children +
    scale.bmi, family = gaussian(link = "identity"), data = trainData)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6793  -0.7253  -0.3032   0.1126   8.0565

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                        17.52550    0.03568 491.203  < 2e-16 ***
scale.age                           1.85695    0.03738  49.676  < 2e-16 ***
smokeryes                           5.81938    0.07905  73.617  < 2e-16 ***
scale.children                      0.52178    0.03522  14.816  < 2e-16 ***
scale.bmi                           0.24511    0.03204   7.651 2.96e-14 ***
scale.age:smokeryes                -1.25868    0.08175 -15.397  < 2e-16 ***
scale.age:scale.children           -0.27959    0.03835  -7.291 4.25e-13 ***
smokeryes:scale.children           -0.50099    0.08248  -6.074 1.46e-09 ***
scale.age:smokeryes:scale.children  0.34113    0.08794   3.879 0.000108 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 2.244142)

    Null deviance: 24893.8  on 2219  degrees of freedom
Residual deviance:  4961.8  on 2211  degrees of freedom
AIC: 8105.5

Number of Fisher Scoring iterations: 2
```

Model Equation:

Y_hat = 17.52 + 1.856*scale.age + 0.245*scale.bmi + 5.819*smoker(yes) + 0.521*scale.children – 1.258*scale.age*smoker(yes) – 0.279*scale.age*scale.children – 0.500*smoker(yes)*scale.children + 0.341*scale.age*smoker(yes)*scale.children

All coefficients are statistically significant, with p-values less than 0.05, indicating that these predictors and interactions have a significant effect on the response variable boxcox.charges.

Model Fit

- The deviance residuals provide an indication of how well the model fits the data. The residuals range from -2.6793 to 8.0565, with the majority of residuals being small in magnitude, indicating a reasonably good fit.
- Null Deviance: 24893.8 on 2219 degrees of freedom; Residual Deviance: 4961.8 on 2211 degrees of freedom. The reduction in deviance from the null model to the fitted model suggests that the model explains a significant portion of the variability in the response variable.
- AIC: 8105.5

Model Diagnostics GLM Model Diagnostics

- Deviance Test: The residual deviance of the model is 4961.797, indicating the level of unexplained variability by the model. The p-value for the deviance test is extremely low (5.51702e-212), indicating that the model fit is significantly better than the null model.

- Pearson’s Chi-Square Test: This test also yields a very low p-value (5.51702e-212), further confirming the model's good fit.
- Type III ANOVA test shows that all predictors and interactions are highly significant.
- The sequential ANOVA confirms that each predictor and interaction significantly improves the model fit.
- The Shapiro-Wilk test indicates that the residuals are not normally distributed ($p < 2.2e-16$). This suggests that there may be issues with the assumptions of the GLM, particularly the assumption of normality of residuals.

4.3 Model Validation: Outputs

Initially, the model was trained using a subset of the available data (trainData). This training phase involved fitting the model to the data to capture the underlying patterns and relationships between the predictor variables and the response variable (boxcox.charges). Multiple models were considered, including a full model (model4) and two subset models (sub_model1 and sub_model2). Each model included different combinations of predictor variables and interaction terms. The subset models aimed to simplify the full model while retaining significant predictors and interaction terms.

To evaluate and compare the models, performance metrics were calculated using the remaining subset of data (testData). The key performance metrics used are:

- Mean Absolute Error (MAE): Measures the average magnitude of errors in predictions, providing a straightforward interpretation of prediction accuracy.
- Mean Squared Error (MSE): Provides a measure of the average squared differences between predicted and actual values, giving more weight to larger errors.
- Root Mean Squared Error (RMSE): The square root of MSE, which is in the same units as the response variable, offering an intuitive understanding of the error magnitude.
- R-squared (R2): Indicates the proportion of variance in the response variable explained by the model, with higher values indicating better explanatory power.
- Mean Absolute Percentage Error (MAPE): Represents the average percentage error, which is useful for understanding the relative accuracy of predictions.

	MAE	MSE	RMSE	R2	MAPE
Model 4	0.8910	2.1222	1.4567	0.8160	0.0450
Sub Model 1	0.9035	2.1410	1.4632	0.8143	0.0458
Sub Model 2	0.9030	2.1505	1.4664	0.8134	0.0460
GLM Model	0.9030	2.1505	1.4664	0.8134	0.0460

Metrics Analysis

- Model 4 has the lowest MSE and RMSE, and the highest R^2, showing strong predictive power. However, this comes with increased complexity, which may reduce interpretability.

- Submodel 1's metrics are very close to Model 4, with slightly higher MSE and RMSE, and a marginally lower R^2. This model is simpler but still retains strong predictive capabilities.

- Submodel 2 has metrics very similar to Submodel 1 and the GLM Model. Despite slightly higher MSE and RMSE compared to Model 4, it provides a good balance of predictive power and simplicity.

- The GLM Model shares identical metrics with Submodel 2, suggesting similar performance levels. However, it may not offer the same simplicity and interpretability benefits.

Plot Analysis

- Model 4: The points are close to the red dashed line, indicating good predictive accuracy. However, there are some deviations, particularly at the higher observed values, which suggest the model may overfit slightly.

- Submodel 1: The scatter plot shows points relatively close to the red dashed line, with slightly more spread compared to Model 4. This indicates good predictive performance but with a few more deviations.

- Submodel 2: The points in this plot are closely aligned with the red dashed line, demonstrating a strong fit with minimal deviations. This indicates that Submodel 2 maintains high predictive accuracy while being simpler.
- GLM Model: This plot also shows points close to the red dashed line but with more noticeable deviations, particularly at the extremes. This indicates slightly less accurate predictions compared to Submodel 2.

5 RESULTS

Despite the negligible differences in performance metrics between the models, Submodel 2 was chosen for its effective balance between predictive power and simplicity. It is easier to interpret and apply practically, making it ideal for real-world applications.

- **Visual Fit**: The scatter plot for Submodel 2 shows a strong alignment of predicted values with observed values, indicating good predictive performance.
- **Simplicity and Interpretability**: Submodel 2 provides a simpler model that is easier to understand and implement, without sacrificing significant predictive power. This makes it a preferred choice for practical applications where model interpretability and ease of communication are crucial.
- **Metrics**: Although Model 4 has slightly better metrics, the differences are minimal. Submodel 2's slightly higher MSE and RMSE are offset by its simplicity and effective balance of predictive power.

6 APPENDIX

6.1 EXPLANATORY DATA ANALYSIS

6.1.1 Dimensions of the Dataset

```
#Dimensions of the dataset
> dim(insurance)
[1] 2772    7
```

6.1.2 Head of the Dataset

```
#Head of the dataset
> head(insurance)
# A tibble: 6 × 7
    age sex      bmi children smoker region    charges
  <dbl> <fct>  <dbl>    <dbl> <fct>  <fct>       <dbl>
1    19 female  27.9        0 yes    southwest  16885.
2    18 male    33.8        1 no     southeast   1726.
3    28 male    33          3 no     southeast   4449.
4    33 male    22.7        0 no     northwest  21984.
5    32 male    28.9        0 no     northwest   3867.
6    31 female  25.7        0 no     southeast   3757.
```

6.1.3 Datatypes of the variables

```
#Datatypes of the features
> lapply(insurance, class)
$age
[1] "numeric"

$sex
[1] "factor"

$bmi
[1] "numeric"

$children
[1] "numeric"

$smoker
[1] "factor"

$region
[1] "factor"

$charges
[1] "numeric"
```

6.1.4 Check for missing values

```
#Check for missing values in the dataset
> any(is.na(insurance))
[1] FALSE
```

6.1.5 Descriptive Statistics of Numerical Variables

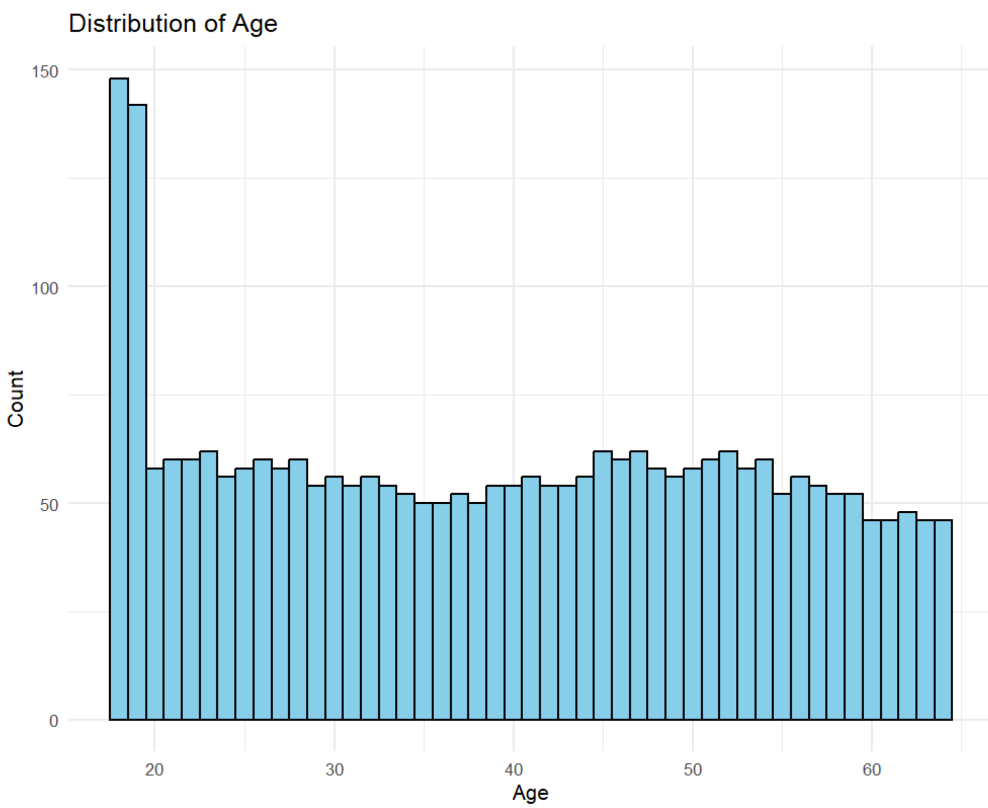
```
#Numerical Variables
> insurance %>% dplyr::select(age, bmi, charges) %>% summary()
      age             bmi           charges
 Min.   :18.00   Min.   :15.96   Min.   : 1122
 1st Qu.:26.00   1st Qu.:26.22   1st Qu.: 4688
 Median :39.00   Median :30.45   Median : 9333
 Mean   :39.11   Mean   :30.70   Mean   :13261
 3rd Qu.:51.00   3rd Qu.:34.77   3rd Qu.:16578
 Max.   :64.00   Max.   :53.13   Max.   :63770
```

6.1.6 Descriptive Statistics of Categorical Variables

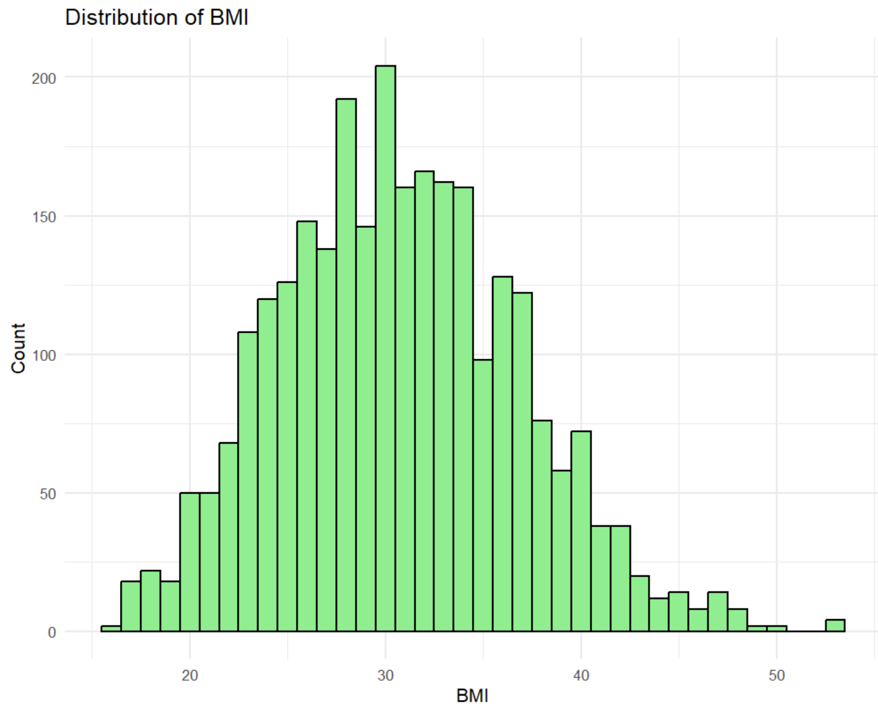
```
#Categorical Variables
> ftable(table(insurance$sex, insurance$smoker, insurance$region))
```

```
            northeast northwest southeast southwest

female no         266       280       290       298
       yes         58        58        74        42
male   no         256       266       284       268
       yes         78        60       118        76
table(insurance$sex)

female   male
  1366   1406
> table(insurance$smoker)

  no  yes
2208  564
> table(insurance$region)

northeast northwest southeast southwest
      658       664       766       684
```

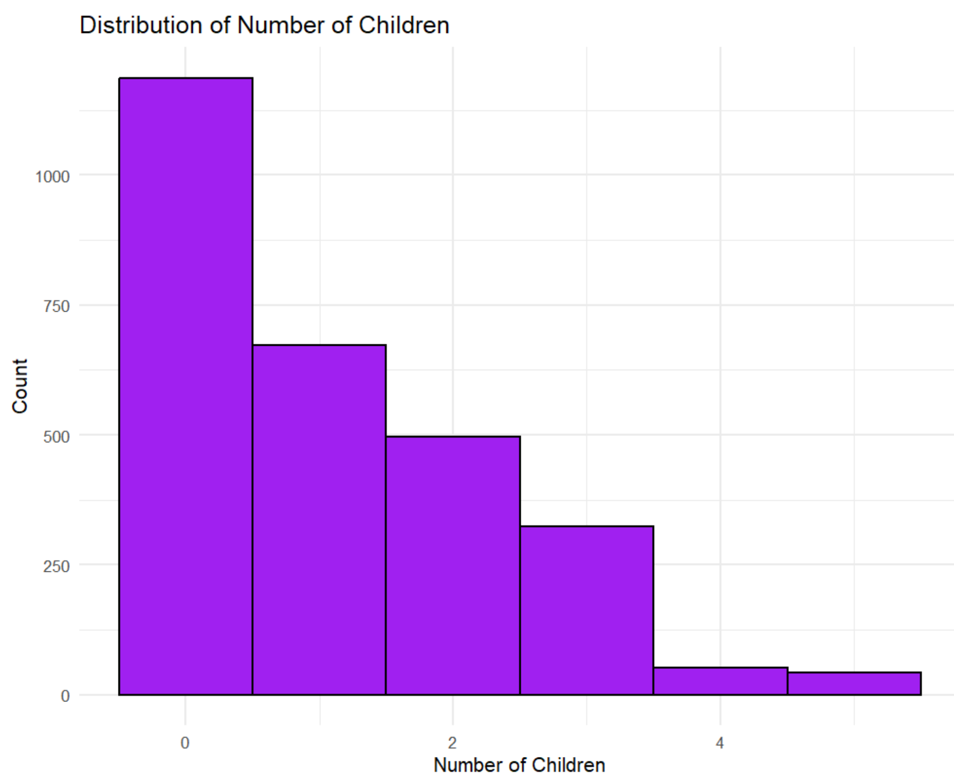
6.1.7 Histogram of Age



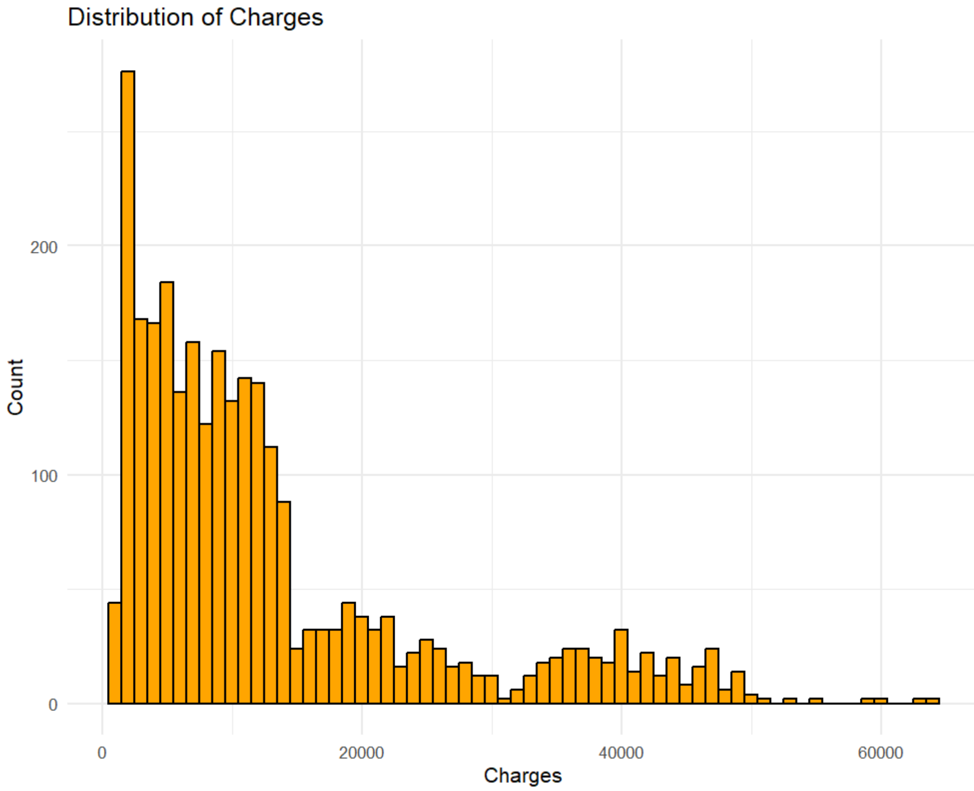
6.1.8 Histogram of BMI



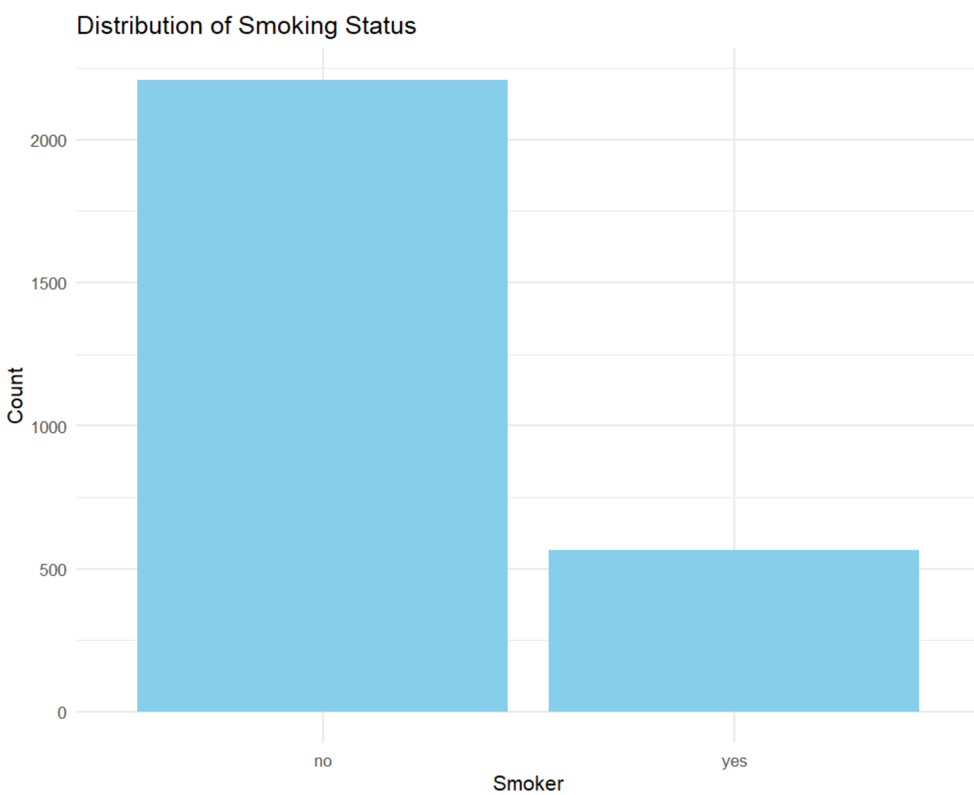
6.1.9 Histogram of Children



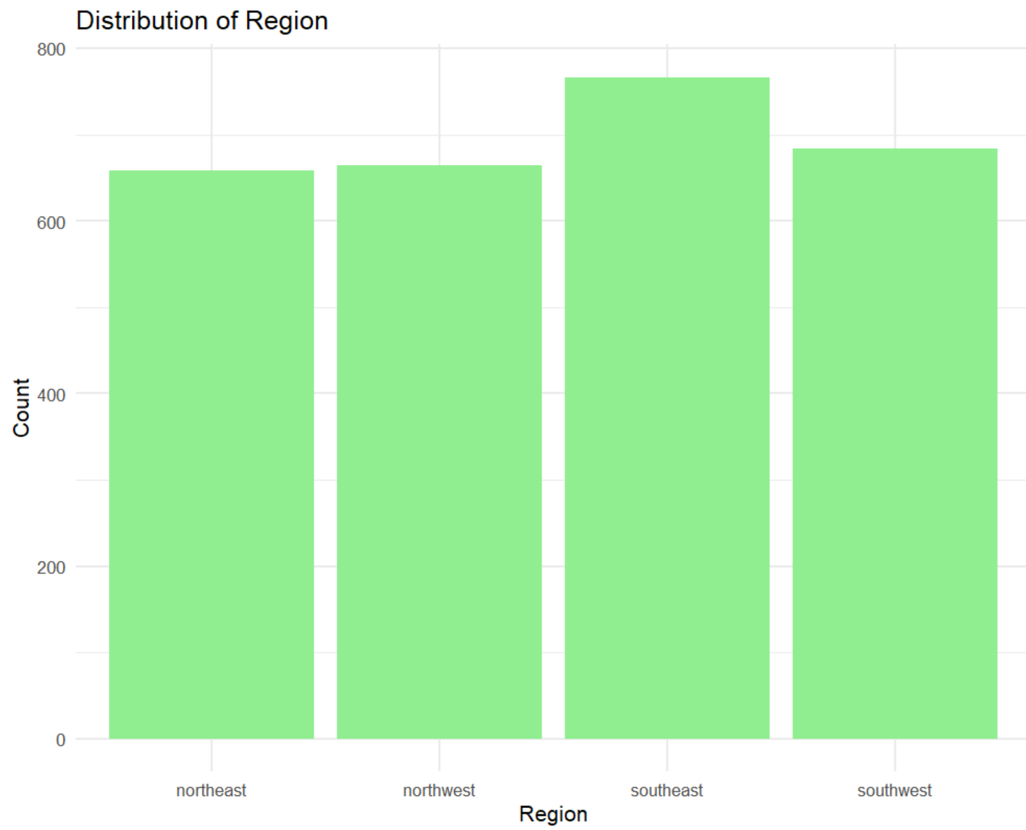
6.1.10 Histogram of Charges



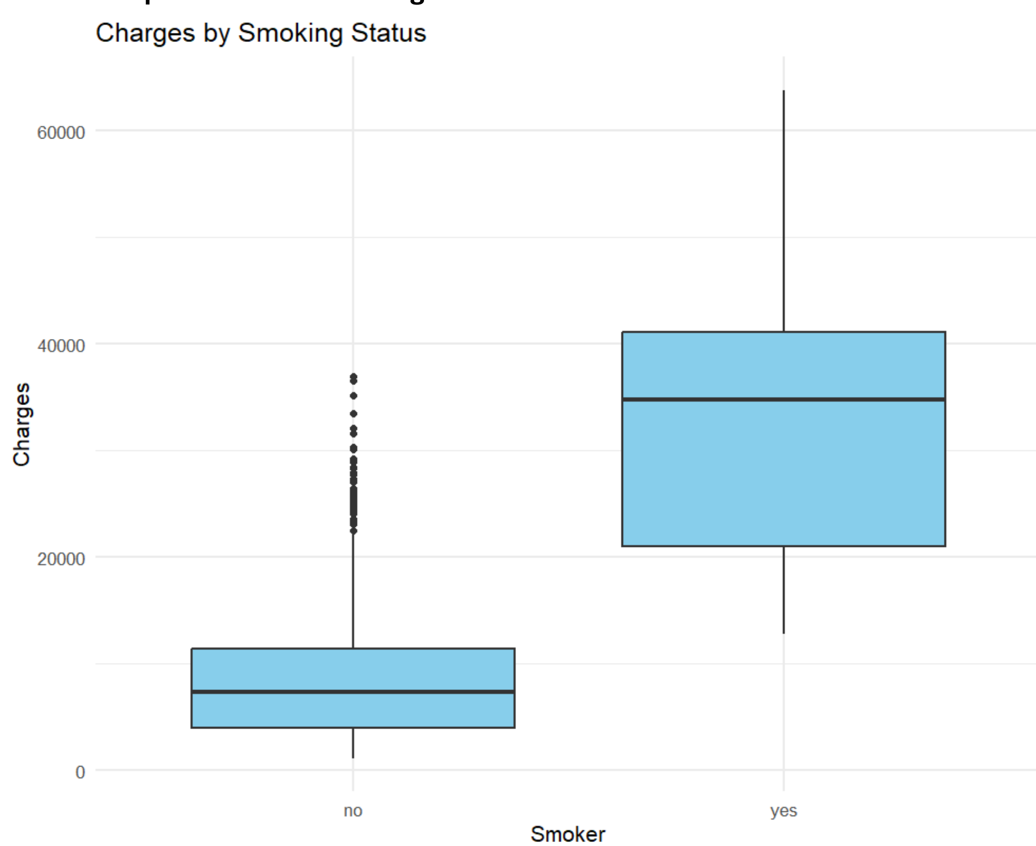
6.1.11 Distribution of Smokers



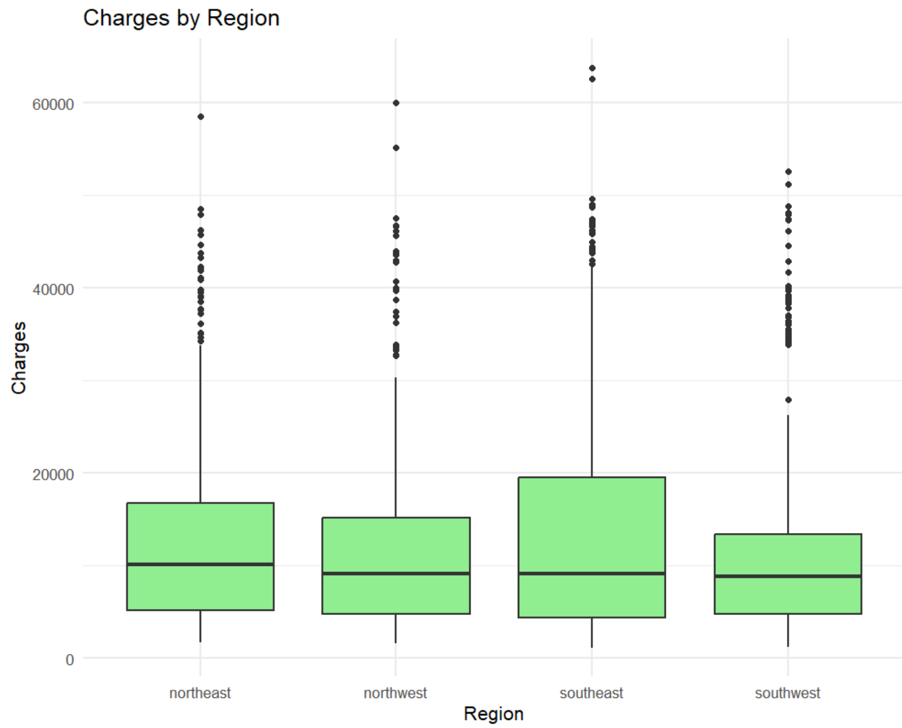
6.1.12 Distribution of Region



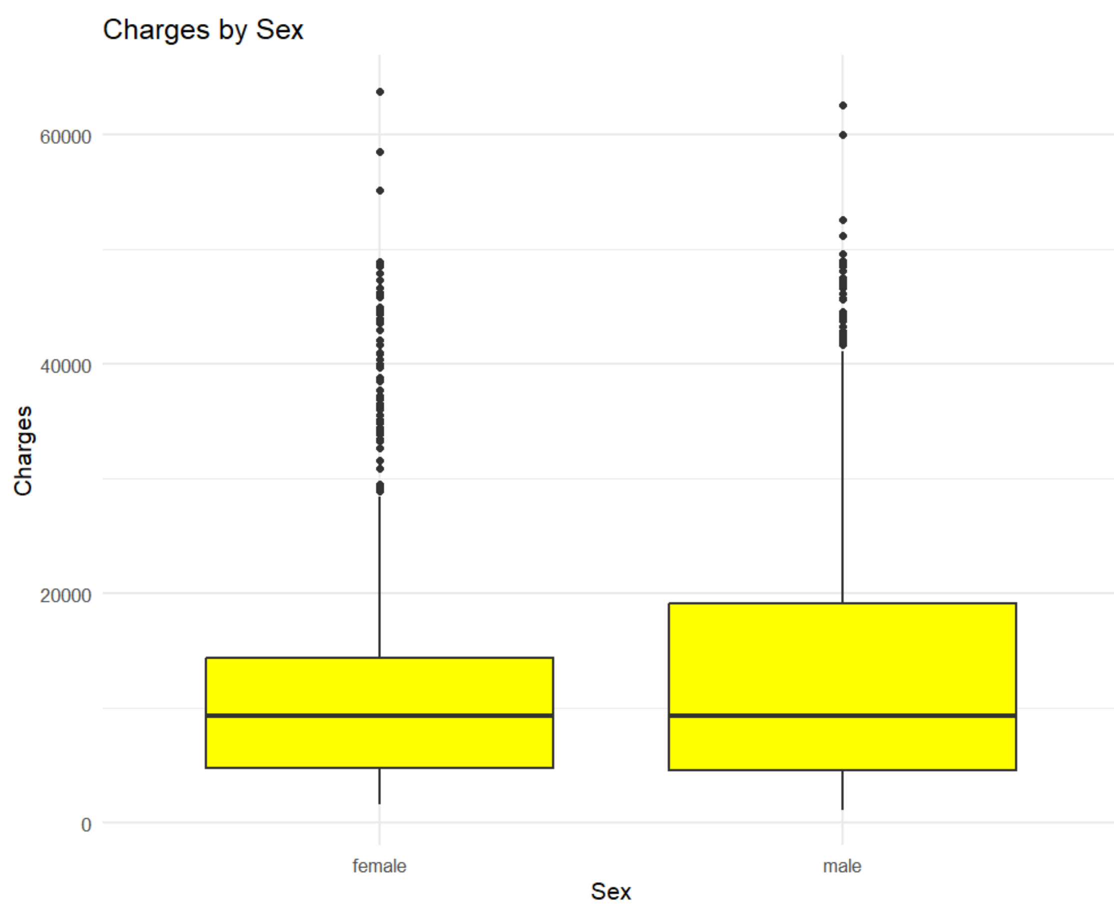
6.1.13 Boxplot Smokers vs Charges



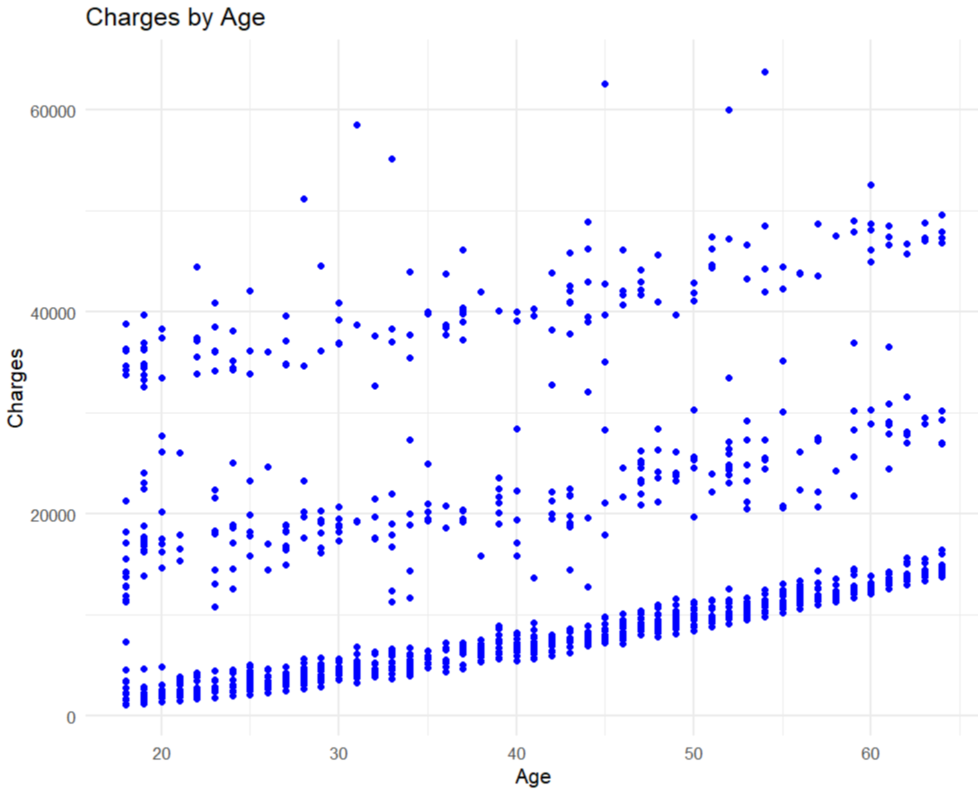
6.1.14 Boxplot Region vs Charges



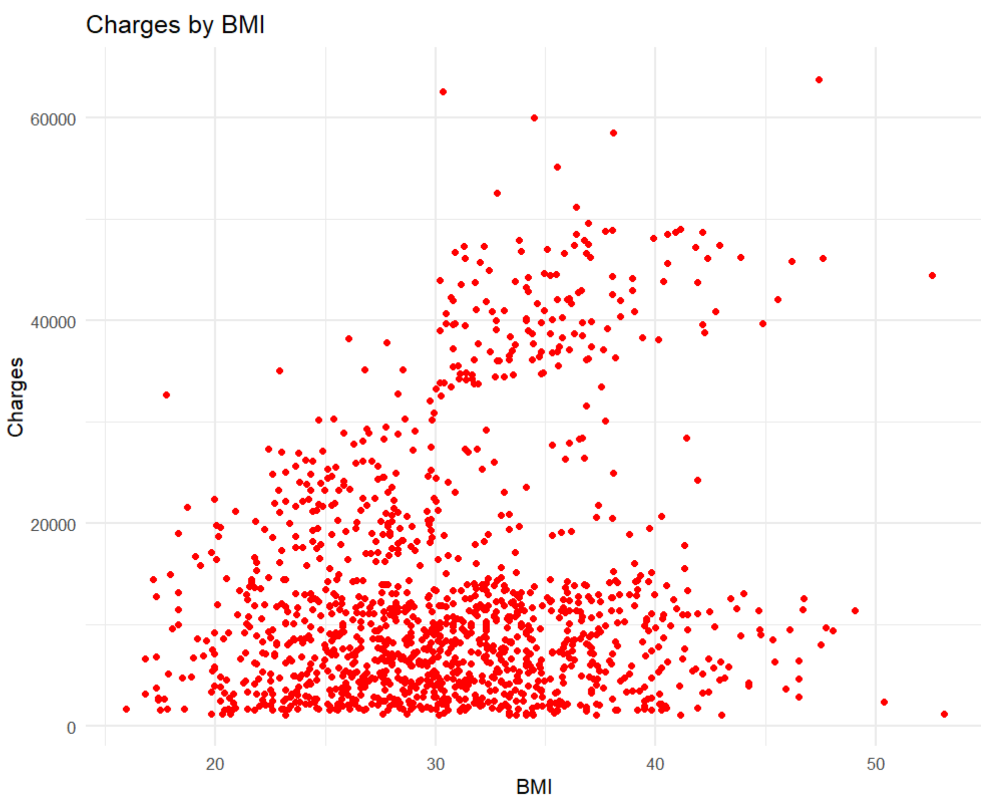
6.1.15 Boxplot Sex vs Charges



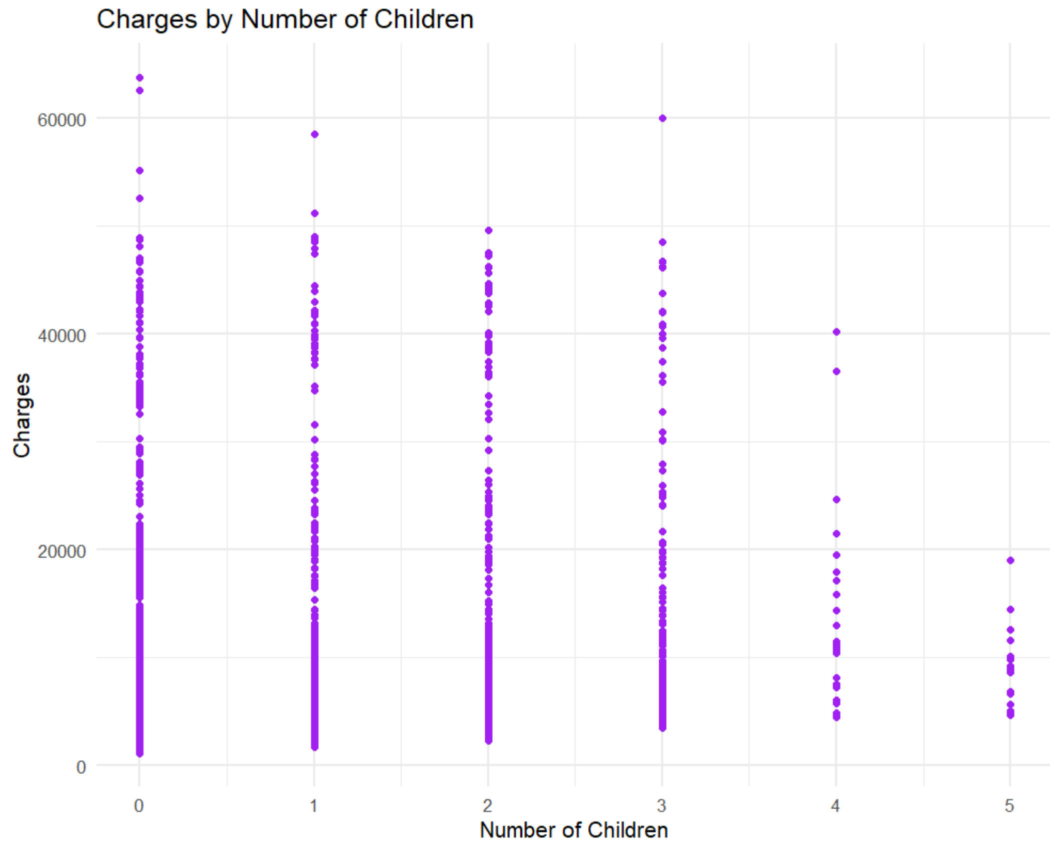
6.1.16 Scatterplot Age vs Charges



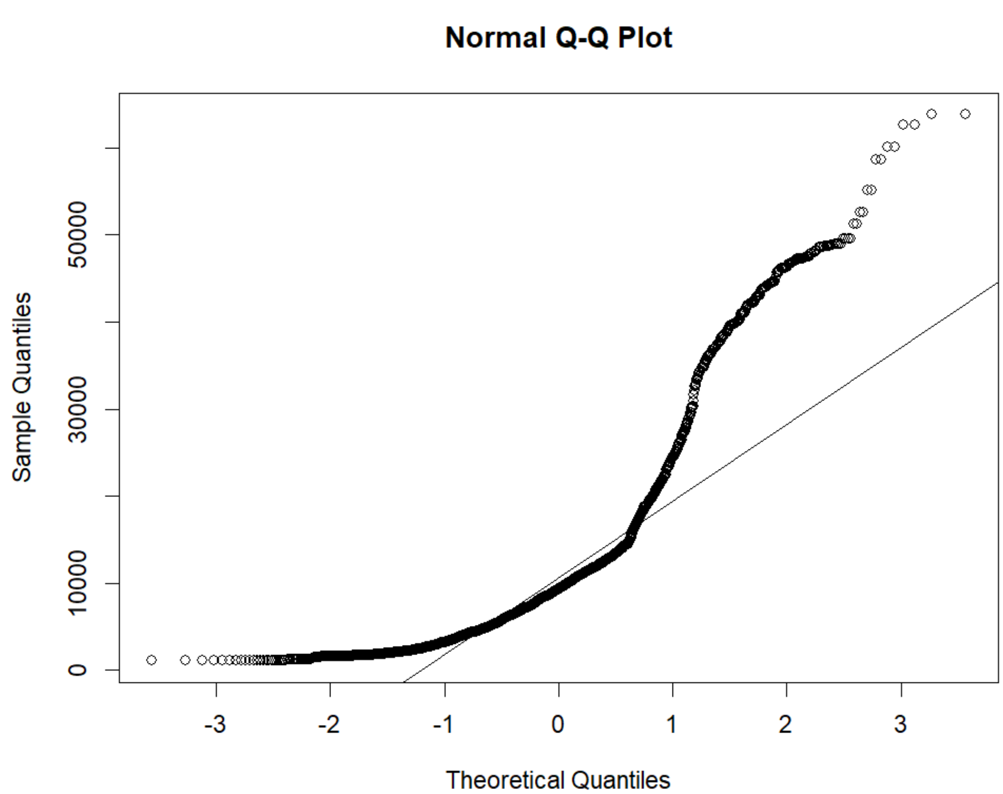
6.1.17 Scatterplot BMI vs Charges



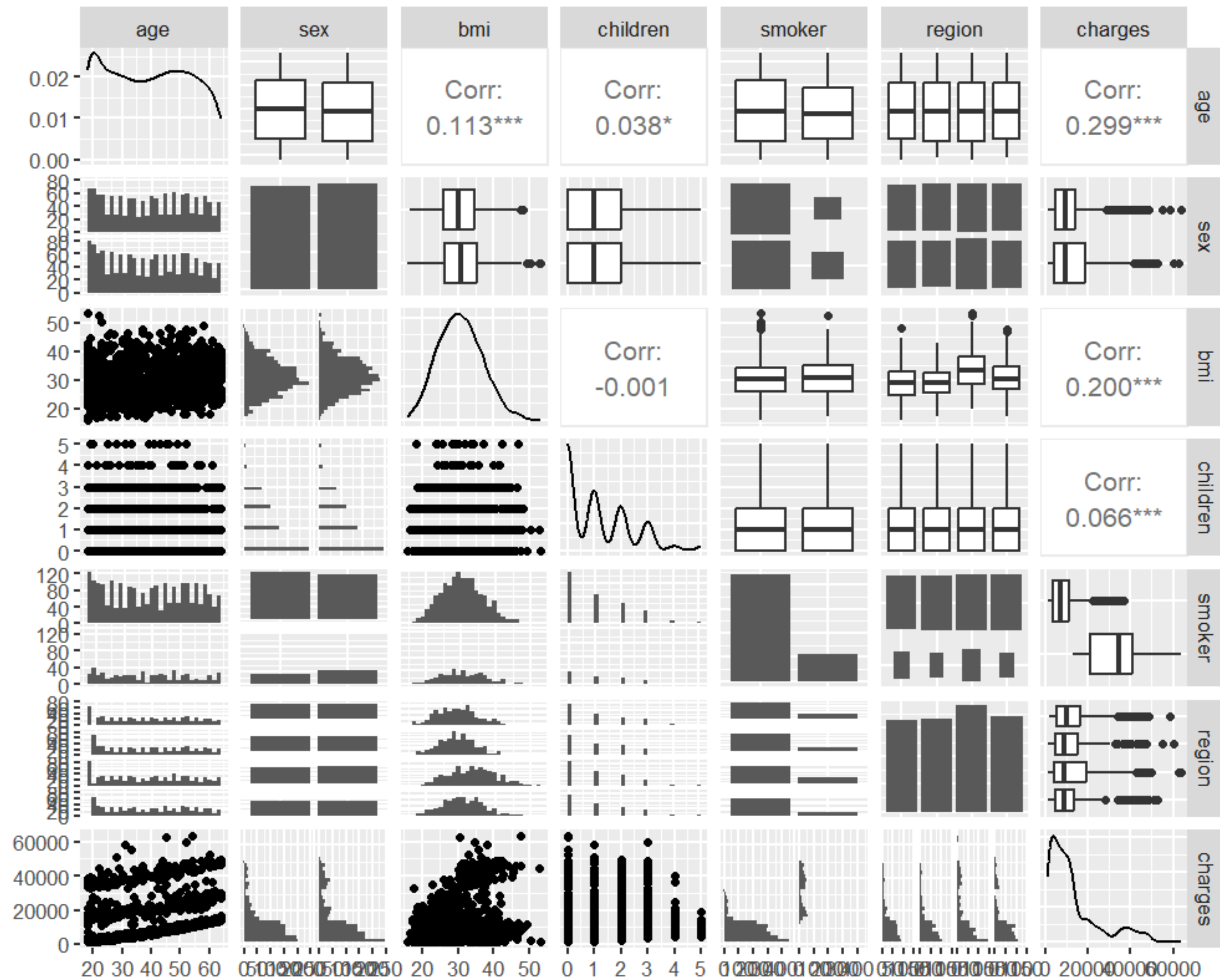
6.1.18 Scatterplot Children vs Charges



6.1.19 QQ plot for Charges



6.1.20 GGPAIRS() Plot



6.2 Model 1 Diagnostics

```
[1] "Variance Homogeneity Test"
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 380.7498, Df = 1, p = < 2.22e-16
[1] "Residual Normality Test"

        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.8941, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.05627569      2.111681   0.008
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
            age         sexmale              bmi        smokeryes        children regionnor
thwest regionsoutheast
         1.0186          1.0107           1.1228           1.0165           1.0039
1.5516          1.7101
regionsouthwest
         1.5584
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
2153 5.053434         4.6955e-07    0.0010424
455  4.252224         2.2047e-05    0.0489440
1559 4.252224         2.2047e-05    0.0489440
[1] "Influential Points"
Potentially influential observations of
         lm(formula = charges ~ age + sex + bmi + smoker + children +      region, data =
trainData) :

     dfb.1_ dfb.age dfb.sxml dfb.bmi dfb.smkr dfb.chld dfb.rgnn dfb.rgnsths dfb.rgnsthw d
ffit   cov.r   cook.d hat
8     0.01   0.09   -0.05    -0.05   -0.02    -0.06     0.09     0.01        0.00
0.18    0.98_*  0.00   0.00
86    0.09  -0.10   -0.07     0.03   -0.03    -0.05    -0.10    -0.10       -0.10
0.19    0.97_*  0.00   0.00
113  -0.01   0.05   -0.05    -0.02   -0.03     0.08     0.00     0.07        0.00
0.15    0.98_*  0.00   0.00
115   0.12  -0.02    0.08    -0.09   -0.05     0.06    -0.11    -0.09       -0.11
0.21_*  0.96_*  0.01   0.00
134   0.00   0.01    0.01    -0.01    0.00    -0.03     0.00     0.00       -0.01       -
0.04    1.01_*  0.00   0.01
179   0.18  -0.07   -0.08    -0.14   -0.05    -0.07     0.00     0.14        0.01
0.26_*  0.94_*  0.01   0.00
197   0.02   0.10   -0.08    -0.06   -0.03    -0.01     0.00     0.02        0.13
0.23_*  0.94_*  0.01   0.00
234   0.02   0.06    0.06    -0.08   -0.04     0.09     0.00     0.09        0.00
0.18    0.98_*  0.00   0.00
236   0.04  -0.04    0.05     0.01   -0.03     0.00    -0.07    -0.08       -0.07
0.12    0.98_*  0.00   0.00
247   0.04  -0.05   -0.05    -0.02   -0.02     0.04     0.00     0.01        0.08
0.14    0.98_*  0.00   0.00
259   0.05  -0.07   -0.06     0.02   -0.03     0.15    -0.10    -0.09       -0.10
0.22_*  0.97_*  0.01   0.01
272   0.06  -0.06   -0.06    -0.02   -0.02    -0.05     0.00     0.01        0.08
0.16    0.98_*  0.00   0.00
282   0.01   0.03    0.06    -0.04   -0.03    -0.06     0.00     0.01        0.09
0.15    0.98_*  0.00   0.00
302  -0.02   0.08    0.05    -0.03   -0.03     0.00     0.00     0.07        0.00
0.14    0.99_*  0.00   0.00
308   0.00   0.06    0.08    -0.06   -0.04     0.05     0.10     0.02        0.00
0.19    0.96_*  0.00   0.00
315   0.04  -0.06    0.05    -0.02   -0.04    -0.04     0.00     0.06        0.00
0.13    0.99_*  0.00   0.00
334   0.00   0.00    0.00     0.00    0.00     0.01     0.00     0.00        0.00
0.01    1.01_*  0.00   0.01
338   0.02  -0.10    0.07     0.04   -0.03    -0.06     0.00    -0.01        0.10
0.21_*  0.96_*  0.00   0.00
```

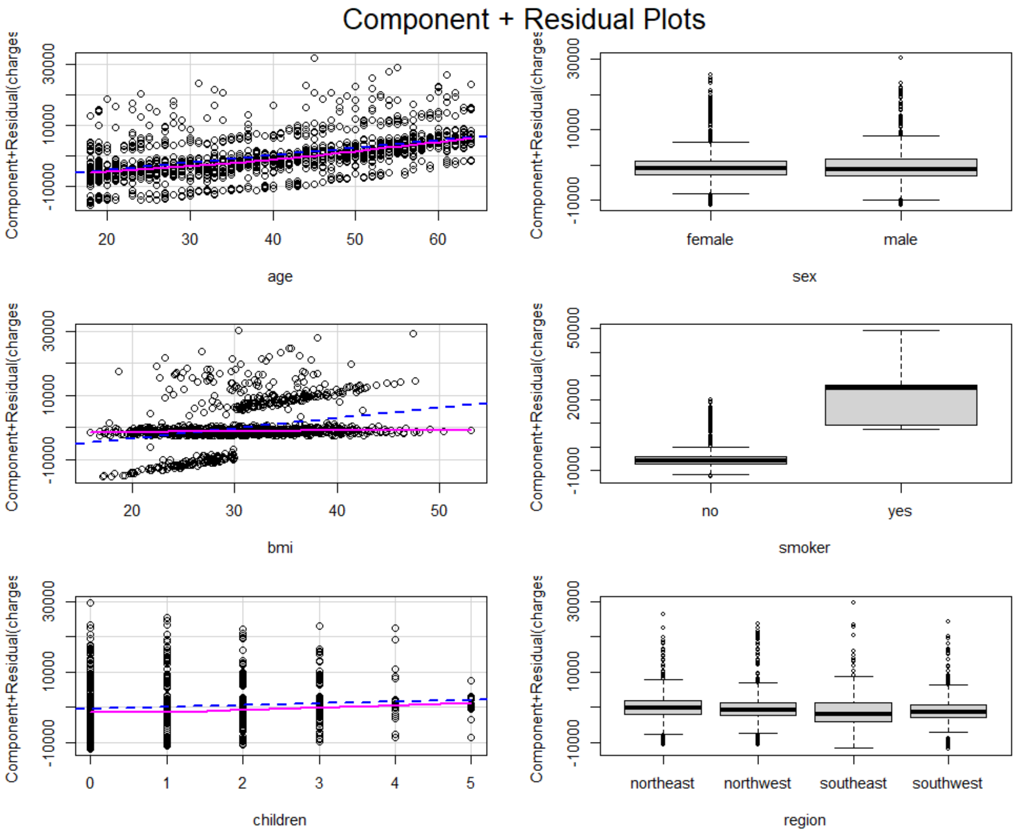
367	0.12	-0.05	-0.07	-0.05	-0.03	0.00	-0.10	-0.08	-0.10	0.17	0.97_*	0.00	0.00
385	-0.03	0.02	-0.06	0.04	0.10	-0.04	0.00	0.05	0.00	0.16	0.99_*	0.00	0.00
409	0.02	-0.11	0.08	0.03	-0.06	0.00	0.00	0.09	0.00	0.21_*	0.95_*	0.01	0.00
413	0.05	0.04	-0.05	-0.02	-0.02	-0.05	-0.08	-0.07	-0.08	0.15	0.98_*	0.00	0.00
419	0.03	-0.11	-0.07	0.03	-0.03	0.05	0.09	-0.01	0.00	0.20_*	0.96_*	0.00	0.00
428	0.00	0.07	0.06	-0.03	-0.04	-0.05	0.00	0.08	0.00	0.16	0.98_*	0.00	0.00
432	-0.18	0.07	-0.11	0.20	0.17	-0.08	0.00	0.05	-0.01	0.34_*	0.95_*	0.01	0.01
451	-0.04	0.08	-0.05	0.06	-0.02	-0.01	-0.08	-0.09	-0.09	0.16	0.98_*	0.00	0.00
455	-0.03	-0.07	-0.12	0.16	0.19	0.00	-0.13	-0.18	-0.14	0.32_*	0.94_*	0.01	0.01
472	-0.09	0.05	-0.07	0.09	-0.02	0.04	0.09	-0.02	-0.01	0.19	0.97_*	0.00	0.00
520	-0.01	0.02	-0.05	0.05	-0.02	0.00	-0.07	-0.08	-0.08	0.13	0.99_*	0.00	0.00
544	0.06	0.04	-0.06	-0.07	-0.02	-0.01	0.00	0.02	0.10	0.17	0.97_*	0.00	0.00
551	0.00	0.05	-0.05	0.03	-0.02	0.04	-0.09	-0.09	-0.09	0.15	0.98_*	0.00	0.00
638	-0.08	-0.01	-0.06	0.12	-0.02	-0.01	0.08	-0.03	-0.01	0.18	0.98_*	0.00	0.00
681	0.04	-0.05	-0.05	0.00	-0.03	0.00	0.00	0.06	0.00	0.12	0.99_*	0.00	0.00
742	0.02	0.02	0.08	0.00	-0.05	0.06	-0.12	-0.11	-0.12	0.19_*	0.96_*	0.00	0.00
775	0.00	0.01	0.00	0.00	0.00	-0.01	0.01	0.01	0.01	-0.02	1.01_*	0.00	0.01
777	0.02	0.03	-0.06	-0.02	-0.03	-0.01	0.09	0.01	0.00	0.16	0.97_*	0.00	0.00
796	0.10	-0.06	0.07	-0.04	-0.04	0.05	-0.10	-0.09	-0.10	0.19	0.97_*	0.00	0.00
799	-0.05	0.11	-0.07	-0.02	-0.04	0.18	-0.01	0.10	-0.01	0.27_*	0.96_*	0.01	0.01
805	0.01	-0.10	-0.07	0.06	-0.03	0.05	0.10	-0.01	-0.01	0.21_*	0.96_*	0.00	0.00
812	0.14	-0.06	0.08	-0.13	-0.04	-0.07	0.11	0.03	0.01	0.24_*	0.96_*	0.01	0.00
824	0.04	-0.10	0.07	-0.02	-0.04	0.05	0.09	0.00	0.00	0.19_*	0.97_*	0.00	0.00
894	0.00	0.00	0.00	-0.01	0.00	0.01	0.00	0.01	0.00	0.02	1.01_*	0.00	0.01
905	-0.06	0.08	0.05	0.01	0.11	-0.06	0.01	-0.01	0.08	0.20_*	0.98_*	0.00	0.01
943	0.04	-0.08	-0.05	0.00	-0.02	0.08	0.07	0.00	0.00	0.16	0.98_*	0.00	0.00
971	-0.12	0.06	0.07	0.06	0.15	0.11	0.11	-0.02	0.00	0.28_*	0.96_*	0.01	0.01
1052	0.13	-0.08	-0.07	-0.04	-0.03	0.06	-0.10	-0.08	-0.10	0.19	0.97_*	0.00	0.00
1083	0.02	0.02	0.08	0.00	-0.05	0.06	-0.12	-0.11	-0.12	0.19_*	0.96_*	0.00	0.00
1089	0.01	-0.10	-0.07	0.06	-0.03	0.05	0.10	-0.01	-0.01	0.21_*	0.96_*	0.00	0.00
1096	0.14	-0.06	0.08	-0.13	-0.04	-0.07	0.11	0.03	0.01	0.24_*	0.96_*	0.01	0.00
1099	0.07	-0.01	0.07	-0.07	-0.04	-0.06	0.09	0.02	0.01	0.18	0.97_*	0.00	0.00
1104	0.01	0.09	-0.05	-0.05	-0.02	-0.06	0.09	0.01	0.00	0.18	0.98_*	0.00	0.00
1123	-0.06	-0.07	0.06	0.08	0.15	-0.01	0.00	-0.03	0.10	0.24_*	0.96_*	0.01	0.00
1149	-0.01	0.12	0.07	-0.06	-0.03	-0.01	0.09	0.02	0.00	0.19	0.98_*	0.00	0.00
1191	0.02	0.10	0.07	-0.02	-0.04	-0.06	-0.09	-0.08	-0.09	0.18	0.97_*	0.00	0.00

```
1210  0.12  -0.02    0.08    -0.09   -0.05     0.06    -0.11    -0.09       -0.11
0.21_*  0.96_*  0.01   0.00
1233  0.00   0.01    0.01    -0.01    0.00    -0.03     0.00     0.00       -0.01       -
0.04    1.01_*  0.00   0.01
1291  0.02   0.10   -0.08    -0.06   -0.03    -0.01     0.00     0.02        0.13
0.23_*  0.94_*  0.01   0.00
1326  0.02   0.06    0.06    -0.08   -0.04     0.09     0.00     0.09        0.00
0.18    0.98_*  0.00   0.00
1342  0.04  -0.05   -0.05    -0.02   -0.02     0.04     0.00     0.01        0.08
0.14    0.98_*  0.00   0.00
1354  0.05  -0.07   -0.06     0.02   -0.03     0.15    -0.10    -0.09       -0.10
0.22_*  0.97_*  0.01   0.01
1369  0.06  -0.06   -0.06    -0.02   -0.02    -0.05     0.00     0.01        0.08
0.16    0.98_*  0.00   0.00
1400 -0.02   0.08    0.05    -0.03   -0.03     0.00     0.00     0.07        0.00
0.14    0.99_*  0.00   0.00
1408  0.00   0.06    0.08    -0.06   -0.04     0.05     0.10     0.02        0.00
0.19    0.96_*  0.00   0.00
1417  0.04  -0.06    0.05    -0.02   -0.04    -0.04     0.00     0.06        0.00
0.13    0.99_*  0.00   0.00
1446  0.02  -0.10    0.07     0.04   -0.03    -0.06     0.00    -0.01        0.10
0.21_*  0.96_*  0.00   0.00
1456 -0.03   0.07   -0.05     0.02   -0.03     0.00     0.00     0.06        0.00
0.13    0.99_*  0.00   0.00
1474  0.12  -0.05   -0.07    -0.05   -0.03     0.00    -0.10    -0.08       -0.10
0.17    0.97_*  0.00   0.00
1489 -0.03   0.02   -0.06     0.04    0.10    -0.04     0.00     0.05        0.00
0.16    0.99_*  0.00   0.00
1511  0.02  -0.11    0.08     0.03   -0.06     0.00     0.00     0.09        0.00
0.21_*  0.95_*  0.01   0.00
1515  0.05   0.04   -0.05    -0.02   -0.02    -0.05    -0.08    -0.07       -0.08
0.15    0.98_*  0.00   0.00
1521  0.03  -0.11   -0.07     0.03   -0.03     0.05     0.09    -0.01        0.00
0.20_*  0.96_*  0.00   0.00
1530  0.00   0.07    0.06    -0.03   -0.04    -0.05     0.00     0.08        0.00
0.16    0.98_*  0.00   0.00
1556 -0.04   0.08   -0.05     0.06   -0.02    -0.01    -0.08    -0.09       -0.09
0.16    0.98_*  0.00   0.00
1559 -0.03  -0.07   -0.12     0.16    0.19     0.00    -0.13    -0.18       -0.14
0.32_*  0.94_*  0.01   0.01
1565  0.08  -0.01   -0.05    -0.08   -0.03     0.00     0.00     0.08        0.01
0.14    0.99_*  0.00   0.00
1624 -0.01   0.02   -0.05     0.05   -0.02     0.00    -0.07    -0.08       -0.08
0.13    0.99_*  0.00   0.00
 [ reached getOption("max.print") -- omitted 24 rows ]
           dfb.1_        dfb.age      dfb.sxml        dfb.bmi      dfb.smkr       dfb.chld
  dfb.rgnn  dfb.rgnsths
8     0.013033432  0.0949233462 -0.052139100 -0.0450727248 -0.022008306 -6.086001e-02  0.
0858400083  0.014976576
86    0.093752943 -0.0997340509 -0.066747594  0.0257445787 -0.029370124 -5.201398e-02 -0.
0968054536 -0.101826598
113  -0.008758281  0.0538389246 -0.050413576 -0.0205653841 -0.029137476  8.155987e-02 -0.
0037201116  0.072458426
115   0.120051250 -0.0182323800  0.081968076 -0.0887330083 -0.050914616  6.025444e-02 -0.
1145472212 -0.087904830
134   0.004760368  0.0138516029  0.008766721 -0.0109207021  0.002772154 -2.642027e-02  0.
0011870835  0.002731017
179   0.178826449 -0.0697737808 -0.081952535 -0.1394542296 -0.051918917 -6.817368e-02  0.
0009078536  0.144147625
197   0.024813991  0.1005381946 -0.078609309 -0.0643442324 -0.028129389 -1.402463e-02 -0.
0006705617  0.021329784
234   0.018296814  0.0589185550  0.058154237 -0.0774341169 -0.041494089  8.614500e-02 -0.
0015181391  0.090646979
236   0.041570940 -0.0356888624  0.049949894  0.0064582189 -0.030897306 -1.105677e-03 -0.
0738977450 -0.075013873
247   0.043967611 -0.0456824497 -0.053359924 -0.0207719900 -0.019420947  4.061994e-02 -0.
0024814115  0.006697291
259   0.054837861 -0.0679462152 -0.064564573  0.0160828186 -0.030557768  1.546666e-01 -0.
0998201567 -0.093948905
272   0.063907275 -0.0610958666 -0.056048437 -0.0190872310 -0.019474990 -5.213023e-02  0.
0001745753  0.005207197
```

```
282   0.014094539  0.0339973001  0.063972517 -0.0360601557 -0.028164424 -5.685572e-02  0.
0025953004  0.009895697
302  -0.018454745  0.0827993852  0.049967423 -0.0304573452 -0.033460308 -4.167373e-03  0.
0006595976  0.068497460
308   0.003438157  0.0609545382  0.079357449 -0.0584808743 -0.039486308  4.624515e-02  0.
0990364897  0.017505360
315   0.041619372 -0.0554155709  0.047591312 -0.0176172929 -0.035809681 -3.886558e-02  0.
0011459681  0.064020968
334   0.001157356  0.0011103304  0.002092623 -0.0033829075 -0.001573930  6.403511e-03 -0.
0001490131  0.003453179
338   0.015428775 -0.1039212780  0.070121969  0.0401924436 -0.033701133 -6.322906e-02  0.
0014561654 -0.013125268
367   0.124970095 -0.0495344416 -0.065385175 -0.0461361748 -0.032646702 -7.580022e-04 -0.
1014076970 -0.084985858
385  -0.034171823  0.0167117469 -0.063250995  0.0413295136  0.100834058 -4.460076e-02  0.
0024332796  0.048921907
409   0.016724299 -0.1097354891  0.077252117  0.0317739313 -0.059699777  2.269562e-03 -0.
0009701238  0.092086960
413   0.053497680  0.0444107921 -0.051861828 -0.0196491705 -0.024185309 -4.907337e-02 -0.
0812497126 -0.073186755
419   0.031488984 -0.1068627362 -0.069722668  0.0305631551 -0.027684864  5.266265e-02  0.
0940853164 -0.007605413
428  -0.001848192  0.0665658799  0.060865993 -0.0342234292 -0.041226020 -5.167704e-02  0.
0020710370  0.082574323
432  -0.184984344  0.0716822267 -0.109707448  0.1974438144  0.172510654 -7.605187e-02  0.
0025196967  0.047269864
451  -0.038300794  0.0762341414 -0.053000439  0.0623788955 -0.020739186 -5.123060e-03 -0.
0807491918 -0.091776826
455  -0.029769129 -0.0705369219 -0.115412042  0.1567326646  0.188841202 -3.670141e-03 -0.
1330342394 -0.181283584
472  -0.089878182  0.0459685277 -0.065420697  0.0892657106 -0.022062491  4.443170e-02  0.
0893526649 -0.021080766
520  -0.010739629  0.0216752389 -0.048823444  0.0543000926 -0.019519506 -2.827275e-03 -0.
0729514019 -0.082990246
544   0.056253648  0.0417497600 -0.058832432 -0.0742941609 -0.022526333 -9.224324e-03 -0.
0002954481  0.022493105
551   0.001087975  0.0475893890 -0.054777747  0.0262182964 -0.023968391  4.163281e-02 -0.
0855145626 -0.085608262
638  -0.082339223 -0.0125617771 -0.058836974  0.1203078558 -0.017791660 -6.936593e-03  0.
0769852581 -0.031221374
681   0.036436736 -0.0462782193 -0.049472941 -0.0045158236 -0.027336708  1.122459e-03 -0.
0016723260  0.061118291
742   0.024444191  0.0248611109  0.078845995  0.0012828136 -0.047206299  5.827183e-02 -0.
1151539412 -0.111161589
775  -0.002565602  0.0065974599 -0.004243319 -0.0013280230  0.002928195 -1.453702e-02  0.
0070048989  0.006772764
777   0.017801445  0.0289122528 -0.061653799 -0.0203307264 -0.025507438 -9.644543e-03  0.
0935458972  0.007830716
796   0.097869565 -0.0629063666  0.069021934 -0.0425819366 -0.044167325  5.440717e-02 -0.
1007822854 -0.087096249
799  -0.048785374  0.1112201460 -0.073494382 -0.0159806011 -0.042128717  1.788344e-01 -0.
0073189126  0.102826811
805   0.008566836 -0.1039871650 -0.073453513  0.0567149419 -0.027906687  5.449055e-02  0.
0976054878 -0.014538804
812   0.139539854 -0.0649784131  0.084427551 -0.1301902383 -0.044714076 -6.905825e-02  0.
1089798968  0.033745875
824   0.041763376 -0.0951725148  0.071018134 -0.0155267733 -0.039258041  5.035147e-02  0.
0942871807  0.003262570
894   0.003834132  0.0002956859 -0.003393898 -0.0058138173 -0.002266476  1.156667e-02 -0.
0003866097  0.006181171
905  -0.060402655  0.0836214586  0.048722544  0.0127926393  0.114445602 -5.750685e-02  0.
0056610839 -0.007431607
943   0.036811227 -0.0791117906 -0.051609621 -0.0009681455 -0.022049273  8.293202e-02  0.
0698329520  0.001183305
971  -0.124478693  0.0635466884  0.066191286  0.0625894471  0.154787941  1.132476e-01  0.
1110202575 -0.021526350
1052  0.126603347 -0.0750904473 -0.065868190 -0.0437764342 -0.033372822  5.525446e-02 -0.
1024570304 -0.084844671
1083  0.024444191  0.0248611109  0.078845995  0.0012828136 -0.047206299  5.827183e-02 -0.
1151539412 -0.111161589
1089  0.008566836 -0.1039871650 -0.073453513  0.0567149419 -0.027906687  5.449055e-02  0.
0976054878 -0.014538804
```

```
1096  0.139539854 -0.0649784131  0.084427551 -0.1301902383 -0.044714076 -6.905825e-02  0.
1089798968  0.033745875
1099  0.067064456 -0.0148598630  0.070245219 -0.0717220073 -0.035725412 -5.956791e-02  0.
0911927967  0.018815466
1104  0.013033432  0.0949233462 -0.052139100 -0.0450727248 -0.022008306 -6.086001e-02  0.
0858400083  0.014976576
1123 -0.060349798 -0.0651349497  0.056473223  0.0824884516  0.148355897 -8.891698e-03  0.
0040526109 -0.029874517
1149 -0.007604369  0.1158777548  0.070494271 -0.0647656197 -0.032772125 -1.256884e-02  0.
0869520805  0.019568530
1191  0.016040436  0.0951174780  0.066038765 -0.0198815274 -0.035977006 -5.683011e-02 -0.
0882371479 -0.082659427
1210  0.120051250 -0.0182323800  0.081968076 -0.0887330083 -0.050914616  6.025444e-02 -0.
1145472212 -0.087904830
1233  0.004760368  0.0138516029  0.008766721 -0.0109207021  0.002772154 -2.642027e-02  0.
0011870835  0.002731017
1291  0.024813991  0.1005381946 -0.078609309 -0.0643442324 -0.028129389 -1.402463e-02 -0.
0006705617  0.021329784
1326  0.018296814  0.0589185550  0.058154237 -0.0774341169 -0.041494089  8.614500e-02 -0.
0015181391  0.090646979
1342  0.043967611 -0.0456824497 -0.053359924 -0.0207719900 -0.019420947  4.061994e-02 -0.
0024814115  0.006697291
1354  0.054837861 -0.0679462152 -0.064564573  0.0160828186 -0.030557768  1.546666e-01 -0.
0998201567 -0.093948905
1369  0.063907275 -0.0610958666 -0.056048437 -0.0190872310 -0.019474990 -5.213023e-02  0.
0001745753  0.005207197
1400 -0.018454745  0.0827993852  0.049967423 -0.0304573452 -0.033460308 -4.167373e-03  0.
0006595976  0.068497460
1408  0.003438157  0.0609545382  0.079357449 -0.0584808743 -0.039486308  4.624515e-02  0.
0990364897  0.017505360
1417  0.041619372 -0.0554155709  0.047591312 -0.0176172929 -0.035809681 -3.886558e-02  0.
0011459681  0.064020968
1446  0.015428775 -0.1039212780  0.070121969  0.0401924436 -0.033701133 -6.322906e-02  0.
0014561654 -0.013125268
1456 -0.034475211  0.0669555878 -0.048786997  0.0167921560 -0.025434830 -3.227300e-03 -0.
0015444208  0.058619123
1474  0.124970095 -0.0495344416 -0.065385175 -0.0461361748 -0.032646702 -7.580022e-04 -0.
1014076970 -0.084985858
1489 -0.034171823  0.0167117469 -0.063250995  0.0413295136  0.100834058 -4.460076e-02  0.
0024332796  0.048921907
1511  0.016724299 -0.1097354891  0.077252117  0.0317739313 -0.059699777  2.269562e-03 -0.
0009701238  0.092086960
1515  0.053497680  0.0444107921 -0.051861828 -0.0196491705 -0.024185309 -4.907337e-02 -0.
0812497126 -0.073186755
1521  0.031488984 -0.1068627362 -0.069722668  0.0305631551 -0.027684864  5.266265e-02  0.
0940853164 -0.007605413
1530 -0.001848192  0.0665658799  0.060865993 -0.0342234292 -0.041226020 -5.167704e-02  0.
0020710370  0.082574323
1556 -0.038300794  0.0762341414 -0.053000439  0.0623788955 -0.020739186 -5.123060e-03 -0.
0807491918 -0.091776826
1559 -0.029769129 -0.0705369219 -0.115412042  0.1567326646  0.188841202 -3.670141e-03 -0.
1330342394 -0.181283584
1565  0.084850525 -0.0141693862 -0.046944029 -0.0802188907 -0.030020216 -6.538363e-06 -0.
0005936340  0.084021159
1624 -0.010739629  0.0216752389 -0.048823444  0.0543000926 -0.019519506 -2.827275e-03 -0.
0729514019 -0.082990246
       dfb.rgnsthw        dffit     cov.r       cook.d          hat
8      2.818738e-03  0.179445531 0.9755182 3.566439e-03 0.003970466
86    -9.692288e-02  0.190786897 0.9707708 4.029340e-03 0.003912836
113   -2.865720e-03  0.152967215 0.9845416 2.594232e-03 0.004005379
115   -1.082873e-01  0.213363585 0.9553517 5.030680e-03 0.003461530
134   -9.886407e-03 -0.036039887 1.0123303 1.443750e-04 0.008744527
179    1.106043e-02  0.262252700 0.9449977 7.590328e-03 0.004279520
197    1.263729e-01  0.229627335 0.9404642 5.816911e-03 0.003110050
234    2.436357e-03  0.176374976 0.9822923 3.447791e-03 0.004720701
236   -7.421027e-02  0.123424315 0.9845785 1.689179e-03 0.002761136
247    7.791569e-02  0.142085056 0.9798133 2.237310e-03 0.002981505
259   -9.896521e-02  0.222723901 0.9747560 5.492660e-03 0.005674332
272    8.426641e-02  0.157940751 0.9774070 2.763642e-03 0.003339365
282    8.667655e-02  0.154269250 0.9765228 2.636448e-03 0.003116355
302    8.311637e-04  0.135757666 0.9877837 2.044153e-03 0.003718231
308    2.176108e-03  0.189051172 0.9630182 3.953141e-03 0.003217590
```

```
315    1.973584e-03  0.130963962 0.9862037 1.902096e-03 0.003273473
334    5.894938e-05  0.008540768 1.0125835 8.108605e-06 0.008432727
338    9.881231e-02  0.207457593 0.9636193 4.760371e-03 0.003869497
367   -9.685540e-02  0.174332115 0.9666958 3.363044e-03 0.003006191
385    1.438668e-03  0.161867645 0.9854422 2.905024e-03 0.004552810
409   -3.502784e-03  0.214940430 0.9477713 5.101015e-03 0.003074871
413   -8.002322e-02  0.145483539 0.9801650 2.345667e-03 0.003147416
419   -4.906004e-03  0.198269530 0.9642373 4.348482e-03 0.003607886
428    2.327531e-03  0.156602223 0.9770208 2.716905e-03 0.003250391
432   -8.795220e-03  0.338783198 0.9485168 1.266781e-02 0.007206745
451   -8.559347e-02  0.163282966 0.9830612 2.955353e-03 0.004248334
455   -1.379195e-01  0.324238611 0.9385059 1.159163e-02 0.005780686
472   -1.099125e-02  0.190409822 0.9682874 4.012394e-03 0.003671675
520   -7.659357e-02  0.128299686 0.9867158 1.825607e-03 0.003227727
544    9.752998e-02  0.170778662 0.9687464 3.228090e-03 0.003039905
551   -8.726081e-02  0.149390725 0.9796605 2.473183e-03 0.003245314
638   -1.139723e-02  0.176690172 0.9794292 3.459154e-03 0.004328890
681   -9.740285e-04  0.117294573 0.9863935 1.525881e-03 0.002722811
742   -1.158052e-01  0.192778641 0.9550926 4.106961e-03 0.002846139
775    6.993939e-03 -0.018911924 1.0122706 3.975730e-05 0.008264860
777   -1.943887e-04  0.156926638 0.9660541 2.725002e-03 0.002432025
796   -9.720020e-02  0.186120350 0.9673728 3.833345e-03 0.003446748
799   -7.265662e-03  0.272285419 0.9608660 8.195778e-03 0.005986180
805   -7.260041e-03  0.206500100 0.9607350 4.715094e-03 0.003614129
812    1.243607e-02  0.243049226 0.9593232 6.529993e-03 0.004741812
824   -1.469825e-05  0.192597956 0.9664945 4.104332e-03 0.003596512
894    4.595335e-05  0.015206874 1.0124019 2.570564e-05 0.008331354
905    8.421137e-02  0.197539463 0.9821167 4.324341e-03 0.005673733
943   -2.846573e-03  0.164482589 0.9836079 2.999079e-03 0.004383056
971   -1.048608e-03  0.277939626 0.9558887 8.535058e-03 0.005682103
1052 -9.757055e-02  0.189820006 0.9678832 3.987415e-03 0.003616349
1083 -1.158052e-01  0.192778641 0.9550926 4.106961e-03 0.002846139
1089 -7.260041e-03  0.206500100 0.9607350 4.715094e-03 0.003614129
1096  1.243607e-02  0.243049226 0.9593232 6.529993e-03 0.004741812
1099  6.943604e-03  0.175310404 0.9723689 3.402920e-03 0.003497033
1104  2.818738e-03  0.179445531 0.9755182 3.566439e-03 0.003970466
1123  1.045799e-01  0.242641307 0.9622115 6.510080e-03 0.004998250
1149  3.879411e-03  0.189749561 0.9756611 3.987652e-03 0.004399457
1191 -8.855824e-02  0.181525321 0.9747497 3.649270e-03 0.003967941
1210 -1.082873e-01  0.213363585 0.9553517 5.030680e-03 0.003461530
1233 -9.886407e-03 -0.036039887 1.0123303 1.443750e-04 0.008744527
1291  1.263729e-01  0.229627335 0.9404642 5.816911e-03 0.003110050
1326  2.436357e-03  0.176374976 0.9822923 3.447791e-03 0.004720701
1342  7.791569e-02  0.142085056 0.9798133 2.237310e-03 0.002981505
1354 -9.896521e-02  0.222723901 0.9747560 5.492660e-03 0.005674332
1369  8.426641e-02  0.157940751 0.9774070 2.763642e-03 0.003339365
1400  8.311637e-04  0.135757666 0.9877837 2.044153e-03 0.003718231
1408  2.176108e-03  0.189051172 0.9630182 3.953141e-03 0.003217590
1417  1.973584e-03  0.130963962 0.9862037 1.902096e-03 0.003273473
1446  9.881231e-02  0.207457593 0.9636193 4.760371e-03 0.003869497
1456 -3.774544e-03  0.131008558 0.9859852 1.903351e-03 0.003246090
1474 -9.685540e-02  0.174332115 0.9666958 3.363044e-03 0.003006191
1489  1.438668e-03  0.161867645 0.9854422 2.905024e-03 0.004552810
1511 -3.502784e-03  0.214940430 0.9477713 5.101015e-03 0.003074871
1515 -8.002322e-02  0.145483539 0.9801650 2.345667e-03 0.003147416
1521 -4.906004e-03  0.198269530 0.9642373 4.348482e-03 0.003607886
1530  2.327531e-03  0.156602223 0.9770208 2.716905e-03 0.003250391
1556 -8.559347e-02  0.163282966 0.9830612 2.955353e-03 0.004248334
1559 -1.379195e-01  0.324238611 0.9385059 1.159163e-02 0.005780686
1565  5.057364e-03  0.138758971 0.9859983 2.135138e-03 0.003585479
1624 -7.659357e-02  0.128299686 0.9867158 1.825607e-03 0.003227727
 [ reached getOption("max.print") -- omitted 24 rows ]
```



Component + Residual Plots
Component+Residual(charges
age
sex
female
male
bmi
smoker
no
yes
children
region
northeast
northwest
southeast
southwest

6.3 MODEL 2 DIAGNOSTICS

```
[1] "Variance Homogeneity Test"
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 56.54749, Df = 1, p = 5.4857e-14
[1] "Residual Normality Test"

        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.81618, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.01801448      2.035987   0.352
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
            age         sexmale             bmi       smokeryes        children regionnor
thwest regionsoutheast
         1.0186          1.0107          1.1228          1.0165          1.0039
1.5516          1.7101
regionsouthwest
         1.5584
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
409  4.938055         8.4811e-07    0.0018828
1511 4.938055         8.4811e-07    0.0018828
179  4.859028         1.2624e-06    0.0028026
338  4.736456         2.3129e-06    0.0051346
1446 4.736456         2.3129e-06    0.0051346
812  4.543185         5.8395e-06    0.0129640
1096 4.543185         5.8395e-06    0.0129640
1930 4.543185         5.8395e-06    0.0129640
2219 4.543185         5.8395e-06    0.0129640
86   4.255599         2.1719e-05    0.0482170
[1] "Influential Points"
Potentially influential observations of
         lm(formula = boxcox.charges ~ age + sex + bmi + smoker + children +      region,
data = trainData) :

     dfb.1_ dfb.age dfb.sxml dfb.bmi dfb.smkr dfb.chld dfb.rgnn dfb.rgnsths dfb.rgnsthw d
ffit   cov.r   cook.d hat
28    0.00    0.00    0.00     0.00    0.00     0.00     0.00     0.00        0.00
0.00    1.01_*  0.00   0.01
86    0.13  -0.14   -0.09     0.04   -0.04    -0.07    -0.14    -0.14       -0.14
0.27_*  0.94_*  0.01   0.00
115   0.12  -0.02    0.08    -0.09   -0.05     0.06    -0.12    -0.09       -0.11
0.22_*  0.95_*  0.01   0.00
118   0.00  -0.04    0.06     0.01   -0.03     0.04     0.08     0.00        0.00
0.15    0.97_*  0.00   0.00
134   0.00   0.01    0.01    -0.01    0.00    -0.02     0.00     0.00       -0.01       -
0.03    1.01_*  0.00   0.01
179   0.22  -0.08   -0.10    -0.17   -0.06    -0.08     0.00     0.18        0.01
0.32_*  0.92_*  0.01   0.00
197   0.02   0.08   -0.06    -0.05   -0.02    -0.01     0.00     0.02        0.09
0.17    0.97_*  0.00   0.00
200  -0.02   0.05    0.05     0.00   -0.02    -0.05     0.07     0.00        0.00
0.13    0.99_*  0.00   0.00
234   0.02   0.05    0.05    -0.07   -0.04     0.08     0.00     0.08        0.00
0.16    0.99_*  0.00   0.00
236   0.06  -0.05    0.07     0.01   -0.04     0.00    -0.10    -0.10       -0.10
0.17    0.96_*  0.00   0.00
246  -0.03  -0.05    0.06     0.05   -0.03     0.04     0.08    -0.01       -0.01
0.16    0.97_*  0.00   0.00
247   0.06  -0.06   -0.07    -0.03   -0.02     0.05     0.00     0.01        0.10
0.18    0.96_*  0.00   0.00
259   0.06  -0.07   -0.07     0.02   -0.03     0.17    -0.11    -0.10       -0.11
0.24_*  0.97_*  0.01   0.01
272   0.09  -0.09   -0.08    -0.03   -0.03    -0.07     0.00     0.01        0.12
0.23_*  0.95_*  0.01   0.00
```

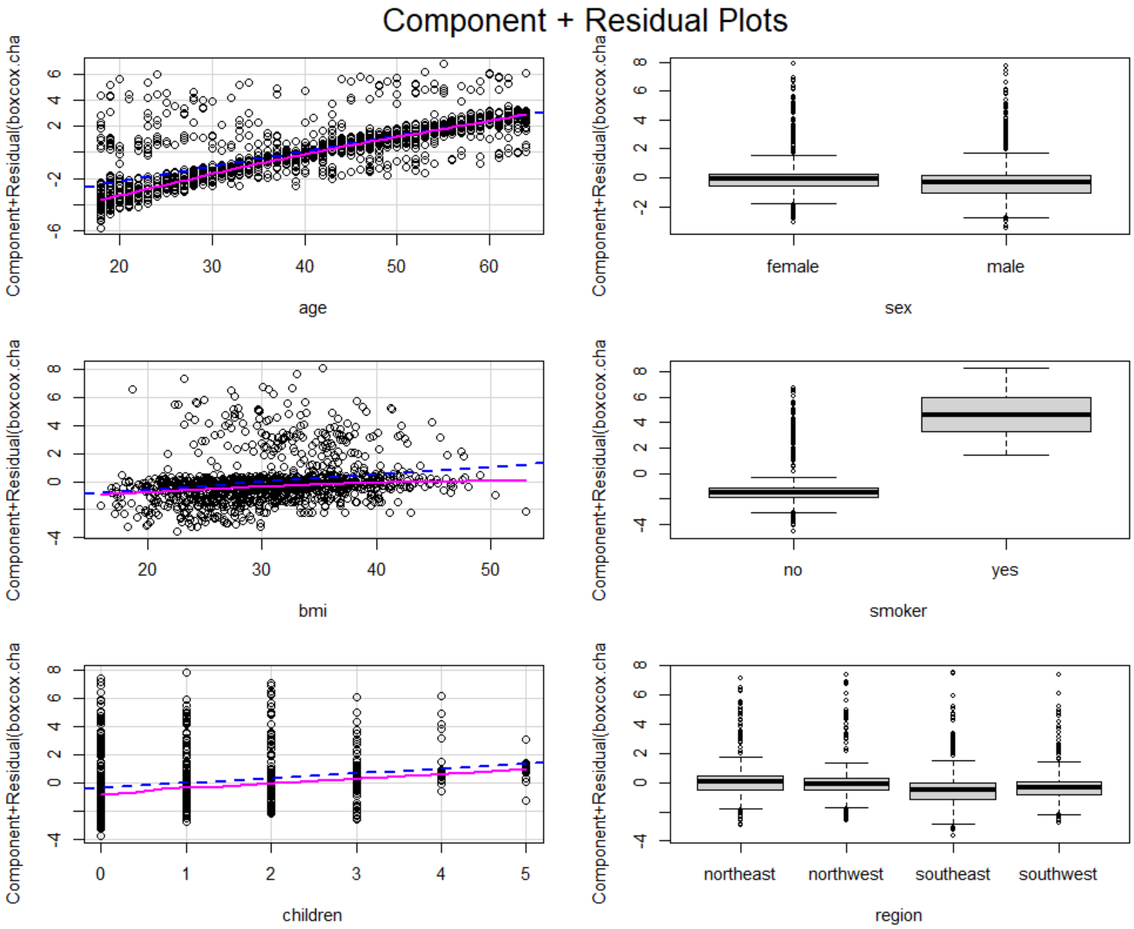
```
281   0.02  -0.11   -0.08     0.07   -0.04    -0.06     0.00     0.07       -0.01
0.21_*  0.96_*  0.00   0.00
282   0.02   0.04    0.07    -0.04   -0.03    -0.06     0.00     0.01        0.10
0.17    0.97_*  0.00   0.00
308   0.00   0.05    0.07    -0.05   -0.03     0.04     0.08     0.01        0.00
0.16    0.98_*  0.00   0.00
315   0.07  -0.09    0.08    -0.03   -0.06    -0.07     0.00     0.11        0.00
0.22_*  0.95_*  0.01   0.00
334   0.00   0.00    0.00     0.00    0.00     0.00     0.00     0.00        0.00
0.00    1.01_*  0.00   0.01
337   0.01  -0.06   -0.06     0.02   -0.02     0.09     0.08     0.00       -0.01
0.17    0.98_*  0.00   0.00
338   0.02  -0.15    0.10     0.06   -0.05    -0.09     0.00    -0.02        0.14
0.30_*  0.92_*  0.01   0.00
367   0.15  -0.06   -0.08    -0.05   -0.04     0.00    -0.12    -0.10       -0.11
0.20_*  0.95_*  0.00   0.00
409   0.02  -0.14    0.10     0.04   -0.08     0.00     0.00     0.12        0.00
0.27_*  0.91_*  0.01   0.00
413   0.05   0.04   -0.05    -0.02   -0.02    -0.05    -0.08    -0.07       -0.08
0.14    0.98_*  0.00   0.00
418   0.05  -0.09   -0.07     0.02   -0.04    -0.05     0.00     0.07        0.00
0.18    0.97_*  0.00   0.00
419   0.04  -0.13   -0.09     0.04   -0.03     0.07     0.12    -0.01       -0.01
0.25_*  0.94_*  0.01   0.00
428   0.00   0.07    0.06    -0.04   -0.04    -0.05     0.00     0.09        0.00
0.16    0.98_*  0.00   0.00
438  -0.01  -0.08   -0.07     0.14   -0.02    -0.05    -0.09    -0.13       -0.10
0.22_*  0.98_*  0.01   0.01
472  -0.07   0.03   -0.05     0.07   -0.02     0.03     0.07    -0.02       -0.01
0.14    0.98_*  0.00   0.00
520  -0.01   0.02   -0.05     0.05   -0.02     0.00    -0.07    -0.08       -0.07
0.13    0.99_*  0.00   0.00
544   0.05   0.04   -0.06    -0.07   -0.02    -0.01     0.00     0.02        0.09
0.16    0.97_*  0.00   0.00
599   0.01  -0.06    0.05     0.04   -0.03     0.13    -0.08    -0.09       -0.08
0.19    0.99_*  0.00   0.01
638  -0.09  -0.01   -0.06     0.13   -0.02    -0.01     0.08    -0.03       -0.01
0.19    0.97_*  0.00   0.00
681   0.06  -0.07   -0.08    -0.01   -0.04     0.00     0.00     0.09        0.00
0.18    0.96_*  0.00   0.00
742   0.02   0.02    0.07     0.00   -0.04     0.05    -0.10    -0.10       -0.10
0.17    0.97_*  0.00   0.00
755   0.04  -0.06    0.06    -0.02   -0.03    -0.01     0.08     0.00        0.00
0.14    0.98_*  0.00   0.00
774   0.06  -0.05   -0.06     0.02   -0.03     0.00    -0.09    -0.09       -0.09
0.15    0.98_*  0.00   0.00
775   0.00   0.01    0.00     0.00    0.00    -0.01     0.01     0.01        0.01          -
0.02    1.01_*  0.00   0.01
777   0.02   0.03   -0.06    -0.02   -0.02    -0.01     0.09     0.01        0.00
0.15    0.97_*  0.00   0.00
796   0.12  -0.08    0.08    -0.05   -0.05     0.07    -0.12    -0.11       -0.12
0.23_*  0.95_*  0.01   0.00
805   0.01  -0.13   -0.09     0.07   -0.03     0.07     0.12    -0.02       -0.01
0.25_*  0.94_*  0.01   0.00
812   0.18  -0.08    0.11    -0.17   -0.06    -0.09     0.14     0.04        0.02
0.31_*  0.93_*  0.01   0.00
824   0.06  -0.13    0.09    -0.02   -0.05     0.07     0.12     0.00        0.00
0.25_*  0.94_*  0.01   0.00
888   0.05  -0.06   -0.07     0.05   -0.03     0.00    -0.10    -0.10       -0.10
0.17    0.97_*  0.00   0.00
892  -0.01  -0.02    0.05     0.02   -0.02    -0.01     0.00    -0.01        0.07
0.11    0.99_*  0.00   0.00
894   0.00   0.00    0.00     0.00    0.00     0.00     0.00     0.00        0.00
0.01    1.01_*  0.00   0.01
898  -0.02  -0.06    0.07     0.06   -0.04    -0.07     0.10    -0.02        0.00
0.20_*  0.96_*  0.00   0.00
913   0.08  -0.06   -0.06    -0.05   -0.03     0.04     0.08     0.02        0.00
0.17    0.98_*  0.00   0.00
917  -0.03  -0.05    0.06     0.06   -0.05     0.00     0.00     0.07       -0.01
0.16    0.97_*  0.00   0.00
938   0.07  -0.07   -0.06    -0.02   -0.02    -0.06     0.00     0.00        0.09
0.18    0.97_*  0.00   0.00
```

```
943   0.05  -0.11   -0.07     0.00   -0.03     0.11     0.10     0.00        0.00
0.22_*  0.96_*  0.01   0.00
959   0.00  -0.10    0.05     0.11   -0.03    -0.05    -0.09    -0.12       -0.10
0.21_*  0.98_*  0.00   0.01
1042  0.07  -0.08    0.06     0.00   -0.04     0.00    -0.08    -0.08       -0.08
0.16    0.98_*  0.00   0.00
1052  0.15  -0.09   -0.08    -0.05   -0.04     0.07    -0.12    -0.10       -0.12
0.23_*  0.95_*  0.01   0.00
1054  0.02  -0.06   -0.05     0.03   -0.02    -0.05     0.00    -0.01        0.07
0.14    0.98_*  0.00   0.00
1083  0.02   0.02    0.07     0.00   -0.04     0.05    -0.10    -0.10       -0.10
0.17    0.97_*  0.00   0.00
1089  0.01  -0.13   -0.09     0.07   -0.03     0.07     0.12    -0.02       -0.01
0.25_*  0.94_*  0.01   0.00
1096  0.18  -0.08    0.11    -0.17   -0.06    -0.09     0.14     0.04        0.02
0.31_*  0.93_*  0.01   0.00
1099  0.08  -0.02    0.09    -0.09   -0.04    -0.07     0.11     0.02        0.01
0.22_*  0.95_*  0.01   0.00
1121  0.00   0.00    0.00     0.00    0.00     0.00     0.00     0.00        0.00
0.00    1.01_*  0.00   0.01
1191  0.01   0.08    0.05    -0.02   -0.03    -0.05    -0.07    -0.07       -0.07
0.15    0.99_*  0.00   0.00
1210  0.12  -0.02    0.08    -0.09   -0.05     0.06    -0.12    -0.09       -0.11
0.22_*  0.95_*  0.01   0.00
1213  0.00  -0.04    0.06     0.01   -0.03     0.04     0.08     0.00        0.00
0.15    0.97_*  0.00   0.00
1233  0.00   0.01    0.01    -0.01    0.00    -0.02     0.00     0.00       -0.01       -
0.03    1.01_*  0.00   0.01
1291  0.02   0.08   -0.06    -0.05   -0.02    -0.01     0.00     0.02        0.09
0.17    0.97_*  0.00   0.00
1326  0.02   0.05    0.05    -0.07   -0.04     0.08     0.00     0.08        0.00
0.16    0.99_*  0.00   0.00
1341 -0.03  -0.05    0.06     0.05   -0.03     0.04     0.08    -0.01       -0.01
0.16    0.97_*  0.00   0.00
1342  0.06  -0.06   -0.07    -0.03   -0.02     0.05     0.00     0.01        0.10
0.18    0.96_*  0.00   0.00
1354  0.06  -0.07   -0.07     0.02   -0.03     0.17    -0.11    -0.10       -0.11
0.24_*  0.97_*  0.01   0.01
1369  0.09  -0.09   -0.08    -0.03   -0.03    -0.07     0.00     0.01        0.12
0.23_*  0.95_*  0.01   0.00
1408  0.00   0.05    0.07    -0.05   -0.03     0.04     0.08     0.01        0.00
0.16    0.98_*  0.00   0.00
1417  0.07  -0.09    0.08    -0.03   -0.06    -0.07     0.00     0.11        0.00
0.22_*  0.95_*  0.01   0.00
1445  0.01  -0.06   -0.06     0.02   -0.02     0.09     0.08     0.00       -0.01
0.17    0.98_*  0.00   0.00
1446  0.02  -0.15    0.10     0.06   -0.05    -0.09     0.00    -0.02        0.14
0.30_*  0.92_*  0.01   0.00
1474  0.15  -0.06   -0.08    -0.05   -0.04     0.00    -0.12    -0.10       -0.11
0.20_*  0.95_*  0.00   0.00
1511  0.02  -0.14    0.10     0.04   -0.08     0.00     0.00     0.12        0.00
0.27_*  0.91_*  0.01   0.00
 [ reached getOption("max.print") -- omitted 36 rows ]
            dfb.1_       dfb.age       dfb.sxml        dfb.bmi       dfb.smkr       dfb.chld
     dfb.rgnn   dfb.rgnsths
28   -0.0003336340  0.0009349667  0.0005932844  5.893115e-05  0.0002252222 -1.901013e-03
 7.405103e-05 -3.323108e-05
86    0.1310672539 -0.1394288831 -0.0933135913  3.599110e-02 -0.0410596342 -7.271590e-02
-1.353347e-01 -1.423543e-01
115   0.1230193211 -0.0186831459  0.0839946033 -9.092679e-02 -0.0521733970  6.174413e-02
-1.173792e-01 -9.007813e-02
118   0.0011488106 -0.0445191038  0.0633698660  6.616635e-03 -0.0336841022  4.330643e-02
 8.390649e-02 -1.961755e-03
134   0.0039650622  0.0115374405  0.0073020807 -9.096200e-03  0.0023090148 -2.200628e-02
 9.887596e-04  2.274751e-03
179   0.2172151085 -0.0847521125 -0.0995452787 -1.693909e-01 -0.0630643465 -8.280852e-02
 1.102742e-03  1.750918e-01
197   0.0185858397  0.0753037572 -0.0588788805 -4.819425e-02 -0.0210690940 -1.050454e-02
-5.022551e-04  1.597615e-02
200  -0.0219244237  0.0523820711  0.0519450070 -3.154972e-03 -0.0241259632 -4.709042e-02
 6.786360e-02  1.456523e-03
```

```
234   0.0164015118  0.0528153916  0.0521302467 -6.941299e-02 -0.0371958643  7.722154e-02
-1.360881e-03  8.125718e-02
236   0.0577951798 -0.0496174542  0.0694442582  8.978722e-03 -0.0429558563 -1.537199e-03
-1.027384e-01 -1.042902e-01
246  -0.0286525707 -0.0500399099  0.0627555365  4.511212e-02 -0.0331621903  4.421560e-02
 8.490462e-02 -1.246514e-02
247   0.0565528325 -0.0587585241 -0.0686335869 -2.671773e-02 -0.0249799691  5.224693e-02
-3.191687e-03  8.614314e-03
259   0.0596935060 -0.0739625456 -0.0702814739  1.750688e-02 -0.0332635201  1.683616e-01
-1.086588e-01 -1.022677e-01
272   0.0913622033 -0.0873429977 -0.0801271639 -2.728721e-02 -0.0278415562 -7.452567e-02
 2.495738e-04  7.444239e-03
281   0.0175014978 -0.1110891418 -0.0776006984  7.117070e-02 -0.0381995792 -5.504946e-02
-1.927529e-03  6.847573e-02
282   0.0158671838  0.0382730797  0.0720182255 -4.059538e-02 -0.0317066129 -6.400636e-02
 2.921707e-03  1.114026e-02
308   0.0028967084  0.0513552871  0.0668600684 -4.927118e-02 -0.0332679197  3.896237e-02
 8.344001e-02  1.474858e-02
315   0.0709134063 -0.0944201380  0.0810887293 -3.001733e-02 -0.0610145304 -6.622134e-02
 1.952564e-03  1.090825e-01
334   0.0006187532  0.0005936121  0.0011187718 -1.808592e-03 -0.0008414648  3.423487e-03
-7.966634e-05  1.846161e-03
337   0.0090662374 -0.0565697422 -0.0566521240  1.939017e-02 -0.0230034441  8.960672e-02
 7.628743e-02 -3.743270e-03
338   0.0219545597 -0.1478760269  0.0997808952  5.719232e-02 -0.0479554305 -8.997255e-02
 2.072068e-03 -1.867676e-02
367   0.1467305276 -0.0581596322 -0.0767703764 -5.416964e-02 -0.0383313130 -8.899894e-04
-1.190653e-01 -9.978403e-02
409   0.0213386984 -0.1400125954  0.0985667394  4.054067e-02 -0.0761715357  2.895756e-03
-1.237791e-03  1.174947e-01
413   0.0513392832  0.0426190117 -0.0497694307 -1.885641e-02 -0.0232095379 -4.709347e-02
-7.797164e-02 -7.023399e-02
418   0.0529449819 -0.0935291842 -0.0672427760  1.572048e-02 -0.0356042443 -4.942042e-02
-1.067935e-03  7.436909e-02
419   0.0395804580 -0.1343224056 -0.0876387481  3.841673e-02 -0.0347988238  6.619496e-02
 1.182617e-01 -9.559715e-03
428  -0.0019058855  0.0686438040  0.0627659888 -3.529175e-02 -0.0425129334 -5.329020e-02
 2.135687e-03  8.515197e-02
438  -0.0145161703 -0.0792247876 -0.0685862440  1.404624e-01 -0.0242423731 -5.066261e-02
-9.427940e-02 -1.287017e-01
472  -0.0681601466  0.0348607585 -0.0496125334  6.769567e-02 -0.0167313420  3.369529e-02
 6.776161e-02 -1.598684e-02
520  -0.0105069215  0.0212055770 -0.0477655315  5.312351e-02 -0.0190965548 -2.766014e-03
-7.137068e-02 -8.119200e-02
544   0.0538999584  0.0400029225 -0.0563708442 -7.118564e-02 -0.0215838168 -8.838372e-03
-2.830864e-04  2.155198e-02
599   0.0058347467 -0.0647750415  0.0489877506  4.489772e-02 -0.0327379181  1.274148e-01
-8.119605e-02 -8.753530e-02
638  -0.0891212060 -0.0135964450 -0.0636831615  1.302172e-01 -0.0192570941 -7.507935e-03
 8.332625e-02 -3.379297e-02
681   0.0557745416 -0.0708391232 -0.0757293565 -6.912474e-03 -0.0418449216  1.718174e-03
-2.559867e-03  9.355516e-02
742   0.0214107844  0.0217759662  0.0690615847  1.123623e-03 -0.0413482236  5.104058e-02
-1.008639e-01 -9.736697e-02
755   0.0365646048 -0.0560266260  0.0595818202 -1.999727e-02 -0.0317076559 -5.482718e-03
 7.927400e-02  4.588557e-03
774   0.0610018517 -0.0549785437 -0.0598238943  2.361778e-02 -0.0265431807 -2.977595e-04
-8.897379e-02 -9.126290e-02
775  -0.0026469433  0.0068066299 -0.0043778519 -1.370127e-03  0.0030210325 -1.499791e-02
 7.226987e-03  6.987492e-03
777   0.0167039188  0.0271297028 -0.0578526089 -1.907726e-02 -0.0239348080 -9.049920e-03
 8.777844e-02  7.347922e-03
796   0.1190810598 -0.0765402077  0.0839812155 -5.181082e-02 -0.0537398108  6.619896e-02
-1.226251e-01 -1.059728e-01
805   0.0103215028 -0.1252859040 -0.0884983232  6.833134e-02 -0.0336225586  6.565135e-02
 1.175971e-01 -1.751665e-02
812   0.1800398824 -0.0838377391  0.1089317920 -1.679766e-01 -0.0576918838 -8.910170e-02
 1.406102e-01  4.354027e-02
824   0.0550039262 -0.1253457571  0.0935335353 -2.044934e-02 -0.0517043063  6.631477e-02
 1.241797e-01  4.296927e-03
888   0.0508750480 -0.0621033821 -0.0658467952  4.537958e-02 -0.0282293582 -2.600872e-04
-9.685126e-02 -1.044509e-01
```

```
892  -0.0116119507 -0.0199426550  0.0484532597  1.839072e-02 -0.0225125390 -5.297901e-03
 1.119144e-04 -5.302411e-03
894   0.0013673614  0.0001054501 -0.0012103614 -2.073374e-03 -0.0008082902  4.125007e-03
-1.378761e-04  2.204383e-03
898  -0.0208882842 -0.0603017314  0.0740981965  5.929345e-02 -0.0375393389 -6.747635e-02
 1.034390e-01 -1.746466e-02
913   0.0795473091 -0.0645754305 -0.0553604186 -5.204428e-02 -0.0256012792  4.394282e-02
 8.122951e-02  1.509772e-02
917  -0.0348950517 -0.0455424588  0.0612828311  5.939608e-02 -0.0459525679  1.616689e-04
-1.058954e-03  6.508329e-02
938   0.0697661996 -0.0735687107 -0.0630683157 -1.632741e-02 -0.0217210528 -5.815193e-02
 1.084417e-04  4.378602e-03
943   0.0503047391 -0.1081109842 -0.0705276279 -1.323029e-03 -0.0301316476  1.133316e-01
 9.543090e-02  1.617057e-03
959  -0.0023133464 -0.1011628300  0.0549500017  1.123103e-01 -0.0345888152 -4.898197e-02
-8.978348e-02 -1.207238e-01
1042  0.0739770584 -0.0838101204  0.0562269557 -1.384467e-04 -0.0361565165  3.897952e-04
-8.430801e-02 -8.419462e-02
1052  0.1519965892 -0.0901515813 -0.0790795863 -5.255682e-02 -0.0400665175  6.633703e-02
-1.230072e-01 -1.018622e-01
1054  0.0181053956 -0.0639171659 -0.0519746737  3.468219e-02 -0.0157414700 -4.575340e-02
-5.659965e-04 -9.519632e-03
1083  0.0214107844  0.0217759662  0.0690615847  1.123623e-03 -0.0413482236  5.104058e-02
-1.008639e-01 -9.736697e-02
1089  0.0103215028 -0.1252859040 -0.0884983232  6.833134e-02 -0.0336225586  6.565135e-02
 1.175971e-01 -1.751665e-02
1096  0.1800398824 -0.0838377391  0.1089317920 -1.679766e-01 -0.0576918838 -8.910170e-02
 1.406102e-01  4.354027e-02
1099  0.0842356377 -0.0186645820  0.0882308035 -9.008571e-02 -0.0448725460 -7.481968e-02
 1.145418e-01  2.363298e-02
1121 -0.0003336340  0.0009349667  0.0005932844  5.893115e-05  0.0002252222 -1.901013e-03
 7.405103e-05 -3.323108e-05
1191  0.0128157983  0.0759958399  0.0527628732 -1.588471e-02 -0.0287444835 -4.540545e-02
-7.049867e-02 -6.604225e-02
1210  0.1230193211 -0.0186831459  0.0839946033 -9.092679e-02 -0.0521733970  6.174413e-02
-1.173792e-01 -9.007813e-02
1213  0.0011488106 -0.0445191038  0.0633698660  6.616635e-03 -0.0336841022  4.330643e-02
 8.390649e-02 -1.961755e-03
1233  0.0039650622  0.0115374405  0.0073020807 -9.096200e-03  0.0023090148 -2.200628e-02
 9.887596e-04  2.274751e-03
1291  0.0185858397  0.0753037572 -0.0588788805 -4.819425e-02 -0.0210690940 -1.050454e-02
-5.022551e-04  1.597615e-02
1326  0.0164015118  0.0528153916  0.0521302467 -6.941299e-02 -0.0371958643  7.722154e-02
-1.360881e-03  8.125718e-02
1341 -0.0286525707 -0.0500399099  0.0627555365  4.511212e-02 -0.0331621903  4.421560e-02
 8.490462e-02 -1.246514e-02
1342  0.0565528325 -0.0587585241 -0.0686335869 -2.671773e-02 -0.0249799691  5.224693e-02
-3.191687e-03  8.614314e-03
1354  0.0596935060 -0.0739625456 -0.0702814739  1.750688e-02 -0.0332635201  1.683616e-01
-1.086588e-01 -1.022677e-01
1369  0.0913622033 -0.0873429977 -0.0801271639 -2.728721e-02 -0.0278415562 -7.452567e-02
 2.495738e-04  7.444239e-03
1408  0.0028967084  0.0513552871  0.0668600684 -4.927118e-02 -0.0332679197  3.896237e-02
 8.344001e-02  1.474858e-02
1417  0.0709134063 -0.0944201380  0.0810887293 -3.001733e-02 -0.0610145304 -6.622134e-02
 1.952564e-03  1.090825e-01
1445  0.0090662374 -0.0565697422 -0.0566521240  1.939017e-02 -0.0230034441  8.960672e-02
 7.628743e-02 -3.743270e-03
1446  0.0219545597 -0.1478760269  0.0997808952  5.719232e-02 -0.0479554305 -8.997255e-02
 2.072068e-03 -1.867676e-02
1474  0.1467305276 -0.0581596322 -0.0767703764 -5.416964e-02 -0.0383313130 -8.899894e-04
-1.190653e-01 -9.978403e-02
1511  0.0213386984 -0.1400125954  0.0985667394  4.054067e-02 -0.0761715357  2.895756e-03
-1.237791e-03  1.174947e-01
       dfb.rgnsthw        dffit     cov.r       cook.d         hat
28   -7.822537e-04 -0.002489066 1.0122220 6.886948e-07 0.008046974
86   -1.354989e-01  0.266721383 0.9366374 7.843777e-03 0.003912836
115  -1.109645e-01  0.218638652 0.9528180 5.280947e-03 0.003461530
118  -2.076654e-03  0.153779537 0.9741103 2.619083e-03 0.002887564
134  -8.234703e-03 -0.030018768 1.0125162 1.001658e-04 0.008744527
179   1.343477e-02  0.318550465 0.9164206 1.116080e-02 0.004279520
197   9.465414e-02  0.171992357 0.9691408 3.274258e-03 0.003110050
```

```
200   4.080195e-04  0.129548360 0.9875858 1.861461e-03 0.003406099
234   2.183984e-03  0.158104920 0.9874459 2.772109e-03 0.004720701
236  -1.031729e-01  0.171594158 0.9642858 3.257423e-03 0.002761136
246  -5.298280e-03  0.162802269 0.9731648 2.935046e-03 0.003121929
247   1.002182e-01  0.182755264 0.9624127 3.694064e-03 0.002981505
259  -1.077281e-01  0.242445096 0.9684227 6.503713e-03 0.005674332
272   1.204677e-01  0.225792994 0.9471125 5.628530e-03 0.003339365
281  -6.072787e-03  0.211207610 0.9636409 4.933964e-03 0.004000406
282   9.757770e-02  0.173671418 0.9684932 3.338249e-03 0.003116355
308   1.833410e-03  0.159278989 0.9756461 2.810148e-03 0.003217590
315   3.362703e-03  0.223143697 0.9472600 5.497358e-03 0.003273473
334   3.151590e-05  0.004566122 1.0126088 2.317653e-06 0.008432727
337  -5.091316e-03  0.167736042 0.9781710 3.117176e-03 0.003804287
338   1.406062e-01  0.295204265 0.9203961 9.589873e-03 0.003869497
367  -1.137204e-01  0.204687716 0.9518700 4.628236e-03 0.003006191
409  -4.469237e-03  0.274244619 0.9124627 8.269221e-03 0.003074871
413  -7.679464e-02  0.139613917 0.9822767 2.160727e-03 0.003147416
418  -1.454571e-03  0.178660225 0.9718849 3.533982e-03 0.003573834
419  -6.166661e-03  0.249217278 0.9399794 6.850977e-03 0.003607886
428   2.400187e-03  0.161490726 0.9751320 2.888553e-03 0.003250391
438  -1.027947e-01  0.218576798 0.9754181 5.290470e-03 0.005586975
472  -8.335343e-03  0.144399465 0.9848533 2.311930e-03 0.003671675
520  -7.493393e-02  0.125519671 0.9875898 1.747521e-03 0.003227727
544   9.344926e-02  0.163633171 0.9718302 2.964657e-03 0.003039905
599  -8.344551e-02  0.190467143 0.9862318 4.021921e-03 0.006097044
638  -1.233598e-02  0.191243499 0.9745463 4.050208e-03 0.004328890
681  -1.490968e-03  0.179545471 0.9596778 3.564419e-03 0.002722811
742  -1.014343e-01  0.168855735 0.9668931 3.155201e-03 0.002846139
755   1.633331e-03  0.144833339 0.9789014 2.324453e-03 0.002994726
774  -8.937995e-02  0.149338635 0.9763637 2.470622e-03 0.002923767
775   7.215679e-03 -0.019511519 1.0122592 4.231820e-05 0.008264860
777  -1.824039e-04  0.147251515 0.9707945 2.400652e-03 0.002432025
796  -1.182666e-01  0.226458640 0.9487047 5.662754e-03 0.003446748
805  -8.747048e-03  0.248795625 0.9403183 6.828083e-03 0.003614129
812   1.604552e-02  0.313591801 0.9278903 1.083044e-02 0.004741812
824  -1.935814e-05  0.253658703 0.9374154 7.095198e-03 0.003596512
888  -9.865972e-02  0.166375015 0.9711640 3.064594e-03 0.003079759
892   6.557343e-02  0.113308369 0.9864008 1.423957e-03 0.002562357
894   1.638828e-05  0.005423207 1.0125010 3.269382e-06 0.008331354
898  -3.551096e-03  0.199625880 0.9606581 4.406468e-03 0.003388676
913   2.156203e-03  0.166765782 0.9759084 3.080522e-03 0.003522057
917  -6.017359e-03  0.163125665 0.9684732 2.945251e-03 0.002773253
938   9.392883e-02  0.178144076 0.9700336 3.512919e-03 0.003391848
943  -3.890012e-03  0.224775276 0.9625665 5.587295e-03 0.004383056
959  -9.700550e-02  0.209001328 0.9768889 4.838014e-03 0.005377010
1042 -8.359423e-02  0.158961138 0.9787851 2.799847e-03 0.003524669
1052 -1.171406e-01  0.227892818 0.9508577 5.736042e-03 0.003616349
1054  7.020118e-02  0.144267107 0.9841461 2.307536e-03 0.003569659
1083 -1.014343e-01  0.168855735 0.9668931 3.155201e-03 0.002846139
1089 -8.747048e-03  0.248795625 0.9403183 6.828083e-03 0.003614129
1096  1.604552e-02  0.313591801 0.9278903 1.083044e-02 0.004741812
1099  8.721444e-03  0.220196875 0.9526412 5.356358e-03 0.003497033
1121 -7.822537e-04 -0.002489066 1.0122220 6.886948e-07 0.008046974
1191 -7.075521e-02  0.145032959 0.9866601 2.332658e-03 0.003967941
1210 -1.109645e-01  0.218638652 0.9528180 5.280947e-03 0.003461530
1213 -2.076654e-03  0.153779537 0.9741103 2.619083e-03 0.002887564
1233 -8.234703e-03 -0.030018768 1.0125162 1.001658e-04 0.008744527
1291  9.465414e-02  0.171992357 0.9691408 3.274258e-03 0.003110050
1326  2.183984e-03  0.158104920 0.9874459 2.772109e-03 0.004720701
1341 -5.298280e-03  0.162802269 0.9731648 2.935046e-03 0.003121929
1342  1.002182e-01  0.182755264 0.9624127 3.694064e-03 0.002981505
1354 -1.077281e-01  0.242445096 0.9684227 6.503713e-03 0.005674332
1369  1.204677e-01  0.225792994 0.9471125 5.628530e-03 0.003339365
1408  1.833410e-03  0.159278989 0.9756461 2.810148e-03 0.003217590
1417  3.362703e-03  0.223143697 0.9472600 5.497358e-03 0.003273473
1445 -5.091316e-03  0.167736042 0.9781710 3.117176e-03 0.003804287
1446  1.406062e-01  0.295204265 0.9203961 9.589873e-03 0.003869497
1474 -1.137204e-01  0.204687716 0.9518700 4.628236e-03 0.003006191
1511 -4.469237e-03  0.274244619 0.9124627 8.269221e-03 0.003074871
 [ reached getOption("max.print") -- omitted 36 rows ]
```



Component + Residual Plots
Component+Residual(boxcox.cha
age
sex
female
male
bmi
smoker
no
yes
children
region
northeast
northwest
southeast
southwest

6.4 MODEL 3 DIAGNOSTICS

```
"ANOVA Table"
Analysis of Variance Table

Response: boxcox.charges
                   Df  Sum Sq Mean Sq   F value    Pr(>F)    
scale.age           1  6054.0  6054.0 2350.1902 < 2.2e-16 ***
sex                 1    21.3    21.3    8.2614  0.004088 ** 
scale.bmi           1   121.4   121.4   47.1291 8.603e-12 ***
smoker              1 12492.4 12492.4 4849.5938 < 2.2e-16 ***
scale.children      1   382.8   382.8  148.6227 < 2.2e-16 ***
region              3   126.4    42.1   16.3612 1.620e-10 ***
Residuals        2211  5695.4     2.6                        
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
[1] "Variance Homogeneity Test"
Non-constant Variance Score Test 
Variance formula: ~ fitted.values 
Chisquare = 56.54749, Df = 1, p = 5.4857e-14
[1] "Residual Normality Test"

	Shapiro-Wilk normality test

data:  residuals(model)
W = 0.81618, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.01801448      2.035987   0.408
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
      scale.age         sexmale       scale.bmi       smokeryes  scale.children regionnor
thwest regionsoutheast 
         1.0186          1.0107          1.1228          1.0165          1.0039          
1.5516          1.7101 
regionsouthwest 
         1.5584 
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
409  4.938055         8.4811e-07    0.0018828
1511 4.938055         8.4811e-07    0.0018828
179  4.859028         1.2624e-06    0.0028026
338  4.736456         2.3129e-06    0.0051346
1446 4.736456         2.3129e-06    0.0051346
812  4.543185         5.8395e-06    0.0129640
1096 4.543185         5.8395e-06    0.0129640
1930 4.543185         5.8395e-06    0.0129640
2219 4.543185         5.8395e-06    0.0129640
86   4.255599         2.1719e-05    0.0482170
[1] "Influential Points"
Potentially influential observations of
	 lm(formula = boxcox.charges ~ scale.age + sex + scale.bmi + smoker +      scale.c
hildren + region, data = trainData) :

     dfb.1_ dfb.scl.g dfb.sxml dfb.scl.b dfb.smkr dfb.scl.c dfb.rgnn dfb.rgnsths dfb.rgns
thw dffit   cov.r   cook.d hat  
28    0.00   0.00      0.00     0.00      0.00      0.00      0.00     0.00        0.00   
     0.00    1.01_*  0.00   0.01
86    0.22  -0.14     -0.09     0.04     -0.04     -0.07     -0.14    -0.14       -0.14   
     0.27_*  0.94_*  0.01   0.00
115   0.11  -0.02      0.08    -0.09     -0.05      0.06     -0.12    -0.09       -0.11   
     0.22_*  0.95_*  0.01   0.00
118  -0.02  -0.04      0.06     0.01     -0.03      0.04      0.08     0.00        0.00   
     0.15    0.97_*  0.00   0.00
134  -0.01   0.01      0.01    -0.01      0.00     -0.02      0.00     0.00       -0.01   
    -0.03    1.01_*  0.00   0.01
179   0.04  -0.08     -0.10    -0.17     -0.06     -0.08      0.00     0.18        0.01   
     0.32_*  0.92_*  0.01   0.00
197   0.02   0.08     -0.06    -0.05     -0.02     -0.01      0.00     0.02        0.09   
     0.17    0.97_*  0.00   0.00
```

200	-0.02	0.05	0.05	0.00	-0.02	-0.05	0.07	0.00	0.00	0.13	0.99_*	0.00	0.00
234	-0.02	0.05	0.05	-0.07	-0.04	0.08	0.00	0.08	0.00	0.16	0.99_*	0.00	0.00
236	0.11	-0.05	0.07	0.01	-0.04	0.00	-0.10	-0.10	-0.10	0.17	0.96_*	0.00	0.00
246	-0.01	-0.05	0.06	0.05	-0.03	0.04	0.08	-0.01	-0.01	0.16	0.97_*	0.00	0.00
247	0.03	-0.06	-0.07	-0.03	-0.02	0.05	0.00	0.01	0.10	0.18	0.96_*	0.00	0.00
259	0.17	-0.07	-0.07	0.02	-0.03	0.17	-0.11	-0.10	-0.11	0.24_*	0.97_*	0.01	0.01
272	0.04	-0.09	-0.08	-0.03	-0.03	-0.07	0.00	0.01	0.12	0.23_*	0.95_*	0.01	0.00
281	0.05	-0.11	-0.08	0.07	-0.04	-0.06	0.00	0.07	-0.01	0.21_*	0.96_*	0.00	0.00
282	-0.03	0.04	0.07	-0.04	-0.03	-0.06	0.00	0.01	0.10	0.17	0.97_*	0.00	0.00
308	-0.03	0.05	0.07	-0.05	-0.03	0.04	0.08	0.01	0.00	0.16	0.98_*	0.00	0.00
315	-0.03	-0.09	0.08	-0.03	-0.06	-0.07	0.00	0.11	0.00	0.22_*	0.95_*	0.01	0.00
334	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.01_*	0.00	0.01
337	0.03	-0.06	-0.06	0.02	-0.02	0.09	0.08	0.00	-0.01	0.17	0.98_*	0.00	0.00
338	-0.03	-0.15	0.10	0.06	-0.05	-0.09	0.00	-0.02	0.14	0.30_*	0.92_*	0.01	0.00
367	0.18	-0.06	-0.08	-0.05	-0.04	0.00	-0.12	-0.10	-0.11	0.20_*	0.95_*	0.00	0.00
409	-0.02	-0.14	0.10	0.04	-0.08	0.00	0.00	0.12	0.00	0.27_*	0.91_*	0.01	0.00
413	0.12	0.04	-0.05	-0.02	-0.02	-0.05	-0.08	-0.07	-0.08	0.14	0.98_*	0.00	0.00
418	0.04	-0.09	-0.07	0.02	-0.04	-0.05	0.00	0.07	0.00	0.18	0.97_*	0.00	0.00
419	0.05	-0.13	-0.09	0.04	-0.03	0.07	0.12	-0.01	-0.01	0.25_*	0.94_*	0.01	0.00
428	-0.02	0.07	0.06	-0.04	-0.04	-0.05	0.00	0.09	0.00	0.16	0.98_*	0.00	0.00
438	0.16	-0.08	-0.07	0.14	-0.02	-0.05	-0.09	-0.13	-0.10	0.22_*	0.98_*	0.01	0.01
472	0.03	0.03	-0.05	0.07	-0.02	0.03	0.07	-0.02	-0.01	0.14	0.98_*	0.00	0.00
520	0.12	0.02	-0.05	0.05	-0.02	0.00	-0.07	-0.08	-0.07	0.13	0.99_*	0.00	0.00
544	0.02	0.04	-0.06	-0.07	-0.02	-0.01	0.00	0.02	0.09	0.16	0.97_*	0.00	0.00
599	0.09	-0.06	0.05	0.04	-0.03	0.13	-0.08	-0.09	-0.08	0.19	0.99_*	0.00	0.01
638	0.05	-0.01	-0.06	0.13	-0.02	-0.01	0.08	-0.03	-0.01	0.19	0.97_*	0.00	0.00
681	0.04	-0.07	-0.08	-0.01	-0.04	0.00	0.00	0.09	0.00	0.18	0.96_*	0.00	0.00
742	0.10	0.02	0.07	0.00	-0.04	0.05	-0.10	-0.10	-0.10	0.17	0.97_*	0.00	0.00
755	-0.02	-0.06	0.06	-0.02	-0.03	-0.01	0.08	0.00	0.00	0.14	0.98_*	0.00	0.00
774	0.14	-0.05	-0.06	0.02	-0.03	0.00	-0.09	-0.09	-0.09	0.15	0.98_*	0.00	0.00
775	-0.01	0.01	0.00	0.00	0.00	-0.01	0.01	0.01	0.01	-0.02	1.01_*	0.00	0.01
777	0.03	0.03	-0.06	-0.02	-0.02	-0.01	0.09	0.01	0.00	0.15	0.97_*	0.00	0.00
796	0.12	-0.08	0.08	-0.05	-0.05	0.07	-0.12	-0.11	-0.12	0.23_*	0.95_*	0.01	0.00
805	0.06	-0.13	-0.09	0.07	-0.03	0.07	0.12	-0.02	-0.01	0.25_*	0.94_*	0.01	0.00
812	-0.06	-0.08	0.11	-0.17	-0.06	-0.09	0.14	0.04	0.02	0.31_*	0.93_*	0.01	0.00
824	-0.03	-0.13	0.09	-0.02	-0.05	0.07	0.12	0.00	0.00	0.25_*	0.94_*	0.01	0.00

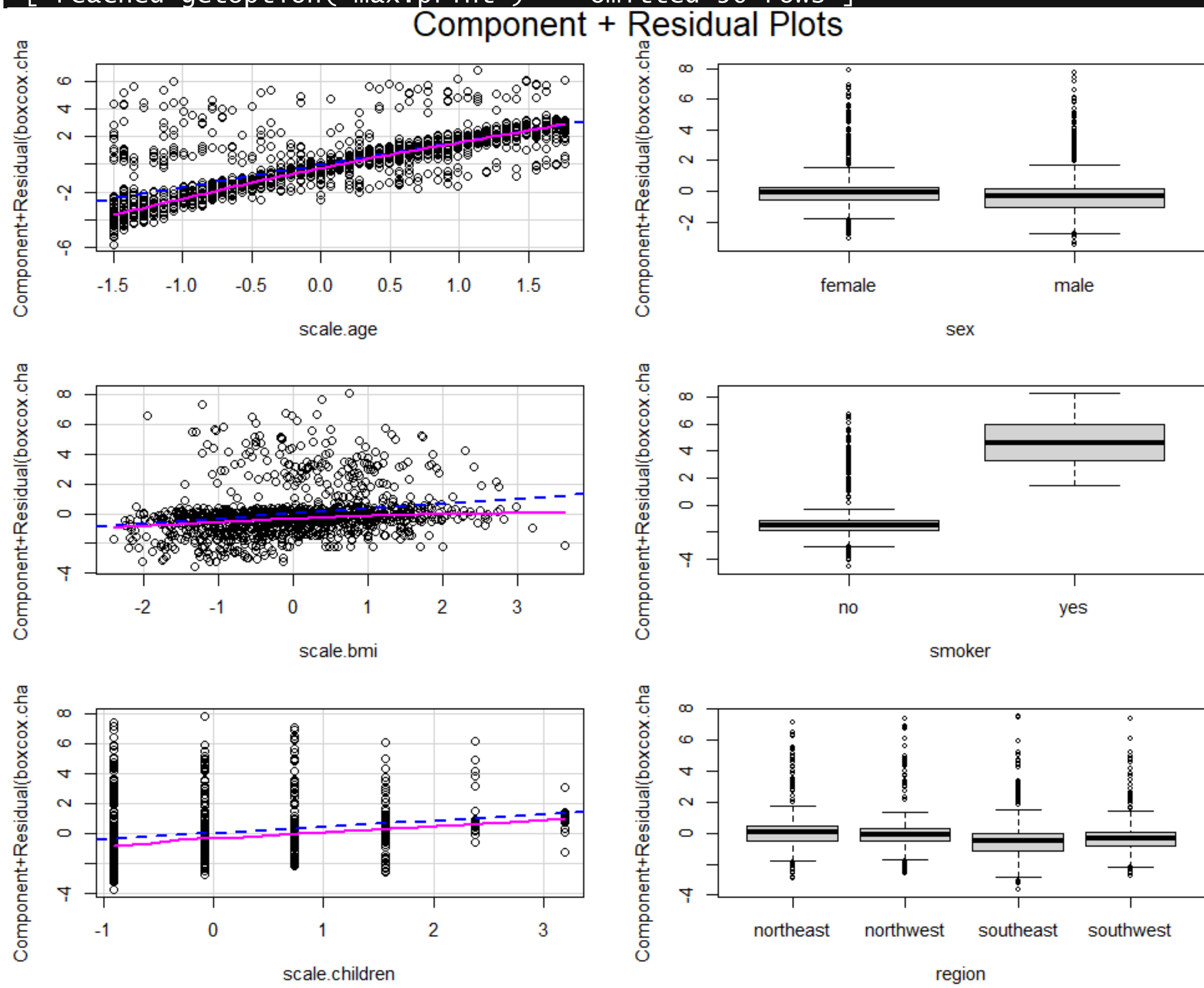
```
888   0.16  -0.06     -0.07     0.05      -0.03      0.00     -0.10    -0.10        -0.10
     0.17    0.97_*  0.00   0.00
892  -0.01  -0.02      0.05     0.02      -0.02     -0.01      0.00    -0.01         0.07
     0.11    0.99_*  0.00   0.00
894   0.00   0.00      0.00     0.00       0.00      0.00      0.00     0.00         0.00
     0.01    1.01_*  0.00   0.01
898  -0.02  -0.06      0.07     0.06      -0.04     -0.07      0.10    -0.02         0.00
     0.20_*  0.96_*  0.00   0.00
913   0.02  -0.06     -0.06    -0.05      -0.03      0.04      0.08     0.02         0.00
     0.17    0.98_*  0.00   0.00
917  -0.01  -0.05      0.06     0.06      -0.05      0.00      0.00     0.07        -0.01
     0.16    0.97_*  0.00   0.00
938   0.03  -0.07     -0.06    -0.02      -0.02     -0.06      0.00     0.00         0.09
     0.18    0.97_*  0.00   0.00
943   0.04  -0.11     -0.07     0.00      -0.03      0.11      0.10     0.00         0.00
     0.22_*  0.96_*  0.01   0.00
959   0.11  -0.10      0.05     0.11      -0.03     -0.05     -0.09    -0.12        -0.10
     0.21_*  0.98_*  0.00   0.01
1042  0.09  -0.08      0.06     0.00      -0.04      0.00     -0.08    -0.08        -0.08
     0.16    0.98_*  0.00   0.00
1052  0.18  -0.09     -0.08    -0.05      -0.04      0.07     -0.12    -0.10        -0.12
     0.23_*  0.95_*  0.01   0.00
1054  0.03  -0.06     -0.05     0.03      -0.02     -0.05      0.00    -0.01         0.07
     0.14    0.98_*  0.00   0.00
1083  0.10   0.02      0.07     0.00      -0.04      0.05     -0.10    -0.10        -0.10
     0.17    0.97_*  0.00   0.00
1089  0.06  -0.13     -0.09     0.07      -0.03      0.07      0.12    -0.02        -0.01
     0.25_*  0.94_*  0.01   0.00
1096 -0.06  -0.08      0.11    -0.17      -0.06     -0.09      0.14     0.04         0.02
     0.31_*  0.93_*  0.01   0.00
1099 -0.04  -0.02      0.09    -0.09      -0.04     -0.07      0.11     0.02         0.01
     0.22_*  0.95_*  0.01   0.00
1121  0.00   0.00      0.00     0.00       0.00      0.00      0.00     0.00         0.00
     0.00    1.01_*  0.00   0.01
1191  0.07   0.08      0.05    -0.02      -0.03     -0.05     -0.07    -0.07        -0.07
     0.15    0.99_*  0.00   0.00
1210  0.11  -0.02      0.08    -0.09      -0.05      0.06     -0.12    -0.09        -0.11
     0.22_*  0.95_*  0.01   0.00
1213 -0.02  -0.04      0.06     0.01      -0.03      0.04      0.08     0.00         0.00
     0.15    0.97_*  0.00   0.00
1233 -0.01   0.01      0.01    -0.01       0.00     -0.02      0.00     0.00        -0.01
    -0.03    1.01_*  0.00   0.01
1291  0.02   0.08     -0.06    -0.05      -0.02     -0.01      0.00     0.02         0.09
     0.17    0.97_*  0.00   0.00
1326 -0.02   0.05      0.05    -0.07      -0.04      0.08      0.00     0.08         0.00
     0.16    0.99_*  0.00   0.00
1341 -0.01  -0.05      0.06     0.05      -0.03      0.04      0.08    -0.01        -0.01
     0.16    0.97_*  0.00   0.00
1342  0.03  -0.06     -0.07    -0.03      -0.02      0.05      0.00     0.01         0.10
     0.18    0.96_*  0.00   0.00
1354  0.17  -0.07     -0.07     0.02      -0.03      0.17     -0.11    -0.10        -0.11
     0.24_*  0.97_*  0.01   0.01
1369  0.04  -0.09     -0.08    -0.03      -0.03     -0.07      0.00     0.01         0.12
     0.23_*  0.95_*  0.01   0.00
1408 -0.03   0.05      0.07    -0.05      -0.03      0.04      0.08     0.01         0.00
     0.16    0.98_*  0.00   0.00
1417 -0.03  -0.09      0.08    -0.03      -0.06     -0.07      0.00     0.11         0.00
     0.22_*  0.95_*  0.01   0.00
1445  0.03  -0.06     -0.06     0.02      -0.02      0.09      0.08     0.00        -0.01
     0.17    0.98_*  0.00   0.00
1446 -0.03  -0.15      0.10     0.06      -0.05     -0.09      0.00    -0.02         0.14
     0.30_*  0.92_*  0.01   0.00
1474  0.18  -0.06     -0.08    -0.05      -0.04      0.00     -0.12    -0.10        -0.11
     0.20_*  0.95_*  0.00   0.00
1511 -0.02  -0.14      0.10     0.04      -0.08      0.00      0.00     0.12         0.00
     0.27_*  0.91_*  0.01   0.00
 [ reached getOption("max.print") -- omitted 36 rows ]
            dfb.1_      dfb.scl.g      dfb.sxml      dfb.scl.b      dfb.smkr      dfb.scl.c
     dfb.rgnn   dfb.rgnsths
28   -0.0003296729  0.0009349667  0.0005932844  5.893115e-05  0.0002252222 -1.901013e-03
 7.405103e-05 -3.323108e-05
```

```
86    0.2165912312 -0.1394288831 -0.0933135913  3.599110e-02 -0.0410596342 -7.271590e-02
-1.353347e-01 -1.423543e-01
115   0.1074301415 -0.0186831459  0.0839946033 -9.092679e-02 -0.0521733970  6.174413e-02
-1.173792e-01 -9.007813e-02
118  -0.0182775249 -0.0445191038  0.0633698660  6.616635e-03 -0.0336841022  4.330643e-02
 8.390649e-02 -1.961755e-03
134  -0.0051275479  0.0115374405  0.0073020807 -9.096200e-03  0.0023090148 -2.200628e-02
 9.887596e-04  2.274751e-03
179   0.0358431601 -0.0847521125 -0.0995452787 -1.693909e-01 -0.0630643465 -8.280852e-02
 1.102742e-03  1.750918e-01
197   0.0239497135  0.0753037572 -0.0588788805 -4.819425e-02 -0.0210690940 -1.050454e-02
-5.022551e-04  1.597615e-02
200  -0.0182595443  0.0523820711  0.0519450070 -3.154972e-03 -0.0241259632 -4.709042e-02
 6.786360e-02  1.456523e-03
234  -0.0217388335  0.0528153916  0.0521302467 -6.941299e-02 -0.0371958643  7.722154e-02
-1.360881e-03  8.125718e-02
236   0.1057903284 -0.0496174542  0.0694442582  8.978722e-03 -0.0429558563 -1.537199e-03
-1.027384e-01 -1.042902e-01
246  -0.0134203680 -0.0500399099  0.0627555365  4.511212e-02 -0.0331621903  4.421560e-02
 8.490462e-02 -1.246514e-02
247   0.0329968642 -0.0587585241 -0.0686335869 -2.671773e-02 -0.0249799691  5.224693e-02
-3.191687e-03  8.614314e-03
259   0.1657194995 -0.0739625456 -0.0702814739  1.750688e-02 -0.0332635201  1.683616e-01
-1.086588e-01 -1.022677e-01
272   0.0369013330 -0.0873429977 -0.0801271639 -2.728721e-02 -0.0278415562 -7.452567e-02
 2.495738e-04  7.444239e-03
281   0.0504532708 -0.1110891418 -0.0776006984  7.117070e-02 -0.0381995792 -5.504946e-02
-1.927529e-03  6.847573e-02
282  -0.0299692086  0.0382730797  0.0720182255 -4.059538e-02 -0.0317066129 -6.400636e-02
 2.921707e-03  1.114026e-02
308  -0.0270698272  0.0513552871  0.0668600684 -4.927118e-02 -0.0332679197  3.896237e-02
 8.344001e-02  1.474858e-02
315  -0.0255728412 -0.0944201380  0.0810887293 -3.001733e-02 -0.0610145304 -6.622134e-02
 1.952564e-03  1.090825e-01
334  -0.0004693515  0.0005936121  0.0011187718 -1.808592e-03 -0.0008414648  3.423487e-03
-7.966634e-05  1.846161e-03
337   0.0334762157 -0.0565697422 -0.0566521240  1.939017e-02 -0.0230034441  8.960672e-02
 7.628743e-02 -3.743270e-03
338  -0.0261418278 -0.1478760269  0.0997808952  5.719232e-02 -0.0479554305 -8.997255e-02
 2.072068e-03 -1.867676e-02
367   0.1782644990 -0.0581596322 -0.0767703764 -5.416964e-02 -0.0383313130 -8.899894e-04
-1.190653e-01 -9.978403e-02
409  -0.0200387232 -0.1400125954  0.0985667394  4.054067e-02 -0.0761715357  2.895756e-03
-1.237791e-03  1.174947e-01
413   0.1188649065  0.0426190117 -0.0497694307 -1.885641e-02 -0.0232095379 -4.709347e-02
-7.797164e-02 -7.023399e-02
418   0.0386665532 -0.0935291842 -0.0672427760  1.572048e-02 -0.0356042443 -4.942042e-02
-1.067935e-03  7.436909e-02
419   0.0518016174 -0.1343224056 -0.0876387481  3.841673e-02 -0.0347988238  6.619496e-02
 1.182617e-01 -9.559715e-03
428  -0.0228496372  0.0686438040  0.0627659888 -3.529175e-02 -0.0425129334 -5.329020e-02
 2.135687e-03  8.515197e-02
438   0.1636421132 -0.0792247876 -0.0685862440  1.404624e-01 -0.0242423731 -5.066261e-02
-9.427940e-02 -1.287017e-01
472   0.0337949413  0.0348607585 -0.0496125334  6.769567e-02 -0.0167313420  3.369529e-02
 6.776161e-02 -1.598684e-02
520   0.1158207013  0.0212055770 -0.0477655315  5.312351e-02 -0.0190965548 -2.766014e-03
-7.137068e-02 -8.119200e-02
544   0.0203350301  0.0400029225 -0.0563708442 -7.118564e-02 -0.0215838168 -8.838372e-03
-2.830864e-04  2.155198e-02
599   0.0869232053 -0.0647750415  0.0489877506  4.489772e-02 -0.0327379181  1.274148e-01
-8.119605e-02 -8.753530e-02
638   0.0476511974 -0.0135964450 -0.0636831615  1.302172e-01 -0.0192570941 -7.507935e-03
 8.332625e-02 -3.379297e-02
681   0.0416164924 -0.0708391232 -0.0757293565 -6.912474e-03 -0.0418449216  1.718174e-03
-2.559867e-03  9.355516e-02
742   0.1012674110  0.0217759662  0.0690615847  1.123623e-03 -0.0413482236  5.104058e-02
-1.008639e-01 -9.736697e-02
755  -0.0209473850 -0.0560266260  0.0595818202 -1.999727e-02 -0.0317076559 -5.482718e-03
 7.927400e-02  4.588557e-03
774   0.1405681863 -0.0549785437 -0.0598238943  2.361778e-02 -0.0265431807 -2.977595e-04
-8.897379e-02 -9.126290e-02
```

```
775  -0.0073929414  0.0068066299 -0.0043778519 -1.370127e-03  0.0030210325 -1.499791e-02
 7.226987e-03  6.987492e-03
777   0.0278785287  0.0271297028 -0.0578526089 -1.907726e-02 -0.0239348080 -9.049920e-03
 8.777844e-02  7.347922e-03
796   0.1182064949 -0.0765402077  0.0839812155 -5.181082e-02 -0.0537398108  6.619896e-02
-1.226251e-01 -1.059728e-01
805   0.0555036360 -0.1252859040 -0.0884983232  6.833134e-02 -0.0336225586  6.565135e-02
 1.175971e-01 -1.751665e-02
812  -0.0554796112 -0.0838377391  0.1089317920 -1.679766e-01 -0.0576918838 -8.910170e-02
 1.406102e-01  4.354027e-02
824  -0.0299573545 -0.1253457571  0.0935335353 -2.044934e-02 -0.0517043063  6.631477e-02
 1.241797e-01  4.296927e-03
888   0.1552560148 -0.0621033821 -0.0658467952  4.537958e-02 -0.0282293582 -2.600872e-04
-9.685126e-02 -1.044509e-01
892  -0.0136637010 -0.0199426550  0.0484532597  1.839072e-02 -0.0225125390 -5.297901e-03
 1.119144e-04 -5.302411e-03
894   0.0005100671  0.0001054501 -0.0012103614 -2.073374e-03 -0.0008082902  4.125007e-03
-1.378761e-04  2.204383e-03
898  -0.0171198869 -0.0603017314  0.0740981965  5.929345e-02 -0.0375393389 -6.747635e-02
 1.034390e-01 -1.746466e-02
913   0.0242236116 -0.0645754305 -0.0553604186 -5.204428e-02 -0.0256012792  4.394282e-02
 8.122951e-02  1.509772e-02
917  -0.0087909507 -0.0455424588  0.0612828311  5.939608e-02 -0.0459525679  1.616689e-04
-1.058954e-03  6.508329e-02
938   0.0296514320 -0.0735687107 -0.0630683157 -1.632741e-02 -0.0217210528 -5.815193e-02
 1.084417e-04  4.378602e-03
943   0.0390759788 -0.1081109842 -0.0705276279 -1.323029e-03 -0.0301316476  1.133316e-01
 9.543090e-02  1.617057e-03
959   0.1063725254 -0.1011628300  0.0549500017  1.123103e-01 -0.0345888152 -4.898197e-02
-8.978348e-02 -1.207238e-01
1042  0.0865226682 -0.0838101204  0.0562269557 -1.384467e-04 -0.0361565165  3.897952e-04
-8.430801e-02 -8.419462e-02
1052  0.1830455778 -0.0901515813 -0.0790795863 -5.255682e-02 -0.0400665175  6.633703e-02
-1.230072e-01 -1.018622e-01
1054  0.0298395733 -0.0639171659 -0.0519746737  3.468219e-02 -0.0157414700 -4.575340e-02
-5.659965e-04 -9.519632e-03
1083  0.1012674110  0.0217759662  0.0690615847  1.123623e-03 -0.0413482236  5.104058e-02
-1.008639e-01 -9.736697e-02
1089  0.0555036360 -0.1252859040 -0.0884983232  6.833134e-02 -0.0336225586  6.565135e-02
 1.175971e-01 -1.751665e-02
1096 -0.0554796112 -0.0838377391  0.1089317920 -1.679766e-01 -0.0576918838 -8.910170e-02
 1.406102e-01  4.354027e-02
1099 -0.0399799784 -0.0186645820  0.0882308035 -9.008571e-02 -0.0448725460 -7.481968e-02
 1.145418e-01  2.363298e-02
1121 -0.0003296729  0.0009349667  0.0005932844  5.893115e-05  0.0002252222 -1.901013e-03
 7.405103e-05 -3.323108e-05
1191  0.0688959204  0.0759958399  0.0527628732 -1.588471e-02 -0.0287444835 -4.540545e-02
-7.049867e-02 -6.604225e-02
1210  0.1074301415 -0.0186831459  0.0839946033 -9.092679e-02 -0.0521733970  6.174413e-02
-1.173792e-01 -9.007813e-02
1213 -0.0182775249 -0.0445191038  0.0633698660  6.616635e-03 -0.0336841022  4.330643e-02
 8.390649e-02 -1.961755e-03
1233 -0.0051275479  0.0115374405  0.0073020807 -9.096200e-03  0.0023090148 -2.200628e-02
 9.887596e-04  2.274751e-03
1291  0.0239497135  0.0753037572 -0.0588788805 -4.819425e-02 -0.0210690940 -1.050454e-02
-5.022551e-04  1.597615e-02
1326 -0.0217388335  0.0528153916  0.0521302467 -6.941299e-02 -0.0371958643  7.722154e-02
-1.360881e-03  8.125718e-02
1341 -0.0134203680 -0.0500399099  0.0627555365  4.511212e-02 -0.0331621903  4.421560e-02
 8.490462e-02 -1.246514e-02
1342  0.0329968642 -0.0587585241 -0.0686335869 -2.671773e-02 -0.0249799691  5.224693e-02
-3.191687e-03  8.614314e-03
1354  0.1657194995 -0.0739625456 -0.0702814739  1.750688e-02 -0.0332635201  1.683616e-01
-1.086588e-01 -1.022677e-01
1369  0.0369013330 -0.0873429977 -0.0801271639 -2.728721e-02 -0.0278415562 -7.452567e-02
 2.495738e-04  7.444239e-03
1408 -0.0270698272  0.0513552871  0.0668600684 -4.927118e-02 -0.0332679197  3.896237e-02
 8.344001e-02  1.474858e-02
1417 -0.0255728412 -0.0944201380  0.0810887293 -3.001733e-02 -0.0610145304 -6.622134e-02
 1.952564e-03  1.090825e-01
1445  0.0334762157 -0.0565697422 -0.0566521240  1.939017e-02 -0.0230034441  8.960672e-02
 7.628743e-02 -3.743270e-03
```

```
1446 -0.0261418278 -0.1478760269  0.0997808952  5.719232e-02 -0.0479554305 -8.997255e-02
 2.072068e-03 -1.867676e-02
1474  0.1782644990 -0.0581596322 -0.0767703764 -5.416964e-02 -0.0383313130 -8.899894e-04
-1.190653e-01 -9.978403e-02
1511 -0.0200387232 -0.1400125954  0.0985667394  4.054067e-02 -0.0761715357  2.895756e-03
-1.237791e-03  1.174947e-01
      dfb.rgnsthw        dffit     cov.r       cook.d         hat
28   -7.822537e-04 -0.002489066 1.0122220 6.886948e-07 0.008046974
86   -1.354989e-01  0.266721383 0.9366374 7.843777e-03 0.003912836
115  -1.109645e-01  0.218638652 0.9528180 5.280947e-03 0.003461530
118  -2.076654e-03  0.153779537 0.9741103 2.619083e-03 0.002887564
134  -8.234703e-03 -0.030018768 1.0125162 1.001658e-04 0.008744527
179   1.343477e-02  0.318550465 0.9164206 1.116080e-02 0.004279520
197   9.465414e-02  0.171992357 0.9691408 3.274258e-03 0.003110050
200   4.080195e-04  0.129548360 0.9875858 1.861461e-03 0.003406099
234   2.183984e-03  0.158104920 0.9874459 2.772109e-03 0.004720701
236  -1.031729e-01  0.171594158 0.9642858 3.257423e-03 0.002761136
246  -5.298280e-03  0.162802269 0.9731648 2.935046e-03 0.003121929
247   1.002182e-01  0.182755264 0.9624127 3.694064e-03 0.002981505
259  -1.077281e-01  0.242445096 0.9684227 6.503713e-03 0.005674332
272   1.204677e-01  0.225792994 0.9471125 5.628530e-03 0.003339365
281  -6.072787e-03  0.211207610 0.9636409 4.933964e-03 0.004000406
282   9.757770e-02  0.173671418 0.9684932 3.338249e-03 0.003116355
308   1.833410e-03  0.159278989 0.9756461 2.810148e-03 0.003217590
315   3.362703e-03  0.223143697 0.9472600 5.497358e-03 0.003273473
334   3.151590e-05  0.004566122 1.0126088 2.317653e-06 0.008432727
337  -5.091316e-03  0.167736042 0.9781710 3.117176e-03 0.003804287
338   1.406062e-01  0.295204265 0.9203961 9.589873e-03 0.003869497
367  -1.137204e-01  0.204687716 0.9518700 4.628236e-03 0.003006191
409  -4.469237e-03  0.274244619 0.9124627 8.269221e-03 0.003074871
413  -7.679464e-02  0.139613917 0.9822767 2.160727e-03 0.003147416
418  -1.454571e-03  0.178660225 0.9718849 3.533982e-03 0.003573834
419  -6.166661e-03  0.249217278 0.9399794 6.850977e-03 0.003607886
428   2.400187e-03  0.161490726 0.9751320 2.888553e-03 0.003250391
438  -1.027947e-01  0.218576798 0.9754181 5.290470e-03 0.005586975
472  -8.335343e-03  0.144399465 0.9848533 2.311930e-03 0.003671675
520  -7.493393e-02  0.125519671 0.9875898 1.747521e-03 0.003227727
544   9.344926e-02  0.163633171 0.9718302 2.964657e-03 0.003039905
599  -8.344551e-02  0.190467143 0.9862318 4.021921e-03 0.006097044
638  -1.233598e-02  0.191243499 0.9745463 4.050208e-03 0.004328890
681  -1.490968e-03  0.179545471 0.9596778 3.564419e-03 0.002722811
742  -1.014343e-01  0.168855735 0.9668931 3.155201e-03 0.002846139
755   1.633331e-03  0.144833339 0.9789014 2.324453e-03 0.002994726
774  -8.937995e-02  0.149338635 0.9763637 2.470622e-03 0.002923767
775   7.215679e-03 -0.019511519 1.0122592 4.231820e-05 0.008264860
777  -1.824039e-04  0.147251515 0.9707945 2.400652e-03 0.002432025
796  -1.182666e-01  0.226458640 0.9487047 5.662754e-03 0.003446748
805  -8.747048e-03  0.248795625 0.9403183 6.828083e-03 0.003614129
812   1.604552e-02  0.313591801 0.9278903 1.083044e-02 0.004741812
824  -1.935814e-05  0.253658703 0.9374154 7.095198e-03 0.003596512
888  -9.865972e-02  0.166375015 0.9711640 3.064594e-03 0.003079759
892   6.557343e-02  0.113308369 0.9864008 1.423957e-03 0.002562357
894   1.638828e-05  0.005423207 1.0125010 3.269382e-06 0.008331354
898  -3.551096e-03  0.199625880 0.9606581 4.406468e-03 0.003388676
913   2.156203e-03  0.166765782 0.9759084 3.080522e-03 0.003522057
917  -6.017359e-03  0.163125665 0.9684732 2.945251e-03 0.002773253
938   9.392883e-02  0.178144076 0.9700336 3.512919e-03 0.003391848
943  -3.890012e-03  0.224775276 0.9625665 5.587295e-03 0.004383056
959  -9.700550e-02  0.209001328 0.9768889 4.838014e-03 0.005377010
1042 -8.359423e-02  0.158961138 0.9787851 2.799847e-03 0.003524669
1052 -1.171406e-01  0.227892818 0.9508577 5.736042e-03 0.003616349
1054  7.020118e-02  0.144267107 0.9841461 2.307536e-03 0.003569659
1083 -1.014343e-01  0.168855735 0.9668931 3.155201e-03 0.002846139
1089 -8.747048e-03  0.248795625 0.9403183 6.828083e-03 0.003614129
1096  1.604552e-02  0.313591801 0.9278903 1.083044e-02 0.004741812
1099  8.721444e-03  0.220196875 0.9526412 5.356358e-03 0.003497033
1121 -7.822537e-04 -0.002489066 1.0122220 6.886948e-07 0.008046974
1191 -7.075521e-02  0.145032959 0.9866601 2.332658e-03 0.003967941
1210 -1.109645e-01  0.218638652 0.9528180 5.280947e-03 0.003461530
1213 -2.076654e-03  0.153779537 0.9741103 2.619083e-03 0.002887564
1233 -8.234703e-03 -0.030018768 1.0125162 1.001658e-04 0.008744527
1291  9.465414e-02  0.171992357 0.9691408 3.274258e-03 0.003110050
```

```
1326  2.183984e-03  0.158104920 0.9874459 2.772109e-03 0.004720701
1341 -5.298280e-03  0.162802269 0.9731648 2.935046e-03 0.003121929
1342  1.002182e-01  0.182755264 0.9624127 3.694064e-03 0.002981505
1354 -1.077281e-01  0.242445096 0.9684227 6.503713e-03 0.005674332
1369  1.204677e-01  0.225792994 0.9471125 5.628530e-03 0.003339365
1408  1.833410e-03  0.159278989 0.9756461 2.810148e-03 0.003217590
1417  3.362703e-03  0.223143697 0.9472600 5.497358e-03 0.003273473
1445 -5.091316e-03  0.167736042 0.9781710 3.117176e-03 0.003804287
1446  1.406062e-01  0.295204265 0.9203961 9.589873e-03 0.003869497
1474 -1.137204e-01  0.204687716 0.9518700 4.628236e-03 0.003006191
1511 -4.469237e-03  0.274244619 0.9124627 8.269221e-03 0.003074871
 [ reached getOption("max.print") -- omitted 36 rows ]
```



6.5 Model 4 Diagnostics

```
"Variance Homogeneity Test"
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 180.911, Df = 1, p = < 2.22e-16
[1] "Residual Normality Test"

        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.73486, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.01647802      2.032565   0.486
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
                         scale.age                          smokeryes
 scale.children
                            1.3891                             1.0244
         1.2280
                         scale.bmi                    regionnorthwest
regionsoutheast
                            1.1271                             1.5541
         1.7226
                   regionsouthwest                            sexmale                 scal
e.age:smokeryes
                            1.5591                             1.0128
         1.3410
          scale.age:scale.children           smokeryes:scale.children scale.age:smokeryes
:scale.children
                            1.3446                             1.2380
         1.3220
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
338  5.689466         1.4437e-08   3.2051e-05
1446 5.689466         1.4437e-08   3.2051e-05
179  5.609488         2.2837e-08   5.0699e-05
409  5.530958         3.5619e-08   7.9074e-05
1511 5.530958         3.5619e-08   7.9074e-05
812  5.305978         1.2328e-07   2.7368e-04
1096 5.305978         1.2328e-07   2.7368e-04
1930 5.305978         1.2328e-07   2.7368e-04
2219 5.305978         1.2328e-07   2.7368e-04
86   5.180013         2.4201e-07   5.3727e-04
[1] "Influential Points"
Potentially influential observations of
         lm(formula = boxcox.charges ~ scale.age * smoker * scale.children +      scale.bm
i + region + sex, data = trainData) :

     dfb.1_ dfb.scl.g dfb.smkr dfb.scl.c dfb.scl.b dfb.rgnn dfb.rgnsths dfb.rgnsthw dfb.s
xml dfb.sc.: dfb.s.:. dfb.sm:.
9     0.00   0.00     -0.03     0.00      0.02      0.00    -0.02        0.00        0.02
    -0.04    0.00      0.03
28   -0.01   0.06      0.01    -0.07      0.00      0.00    -0.01       -0.03        0.02
    -0.03    0.12      0.03
34   -0.01   0.00      0.04     0.00      0.03      0.00    -0.01        0.02        0.01
     0.04    0.00     -0.04
49   -0.01   0.00      0.02     0.00      0.01      0.01     0.00        0.00        0.01
     0.03    0.00      0.01
74    0.00   0.00      0.02     0.00      0.00      0.01     0.00        0.00       -0.01
     0.02    0.00     -0.02
81    0.00   0.00      0.03     0.00      0.00      0.00     0.00        0.02       -0.02
     0.06    0.00      0.02
86    0.27  -0.15     -0.04    -0.10      0.05     -0.17    -0.18       -0.16       -0.11
     0.07    0.13      0.05
87    0.00   0.00     -0.02     0.00      0.01      0.00    -0.02        0.00        0.02
    -0.05    0.00     -0.04
92   -0.01   0.00      0.04     0.00      0.00      0.00     0.01        0.00        0.01
     0.05    0.00     -0.04
```

115	0.11	-0.03	-0.06	0.07	-0.10	-0.12	-0.09	-0.12	0.09	0.02	-0.03	-0.03
118	-0.02	-0.07	-0.04	0.05	0.01	0.09	0.00	0.00	0.07	0.03	-0.06	-0.03
134	-0.02	0.08	0.01	-0.09	-0.04	0.00	0.00	-0.03	0.03	-0.04	0.14	0.04
179	0.04	-0.08	-0.07	-0.11	-0.19	0.00	0.19	0.02	-0.11	0.04	0.08	0.05
197	0.02	0.09	-0.02	-0.01	-0.05	0.00	0.02	0.10	-0.06	-0.03	0.02	0.00
199	-0.03	0.00	-0.05	0.00	0.00	0.03	0.03	0.02	0.02	-0.06	0.00	0.05
204	-0.02	-0.01	-0.06	0.00	0.05	0.04	0.03	0.04	-0.03	0.11	0.00	-0.05
209	-0.01	0.00	0.03	0.00	0.01	0.02	0.00	0.00	0.01	0.02	0.00	-0.03
214	-0.02	0.00	-0.03	0.00	0.02	0.03	0.02	0.03	-0.01	-0.05	0.00	-0.07
234	-0.02	0.09	-0.04	0.09	-0.08	0.00	0.08	0.00	0.06	-0.04	0.10	-0.04
236	0.12	-0.06	-0.05	0.00	0.01	-0.12	-0.12	-0.12	0.08	0.03	-0.01	0.00
243	-0.04	0.00	-0.04	0.00	0.03	0.04	0.02	0.03	0.03	-0.06	0.00	-0.07
246	-0.01	-0.07	-0.03	0.05	0.05	0.09	-0.01	0.00	0.06	0.03	-0.06	-0.03
247	0.04	-0.09	-0.03	0.06	-0.03	0.00	0.01	0.11	-0.08	0.05	-0.08	-0.03
259	0.16	-0.13	-0.03	0.18	0.02	-0.10	-0.09	-0.10	-0.07	0.06	-0.19	-0.07
272	0.05	-0.09	-0.03	-0.10	-0.03	0.00	0.00	0.14	-0.10	0.05	0.08	0.04
280	0.08	-0.02	-0.03	-0.04	0.05	-0.07	-0.08	-0.07	0.04	0.01	0.02	0.02
281	0.06	-0.12	-0.04	-0.08	0.09	-0.01	0.08	-0.01	-0.10	0.05	0.08	0.04
282	-0.03	0.03	-0.03	-0.07	-0.04	0.00	0.01	0.10	0.08	-0.01	-0.02	0.03
308	-0.03	0.08	-0.03	0.05	-0.05	0.09	0.01	0.00	0.07	-0.04	0.06	-0.03
315	-0.03	-0.10	-0.07	-0.09	-0.03	0.00	0.12	0.00	0.10	0.04	0.09	0.04
327	0.01	0.03	0.01	-0.05	0.01	0.00	-0.01	-0.02	-0.01	-0.02	0.06	0.02
331	0.00	0.00	-0.05	0.00	0.00	-0.03	0.00	0.00	0.02	-0.05	0.00	0.04
332	-0.01	0.00	0.04	0.00	0.00	0.00	0.02	0.00	0.01	0.05	0.00	-0.04
335	0.12	-0.07	-0.02	-0.05	0.01	-0.08	-0.08	-0.08	-0.05	0.03	0.06	0.02
337	0.03	-0.09	-0.02	0.10	0.02	0.08	0.00	0.00	-0.06	0.04	-0.11	-0.04
338	-0.03	-0.16	-0.05	-0.12	0.07	0.00	-0.04	0.17	0.12	0.07	0.14	0.05
343	-0.01	-0.02	0.01	-0.05	-0.03	0.00	-0.01	0.00	0.01	0.01	-0.05	0.02
367	0.20	-0.08	-0.04	0.00	-0.06	-0.13	-0.11	-0.13	-0.09	0.04	-0.01	0.01
373	0.01	0.00	-0.02	0.00	0.01	0.00	0.00	-0.02	-0.01	-0.04	0.00	-0.04
409	-0.02	-0.18	-0.08	0.00	0.05	0.00	0.13	0.00	0.11	0.08	-0.05	0.00
413	0.12	0.04	-0.03	-0.05	-0.02	-0.08	-0.07	-0.08	-0.05	-0.02	-0.03	0.03
418	0.05	-0.10	-0.04	-0.07	0.02	-0.01	0.09	0.00	-0.08	0.05	0.08	0.04
419	0.06	-0.20	-0.04	0.08	0.05	0.13	0.00	0.00	-0.10	0.09	-0.17	-0.04
428	-0.02	0.06	-0.04	-0.06	-0.04	0.00	0.09	0.00	0.06	-0.03	-0.05	0.03
438	0.20	-0.09	-0.03	-0.07	0.17	-0.12	-0.16	-0.13	-0.08	0.04	0.06	0.04

```
472   0.04   0.06     -0.02      0.04       0.07       0.07    -0.02        -0.01        -0.05
    -0.03     0.06    -0.02
506   0.00   0.00      0.00      0.00       0.00       0.00     0.00         0.00         0.00
     0.00    -0.01     0.00
520   0.12   0.03     -0.02      0.00       0.06      -0.08    -0.09        -0.08        -0.05
    -0.01     0.01     0.00
521  -0.01  -0.01      0.00     -0.01       0.00       0.01     0.01         0.01         0.00
     0.00    -0.01     0.00
525   0.00   0.00     -0.03      0.00       0.02       0.00    -0.02         0.00         0.02
    -0.04     0.00     0.03
529   0.01   0.00      0.02      0.00       0.01      -0.01    -0.02        -0.01         0.01
     0.03     0.00    -0.02
535  -0.01   0.00      0.02      0.00       0.00       0.01     0.00         0.00         0.01
     0.04     0.00     0.04
544   0.02   0.05     -0.02     -0.01      -0.08       0.00     0.02         0.10        -0.06
    -0.02     0.01     0.00
586   0.02   0.00      0.04      0.00       0.01      -0.03    -0.03        -0.03         0.01
    -0.06     0.00     0.06
599   0.08  -0.10     -0.03      0.13       0.04      -0.07    -0.07        -0.07         0.04
     0.05    -0.16    -0.05
638   0.05  -0.02     -0.02     -0.01       0.14       0.09    -0.04        -0.01        -0.07
     0.00    -0.01     0.00
681   0.05  -0.09     -0.05      0.00      -0.01       0.00     0.11         0.00        -0.09
     0.04    -0.03     0.01
702   0.02   0.00      0.01      0.00       0.01      -0.01    -0.01        -0.01        -0.01
     0.02     0.00     0.03
     dfb.s.:: dffit   cov.r   cook.d hat
9     0.04    -0.08    1.02_*  0.00   0.02
28   -0.05    -0.14    1.03_*  0.00   0.03_*
34   -0.04     0.10    1.02_*  0.00   0.02
49    0.03     0.06    1.02_*  0.00   0.01
74   -0.02     0.05    1.02_*  0.00   0.01
81    0.05     0.09    1.02_*  0.00   0.02_*
86   -0.06     0.36_*  0.86_*  0.01   0.00
87   -0.08    -0.11    1.04_*  0.00   0.04_*
92   -0.05     0.10    1.02_*  0.00   0.02
115   0.02     0.24_*  0.92_*  0.00   0.00
118   0.02     0.18    0.96_*  0.00   0.00
134  -0.06    -0.18    1.02_*  0.00   0.02_*
179  -0.02     0.38_*  0.84_*  0.01   0.00
197  -0.01     0.18    0.95_*  0.00   0.00
199   0.06    -0.12    1.02_*  0.00   0.02_*
204   0.10    -0.19    1.01    0.00   0.02_*
209  -0.02     0.06    1.02_*  0.00   0.01
214  -0.07    -0.13    1.02_*  0.00   0.02_*
234  -0.04     0.20    0.98_*  0.00   0.01
236   0.00     0.20    0.93_*  0.00   0.00
243  -0.09    -0.15    1.02_*  0.00   0.02_*
246   0.02     0.19    0.95_*  0.00   0.00
247   0.03     0.22    0.94_*  0.00   0.00
259   0.08     0.31_*  0.96_*  0.01   0.01
272  -0.03     0.29_*  0.89_*  0.01   0.00
280  -0.01     0.13    0.98_*  0.00   0.00
281  -0.03     0.29_*  0.91_*  0.01   0.00
282   0.01     0.19    0.95_*  0.00   0.00
308  -0.03     0.18    0.96_*  0.00   0.00
315  -0.03     0.29_*  0.89_*  0.01   0.00
327  -0.03    -0.09    1.02_*  0.00   0.02
331   0.05    -0.12    1.02_*  0.00   0.02_*
332  -0.05     0.11    1.02_*  0.00   0.02_*
335  -0.03     0.17    0.98_*  0.00   0.00
337   0.05     0.21    0.97_*  0.00   0.01
338  -0.07     0.40_*  0.84_*  0.01   0.00
343   0.02    -0.08    1.02_*  0.00   0.02
367   0.01     0.23_*  0.91_*  0.00   0.00
373  -0.06    -0.09    1.03_*  0.00   0.02_*
409   0.03     0.32_*  0.84_*  0.01   0.00
413   0.01     0.15    0.97_*  0.00   0.00
418  -0.03     0.25_*  0.93_*  0.00   0.00
419   0.07     0.33_*  0.90_*  0.01   0.01
428   0.03     0.17    0.97_*  0.00   0.00
438  -0.03     0.28_*  0.94_*  0.01   0.01
```

```
472  -0.03     0.17    0.97_*  0.00   0.00
506   0.00    -0.01    1.02_*  0.00   0.01
520  -0.01     0.13    0.98_*  0.00   0.00
521   0.01    -0.02    1.02_*  0.00   0.01
525   0.04    -0.09    1.02_*  0.00   0.02_*
529  -0.03     0.06    1.02_*  0.00   0.02_*
535   0.06     0.09    1.04_*  0.00   0.03_*
544   0.00     0.18    0.95_*  0.00   0.00
586  -0.08     0.14    1.03_*  0.00   0.02_*
599   0.07     0.24_*  0.99    0.00   0.01
638  -0.01     0.21    0.95_*  0.00   0.00
681   0.02     0.21    0.92_*  0.00   0.00
702   0.03     0.05    1.02_*  0.00   0.02_*
 [ reached getOption("max.print") -- omitted 152 rows ]
            dfb.1_     dfb.scl.g     dfb.smkr     dfb.scl.c     dfb.scl.b      dfb.rgnn
 dfb.rgnsths   dfb.rgnsthw
9     0.0011081025 -1.605472e-03 -0.031487870 -3.731158e-04  1.534383e-02 -1.726851e-03 -
0.0180408023 -0.0035336802
28   -0.0102031421  6.480212e-02  0.008307992 -6.927808e-02  3.900752e-04 -1.428134e-03 -
0.0077200712 -0.0269164096
34   -0.0098550866 -3.841228e-03  0.039849270 -9.311070e-04  2.721592e-02  2.003581e-03 -
0.0116587951  0.0237606908
49   -0.0064278804 -1.237862e-03  0.019791824 -3.827699e-04  1.233246e-02  1.430445e-02 -
0.0039846217  0.0002349529
74    0.0004639627 -9.197497e-04  0.021209794 -4.586142e-04  3.999007e-03  1.294869e-02 -
0.0028395151  0.0011109424
81    0.0010932753 -1.409022e-03  0.026982019 -6.429940e-04  1.505520e-04 -1.348265e-03 -
0.0004179348  0.0207692365
86    0.2670569182 -1.490637e-01 -0.044762571 -1.045609e-01  4.657924e-02 -1.705745e-01 -
0.1846990981 -0.1649697465
87   -0.0008906127 -1.120325e-03 -0.018046219 -5.906018e-04  1.101461e-02  2.300646e-03 -
0.0195736548 -0.0017087322
92   -0.0117650260  2.623749e-04  0.036243426  5.068520e-04  4.868022e-03  2.272966e-03
0.0147012873  0.0025830210
115   0.1144703449 -3.021050e-02 -0.055486774  7.314048e-02 -9.676780e-02 -1.241040e-01 -
0.0943852975 -0.1179606350
118  -0.0194727189 -6.552342e-02 -0.035070533  5.310895e-02  8.343595e-03  9.087557e-02
0.0005285291 -0.0011560865
134  -0.0180676789  8.456129e-02  0.009446058 -8.963700e-02 -3.507082e-02 -1.168668e-03
0.0001627928 -0.0318769010
179   0.0436197821 -8.420464e-02 -0.071335570 -1.107245e-01 -1.942611e-01 -3.004264e-03
0.1949226200  0.0153653718
197   0.0242094081  8.658415e-02 -0.023037195 -1.111173e-02 -5.162024e-02 -7.182551e-04
0.0169989272  0.0976943353
199  -0.0345131532  7.753847e-04 -0.046748264  7.178217e-07  1.324266e-03  2.581627e-02
0.0317961529  0.0242871136
204  -0.0243568502 -7.839682e-03 -0.058984946 -6.949075e-04  5.381236e-02  4.189529e-02
0.0333636264  0.0399508986
209  -0.0080206930 -7.812406e-04  0.025249815 -5.113390e-04  8.684188e-03  1.678550e-02 -
0.0048363631  0.0012429338
214  -0.0194276698 -3.053012e-03 -0.030893821 -5.532559e-04  1.952828e-02  3.139227e-02
0.0235562634  0.0262361163
234  -0.0243435024  9.060670e-02 -0.039145684  9.059865e-02 -7.710915e-02 -3.391766e-03
0.0819707899  0.0011956658
236   0.1196530393 -6.213949e-02 -0.047035932 -3.004842e-03  1.140928e-02 -1.156395e-01 -
0.1177545370 -0.1155132656
243  -0.0399319777 -2.335789e-03 -0.037883568 -4.817325e-04  2.691081e-02  3.619217e-02
0.0245109938  0.0294708115
246  -0.0143250625 -7.206377e-02 -0.034443061  5.416826e-02  4.940657e-02  9.229622e-02 -
0.0106628804 -0.0045452666
247   0.0357569819 -9.110046e-02 -0.026525244  6.245455e-02 -2.687944e-02 -1.149641e-03
0.0125415464  0.1096127942
259   0.1570835556 -1.289874e-01 -0.032745412  1.775078e-01  1.931183e-02 -9.682344e-02 -
0.0863274512 -0.1000151412
272   0.0465722702 -9.439812e-02 -0.030106541 -1.008961e-01 -3.054633e-02 -3.743242e-03
0.0011481505  0.1440901973
280   0.0756241497 -2.056137e-02 -0.025700905 -4.431092e-02  4.905836e-02 -6.850599e-02 -
0.0824232488 -0.0711583536
281   0.0648192022 -1.227668e-01 -0.043905575 -8.105385e-02  8.979908e-02 -6.546694e-03
0.0763351550 -0.0074272338
282  -0.0318061489  3.489147e-02 -0.034871361 -7.307837e-02 -4.312351e-02  3.398505e-03
0.0137161350  0.1035291903
```

```
308  -0.0294684854  7.566646e-02 -0.034632508  4.800793e-02 -5.390915e-02  8.707407e-02
0.0130564323  0.0010946235
315  -0.0282159160 -9.561510e-02 -0.069903283 -9.053804e-02 -3.431422e-02 -1.822005e-03
0.1218321377  0.0038741299
327   0.0063708329  3.255304e-02  0.008700425 -5.242842e-02  1.465477e-02 -1.003388e-03 -
0.0079006399 -0.0215086296
331   0.0006402369  3.689028e-04 -0.046163950  9.770203e-04  4.346482e-03 -2.763378e-02
0.0032303708 -0.0036529570
332  -0.0122691508  6.844368e-04  0.036573905  5.117573e-04  1.501854e-03  2.338782e-03
0.0155958962  0.0029128607
335   0.1241175226 -6.886275e-02 -0.021284081 -4.902802e-02  1.339088e-02 -7.989879e-02 -
0.0843200041 -0.0766872415
337   0.0327946252 -9.107666e-02 -0.022979841  1.000799e-01  2.114093e-02  7.982322e-02
0.0023495691 -0.0033072812
338  -0.0278084681 -1.562503e-01 -0.051599467 -1.216636e-01  7.176198e-02 -3.708762e-03 -
0.0354152642  0.1689859379
343  -0.0122996396 -2.333664e-02  0.006278916 -4.916225e-02 -2.928550e-02  3.254748e-03 -
0.0072092607  0.0047662852
367   0.2009703291 -7.583837e-02 -0.042013767 -3.702644e-03 -5.916568e-02 -1.337538e-01 -
0.1121946227 -0.1269451719
373   0.0115853744 -1.544520e-03 -0.021836413  4.355926e-04  1.486018e-02  1.415040e-03 -
0.0041308379 -0.0209471989
409  -0.0208853126 -1.794615e-01 -0.083587355  6.598498e-04  4.852537e-02 -8.935350e-04
0.1314715137 -0.0033548406
413   0.1214878901  3.703050e-02 -0.025266127 -5.363366e-02 -1.943541e-02 -7.908417e-02 -
0.0689488412 -0.0788222734
418   0.0504534890 -1.035306e-01 -0.041753165 -7.318410e-02  2.118813e-02 -5.116188e-03
0.0855031421 -0.0017777411
419   0.0560852810 -2.020338e-01 -0.036601999  8.010331e-02  4.563685e-02  1.318842e-01 -
0.0025118893 -0.0036556717
428  -0.0230072913  5.895603e-02 -0.044394703 -5.682716e-02 -3.533116e-02  3.497857e-03
0.0879749275  0.0020381402
438   0.2036646968 -8.888268e-02 -0.026817596 -7.366481e-02  1.749091e-01 -1.190888e-01 -
0.1641789802 -0.1266880905
472   0.0354490034  5.698031e-02 -0.017060401  4.025905e-02  7.069981e-02  7.134535e-02 -
0.0196900697 -0.0096516795
506   0.0004564842 -3.679253e-03  0.001061358 -4.051806e-03  4.506494e-04  2.328902e-04 -
0.0017725514  0.0001818367
520   0.1228112208  2.632822e-02 -0.020381355 -4.488772e-03  5.594761e-02 -7.587495e-02 -
0.0860031614 -0.0798799029
521  -0.0084987520 -8.850500e-03  0.001492029 -9.489223e-03  9.017688e-04  6.143130e-03
0.0057626225  0.0058886025
525   0.0022068395 -2.660789e-03 -0.032880603 -3.963034e-04  2.392297e-02 -1.915607e-03 -
0.0207367865 -0.0043920120
529   0.0107303007 -4.382789e-04  0.023052115  5.920878e-05  5.647408e-03 -1.335144e-02 -
0.0186605885 -0.0131741084
535  -0.0075946842  2.462260e-04  0.016374128 -2.853291e-04  5.155467e-04  1.333609e-02
0.0003408626  0.0009036848
544   0.0213584932  4.546710e-02 -0.023639815 -9.700546e-03 -7.651191e-02 -4.716037e-04
0.0231245012  0.0992034977
586   0.0217972971 -1.186576e-03  0.038768383  3.093778e-04  1.449223e-02 -2.651228e-02 -
0.0335912745 -0.0281846792
599   0.0768917437 -1.044333e-01 -0.029952252  1.266479e-01  4.230499e-02 -6.708610e-02 -
0.0692311913 -0.0723643855
638   0.0521501031 -1.604251e-02 -0.020214231 -8.670423e-03  1.424709e-01  9.103214e-02 -
0.0369259142 -0.0133975227
681   0.0473730321 -9.323096e-02 -0.046435930 -9.431505e-04 -6.148585e-03 -2.523483e-03
0.1051212263 -0.0007606232
702   0.0174600377 -1.734043e-03  0.014890114  1.264574e-04  1.082898e-02 -1.347170e-02 -
0.0147610757 -0.0125604362
        dfb.sxml      dfb.sc.:      dfb.s.:.      dfb.sm:.      dfb.s.::        dffit
  cov.r        cook.d         hat
9     0.015231001 -0.0373262052  0.0014493367  0.0322624484  0.0364486605 -0.084746959 1.
0212934 5.526141e-04 0.017412825
28    0.022766917 -0.0305213475  0.1153325519  0.0293621566 -0.0506299610 -0.144532086 1.
0276215 1.607038e-03 0.025687858
34    0.013245315  0.0429229071  0.0014394195 -0.0384158437 -0.0443500567  0.100654399 1.
0203488 7.794832e-04 0.017458317
49    0.007753347  0.0336362475  0.0001421232  0.0149763648  0.0281486063  0.058854240 1.
0184039 2.665385e-04 0.013717498
74   -0.009490246  0.0191888222 -0.0001950422 -0.0198970179 -0.0199313259  0.047586421 1.
0191392 1.742561e-04 0.013914627
```

```
81   -0.017170669  0.0578501021  0.0002614386  0.0196170865  0.0491251064  0.090536760 1.
0242545 6.307033e-04 0.020208191
86   -0.112471717  0.0691115630  0.1276594857  0.0495618269 -0.0581488986  0.364325990 0.
8638601 1.009214e-02 0.004922384
87    0.016856590 -0.0513968111  0.0018088968 -0.0422573801 -0.0787226785 -0.112091313 1.
0420820 9.667903e-04 0.036591972
92    0.013490909  0.0465125476 -0.0009286557 -0.0395135160 -0.0471901695  0.101960471 1.
0203069 7.998383e-04 0.017496600
115   0.089124578  0.0162808471 -0.0302302401 -0.0310884641  0.0154620446  0.238057106 0.
9204070 4.330401e-03 0.003587604
118   0.065317148  0.0279030827 -0.0574437489 -0.0270823074  0.0212080091  0.176044108 0.
9563505 2.375176e-03 0.003370076
134   0.030900746 -0.0388581754  0.1425888437  0.0377588590 -0.0611728423 -0.184505344 1.
0235054 2.618236e-03 0.024885288
179  -0.114697672  0.0417134830  0.0792513943  0.0532895557 -0.0184618156  0.384050138 0.
8406720 1.119124e-02 0.004665503
197  -0.059945752 -0.0348318555  0.0170341837  0.0043042468 -0.0052859702  0.182940659 0.
9501228 2.563652e-03 0.003251820
199   0.021240684 -0.0553250742 -0.0014068487  0.0450326381  0.0562459816 -0.121521540 1.
0203804 1.136076e-03 0.018678455
204  -0.028676060  0.1134295462 -0.0041938823 -0.0481136888  0.0986007025 -0.191403377 1.
0117572 2.816770e-03 0.017649861
209   0.009650032  0.0227652268  0.0004087788 -0.0255628574 -0.0248415756  0.059151679 1.
0178365 2.692380e-04 0.013236296
214  -0.014668749 -0.0463803148 -0.0023083629 -0.0655548901 -0.0715740910 -0.129139741 1.
0203756 1.282936e-03 0.019131508
234   0.060364738 -0.0413559978  0.0997113231 -0.0378395211 -0.0376913632  0.202809508 0.
9776729 3.156826e-03 0.006809641
236   0.077124229  0.0280872369 -0.0068267985  0.0019728879  0.0008747092  0.195287070 0.
9320864 2.917155e-03 0.002826330
243   0.028250066 -0.0578121207 -0.0012048251 -0.0749807434 -0.0883244593 -0.154080629 1.
0213035 1.826146e-03 0.021363762
246   0.064813998  0.0297380230 -0.0591376290 -0.0273518165  0.0202227001  0.186326602 0.
9543494 2.660262e-03 0.003616805
247  -0.076742562  0.0463656592 -0.0751012415 -0.0267865124  0.0340427915  0.215694994 0.
9352744 3.559442e-03 0.003566257
259  -0.072640280  0.0592235659 -0.1949132524 -0.0722310585  0.0847905486  0.312448155 0.
9620725 7.481241e-03 0.010339325
272  -0.095465276  0.0495237122  0.0806032898  0.0424597934 -0.0342892515  0.292183269 0.
8889072 6.505825e-03 0.003908983
280   0.044500161  0.0080794866  0.0185417293  0.0195368613 -0.0111099731  0.130538906 0.
9821908 1.308623e-03 0.003634203
281  -0.095464229  0.0522984907  0.0840793680  0.0405428624 -0.0342811053  0.285241415 0.
9149989 6.213756e-03 0.004776866
282   0.075567004 -0.0123769528 -0.0218337499  0.0281749865  0.0098672419  0.188661258 0.
9464566 2.725681e-03 0.003249672
308   0.073610476 -0.0350143182  0.0588430101 -0.0253642798 -0.0270824168  0.182373032 0.
9591173 2.549509e-03 0.003789094
315   0.097021447  0.0413993633  0.0863692504  0.0407700089 -0.0315085549  0.287547382 0.
8904459 6.301887e-03 0.003840583
327  -0.013661978 -0.0157296117  0.0607245128  0.0228993174 -0.0269619222 -0.089589264 1.
0205213 6.175528e-04 0.017008836
331   0.019449190 -0.0525492954  0.0001988703  0.0442831230  0.0538630820 -0.117381142 1.
0209768 1.060009e-03 0.018901167
332   0.013725007  0.0485142942 -0.0008851234 -0.0400572528 -0.0491488080  0.105135214 1.
0212260 8.504208e-04 0.018412616
335  -0.052360780  0.0321640009  0.0599748434  0.0231825758 -0.0269772824  0.170022067 0.
9763131 2.218724e-03 0.004872147
337  -0.059666652  0.0402984002 -0.1143204757 -0.0442792016  0.0472356132  0.208348408 0.
9695194 3.329693e-03 0.005932160
338   0.121073937  0.0743996736  0.1377554575  0.0473479200 -0.0659949971  0.396251671 0.
8364065 1.190884e-02 0.004827226
343   0.013326799  0.0117650497 -0.0503966698  0.0199150214  0.0221733179 -0.084377076 1.
0201598 5.477956e-04 0.016440676
367  -0.087557370  0.0377391862 -0.0141791129  0.0051948449  0.0084965283  0.233281951 0.
9104395 4.155093e-03 0.003097068
373  -0.009996909 -0.0369044402 -0.0011653376 -0.0435153742 -0.0553097443 -0.093089053 1.
0266572 6.667703e-04 0.022391316
409   0.106888989  0.0766352326 -0.0503324301  0.0032056216  0.0261679182  0.319513175 0.
8438214 7.749075e-03 0.003326059
413  -0.051975185 -0.0157180230 -0.0306019682  0.0252235921  0.0141732206  0.149479706 0.
9716953 1.714536e-03 0.003443260
```

```
418  -0.083536018  0.0455006132  0.0767148729  0.0362076056 -0.0288875116  0.245814843 0.
9312004 4.621111e-03 0.004335642
419  -0.100391129  0.0900673868 -0.1725867344 -0.0368199084  0.0703266342  0.327665450 0.
9009104 8.189370e-03 0.005416008
428   0.060105714 -0.0283836476 -0.0502983898  0.0256450436  0.0265916940  0.171795335 0.
9651230 2.263421e-03 0.003824540
438  -0.084398585  0.0373035266  0.0580570773  0.0358987080 -0.0326529173  0.282654839 0.
9381545 6.112667e-03 0.006147122
472  -0.050180515 -0.0287130584  0.0551187925 -0.0185696026 -0.0288621673  0.166433727 0.
9726186 2.125537e-03 0.004277669
506  -0.001817527  0.0017288132 -0.0057458869  0.0016950376  0.0024118002 -0.008191481 1.
0186995 5.163894e-06 0.012586278
520  -0.050251102 -0.0128955504  0.0072171318  0.0045509642 -0.0053868481  0.134118630 0.
9771218 1.380869e-03 0.003266295
521   0.002896852  0.0039684131 -0.0149509625  0.0039014140  0.0065020013 -0.020439511 1.
0198995 3.215057e-05 0.013889825
525   0.015436511 -0.0417397009  0.0013986427  0.0341127950  0.0407366808 -0.094327357 1.
0237981 6.846096e-04 0.019994782
529   0.008455828  0.0276303172  0.0013028844 -0.0238270586 -0.0291574638  0.062101064 1.
0227024 2.967627e-04 0.017662924
535   0.007469344  0.0420057598  0.0001466920  0.0352036724  0.0620976544  0.091196449 1.
0385562 6.399722e-04 0.032832393
544  -0.059472273 -0.0153215115  0.0083760189  0.0035308938 -0.0005339099  0.175269247 0.
9515687 2.353460e-03 0.003072186
586   0.014466977 -0.0615191021  0.0020420257  0.0584643047 -0.0834112255  0.135668109 1.
0270175 1.416009e-03 0.024732625
599   0.039158224  0.0463806160 -0.1566589216 -0.0530504533  0.0661157216  0.238451086 0.
9903236 4.366449e-03 0.011961237
638  -0.069664179  0.0026332926 -0.0055844882  0.0023739410 -0.0057247341  0.209002712 0.
9525215 3.346494e-03 0.004342632
681  -0.086786590  0.0412222034 -0.0332045708  0.0056061396  0.0207005682  0.207228188 0.
9249035 3.282838e-03 0.002895465
702  -0.011265540  0.0185843547  0.0002230283  0.0276096952  0.0270976272  0.052904365 1.
0241849 2.153811e-04 0.018694191
 [ reached getOption("max.print") -- omitted 152 rows ]
```

6.6 SUBMODEL 1 DIAGNOSTICS

```
sub_model1 <- lm(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi + regio
n, trainData)
> model_diagnostics(sub_model1)
[1] "Model Summary"

Call:
lm(formula = boxcox.charges ~ scale.age * smoker * scale.children +
    scale.bmi + region, data = trainData)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6857 -0.7284 -0.3237  0.1209  8.3222

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                        17.80622    0.06779 262.652  < 2e-16 ***
scale.age                           1.85080    0.03706  49.942  < 2e-16 ***
smokeryes                           5.82967    0.07864  74.131  < 2e-16 ***
scale.children                      0.52123    0.03488  14.945  < 2e-16 ***
scale.bmi                           0.30173    0.03338   9.040  < 2e-16 ***
regionnorthwest                    -0.09282    0.09096  -1.020  0.30765
regionsoutheast                    -0.50144    0.09228  -5.434 6.12e-08 ***
regionsouthwest                    -0.49865    0.09169  -5.439 5.96e-08 ***
scale.age:smokeryes                -1.25945    0.08099 -15.551  < 2e-16 ***
scale.age:scale.children           -0.27087    0.03804  -7.121 1.45e-12 ***
smokeryes:scale.children           -0.50458    0.08165  -6.180 7.62e-10 ***
scale.age:smokeryes:scale.children  0.31127    0.08728   3.566  0.00037 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.483 on 2208 degrees of freedom
Multiple R-squared:  0.805,    Adjusted R-squared:  0.8041
F-statistic: 828.8 on 11 and 2208 DF,  p-value: < 2.2e-16

[1] "ANOVA Table"
Analysis of Variance Table

Response: boxcox.charges
                                  Df  Sum Sq Mean Sq  F value    Pr(>F)
scale.age                          1  6054.0  6054.0 2754.215 < 2.2e-16 ***
smoker                             1 12469.2 12469.2 5672.740 < 2.2e-16 ***
scale.children                     1   376.3   376.3  171.214 < 2.2e-16 ***
scale.bmi                          1   143.5   143.5   65.269 1.064e-15 ***
region                             3   126.7    42.2   19.213 2.684e-12 ***
scale.age:smoker                   1   674.6   674.6  306.899 < 2.2e-16 ***
scale.age:scale.children           1    79.5    79.5   36.172 2.109e-09 ***
smoker:scale.children              1    88.7    88.7   40.347 2.574e-10 ***
scale.age:smoker:scale.children    1    28.0    28.0   12.718 0.0003699 ***
Residuals                       2208  4853.4     2.2
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
[1] "Variance Homogeneity Test"
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 189.4136, Df = 1, p = < 2.22e-16
[1] "Residual Normality Test"

        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.74249, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.01484914      2.029364   0.512
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
                         scale.age                          smokeryes
 scale.children
```

```
                             1.3865                                   1.0178
         1.2280
                          scale.bmi                          regionnorthwest
regionsoutheast
                             1.1246                                   1.5537
         1.7226
                    regionsouthwest                     scale.age:smokeryes            scale.age
:scale.children
                             1.5589                                   1.3400
         1.3432
          smokeryes:scale.children scale.age:smokeryes:scale.children
                             1.2375                                   1.3213
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
179  5.664666         1.6653e-08   3.6970e-05
338  5.602889         2.3711e-08   5.2637e-05
1446 5.602889         2.3711e-08   5.2637e-05
409  5.451716         5.5452e-08   1.2310e-04
1511 5.451716         5.5452e-08   1.2310e-04
86   5.240224         1.7563e-07   3.8989e-04
812  5.210444         2.0589e-07   4.5708e-04
1096 5.210444         2.0589e-07   4.5708e-04
1930 5.210444         2.0589e-07   4.5708e-04
2219 5.210444         2.0589e-07   4.5708e-04
[1] "Influential Points"
Potentially influential observations of
         lm(formula = boxcox.charges ~ scale.age * smoker * scale.children +      scale.bm
i + region, data = trainData) :

     dfb.1_ dfb.scl.g dfb.smkr dfb.scl.c dfb.scl.b dfb.rgnn dfb.rgnsths dfb.rgnsthw dfb.s
c.: dfb.s.:. dfb.sm:. dfb.s.::
9     0.01   0.00     -0.03     0.00      0.01      0.00    -0.02        0.00       -0.03
     0.00     0.03     0.03
28    0.00   0.06      0.01    -0.06      0.00      0.00    -0.01       -0.02       -0.03
     0.10     0.03    -0.05
34    0.00   0.00      0.04     0.00      0.03      0.00    -0.01        0.02        0.04
     0.00    -0.04    -0.04
49    0.00   0.00      0.02     0.00      0.01      0.01     0.00        0.00        0.03
     0.00     0.01     0.03
74    0.00   0.00      0.02     0.00      0.00      0.02     0.00        0.00        0.02
     0.00    -0.02    -0.02
81   -0.01   0.00      0.03     0.00      0.00      0.00     0.00        0.02        0.07
     0.00     0.02     0.06
86    0.24  -0.15     -0.05    -0.11      0.04     -0.17    -0.19       -0.17        0.07
     0.13     0.05    -0.06
87    0.01   0.00     -0.01     0.00      0.01      0.00    -0.02        0.00       -0.04
     0.00    -0.03    -0.06
92   -0.01   0.00      0.03     0.00      0.01      0.00     0.01        0.00        0.04
     0.00    -0.04    -0.04
115   0.16  -0.03     -0.05     0.07     -0.09     -0.12    -0.09       -0.12        0.02
    -0.03    -0.03     0.02
118   0.01  -0.07     -0.03     0.05      0.01      0.09     0.00        0.00        0.03
    -0.06    -0.03     0.02
134  -0.01   0.08      0.01    -0.08     -0.03      0.00     0.00       -0.03       -0.03
     0.13     0.04    -0.06
179  -0.01  -0.08     -0.08    -0.11     -0.20      0.00     0.20        0.02        0.04
     0.08     0.05    -0.02
197   0.00   0.09     -0.03    -0.01     -0.06      0.00     0.02        0.10       -0.04
     0.02     0.00    -0.01
199  -0.03   0.00     -0.04     0.00      0.00      0.02     0.03        0.02       -0.05
     0.00     0.04     0.05
204  -0.04  -0.01     -0.06     0.00      0.05      0.04     0.03        0.04        0.12
     0.00    -0.05     0.10
209   0.00   0.00      0.02     0.00      0.01      0.01     0.00        0.00        0.02
     0.00    -0.02    -0.02
214  -0.03   0.00     -0.03     0.00      0.02      0.03     0.02        0.03       -0.05
     0.00    -0.07    -0.08
234   0.00   0.08     -0.03     0.09     -0.07      0.00     0.08        0.00       -0.04
     0.09    -0.04    -0.04
236   0.16  -0.06     -0.04     0.00      0.01     -0.11    -0.12       -0.11        0.03
    -0.01     0.00     0.00
```

241	0.00	0.00	0.03	0.00	0.02	0.02	-0.01	0.00	-0.03	0.00	0.05	-0.04
243	-0.03	0.00	-0.03	0.00	0.03	0.03	0.02	0.03	-0.05	0.00	-0.07	-0.08
246	0.01	-0.07	-0.03	0.05	0.05	0.09	-0.01	-0.01	0.03	-0.06	-0.03	0.02
247	0.00	-0.09	-0.03	0.06	-0.03	0.00	0.01	0.11	0.05	-0.07	-0.03	0.03
259	0.14	-0.13	-0.04	0.18	0.02	-0.10	-0.09	-0.10	0.06	-0.20	-0.08	0.09
272	0.01	-0.09	-0.04	-0.10	-0.04	0.00	0.00	0.15	0.05	0.08	0.04	-0.04
281	0.03	-0.12	-0.05	-0.08	0.09	-0.01	0.08	-0.01	0.05	0.09	0.04	-0.04
282	0.00	0.03	-0.03	-0.07	-0.04	0.00	0.01	0.10	-0.01	-0.02	0.03	0.01
304	0.01	0.00	0.02	0.00	0.00	-0.01	-0.01	-0.01	0.02	0.00	-0.02	-0.02
308	0.00	0.07	-0.03	0.05	-0.05	0.08	0.01	0.00	-0.03	0.05	-0.02	-0.02
315	0.01	-0.10	-0.06	-0.09	-0.03	0.00	0.12	0.00	0.04	0.08	0.04	-0.03
327	0.00	0.04	0.01	-0.06	0.02	0.00	-0.01	-0.02	-0.02	0.07	0.02	-0.03
331	0.01	0.00	-0.04	0.00	0.00	-0.03	0.00	0.00	-0.05	0.00	0.04	0.05
332	-0.01	0.00	0.03	0.00	0.00	0.00	0.01	0.00	0.04	0.00	-0.04	-0.04
335	0.12	-0.07	-0.03	-0.05	0.01	-0.08	-0.09	-0.08	0.03	0.06	0.02	-0.03
337	0.01	-0.09	-0.03	0.10	0.02	0.08	0.00	0.00	0.04	-0.12	-0.05	0.05
338	0.03	-0.16	-0.04	-0.12	0.08	-0.01	-0.03	0.17	0.08	0.13	0.05	-0.06
343	-0.01	-0.02	0.01	-0.04	-0.03	0.00	-0.01	0.00	0.01	-0.05	0.02	0.02
367	0.18	-0.07	-0.05	0.00	-0.06	-0.13	-0.11	-0.13	0.04	-0.01	0.00	0.01
373	0.01	0.00	-0.02	0.00	0.02	0.00	0.00	-0.02	-0.04	0.00	-0.05	-0.06
409	0.03	-0.18	-0.07	0.00	0.05	0.00	0.13	0.00	0.08	-0.05	0.01	0.03
413	0.11	0.04	-0.03	-0.05	-0.02	-0.08	-0.07	-0.08	-0.02	-0.03	0.02	0.01
418	0.02	-0.10	-0.05	-0.07	0.02	0.00	0.09	0.00	0.04	0.08	0.04	-0.03
419	0.02	-0.20	-0.05	0.08	0.04	0.14	0.00	0.00	0.09	-0.17	-0.04	0.07
428	0.00	0.05	-0.04	-0.06	-0.03	0.00	0.09	0.00	-0.03	-0.05	0.03	0.03
438	0.19	-0.09	-0.03	-0.07	0.17	-0.12	-0.17	-0.13	0.04	0.06	0.03	-0.04
447	0.00	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.01	0.00
472	0.02	0.06	-0.02	0.04	0.07	0.07	-0.02	-0.01	-0.03	0.06	-0.02	-0.03
475	0.00	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.01	-0.02	0.00	0.01
479	0.01	0.00	-0.05	0.00	0.02	-0.03	0.00	0.00	-0.05	0.00	0.05	0.05
506	0.00	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.01
520	0.12	0.03	-0.03	0.00	0.06	-0.08	-0.09	-0.08	-0.01	0.01	0.00	-0.01
521	-0.01	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00
525	0.01	0.00	-0.03	0.00	0.02	0.00	-0.02	0.00	-0.04	0.00	0.03	0.04
529	0.01	0.00	0.02	0.00	0.01	-0.01	-0.02	-0.01	0.02	0.00	-0.02	-0.03
535	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.04	0.00	0.03	0.06

```
544   0.00   0.05     -0.03    -0.01      -0.08       0.00      0.02         0.10        -0.02
     0.01     0.00     0.00
551   0.10   0.06     -0.02     0.04       0.02      -0.07     -0.08        -0.07        -0.03
     0.06    -0.01    -0.03
586   0.03   0.00      0.04     0.00       0.01      -0.02     -0.03        -0.03        -0.06
     0.00     0.05    -0.08
599   0.10  -0.10     -0.03     0.12       0.04      -0.07     -0.07        -0.07         0.05
    -0.15    -0.05     0.06
638   0.03  -0.01     -0.03    -0.01       0.14       0.09     -0.04        -0.01         0.00
     0.00     0.00    -0.01
681   0.01  -0.09     -0.05     0.00      -0.01       0.00      0.11         0.00         0.04
    -0.03     0.00     0.02
     dffit    cov.r   cook.d hat
9    -0.07     1.02_*  0.00   0.02_*
28   -0.13     1.03_*  0.00   0.03_*
34    0.09     1.02_*  0.00   0.02_*
49    0.05     1.02_*  0.00   0.01
74    0.06     1.02_*  0.00   0.01
81    0.10     1.02_*  0.00   0.02_*
86    0.35_*  0.87_*  0.01   0.00
87   -0.09     1.04_*  0.00   0.04_*
92    0.09     1.02_*  0.00   0.02_*
115   0.22     0.93_*  0.00   0.00
118   0.16     0.96_*  0.00   0.00
134  -0.17     1.02_*  0.00   0.02_*
179   0.37_*  0.85_*  0.01   0.00
197   0.18     0.95_*  0.00   0.00
199  -0.11     1.02_*  0.00   0.02_*
204  -0.20     1.01    0.00   0.02_*
209   0.05     1.02_*  0.00   0.01
214  -0.14     1.02_*  0.00   0.02_*
234   0.19     0.98_*  0.00   0.01
236   0.18     0.94_*  0.00   0.00
241   0.08     1.02_*  0.00   0.01
243  -0.14     1.02_*  0.00   0.02_*
246   0.17     0.96_*  0.00   0.00
247   0.21     0.94_*  0.00   0.00
259   0.31_*  0.96_*  0.01   0.01
272   0.28_*  0.89_*  0.01   0.00
281   0.27_*  0.92_*  0.01   0.00
282   0.17     0.95_*  0.00   0.00
304   0.04     1.02_*  0.00   0.01
308   0.16     0.96_*  0.00   0.00
315   0.27_*  0.90_*  0.01   0.00
327  -0.10     1.02_*  0.00   0.02_*
331  -0.10     1.02_*  0.00   0.02_*
332   0.10     1.02_*  0.00   0.02_*
335   0.17     0.98_*  0.00   0.00
337   0.20     0.97_*  0.00   0.01
338   0.37_*  0.85_*  0.01   0.00
343  -0.07     1.02_*  0.00   0.02
367   0.22     0.91_*  0.00   0.00
373  -0.10     1.03_*  0.00   0.02_*
409   0.30_*  0.86_*  0.01   0.00
413   0.14     0.97_*  0.00   0.00
418   0.24_*  0.93_*  0.00   0.00
419   0.32_*  0.91_*  0.01   0.00
428   0.16     0.97_*  0.00   0.00
438   0.27_*  0.94_*  0.01   0.01
447  -0.02     1.02_*  0.00   0.01
472   0.16     0.97_*  0.00   0.00
475  -0.02     1.02_*  0.00   0.01
479  -0.11     1.02_*  0.00   0.02
506  -0.02     1.02_*  0.00   0.01
520   0.13     0.98_*  0.00   0.00
521  -0.01     1.02_*  0.00   0.01
525  -0.08     1.02_*  0.00   0.02_*
529   0.05     1.02_*  0.00   0.02_*
535   0.08     1.04_*  0.00   0.03_*
544   0.17     0.95_*  0.00   0.00
551   0.13     0.98_*  0.00   0.00
586   0.13     1.03_*  0.00   0.02_*
```

```
599   0.23_*  0.99    0.00   0.01
638   0.20    0.95_*  0.00   0.00
681   0.19    0.93_*  0.00   0.00
 [ reached getOption("max.print") -- omitted 158 rows ]
            dfb.1_     dfb.scl.g      dfb.smkr     dfb.scl.c     dfb.scl.b      dfb.rgnn
 dfb.rgnsths   dfb.rgnsthw
9     7.218047e-03 -1.987509e-03 -0.026489223 -3.739249e-04  0.0140261999 -1.718710e-03 -
0.0157248930 -3.223542e-03
28   -6.335704e-04  5.758624e-02  0.009172818 -6.254525e-02  0.0013166357 -1.615291e-03 -
0.0069378218 -2.448945e-02
34   -4.350892e-03 -4.075615e-03  0.037777299 -9.013131e-04  0.0256477804  1.649812e-03 -
0.0107158595  2.174002e-02
49   -3.123923e-03 -1.413362e-03  0.018314508 -3.675118e-04  0.0113653041  1.268182e-02 -
0.0035549971  1.368941e-04
74   -4.633405e-03 -5.997284e-04  0.024478869 -5.060500e-04  0.0042453891  1.563478e-02 -
0.0034015402  1.445398e-03
81   -7.691663e-03 -7.457120e-04  0.029287620 -6.619197e-04 -0.0007480204 -1.225440e-03 -
0.0004990810  2.389141e-02
86    2.449758e-01 -1.458959e-01 -0.054600099 -1.053379e-01  0.0418173010 -1.707247e-01 -
0.1869328256 -1.656561e-01
87    5.648726e-03 -1.543166e-03 -0.013862309 -5.397631e-04  0.0097844124  1.682381e-03 -
0.0161885149 -1.562055e-03
92   -6.185916e-03 -3.053980e-04  0.034480967  4.214407e-04  0.0050705864  1.894972e-03
0.0135500226  2.247203e-03
115   1.637641e-01 -3.342215e-02 -0.047418687  7.123095e-02 -0.0906692103 -1.228073e-01 -
0.0922300253 -1.163192e-01
118   8.550186e-03 -6.664799e-02 -0.029119339  5.146871e-02  0.0111184502  8.746194e-02
0.0005875941 -1.796891e-03
134  -5.155689e-03  7.662985e-02  0.011014542 -8.258132e-02 -0.0309693353 -1.528804e-03
0.0001824718 -2.963260e-02
179  -5.085201e-03 -7.999612e-02 -0.081602502 -1.113693e-01 -0.2017878130 -1.187151e-03
0.1966658963  1.673694e-02
197  -1.094684e-03  9.081735e-02 -0.028422254 -1.107751e-02 -0.0554138729  2.414778e-04
0.0172159737  9.999357e-02
199  -2.555630e-02 -1.446407e-04 -0.040952034 -6.944646e-05  0.0021059134  2.309331e-02
0.0288389848  2.180935e-02
204  -4.188484e-02 -6.873985e-03 -0.064199446 -6.165011e-04  0.0548378754  4.422346e-02
0.0347987094  4.202897e-02
209  -3.822712e-03 -1.055770e-03  0.022836984 -4.779635e-04  0.0080002694  1.454822e-02 -
0.0042202813  9.980324e-04
214  -2.973560e-02 -2.540926e-03 -0.033929647 -5.270530e-04  0.0198872686  3.335181e-02
0.0248206924  2.782904e-02
234   1.081867e-03  8.483170e-02 -0.033149218  8.709055e-02 -0.0716523077 -4.196225e-03
0.0790548003  5.379967e-04
236   1.638924e-01 -6.415990e-02 -0.040066682 -3.213696e-03  0.0147155646 -1.143226e-01 -
0.1150875026 -1.137848e-01
241  -4.626735e-03 -2.088679e-03  0.032747519 -5.965335e-04  0.0170306730  2.119660e-02 -
0.0062384255 -1.182670e-03
243  -2.816591e-02 -3.262982e-03 -0.032539650 -5.324522e-04  0.0257460183  3.255749e-02
0.0223540149  2.657159e-02
246   1.386243e-02 -7.308844e-02 -0.028580872  5.256574e-02  0.0511753841  8.896125e-02 -
0.0103200337 -5.097373e-03
247   3.937712e-03 -8.930245e-02 -0.033363533  6.380211e-02 -0.0310338864  7.510249e-05
0.0126652894  1.123087e-01
259   1.426935e-01 -1.288215e-01 -0.039608988  1.818808e-01  0.0162882260 -9.788808e-02 -
0.0884074628 -1.015469e-01
272   7.207351e-03 -9.137278e-02 -0.038366357 -1.017617e-01 -0.0354879336 -2.248347e-03
0.0010509401  1.468561e-01
281   2.766734e-02 -1.205873e-01 -0.052584652 -8.200364e-02  0.0867855931 -5.106136e-03
0.0774503284 -6.522753e-03
282  -6.072816e-05  3.077693e-02 -0.028136822 -7.145177e-02 -0.0385933736  2.137152e-03
0.0134450585  1.000728e-01
304   1.296653e-02 -4.182396e-04  0.019146852  4.728340e-05  0.0030707963 -1.141153e-02 -
0.0148175419 -1.118385e-02
308   1.558038e-03  7.025102e-02 -0.027911808  4.626118e-02 -0.0489454555  8.324985e-02
0.0127339334  3.075646e-04
315   1.357913e-02 -9.818651e-02 -0.061181407 -8.925969e-02 -0.0292588438 -3.308705e-03
0.1197538676  2.798674e-03
327   7.611361e-04  3.634134e-02  0.008351061 -5.735842e-02  0.0153627483 -8.603711e-04 -
0.0086689156 -2.339668e-02
331   8.807017e-03 -4.421862e-04 -0.040581229  8.218903e-04  0.0047751523 -2.534711e-02
0.0029500869 -3.499526e-03
```

```
332  -6.593235e-03  7.403436e-05  0.034780522  4.248900e-04  0.0019766905  1.950845e-03
0.0143644094  2.546941e-03
335   1.155479e-01 -6.840805e-02 -0.026264620 -5.014632e-02  0.0112389718 -8.119291e-02 -
0.0866412712 -7.817880e-02
337   8.722727e-03 -9.071239e-02 -0.028555077  1.027628e-01  0.0188122891  8.277015e-02
0.0023370456 -2.743220e-03
338   2.502037e-02 -1.592384e-01 -0.041347228 -1.202199e-01  0.0763322831 -5.552636e-03 -
0.0347314537  1.651273e-01
343  -6.597250e-03 -2.139581e-02  0.006589395 -4.396955e-02 -0.0256357002  2.718591e-03 -
0.0064277085  4.133123e-03
367   1.834692e-01 -7.306711e-02 -0.049908470 -3.431389e-03 -0.0642345244 -1.342346e-01 -
0.1138731707 -1.277985e-01
373   8.831348e-03 -1.199765e-03 -0.024648473  5.121904e-04  0.0156337207  1.708703e-03 -
0.0044950716 -2.261767e-02
409   2.609016e-02 -1.816694e-01 -0.074140683  2.667095e-04  0.0528228720 -2.560256e-03
0.1296849020 -4.418654e-03
413   1.124580e-01  4.030400e-02 -0.030254614 -5.473195e-02 -0.0224261260 -8.015048e-02 -
0.0706706732 -8.016313e-02
418   1.721568e-02 -1.016870e-01 -0.049475606 -7.414627e-02  0.0175871387 -3.850626e-03
0.0868906196 -9.113035e-04
419   1.554332e-02 -2.006382e-01 -0.045464892  8.162112e-02  0.0415570882  1.354124e-01 -
0.0026648893 -2.632952e-03
428   2.403947e-03  5.474102e-02 -0.038544348 -5.540246e-02 -0.0316083839  2.466743e-03
0.0855068621  1.363123e-03
438   1.888856e-01 -8.687292e-02 -0.034357952 -7.475379e-02  0.1743931419 -1.199989e-01 -
0.1674018548 -1.281977e-01
447  -7.514422e-04 -5.058909e-03  0.001266735 -1.252496e-02 -0.0003782232  5.920121e-04
0.0005728778 -4.614291e-03
472   1.619894e-02  6.062451e-02 -0.021658228  4.138288e-02  0.0700119679  7.383312e-02 -
0.0202071704 -9.334559e-03
475  -1.350687e-03 -1.254488e-02  0.002450729 -7.354748e-03 -0.0026757211  6.968456e-04 -
0.0040198942  8.364376e-04
479   1.177953e-02 -1.904722e-03 -0.045665713  9.083918e-04  0.0166579140 -2.958899e-02 -
0.0001282942 -4.806366e-03
506  -7.250928e-04 -7.717903e-03  0.001966820 -8.665754e-03  0.0007835329  5.610476e-04 -
0.0038016793  4.306704e-04
520   1.151792e-01  2.936014e-02 -0.025183149 -4.425291e-03  0.0551369942 -7.716663e-02 -
0.0884488329 -8.155549e-02
521  -5.403530e-03 -6.051646e-03  0.001165439 -6.397190e-03  0.0006995692  4.106207e-03
0.0038827959  3.945008e-03
525   8.427089e-03 -2.941346e-03 -0.027924594 -3.980984e-04  0.0217073555 -1.901993e-03 -
0.0182274980 -4.007468e-03
529   1.363978e-02 -7.029877e-04  0.020633008  2.462731e-05  0.0052436294 -1.168848e-02 -
0.0161626004 -1.149452e-02
535  -4.401794e-03 -7.402573e-05  0.015264658 -2.801290e-04  0.0007771475  1.184829e-02
0.0003132207  7.393483e-04
544  -4.068089e-03  4.896338e-02 -0.029009621 -9.648037e-03 -0.0807630890  4.848410e-04
0.0234548091  1.015633e-01
551   1.037583e-01  6.190658e-02 -0.023521543  3.799261e-02  0.0186532765 -7.320837e-02 -
0.0753301173 -7.456710e-02
586   2.860743e-02 -1.699629e-03  0.037308866  2.390597e-04  0.0141441379 -2.490824e-02 -
0.0312673248 -2.639148e-02
599   9.974194e-02 -1.030348e-01 -0.026073213  1.226668e-01  0.0428501978 -6.566380e-02 -
0.0670865022 -7.057306e-02
638   2.573008e-02 -1.325492e-02 -0.026432498 -8.589063e-03  0.1422071063  9.404045e-02 -
0.0377633527 -1.292180e-02
681   1.222490e-02 -9.100762e-02 -0.054495162 -6.377162e-04 -0.0104052950 -1.159250e-03
0.1067928730  1.583807e-04
          dfb.sc.:       dfb.s.:.       dfb.sm:.       dfb.s.::        dffit     cov.r
cook.d          hat
9    -0.0322192097  8.321287e-04  0.0284252520  0.0321073603 -0.07273570 1.0209686 4.4101
16e-04 0.016850383
28   -0.0269776274  1.033849e-01  0.0269012024 -0.0452045044 -0.12870175 1.0276770 1.3805
67e-03 0.025050464
34    0.0398465567  9.281862e-04 -0.0351135953 -0.0405526573  0.09182906 1.0203181 7.0287
93e-04 0.017156001
49    0.0302750031 -9.877866e-05  0.0135358213  0.0253321123  0.05216755 1.0180858 2.2687
00e-04 0.013479431
74    0.0226013917  1.361890e-04 -0.0239800690 -0.0240455695  0.05564625 1.0178007 2.5813
23e-04 0.013361200
81    0.0654622498  9.363501e-04  0.0219745982  0.0555917212  0.10136612 1.0225485 8.5644
48e-04 0.019481329
```

```
86    0.0668520180  1.328884e-01  0.0478150747 -0.0613832346  0.35047607 0.8706754 1.0114
91e-02 0.004453267
87   -0.0421888029  1.043563e-03 -0.0347063271 -0.0648752825 -0.09174232 1.0414583 7.0163
34e-04 0.035764447
92    0.0431759780 -1.260271e-03 -0.0361348749 -0.0431805432  0.09305142 1.0202875 7.2171
23e-04 0.017190282
115   0.0182865651 -3.242888e-02 -0.0286386268  0.0170893135  0.21593776 0.9294783 3.8611
57e-03 0.003084756
118   0.0288897739 -5.802147e-02 -0.0250716372  0.0220800813  0.15914114 0.9619368 2.1031
68e-03 0.002906148
134  -0.0349995905  1.303426e-01  0.0353303568 -0.0556616486 -0.16737011 1.0240622 2.3342
59e-03 0.024187276
179   0.0390031275  8.383048e-02  0.0514504140 -0.0212477428  0.37005080 0.8491477 1.1253
02e-02 0.004249371
197  -0.0370792403  1.931759e-02  0.0031338148 -0.0067470980  0.17575956 0.9519750 2.5631
28e-03 0.002902661
199  -0.0496395102 -1.903039e-03  0.0412144209  0.0514213851 -0.10843948 1.0204496 9.8008
76e-04 0.018107803
204   0.1176601264 -3.405264e-03 -0.0508542755  0.1022963986 -0.19757812 1.0108147 3.2512
91e-03 0.017253691
209   0.0201461080  8.266836e-05 -0.0221901964 -0.0215426756  0.05104214 1.0174693 2.1718
71e-04 0.012884015
214  -0.0493403635 -1.931255e-03 -0.0694531581 -0.0758367536 -0.13528765 1.0195203 1.5252
63e-03 0.018884668
234  -0.0382947580  9.423826e-02 -0.0352830408 -0.0350211960  0.18657194 0.9815658 2.8947
61e-03 0.006206367
236   0.0295188674 -9.139005e-03  0.0034695238  0.0025517052  0.17548109 0.9398906 2.5524
04e-03 0.002385514
241  -0.0271445524 -1.305806e-04  0.0489339838 -0.0354308105  0.08484952 1.0166650 6.0008
53e-04 0.013793968
243  -0.0519791938 -1.936419e-03 -0.0677810061 -0.0798870942 -0.13795655 1.0216203 1.5860
71e-03 0.020645603
246   0.0307008506 -5.972807e-02 -0.0253749546  0.0211354037  0.17026307 0.9601166 2.4069
74e-03 0.003179169
247   0.0450642539 -7.387649e-02 -0.0288408954  0.0328761961  0.20501179 0.9376660 3.4828
45e-03 0.003114812
259   0.0586069942 -1.971005e-01 -0.0754319871  0.0851006223  0.31090795 0.9628764 8.0233
84e-03 0.009780480
272   0.0475300442  8.476711e-02  0.0410089928 -0.0368836813  0.27947747 0.8943461 6.4468
06e-03 0.003491687
281   0.0505387997  8.863677e-02  0.0392233690 -0.0370215260  0.27311138 0.9183812 6.1696
86e-03 0.004241810
282  -0.0100726139 -2.367756e-02  0.0289522382  0.0112686648  0.16838328 0.9530858 2.3527
66e-03 0.002728311
304   0.0168531696  8.498635e-04 -0.0189043061 -0.0178775747  0.04277468 1.0165168 1.5253
12e-04 0.011715887
308  -0.0320156507  5.472577e-02 -0.0231278857 -0.0246463402  0.16168812 0.9646618 2.1714
90e-03 0.003171798
315   0.0432437487  8.175828e-02  0.0419906598 -0.0288084302  0.26582396 0.9020026 5.8364
79e-03 0.003403349
327  -0.0176352403  6.702012e-02  0.0247761391 -0.0298688696 -0.09695862 1.0193484 7.8357
19e-04 0.016613296
331  -0.0472009957 -3.943789e-04  0.0405312638  0.0492604531 -0.10498816 1.0210070 9.1871
40e-04 0.018382253
332   0.0449956187 -1.226378e-03 -0.0366068866 -0.0449496415  0.09590774 1.0212094 7.6669
85e-04 0.018098822
335   0.0315860035  6.336666e-02  0.0227115084 -0.0289159007  0.16609305 0.9763259 2.2934
78e-03 0.004410062
337   0.0396529531 -1.152001e-01 -0.0466244916  0.0470343136  0.20452806 0.9698911 3.4755
26e-03 0.005445645
338   0.0764967664  1.318159e-01  0.0490593087 -0.0623124363  0.37147603 0.8524308 1.1343
40e-02 0.004376558
343   0.0108373265 -4.544078e-02  0.0180405528  0.0200843166 -0.07444366 1.0199433 4.6195
96e-04 0.016030545
367   0.0358856889 -1.149255e-02  0.0034569483  0.0066219217  0.21926529 0.9144018 3.9757
91e-03 0.002660780
373  -0.0403545906 -9.117206e-04 -0.0474521459 -0.0602787026 -0.10043186 1.0257201 8.4075
76e-04 0.022133082
409   0.0783953685 -5.306116e-02  0.0053095077  0.0281629715  0.29673435 0.8588776 7.2433
86e-03 0.002953823
413  -0.0175398493 -2.962107e-02  0.0247502096  0.0133215695  0.14352816 0.9720311 1.7122
08e-03 0.003026969
```

```
418   0.0440127169  8.085834e-02  0.0351097469 -0.0313074050  0.23520083 0.9336995 4.5822
07e-03 0.003834934
419   0.0886399197 -1.718314e-01 -0.0394327856  0.0690618678  0.31637166 0.9053000 8.2686
62e-03 0.004907604
428  -0.0259993697 -5.077713e-02  0.0261024963  0.0271442971  0.15629987 0.9698831 2.0300
54e-03 0.003356387
438   0.0357046226  6.199589e-02  0.0348341995 -0.0352162214  0.27489318 0.9401652 6.2619
63e-03 0.005599061
447   0.0021612692 -1.061020e-02  0.0054082167  0.0046765144 -0.01832503 1.0170164 2.7996
22e-05 0.011526840
472  -0.0307783952  5.810434e-02 -0.0200534703 -0.0306955017  0.16238039 0.9730357 2.1915
71e-03 0.003888808
475   0.0059332626 -1.633990e-02  0.0029259450  0.0069683294 -0.02210975 1.0169307 4.0754
45e-05 0.011514321
479  -0.0459192137 -5.132308e-04  0.0457700408  0.0483671426 -0.11094198 1.0165087 1.0257
70e-03 0.015166544
506   0.0035997091 -1.218862e-02  0.0035524457  0.0050803840 -0.01711058 1.0174947 2.4408
45e-05 0.011966645
520  -0.0146528676  9.104343e-03  0.0036242265 -0.0066988272  0.12779903 0.9770180 1.3580
97e-03 0.002807764
521   0.0027259842 -1.013690e-02  0.0026675708  0.0044234128 -0.01362503 1.0192501 1.5477
04e-05 0.013610822
525  -0.0363669698  7.885492e-04  0.0302939745  0.0361522868 -0.08186534 1.0235235 5.5866
21e-04 0.019459307
529   0.0241506719  8.908649e-04 -0.0205029605 -0.0251091023  0.05331473 1.0222923 2.3696
17e-04 0.017335450
535   0.0378454882 -8.662313e-05  0.0316972602  0.0558244882  0.08146735 1.0382342 5.5327
87e-04 0.032612145
544  -0.0172259308  1.049697e-02  0.0023579920 -0.0019034821  0.16772143 0.9533031 2.3343
53e-03 0.002718461
551  -0.0285314338  5.850349e-02 -0.0147907492 -0.0267403118  0.13475191 0.9822849 1.5104
61e-03 0.003644506
586  -0.0569483959  1.463649e-03  0.0547327686 -0.0773957068  0.12562424 1.0271298 1.3153
41e-03 0.024451389
599   0.0460147881 -1.532225e-01 -0.0506762517  0.0649790399  0.22807537 0.9931717 4.3281
66e-03 0.011638667
638   0.0007683793 -3.385114e-03  0.0009722209 -0.0074416457  0.20108355 0.9538040 3.3552
13e-03 0.003860164
681   0.0395495516 -3.090662e-02  0.0039008614  0.0190709438  0.19135364 0.9276767 3.0317
18e-03 0.002387627
 [ reached getOption("max.print") -- omitted 158 rows ]
> press(sub_model1)
[1] 4903.566
```

6.7 SUBMODEL 2 DIAGNOSTICS

```
"Variance Homogeneity Test"
Non-constant Variance Score Test 
Variance formula: ~ fitted.values 
Chisquare = 212.9258, Df = 1, p = < 2.22e-16
[1] "Residual Normality Test"

	Shapiro-Wilk normality test

data:  residuals(model)
W = 0.74582, p-value < 2.2e-16

[1] "Test for Auto Correlated Errors"
 lag Autocorrelation D-W Statistic p-value
   1     -0.01096771      2.021586   0.612
 Alternative hypothesis: rho != 0
[1] "Test for Multi-Collinearity"
               scale.age                smokeryes           scale.children 
 scale.bmi      scale.age:smokeryes 
                 1.7308                   1.0020                   1.2264 
    1.0146                   1.2648 
      scale.age:children smokeryes:scale.children 
                 1.4908                   1.2343 
[1] "Test for Outliers"
     rstudent unadjusted p-value Bonferroni p
179  5.368501         8.7700e-08   0.00019469
338  5.363152         9.0304e-08   0.00020047
1446 5.363152         9.0304e-08   0.00020047
86   5.325086         1.1112e-07   0.00024669
409  5.271545         1.4843e-07   0.00032951
1511 5.271545         1.4843e-07   0.00032951
812  5.172000         2.5245e-07   0.00056043
1096 5.172000         2.5245e-07   0.00056043
1930 5.172000         2.5245e-07   0.00056043
2219 5.172000         2.5245e-07   0.00056043
[1] "Influential Points"
Potentially influential observations of
	 lm(formula = boxcox.charges ~ scale.age + smoker + scale.children +      scale.bm
i + scale.age:smoker + scale.age:children + smoker:scale.children,      data = trainData)
 :

     dfb.1_ dfb.scl.g dfb.smkr dfb.scl.c dfb.scl.b dfb.scl.g:s dfb.scl.g:c dfb.sm:. dffit
   cov.r   cook.d hat    
9      0.00  -0.02     -0.04     0.00      0.02     -0.07        0.03        0.05    -0.10
    1.01_*  0.00   0.01_*
28   -0.02  -0.02      0.01    -0.06      0.00     -0.02        0.09        0.03    -0.11
    1.02_*  0.00   0.02_*
31     0.00   0.00      0.00    -0.01      0.00      0.00       -0.01        0.00    -0.02
    1.01_*  0.00   0.01  
34     0.00   0.00      0.02     0.00      0.01      0.03       -0.01       -0.02     0.04
    1.01_*  0.00   0.01_*
51     0.00   0.01      0.03     0.00      0.01     -0.04       -0.02        0.03     0.07
    1.01_*  0.00   0.01  
61    -0.01  -0.01      0.01    -0.04      0.00      0.00        0.03        0.02    -0.05
    1.01_*  0.00   0.01  
74     0.00   0.01      0.02     0.00      0.00      0.03       -0.01       -0.02     0.04
    1.01_*  0.00   0.01  
81     0.00  -0.02      0.04     0.00      0.00      0.06        0.03        0.02     0.09
    1.01_*  0.00   0.01_*
86     0.13  -0.22     -0.06    -0.11      0.00      0.08        0.11        0.05     0.28
_*  0.91_*  0.01   0.00  
87     0.00   0.00     -0.01     0.00      0.00     -0.01       -0.01       -0.01    -0.02
    1.02_*  0.00   0.02_*
92     0.00   0.01      0.02     0.00      0.00      0.03       -0.01       -0.02     0.04
    1.02_*  0.00   0.01_*
115    0.10   0.00     -0.04     0.07     -0.11      0.01       -0.03       -0.03     0.17
    0.95_*  0.00   0.00  
118    0.08  -0.02     -0.03     0.06      0.00      0.03       -0.05       -0.02     0.12
    0.97_*  0.00   0.00  
134   -0.03  -0.02      0.01    -0.08     -0.03     -0.02        0.11        0.03    -0.14
    1.02_*  0.00   0.02_*
```

150	0.00	0.00	0.03	0.00	0.03	-0.01	-0.01	0.06	0.08	1.01_*	0.00	0.01
179	0.13	-0.15	-0.06	-0.11	-0.13	0.05	0.09	0.05	0.27_*	0.91_*	0.01	0.00
197	0.07	0.07	-0.03	-0.01	-0.05	-0.04	0.01	0.00	0.12	0.97_*	0.00	0.00
199	0.00	-0.01	-0.04	0.00	0.01	-0.07	0.03	0.04	-0.09	1.01_*	0.00	0.01_*
200	0.05	0.07	-0.02	-0.05	-0.01	-0.02	-0.03	0.02	0.10	0.99_*	0.00	0.00
204	0.00	-0.03	-0.06	0.00	0.07	0.10	0.05	-0.06	-0.17	1.01	0.00	0.01_*
209	0.00	0.00	0.02	0.00	0.00	0.02	-0.01	-0.02	0.04	1.01_*	0.00	0.01
210	0.00	0.00	0.00	-0.01	0.00	0.00	-0.01	0.00	-0.01	1.01_*	0.00	0.01
214	0.00	0.01	-0.02	0.00	0.01	-0.02	-0.02	-0.04	-0.06	1.01_*	0.00	0.01_*
216	0.00	-0.01	0.04	0.00	0.03	0.01	0.02	0.06	0.09	1.01_*	0.00	0.01
236	0.09	-0.05	-0.04	-0.01	-0.01	0.03	-0.01	0.00	0.11	0.96_*	0.00	0.00
241	0.00	0.01	0.03	0.00	0.01	-0.02	-0.02	0.05	0.07	1.01_*	0.00	0.01
243	0.00	0.01	-0.02	0.00	0.01	-0.02	-0.02	-0.03	-0.06	1.02_*	0.00	0.01_*
246	0.08	-0.02	-0.03	0.06	0.04	0.03	-0.06	-0.03	0.13	0.97_*	0.00	0.00
247	0.09	-0.02	-0.04	0.06	-0.03	0.03	-0.07	-0.03	0.14	0.96_*	0.00	0.00
259	0.09	0.03	-0.03	0.19	0.00	0.04	-0.19	-0.09	0.29_*	0.97_*	0.01	0.01
272	0.11	-0.13	-0.05	-0.09	-0.04	0.05	0.07	0.04	0.20_*	0.94_*	0.00	0.00
280	0.05	-0.03	-0.02	-0.05	0.04	0.01	0.01	0.02	0.09	0.99_*	0.00	0.00
281	0.09	-0.17	-0.04	-0.08	0.12	0.06	0.08	0.04	0.24_*	0.95_*	0.01	0.00
282	0.07	0.05	-0.03	-0.06	-0.04	-0.02	-0.02	0.03	0.11	0.97_*	0.00	0.00
304	0.00	0.00	0.02	0.00	0.00	0.02	-0.01	-0.02	0.04	1.01_*	0.00	0.01
308	0.07	0.02	-0.03	0.05	-0.06	-0.03	0.05	-0.02	0.13	0.98_*	0.00	0.00
315	0.10	-0.15	-0.05	-0.09	0.02	0.06	0.08	0.04	0.21_*	0.94_*	0.01	0.00
327	-0.02	-0.02	0.01	-0.06	0.01	-0.01	0.06	0.03	-0.09	1.02_*	0.00	0.01_*
331	0.00	-0.01	-0.04	0.00	0.01	-0.07	0.03	0.04	-0.10	1.01_*	0.00	0.01_*
332	0.00	0.01	0.02	0.00	0.00	0.03	-0.01	-0.02	0.04	1.02_*	0.00	0.01_*
334	-0.01	0.00	0.00	-0.03	0.01	0.00	-0.01	0.01	-0.03	1.01_*	0.00	0.01
335	0.06	-0.11	-0.03	-0.05	-0.01	0.04	0.06	0.02	0.14	0.98_*	0.00	0.00
337	0.07	0.00	-0.03	0.11	0.01	0.03	-0.11	-0.05	0.18	0.98_*	0.00	0.00
338	0.13	-0.21	-0.06	-0.11	0.06	0.08	0.11	0.05	0.28_*	0.91_*	0.01	0.00
343	-0.02	0.02	0.01	-0.06	-0.04	0.01	-0.05	0.02	-0.09	1.02_*	0.00	0.01_*
367	0.10	-0.05	-0.05	-0.01	-0.09	0.03	-0.01	0.00	0.16	0.94_*	0.00	0.00
372	0.06	0.04	-0.02	-0.05	0.02	-0.01	-0.02	0.02	0.08	0.99_*	0.00	0.00
373	0.00	0.01	-0.02	0.00	0.01	-0.02	-0.02	-0.03	-0.06	1.02_*	0.00	0.01_*
379	-0.01	0.00	0.01	-0.02	-0.02	0.01	-0.04	0.01	-0.05	1.01_*	0.00	0.01
409	0.13	-0.13	-0.06	0.00	0.10	0.08	-0.04	0.00	0.24_*	0.91_*	0.01	0.00

```
413   0.07   0.06     -0.03    -0.06     -0.04     -0.02      -0.03       0.02     0.11
    0.98_*  0.00   0.00
418   0.08  -0.15     -0.04    -0.07      0.05      0.06       0.07       0.03     0.19
_*  0.96_*  0.00   0.00
419   0.12  -0.06     -0.05     0.09      0.02      0.08      -0.16      -0.04     0.26
_*  0.93_*  0.01   0.00
428   0.06   0.07     -0.03    -0.05      0.00     -0.03      -0.04       0.02     0.11
    0.98_*  0.00   0.00
432   0.00   0.01      0.06     0.00      0.07      0.07      -0.03      -0.06     0.14
    1.01    0.00   0.01_*
438   0.09  -0.12     -0.04    -0.08      0.15      0.04       0.05       0.04     0.22
_*  0.95_*  0.01   0.00
447  -0.01   0.01      0.00    -0.03      0.00      0.00      -0.02       0.01    -0.04
    1.01_*  0.00   0.01
472   0.06   0.02     -0.03     0.05      0.05     -0.03       0.05      -0.02     0.12
    0.98_*  0.00   0.00
475  -0.01   0.00      0.01    -0.02     -0.01      0.01      -0.03       0.01    -0.04
    1.01_*  0.00   0.01
479   0.00  -0.02     -0.04     0.00      0.03     -0.06       0.02       0.05    -0.10
    1.01_*  0.00   0.01_*
506  -0.01   0.01      0.00    -0.02      0.00      0.01      -0.03       0.01    -0.04
    1.01_*  0.00   0.01
520   0.06   0.03     -0.03    -0.01      0.04     -0.02       0.00       0.00     0.08
    0.98_*  0.00   0.00
521   0.00   0.00      0.00     0.00      0.00      0.00      -0.01       0.00    -0.01
    1.01_*  0.00   0.01
525   0.00  -0.02     -0.05     0.00      0.03     -0.08       0.03       0.05    -0.12
    1.01_*  0.00   0.01_*
529   0.00   0.01      0.02     0.00      0.00      0.03      -0.01      -0.02     0.04
    1.02_*  0.00   0.01_*
535   0.00  -0.03      0.04     0.00      0.00      0.06       0.06       0.07     0.13
    1.02_*  0.00   0.02_*
544   0.07   0.04     -0.03    -0.01     -0.07     -0.02       0.01       0.00     0.11
    0.97_*  0.00   0.00
551   0.05   0.02     -0.03     0.04      0.00     -0.03       0.05      -0.02     0.10
    0.99_*  0.00   0.00
574   0.00   0.01      0.03     0.00      0.02     -0.02      -0.02       0.05     0.08
    1.01_*  0.00   0.01_*
586   0.00   0.01      0.03     0.00      0.01     -0.03      -0.03       0.05     0.07
    1.02_*  0.00   0.01_*
599   0.06   0.02     -0.02     0.14      0.04      0.04      -0.16      -0.06     0.23
_*  0.99    0.01   0.01
638   0.08   0.00     -0.04    -0.01      0.12      0.00      -0.01       0.00     0.15
    0.97_*  0.00   0.00
678   0.00   0.01     -0.02     0.00      0.01     -0.01      -0.02      -0.04    -0.06
    1.01_*  0.00   0.01
681   0.09  -0.07     -0.04     0.00      0.03      0.04      -0.02       0.00     0.13
    0.96_*  0.00   0.00
702   0.00  -0.02      0.04     0.00      0.02      0.03       0.03       0.06     0.10
    1.01_*  0.00   0.01_*
708   0.00  -0.02     -0.04     0.00      0.01     -0.07       0.03       0.04    -0.10
    1.01_*  0.00   0.01_*
718   0.00   0.00      0.02     0.00      0.01      0.03      -0.01      -0.02     0.04
    1.02_*  0.00   0.01_*
738  -0.01   0.00      0.00    -0.03      0.01      0.00      -0.02       0.01    -0.04
    1.01_*  0.00   0.01
742   0.08   0.01     -0.04     0.06     -0.02     -0.01       0.03      -0.02     0.11
    0.96_*  0.00   0.00
749   0.00  -0.02     -0.04     0.00      0.00      0.05       0.04      -0.07    -0.11
    1.01_*  0.00   0.01_*
755   0.07  -0.05     -0.03     0.00     -0.03      0.03      -0.01       0.00     0.11
    0.97_*  0.00   0.00
774   0.08  -0.05     -0.04     0.00      0.01      0.03      -0.02       0.00     0.11
    0.96_*  0.00   0.00
775  -0.02  -0.02      0.01    -0.06      0.00     -0.01       0.08       0.03    -0.10
    1.02_*  0.00   0.02_*
 [ reached getOption("max.print") -- omitted 179 rows ]
            dfb.1_     dfb.scl.g      dfb.smkr     dfb.scl.c     dfb.scl.b  dfb.scl.g:s
dfb.scl.g:c      dfb.sm:.
9    -8.116959e-04 -0.0153310499 -0.0430310469 -1.711515e-04  1.585965e-02 -0.069887731
0.026644714  0.0453850866
```

```
28   -2.119520e-02 -0.0220684844  0.0067638079 -5.892297e-02 -8.800001e-04 -0.015052230
0.087694011  0.0272016053
31   -4.309640e-03  0.0006581464  0.0023837427 -7.006837e-03 -2.004205e-04  0.003765473 -
0.013494990  0.0026796196
34    3.056864e-04  0.0046889750  0.0184020001  1.647149e-04  1.219314e-02  0.027951224 -
0.011221824 -0.0189604202
51    6.849885e-04  0.0118254930  0.0279384640  2.472458e-04  5.244478e-03 -0.044828875 -
0.023703014  0.0263462275
61   -1.487606e-02 -0.0065235107  0.0058648383 -4.420180e-02  2.365423e-03 -0.004909685
0.027220896  0.0195578395
74    2.966364e-04  0.0052509157  0.0196065952  9.506898e-05  9.656962e-05  0.025901235 -
0.010122530 -0.0202441937
81   -9.334895e-04 -0.0164922592  0.0370337966 -3.021247e-04 -8.385373e-04  0.061618693
0.031889683  0.0227464124
86    1.250803e-01 -0.2184382367 -0.0602013725 -1.087042e-01  4.294492e-03  0.080620263
0.112459952  0.0491289296
87    2.744100e-04  0.0048112013 -0.0062284501  9.222258e-05  8.642719e-04 -0.008811614 -
0.009414720 -0.0096389736
92    3.093699e-04  0.0051961652  0.0163279668  1.250437e-04  4.792404e-03  0.027683825 -
0.010863711 -0.0170801984
115   9.663620e-02 -0.0042439693 -0.0431402309  7.017116e-02 -1.095347e-01  0.014805539 -
0.032387234 -0.0323605431
118   7.513013e-02 -0.0180794859 -0.0322915019  5.523448e-02 -2.622495e-03  0.025796117 -
0.053288429 -0.0248805519
134  -2.704801e-02 -0.0248270293  0.0086310759 -7.580610e-02 -2.928140e-02 -0.019164007
0.108289396  0.0344749300
150   1.752226e-04  0.0011367975  0.0327591486  2.377685e-04  3.296323e-02 -0.009766295 -
0.008130186  0.0567018550
179   1.265976e-01 -0.1469033011 -0.0605724820 -1.114440e-01 -1.290778e-01  0.054504911
0.085442172  0.0477043915
197   7.091949e-02  0.0658607673 -0.0327850387 -8.244559e-03 -5.092550e-02 -0.037956229
0.014985272  0.0031184098
199  -7.523341e-04 -0.0139308161 -0.0387088514 -1.844245e-04  1.002692e-02 -0.065288863
0.025001535  0.0408701861
200   5.162325e-02  0.0654324591 -0.0223136282 -4.686456e-02 -1.026387e-02 -0.023890167 -
0.033807826  0.0190970181
204  -1.610196e-03 -0.0326153460 -0.0626828039 -1.359511e-04  6.834627e-02  0.103381069
0.050445389 -0.0586145156
209   2.742431e-04  0.0046352819  0.0192156961  1.081558e-04  3.760848e-03  0.024241313 -
0.009598355 -0.0197151667
210  -2.986018e-03  0.0011608414  0.0015870462 -7.331486e-03  3.454508e-03  0.001397388 -
0.006880586  0.0030254528
214   5.753877e-04  0.0094524209 -0.0219956919  2.521383e-04  1.245975e-02 -0.017477796 -
0.020436896 -0.0354158262
216  -5.373119e-04 -0.0115936858  0.0354879881  2.027210e-05  3.469967e-02  0.014621774
0.016055943  0.0587998364
236   8.983149e-02 -0.0465567973 -0.0402274452 -5.138778e-03 -6.766843e-03  0.028914321 -
0.013953799  0.0017735425
241   4.546387e-04  0.0074030449  0.0277170782  2.052998e-04  1.094562e-02 -0.018874260 -
0.016220015  0.0490015079
243   5.757010e-04  0.0093267882 -0.0205172371  2.643633e-04  1.465670e-02 -0.017655663 -
0.020591180 -0.0328809429
246   7.725579e-02 -0.0214922470 -0.0329911271  5.702747e-02  3.671922e-02  0.027549625 -
0.056354504 -0.0251366857
247   8.588067e-02 -0.0207005230 -0.0368741313  6.302978e-02 -2.915598e-02  0.031541054 -
0.065471921 -0.0288863892
259   8.756288e-02  0.0326418539 -0.0333285904  1.930121e-01  3.819259e-03  0.042066347 -
0.194984766 -0.0869541595
272   1.053189e-01 -0.1272735514 -0.0500241206 -9.239952e-02 -3.932062e-02  0.047117183
0.068460831  0.0404973056
280   5.358445e-02 -0.0285197785 -0.0244626955 -4.725735e-02  3.721867e-02  0.010582894
0.011509121  0.0210151257
281   9.383822e-02 -0.1723553812 -0.0445534418 -8.103090e-02  1.166121e-01  0.063454335
0.080028330  0.0382093651
282   7.197488e-02  0.0456169026 -0.0320218810 -6.484691e-02 -3.621210e-02 -0.016518086 -
0.021652792  0.0266810861
304   2.633477e-04  0.0047005006  0.0191303390  8.080980e-05 -5.648243e-04  0.022840239 -
0.008944050 -0.0196419886
308   6.700367e-02  0.0223193030 -0.0322524262  4.814102e-02 -6.144431e-02 -0.026083218
0.052120324 -0.0202781460
315   1.014960e-01 -0.1523195643 -0.0483540145 -8.847846e-02  1.515566e-02  0.056241686
0.077349199  0.0398318235
```

```
327  -2.111501e-02 -0.0168122555  0.0075686555 -6.054132e-02  1.389655e-02 -0.010814406
0.063475539  0.0275511346
331  -7.864221e-04 -0.0146985555 -0.0403574839 -1.801607e-04  1.276787e-02 -0.068032617
0.025984885  0.0426517664
332   3.102929e-04  0.0053094713  0.0157572683  1.163771e-04  3.168811e-03  0.027641758 -
0.010789068 -0.0165567905
334  -8.859422e-03  0.0032313284  0.0044589671 -2.866830e-02  9.336394e-03  0.001955772 -
0.012774977  0.0120958077
335   6.137649e-02 -0.1065424569 -0.0295870774 -5.338048e-02 -6.512814e-03  0.039334242
0.055513814  0.0240046545
337   7.364342e-02  0.0020671772 -0.0297988577  1.098439e-01  7.824616e-03  0.031727330 -
0.108467292 -0.0493446188
338   1.259980e-01 -0.2134471725 -0.0601681284 -1.093841e-01  5.914036e-02  0.078718639
0.105609306  0.0500711942
343  -1.641744e-02  0.0185066366  0.0089731318 -5.618715e-02 -3.832173e-02  0.007314065 -
0.052913340  0.0222250675
367   1.048953e-01 -0.0539814413 -0.0473600157 -6.197594e-03 -8.728543e-02  0.034992024 -
0.014525704  0.0010233773
372   5.515112e-02  0.0361540305 -0.0242125188 -4.953675e-02  1.804283e-02 -0.013156400 -
0.020756475  0.0209593952
373   6.456707e-04  0.0107913129 -0.0214028450  2.659033e-04  1.089541e-02 -0.019979548 -
0.022731514 -0.0341576788
379  -1.088416e-02  0.0028017763  0.0059967409 -1.784052e-02 -1.508193e-02  0.009622150 -
0.035234150  0.0065872501
409   1.282101e-01 -0.1346162020 -0.0563306499 -4.362247e-03  1.035824e-01  0.080121944 -
0.038938385  0.0024935986
413   6.573649e-02  0.0635403702 -0.0289044038 -5.951535e-02 -3.689360e-02 -0.023121584 -
0.030891812  0.0241783626
418   8.411811e-02 -0.1504050966 -0.0402337493 -7.288951e-02  4.973445e-02  0.055445659
0.073836994  0.0335982764
419   1.168745e-01 -0.0560474975 -0.0477016637  8.666414e-02  2.136467e-02  0.077125280 -
0.158254408 -0.0403892074
428   6.053264e-02  0.0709075060 -0.0261823062 -5.483545e-02  5.503674e-06 -0.025890586 -
0.037538742  0.0225693379
432   6.992872e-04  0.0080088225  0.0584136142  6.279873e-04  7.340542e-02  0.067452834 -
0.028645803 -0.0584721717
438   9.057196e-02 -0.1153999051 -0.0419053411 -7.871793e-02  1.489163e-01  0.042472503
0.047853937  0.0369533637
447  -8.966674e-03  0.0065364472  0.0047679377 -2.988894e-02 -1.608507e-04  0.003286898 -
0.022090455  0.0122616482
472   6.197033e-02  0.0156634667 -0.0295212426  4.510388e-02  5.495476e-02 -0.025752178
0.053057932 -0.0172387461
475  -9.210139e-03  0.0021729973  0.0051393030 -1.511599e-02 -9.284110e-03  0.008636742 -
0.031381228  0.0055975381
479  -7.617638e-04 -0.0150256519 -0.0447982846 -1.016818e-04  2.556359e-02 -0.064325888
0.024307603  0.0469910481
506  -8.854269e-03  0.0061395914  0.0049484332 -2.237816e-02 -1.820737e-03  0.005789025 -
0.029490887  0.0088555707
520   6.093999e-02  0.0268881695 -0.0275243255 -5.706281e-03  4.185120e-02 -0.016350532
0.003822138  0.0031457224
521  -1.514646e-03  0.0010246382  0.0008638999 -3.838942e-03  9.044908e-04  0.001090343 -
0.005428352  0.0015285918
525  -9.427196e-04 -0.0184072057 -0.0461698421 -1.431910e-04  2.849167e-02 -0.080402283
0.030287369  0.0491808112
529   3.246517e-04  0.0056859835  0.0175655804  1.096714e-04  1.124664e-03  0.028652623 -
0.011145136 -0.0183923525
535  -1.811721e-03 -0.0320645355  0.0429425795 -5.811635e-04 -6.849687e-04  0.057969663
0.061830053  0.0667699545
544   7.031224e-02  0.0360665915 -0.0323844886 -7.062893e-03 -7.319780e-02 -0.019661199
0.008203418  0.0021273590
551   5.405812e-02  0.0182284777 -0.0261139638  3.906165e-02  4.807208e-03 -0.025523446
0.051956796 -0.0154224549
574   5.408096e-04  0.0081576078  0.0297522864  3.041592e-04  2.388209e-02 -0.022095191 -
0.020004269  0.0528452836
586   8.571793e-04  0.0148743146  0.0253108565  3.023589e-04  5.287441e-03 -0.032552901 -
0.029578061  0.0463190539
599   6.292289e-02  0.0245497455 -0.0231398901  1.378292e-01  3.764205e-02  0.035074910 -
0.160773931 -0.0620762543
638   8.021479e-02 -0.0044281288 -0.0355815319 -5.767571e-03  1.245063e-01  0.001114705 -
0.006150235  0.0041621872
678   4.771966e-04  0.0076148157 -0.0237065650  2.298632e-04  1.409366e-02 -0.013883287 -
0.017195074 -0.0386827053
```

```
681   9.024123e-02 -0.0682699704 -0.0400654650 -4.223024e-03  2.971391e-02  0.041286513 -
0.020034199  0.0017803356
702  -9.904082e-04 -0.0187525069  0.0398729396 -2.045828e-04  2.012180e-02  0.029839573
0.032460770  0.0645852929
708  -8.119028e-04 -0.0151659086 -0.0402771599 -1.868199e-04  1.303263e-02 -0.070325114
0.026836550  0.0426946794
718   3.080600e-04  0.0046370178  0.0186199939  1.741613e-04  1.376767e-02  0.028306677 -
0.011405685 -0.0191598301
738  -9.247905e-03  0.0041521676  0.0047216817 -3.014062e-02  7.900257e-03  0.002384897 -
0.015660380  0.0126371217
742   8.080034e-02  0.0104011008 -0.0374538099  5.867168e-02 -1.578875e-02 -0.012891644
0.025662333 -0.0247170535
749  -1.203775e-03 -0.0212832483 -0.0405675693 -3.881461e-04 -8.173676e-04  0.046276456
0.041105882 -0.0735909377
755   7.279158e-02 -0.0546828433 -0.0325927533 -3.550472e-03 -3.219585e-02  0.034103342 -
0.014881231  0.0006975380
774   8.151728e-02 -0.0516799376 -0.0363698052 -4.259256e-03  6.340615e-03  0.031672843 -
0.015227204  0.0015617785
775  -2.079490e-02 -0.0203945144  0.0067348390 -5.814707e-02 -4.518996e-03 -0.014261603
0.082497880  0.0266959536
            dffit     cov.r       cook.d          hat
9    -0.103897294 1.0129051 1.349411e-03 0.0122721290
28   -0.109883895 1.0228174 1.509603e-03 0.0207826988
31   -0.017752627 1.0114553 3.941158e-05 0.0078849952
34    0.043424322 1.0149340 2.357986e-04 0.0117127065
51    0.070407265 1.0117288 6.197878e-04 0.0098057490
61   -0.054229979 1.0112233 3.677220e-04 0.0087153600
74    0.041519274 1.0122440 2.155606e-04 0.0091827841
81    0.091888548 1.0129011 1.055575e-03 0.0117099814
86    0.276929492 0.9087993 9.469135e-03 0.0026971940
87   -0.019227159 1.0207456 4.623091e-05 0.0168503949
92    0.039992379 1.0161901 2.000031e-04 0.0128063194
115   0.168890763 0.9495347 3.541711e-03 0.0017947576
118   0.118028713 0.9717438 1.734796e-03 0.0014872921
134  -0.139561689 1.0209208 2.434748e-03 0.0202824588
150   0.080757095 1.0114052 8.153443e-04 0.0100069451
179   0.265610492 0.9070612 8.709081e-03 0.0024418704
197   0.120598615 0.9724634 1.811310e-03 0.0015828874
199  -0.094958110 1.0140391 1.127285e-03 0.0127647473
200   0.095053216 0.9888195 1.127529e-03 0.0019398028
204  -0.172780738 1.0060326 3.729031e-03 0.0115601999
209   0.039863899 1.0118522 1.987150e-04 0.0087757147
210  -0.011386244 1.0108981 1.621302e-05 0.0072593489
214  -0.057075911 1.0142692 4.073400e-04 0.0114976101
216   0.087799515 1.0115178 9.637113e-04 0.0104265723
236   0.110786578 0.9553495 1.525305e-03 0.0008814991
241   0.068542710 1.0118505 5.874027e-04 0.0098273394
243  -0.055060504 1.0152475 3.790871e-04 0.0123179780
246   0.126818238 0.9699756 2.002306e-03 0.0016225868
247   0.142939885 0.9621987 2.541170e-03 0.0016791667
259   0.294418605 0.9708379 1.078499e-02 0.0076012662
272   0.197176504 0.9357621 4.818484e-03 0.0019424359
280   0.085242026 0.9865746 9.065817e-04 0.0014137751
281   0.238395806 0.9508200 7.056295e-03 0.0035446859
282   0.111460722 0.9724532 1.547260e-03 0.0013611127
304   0.038629730 1.0113045 1.866012e-04 0.0082394276
308   0.127067692 0.9757424 2.011602e-03 0.0019329409
315   0.205294413 0.9406166 5.226586e-03 0.0022581010
327  -0.092795401 1.0153925 1.076560e-03 0.0137897835
331  -0.099305634 1.0138867 1.232836e-03 0.0128345850
332   0.039332607 1.0168049 1.934592e-04 0.0133706403
334  -0.034215191 1.0111084 1.463914e-04 0.0079281504
335   0.136013344 0.9818297 2.306379e-03 0.0027027423
337   0.178995388 0.9770032 3.991408e-03 0.0037607685
338   0.276962481 0.9074454 9.469671e-03 0.0026597760
343  -0.090636968 1.0161818 1.027085e-03 0.0143691758
367   0.159463845 0.9388344 3.153081e-03 0.0013439479
372   0.084699805 0.9857140 8.949958e-04 0.0013383661
373  -0.057522862 1.0152849 4.137478e-04 0.0124254805
379  -0.049002104 1.0121172 3.002518e-04 0.0093080170
409   0.238498632 0.9100462 7.025120e-03 0.0020427155
413   0.113003240 0.9782084 1.591491e-03 0.0016845174
```

```
418   0.192820913 0.9609907 4.622758e-03 0.0028887275
419   0.263821970 0.9320839 8.620607e-03 0.0032363046
428   0.107512896 0.9824487 1.441358e-03 0.0018024386
432   0.135846963 1.0087476 2.306116e-03 0.0110207854
438   0.219525515 0.9546171 5.986624e-03 0.0032489773
447  -0.039121670 1.0126306 1.913858e-04 0.0094682295
472   0.124921126 0.9798550 1.945181e-03 0.0021532395
475  -0.041983675 1.0125265 2.204099e-04 0.0094544654
479  -0.103033190 1.0111536 1.327002e-03 0.0108995094
506  -0.040265283 1.0127912 2.027380e-04 0.0096479112
520   0.084141868 0.9807114 8.827130e-04 0.0010578896
521  -0.007239635 1.0142860 6.554488e-06 0.0105294990
525  -0.118879447 1.0143366 1.766549e-03 0.0140933958
529   0.041825462 1.0152790 2.187558e-04 0.0119948462
535   0.129332000 1.0161656 2.090819e-03 0.0160085529
544   0.109479921 0.9733611 1.492931e-03 0.0013519647
551   0.102192247 0.9856275 1.302739e-03 0.0018851873
574   0.077716409 1.0127685 7.551357e-04 0.0109781192
586   0.074909619 1.0160296 7.016200e-04 0.0136571311
599   0.228945473 0.9921403 6.538075e-03 0.0090924549
638   0.148416477 0.9650944 2.740568e-03 0.0019314271
678  -0.058513958 1.0128571 4.281153e-04 0.0102951570
681   0.130880642 0.9556540 2.128784e-03 0.0012291155
702   0.100436569 1.0113724 1.260988e-03 0.0109321131
708  -0.101355945 1.0146081 1.284273e-03 0.0135071779
718   0.044343723 1.0151545 2.458888e-04 0.0119402211
738  -0.036501629 1.0112741 1.666092e-04 0.0081470680
742   0.108576542 0.9629400 1.466484e-03 0.0010027122
749  -0.113904744 1.0123696 1.621759e-03 0.0123520471
755   0.111150289 0.9727746 1.538716e-03 0.0013670398
774   0.107587000 0.9641259 1.440095e-03 0.0010137566
775  -0.104936062 1.0214501 1.376724e-03 0.0194256482
 [ reached getOption("max.print") -- omitted 179 rows ]
> press(sub_model2)
[1] 5030.085
```

6.8 GLM Model Diagnostics

```
summary(glm_model)

Call:
glm(formula = boxcox.charges ~ scale.age * smoker * scale.children +
    scale.bmi, family = gaussian(link = "identity"), data = trainData)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6793  -0.7253  -0.3032   0.1126   8.0565

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                       17.52550    0.03568 491.203  < 2e-16 ***
scale.age                          1.85695    0.03738  49.676  < 2e-16 ***
smokeryes                          5.81938    0.07905  73.617  < 2e-16 ***
scale.children                     0.52178    0.03522  14.816  < 2e-16 ***
scale.bmi                          0.24511    0.03204   7.651 2.96e-14 ***
scale.age:smokeryes               -1.25868    0.08175 -15.397  < 2e-16 ***
scale.age:scale.children          -0.27959    0.03835  -7.291 4.25e-13 ***
smokeryes:scale.children          -0.50099    0.08248  -6.074 1.46e-09 ***
scale.age:smokeryes:scale.children  0.34113    0.08794   3.879 0.000108 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 2.244142)

    Null deviance: 24893.8  on 2219  degrees of freedom
Residual deviance:  4961.8  on 2211  degrees of freedom
AIC: 8105.5

Number of Fisher Scoring iterations: 2

> # Deviance of the model
> deviance(glm_model)
[1] 4961.797
> # p-value for the deviance test
> pchisq(glm_model$deviance, df=glm_model$df.residual, lower.tail=FALSE)
[1] 5.51702e-212
> # Pearson's Chi-Square Test
> pchisq(sum(resid(glm_model, type="pearson")^2), glm_model$df.residual, lower.tail = FAL
SE)
[1] 5.51702e-212
> # Hosmer-Lemeshow Test
> hoslem.test(trainData$boxcox.charges, fitted(glm_model))

        Hosmer and Lemeshow goodness of fit (GOF) test

data:  trainData$boxcox.charges, fitted(glm_model)
X-squared = -0.24787, df = 8, p-value = 1

> #Anova Test
> Anova(glm_model, type = 3)
Analysis of Deviance Table (Type III tests)

Response: boxcox.charges
                               LR Chisq Df Pr(>Chisq)
scale.age                        2467.7  1  < 2.2e-16 ***
smoker                           5419.4  1  < 2.2e-16 ***
scale.children                    219.5  1  < 2.2e-16 ***
scale.bmi                          58.5  1  1.996e-14 ***
scale.age:smoker                  237.1  1  < 2.2e-16 ***
scale.age:scale.children           53.2  1  3.069e-13 ***
smoker:scale.children              36.9  1  1.247e-09 ***
scale.age:smoker:scale.children    15.0  1  0.0001049 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(glm_model, test = "Chisq")
Analysis of Deviance Table

Model: gaussian, link: identity
```

```
Response: boxcox.charges

Terms added sequentially (first to last)


                                Df Deviance Resid. Df Resid. Dev  Pr(>Chi)    
NULL                                             2219    24893.8              
scale.age                        1   6054.0      2218    18839.8 < 2.2e-16 ***
smoker                           1  12469.2      2217     6370.6 < 2.2e-16 ***
scale.children                   1    376.3      2216     5994.3 < 2.2e-16 ***
scale.bmi                        1    143.5      2215     5850.8 1.290e-15 ***
scale.age:smoker                 1    685.0      2214     5165.8 < 2.2e-16 ***
scale.age:scale.children         1     82.4      2213     5083.4 1.371e-09 ***
smoker:scale.children            1     87.9      2212     4995.6 3.911e-10 ***
scale.age:smoker:scale.children  1     33.8      2211     4961.8 0.0001049 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> #Residual Normality
> shapiro.test(resid(glm_model, type="deviance"))

	Shapiro-Wilk normality test

data:  resid(glm_model, type = "deviance")
W = 0.74418, p-value < 2.2e-16
```

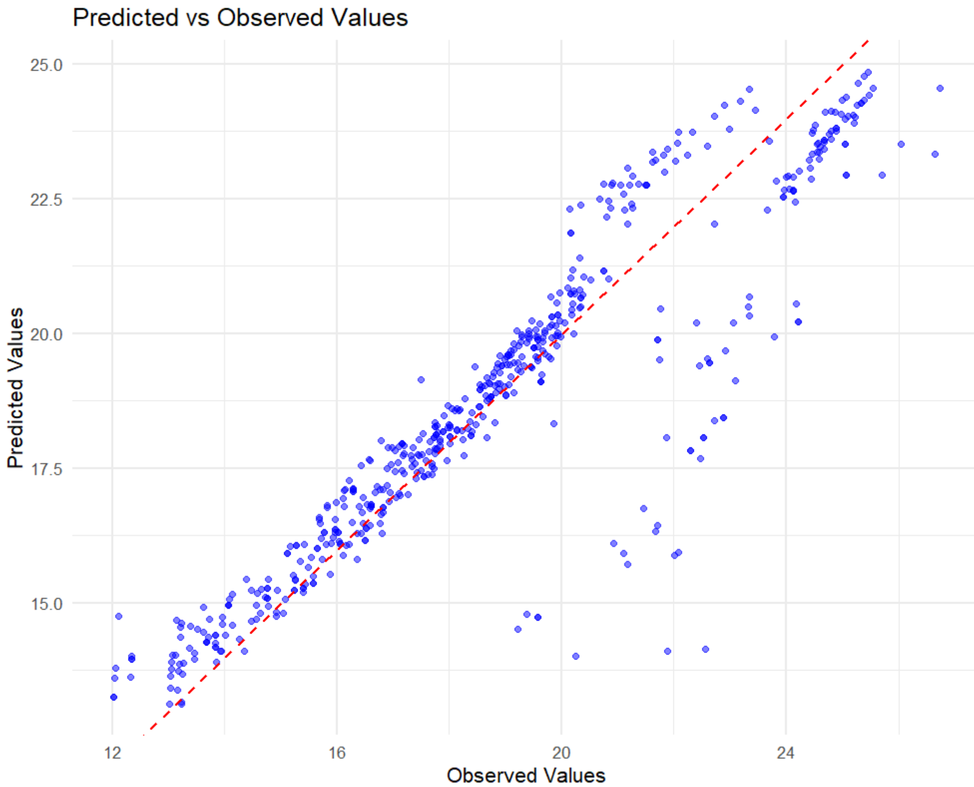
6.9 Model Validation

6.9.1 Performance Metrics

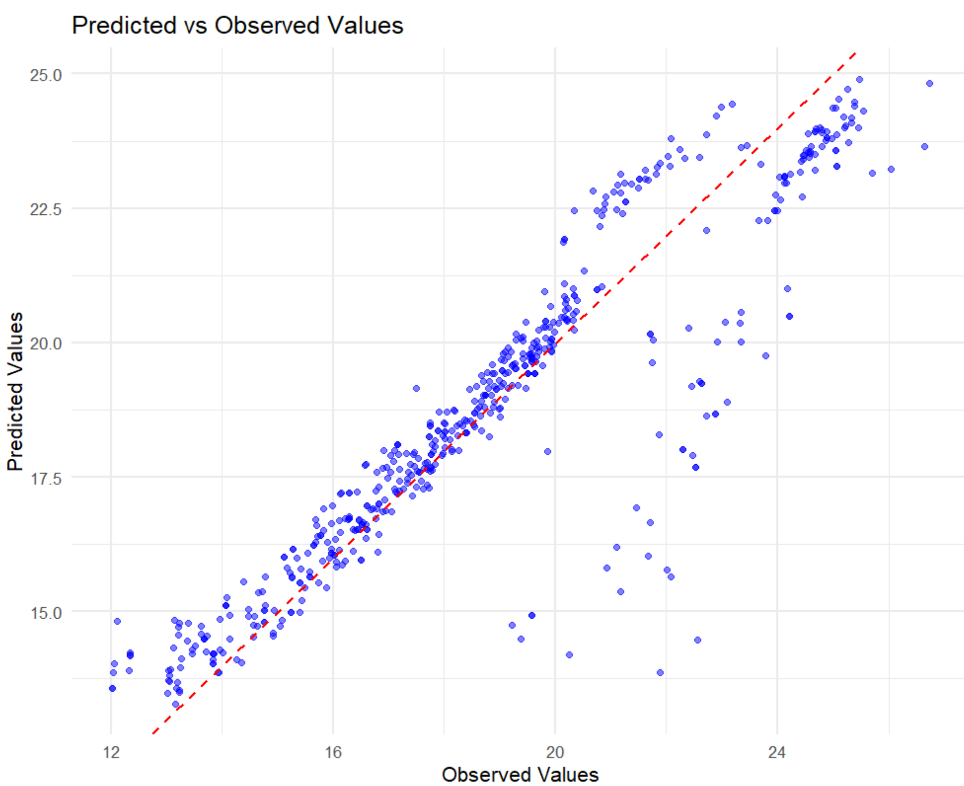
```
performance_metrics(sub_model1, testData, "boxcox.charges")
                                 Metric     Value
1             Mean Absolute Error (MAE) 0.9035586
2              Mean Squared Error (MSE) 2.1410181
3        Root Mean Squared Error (RMSE) 1.4632218
4                        R-squared (R2) 0.8143815
5 Mean Absolute Percentage Error (MAPE) 0.0458562
> performance_metrics(sub_model2, testData, "boxcox.charges")
                                 Metric      Value
1             Mean Absolute Error (MAE) 0.90302064
2              Mean Squared Error (MSE) 2.15054628
3        Root Mean Squared Error (RMSE) 1.46647410
4                        R-squared (R2) 0.81343767
5 Mean Absolute Percentage Error (MAPE) 0.04608643
> performance_metrics(model4, testData, "boxcox.charges")
                                 Metric      Value
1             Mean Absolute Error (MAE) 0.89103013
2              Mean Squared Error (MSE) 2.12220327
3        Root Mean Squared Error (RMSE) 1.45677839
4                        R-squared (R2) 0.81607205
5 Mean Absolute Percentage Error (MAPE) 0.04505584
> performance_metrics(glm_model, testData, "boxcox.charges")
                                 Metric      Value
1             Mean Absolute Error (MAE) 0.90302064
2              Mean Squared Error (MSE) 2.15054628
3        Root Mean Squared Error (RMSE) 1.46647410
4                        R-squared (R2) 0.81343767
5 Mean Absolute Percentage Error (MAPE) 0.04608643
```

6.9.2 Predicted vs Observed Values Scatterplots

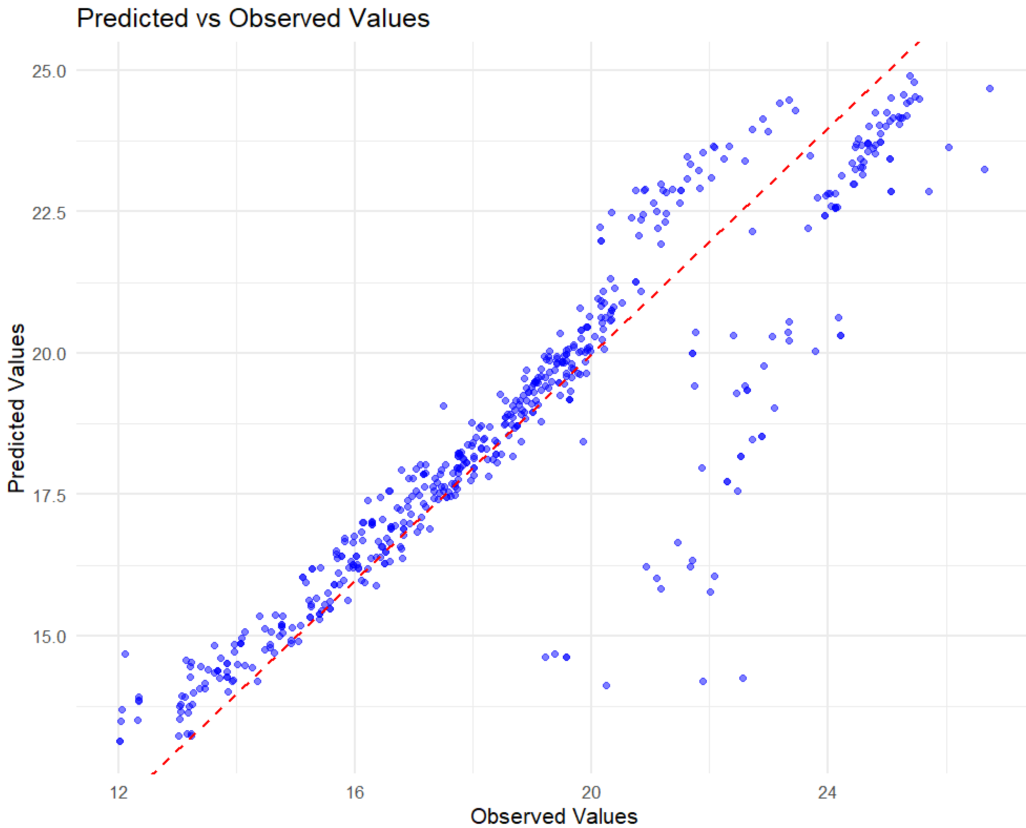
6.9.2.1 Submodel 1



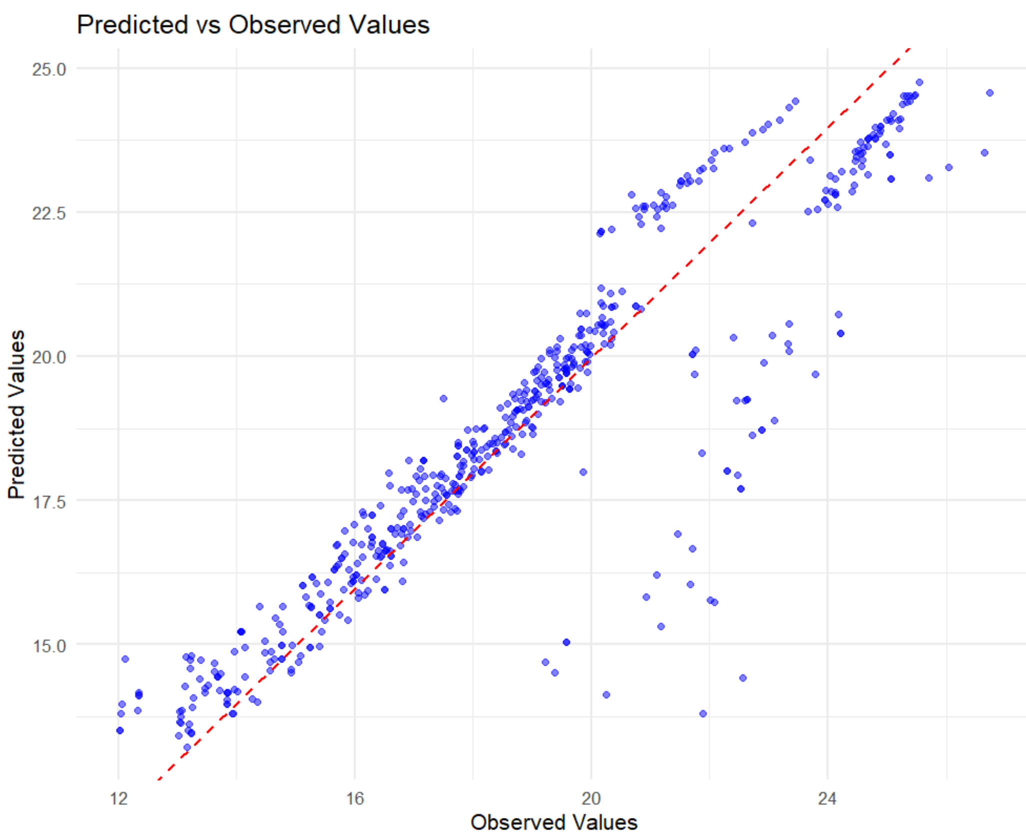
6.9.2.2 Submodel 2



6.9.2.3 Model 4



6.9.2.4 GLM Model



7 Codes

```
#Load the libraries
library(readr)
library(dplyr)
library(magrittr)
library(ggplot2)
library(tidyr)
library(GGally)
library(car)
library(caret)
library(leaps)
library(MASS)
library(DAAG)
library(Metrics)
library(ResourceSelection)

setwd("C:/Users/Asus/Desktop/Regression Analysis/Project")

#Read in the dataset
insurance <- read_csv("medical_insurance.csv")

####Exploratory Data Analysis####

#Dimensions of the dataset
dim(insurance)
#Head of the dataset
head(insurance)
#Structure of the dataset
str(insurance)
#Datatypes of the features
lapply(insurance, class)

#Change the datatype of categorical data
insurance$sex <- as.factor(insurance$sex)
insurance$smoker <- as.factor(insurance$smoker)
insurance$region <- as.factor(insurance$region)

#Check for missing values in the dataset
any(is.na(insurance))
```

```
###Descriptive Statistics###

#Numerical Variables
insurance %>% dplyr::select(age, bmi, charges) %>% summary()
#Categorical Variables
ftable(table(insurance$sex, insurance$smoker, insurance$region))

table(insurance$sex)
table(insurance$smoker)
table(insurance$region)

#Check for Correlation in the Variables
ggpairs(insurance)

# Histogram for age
ggplot(insurance, aes(x = age)) +
  geom_histogram(binwidth = 1, fill = "skyblue", color = "black") +
  labs(title = "Distribution of Age", x = "Age", y = "Count") +
  theme_minimal()

# Histogram for bmi
ggplot(insurance, aes(x = bmi)) +
  geom_histogram(binwidth = 1, fill = "lightgreen", color = "black") +
  labs(title = "Distribution of BMI", x = "BMI", y = "Count") +
  theme_minimal()

# Histogram for children
ggplot(insurance, aes(x = children)) +
  geom_histogram(binwidth = 1, fill = "purple", color = "black") +
  labs(title = "Distribution of Number of Children", x = "Number of Children", y =
"Count") +
  theme_minimal()

# Histogram for charges
ggplot(insurance, aes(x = charges)) +
  geom_histogram(binwidth = 1000, fill = "orange", color = "black") +
  labs(title = "Distribution of Charges", x = "Charges", y = "Count") +
  theme_minimal()
```

```
#Bar plot for smoker
ggplot(insurance, aes(x = smoker)) +
  geom_bar(fill = "skyblue") +
  labs(title = "Distribution of Smoking Status", x = "Smoker", y = "Count") +
  theme_minimal()

#Bar plot for region
ggplot(insurance, aes(x = region)) +
  geom_bar(fill = "lightgreen") +
  labs(title = "Distribution of Region", x = "Region", y = "Count") +
  theme_minimal()

# Box plot for smoker vs. charges
ggplot(insurance, aes(x = smoker, y = charges)) +
  geom_boxplot(fill = "skyblue") +
  labs(title = "Charges by Smoking Status", x = "Smoker", y = "Charges") +
  theme_minimal()

# Box plot for region vs. charges
ggplot(insurance, aes(x = region, y = charges)) +
  geom_boxplot(fill = "lightgreen") +
  labs(title = "Charges by Region", x = "Region", y = "Charges") +
  theme_minimal()

# Box plot for sex vs. charges
ggplot(insurance, aes(x = sex, y = charges)) +
  geom_boxplot(fill = "yellow") +
  labs(title = "Charges by Sex", x = "Sex", y = "Charges") +
  theme_minimal()

# Scatterplot for age vs. charges
ggplot(insurance, aes(x = age, y = charges)) +
  geom_point(color = "blue") +
  labs(title = "Charges by Age", x = "Age", y = "Charges") +
  theme_minimal()

# Scatterplot for BMI vs. charges
ggplot(insurance, aes(x = bmi, y = charges)) +
  geom_point(color = "red") +
  labs(title = "Charges by BMI", x = "BMI", y = "Charges") +
```

```
  theme_minimal()

# Scatterplot for children vs. charges
ggplot(insurance, aes(x = children, y = charges)) +
  geom_point(color = "purple") +
  labs(title = "Charges by Number of Children", x = "Number of Children", y = "Charges")
+
  theme_minimal()

#Normal QQ Plot of Charges
qqnorm(insurance$charges)
qqline(insurance$charges)

#Split Training Set (80%) and Test Set (20%) of Data
set.seed(1220)
trainIndex <- createDataPartition(insurance$charges, p = 0.8,
                                  list = FALSE,
                                  times = 1)
trainData <- insurance[ trainIndex,]
testData  <- insurance[-trainIndex,]


#Create a function for Model Diagnostic Tests
model_diagnostics <- function(model) {
  #Plot Model
  par(mfrow = c(2,2))
  plot(model)
  par(mfrow = c(1, 1))
  #Model Summary
  print("Model Summary")
  print(summary(model))
  #ANOVA Table
  print("ANOVA Table")
  print(anova(model))
  #NCV Test
  print("Variance Homogeneity Test")
  print(ncvTest(model))
  #Normality Test
  print("Residual Normality Test")
  print(shapiro.test(residuals(model)))
```

```
  #Autocorrelation Test
  print("Test for Auto Correlated Errors")
  print(durbinWatsonTest(model))
  #MultiCollinearity Test
  print("Test for Multi-Collinearity")
  print(vif(model))
  #Outlier Test
  print("Test for Outliers")
  print(outlierTest(model))
  #Influencial Points Test
  print("Influential Points")
  print(summary(influence.measures(model)))
  #Linearity Plot
  print("Linearity Test")
  print(crPlots(model))
}

#Create full model with all variables
model1 <- lm(charges~age+sex+bmi+smoker+children+region,trainData)
model_diagnostics(model1)

#Use transformation on the dependent variable 'charges'
#Find Lambda
b <- boxcox(model1)
lambda <- b$x[which.max(b$y)]
#Create column with boxcox transformed
trainData$boxcox.charges <- ((trainData$charges^lambda)-1)/lambda

ggplot(trainData, aes(x = boxcox.charges)) + geom_histogram()

#Check Model after Transformation
model2 <- lm(boxcox.charges~age+sex+bmi+smoker+children+region, trainData)
model_diagnostics(model2)



#Build model after scaling the variables

#Scaling on the age and bmi variable
trainData <- trainData %>% mutate(scale.age = scale(age),
                                  scale.bmi = scale(bmi),
```

```
                                  scale.children = scale(children))




#Fit Model after scaling of the residuals

model3 <- lm(boxcox.charges~scale.age+sex+scale.bmi+smoker+scale.children+region,
trainData)

model_diagnostics(model3)




#Use interaction terms in the model

#Update the model diagnostics function by removing the linearity test as it is no longer
needed

model_diagnostics <- function(model) {

  #Plot Model

  par(mfrow = c(2,2))

  plot(model)

  par(mfrow = c(1, 1))

  #Model Summary

  print("Model Summary")

  print(summary(model))

  #ANOVA Table

  print("ANOVA Table")

  print(anova(model))

  #NCV Test

  print("Variance Homogeneity Test")

  print(ncvTest(model))

  #Normality Test

  print("Residual Normality Test")

  print(shapiro.test(residuals(model)))

  #Autocorrelation Test

  print("Test for Auto Correlated Errors")

  print(durbinWatsonTest(model))

  #MultiCollinearity Test

  print("Test for Multi-Collinearity")

  print(vif(model))

  #Outlier Test

  print("Test for Outliers")

  print(outlierTest(model))

  #Influencial Points Test
```

```
  print("Influential Points")

  print(summary(influence.measures(model)))

}


model4 <- lm(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi + region +
sex, data = trainData)

model_diagnostics(model4)



all_model <- regsubsets(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi
+ region + sex, data = trainData)

summary(all_model)

plot(all_model, scale = "bic")

plot(all_model, scale = "adjr2")



sub_model1 <- lm(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi +
region, trainData)

model_diagnostics(sub_model1)

press(sub_model1)


sub_model2 <- lm(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi,
trainData)

model_diagnostics(sub_model2)

press(sub_model2)


#GLM Model using Sub_model2

glm_model <- glm(boxcox.charges ~ scale.age * smoker * scale.children + scale.bmi, family
= gaussian(link = "identity"), data = trainData)

summary(glm_model)

# Deviance of the model

deviance(glm_model)

# p-value for the deviance test

pchisq(glm_model$deviance, df=glm_model$df.residual, lower.tail=FALSE)

# Pearson's Chi-Square Test

pchisq(sum(resid(glm_model, type="pearson")^2), glm_model$df.residual, lower.tail =
FALSE)

# Hosmer-Lemeshow Test

hoslem.test(trainData$boxcox.charges, fitted(glm_model))

#Anova Test

Anova(glm_model, type = 3)

anova(glm_model, test = "Chisq")

#Residual Normality
```

```
shapiro.test(resid(glm_model, type="deviance"))

#Prepare Test data
testData$boxcox.charges <- ((testData$charges^lambda)-1)/lambda

testData <- testData %>% mutate(scale.age = scale(age),
                                scale.bmi = scale(bmi),
                                scale.children = scale(children))

#Create Function for Performance Metrics

performance_metrics <- function(model, testData, response_var) {

   # Make predictions on the test data
  predictions <- predict(model, newdata = testData)

  # Extract actual values
  actuals <- testData[[response_var]]

  # Calculate performance metrics
  mae_value <- mae(actuals, predictions)
  mse_value <- mse(actuals, predictions)
  rmse_value <- rmse(actuals, predictions)
  r2_value <- R2(predictions, actuals)
  mape_value <- mape(actuals, predictions)

  # Create a table of performance metrics
  performance_metrics <- data.frame(
    Metric = c("Mean Absolute Error (MAE)", "Mean Squared Error (MSE)",
               "Root Mean Squared Error (RMSE)", "R-squared (R2)",
               "Mean Absolute Percentage Error (MAPE)"),
    Value = c(mae_value, mse_value, rmse_value, r2_value, mape_value)
  )

  return(performance_metrics)
}

performance_metrics(sub_model1, testData, "boxcox.charges")

performance_metrics(sub_model2, testData, "boxcox.charges")
```

```r
performance_metrics(model4, testData, "boxcox.charges")

performance_metrics(glm_model, testData, "boxcox.charges")

scatter_predictvsobserved <- function(model, test_data, response_var) {
  # Generate predictions
  predictions <- predict(model, newdata = test_data)

  # Create a data frame for plotting
  plot_data <- data.frame(
    Observed = test_data[[response_var]],
    Predicted = predictions
  )

  # Create the plot
  ggplot(plot_data, aes(x = Observed, y = Predicted)) +
    geom_point(color = "blue", alpha = 0.5) +
    geom_abline(intercept = 0, slope = 1, color = "red", linetype = "dashed") +
    labs(
      title = "Predicted vs Observed Values",
      x = "Observed Values",
      y = "Predicted Values"
    ) +
    theme_minimal()
}

scatter_predictvsobserved(sub_model1, testData, "boxcox.charges")
scatter_predictvsobserved(sub_model2, testData, "boxcox.charges")
scatter_predictvsobserved(model4, testData, "boxcox.charges")
scatter_predictvsobserved(glm_model, testData, "boxcox.charges")
```