How High Can You Go?

An Analysis of SBP

Zhe Fan Li,1005057368 (Data Cleaning, Model Assumptions, Outliers, Research)
Arib Shaikh, 1005422348 (Introduction, Dataset Description, Model Interpretation, Research)
Joseph Wang, 1005233141 (Model Selection, Significance Testing, Model Interpretation, Model Validatio
Shruti Sood, 1005593715 (Introduction, Data Visualizaiton, Conclusion, Research)

12/04/2021

Contents

Abstract

High blood pressure is a major risk factor for heart disease and stroke, both of which are leading causes of death in the US.(2013).

Total blood pressure readings are determined by measuring your systolic and diastolic blood pressures. Systolic blood pressure, the top number, measures the force your heart exerts on the walls of your arteries

each time it beats. Diastolic blood pressure, the bottom number, measures the force your heart exerts on the walls of your arteries in between beats (2012).

The goal of this case study is to identify the social and health factors that have a significant impact on Systolic Blood Pressure in BloodPressure.xlsx dataset.

Introduction

Hypothesis

By primary analysis of the correlation and scatter-plot matrices, it is hypothesized that health factors such as smoking, BMI, salt, alcohol consumption and stress will have a significant positive relation with SBP. Whereas, exercising, and hypertension treatment have a significant negative relation with SBP.

Data Cleaning

```
dat$gender <- ifelse(dat$gender=="M", 1, 0)
dat$married <- ifelse(dat$married=="Y", 1, 0)
dat$smoke <- ifelse(dat$smoke=="Y", 1, 0)
sum(is.na(dat)) == 0
```

To initially clean the data, all alphabetical categorical variables are turned into binary categorical variables. The data is also checked for NA values. There are none, so no measures need to be taken.

Methods

The problem of model selection can be solved by finding the most correlated subset of quantitative variables with SBP, and interacting them with the most correlated subset of categorical variables with SBP.

After the initial model is formed with stepAIC, individual variables will be tested for significance to polish the model. The model coefficients and assumptions will be analyzed and any remedial measures will be applied, if needed.

The coefficient and the significance will determine the importance of a variable, for a healthy SBP.

Left Out Variables

To identify and remove multicollinear terms, the alias() function is used. It is discovered that the unable childbearing female is a linear combination of gender and able bearing female, therefore chldbear is unused in favor of the preexisting gender variable.

```
fit = lm(sbp ~ factor(gender) + factor(married) + factor(smoke) +
factor(exercise) + age + weight + height + factor(overwt) + factor(race) +
factor(alcohol) + factor(trt) + bmi + factor(stress) + factor(salt) +
factor(chldbear) + factor(income) + factor(educatn), data= dat)
# from alias, childbearing is a linear combination of gender
alias(fit)$Complete
```

```
##                     (Intercept) factor(gender)1 factor(married)1 factor(smoke)1
## factor(chldbear)3  1          -1               0                0
##                     factor(exercise)2 factor(exercise)3 age weight height
## factor(chldbear)3  0                 0                 0   0      0
##                     factor(overwt)2 factor(overwt)3 factor(race)2 factor(race)3
## factor(chldbear)3  0               0               0             0
##                     factor(race)4 factor(alcohol)2 factor(alcohol)3 factor(trt)1
## factor(chldbear)3  0             0                0                0
##                     bmi factor(stress)2 factor(stress)3 factor(salt)2
## factor(chldbear)3  0   0               0               0
##                     factor(salt)3 factor(chldbear)2 factor(income)2
## factor(chldbear)3  0             -1                 0
##                     factor(income)3 factor(educatn)2 factor(educatn)3
## factor(chldbear)3  0               0                0
```

```
# fit after removing chldbearing
fit = lm(sbp ~ factor(gender) + factor(married) + factor(smoke) +
factor(exercise) + age + weight + height + factor(overwt) + factor(race) +
factor(alcohol) + factor(trt) + bmi + factor(stress) + factor(salt) +
factor(income) + factor(educatn), data= dat)
# from this vif, weight/height/overwt and bmi are highly multicollinear
vif(fit)[c("weight", "height", "factor(overwt)", "bmi"),]
```

```
##                       GVIF Df GVIF^(1/(2*Df))
## weight         27.713811  1        5.264391
## height         16.707595  1        4.087492
## factor(overwt)  6.587666  2        1.602075
## bmi            40.269437  1        6.345820
```

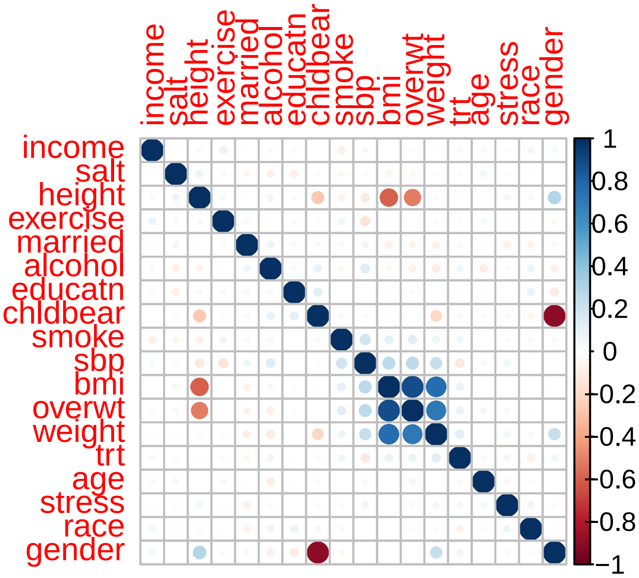
VIF is used to identify multicollinearity; it is discovered weight, height, overweight, and BMI are multicollinear. Is decided that BMI is sufficient in describing the other 3 factors since BMI is a function of weight and height and overwt is a function of weight.

Description of Dataset

The first step to selecting a model is to determine which variables are correlated with the response variable and the degree of intercorrelation between the predictor variables.

Correlation Matrix

A correlation matrix is used to find the pairwise correlation between the variables.



First, parameters that are correlated with the response variable, SBP, are identified.

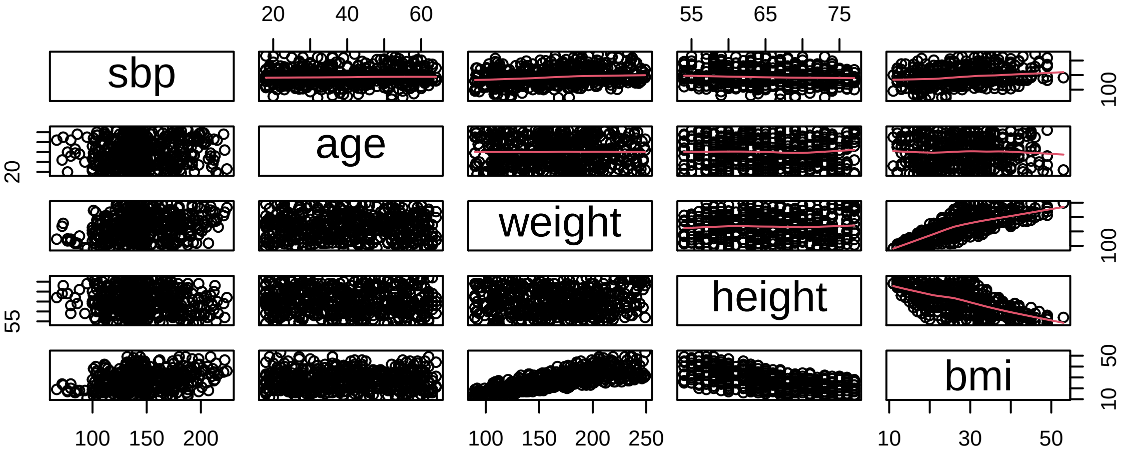
Variables that have a significant positive relationships with the response variable, SBP: smoke, weight, overwt, alcohol, BMI.

Variables that have a significant negative relationships with the response variable, SBP: exercise, height, trt.

Social variables do not seem to have a relationship with SBP, while health variables have a more significant relationship. Apart from using these initial variables in the hypothesis, these serve as a basis to find other correlated variables.

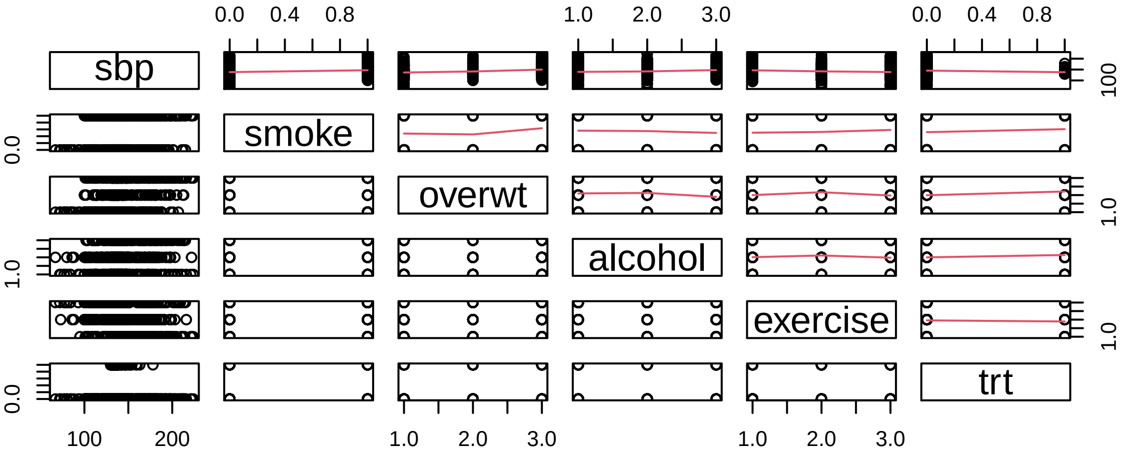
Pairwise Scatter Plot

Significant Quantitative Variables



There is a linear relationship between weight/height and BMI as expected. There is also a slight linear relationship between weight/BMI and SBP.

Significant Categorical Variables

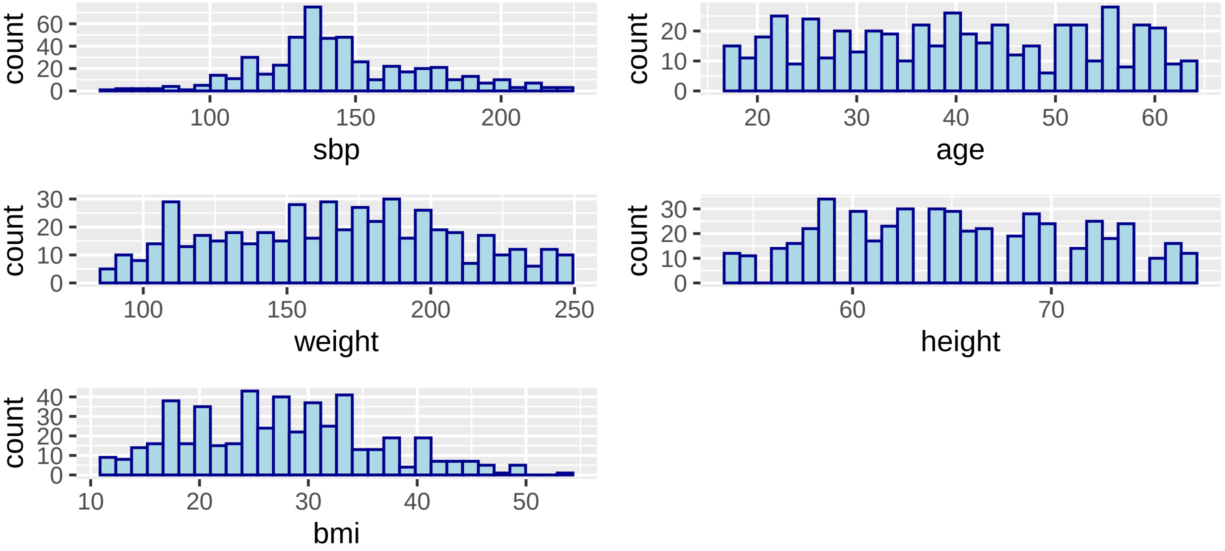


There seems to be a slight positive linear relationship between smoke/overweight/alcohol and SBP and a slight negative relationship between exercise/trt vs SBP.

There is not a lot of overall correlation within the data as seen from the correlation and scatter-plot matrices.

Histograms

Significant Quantitative Variables



Histogram 1 - SBP: The histogram for SBP seems to follow a normal distribution with average around 145. The average for this data set shows that most people in this data set have an extremely unhealthy SBP which could result in bias.

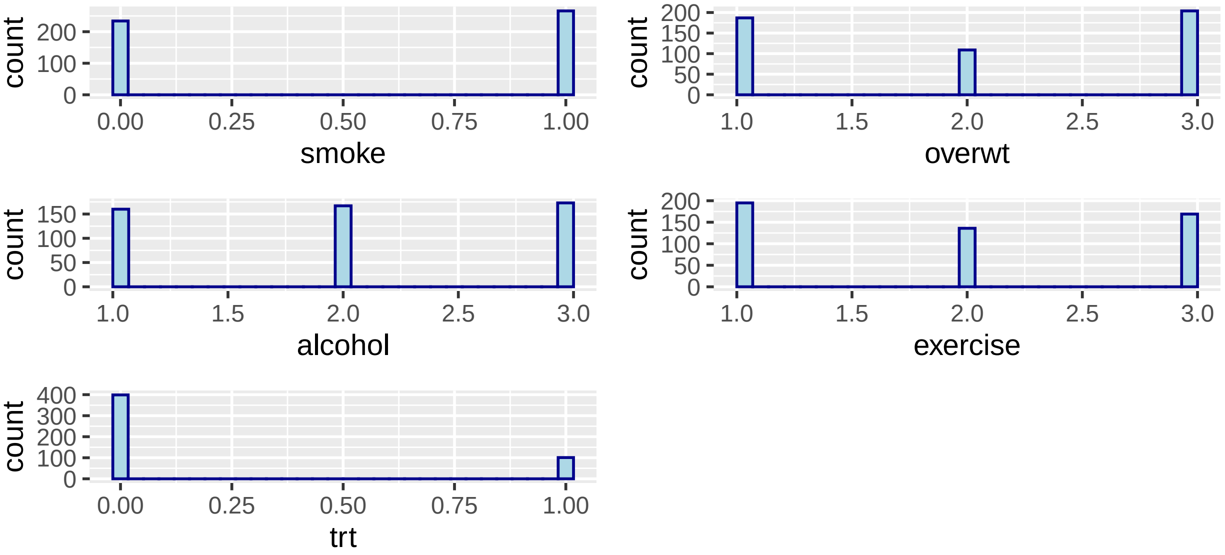
Histogram 2 - Age: The age histogram is almost uniformly distributed. This means our data set accurately represents people of all ages and is not biased towards any certain age group.

Histogram 3 - Weight: The weight histogram is roughly uniform.

Histogram 4 - Height: The height histogram is also roughly uniform.

Histogram 5 - BMI: The BMI histogram is slightly right skewed with mean around 37. BMI 30 or above is considered to be obese.

Significant Categorical Variables



Histogram 6 - Smoke: There are almost equal amounts of smokers and non-smokers.

Histogram 7 - Overweight: There are almost an equal amount of Normal Weight and Obese people in this data set.

Histogram 8 - Alcohol: All three categories have around equal distribution.

Histogram 9 - Exercise: According to the histogram, the number of people doing a low level of exercise is the highest.

Histogram 10 - Treatment of Hypertension: In this data set, 4 times more people are not undergoing hypertension treatment compared to people undergoing hypertension treatment.

To summarize, people from in this data are on the heavier side with higher than average SBP and BMI. There aren't any obvious bias within the data.

Model Selection

Model Building

For model selection, backwards stepAIC is used to find the optimal subset of covariates based on a "full model". The variables in the "full model" are selected based on earlier analysis on the correlation matrix and pairwise plot matrix.

Quantitative Variables for stepAIC

Since BMI is a function of weight and height, BMI is used instead of both weight and height.

Even though age did not seem to have a relationship with SBP in the pairwise plots, it is used in the model as it still could have an interactive relationship with one of the other variables.

Categorical Variables for stepAIC

All categorical variables except chldbear and overwt are used along with their interaction terms to find the model with the lowest AIC. Individual variables are tested for significance after.

```
initialModel = sbp ~ age*(factor(exercise)+factor(stress)+factor(salt)+
factor(alcohol)+factor(smoke)+factor(gender)+factor(trt)+factor(married)+
factor(race)+factor(income)+factor(educatn)) + bmi*(factor(exercise)+
factor(stress)+factor(salt)+factor(alcohol)+factor(smoke)+factor(gender)+
factor(trt)+factor(married)+factor(race)+factor(income)+factor(educatn))
fit <- lm(data = dat, formula = initialModel)

stepAIC(fit, direction="backward")
```

The model with the lowest AIC has the formula = "sbp ~ age + factor(exercise) + factor(stress) + factor(salt) + factor(alcohol) + factor(smoke) + factor(gender) + factor(trt) + factor(married) + bmi + age:factor(stress) + age:factor(gender) + factor(exercise):bmi + factor(salt):bmi + factor(trt):bmi"

```
AICModel = sbp ~ age + factor(exercise) + factor(stress) + factor(salt) +
factor(alcohol) + factor(smoke) + factor(gender) + factor(trt) + factor(married) +
bmi + age:factor(stress) + age:factor(gender) + factor(exercise):bmi +
factor(salt):bmi + factor(trt):bmi
fit <- lm(formula = AICModel, data = dat)

summary(fit)$coefficients
```

```
##                           Estimate Std. Error     t value     Pr(>|t|)
## (Intercept)           104.43132254 10.9758815  9.51461828 9.058859e-20
## age                     0.09029192  0.1689614  0.53439380 5.933181e-01
## factor(exercise)2     -14.39095258  9.4590318 -1.52139806 1.288226e-01
## factor(exercise)3     -31.89574073  9.0879941 -3.50965685 4.912415e-04
## factor(stress)2        -0.37802703  8.8168091 -0.04287572 9.658185e-01
## factor(stress)3       -10.61952177  8.7656186 -1.21149712 2.263048e-01
## factor(salt)2          -0.87597226  9.4102651 -0.09308688 9.258736e-01
## factor(salt)3          17.10852547  9.0676180  1.88677176 5.979808e-02
## factor(alcohol)2        1.61010806  2.7997390  0.57509221 5.655001e-01
## factor(alcohol)3       11.42735105  2.7978662  4.08430940 5.184597e-05
## factor(smoke)1         11.24000374  2.2751947  4.94023824 1.081526e-06
## factor(gender)1        14.58583665  7.1560640  2.03824850 4.207613e-02
## factor(trt)1           20.35288879 10.3098878  1.97411351 4.894534e-02
## factor(married)1        3.72367169  2.2574967  1.64946936 9.970986e-02
## bmi                     1.05592170  0.2680958  3.93859865 9.419148e-05
## age:factor(stress)2     0.06159471  0.2092560  0.29435097 7.686178e-01
## age:factor(stress)3     0.40295336  0.2084617  1.93298483 5.382960e-02
## age:factor(gender)1    -0.32344357  0.1683760 -1.92095957 5.533252e-02
## factor(exercise)2:bmi   0.18961065  0.3178883  0.59646948 5.511446e-01
## factor(exercise)3:bmi   0.79869760  0.3211235  2.48719792 1.321610e-02
## factor(salt)2:bmi       0.08413210  0.3212546  0.26188605 7.935223e-01
## factor(salt)3:bmi      -0.59752595  0.3147608 -1.89834916 5.825341e-02
## factor(trt)1:bmi       -1.18828602  0.3422897 -3.47157950 5.644562e-04
```

```r
adjrsq = glance(fit)$adj.r.squared
pval = glance(fit)$p.value
adjrsq
```

```
## [1] 0.2180259
```

```r
pval
```

```
##        value
## 1.773958e-19
```

The model is overall very significant with p-val = $1.7739582 \times 10^{-19}$, but it does not have a high adjusted R^2 value = 0.2180259. The model does not fit the data very well; this could be because there exists a variable that explains SBP that is not within the dataset.

There are variables with high p-values in the reduced model, namely age and stress. An F-test can be done to see if the main effects of age and stress are significant within the model.

```r
significantModel = sbp ~ factor(exercise) + factor(salt) + factor(alcohol) +
factor(smoke) + factor(gender) + factor(trt) + factor(married) + bmi +
age:factor(stress) + age:factor(gender) + factor(exercise):bmi +
factor(salt):bmi + factor(trt):bmi
fit2 <- lm(formula = significantModel, data = dat)

pval = anova(fit2,fit)[2,"Pr(>F)"]
pval
```

```
## [1] 0.3859821
```

The p-val = 0.3859821 > alpha = 0.05 which means that H0:the main effects of age and stress are not significant within that model, cannot be rejected with 95% confidence.

Their interaction terms are still kept since they are still significant.

```r
adjrsq = glance(fit2)$adj.r.squared
adjrsq
```

```
## [1] 0.2181765
```

The new model has a slightly higher adjusted R^2, (0.2181765), than the model with age and stress.

Model Interpretation

```
summary(fit2)
```

```
## 
## Call:
## lm(formula = significantModel, data = dat)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -76.152 -16.306  -1.864  14.467  70.112 
## 
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)    
## (Intercept)           101.4905     9.6296  10.539  < 2e-16 ***
## factor(exercise)2     -14.9008     9.4509  -1.577 0.115536    
## factor(exercise)3     -31.9906     9.0500  -3.535 0.000448 ***
## factor(salt)2          -0.8487     9.4026  -0.090 0.928113    
## factor(salt)3          17.3068     9.0656   1.909 0.056851 .  
## factor(alcohol)2        1.6577     2.7869   0.595 0.552244    
## factor(alcohol)3       11.4361     2.7863   4.104 4.76e-05 ***
## factor(smoke)1         11.2001     2.2727   4.928 1.15e-06 ***
## factor(gender)1        13.4850     7.1093   1.897 0.058455 .  
## factor(trt)1           20.9926    10.2851   2.041 0.041790 *  
## factor(married)1        3.7263     2.2567   1.651 0.099349 .  
## bmi                     1.0465     0.2680   3.905 0.000108 ***
## age:factor(stress)1     0.1633     0.1239   1.318 0.188209    
## age:factor(stress)2     0.2145     0.1200   1.788 0.074390 .  
## age:factor(stress)3     0.3277     0.1197   2.736 0.006444 ** 
## factor(gender)1:age    -0.3009     0.1675  -1.796 0.073129 .  
## factor(exercise)2:bmi   0.2035     0.3177   0.640 0.522203    
## factor(exercise)3:bmi   0.8011     0.3204   2.501 0.012732 *  
## factor(salt)2:bmi       0.0855     0.3211   0.266 0.790175    
## factor(salt)3:bmi      -0.6026     0.3147  -1.915 0.056099 .  
## factor(trt)1:bmi       -1.2029     0.3416  -3.522 0.000470 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 24.75 on 479 degrees of freedom
## Multiple R-squared:  0.2495, Adjusted R-squared:  0.2182 
## F-statistic: 7.963 on 20 and 479 DF,  p-value: < 2.2e-16
```

There are still variables will high p-values like medium use of alcohol(p=5.522437e-01), but that doesn't mean the variable is useless; a high use of alcohol is very significant within our model(p=4.764564e-05) so the variable must not be removed.

Quantitative Variables Analysis

BMI, as predicted, has a positive effect on SBP. With everything else kept constant, every unit increase of BMI increases the mean SBP by 1.04648264 with a significant p-value = 1.077199e-04.

A high level of exercise increases the effect of BMI on mean SBP by 0.80107948 with a significant p-value = 1.273243e-02 with everything else kept constant. It is surprising that a high level of exercise will interact positively with the effect of BMI on SBP instead of negatively.

Treatment for hypertension and high levels of salt both decrease the effect of BMI on mean SBP by 1.20293542 and 0.60263588 with significant p-values = 4.699212e-04, 5.609918e-02 respectively, with everything else kept constant. This is also surprising as their effects are flipped.

High levels of stress has a positive interaction with age on SBP; for high levels stress, the effect of age on mean SBP is increased by 0.32766193 with a significant p-value 6.444475e-03, with everything else kept constant.

Gender(male) has negative interaction with age on SBP; for males, the effect of age on mean SBP is decreased by 0.30089780 with a significant p-value 7.312898e-02 with everything else kept constant.

Categorical Variables

As predicted, exercise has a negative effect on SBP. Mean SBP decreases by 31.99059446 for people with high levels of exercise with a significant p-value = 4.475478e-04 with everything else kept constant.

Also as seen in previous research, high levels of salt intake, alcohol use, smoking and treatment for hypertension all have a positive effect on SBP. Mean SBP increases by 17.30677653, 11.43612223, 11.20009325, 20.99262929 with significant p-values = 5.685063e-02, 4.764564e-05, 1.145981e-06, 4.179035e-02 respectively with everything else kept constant.

Finally, social factors like gender and marriage also have a positive effect on SBP. Mean SBP increases by 13.48497389 for males and 3.72626604 for married people with significant p-values = 5.845514e-02, 9.934916e-02 respectively with everything else kept constant.

As mentioned earlier, categorical variables have high p-values for some categories and a low p-value for another category. This means that the category with the low p-value could have a particular high effect on SBP while other categories may not. Overall, the variable is still significant and holds information about the levels of SBP and cannot be removed.

Most Significant Variables

Therefore, the most significant (p-value < 0.05) variables are: BMI, BMI:exercise, trt:BMI, age:stress, exercise, alcohol, smoke, salt, trt.

Looking at the their coefficients, increasing exercise, not smoking, reducing alcohol and salt intake, have the largest effects on decreasing SBP. Reducing BMI also has a large effect on decreasing SBP if it can be reduced a lot.

Treatment for hypertension has an unexpected relationship, possibly from multicolinearity.

Model Validation

To validate the selected model, the data will be split 60% training and 40% validation.

```r
set.seed(123)
n = length(dat$sbp)*0.6
datIdx <- sample(1:length(dat$sbp), n, replace=FALSE)
trainingDat <- dat[datIdx,]
validationDat <- dat[-datIdx,]

fitTrain <- lm(formula = significantModel, data = trainingDat)
fitValidate <- lm(formula = significantModel, data = validationDat)

predictTrainWVal <- predict(fitValidate, trainingDat)
MSPR <- sum((trainingDat$sbp - predictTrainWVal)^2)/n
MSE <- anova(fitTrain)["Residuals", "Mean Sq"]
MSPR
```

```
## [1] 717.2913
```

```
MSE
```

```
## [1] 627.0063
```

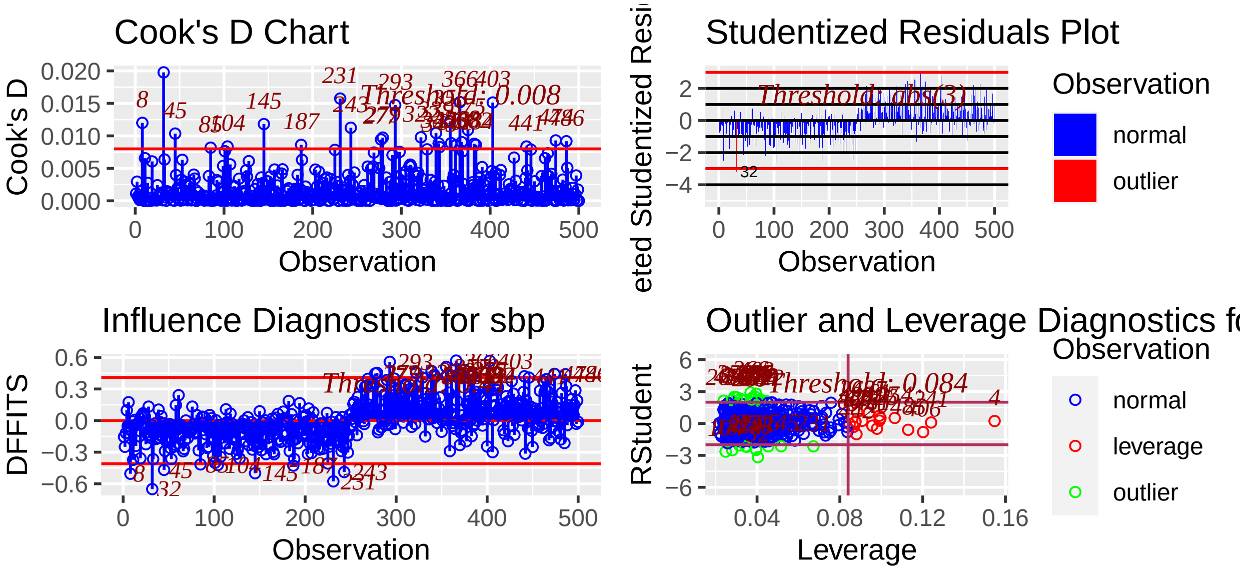
The MSE of the training model is 627.0063461 and the MSPR of the prediction variable is 717.2912629. The MSE and MSPR are similar so the model is not too overfit to the data.

Model Diagnostics

Outliers and Influential Points

To find outliers, 4 detection techniques were used. Cook's distance, deleted studentized residuals, DIFFITS and leverage.

Since this dataset contains 500 observations, the threshold for large sized datasets is used. This means Cook's distance's threshold is the 8th percentile, deleted studentized residuals' threshold is greater than 3 or smaller than -3, DIFFITS's residual is $2 \cdot \sqrt{\frac{p'}{n}}$, and leverage's threshold is $2 \cdot \frac{p'}{n}$



By taking an intersection of outliers from all techniques with non-null outputs, a confident list of all outliers can be created.

```
all_outliers=intersect(cook_outliers, dffits_outliers)
all_outliers
```

```
##  [1]   8  32  45  85 104 145 187 231 243 277 279 293 322 339 342 343 355 356 361
## [20] 366 368 375 382 384 403 441 474 486
```

```
striped_data = data[-all_outliers, ]
adjrsq = summary(lm(formula = model, data = striped_data))$adj.r.squared
adjrsq
```

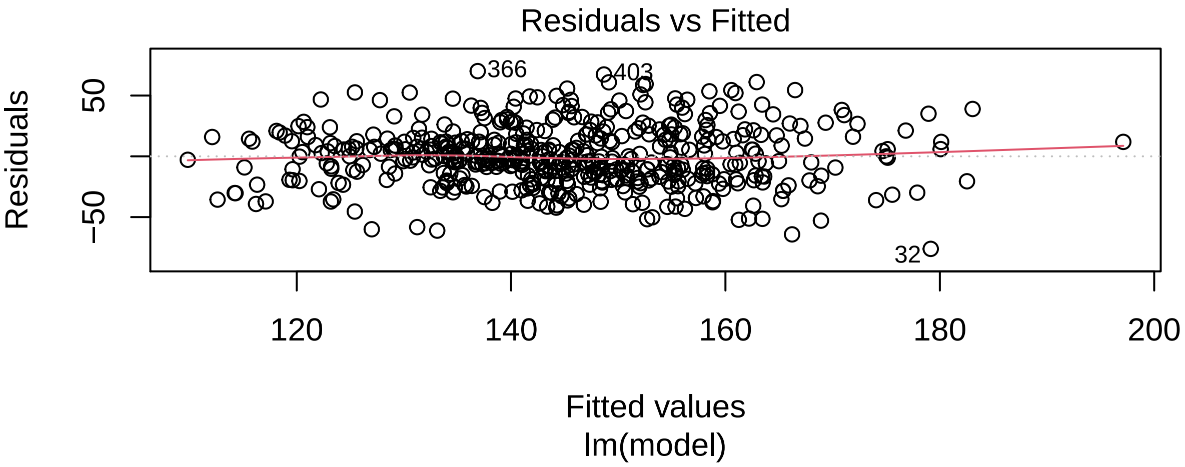
```
## [1] 0.2825429
```

When a model is fitted with the list of outliers taken out, there is a significantly higher adjusted R^2 value = 0.2825429. This means that these points can be called influential points.

Model Assumptions

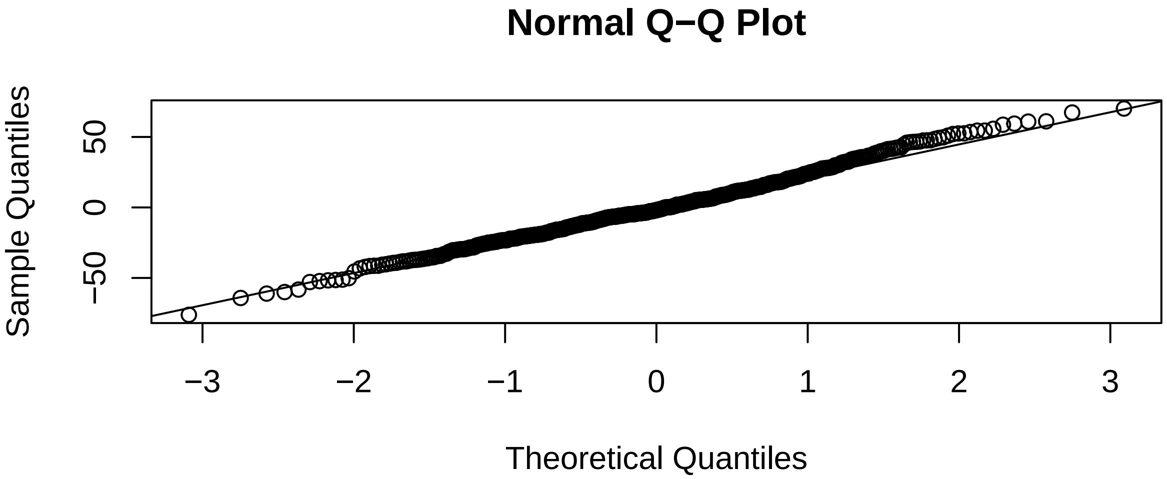
Linearity of Error Terms

The linearity of the data can be seen through the residual vs fitted plot.



The residuals seem to linear with a slight downwards trend. But overall, we cannot reject linearity.

Normality of Error Terms



```
shapiro.test(fit$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit$residuals
## W = 0.99446, p-value = 0.0672
```

Since the QQ plot is relatively linear, and the p-value = 0.0672 is greater than alpha=0.05, we are not able to reject H0: the residuals are normal.

The p-value is very close to being less than 0.05, to be more confident of normality, a box-cox transformation is done.

```
result = boxcox(fit, plotit = FALSE)
lambda = result$x[which.max(result$y)]

transformedModel = sbp^lambda ~ factor(exercise) + factor(salt) +
factor(alcohol) + factor(smoke) + factor(gender) + factor(trt) + factor(married) +
bmi + age:factor(stress) + age:factor(gender) + factor(exercise):bmi +
factor(salt):bmi + factor(trt):bmi

fit = lm(formula = transformedModel, data = data)
shapiro.test(fit$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit$residuals
## W = 0.99601, p-value = 0.239
```

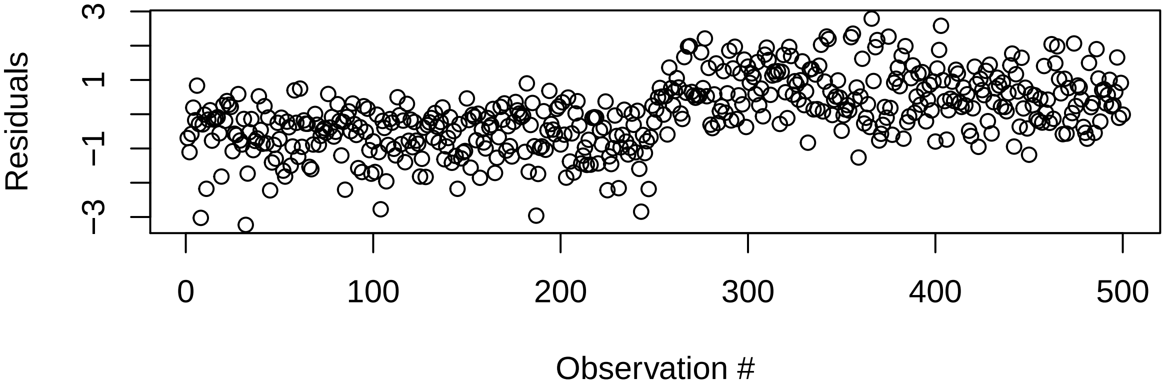
There is an increase in confidence of normality after the transformation shown in the p-value = 0.239. The lambda value used was 0.5.

Final Model

Our final model is the transformedModel = sbp^lambda ~ factor(exercise) + factor(salt) + factor(alcohol) + factor(smoke) + factor(gender) + factor(trt) + factor(married) + bmi + age:factor(stress) + age:factor(gender) + factor(exercise):bmi + factor(salt):bmi + factor(trt):bmi

Independent Error Terms

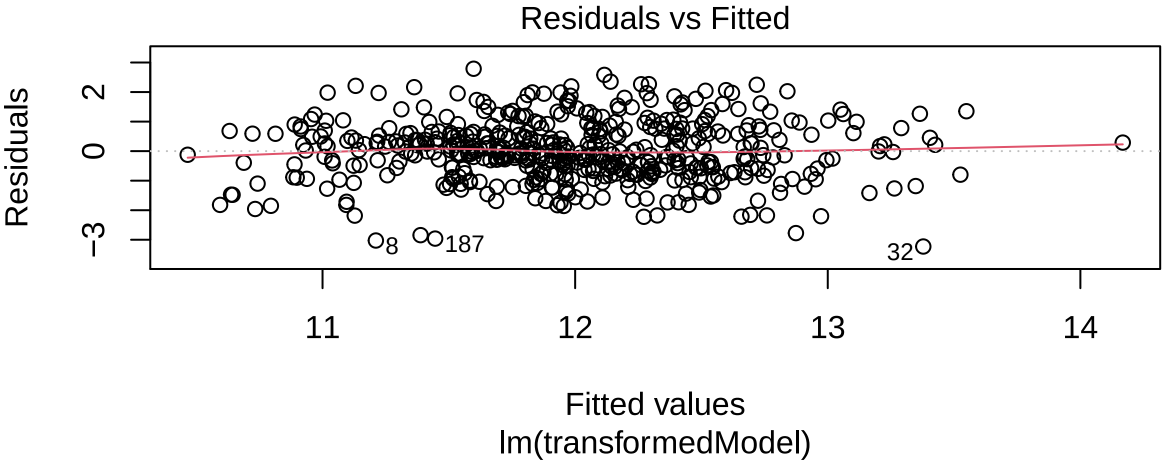
Since there are no variables that had to do with time, the observation number was used to test independent error terms.



Since there is no obvious deviate pattern in the residual - observation plot, we can not reject independent error terms.

Constant Error Variance

A BP-test was used to test constant variance.



```
p = ols_test_breusch_pagan(fit)$p
p
```

```
## [1] 0.1192329
```

Breusch-Pagan test gives a p-value = 0.1192329 > 0.05, therefore we can not reject homoscedasticity.

Conclusion

Findings summary

The final model chosen is the transformedModel = sbp^0.5 ~ factor(exercise) + factor(salt) + factor(alcohol) + factor(smoke) + factor(gender) + factor(trt) + factor(married) + bmi + age:factor(stress) + age:factor(gender) + factor(exercise):bmi + factor(salt):bmi + factor(trt):bmi

The final model has a small p-value which means its results are significant but has a small adjusted R^2 which means that the model doesn't fit the data well.

The most significant (p-value < 0.05) variables found are: BMI, BMI:exercise, trt:BMI, age:stress, exercise, alcohol, smoke, salt, trt.

The regression model shows that out of the quantitative variables, only BMI has an effect(positive) on SBP. It was also found that high levels of exercise increases the effect of BMI on SBP, a decreasing effect was expected. Treatment for hypertension decreased the effect of BMI on SBP.

The model also shows that high levels of stress when interacting with age has a positive effect on SBP. Finally, high levels of salt intake, alcohol use, smoking and treatment for hypertension all have a positive effect on SBP, while exercise has a negative effect on SBP.

Overall, based on the magnitude of the coefficients, reducing bmi, increasing exercise, not smoking, reducing alcohol and lowering salt intake have the largest effects on decreasing SBP.

From validation, 627.0 = MSE ≈ MSPR = 717.29, which means that the model is not overfitting and can generalize okay.

The model also satisfies all the assumptions such as Constant Variance, Error Normality, Independence of Errors and Model Linearity, thus no remedial measures are needed.

The aim of this case study was to find the health and social factors that affect SBP in the BloodPressure.xlsx dataset. It was hypothesised that health factors such as smoking, BMI, salt, alcohol consumption and stress will have a significant positive relation with SBP. Whereas, exercising, and hypertension treatment have a significant negative relation with SBP. It can be concluded that the hypothesis is satisfied given results from the model.

Limitations

The model has a low R-squared value which means it does not fit the data very well. This could be because there exists a variable that explains SBP not within the dataset.

Some of the observations are also unexpected. With an increase in exercise, it was expected the effect on BMI on SBP levels would go down, but it went up instead. Similarly with an increase in treatment for hypertension, it was expected that SBP levels would go down, but instead it went up. A possible explanation is that people with high SBP are either not aware of it or are not seeking treatment for hypertension, this is observed in the trt column in the dataset.

Controlling SBP

Based on further research, the next step for controlling SBP is:

Having high SBP is very risky. According to the Centers for Disease Control and Prevention, when your blood pressure is high: You are 4 times more likely to die from a stroke You are 3 times more likely to die from heart disease. Even blood pressure that is slightly high can put you at greater risk. (2012)

As stated in the article in Harvard Health Publishing, few simple steps can help you lower SBP levels:

Weight loss: If you are overweight, every 2 pounds of weight loss can reduce SBP by 1 mm Hg. Losing 10 pounds can drop your SBP by 5 mm Hg (2013).

Diet: A low-fat diet rich in fruits and vegetables and that includes low-fat dairy can reduce SBP by 8 to 14 mm Hg (2013).

Sodium: Limiting daily sodium intake to 6 grams can reduce SBP by 2 to 8 mm Hg (2013).

Exercise: Thirty minutes of physical activity most days of the week can reduce SBP by 4 to 9 mm Hg (2013).

Moderate alcohol: In women, one drink a day may lower SBP by 2 to 4 mm Hg. However, there may be a very small increase in the risk of breast cancer in women who have one drink a day (2013).

Following a healthier lifestyle day to day can help you live a longer, healthier life.

References

Centers for Disease Control and Prevention. (2012, September). Vital Signs.Getting Blood Pressure Under Control. https://www.cdc.gov/vitalsigns/hypertension/index.html

Harvard Health Publishing.(2013, May). Harvard Medical School. Ask the doctor: Will lifestyle changes help with systolic hypertension? https://www.health.harvard.edu/heart-health/will-lifestyle-changes-help-with-systolic-hypertension