HW3_priyamm2

February 10, 2022

1 Homework 3

Name: Priyam Mazumdar

NetID: priyamm2

[1]:
```
# Load Required Packages
import numpy as np
import pandas as pd
from datetime import timedelta
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
import seaborn as sns
from sklearn.preprocessing import StandardScaler
```

1.1 Question 1 [50 Points] A Simulation Study

Let's use a simulation study to confirm the bias-variance trade-off of linear regressions. Consider the following model.

$$Y = \sum_{j}^{p} 0.8^j \times X_j + \epsilon$$

All the covariates and the error term follow i.i.d. standard Gaussian distribution. The true model involves all the variables; however, larger indexes do not significantly contribute to the variation. Hence, there could be a benefit in using a smaller subset for prediction purposes. Let's confirm that with a simulation study. - Generate 100 samples of covariates X with $p = 30$ by the following code.

- Then the study essentially **repeats** the following steps 100 times. Begin with another fixed random seed before your loop.
  - Using the fixed covariates X, generate 100 training outcomes Y_{train} and 100 testing outcomes Y_{test} independently.
  - Consider using only the first j variables to fit the linear regression (**NO intercept term**). Let j ranges from 1 to 30. Calculate and record the corresponding prediction error by comparing your prediction with the outcomes for testing data.

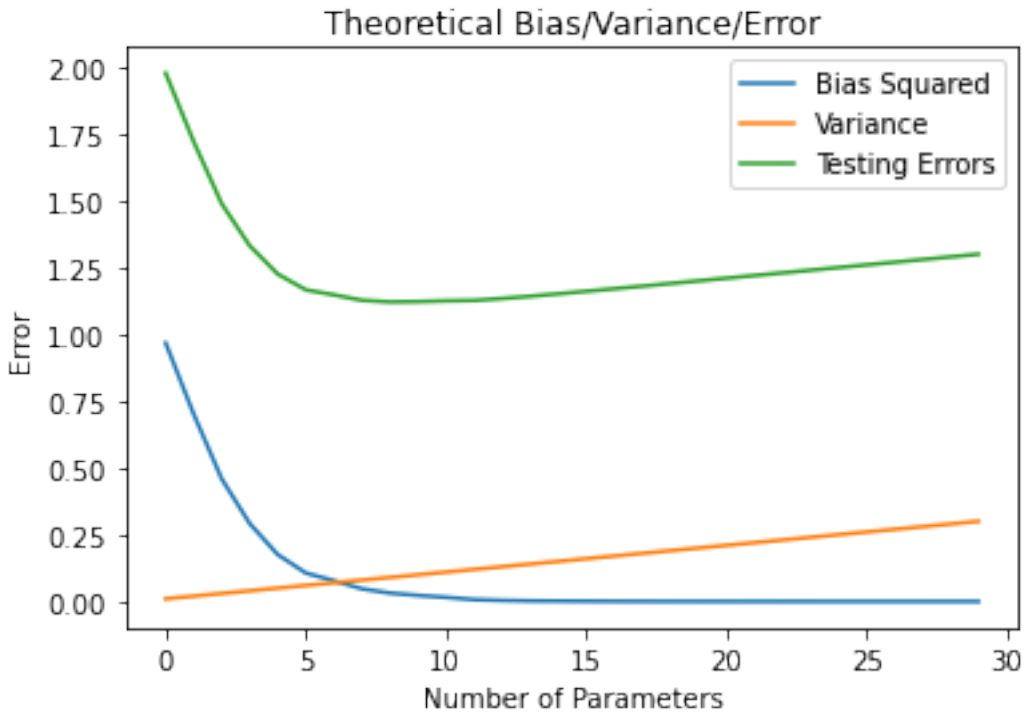
Without running the simulation, for each j value, we also have the theoretical decomposition of the testing error based on the lecture. Suppose you know the true model, covariates X and the distribution of random noise.

a) [15 pts] Please calculate the bias^2 , variance (of the prediction) and testing error for each j based on the theoretical formulas. Plot the 3 lines on the same figure, using the `number of variables` as the x-axis and `bias^2`, `variance`, `theoretical testing error` as the y-axis. Label each line.

- $\text{Bias}^2 = \frac{1}{n}\|E(Y_{\text{pred}}) - Y_{\text{true}}\|^2$, where Y is an $n \times 1$ vector.
- $\text{Var} = \frac{1}{n}E\|Y_{\text{pred}} - E(Y_{\text{pred}})\|^2$.

[2]:
```
# Generate Data
np.random.seed(542)
n = 100
p = 30
b = np.array([0.8 ** i for i in range(1, p+1)]).reshape(p,1)
X = np.random.standard_normal(size=(n,p))

# Calculate Bias
H = X.dot(np.linalg.inv(X.T.dot(X))).dot(X.T)
u = X.dot(b)

bias_values = []
variance_values = []
testing_errors = []
for idx in range(len(b)):
    current_x = X[:, :idx+1]
    current_h = current_x.dot(np.linalg.inv(current_x.T.dot(current_x))).
 ↪dot(current_x.T)
    bias2 = ((u - current_h.dot(u))**2).mean()
    bias_values.append(bias2)
    variance_values.append((idx+1)/n)
    testing_errors.append(1 + bias2 + (idx+1)/n)

plt.plot(bias_values, label="Bias Squared")
plt.plot(variance_values, label="Variance")
plt.plot(testing_errors, label="Testing Errors")
plt.title("Theoretical Bias/Variance/Error")
plt.xlabel("Number of Parameters")
plt.ylabel("Error")
plt.legend()
```

[2]:
```
<matplotlib.legend.Legend at 0x7f7c7260e670>
```



b) [5 pts] Report the theoretical testing error with $p = 30$, $\frac{1}{n}E\|Y_{\text{test}} - Y_{\text{pred}}\|^2$.

[3]:
```
print("Theoretical Testing Error with P = 30:", testing_errors[-1])
```

```
Theoretical Testing Error with P = 30: 1.3
```

After finishing the simulation:

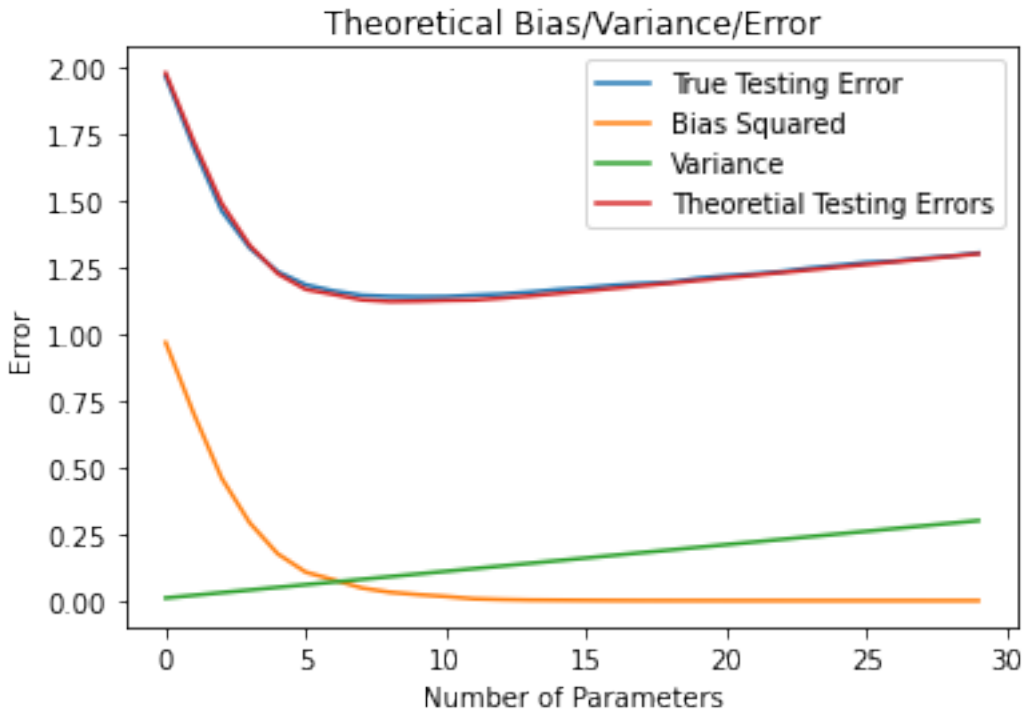
c) [20 pts] Perform the simulation. Report the averaged (empirical) prediction error with $p = 30$. Note that 100 times simulation approximates the E operation. Plot `pred err` in the **same figure** of question a. Label your line. Does your empirical testing error match our theoretical analysis? Comment on your findings.

[4]:
```
mean_errors = []
y_trains = []
y_tests = []

# Create a list of 100 samples of y trains and test based on X and b
for i in range(100):
    y_trains.append(X.dot(b)+np.random.standard_normal(size=(100,1)))
    y_tests.append(X.dot(b)+np.random.standard_normal(size=(100,1)))

for i in range(1,p+1):
    errors = []
    for j in range(100):
```

```
        y_train, y_test = y_trains[j], y_tests[j]
        x_sub = X[:,:i]
        model = LinearRegression().fit(x_sub,y_train)
        errors.append(mean_squared_error(y_test,model.predict(x_sub)))
    mean_errors.append(np.mean(errors))
```

[5]:
```
plt.plot(mean_errors, label="True Testing Error")
plt.plot(bias_values, label="Bias Squared")
plt.plot(variance_values, label="Variance")
plt.plot(testing_errors, label="Theoretial Testing Errors")
plt.title("Theoretical Bias/Variance/Error")
plt.xlabel("Number of Parameters")
plt.ylabel("Error")
plt.legend()
```

[5]: <matplotlib.legend.Legend at 0x7f7c30273370>



We can see that our true testing error calculated from simulation and the theoretical testing error are nearly identical. This shows that given our Betas, we can accurately estimate the testing error based on the number of predictors used.

We aso see that our testing error was minimized around 9 parameters, and adding more paramters after that only hurt our model performance.

d) [10 pts] Evaluate the bias^2 for model $p = 5$ without theoretical formulas. You can still assume you know the true outcomes while using the average results to approximate the E operation. Compare the empirical value with the theoretical one.

[6]:
```
bias_list = []

for sim in range(100):
    y_train = y_trains[sim]
    y_test = y_tests[sim]
    x_sub = X[:, :5]
    model=LinearRegression().fit(x_sub, y_train)
    y_pred = model.predict(x_sub)
    bias = np.mean((y_test-y_pred)**2)

print("Theoretical Bias:", bias_values[4])
print("Simulated Bias:", np.array(bias).mean())
```

```
Theoretical Bias: 0.17563088770223498
Simulated Bias: 1.2291153994161765
```

1.2 Question 2 [50 Points] Bitcoin price prediction

For this question, we will use the Bitcoin data provided on the course website. The data were posted originally on Kaggle (link). Make sure that you read relevant information from the Kaggle website. Our data is the `bitcoin_dataset.csv` file. You should use a training/testing split such that your training data is constructed using only information up to 12/31/2016, and your testing data is constructed using only information starting from 01/01/2017. The goal of our analysis is to predict the `btc_market_price`. Since this is longitudinal data, we will use the information from previous days to predict the market price at a future day. In particular, on each calendar day (say, day 1), we use the information from three days onward (days 1, 2, and 3) to predict the market price on the 7th day.

Hence you need to first reconstruct the data properly to fit this purpose. This is mainly to put the outcome (of day 7) and the covariates (of the previous days) into the same row. Note that you may face missing data, categorical predictors, outliers, scaling issues, computational issues, and maybe others for this question. Use your best judgment to deal with them. There is no general "best answer''. Hence the grading will be based on whether you provided reasoning for your decision and whether you carried out the analysis correctly.

a. [25 Points] Data Construction. Data pre-processing is usually the most time-consuming and difficult part of an analysis. We will use this example as a practice. Construct your data appropriately such that further analysis can be performed. Make sure that you consider the following:

- The data is appropriate for our analysis goal: each row contains the outcome on the seventh day and the covariates based on the first three days. The covariates are not limited to the price.
- Missing data is addressed (you can remove variables, remove observations, impute values or propose your own method)

- You may process the covariates and/or outcome by considering centering, scaling, transformation, removing outliers, etc. However, these are your choice.

For each of the above tasks, make sure that you **clearly document your choice**. In the end, provide a summary table/figure of your data. You can consider using boxplots, quantiles, histograms, or any method that is easy for readers to understand. You are required to pick at least one method to present.

[7]:
```
# Load Data
data = pd.read_csv("bitcoin.csv")

# Check for NAs
missing_values = data.isnull().sum()[data.isnull().sum()>0]
```

We can see that we have 21 missing values in our data in the btc_trade_volume column so we can use interpolation by filling an NA with the average of the point before and after. Now we can apply some light post processing to column names, realign our data and split into train and test.

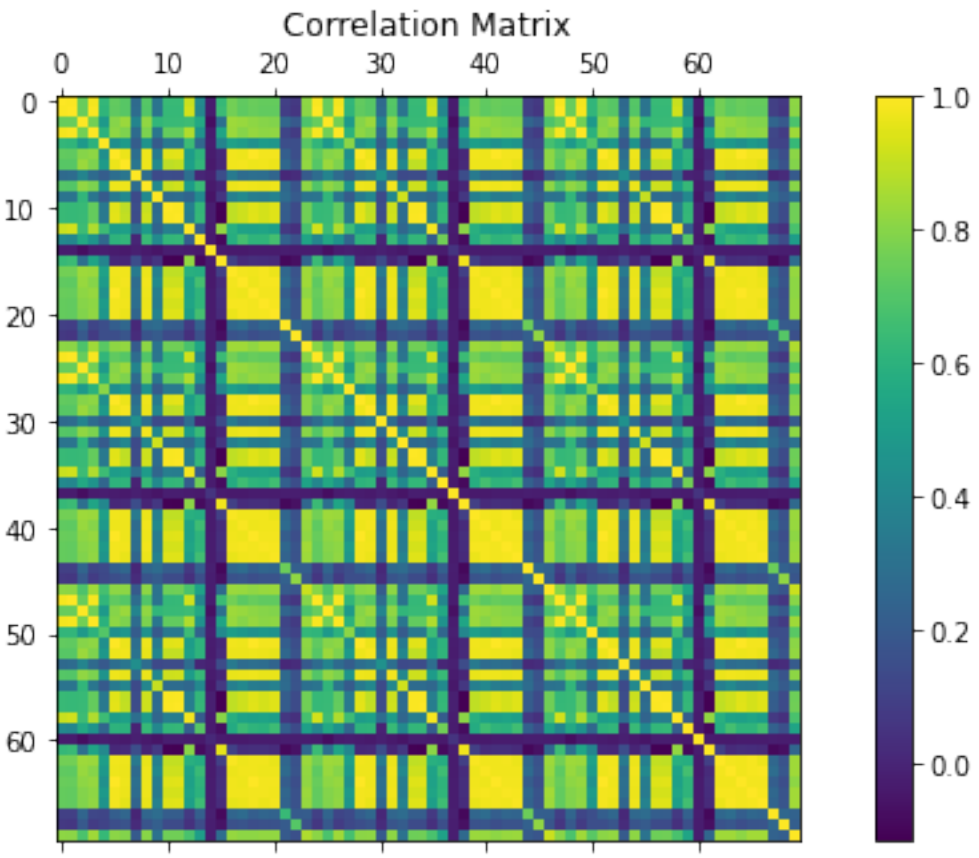
[8]:
```
# Interpolate
data["btc_trade_volume"] = data["btc_trade_volume"].interpolate()

# Preprocess Columns
data = data.rename(columns={"Date":"date"})
data["date"] = pd.to_datetime(data["date"])
data = data.sort_values(by="date", ascending=False).reset_index(drop=True)

# Split train and test data
train_data = data[data["date"] <= "2016-12-31"].reset_index(drop=True)
test_data = data[data["date"] > "2016-12-31"].reset_index(drop=True)

# Get target price chart
price_chart = train_data[["date", "btc_market_price"]]

def alignData(dataframe):
    # Pull price chart out
    price_chart = dataframe[["date", "btc_market_price"]]

    # Set the date as index
    dataframe = dataframe.set_index("date")

    # Go back 4, 5, and 6, days (5, 6, and 7 with 0 index) and
    # shift dataframe to align idx days previous to current day
    # store dataframe in list

    dataframe_list = []
    for shift in [-5, -6, -7]:
        shifted = dataframe.shift(shift)
        dataframe_list.append(shifted)
```

```
    # Loop through shifted dataframes and rename columns indicating # of days␣
↪previous
    for idx, data in enumerate(dataframe_list):
        days_back = 4 + idx
        cols = [f"{days_back}_days_prev_{col}" for col in data.columns]
        data.columns = cols
        dataframe_list[idx] = data

    # Concatenate all dataframes in list and merge on price chart
    data_concat = pd.concat(dataframe_list, axis=1).reset_index()
    merged_data = pd.merge(price_chart, data_concat, on="date", how="left")
    merged_data = merged_data.rename(columns={"btc_market_price":
↪"target_price"})
    merged_data = merged_data.set_index("date")
    merged_data = merged_data.loc[:,~merged_data.columns.duplicated()]
    return merged_data

train_data = alignData(train_data)
test_data = alignData(test_data)
```

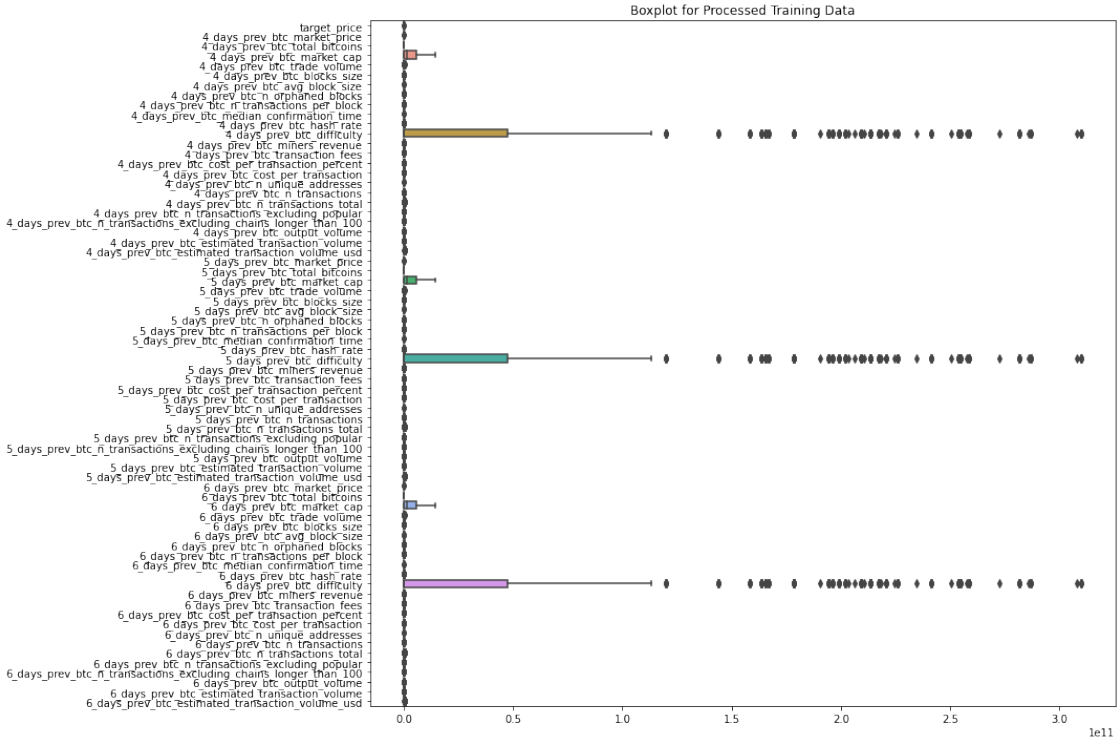
1.2.1 Feature Engineering

We can expect that because most columns are just a shifted version of the original, we will see some heavy correlation between them. Do handle this we want to first check the correlation matrix and a boxplot of all the columns to make sure we have well adjusted data.

[9]:
```
f = plt.figure(figsize=(10, 5))
plt.matshow(train_data.corr(), fignum=f.number)
plt.title("Correlation Matrix")
cb = plt.colorbar()
```



[10]:
```
plt.figure(figsize=(15,10))
ax = sns.boxplot(data=train_data,orient='h')
plt.title('Boxplot for Processed Training Data')
plt.tight_layout()
plt.show()
```



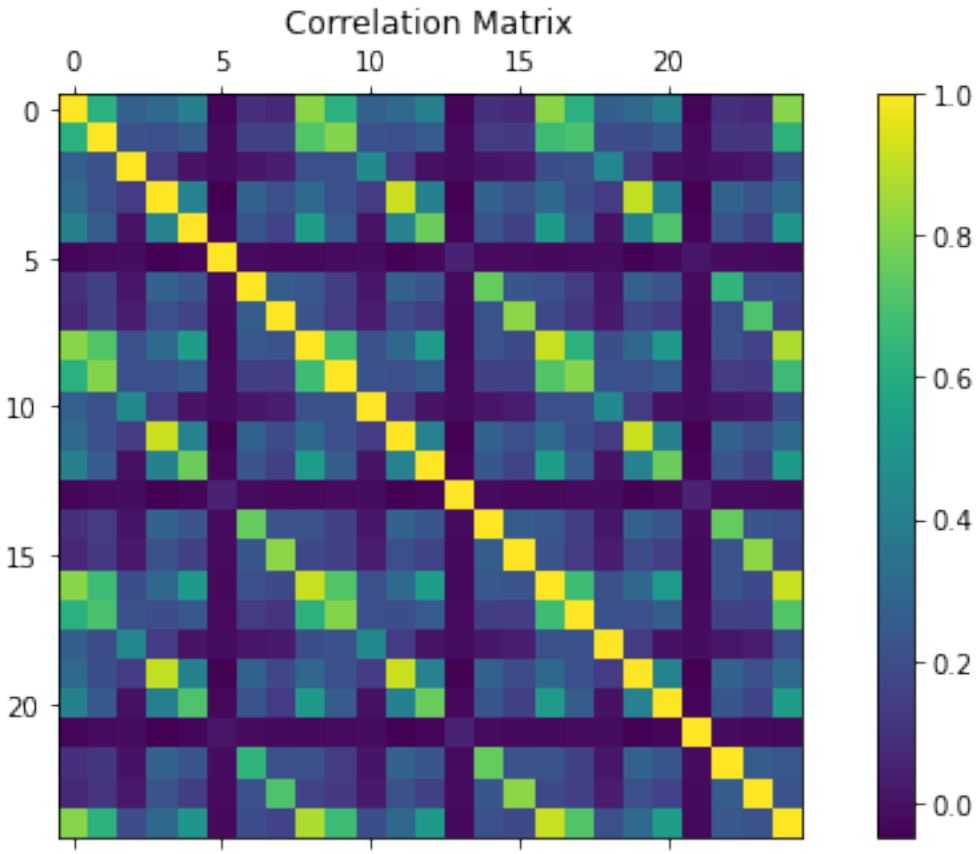
Looking at our data we can see a couple of glaring issue: - High correlation between many pairs of variables - No proper scaling between variables

To fix this we will take our data and perform the removal of columns that have high correlation (greater than 0.95) as well as normalize all the columns to alleviate this.

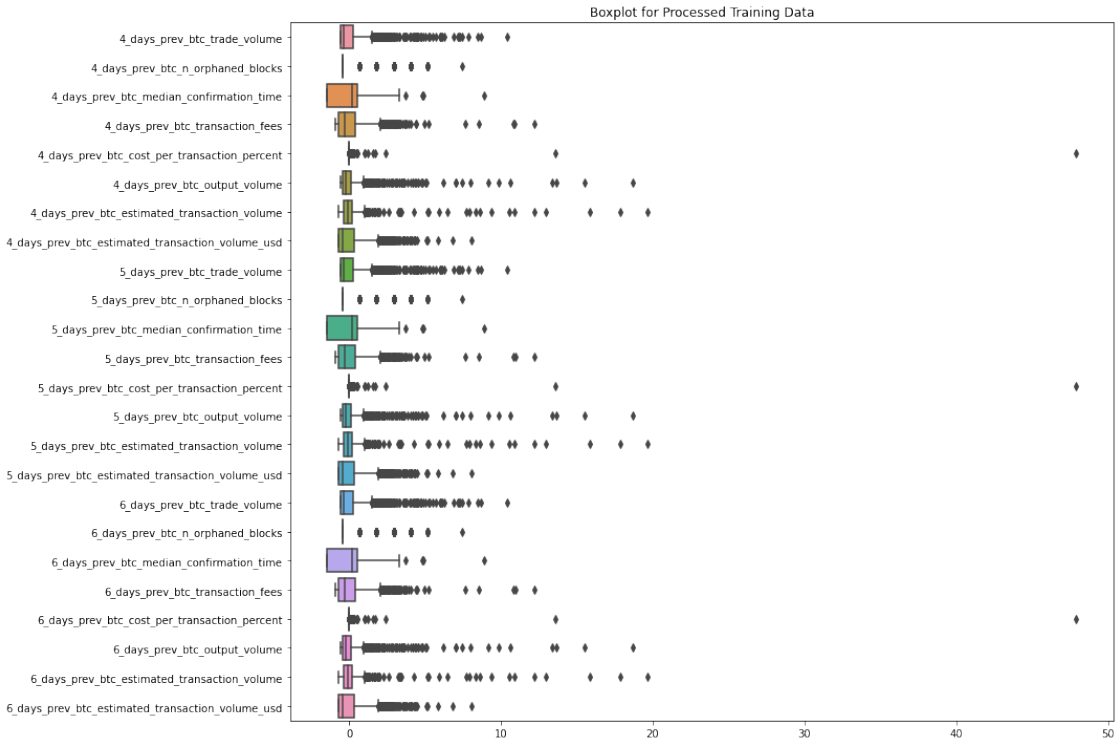
[11]:
```
# Store targets
target_train = train_data["target_price"].reset_index()
target_test = test_data["target_price"].reset_index()

# Drop columns with greater than 0.95 correlation
correlation_table = train_data.corr()
uncorrelated_cols = ~(correlation_table.mask(np.eye(len(correlation_table),␣
↪dtype=bool)).abs() > 0.95).any()
train_data_uncorr = train_data[list(uncorrelated_cols[uncorrelated_cols ==␣
↪True].reset_index()["index"])]
test_data_uncorr = test_data[list(uncorrelated_cols[uncorrelated_cols == True].
↪reset_index()["index"])]

scaler = StandardScaler()
# Scale Training Data
scale_train = pd.DataFrame(scaler.fit_transform(train_data_uncorr))
scale_train.columns = train_data_uncorr.columns
scale_train_data = pd.concat([target_train, scale_train], axis=1)
```

```
# Scale testing data
scale_test = pd.DataFrame(scaler.transform(test_data_uncorr))
scale_test.columns = train_data_uncorr.columns
scale_test_data = pd.concat([target_test, scale_test], axis=1)
```

[12]:
```
f = plt.figure(figsize=(10, 5))
plt.matshow(scale_train_data.corr(), fignum=f.number)
plt.title("Correlation Matrix")
cb = plt.colorbar()
```



[13]:
```
plt.figure(figsize=(15,10))
ax = sns.boxplot(data=scale_train_data.loc[:,scale_train_data.columns !=␣
 ↪"target_price"],orient='h')
plt.title('Boxplot for Processed Training Data')
plt.tight_layout()
plt.show()
```



We can now see that we have much fewer predictors and all are well scaled to be used for prediction purposes. Becuase python does not have a package for Forward/Backward elimination, I will be storing my training data and doing the remaining analysis in R.

[14]:
```
scale_train_data.set_index("date").to_csv("train_data.csv", index=False)
scale_test_data.set_index("date").to_csv("test_data.csv", index=False)
```

b. [20 Points] Model Selection Criterion. Use AIC and BIC criteria to select the best model and report the result from each of them. Use the forward selection for AIC and backward selection for BIC. Report the following mean squared error from **both training and testing data**.

- The mean squared error: $n^{-1}\sum_i(Y_i - \widehat{Y}_i)^2$
- Since these quantities can be affected by scaling and transformation, make sure that you **state any modifications applied to the outcome variable**. Compare the training data errors and testing data errors. Which model works better? Provide a summary of your results.

```
### Load packages
library(MASS)

### Load Data
training_data = read.csv("train_data.csv")
testing_data = read.csv("test_data.csv")
```

```
model = lm(target_price~., data=training_data)
AIC_model = step(model, direction="forward")
```

```
## Start:  AIC=24285.17
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_cost_per_transaction_percent + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd
##     X5_days_prev_btc_trade_volume + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_median_confirmation_time + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_cost_per_transaction_percent + X5_days_prev_btc_output_volume +
##     X5_days_prev_btc_estimated_transaction_volume + X5_days_prev_btc_estimated_transaction_volume_usd
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_cost_per_transaction_percent + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
```

```
y_pred_train <- predict(AIC_model, training_data)
train_mse <- mean((training_data$target_price - y_pred_train)^2, na.rm=TRUE)

y_pred_test <- predict(AIC_model, testing_data)
test_mse <- mean((testing_data$target_price - y_pred_test)^ 2, na.rm=TRUE)

print(paste("Training Error:", train_mse))
```

```
## [1] "Training Error: 16411.159096649"
```

```
print(paste("Testing Error:", test_mse))
```

```
## [1] "Testing Error: 2917743.26443371"
```

```
BIC_model = step(model, direction="backward", k=log(nrow(training_data)))
```

```
## Start:  AIC=24430.82
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_cost_per_transaction_percent + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd
##     X5_days_prev_btc_trade_volume + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_median_confirmation_time + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_cost_per_transaction_percent + X5_days_prev_btc_output_volume +
##     X5_days_prev_btc_estimated_transaction_volume + X5_days_prev_btc_estimated_transaction_volume_usd
```

```
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_cost_per_transaction_percent + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X5_days_prev_btc_median_confirmation_time      1      5159 40983823 24423
## - X5_days_prev_btc_cost_per_transaction_percent  1      5426 40984090 24423
## - X4_days_prev_btc_cost_per_transaction_percent  1      6017 40984682 24423
## - X6_days_prev_btc_cost_per_transaction_percent  1      8100 40986764 24424
## - X5_days_prev_btc_estimated_transaction_volume  1      8395 40987059 24424
## - X5_days_prev_btc_output_volume                 1     22616 41001280 24424
## - X4_days_prev_btc_transaction_fees              1     30665 41009329 24425
## - X6_days_prev_btc_median_confirmation_time      1     31688 41010352 24425
## - X6_days_prev_btc_transaction_fees              1     32590 41011255 24425
## - X5_days_prev_btc_trade_volume                  1     34933 41013597 24425
## - X4_days_prev_btc_trade_volume                  1     50833 41029498 24426
## - X5_days_prev_btc_transaction_fees              1     65988 41044652 24427
## - X6_days_prev_btc_trade_volume                  1     66789 41045453 24427
## - X4_days_prev_btc_median_confirmation_time      1     83521 41062185 24428
## - X4_days_prev_btc_estimated_transaction_volume  1     93174 41071838 24429
## - X6_days_prev_btc_estimated_transaction_volume  1     94028 41072692 24429
## - X4_days_prev_btc_output_volume                 1    105538 41084202 24429
## <none>                                                      40978664 24431
## - X6_days_prev_btc_n_orphaned_blocks             1    138627 41117291 24431
## - X5_days_prev_btc_n_orphaned_blocks             1    146005 41124669 24432
## - X6_days_prev_btc_output_volume                 1    191628 41170292 24435
## - X4_days_prev_btc_n_orphaned_blocks             1    202117 41180782 24435
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    534868 41513532 24455
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2206653 43185317 24554
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2405491 43384155 24565
##
## Step:  AIC=24423.3
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_cost_per_transaction_percent + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd +
##     X5_days_prev_btc_trade_volume + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_cost_per_transaction_percent +
##     X5_days_prev_btc_output_volume + X5_days_prev_btc_estimated_transaction_volume +
##     X5_days_prev_btc_estimated_transaction_volume_usd + X6_days_prev_btc_trade_volume +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_median_confirmation_time +
##     X6_days_prev_btc_transaction_fees + X6_days_prev_btc_cost_per_transaction_percent +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X5_days_prev_btc_cost_per_transaction_percent  1      5459 40989282 24416
## - X4_days_prev_btc_cost_per_transaction_percent  1      6054 40989877 24416
## - X5_days_prev_btc_estimated_transaction_volume  1      7770 40991593 24416
## - X6_days_prev_btc_cost_per_transaction_percent  1      8148 40991972 24416
## - X5_days_prev_btc_output_volume                 1     22287 41006110 24417
## - X4_days_prev_btc_transaction_fees              1     30230 41014053 24417
## - X6_days_prev_btc_transaction_fees              1     32240 41016063 24417
```

```
## - X5_days_prev_btc_trade_volume                       1     35113 41018936 24418
## - X4_days_prev_btc_trade_volume                       1     51088 41034911 24419
## - X6_days_prev_btc_median_confirmation_time           1     61911 41045734 24419
## - X5_days_prev_btc_transaction_fees                   1     65416 41049239 24420
## - X6_days_prev_btc_trade_volume                       1     67217 41051040 24420
## - X4_days_prev_btc_estimated_transaction_volume       1     94407 41078230 24421
## - X6_days_prev_btc_estimated_transaction_volume       1     95388 41079211 24421
## - X4_days_prev_btc_output_volume                      1    105397 41089220 24422
## <none>                                                     40983823 24423
## - X6_days_prev_btc_n_orphaned_blocks                  1    139149 41122972 24424
## - X4_days_prev_btc_median_confirmation_time           1    142495 41126318 24424
## - X5_days_prev_btc_n_orphaned_blocks                  1    146130 41129953 24424
## - X6_days_prev_btc_output_volume                      1    191924 41175747 24427
## - X4_days_prev_btc_n_orphaned_blocks                  1    203246 41187069 24428
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    534570 41518393 24448
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2205939 43189762 24546
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2405280 43389103 24558
##
## Step:  AIC=24415.81
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_cost_per_transaction_percent + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd +
##     X5_days_prev_btc_trade_volume + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_output_volume +
##     X5_days_prev_btc_estimated_transaction_volume + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_cost_per_transaction_percent + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                    Df Sum of Sq      RSS   AIC
## - X4_days_prev_btc_cost_per_transaction_percent     1      6698 40995980 24408
## - X5_days_prev_btc_estimated_transaction_volume     1      7748 40997029 24408
## - X6_days_prev_btc_cost_per_transaction_percent     1      8866 40998148 24408
## - X5_days_prev_btc_output_volume                    1     22280 41011561 24409
## - X4_days_prev_btc_transaction_fees                 1     30144 41019425 24410
## - X6_days_prev_btc_transaction_fees                 1     32140 41021422 24410
## - X5_days_prev_btc_trade_volume                     1     35087 41024369 24410
## - X4_days_prev_btc_trade_volume                     1     51166 41040448 24411
## - X6_days_prev_btc_median_confirmation_time         1     62159 41051441 24412
## - X5_days_prev_btc_transaction_fees                 1     65322 41054604 24412
## - X6_days_prev_btc_trade_volume                     1     67306 41056588 24412
## - X4_days_prev_btc_estimated_transaction_volume     1     94277 41083558 24414
## - X6_days_prev_btc_estimated_transaction_volume     1     95342 41084624 24414
## - X4_days_prev_btc_output_volume                    1    105381 41094663 24414
## <none>                                                   40989282 24416
## - X6_days_prev_btc_n_orphaned_blocks                1    139350 41128631 24416
## - X4_days_prev_btc_median_confirmation_time         1    142857 41132138 24417
## - X5_days_prev_btc_n_orphaned_blocks                1    146325 41135606 24417
## - X6_days_prev_btc_output_volume                    1    191872 41181153 24420
## - X4_days_prev_btc_n_orphaned_blocks                1    203492 41192774 24420
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    534419 41523700 24440
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2205802 43195084 24539
```

```
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2404969 43394251 24550
##
## Step:  AIC=24408.39
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_trade_volume +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_output_volume + X5_days_prev_btc_estimated_transaction_volume +
##     X5_days_prev_btc_estimated_transaction_volume_usd + X6_days_prev_btc_trade_volume +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_median_confirmation_time +
##     X6_days_prev_btc_transaction_fees + X6_days_prev_btc_cost_per_transaction_percent +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X5_days_prev_btc_estimated_transaction_volume      1      7737 41003717 24401
## - X6_days_prev_btc_cost_per_transaction_percent      1      9033 41005013 24401
## - X5_days_prev_btc_output_volume                     1     22272 41018251 24402
## - X4_days_prev_btc_transaction_fees                  1     30038 41026018 24402
## - X6_days_prev_btc_transaction_fees                  1     32018 41027997 24402
## - X5_days_prev_btc_trade_volume                      1     35056 41031036 24403
## - X4_days_prev_btc_trade_volume                      1     51263 41047243 24404
## - X6_days_prev_btc_median_confirmation_time          1     62464 41058444 24404
## - X5_days_prev_btc_transaction_fees                  1     65207 41061186 24404
## - X6_days_prev_btc_trade_volume                      1     67415 41063395 24405
## - X4_days_prev_btc_estimated_transaction_volume      1     94063 41090043 24406
## - X6_days_prev_btc_estimated_transaction_volume      1     95267 41091247 24406
## - X4_days_prev_btc_output_volume                     1    105359 41101339 24407
## <none>                                                     40995980 24408
## - X6_days_prev_btc_n_orphaned_blocks                 1    139594 41135574 24409
## - X4_days_prev_btc_median_confirmation_time          1    143303 41139283 24409
## - X5_days_prev_btc_n_orphaned_blocks                 1    146564 41142544 24410
## - X6_days_prev_btc_output_volume                     1    191808 41187788 24412
## - X4_days_prev_btc_n_orphaned_blocks                 1    203795 41199775 24413
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    534257 41530237 24433
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2205609 43201589 24531
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2404557 43400537 24543
##
## Step:  AIC=24401.04
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_trade_volume +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_output_volume + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_cost_per_transaction_percent + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X6_days_prev_btc_cost_per_transaction_percent      1      9019 41012736 24394
## - X5_days_prev_btc_output_volume                     1     25650 41029367 24395
```

```
## - X4_days_prev_btc_transaction_fees                  1     29662 41033379 24395
## - X6_days_prev_btc_transaction_fees                  1     32050 41035767 24395
## - X5_days_prev_btc_trade_volume                      1     35385 41039102 24395
## - X4_days_prev_btc_trade_volume                      1     51668 41055385 24396
## - X6_days_prev_btc_median_confirmation_time          1     62433 41066150 24397
## - X5_days_prev_btc_transaction_fees                  1     65835 41069553 24397
## - X6_days_prev_btc_trade_volume                      1     67440 41071157 24397
## - X4_days_prev_btc_output_volume                     1    103454 41107171 24400
## <none>                                                     41003717 24401
## - X6_days_prev_btc_n_orphaned_blocks                 1    139618 41143335 24402
## - X4_days_prev_btc_median_confirmation_time          1    143409 41147126 24402
## - X5_days_prev_btc_n_orphaned_blocks                 1    146767 41150484 24402
## - X6_days_prev_btc_output_volume                     1    187052 41190769 24405
## - X4_days_prev_btc_estimated_transaction_volume      1    197549 41201266 24405
## - X6_days_prev_btc_estimated_transaction_volume      1    199428 41203145 24405
## - X4_days_prev_btc_n_orphaned_blocks                 1    203390 41207107 24406
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    541993 41545710 24426
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2304691 43308408 24530
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2517653 43521370 24542
##
## Step:  AIC=24393.76
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_trade_volume +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_output_volume + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                    Df Sum of Sq      RSS   AIC
## - X5_days_prev_btc_output_volume                    1     25635 41038371 24388
## - X4_days_prev_btc_transaction_fees                 1     29523 41042259 24388
## - X6_days_prev_btc_transaction_fees                 1     31886 41044622 24388
## - X5_days_prev_btc_trade_volume                     1     35343 41048080 24388
## - X4_days_prev_btc_trade_volume                     1     51797 41064533 24389
## - X6_days_prev_btc_median_confirmation_time         1     62846 41075582 24390
## - X5_days_prev_btc_transaction_fees                 1     65679 41078416 24390
## - X6_days_prev_btc_trade_volume                     1     67588 41080324 24390
## - X4_days_prev_btc_output_volume                    1    103428 41116165 24392
## <none>                                                    41012736 24394
## - X6_days_prev_btc_n_orphaned_blocks                1    139948 41152684 24394
## - X4_days_prev_btc_median_confirmation_time         1    144001 41156737 24395
## - X5_days_prev_btc_n_orphaned_blocks                1    147087 41159824 24395
## - X6_days_prev_btc_output_volume                    1    186970 41199706 24397
## - X4_days_prev_btc_estimated_transaction_volume     1    197080 41209816 24398
## - X6_days_prev_btc_estimated_transaction_volume     1    199102 41211839 24398
## - X4_days_prev_btc_n_orphaned_blocks                1    203796 41216532 24398
## - X5_days_prev_btc_estimated_transaction_volume_usd 1    541799 41554535 24419
## - X6_days_prev_btc_estimated_transaction_volume_usd 1   2304361 43317097 24522
## - X4_days_prev_btc_estimated_transaction_volume_usd 1   2517054 43529790 24535
##
```

```
## Step:  AIC=24387.5
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_transaction_fees +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_trade_volume +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_estimated_transaction_volume_usd + X6_days_prev_btc_trade_volume +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_median_confirmation_time +
##     X6_days_prev_btc_transaction_fees + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X4_days_prev_btc_transaction_fees             1     28930 41067301 24381
## - X6_days_prev_btc_transaction_fees             1     32641 41071013 24382
## - X5_days_prev_btc_trade_volume                 1     35613 41073985 24382
## - X4_days_prev_btc_trade_volume                 1     52800 41091172 24383
## - X6_days_prev_btc_median_confirmation_time     1     61638 41100010 24383
## - X5_days_prev_btc_transaction_fees             1     66471 41104843 24384
## - X6_days_prev_btc_trade_volume                 1     66521 41104892 24384
## <none>                                                    41038371 24388
## - X6_days_prev_btc_n_orphaned_blocks            1    138911 41177283 24388
## - X4_days_prev_btc_median_confirmation_time     1    144464 41182835 24388
## - X5_days_prev_btc_n_orphaned_blocks            1    148352 41186724 24389
## - X4_days_prev_btc_estimated_transaction_volume 1    195103 41233474 24392
## - X6_days_prev_btc_estimated_transaction_volume 1    203787 41242158 24392
## - X4_days_prev_btc_n_orphaned_blocks            1    203897 41242268 24392
## - X4_days_prev_btc_output_volume                1    228219 41266591 24394
## - X6_days_prev_btc_output_volume                1    366386 41404757 24402
## - X5_days_prev_btc_estimated_transaction_volume_usd 1    519874 41558245 24411
## - X6_days_prev_btc_estimated_transaction_volume_usd 1   2383089 43421460 24521
## - X4_days_prev_btc_estimated_transaction_volume_usd 1   2543655 43582027 24530
##
## Step:  AIC=24381.43
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd +
##     X5_days_prev_btc_trade_volume + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X5_days_prev_btc_trade_volume                 1     34335 41101636 24376
## - X6_days_prev_btc_transaction_fees             1     52233 41119535 24377
## - X4_days_prev_btc_trade_volume                 1     55825 41123126 24377
## - X6_days_prev_btc_median_confirmation_time     1     61204 41128506 24377
## - X6_days_prev_btc_trade_volume                 1     68972 41136273 24378
## <none>                                                    41067301 24381
## - X5_days_prev_btc_transaction_fees             1    129783 41197085 24382
## - X4_days_prev_btc_median_confirmation_time     1    137295 41204596 24382
## - X6_days_prev_btc_n_orphaned_blocks            1    142432 41209734 24382
## - X5_days_prev_btc_n_orphaned_blocks            1    152096 41219397 24383
```

```
## - X4_days_prev_btc_estimated_transaction_volume      1    196444 41263746 24386
## - X6_days_prev_btc_estimated_transaction_volume      1    202302 41269603 24386
## - X4_days_prev_btc_n_orphaned_blocks                 1    207448 41274750 24386
## - X4_days_prev_btc_output_volume                     1    230422 41297724 24388
## - X6_days_prev_btc_output_volume                     1    367272 41434574 24396
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    516416 41583718 24405
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2397085 43464387 24515
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2517406 43584707 24522
##
## Step:  AIC=24375.69
## target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_estimated_transaction_volume_usd + X6_days_prev_btc_trade_volume +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_median_confirmation_time +
##     X6_days_prev_btc_transaction_fees + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X4_days_prev_btc_trade_volume                 1     28540 41130176 24370
## - X6_days_prev_btc_trade_volume                 1     39317 41140954 24370
## - X6_days_prev_btc_transaction_fees             1     49895 41151532 24371
## - X6_days_prev_btc_median_confirmation_time      1     60747 41162384 24372
## - X5_days_prev_btc_transaction_fees             1    125910 41227546 24376
## <none>                                                      41101636 24376
## - X4_days_prev_btc_median_confirmation_time      1    137204 41238841 24376
## - X6_days_prev_btc_n_orphaned_blocks            1    139700 41241337 24376
## - X5_days_prev_btc_n_orphaned_blocks            1    150755 41252391 24377
## - X4_days_prev_btc_estimated_transaction_volume      1    197699 41299335 24380
## - X6_days_prev_btc_estimated_transaction_volume      1    201743 41303379 24380
## - X4_days_prev_btc_n_orphaned_blocks            1    210520 41312156 24381
## - X4_days_prev_btc_output_volume                1    233031 41334667 24382
## - X6_days_prev_btc_output_volume                1    364437 41466073 24390
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    482081 41583718 24397
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2485756 43587392 24514
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   2571770 43673406 24519
##
## Step:  AIC=24369.6
## target_price ~ X4_days_prev_btc_n_orphaned_blocks + X4_days_prev_btc_median_confirmation_time +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_transaction_fees +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X6_days_prev_btc_transaction_fees             1     54124 41184300 24365
## - X6_days_prev_btc_median_confirmation_time      1     61944 41192120 24366
## - X6_days_prev_btc_trade_volume                 1     98148 41228324 24368
## <none>                                                      41130176 24370
## - X5_days_prev_btc_transaction_fees             1    134943 41265120 24370
```

```
## - X4_days_prev_btc_median_confirmation_time          1    137801 41267977 24370
## - X6_days_prev_btc_n_orphaned_blocks                 1    137964 41268141 24370
## - X5_days_prev_btc_n_orphaned_blocks                 1    153943 41284119 24371
## - X4_days_prev_btc_estimated_transaction_volume      1    196354 41326530 24374
## - X6_days_prev_btc_estimated_transaction_volume      1    202936 41333113 24374
## - X4_days_prev_btc_n_orphaned_blocks                 1    214263 41344439 24375
## - X4_days_prev_btc_output_volume                     1    229930 41360106 24376
## - X6_days_prev_btc_output_volume                     1    373083 41503259 24384
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    490870 41621046 24391
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2464786 43594962 24507
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   3266680 44396856 24553
##
## Step:  AIC=24365.06
## target_price ~ X4_days_prev_btc_n_orphaned_blocks + X4_days_prev_btc_median_confirmation_time +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_median_confirmation_time + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X6_days_prev_btc_median_confirmation_time      1     54182 41238482 24360
## - X6_days_prev_btc_trade_volume                  1    115056 41299356 24364
## <none>                                                       41184300 24365
## - X4_days_prev_btc_median_confirmation_time      1    136777 41321077 24366
## - X6_days_prev_btc_n_orphaned_blocks             1    144039 41328339 24366
## - X5_days_prev_btc_n_orphaned_blocks             1    162700 41347000 24367
## - X4_days_prev_btc_estimated_transaction_volume  1    192941 41377241 24369
## - X6_days_prev_btc_estimated_transaction_volume  1    205065 41389366 24370
## - X4_days_prev_btc_n_orphaned_blocks             1    222881 41407181 24371
## - X4_days_prev_btc_output_volume                 1    231332 41415632 24371
## - X6_days_prev_btc_output_volume                 1    378091 41562391 24380
## - X5_days_prev_btc_transaction_fees              1    386371 41570671 24380
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    487497 41671797 24387
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2410676 43594976 24499
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   3261877 44446177 24548
##
## Step:  AIC=24360.52
## target_price ~ X4_days_prev_btc_n_orphaned_blocks + X4_days_prev_btc_median_confirmation_time +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_trade_volume + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                Df Sum of Sq      RSS   AIC
## - X6_days_prev_btc_trade_volume                  1    118433 41356915 24360
## <none>                                                       41238482 24360
## - X6_days_prev_btc_n_orphaned_blocks             1    147172 41385654 24362
## - X5_days_prev_btc_n_orphaned_blocks             1    166861 41405343 24363
## - X4_days_prev_btc_estimated_transaction_volume  1    200177 41438660 24365
## - X6_days_prev_btc_estimated_transaction_volume  1    202745 41441227 24365
```

```
## - X4_days_prev_btc_output_volume                      1    224713 41463195 24366
## - X4_days_prev_btc_n_orphaned_blocks                  1    226896 41465379 24366
## - X5_days_prev_btc_transaction_fees                   1    362084 41600566 24374
## - X6_days_prev_btc_output_volume                      1    372327 41610809 24375
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    487536 41726018 24382
## - X4_days_prev_btc_median_confirmation_time           1   1374787 42613269 24435
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   2450374 43688856 24497
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   3223781 44462264 24541
##
## Step:  AIC=24359.85
## target_price ~ X4_days_prev_btc_n_orphaned_blocks + X4_days_prev_btc_median_confirmation_time +
##     X4_days_prev_btc_output_volume + X4_days_prev_btc_estimated_transaction_volume +
##     X4_days_prev_btc_estimated_transaction_volume_usd + X5_days_prev_btc_n_orphaned_blocks +
##     X5_days_prev_btc_transaction_fees + X5_days_prev_btc_estimated_transaction_volume_usd +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_output_volume +
##     X6_days_prev_btc_estimated_transaction_volume + X6_days_prev_btc_estimated_transaction_volume_usd
##
##                                                    Df Sum of Sq      RSS   AIC
## <none>                                                          41356915 24360
## - X6_days_prev_btc_n_orphaned_blocks                1    156548 41513463 24362
## - X5_days_prev_btc_n_orphaned_blocks                1    173203 41530119 24362
## - X6_days_prev_btc_estimated_transaction_volume     1    201030 41557945 24364
## - X4_days_prev_btc_estimated_transaction_volume     1    202020 41558935 24364
## - X4_days_prev_btc_output_volume                    1    218977 41575893 24365
## - X4_days_prev_btc_n_orphaned_blocks                1    237875 41594791 24366
## - X6_days_prev_btc_output_volume                    1    377004 41733920 24375
## - X5_days_prev_btc_transaction_fees                 1    467117 41824033 24380
## - X5_days_prev_btc_estimated_transaction_volume_usd  1    551299 41908214 24385
## - X4_days_prev_btc_median_confirmation_time          1   1394004 42750919 24435
## - X6_days_prev_btc_estimated_transaction_volume_usd  1   3168932 44525848 24536
## - X4_days_prev_btc_estimated_transaction_volume_usd  1   3187774 44544689 24537
```

```
summary(BIC_model)
```

```
##
## Call:
## lm(formula = target_price ~ X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_median_confirmation_time + X4_days_prev_btc_output_volume +
##     X4_days_prev_btc_estimated_transaction_volume + X4_days_prev_btc_estimated_transaction_volume_usd +
##     X5_days_prev_btc_n_orphaned_blocks + X5_days_prev_btc_transaction_fees +
##     X5_days_prev_btc_estimated_transaction_volume_usd + X6_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_output_volume + X6_days_prev_btc_estimated_transaction_volume +
##     X6_days_prev_btc_estimated_transaction_volume_usd, data = training_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -699.28  -56.82  -36.77   26.02  612.02
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                       229.608      2.582  88.920
## X4_days_prev_btc_n_orphaned_blocks                 11.539      3.053   3.780
## X4_days_prev_btc_median_confirmation_time          27.404      2.995   9.150
## X4_days_prev_btc_output_volume                    -12.517      3.451  -3.627
## X4_days_prev_btc_estimated_transaction_volume     -13.394      3.845  -3.483
```

```
## X4_days_prev_btc_estimated_transaction_volume_usd   93.085      6.727  13.837
## X5_days_prev_btc_n_orphaned_blocks                   9.899      3.069   3.225
## X5_days_prev_btc_transaction_fees                  -17.594      3.322  -5.297
## X5_days_prev_btc_estimated_transaction_volume_usd   45.018      7.823   5.754
## X6_days_prev_btc_n_orphaned_blocks                   9.360      3.053   3.066
## X6_days_prev_btc_output_volume                     -16.436      3.454  -4.759
## X6_days_prev_btc_estimated_transaction_volume      -13.409      3.859  -3.475
## X6_days_prev_btc_estimated_transaction_volume_usd   92.317      6.692  13.796
##                                                   Pr(>|t|)
## (Intercept)                                        < 2e-16 ***
## X4_days_prev_btc_n_orphaned_blocks                0.000161 ***
## X4_days_prev_btc_median_confirmation_time          < 2e-16 ***
## X4_days_prev_btc_output_volume                    0.000293 ***
## X4_days_prev_btc_estimated_transaction_volume     0.000504 ***
## X4_days_prev_btc_estimated_transaction_volume_usd  < 2e-16 ***
## X5_days_prev_btc_n_orphaned_blocks                0.001274 **
## X5_days_prev_btc_transaction_fees                 1.28e-07 ***
## X5_days_prev_btc_estimated_transaction_volume_usd 9.77e-09 ***
## X6_days_prev_btc_n_orphaned_blocks                0.002190 **
## X6_days_prev_btc_output_volume                    2.06e-06 ***
## X6_days_prev_btc_estimated_transaction_volume     0.000520 ***
## X6_days_prev_btc_estimated_transaction_volume_usd  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 129 on 2484 degrees of freedom
##   (7 observations deleted due to missingness)
## Multiple R-squared:  0.7485, Adjusted R-squared:  0.7473
## F-statistic:   616 on 12 and 2484 DF,  p-value: < 2.2e-16
```

```
y_pred_train <- predict(BIC_model, training_data)
train_mse <- mean((training_data$target_price - y_pred_train)^2, na.rm=TRUE)

y_pred_test <- predict(BIC_model, testing_data)
test_mse <- mean((testing_data$target_price - y_pred_test)^ 2, na.rm=TRUE)

print(paste("Training Error:", train_mse))
```

```
## [1] "Training Error: 16562.6413481487"
```

```
print(paste("Testing Error:", test_mse))
```

```
## [1] "Testing Error: 3296021.15984621"
```

From our predictions, we see that our training and testing errors are very high despite optimization with AIC and BIC variable selection. This is most likely because a linear model may not be the correct way to model such high variability time series data. For model improvement it would be better to start testing different model architectures such as Arima to take advantage of the time component of our data. Linear Regression prediction has no concept of moving averages and we are just using 3 days of data to predict the 7th, rather than looking at the general trends within the data.

Alhough poor performance, we got better predictive ability with a slightly lower testing error wih our BIC Model

c. [10 Points] Best Subset Selection. Fit the best subset selection to the dataset and report the best model of each model size (up to 7 variables, excluding the intercept) and their prediction errors. Make sure that you simplify your output to only present the essential information. If the algorithm cannot handle

this many variables, then consider using just day 1 and 2 information. You can use `leaps` package for subset selection.

```
library(leaps)
subset_model <- regsubsets(target_price~., data=training_data, nvmax=7)
summary(subset_model)
```

```
## Subset selection object
## Call: regsubsets.formula(target_price ~ ., data = training_data, nvmax = 7)
## 24 Variables  (and intercept)
##                                              Forced in Forced out
## X4_days_prev_btc_trade_volume                    FALSE      FALSE
## X4_days_prev_btc_n_orphaned_blocks               FALSE      FALSE
## X4_days_prev_btc_median_confirmation_time        FALSE      FALSE
## X4_days_prev_btc_transaction_fees                FALSE      FALSE
## X4_days_prev_btc_cost_per_transaction_percent    FALSE      FALSE
## X4_days_prev_btc_output_volume                   FALSE      FALSE
## X4_days_prev_btc_estimated_transaction_volume    FALSE      FALSE
## X4_days_prev_btc_estimated_transaction_volume_usd FALSE     FALSE
## X5_days_prev_btc_trade_volume                    FALSE      FALSE
## X5_days_prev_btc_n_orphaned_blocks               FALSE      FALSE
## X5_days_prev_btc_median_confirmation_time        FALSE      FALSE
## X5_days_prev_btc_transaction_fees                FALSE      FALSE
## X5_days_prev_btc_cost_per_transaction_percent    FALSE      FALSE
## X5_days_prev_btc_output_volume                   FALSE      FALSE
## X5_days_prev_btc_estimated_transaction_volume    FALSE      FALSE
## X5_days_prev_btc_estimated_transaction_volume_usd FALSE     FALSE
## X6_days_prev_btc_trade_volume                    FALSE      FALSE
## X6_days_prev_btc_n_orphaned_blocks               FALSE      FALSE
## X6_days_prev_btc_median_confirmation_time        FALSE      FALSE
## X6_days_prev_btc_transaction_fees                FALSE      FALSE
## X6_days_prev_btc_cost_per_transaction_percent    FALSE      FALSE
## X6_days_prev_btc_output_volume                   FALSE      FALSE
## X6_days_prev_btc_estimated_transaction_volume    FALSE      FALSE
## X6_days_prev_btc_estimated_transaction_volume_usd FALSE     FALSE
## 1 subsets of each size up to 7
## Selection Algorithm: exhaustive
##          X4_days_prev_btc_trade_volume X4_days_prev_btc_n_orphaned_blocks
## 1  ( 1 ) " "                           " "
## 2  ( 1 ) " "                           " "
## 3  ( 1 ) " "                           " "
## 4  ( 1 ) " "                           "*"
## 5  ( 1 ) " "                           "*"
## 6  ( 1 ) " "                           "*"
## 7  ( 1 ) " "                           "*"
##          X4_days_prev_btc_median_confirmation_time
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) "*"
## 7  ( 1 ) "*"
##          X4_days_prev_btc_transaction_fees
## 1  ( 1 ) " "
```

```
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X4_days_prev_btc_cost_per_transaction_percent
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X4_days_prev_btc_output_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X4_days_prev_btc_estimated_transaction_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) "*"
## 5  ( 1 ) "*"
## 6  ( 1 ) "*"
## 7  ( 1 ) "*"
##          X4_days_prev_btc_estimated_transaction_volume_usd
## 1  ( 1 ) "*"
## 2  ( 1 ) "*"
## 3  ( 1 ) "*"
## 4  ( 1 ) "*"
## 5  ( 1 ) "*"
## 6  ( 1 ) "*"
## 7  ( 1 ) "*"
##          X5_days_prev_btc_trade_volume X5_days_prev_btc_n_orphaned_blocks
## 1  ( 1 ) " "                           " "
## 2  ( 1 ) " "                           " "
## 3  ( 1 ) " "                           " "
## 4  ( 1 ) " "                           " "
## 5  ( 1 ) " "                           " "
## 6  ( 1 ) " "                           " "
## 7  ( 1 ) " "                           " "
##          X5_days_prev_btc_median_confirmation_time
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
```

```
##          X5_days_prev_btc_transaction_fees
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) "*"
##          X5_days_prev_btc_cost_per_transaction_percent
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X5_days_prev_btc_output_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X5_days_prev_btc_estimated_transaction_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X5_days_prev_btc_estimated_transaction_volume_usd
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X6_days_prev_btc_trade_volume X6_days_prev_btc_n_orphaned_blocks
## 1  ( 1 ) " "                           " "
## 2  ( 1 ) " "                           " "
## 3  ( 1 ) " "                           " "
## 4  ( 1 ) " "                           " "
## 5  ( 1 ) " "                           " "
## 6  ( 1 ) " "                           " "
## 7  ( 1 ) " "                           " "
##          X6_days_prev_btc_median_confirmation_time
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
```

```
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X6_days_prev_btc_transaction_fees
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X6_days_prev_btc_cost_per_transaction_percent
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X6_days_prev_btc_output_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) "*"
## 4  ( 1 ) " "
## 5  ( 1 ) "*"
## 6  ( 1 ) "*"
## 7  ( 1 ) "*"
##          X6_days_prev_btc_estimated_transaction_volume
## 1  ( 1 ) " "
## 2  ( 1 ) " "
## 3  ( 1 ) " "
## 4  ( 1 ) " "
## 5  ( 1 ) " "
## 6  ( 1 ) " "
## 7  ( 1 ) " "
##          X6_days_prev_btc_estimated_transaction_volume_usd
## 1  ( 1 ) " "
## 2  ( 1 ) "*"
## 3  ( 1 ) "*"
## 4  ( 1 ) "*"
## 5  ( 1 ) "*"
## 6  ( 1 ) "*"
## 7  ( 1 ) "*"
```

```
best_subset <- lm(target_price~X4_days_prev_btc_trade_volume+X4_days_prev_btc_n_orphaned_blocks+X4_days_
                    X5_days_prev_btc_n_orphaned_blocks+X6_days_prev_btc_n_orphaned_blocks+X6_days_prev_b

summary(best_subset)
```

```
##
## Call:
## lm(formula = target_price ~ X4_days_prev_btc_trade_volume + X4_days_prev_btc_n_orphaned_blocks +
##     X4_days_prev_btc_transaction_fees + X5_days_prev_btc_n_orphaned_blocks +
##     X6_days_prev_btc_n_orphaned_blocks + X6_days_prev_btc_transaction_fees,
##     data = training_data)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1294.08   -92.53   -44.59    63.21   765.64
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        229.593      3.608  63.637  < 2e-16 ***
## X4_days_prev_btc_trade_volume      123.353      3.898  31.642  < 2e-16 ***
## X4_days_prev_btc_n_orphaned_blocks  23.136      4.248   5.446 5.67e-08 ***
## X4_days_prev_btc_transaction_fees   38.899      5.140   7.567 5.33e-14 ***
## X5_days_prev_btc_n_orphaned_blocks  21.491      4.275   5.028 5.32e-07 ***
## X6_days_prev_btc_n_orphaned_blocks  24.117      4.240   5.689 1.43e-08 ***
## X6_days_prev_btc_transaction_fees   47.214      5.125   9.212  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 180.3 on 2490 degrees of freedom
##   (7 observations deleted due to missingness)
## Multiple R-squared:  0.5078, Adjusted R-squared:  0.5066
## F-statistic: 428.1 on 6 and 2490 DF,  p-value: < 2.2e-16
```

```
y_pred_train <- predict(best_subset, training_data)
train_mse <- mean((training_data$target_price - y_pred_train)^2, na.rm=TRUE)

y_pred_test <- predict(best_subset, testing_data)
test_mse <- mean((testing_data$target_price - y_pred_test)^ 2, na.rm=TRUE)

print(paste("Training Error:", train_mse))
```

```
## [1] "Training Error: 32411.5682936696"
```

```
print(paste("Testing Error:", test_mse))
```

```
## [1] "Testing Error: 8453463.49940621"
```

We see a similar trend here that we have a very large errors again for reasons similar to the previous problem. We can also see that some of our predictors have very high beta values, we could potentially get better results if we used regularization to stop overweighting on these values.