SESSION 13: Decision Tree Based Models

Assignment 3

Data Analytics

Table of Contents

1. Problem Statement

Use the given link below:
https://archive.ics.uci.edu/ml/machine-learning-databases/00304/

Problem- prediction of the number of comments in the upcoming 24 hours on those blogs, the train data was generated from different base times that may temporally overlap. Therefore, if you simply split the train into disjoint partitions, the underlying time intervals may overlap. Therefore, the you should use the provided, temporally disjoint train and test splits to ensure that the evaluation is fair.

a) Interpret the final model coefficients.
b) Plot the model result and compare it with assumptions of the model.

2. Solution

a) Interpret the final model coefficients.

The R-script for the given problem is as follows:

```
library(data.table)
library(foreach)
library(readr)
library(dplyr)

setwd("F:/ACADGILD - Online Course/1. DATA SETS/BlogFeedback")
getwd()

blogData_train <- read_csv("F:/ACADGILD - Online Course/1. DATA
SETS/BlogFeedback/blogData_train.csv")

# retrieve filenames of test sets
test_filenames = list.files(pattern = "blogData_test")

# load and combine dataset
train = fread("blogData_train.csv")
fbtest = foreach(i = 1:length(test_filenames), .combine = rbind) %do%
  { temp = fread(test_filenames[i], header = F)

}
```

```r
# Assign variable names to the train and test data set

colnames(blogData_train) <-

c("plikes","checkin","talking","category","d5","d6","d7","d8","d9","d10","d11","d12",

"d13","d14","d15","d16","d17","d18","d19","d20","d21","d22","d23","d24","d25","d26",

"d27","d28","d29","cc1","cc2","cc3","cc4","cc5","basetime","postlength","postshre",

"postpromo","Hhrs","sun","mon","tue","wed","thu","fri","sat","basesun","basemon",
                "basetue","basewed","basethu","basefri","basesat","target")
dim(blogData_train)
dim(fbtest)
View(blogData_train)
View(fbtest)
str(blogData_train)
str(fbtest)

train <- blogData_train; test <- fbtest
head(train); head(test)

# making the data tidy by constructing single collumn for post publish day
train$pubday<- ifelse(train$sun ==1, 1, ifelse(train$mon ==1, 2, ifelse(train$tue ==1,
                                3, ifelse(train$wed ==1, 4, ifelse(train$thu
==1, 5, ifelse(train$fri ==1, 6,

ifelse(train$sat ==1, 7, NA)))))))
# making the data tidy by constructing single collumn for base day
train$baseday<- ifelse(train$basesun ==1, 1, ifelse(train$basemon ==1,
2, ifelse(train$basetue ==1, 3,
                                        ifelse(train$basewed ==1, 4,
ifelse(train$basethu ==1, 5,
library(MASS)
final_model <- lm(target ~ checkin + talking + d5 + d6 + d7 + d8 + d9 + d10 + d11 +
            d12 + d13 + d16 + d17 + d19 + d20 + d21 + d22 + d23 + d24 +
            cc1 + cc2 + cc3 + cc4 + basetime + postshre + Hhrs + wed +
            thu + fri + basemon + basewed, data = train)
summary(final_model)

# a interpret the final model coefficients
summary(final_model) coef(final_model) #
coefficients of the model
```

The output of the R-Script (from Console window) is given as follows:

```
> library(data.table)
> library(foreach)
> library(readr)
> library(dplyr)
>
>
> setwd("F:/ACADGILD - Online Course/1. DATA
SETS/BlogFeedback")
> getwd()
[1] " F:/ACADGILD - Online Course/1. DATA
SETS/BlogFeedback "
>
> blogData_train <- read_csv("F:/ACADGILD - Online Course/1.
DATA SETS/BlogFeedback /blogData_train.csv")
Parsed with column specification:
cols(
   .default = col_double()
)
See spec(...) for full column specifications.
|============================================================================
 =============| 100%   62 MB
> # retrieve filenames of test sets
> test_filenames = list.files(pattern = "blogData_test")
>
> # load and combine dataset
> train = fread("blogData_train.csv")
> fbtest = foreach(i = 1:length(test_filenames), .combine = rbind) %do% {
+ temp = fread(test_filenames[i], header = F)
+ }
>
> # Assign variable names to the train and test data set
> colnames(blogData_train) <-
c("plikes","checkin","talking","category","d5","d6","d7","d8","d9","d10","d11
","d12",
+
"d13","d14","d15","d16","d17","d18","d19","d20","d21","d22","d23","d24","d25"
,"d26",
+
"d27","d28","d29","cc1","cc2","cc3","cc4","cc5","basetime","postlength","post
shre",
+
"postpromo","Hhrs","sun","mon","tue","wed","thu","fri","sat","basesun","basem
on",
+
"basetue","basewed","basethu","basefri","basesat","target")
> dim(blogData_train)
[1] 52396   281
> dim(fbtest)
[1] 7624 281
> View(blogData_train)
> View(fbtest)
> str(blogData_train)
Classes 'spec_tbl_df', 'tbl_df', 'tbl' and 'data.frame':    52396 obs. of
281 variables:
 $ plikes    : num   40.3  40.3 40.3   40.3 40.3 ...
 $ checkin   : num   53.8   53.8 53.8   53.8 53.8 ...
 $ talking   : num   0 0  0 0 0 0 0  0 0 ...
 $ category  : num   401   401 401 401  401 401  401 401 401 401 ...
```

```
 $ d5         : num   15 15   15 15 15   15 15 15 15 15 ...
 $ d6         : num   15.5 15.5 15.5   15.5 15.5 ...
 $ d7         : num   32.4 32.4 32.4   32.4 32.4 ...
 $ d8         : num   0 0 0  0 0 0 0 0   0 0 ...
 $ d9         : num   377  377 377 377   377 377 377 377 377 377 ...
 $ d10        : num   3 3 3   3 3 3 3 3   3 3 ...
 $ d11        : num   14 14   14 14 14 ...
 $ d12        : num   32.6 32.6 32.6   32.6 32.6 ...
 $ d13        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ d14        : num   377  377 377 377   377 377 377 377 377 377 ...
 $ d15        : num   2 2 2   2 2 2 2 2   2 2 ...
 $ d16        : num   34.6 34.6 34.6   34.6 34.6 ...
 $ d17        : num   48.5 48.5 48.5   48.5 48.5 ...
 $ d18        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ d19        : num   378  378 378 378   378 378 378 378 378 378 ...
 $ d20        : num   12 12   12 12 12   12 12 12 12 12 ...
 $ d21        : num   1.48 1.48 1.48   1.48 1.48 ...
 $ d22        : num   46.2 46.2 46.2   46.2 46.2 ...
 $ d23        : num   -356 -356 -356   -356 -356 -356 -356 -356 -356 -356 ...
 $ d24        : num   377  377 377 377   377 377 377 377 377 377 ...
 $ d25        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ d26        : num   1.08 1.08 1.08   1.08 1.08 ...
 $ d27        : num   1.8 1.8  1.8 1.8   1.8 ...
 $ d28        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ d29        : num   11 11   11 11 11   11 11 11 11 11 ...
 $ cc1        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ cc2        : num   0.4 0.4  0.4 0.4   0.4 ...
 $ cc3        : num   1.08 1.08 1.08   1.08 1.08 ...
 $ cc4        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ cc5        : num   9 9 9   9 9 9 9 9   9 9 ...
 $ basetime   : num   0 0 0  0 0 0 0 0   0 0 ...
 $ postlength : num   0.378  0.378 0.378   0.378 0.378 ...
 $ postshre   : num   1.07 1.07 1.07   1.07 1.07 ...
 $ postpromo  : num   0 0 0  0 0 0 0 0   0 0 ...
 $ Hhrs       : num   9 9 9   9 9 9 9 9   9 9 ...
 $ sun        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ mon        : num   0.973  0.973 0.973   0.973 0.973 ...
 $ tue        : num   1.7  1.7 1.7 1.7   1.7 ...
 $ wed        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ thu        : num   10 10   10 10 10   10 10 10 10 10 ...
 $ fri        : num   0 0 0  0 0 0 0 0   0 0 ...
 $ sat        : num   0.0229 0.0229 0.0229 0.0229 0.0229 ...
 $ basesun    : num   1.52 1.52 1.52   1.52 1.52 ...
 $ basemon    : num   -8 -8 -8 -8 -8   -8 -8 -8 -8 -8 ...
 $ basetue    : num   9 9 9   9 9 9 9 9   9 9 ...
 $ basewed    : num   0 0 0  0 0 0 0 0   0 0 ...
 $ basethu    : num   6 6 2   3 6 6 3 30 30 0 ...
 $ basefri    : num   2 2 2   1 0 0 1 27 27 0 ...
 $ basesat    : num   4 4 0   2 2 2 2 1   1 0 ...
 $ target     : num   5 5 2   2 5 5 2 2   2 0 ...
 $ NA         : num   -2 -2 2 -1 -2 -2 -1 26 26 0 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 2 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 2 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 0 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 2 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 2 ...
 $ NA         : num   35 35   10 34 59   59 34 58 58 11 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 0 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 0 ...
 $ NA         : num   0 0 0  0 0 0 0 0   0 0 ...
```

```
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
$ NA            : num   0 0 0 0 0 0 0 0 0 00 ...
 [list output truncated]
- attr(*, "spec")=
 .. cols(
 ..    `40.30467` = col_double(),
 ..    `53.845657` = col_double(),
 ..    `0.0` = col_double(),
 ..    `401.0` = col_double(),
 ..    `15.0` = col_double(),
 ..    `15.52416` = col_double(),
 ..    `32.44188` = col_double(),
 ..    `0.0_1` = col_double(),
 ..    `377.0` = col_double(),
 ..    `3.0` = col_double(),
 ..    `14.044226` = col_double(),
 ..    `32.615417` = col_double(),
 ..    `0.0_2` = col_double(),
 ..    `377.0_1` = col_double(),
 ..    `2.0` = col_double(),
 ..    `34.567566` = col_double(),
 ..    `48.475178` = col_double(),
 ..    `0.0_3` = col_double(),
 ..    `378.0` = col_double(),
 ..    `12.0` = col_double(),
 ..    `1.4799345` = col_double(),
 ..    `46.18691` = col_double(),
```

```
..    `-356.0` = col_double(),
..    `377.0_2` = col_double(),
..    `0.0_4` = col_double(),
..    `1.0761671` = col_double(),
..    `1.795416` = col_double(),
..    `0.0_5` = col_double(),
..    `11.0` = col_double(),
..    `0.0_6` = col_double(),
..    `0.4004914` = col_double(),
..    `1.0780969` = col_double(),
..    `0.0_7` = col_double(),
..    `9.0` = col_double(),
..    `0.0_8` = col_double(),
..    `0.37755936` = col_double(),
..    `1.07421` = col_double(),
..    `0.0_9` = col_double(),
..    `9.0_1` = col_double(),
..    `0.0_10` = col_double(),
..    `0.972973` = col_double(),
..    `1.704671` = col_double(),
..    `0.0_11` = col_double(),
..    `10.0` = col_double(),
..    `0.0_12` = col_double(),
..    `0.022932023` = col_double(),
..    `1.521174` = col_double(),
..    `-8.0` = col_double(),
..    `9.0_2` = col_double(),
..    `0.0_13` = col_double(),
..    `2.0_1` = col_double(),
..    `2.0_2` = col_double(),
..    `0.0_14` = col_double(),
..    `2.0_3` = col_double(),
..    `2.0_4` = col_double(),
..    `0.0_15` = col_double(),
..    `0.0_16` = col_double(),
..    `0.0_17` = col_double(),
..    `0.0_18` = col_double(),
..    `0.0_19` = col_double(),
..    `10.0_1` = col_double(),
..    `0.0_20` = col_double(),
..    `0.0_21` = col_double(),
..    `0.0_22` = col_double(),
..    `0.0_23` = col_double(),
..    `0.0_24` = col_double(),
..    `0.0_25` = col_double(),
..    `0.0_26` = col_double(),
..    `0.0_27` = col_double(),
..    `0.0_28` = col_double(),
..    `0.0_29` = col_double(),
..    `0.0_30` = col_double(),
..    `0.0_31` = col_double(),
..    `0.0_32` = col_double(),
..    `0.0_33` = col_double(),
..    `0.0_34` = col_double(),
..    `0.0_35` = col_double(),
..    `0.0_36` = col_double(),
..    `0.0_37` = col_double(),
..    `0.0_38` = col_double(),
..    `0.0_39` = col_double(),
..    `0.0_40` = col_double(),
```

```
..    `0.0_41`  = col_double(),
..    `0.0_42`  = col_double(),
..    `0.0_43`  = col_double(),
..    `0.0_44`  = col_double(),
..    `0.0_45`  = col_double(),
..    `0.0_46`  = col_double(),
..    `0.0_47`  = col_double(),
..    `0.0_48`  = col_double(),
..    `0.0_49`  = col_double(),
..    `0.0_50`  = col_double(),
..    `0.0_51`  = col_double(),
..    `0.0_52`  = col_double(),
..    `0.0_53`  = col_double(),
..    `0.0_54`  = col_double(),
..    `0.0_55`  = col_double(),
..    `0.0_56`  = col_double(),
..    `0.0_57`  = col_double(),
..    `0.0_58`  = col_double(),
..    `0.0_59`  = col_double(),
..    `0.0_60`  = col_double(),
..    `0.0_61`  = col_double(),
..    `0.0_62`  = col_double(),
..    `0.0_63`  = col_double(),
..    `0.0_64`  = col_double(),
..    `0.0_65`  = col_double(),
..    `0.0_66`  = col_double(),
..    `0.0_67`  = col_double(),
..    `0.0_68`  = col_double(),
..    `0.0_69`  = col_double(),
..    `0.0_70`  = col_double(),
..    `0.0_71`  = col_double(),
..    `0.0_72`  = col_double(),
..    `0.0_73`  = col_double(),
..    `0.0_74`  = col_double(),
..    `0.0_75`  = col_double(),
..    `0.0_76`  = col_double(),
..    `0.0_77`  = col_double(),
..    `0.0_78`  = col_double(),
..    `0.0_79`  = col_double(),
..    `0.0_80`  = col_double(),
..    `0.0_81`  = col_double(),
..    `0.0_82`  = col_double(),
..    `0.0_83`  = col_double(),
..    `0.0_84`  = col_double(),
..    `0.0_85`  = col_double(),
..    `0.0_86`  = col_double(),
..    `0.0_87`  = col_double(),
..    `0.0_88`  = col_double(),
..    `0.0_89`  = col_double(),
..    `0.0_90`  = col_double(),
..    `0.0_91`  = col_double(),
..    `0.0_92`  = col_double(),
..    `0.0_93`  = col_double(),
..    `0.0_94`  = col_double(),
..    `0.0_95`  = col_double(),
..    `0.0_96`  = col_double(),
..    `0.0_97`  = col_double(),
..    `0.0_98`  = col_double(),
..    `0.0_99`  = col_double(),
..    `0.0_100` = col_double(),
```

```
..    `0.0_101`  = col_double(),
..    `0.0_102`  = col_double(),
..    `0.0_103`  = col_double(),
..    `0.0_104`  = col_double(),
..    `0.0_105`  = col_double(),
..    `0.0_106`  = col_double(),
..    `0.0_107`  = col_double(),
..    `0.0_108`  = col_double(),
..    `0.0_109`  = col_double(),
..    `0.0_110`  = col_double(),
..    `0.0_111`  = col_double(),
..    `0.0_112`  = col_double(),
..    `0.0_113`  = col_double(),
..    `0.0_114`  = col_double(),
..    `0.0_115`  = col_double(),
..    `0.0_116`  = col_double(),
..    `0.0_117`  = col_double(),
..    `0.0_118`  = col_double(),
..    `0.0_119`  = col_double(),
..    `0.0_120`  = col_double(),
..    `0.0_121`  = col_double(),
..    `0.0_122`  = col_double(),
..    `0.0_123`  = col_double(),
..    `0.0_124`  = col_double(),
..    `0.0_125`  = col_double(),
..    `0.0_126`  = col_double(),
..    `0.0_127`  = col_double(),
..    `0.0_128`  = col_double(),
..    `0.0_129`  = col_double(),
..    `0.0_130`  = col_double(),
..    `0.0_131`  = col_double(),
..    `0.0_132`  = col_double(),
..    `0.0_133`  = col_double(),
..    `0.0_134`  = col_double(),
..    `0.0_135`  = col_double(),
..    `0.0_136`  = col_double(),
..    `0.0_137`  = col_double(),
..    `0.0_138`  = col_double(),
..    `0.0_139`  = col_double(),
..    `0.0_140`  = col_double(),
..    `0.0_141`  = col_double(),
..    `0.0_142`  = col_double(),
..    `0.0_143`  = col_double(),
..    `0.0_144`  = col_double(),
..    `0.0_145`  = col_double(),
..    `0.0_146`  = col_double(),
..    `0.0_147`  = col_double(),
..    `0.0_148`  = col_double(),
..    `0.0_149`  = col_double(),
..    `0.0_150`  = col_double(),
..    `0.0_151`  = col_double(),
..    `0.0_152`  = col_double(),
..    `0.0_153`  = col_double(),
..    `0.0_154`  = col_double(),
..    `0.0_155`  = col_double(),
..    `0.0_156`  = col_double(),
..    `0.0_157`  = col_double(),
..    `0.0_158`  = col_double(),
..    `0.0_159`  = col_double(),
..    `0.0_160`  = col_double(),
```

```
..    `0.0_161`  = col_double(),
..    `0.0_162`  = col_double(),
..    `0.0_163`  = col_double(),
..    `0.0_164`  = col_double(),
..    `0.0_165`  = col_double(),
..    `0.0_166`  = col_double(),
..    `0.0_167`  = col_double(),
..    `0.0_168`  = col_double(),
..    `0.0_169`  = col_double(),
..    `0.0_170`  = col_double(),
..    `0.0_171`  = col_double(),
..    `0.0_172`  = col_double(),
..    `0.0_173`  = col_double(),
..    `0.0_174`  = col_double(),
..    `0.0_175`  = col_double(),
..    `0.0_176`  = col_double(),
..    `0.0_177`  = col_double(),
..    `0.0_178`  = col_double(),
..    `0.0_179`  = col_double(),
..    `0.0_180`  = col_double(),
..    `0.0_181`  = col_double(),
..    `0.0_182`  = col_double(),
..    `0.0_183`  = col_double(),
..    `0.0_184`  = col_double(),
..    `0.0_185`  = col_double(),
..    `0.0_186`  = col_double(),
..    `0.0_187`  = col_double(),
..    `0.0_188`  = col_double(),
..    `0.0_189`  = col_double(),
..    `0.0_190`  = col_double(),
..    `0.0_191`  = col_double(),
..    `0.0_192`  = col_double(),
..    `0.0_193`  = col_double(),
..    `0.0_194`  = col_double(),
..    `0.0_195`  = col_double(),
..    `0.0_196`  = col_double(),
..    `0.0_197`  = col_double(),
..    `0.0_198`  = col_double(),
..    `0.0_199`  = col_double(),
..    `0.0_200`  = col_double(),
..    `0.0_201`  = col_double(),
..    `0.0_202`  = col_double(),
..    `0.0_203`  = col_double(),
..    `0.0_204`  = col_double(),
..    `0.0_205`  = col_double(),
..    `0.0_206`  = col_double(),
..    `0.0_207`  = col_double(),
..    `0.0_208`  = col_double(),
..    `0.0_209`  = col_double(),
..    `0.0_210`  = col_double(),
..    `0.0_211`  = col_double(),
..    `0.0_212`  = col_double(),
..    `0.0_213`  = col_double(),
..    `0.0_214`  = col_double(),
..    `0.0_215`  = col_double(),
..    `0.0_216`  = col_double(),
..    `0.0_217`  = col_double(),
..    `0.0_218`  = col_double(),
..    `0.0_219`  = col_double(),
..    `0.0_220`  = col_double(),
```

```
..    `0.0_221` = col_double(),
..    `0.0_222` = col_double(),
..    `0.0_223` = col_double(),
..    `0.0_224` = col_double(),
..    `1.0`   = col_double(),
..    `0.0_225` = col_double(),
..    `0.0_226` = col_double(),
..    `0.0_227` = col_double(),
..    `0.0_228` = col_double(),
..    `0.0_229` = col_double(),
..    `1.0_1` = col_double(),
..    `0.0_230` = col_double(),
..    `0.0_231` = col_double(),
..    `0.0_232` = col_double(),
..    `0.0_233` = col_double(),
..    `0.0_234` = col_double(),
..    `0.0_235` = col_double(),
..    `0.0_236` = col_double(),
..    `1.0_2` = col_double()
.. )
> str(fbtest)
Classes 'data.table' and 'data.frame':7624 obs. of  281 variables:
 $ v1   : num 10.63 43.44 1.73 27.23 4.5 ...
 $ v144: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v145: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v2   : num 17.88 75.59 3.04 45.97 6.68 ...
 $ v3   : num 1 0 0 0 0 0 0 0 0 0 ...
 $ v142: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v143: num 0 0 1 1 1 1 1 1 0 1 ...
 $ v4   : num 259 634 9 371 18 ...
 $ v5   : num 5 20 0 14 0.5 28 1 87 7.5 0 ...
 $ v6   : num 4.018 15.999 0.733 10.784 3 ...
 $ v7   : num 10.4 44.56 1.53 24.21 4 ...
 $ v146: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v147: num 0 0 0 0 0 1 0 0 0 0 ...
 $ v8   : num 0 0 0 0 0 0 0 0 0 0 ...
 $ v9   : num 235 473 5 228 10 725 179 491 174 0 ...
 $ v148: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v149: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v10  : num 1 2 0 4 0.5 16 0 19.5 1.5 0 ...
 $ v11  : num 3.817 15.47 0.667 9.998 1.333 ...
 $ v150: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v151: num 0 1 1 0 0 1 1 0 0 1 ...
 $ v12  : num 10.3 44.69 1.53 24.4 2.56 ...
 $ v13  : num 0 0 0 0 0 0 0 0 0 0 ...
 $ v152: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v153: num 0 0 1 0 0 1 0 0 0 0 ...
 $ v14  : num 235 473 5 228 7 725 179 491 174 0 ...
 $ v15  : num 1 1 0 2 0 3 0 14 1 0 ...
 $ v154: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v155: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v16  : num 9.78 40.97 1.13 22.56 2.83 ...
 $ v17  : num 16.07 70.31 1.82 39.76 3.67 ...
 $ v156: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v157: num 0 0 0 0 0 0 0 0 0 0 ...
 $ v18  : num 1 0 0 0 0 0 0 0 0 0 ...
 $ v19  : num 192 479 5 337 8 913 189 786 186 0 ...
 $ v158: num 0 0 1 1 0 1 1 0 0 1 ...
 $ v159: num 0 0 1 0 0 1 0 0 0 0 ...
 $ v20  : num 5 18 0 10 0.5 26 0 74 5.5 0 ...
```

```
$ v21 : num  0.201 0.5289 0.0667 0.7866 1.6667 ...
$ v160: num  0 0 0 0 0 0 0 0 0 0 ...
$ v161: num  0 0 0 0 0 0 0 0 0 0 ...
$ v22 : num  13.95 62.13 1.73 30.36 2.21 ...
$ v23 : num  -229 -461 -5 -156 0 -519 -178 -418 -161 0 ...
$ v162: num  0 0 0 0 0 0 0 0 0 0 ...
$ v163: num  0 0 0 0 0 0 0 0 0 0 ...
$ v24 : num  217 473 4 228 6 725 170 491 174 0 ...
$ v25 : num  0 0 0 0 0.5 2 0  -3 0 0 ...
$ v164: num  0 0 0 0 0 0 0 0 0 0 ...
$ v165: num  0 0 0 0 0 0 0 0 0 0 ...
$ v26 : num  0.252 0.193 0.333 0.11 0 ...
$ v27 : num  0.904 0.458 0.699 0.356 0 ...
$ v166: num  0 0 0 0 0 0 0 0 0 0 ...
$ v167: num  0 0 0 0 0 0 0 0 0 0 ...
$ v28 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v29 : num  14 2 2 2 0 0 6 0 1 0 ...
$ v168: num  0 0 0 0 0 0 0 0 0 0 ...
$ v169: num  0 0 0 0 0 0 0 0 0 0 ...
$ v30 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v31 : num  0.0944 0.0733 0.1333 0.0432 0 ...
$ v170: num  0 0 1 0 0 1 0 0 0 0 ...
$ v171: num  0 0 0 0 0 0 0 0 0 0 ...
$ v32 : num  0.507 0.286 0.34 0.215 0 ...
$ v33 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v172: num  0 0 0 0 0 0 0 0 0 0 ...
$ v173: num  0 0 0 0 0 0 0 0 0 0 ...
$ v34 : num  12 2 1 2 0 0 5 0 1 0 ...
$ v35 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v174: num  0 0 0 0 0 0 0 0 1 0 ...
$ v175: num  0 0 0 0 0 0 0 0 0 0 ...
$ v36 : num  0.0919 0.0677 0.1333 0.0408 0 ...
$ v37 : num  0.504 0.278 0.34 0.21 0 ...
$ v176: num  0 0 0 0 0 0 0 0 0 0 ...
$ v177: num  0 0 0 0 0 0 0 0 0 0 ...
$ v38 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v39 : num  12 2 1 2 0 0 5 0 1 0 ...
$ v178: num  0 0 0 0 0 0 0 0 0 0 ...
$ v179: num  0 0 0 0 0 0 0 0 0 0 ...
$ v40 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v41 : num  0.2335 0.1763 0.2 0.0983 0 ...
$ v180: num  0 0 1 0 0 1 1 0 0 0 ...
$ v181: num  0 0 1 0 0 0 0 0 0 0 ...
$ v42 : num  0.855 0.43 0.4 0.321 0 ...
$ v43 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v182: num  0 0 0 0 0 0 0 0 0 0 ...
$ v183: num  0 0 0 0 0 1 0 0 0 0 ...
$ v44 : num  13 2 1 2 0 0 5 0 1 0 ...
$ v45 : num  0 0 0 0 0 0 0 0 0 0 ...
$ v184: num  0 0 0 0 0 0 0 0 0 0 ...
$ v185: num  0 0 0 0 0 0 0 0 0 0 ...
$ v46 : num  0.00245 0.00564 0 0.0024 0 ...
$ v47 : num  0.675 0.404 0.365 0.29 0 ...
$ v186: num  0 0 0 0 0 0 0 0 0 0 ...
$ v187: num  0 0 0 0 0 0 0 0 0 0 ...
$ v48 : num  -10 -2 -1 -2 0 0 -5 0 -1 0 ...
$ v49 : num  12 2 1 2 0 0 5 0 1 0 ...
$ v188: num  0 0 0 0 0 0 0 0 0 0 ...
$ v189: num  0 0 0 0 0 0 0 0 0 0 ...
$ v50 : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
  $ V51 : num  35 21 2 3 0 12 103 61 7 0 ...
   [list output truncated]
 - attr(*, ".internal.selfref")=<externalptr>
>
> train <- blogData_train; test <- fbtest
> head(train); head(test)
# A tibble: 6 x 281
   plikes checkin talking category    d5    d6    d7    d8    d9   d10   d11
d12   d13   d14    d15
    <dbl>   <dbl>   <dbl>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
<dbl> <dbl>  <dbl> <dbl>
1    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
2    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
3    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
4    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
5    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
6    40.3    53.8       0      401    15  15.5  32.4     0   377     3  14.0
32.6     0   377      2
# ... with 266 more variables: d16 <dbl>, d17 <dbl>, d18 <dbl>, d19 <dbl>,
d20 <dbl>, d21 <dbl>,
#   d22 <dbl>, d23 <dbl>, d24 <dbl>, d25 <dbl>, d26 <dbl>, d27 <dbl>, d28
<dbl>, d29 <dbl>, cc1 <dbl>,
#   cc2 <dbl>, cc3 <dbl>, cc4 <dbl>, cc5 <dbl>, basetime <dbl>, postlength
<dbl>, postshre <dbl>,
#   postpromo <dbl>, Hhrs <dbl>, sun <dbl>, mon <dbl>, tue <dbl>, wed <dbl>,
thu <dbl>, fri <dbl>,
#   sat <dbl>, basesun <dbl>, basemon <dbl>, basetue <dbl>, basewed <dbl>,
basethu <dbl>,
#   basefri <dbl>, basesat <dbl>, target <dbl>, NA <dbl>, NA <dbl>, NA <dbl>,
NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA
<dbl>, NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA
<dbl>, NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA
<dbl>, NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA
<dbl>, NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA
<dbl>, NA <dbl>, NA <dbl>,
#   NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, NA <dbl>, ...
            V1 V144 V145       V2 V3 V142 V143   V4   V5         V6        V7
V146 V147 V8 V9  V148
1:   10.630660    0    0 17.882992  1    0    0  259  5.0  4.0182760   10.39679
0      0  0 235     0
2:   43.435825    0    0 75.590485  0    0    0  634 20.0 15.9985895   44.56087
0      0  0 473     0
3:    1.733333    0    0  3.043390  0    0    1    9  0.0  0.7333333    1.52607
0      0  0   5     0
    V149 V10       V11 V150 V151       V12 V13 V152 V153 V14 V15 V154 V155
V16         V17 V156
1:     0 1.0 3.8172395    0    0 10.297346   0    0    0 235   1    0    0
9.776869   16.073494      0
2:     0 2.0 15.4696760   0    1 44.685085   0    0    0 473   1    0    0
40.971790    70.307840      0
```

```
3:       0 0.0   0.6666667      0    1    1.534782    0    0    1    5   0     0     0
1.133333    1.820867     0
   V157 V18    V19  V158 V159   V20           V21 V160 V161       V22    V23 V162
V163 V24 V25 V164 V165
1:      0    1 192    0     0    5.0   0.20103656    0     0    13.948867  -229    0
0 217   0.0    0      0
2:      0    0 479    0     0 18.0   0.52891400    0     0    62.134968  -461    0
0 473   0.0    0      0
3:      0    0   5    1     1    0.0   0.06666667    0     0    1.730767    -5    0
0   4  0.0    0      0
         V26          V27 V166   V167 V28 V29 V168 V169  V30         V31   V170 V171
V32 V33 V172 V173 V34
1: 0.2517731 0.9038038     0     0    0  14     0    0  0 0.09438080     0    0
0.5067316     0    0      0   12
2: 0.1932299 0.4576994     0     0    0   2     0    0  0 0.07334273     0    0
0.2864750     0    0      0    2
3: 0.3333333 0.6992059     0     0    0   2     0    0  0 0.13333334     1    0
0.3399347     0    0      0    1
   V35 V174 V175         V36          V37 V176 V177 V38   V39 V178 V179   V40
V41 V180 V181       V42
1:    0    0     0 0.09192581 0.5042160     0    0    0   12     0    0     0
0.23349700     0    0 0.8547111
2:    0    0     0 0.06770099 0.2778884     0    0    0    2     0    0     0
0.17630465     0    0 0.4297832
3:    0    0     0 0.13333334 0.3399347     0    0    0    1     0    0     0
0.20000000     1    1 0.4000000
   V43 V182 V183   V44 V45 V184 V185         V46         V47 V186 V187   V48 V49
V188   V189 V50 V51 V190 V191
1:    0    0     0   13   0     0    0 0.002454992 0.6747285    0     0 -10   12
0    0    0   35     0     0
2:    0    0     0    2   0     0    0 0.005641749 0.4044489    0     0  -2    2
0    0    0   21     0     0
3:    0    0     0    1   0     0    0 0.000000000 0.3651484    0     0  -1    1
0    0    0    2     0     1
   V52   V53 V192 V193 V54 V55 V194 V195   V56 V57 V196 V197  V58 V59 V198 V199
V60 V61 V200 V201     V62 V63
1:   35    0     0    0  35   35    0    0    0  0    0    0    0   0    0    0
0    9    0     0     0    0
2:    0    2     0    0  21   -2    0    0    0  0    0    0    0   0    0    0
0   62    0     0   696    0
3:    2    0     0    0   2    2    0    0    2  2    0    0    0   2    0    0
2   13    0     0  8361    0
   V202 V203 V64   V65 V204  V205 V66 V67 V206 V207 V68 V69 V208 V209  V70 V71
V210   V211 V72 V73 V212 V213
1:      0    0   0    0    0     0   0   0    0     0   0   0    0    0    0    0
0    0    0    0     0    0
2:      0    0   0    0    0     0   0   0    0     0   0   1    0    0    0    0
0    0    0    0     0    0
3:      1    0   1    0    0     0   0   1    0     1   1   1    0    0    0    0
1    0    0    0     0    1
   V74   V75 V214 V215 V76 V77 V216 V217   V78 V79 V218 V219  V80 V81 V220 V221
V82 V83 V222 V223 V84 V85
1:    0    0     0    0   0    0    0    0    0   0    0    0    0    0    0    0
0    0    0    0    0    0
2:    0    0     0    0   0    0    0    0    0   0    0    0    0    0    0    0
0    0    0    0    0    0
3:    0    0     0    0   0    0    0    0    0   1    0    0    0    0    0    0
0    0    0    0    0    0
   V224 V225 V86   V87 V226  V227 V88 V89 V228 V229 V90 V91 V230 V231   V92 V93
V232   V233 V94 V95 V234 V235
```

```
1:    0     0  0   0     0     0   0  0     0    0    0   0    0    0   0   0
0    0   0    0     0    0
2:    0     0  0   0     1     0   0  0     0    0    0   0    0    0   0   0
1    0   0    0     0    0
3:    0     0  0   0     1     0   0  0     1    0    0   0    0    0   0   0
1    0   0    0     0    0
   V96 V97 V236 V237 V98 V99 V238 V239 V100 V101 V240 V241 V102 V103  V242
V243 V104 V105 V244 V245  V106
1:   0   0    0    0   0   0    0    0    0    0    0    0    0    0    0
0    0    0    0    0    0
2:   0   0    0    0   0   0    0    0    0    1    0    0    0    0    0
0    0    0    0    0    0
3:   1   0    0    0   0   0    0    0    0    1    0    0    1    0    0
0    0    0    0    0    0
   V107  V246 V247 V108 V109 V248 V249 V110 V111  V250 V251 V112 V113 V252 V253
V114 V115 V254 V255 V116
1:    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    0    0
2:    0    1    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    0    0
3:    0    1    0    0    0    1    0    0    0    0    0    0    0    0    0
0    0    1    0    0
   V117  V256 V257 V118 V119 V258 V259 V120 V121  V260 V261 V122 V123 V262 V263
V124 V125 V264 V265 V126
1:    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    1    0
2:    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    1    0
3:    0    0    0    0    0    0    0    1    0    0    0    1    0    0    0
0    0    0    1    0
   V127  V266 V267 V128 V129 V268 V269 V130 V131  V270 V271 V132 V133 V272 V273
V134 V135 V274 V275 V136
1:    0    0    0    0    0    0    0    0    0    0    1    0    0    0    0
0    0    0    0    0
2:    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    0    0
3:    0    0    0    0    0    0    0    0    0    0    1    0    0    0    0
0    0    0    0    0
   V137  V276 V277 V138 V139 V278 V279 V140 V141  V280 V281
1:    0    0    0    0    0    0    0    0    0    0    4
2:    0    1    0    0    1    0    0    0    0    0    0
3:    0    0    0    0    0    0    0    0    0    0    1
 [ reached getOption("max.print") -- omitted 3 rows ]
>
> # making the data tidy by constructing single collumn for post publish day
> train$pubday<- ifelse(train$sun ==1, 1, ifelse(train$mon ==1,
2, ifelse(train$tue ==1, 3,
+
ifelse(train$wed ==1, 4, ifelse(train$thu ==1, 5, ifelse(train$fri ==1, 6,
+
ifelse(train$sat ==1, 7, NA)))))))
> # making the data tidy by constructing single collumn for base day
> train$baseday<- ifelse(train$basesun ==1, 1, ifelse(train$basemon ==1, 2,
ifelse(train$basetue ==1, 3,
+
ifelse(train$basewed ==1, 4, ifelse(train$basethu ==1, 5,
+
ifelse(train$basefri ==1, 6, ifelse(train$basesat ==1, 7, NA)))))))
>
```

```
> # # clean dataset, impute missing values and perform exploratory data
analysis
> #
> # distinct(train)    # removing overlapping observations if any
> # dim(train)
>
>
> library(MASS)
>
> final_model <- lm(target ~ checkin + talking + d5 + d6 + d7 + d8 + d9 + d10
+ d11 +
+                    d12 + d13 + d16 + d17 + d19 + d20 + d21 + d22 + d23 +
d24 +
+                    cc1 + cc2 + cc3 + cc4 + basetime + postshre + Hhrs +
wed +
+                    thu + fri + basemon + basewed, data = train)
> summary(final_model)

Call:
lm(formula = target ~ checkin +  talking + d5 + d6 + d7 + d8 +
    d9 + d10 + d11 + d12 + d13 + d16 + d17 + d19 + d20 + d21 +
    d22 + d23 + d24 + cc1 + cc2 + cc3 + cc4 + basetime + postshre +
    Hhrs + wed + thu + fri + basemon + basewed, data = train)

Residuals:
    Min      1Q  Median      3Q     Max
-561.78  -13.04   -1.83    0.00 1743.64

Coefficients: (2 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.947e-04   5.171e-01  -0.001    0.999
checkin      3.892e-05   1.692e-01   0.000    1.000
talking      1.700e-04   1.203e-01   0.001    0.999
d5           1.263e-05   1.282e-01   0.000    1.000
d6          -9.984e+02   6.649e+05  -0.002    0.999
d7          -1.411e-03   5.473e-01  -0.003    0.998
d8           4.528e-04   3.698e+00   0.000    1.000
d9           3.487e-05   2.347e-02   0.001    0.999
d10         -3.316e-04   1.752e-01  -0.002    0.998
d11          9.984e+02   6.649e+05   0.002    0.999
d12          3.521e-04   3.883e-01   0.001    0.999
d13                 NA          NA      NA       NA
d16          9.999e-01   1.997e-01   5.007 5.55e-07 ***
d17          5.831e-05   1.330e-01   0.000    1.000
d19         -1.190e-05   1.008e-02  -0.001    0.999
d20         -8.603e-05   1.488e-01  -0.001    1.000
d21          9.984e+02   6.649e+05   0.002    0.999
d22          5.252e-04   2.758e-01   0.002    0.998
d23          1.633e-05   1.088e-02   0.002    0.999
d24         -1.133e-06   1.780e-02   0.000    1.000
cc1         -7.536e-03   1.932e+00  -0.004    0.997
cc2          1.402e-02   8.699e+00   0.002    0.999
cc3          2.395e-04   1.436e+01   0.000    1.000
cc4                 NA          NA      NA       NA
basetime    -8.246e-03   1.027e+01  -0.001    0.999
postshre     2.803e-03   1.443e+01   0.000    1.000
Hhrs        -8.483e-04   8.746e-01  -0.001    0.999
wed          8.755e-04   4.810e+00   0.000    1.000
thu          3.968e-04   3.294e-01   0.001    0.999
fri          4.796e-04   1.784e+00   0.000    1.000
```

```
basemon        -2.404e-04 8.184e-01    0.000    1.000
basewed         4.229e-03 2.081e+01    0.000    1.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 67.27 on 52366 degrees of freedom
Multiple R-squared:  0.4976,   Adjusted R-squared:  0.4973
F-statistic:  1788 on 29 and 52366 DF,  p-value: < 2.2e-16


>
>
> # a interpret the final model coefficients
> summary(final_model)

Call:
lm(formula = target ~ checkin + talking + d5 + d6 + d7 + d8 +
    d9 + d10 + d11 + d12 + d13 + d16 + d17 + d19 + d20 + d21 +
    d22 + d23 + d24 + cc1 + cc2 + cc3 + cc4 + basetime + postshre +
    Hhrs + wed + thu + fri + basemon + basewed, data = train)

Residuals:
    Min      1Q  Median      3Q     Max
-561.78  -13.04   -1.83    0.00 1743.64


Coefficients: (2 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.947e-04  5.171e-01  -0.001    0.999
checkin      3.892e-05  1.692e-01   0.000    1.000
talking      1.700e-04  1.203e-01   0.001    0.999
d5           1.263e-05  1.282e-01   0.000    1.000
d6          -9.984e+02  6.649e+05  -0.002    0.999
d7          -1.411e-03  5.473e-01  -0.003    0.998
d8           4.528e-04  3.698e+00   0.000    1.000
d9           3.487e-05  2.347e-02   0.001    0.999
d10         -3.316e-04  1.752e-01  -0.002    0.998
d11          9.984e+02  6.649e+05   0.002    0.999
d12          3.521e-04  3.883e-01   0.001    0.999
d13                 NA         NA      NA       NA
d16          9.999e-01  1.997e-01   5.007 5.55e-07 ***
d17          5.831e-05  1.330e-01   0.000    1.000
d19         -1.190e-05  1.008e-02  -0.001    0.999
d20         -8.603e-05  1.488e-01  -0.001    1.000
d21          9.984e+02  6.649e+05   0.002    0.999
d22          5.252e-04  2.758e-01   0.002    0.998
d23          1.633e-05  1.088e-02   0.002    0.999
d24         -1.133e-06  1.780e-02   0.000    1.000
cc1         -7.536e-03  1.932e+00  -0.004    0.997
cc2          1.402e-02  8.699e+00   0.002    0.999
cc3          2.395e-04  1.436e+01   0.000    1.000
cc4                 NA         NA      NA       NA
basetime    -8.246e-03  1.027e+01  -0.001    0.999
postshre     2.803e-03  1.443e+01   0.000    1.000
Hhrs        -8.483e-04  8.746e-01  -0.001    0.999
wed          8.755e-04  4.810e+00   0.000    1.000
thu          3.968e-04  3.294e-01   0.001    0.999
fri          4.796e-04  1.784e+00   0.000    1.000
basemon     -2.404e-04  8.184e-01   0.000    1.000
basewed      4.229e-03  2.081e+01   0.000    1.000
---
Signif. codes:   0 '***'0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

18

```
Residual standard error: 67.27 on 52366 degrees of freedom
Multiple R-squared:  0.4976,   Adjusted R-squared:   0.4973
F-statistic:  1788 on 29 and 52366 DF,  p-value: < 2.2e-16
```

```
> coef(final_model) # coefficients of the model
   (Intercept)          checkin          talking              d5              d6
d7            d8
-4.946570e-04   3.891451e-05  1.700457e-04 1.262629e-05 -9.983858e+02 -
1.410870e-03  4.528305e-04
           d9             d10              d11             d12             d13
d16           d17
  3.486801e-05  -3.316009e-04  9.983864e+02  3.520834e-04             NA
9.999110e-01  5.831097e-05
           d19             d20              d21             d22             d23
d24           cc1
-1.189822e-05  -8.602563e-05  9.983873e+02  5.251878e-04 1.632576e-05 -
1.133489e-06 -7.535792e-03
           cc2             cc3              cc4         basetime         postshre
Hhrs          wed
  1.402018e-02   2.395188e-04                NA -8.245685e-03 2.802928e-03 -
8.483031e-04  8.754786e-04
           thu             fri          basemon          basewed
  3.967947e-04   4.795834e-04 -2.404461e-04   4.228575e-03
> #  talking               d5               d7               d8              d10
d11
> # -1.858115e-05 -4.759496e-01   8.609203e-01   1.675394e-01 -1.239555e-01 -
2.236221e-03
 > # d12              d13              d16              d17              d19             d20
 d22
> # 1.612318e-01   1.276223e-01  1.114969e-02   1.085186e-01 -1.165972e-01
4.201675e-01 -8.837498e-01
> # d23             cc1              cc2              cc3              cc4         basetime
postshre
> # -2.159461e-01   4.338324e-02   2.196493e-01 -2.272725e-02 -6.728051e-02 -
1.933110e-01   2.921963e-03
> # Hhrs
> # 3.880629e-01
```

b) Plot the model result and compare it with assumptions of the model.

The R-script for the given problem is as follows:

# b plot the model result and compare it with assumptions of the model
par(mfrow=c(2,2))
plot(final_model)

The output of the R-Script (from Console window) is given as follows:

```
> # b plot the model result and compare it with assumptions of the
model
> par(mfrow=c(2,2))
> plot(final_model)
```
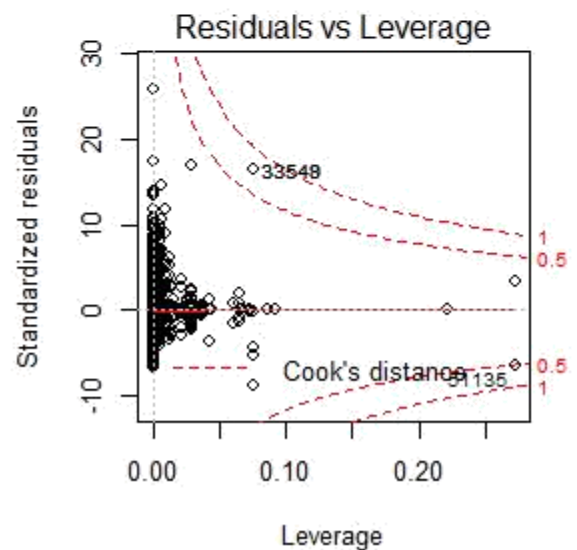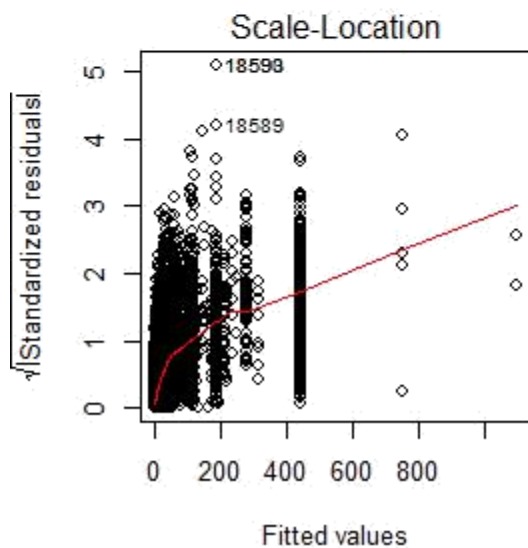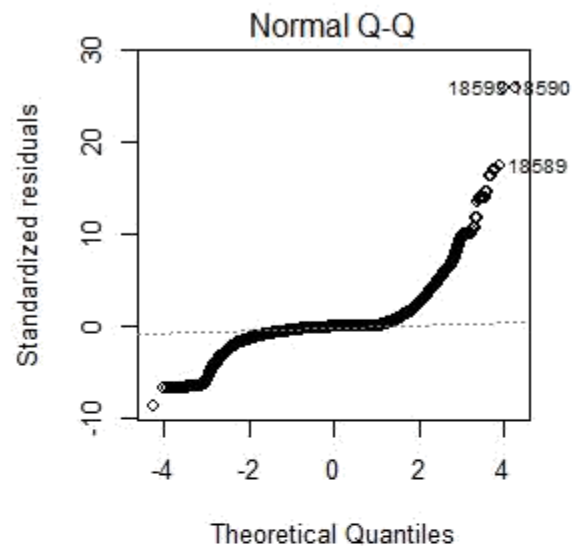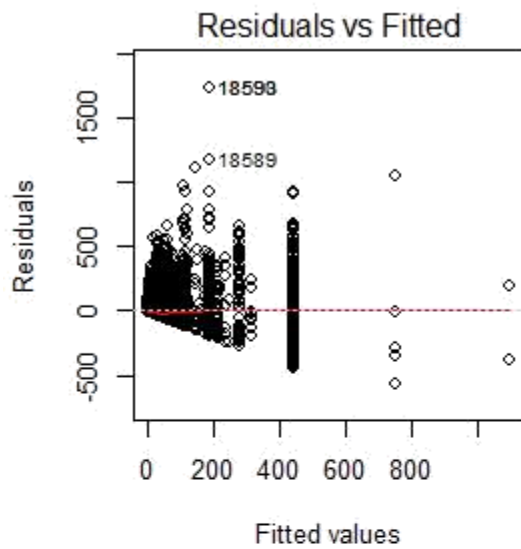
Conclusion/Interpretation:

- Model does not pass the test of normality
- The data is heteroscadatic
- Observations shown may have the leverage or potential for influencing the model