Biostatistics Project

December 15, 2017

1 Final Exam

CS 555 Multivariate Data Analysis, Fall 2017 Due on December 15, 2017 100 total points

Use R to perform all necessary calculations. Attach your code and output. Give interpretation and discuss all relevant statistical measures.

2 Problem 1.

(25 points) We are interested in comparing four characteristics (sepal length, sepal width, petal length and petal width) of three flower species Setosa, Versicolor and Virginica. We want to analyze the dataset "iris" available in the package MVN. Perform all necessary data analysis steps and write a section summarizing the findings.

2.1 Load data and basic analysis

```
In [2]: # load data
        library(car)
        library(MVN)
        data("iris")
        head(iris)
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa

```
In [3]: # understand the data
        dim(iris)
        table(iris$Species)
        summary(iris)
        str(iris)
```

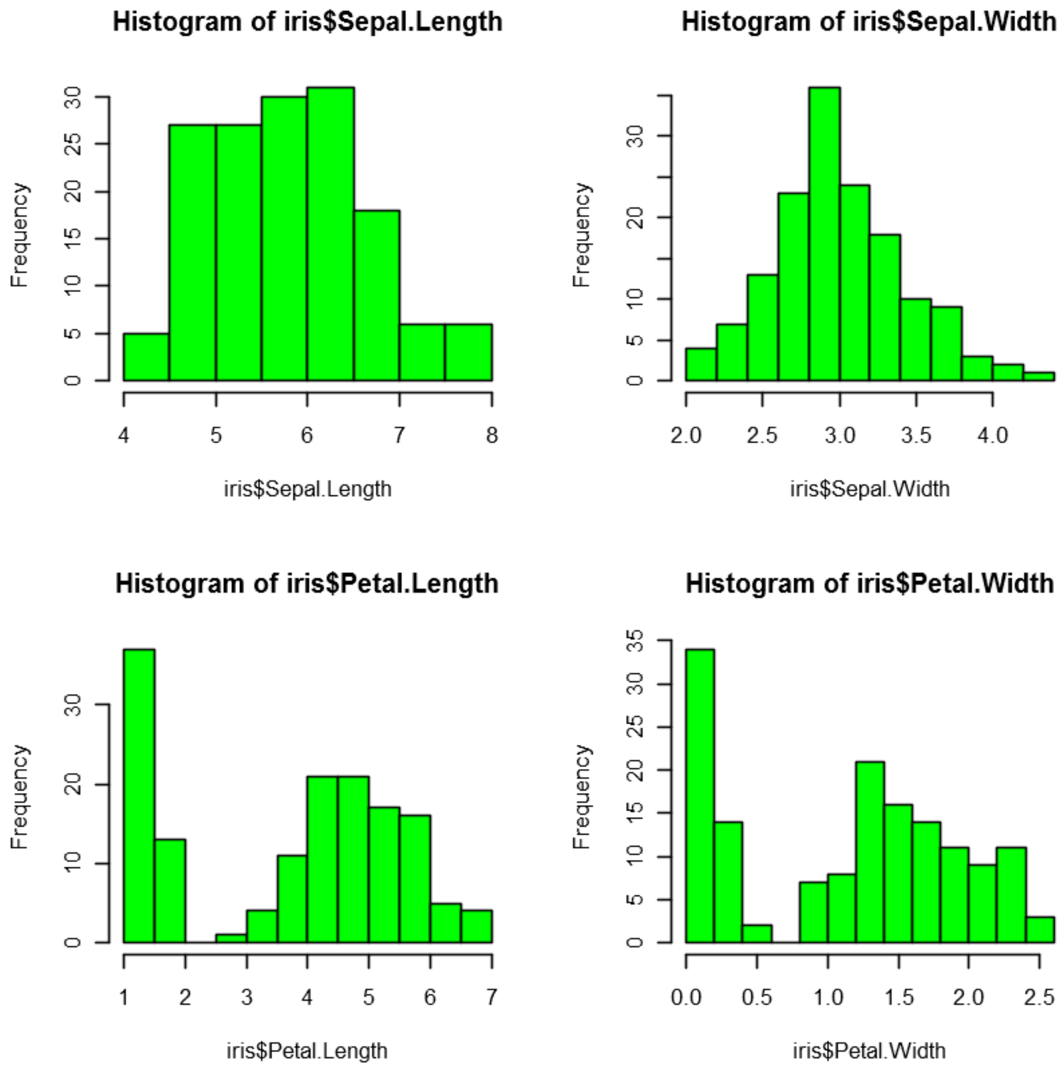
1. 150 2. 5

```
    setosa versicolor  virginica 
        50         50         50 
```

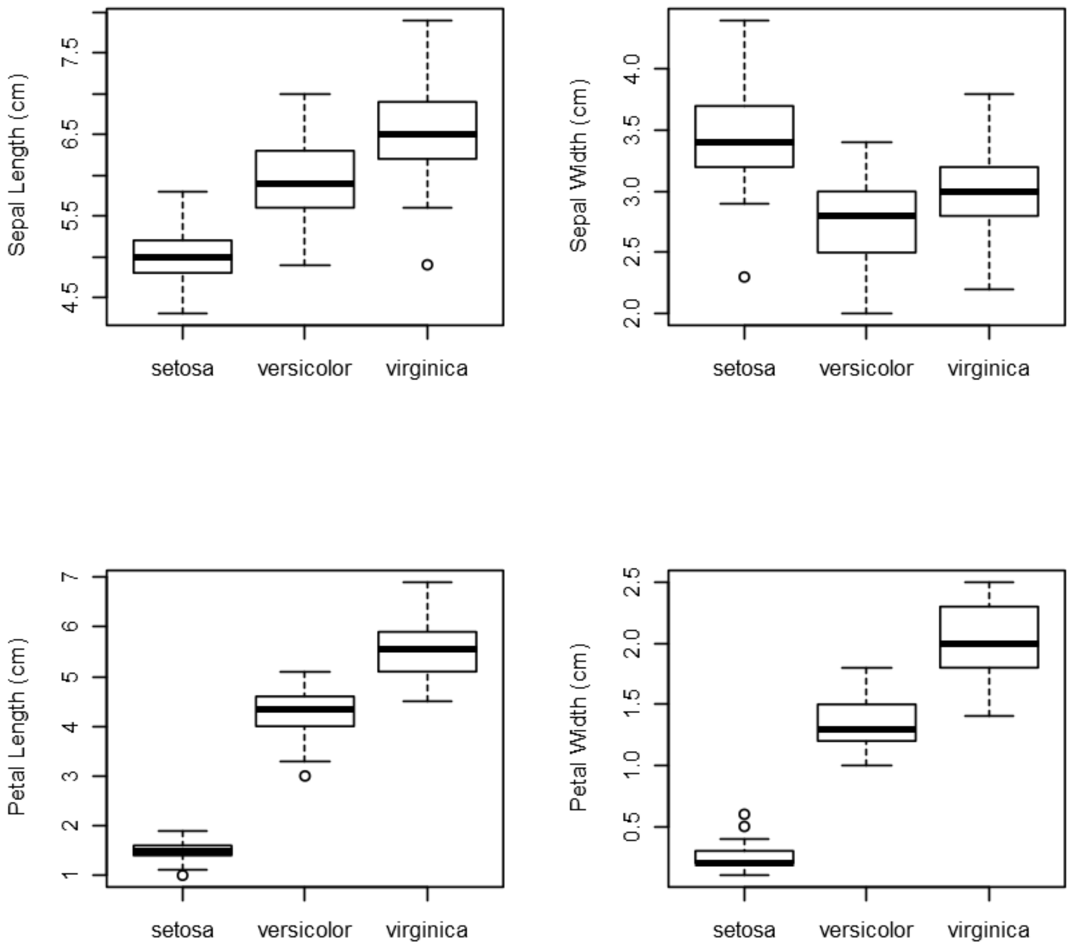
```
  Sepal.Length    Sepal.Width     Petal.Length    Petal.Width   
 Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100  
 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300  
 Median :5.800   Median :3.000   Median :4.350   Median :1.300  
 Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199  
 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800  
 Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500  
       Species  
 setosa    :50  
 versicolor:50  
 virginica :50  
```

```
'data.frame':	150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

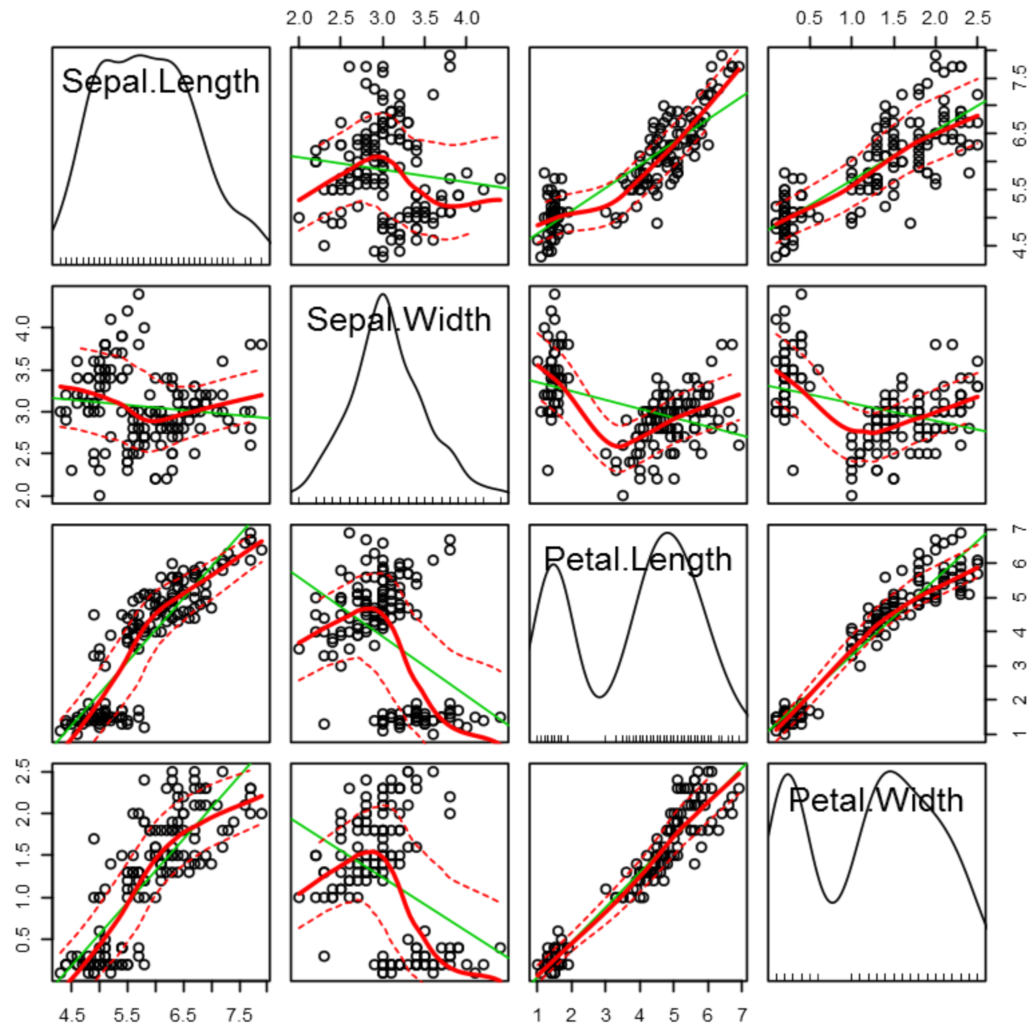
In [13]:
```
# histgram
par(mfrow=c(2,2))
hist(iris$Sepal.Length,col='green')
hist(iris$Sepal.Width,col='green')
hist(iris$Petal.Length,col='green')
hist(iris$Petal.Width,col='green')
```



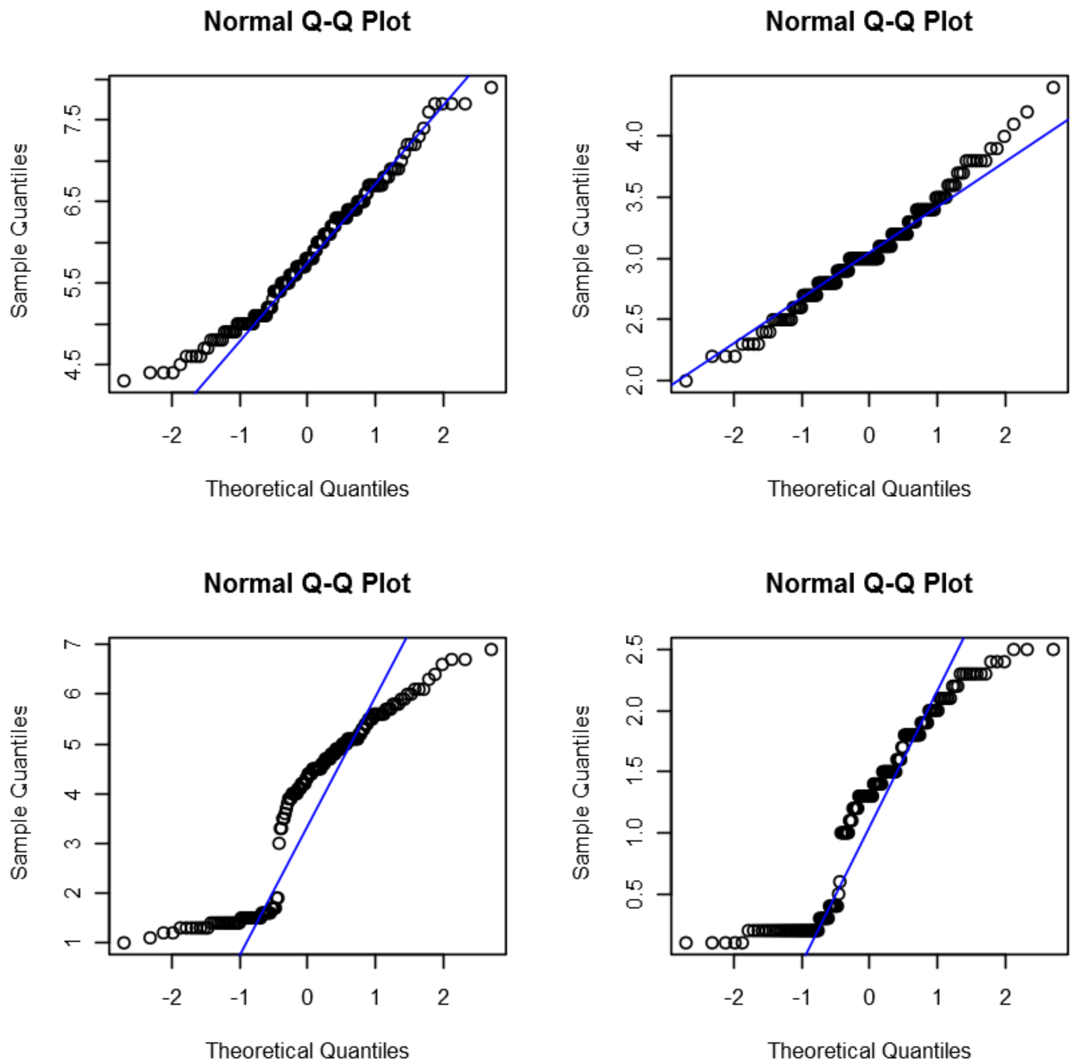

```
In [14]: # boxplot
         par(mfrow=c(2,2))
         boxplot(iris$Sepal.Length~iris$Species,ylab="Sepal Length (cm)")
         boxplot(iris$Sepal.Width~iris$Species,ylab="Sepal Width (cm)")
         boxplot(iris$Petal.Length~iris$Species,ylab="Petal Length (cm)")
         boxplot(iris$Petal.Width~iris$Species,ylab="Petal Width (cm)")
```



In [15]:
```
# scatter plot
par(mfrow=c(1,1))
scatterplotMatrix(~Sepal.Length + Sepal.Width + Petal.Length + Petal.Width, data=iris
```



In [34]:
```
# QQ Plot
par(mfrow=c(2,2))
qqnorm(iris$Sepal.Length);qqline(iris$Sepal.Length,col='blue')
qqnorm(iris$Sepal.Width);qqline(iris$Sepal.Width,col='blue')
qqnorm(iris$Petal.Length);qqline(iris$Petal.Length,col='blue')
qqnorm(iris$Petal.Width);qqline(iris$Petal.Width,col='blue')
```



In [18]:
```
# correlation
cor(iris[,c(1:4)])

# Petal Length is highly correlated with Sepal.Length
# Petal Width is highly correlated with Sepal.Length
# Petal.Width is highly correlated with Petal.Length
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	1.0000000	-0.1175698	0.8717538	0.8179411
Sepal.Width	-0.1175698	1.0000000	-0.4284401	-0.3661259
Petal.Length	0.8717538	-0.4284401	1.0000000	0.9628654
Petal.Width	0.8179411	-0.3661259	0.9628654	1.0000000

2.2 Normality Test

```
In [5]: # check normality
        library(nortest)
        n <- dim(iris)[2]
        name <- names(iris)
        for (i in 1:(n-1)){
            print(name[i])
            print(ks.test(iris[,i],"pnorm"))
            print(shapiro.test(iris[,i]))
            print(pearson.test(iris[,i]))
            print(ad.test(iris[,i]))
            print(cvm.test(iris[,i]))
        }
        # Sepal.Length: Not Normal distributed, only Pearson chi-square test p-value=0.1352>0.
        # Sepal.Width: Not Normal distributed, only Shapiro-Wilk test p-value=0.1012>0.05
        # Petal.Length: Not Normal distributed, all p-value<0.05
        # Petal.Width: Not Normal distributed, all p-value<0.05
```

```
[1] "Sepal.Length"


Warning message in ks.test(iris[, i], "pnorm"):
"ties should not be present for the Kolmogorov-Smirnov test"


        One-sample Kolmogorov-Smirnov test

data:  iris[, i]
D = 0.99999, p-value < 2.2e-16
alternative hypothesis: two-sided


        Shapiro-Wilk normality test

data:  iris[, i]
W = 0.97609, p-value = 0.01018


        Pearson chi-square normality test

data:  iris[, i]
P = 17.4, p-value = 0.1352


        Anderson-Darling normality test

data:  iris[, i]
A = 0.8892, p-value = 0.02251
```

```
        Cramer-von Mises normality test

data:  iris[, i]
W = 0.1274, p-value = 0.04706

[1] "Sepal.Width"


Warning message in ks.test(iris[, i], "pnorm"):
"ties should not be present for the Kolmogorov-Smirnov test"


        One-sample Kolmogorov-Smirnov test

data:  iris[, i]
D = 0.97943, p-value < 2.2e-16
alternative hypothesis: two-sided


        Shapiro-Wilk normality test

data:  iris[, i]
W = 0.98492, p-value = 0.1012


        Pearson chi-square normality test

data:  iris[, i]
P = 46.2, p-value = 6.409e-06


        Anderson-Darling normality test

data:  iris[, i]
A = 0.90796, p-value = 0.02023


        Cramer-von Mises normality test

data:  iris[, i]
W = 0.18065, p-value = 0.009336

[1] "Petal.Length"


Warning message in ks.test(iris[, i], "pnorm"):
"ties should not be present for the Kolmogorov-Smirnov test"
```

```
        One-sample Kolmogorov-Smirnov test

data:  iris[, i]
D = 0.87653, p-value < 2.2e-16
alternative hypothesis: two-sided


        Shapiro-Wilk normality test

data:  iris[, i]
W = 0.87627, p-value = 7.412e-10


        Pearson chi-square normality test

data:  iris[, i]
P = 192.8, p-value < 2.2e-16


        Anderson-Darling normality test

data:  iris[, i]
A = 7.6785, p-value < 2.2e-16



Warning message in cvm.test(iris[, i]):
"p-value is smaller than 7.37e-10, cannot be computed more accurately"


        Cramer-von Mises normality test

data:  iris[, i]
W = 1.2223, p-value = 7.37e-10

[1] "Petal.Width"


Warning message in ks.test(iris[, i], "pnorm"):
"ties should not be present for the Kolmogorov-Smirnov test"


        One-sample Kolmogorov-Smirnov test

data:  iris[, i]
D = 0.54593, p-value < 2.2e-16
alternative hypothesis: two-sided
```

```
        Shapiro-Wilk normality test

data:  iris[, i]
W = 0.90183, p-value = 1.68e-08


        Pearson chi-square normality test

data:  iris[, i]
P = 155.6, p-value < 2.2e-16


        Anderson-Darling normality test

data:  iris[, i]
A = 5.1057, p-value = 1.125e-12


        Cramer-von Mises normality test

data:  iris[, i]
W = 0.72156, p-value = 4.338e-08
```

2.3 Equality of Variance test

```
In [57]: # check equality of variance
         # Not normal distributed
         lapply(iris[,c(1:4)],function(x) fligner.test(x~iris$Species))
         #Sepal.Length, Petal.Length, p-value<0.05, Petal.Width:  reject H0, variance are not
         #Sepal.Width: p-value>0.05, Not reject H0, variance are equal.

         # normal distributed
         # bartlett.test get the same results as fligner.test
         lapply(iris[,c(1:4)],function(x) bartlett.test(x~iris$Species))
```

```
$Sepal.Length

Fligner-Killeen test of homogeneity of variances

data:  x by iris$Species
Fligner-Killeen:med chi-squared = 11.618, df = 2, p-value = 0.003


$Sepal.Width
```

```
Fligner-Killeen test of homogeneity of variances

data:  x by iris$Species
Fligner-Killeen:med chi-squared = 0.9122, df = 2, p-value = 0.6338


$Petal.Length

Fligner-Killeen test of homogeneity of variances

data:  x by iris$Species
Fligner-Killeen:med chi-squared = 34.708, df = 2, p-value = 2.906e-08


$Petal.Width

Fligner-Killeen test of homogeneity of variances

data:  x by iris$Species
Fligner-Killeen:med chi-squared = 29.387, df = 2, p-value = 4.157e-07




$Sepal.Length

Bartlett test of homogeneity of variances

data:  x by iris$Species
Bartlett's K-squared = 16.006, df = 2, p-value = 0.0003345


$Sepal.Width

Bartlett test of homogeneity of variances

data:  x by iris$Species
Bartlett's K-squared = 2.0911, df = 2, p-value = 0.3515


$Petal.Length

Bartlett test of homogeneity of variances

data:  x by iris$Species
Bartlett's K-squared = 55.423, df = 2, p-value = 9.229e-13
```

```
$Petal.Width

Bartlett test of homogeneity of variances

data:  x by iris$Species
Bartlett's K-squared = 39.213, df = 2, p-value = 3.055e-09
```

From the Normality test and equality of variance test. The iris dataset are Not Normal distributed, and the variance is NOT equal. So I will use Kruskal-Wallis test to check whether there is difference in three different species. If there is difference, then use MANN-Whitney-Wilcoxon test to do pairwised test.

2.4 Kruskal-Wallis test

```
In [50]: # Kruskal-Wallis test
         kruskal.test(iris$Sepal.Length+iris$Sepal.Width+iris$Petal.Length+iris$Petal.Width~iri
         # There is difference in these three Species: setosa,versicolor,virginica.

         lapply(iris[,c(1:4)],function(x) kruskal.test(x~iris$Species)) #p-value<0.05, reject 
         # There is difference of Sepal Length in three Species: setosa,versicolor,virginica.
         # There is difference of Sepal Width in three Species: setosa,versicolor,virginica.
         # There is difference of Petal Length in three Species: setosa,versicolor,virginica.
         # There is difference of Petal Width in three Species: setosa,versicolor,virginica.
```

```
Kruskal-Wallis rank sum test

data:  iris$Sepal.Length + iris$Sepal.Width + iris$Petal.Length + iris$Petal.Width by iris$Spec
Kruskal-Wallis chi-squared = 124.17, df = 2, p-value < 2.2e-16
```

```
$Sepal.Length

Kruskal-Wallis rank sum test

data:  x by iris$Species
Kruskal-Wallis chi-squared = 96.937, df = 2, p-value < 2.2e-16


$Sepal.Width

Kruskal-Wallis rank sum test

data:  x by iris$Species
Kruskal-Wallis chi-squared = 63.571, df = 2, p-value = 1.569e-14
```

```
$Petal.Length

Kruskal-Wallis rank sum test

data:  x by iris$Species
Kruskal-Wallis chi-squared = 130.41, df = 2, p-value < 2.2e-16


$Petal.Width

Kruskal-Wallis rank sum test

data:  x by iris$Species
Kruskal-Wallis chi-squared = 131.19, df = 2, p-value < 2.2e-16
```

```
In [51]: # Even though the iris dataset is NOT Normal distributed, and the variance are not eq
         #I still want to check the result of ANOVA test. It gets the same results as  Kruskal

         # ANOVA test
         summary(aov(iris$Sepal.Length+iris$Sepal.Width+iris$Petal.Length+iris$Petal.Width~iris
         # There is difference in these three Species: setosa,versicolor,virginica.

         lapply(iris[,c(1:4)],function(x) summary(aov(x~iris$Species))) #p-value<0.05, reject
```

```
              Df Sum Sq Mean Sq F value Pr(>F)
iris$Species   2 1238.4   619.2   422.4 <2e-16 ***
Residuals    147  215.5     1.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


$Sepal.Length
              Df Sum Sq Mean Sq F value Pr(>F)
iris$Species   2  63.21  31.606   119.3 <2e-16 ***
Residuals    147  38.96   0.265
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Sepal.Width
              Df Sum Sq Mean Sq F value Pr(>F)
iris$Species   2  11.35   5.672   49.16 <2e-16 ***
Residuals    147  16.96   0.115
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Petal.Length
              Df Sum Sq Mean Sq F value Pr(>F)
iris$Species   2  437.1  218.55    1180 <2e-16 ***
Residuals    147   27.2    0.19
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Petal.Width
              Df Sum Sq Mean Sq F value Pr(>F)
iris$Species   2  80.41   40.21     960 <2e-16 ***
Residuals    147   6.16    0.04
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.5 pairwised test: MANN-Whitney-Wilcoxon test

In [72]:
```
# pairwised test: MANN-Whitney-Wilcoxon test
pairwise.wilcox.test(iris$Sepal.Length+iris$Sepal.Width+iris$Petal.Length+iris$Petal.W
# p-value < 0.05, reject H0.
# There are differences between setosa and versicolor, setosa and virginica, versicol

lapply(iris[,c(1:4)], function(x) pairwise.wilcox.test(x, iris$Species, paired=FALSE)]
# all of p-values < 0.05, reject H0.
```

```
Pairwise comparisons using Wilcoxon rank sum test

data:  iris$Sepal.Length + iris$Sepal.Width + iris$Petal.Length + iris$Petal.Width and iris$Spe

           setosa  versicolor
versicolor < 2e-16 -
virginica  < 2e-16 6.8e-14

P value adjustment method: holm


$Sepal.Length

Pairwise comparisons using Wilcoxon rank sum test

data:  x and iris$Species

           setosa  versicolor
versicolor 1.7e-13 -
```

```
virginica  < 2e-16 5.9e-07

P value adjustment method: holm

$Sepal.Width

Pairwise comparisons using Wilcoxon rank sum test

data:  x and iris$Species

           setosa  versicolor
versicolor 6.4e-13 -
virginica  1.4e-08 0.0046

P value adjustment method: holm

$Petal.Length

Pairwise comparisons using Wilcoxon rank sum test

data:  x and iris$Species

           setosa versicolor
versicolor <2e-16 -
virginica  <2e-16 <2e-16

P value adjustment method: holm

$Petal.Width

Pairwise comparisons using Wilcoxon rank sum test

data:  x and iris$Species

           setosa versicolor
versicolor <2e-16 -
virginica  <2e-16 <2e-16

P value adjustment method: holm
```

In [74]:
```
# Even though the iris dataset is NOT Normal distributed, and the variance are not eq
# I still want to use t-test to check the difference. It gets the same results as MAN
pairwise.t.test(iris$Sepal.Length+iris$Sepal.Width+iris$Petal.Length+iris$Petal.Width

lapply(iris[,c(1:4)], function(x) pairwise.t.test(x, iris$Species, paired=FALSE))
```

```
Pairwise comparisons using t tests with pooled SD

data:  iris$Sepal.Length + iris$Sepal.Width + iris$Petal.Length + iris$Petal.Width and iris$Spe

           setosa versicolor
versicolor <2e-16 -
virginica  <2e-16 <2e-16

P value adjustment method: holm


$Sepal.Length

Pairwise comparisons using t tests with pooled SD

data:  x and iris$Species

           setosa  versicolor
versicolor 1.8e-15 -
virginica  < 2e-16 2.8e-09

P value adjustment method: holm

$Sepal.Width

Pairwise comparisons using t tests with pooled SD

data:  x and iris$Species

           setosa  versicolor
versicolor < 2e-16 -
virginica  9.1e-10 0.0031

P value adjustment method: holm

$Petal.Length

Pairwise comparisons using t tests with pooled SD

data:  x and iris$Species

           setosa versicolor
versicolor <2e-16 -
virginica  <2e-16 <2e-16

P value adjustment method: holm

$Petal.Width
```

```
Pairwise comparisons using t tests with pooled SD

data:  x and iris$Species

           setosa versicolor
versicolor <2e-16 -
virginica  <2e-16 <2e-16

P value adjustment method: holm
```

2.6 Conclusion

From the above tests, we can tell that all the three species setosa, versicolor, and virginica are different with each other in all of their sepal length, sepal width, petal length and petal width.

3 Problem 2.

(25 points) We are interested in finding the important predictors of accumulated wealth at the time of retirement, assess their adjusted effect sizes (in direction and magnitude) and use the best linear regression model for interpretation and prediction. We want to analyze the Pension.txt dataset (available on Blackboard) that contains 194 observations on 17 variables: pyears - years of employment, prftshr - indicator for profit sharing company, choice - indicator for company giving a choice to contribute, female, married, age, educ - years of education, finc25, finc35, finc50, finc75, finc100, finc101- indicators for 25, 35, 50, 75, 100 and 101 levels of retirement contribution, wealth89 - wealth in thousands of dollars, race, stckin89 - percent of the portfolio in stock, irain89 - percent of the portfolio in IRA. Perform all necessary data analysis steps and write a section summarizing the findings.

3.1 Load data and basic analysis

```
In [148]: # load dataset
          pension_data <- read.table('Pension.txt',header=T)
          dim(pension_data)
          pension_data <- na.omit(pension_data)
          dim(pension_data)
          head(pension_data)
```

1. 194 2. 17
1. 191 2. 17

pyears	prftshr	choice	female	married	age	educ	finc25	finc35	finc50	finc75	finc100	finc
1	0	1	0	1	64	12	0	0	1	0	0	0
6	1	1	1	1	56	13	0	0	0	1	0	0
25	1	1	0	1	56	12	0	0	0	1	0	0
20	1	0	1	1	63	12	1	0	0	0	0	0
35	0	1	0	1	67	12	0	1	0	0	0	0
13	1	0	0	1	64	11	0	0	0	1	0	0

In [446]:
```
# 17 variables:
# prftshr - indicator for profit sharing company, 2 levels: 0 1
# choice - indicator for company giving a choice to contribute, 2 levels: 0 1
# female, 2 levels: 0 1
# married, 2 levels: 0 1
# finc25, finc35, finc50, finc75, finc100, finc101- indicators for 25, 35, 50, 75, 1
# race, 2 levels 0 1
# stckin89 - percent of the portfolio in stock, 2 levels 0 1
# irain89 - percent of the portfolio in IRA, 2 levels 0 1

# pyears - years of employment
# age
# educ - years of education
# wealth89 - wealth in thousands of dollars
str(pension_data)
summary(pension_data)
lapply(pension_data[, c(2:5,8:13,15:17)], function(x) unique(x))
```

```
'data.frame':	191 obs. of  17 variables:
 $ pyears  : int  1 6 25 20 35 13 2 10 26 5 ...
 $ prftshr : Factor w/ 2 levels "0","1": 1 2 2 2 1 2 1 2 1 1 ...
 $ choice  : Factor w/ 2 levels "0","1": 2 2 2 1 2 1 2 1 1 2 ...
 $ female  : Factor w/ 2 levels "0","1": 1 2 1 2 1 1 2 2 2 1 ...
 $ married : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ age     : int  64 56 56 63 67 64 64 64 69 60 ...
 $ educ    : int  12 13 12 12 12 11 12 12 12 14 ...
 $ finc25  : Factor w/ 2 levels "0","1": 1 1 1 2 1 1 2 2 1 1 ...
 $ finc35  : Factor w/ 2 levels "0","1": 1 1 1 1 2 1 1 1 2 1 ...
 $ finc50  : Factor w/ 2 levels "0","1": 2 1 1 1 1 1 1 1 1 1 ...
 $ finc75  : Factor w/ 2 levels "0","1": 1 2 2 1 1 2 1 1 1 1 ...
 $ finc100 : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
 $ finc101 : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 2 ...
 $ wealth89: num  77.9 154.9 154.9 232.5 179 ...
 $ race    : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
 $ stckin89: Factor w/ 2 levels "0","1": 2 2 2 2 1 2 1 1 1 2 ...
 $ irain89 : Factor w/ 2 levels "0","1": 2 2 2 2 2 1 2 2 1 2 ...
 - attr(*, "na.action")=Class 'omit'  Named int [1:3] 39 52 173
  .. ..- attr(*, "names")= chr [1:3] "39" "52" "173"
```

```
    pyears      prftshr choice  female  married      age               educ
```

```
 Min.   : 0.0   0:151   0: 74   0: 75   0: 47   Min.   :54.00   Min.   : 8.00
 1st Qu.: 4.0   1: 40   1:117   1:116   1:144   1st Qu.:57.00   1st Qu.:12.00
 Median : 9.0                                   Median :60.00   Median :12.00
 Mean   :11.3                                   Mean   :60.52   Mean   :13.53
 3rd Qu.:16.0                                   3rd Qu.:64.00   3rd Qu.:16.00
 Max.   :45.0                                   Max.   :73.00   Max.   :18.00
 finc25  finc35  finc50  finc75  finc100 finc101    wealth89      race
 0:151   0:157   0:146   0:165   0:165   0:181   Min.   :  -6.3   0:169
 1: 40   1: 34   1: 45   1: 26   1: 26   1: 10   1st Qu.:  65.8   1: 22
                                                 Median : 140.0
                                                 Mean   : 212.0
                                                 3rd Qu.: 253.4
                                                 Max.   :1485.0
 stckin89 irain89
 0:126    0:93
 1: 65    1:98
```

$prftshr 1. 0 2. 1

$choice 1. 1 2. 0

$female 1. 0 2. 1

$married 1. 1 2. 0

$finc25 1. 0 2. 1

$finc35 1. 0 2. 1

$finc50 1. 1 2. 0

$finc75 1. 0 2. 1

$finc100 1. 0 2. 1

$finc101 1. 0 2. 1

$race 1. 0 2. 1

$stckin89 1. 1 2. 0

$irain89 1. 1 2. 0

In [150]:
```
# change some variables to factors
pension_data[, c(2:5,8:13,15:17)] <- lapply(pension_data[, c(2:5,8:13,15:17)], funct
str(pension_data)
summary(pension_data)
```

```
'data.frame':	191 obs. of  17 variables:
 $ pyears  : int  1 6 25 20 35 13 2 10 26 5 ...
 $ prftshr : Factor w/ 2 levels "0","1": 1 2 2 2 1 2 1 2 1 1 ...
 $ choice  : Factor w/ 2 levels "0","1": 2 2 2 1 2 1 2 1 1 2 ...
 $ female  : Factor w/ 2 levels "0","1": 1 2 1 2 1 1 2 2 2 1 ...
 $ married : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ age     : int  64 56 56 63 67 64 64 64 69 60 ...
 $ educ    : int  12 13 12 12 12 11 12 12 12 14 ...
 $ finc25  : Factor w/ 2 levels "0","1": 1 1 1 2 1 1 2 2 1 1 ...
 $ finc35  : Factor w/ 2 levels "0","1": 1 1 1 1 2 1 1 1 2 1 ...
 $ finc50  : Factor w/ 2 levels "0","1": 2 1 1 1 1 1 1 1 1 1 ...
 $ finc75  : Factor w/ 2 levels "0","1": 1 2 2 1 1 2 1 1 1 1 ...
 $ finc100 : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
 $ finc101 : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 2 ...
 $ wealth89: num  77.9 154.9 154.9 232.5 179 ...
 $ race    : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
 $ stckin89: Factor w/ 2 levels "0","1": 2 2 2 2 1 2 1 1 1 2 ...
 $ irain89 : Factor w/ 2 levels "0","1": 2 2 2 2 2 1 2 2 1 2 ...
 - attr(*, "na.action")=Class 'omit'  Named int [1:3] 39 52 173
  .. ..- attr(*, "names")= chr [1:3] "39" "52" "173"
```

```
     pyears      prftshr choice  female  married      age             educ      
 Min.   : 0.0   0:151   0: 74   0: 75   0: 47   Min.   :54.00   Min.   : 8.00  
 1st Qu.: 4.0   1: 40   1:117   1:116   1:144   1st Qu.:57.00   1st Qu.:12.00  
 Median : 9.0                                   Median :60.00   Median :12.00  
 Mean   :11.3                                   Mean   :60.52   Mean   :13.53  
 3rd Qu.:16.0                                   3rd Qu.:64.00   3rd Qu.:16.00  
 Max.   :45.0                                   Max.   :73.00   Max.   :18.00  
 finc25  finc35  finc50  finc75  finc100 finc101    wealth89       race   
 0:151   0:157   0:146   0:165   0:165   0:181    Min.   :  -6.3   0:169  
 1: 40   1: 34   1: 45   1: 26   1: 26   1: 10    1st Qu.:  65.8   1: 22  
                                                  Median : 140.0          
                                                  Mean   : 212.0          
                                                  3rd Qu.: 253.4          
                                                  Max.   :1485.0          
 stckin89 irain89
 0:126    0:93   
 1: 65    1:98   
```

```
In [151]: # normalize variables: pyears, age, educ, wealth89
          #pension_data[,c(1,6,7,14)] <- lapply(pension_data[,c(1,6,7,14)],function(x) ((x-min
          #summary(pension_data)
```

3.2 Split the dataset into training and testing dataset

```
In [166]: # train dataset: 80%, test dataset: 20%
          (n <- dim(pension_data)[1])
          trainID <- sample(1:n,n*0.8)
          train <- pension_data[trainID,]
          test <- pension_data[-trainID,]
```

191

3.3 Build Models

```
In [244]: library(MASS)
          lm.fit <- glm(wealth89~., data=train, family=gaussian)
          summary(lm.fit) # AIC: 2086.5
          # Best model 1: Smallest AIC value
          lm.best1 <- stepAIC(lm.fit, direction = "both", trace=F)
          summary(lm.best1)  # AIC: 2072.9
```

```
Call:
glm(formula = wealth89 ~ ., family = gaussian, data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-323.12  -120.18   -37.47    67.09  1153.27

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -993.523    366.323  -2.712  0.00756 **
pyears        -1.068      2.207  -0.484  0.62918
prftshr1     -39.447     45.712  -0.863  0.38970
choice1      -35.341     39.596  -0.893  0.37369
female1       18.388     47.421   0.388  0.69881
married1      29.051     50.717   0.573  0.56773
age           14.572      5.123   2.845  0.00514 **
educ          14.698      7.746   1.897  0.05990 .
finc251        1.555     84.441   0.018  0.98534
finc351       42.812     87.199   0.491  0.62424
finc501       46.297     85.708   0.540  0.58997
finc751      148.742     97.188   1.530  0.12824
finc1001      95.619     95.286   1.003  0.31742
finc1011     250.303    106.848   2.343  0.02061 *
race1        -63.618     60.042  -1.060  0.29124
stckin891     95.298     42.546   2.240  0.02673 *
irain891      77.280     41.735   1.852  0.06625 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for gaussian family taken to be 47610.02)

    Null deviance: 9318957  on 151  degrees of freedom
Residual deviance: 6427352  on 135  degrees of freedom
AIC: 2086.5

Number of Fisher Scoring iterations: 2




Call:
glm(formula = wealth89 ~ age + educ + finc75 + finc101 + stckin89 +
    irain89, family = gaussian, data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-333.23  -126.33   -41.41    60.88  1185.08

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1007.733    282.800  -3.563 0.000496 ***
age            14.220      4.303   3.304 0.001200 **
educ           18.411      6.935   2.655 0.008826 **
finc751        97.291     55.670   1.748 0.082647 .
finc1011      222.799     75.823   2.938 0.003840 **
stckin891      98.490     40.350   2.441 0.015857 *
irain891       92.367     37.530   2.461 0.015021 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 46239.43)

    Null deviance: 9318957  on 151  degrees of freedom
Residual deviance: 6704717  on 145  degrees of freedom
AIC: 2072.9

Number of Fisher Scoring iterations: 2
```

In [248]:
```
# Best model: Smallest AIC value, add two way of interaction
lm.best2 <- stepAIC(lm.fit, ~.^2, direction = "both", trace=F)
summary(lm.best2)  # AIC: 2003.9
# race is not significant, and it isn't interactive with other variables. remove it.
```

```
Call:
```

```
glm(formula = wealth89 ~ pyears + choice + female + married +
    age + educ + finc50 + finc75 + finc100 + finc101 + race +
    stckin89 + irain89 + married:finc75 + age:finc101 + female:finc101 +
    female:irain89 + pyears:finc50 + age:stckin89 + educ:irain89 +
    educ:stckin89 + female:age + pyears:age + choice:finc50 +
    female:finc100 + age:finc100 + pyears:finc75 + choice:finc100 +
    choice:educ + finc100:irain89, family = gaussian, data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-310.07   -65.77   -18.42    56.45   943.22

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        -813.7271   611.4582  -1.331 0.185758
pyears               51.4753    26.2552   1.961 0.052227 .
choice1            -285.0433   158.4464  -1.799 0.074513 .
female1             937.7065   484.7226   1.935 0.055383 .
married1             54.3300    38.7141   1.403 0.163069
age                  13.1427     9.6119   1.367 0.174058
educ                  0.7784    10.8535   0.072 0.942943
finc501              31.2477    71.0050   0.440 0.660666
finc751            1109.2483   181.4240   6.114 1.23e-08 ***
finc1001          -2620.5919   823.3630  -3.183 0.001854 **
finc1011          -6149.0870   953.7124  -6.448 2.44e-09 ***
race1               -71.7749    45.2781  -1.585 0.115531
stckin891         -1489.1391   552.2316  -2.697 0.008004 **
irain891            603.1034   165.7642   3.638 0.000405 ***
married1:finc751  -1020.0467   176.5700  -5.777 6.03e-08 ***
age:finc1011         97.4925    15.0827   6.464 2.25e-09 ***
female1:finc1011    701.3971   150.2392   4.669 7.91e-06 ***
female1:irain891   -203.7372    59.4831  -3.425 0.000840 ***
pyears:finc501        6.3611     3.4606   1.838 0.068497 .
age:stckin891        18.0072     8.3033   2.169 0.032064 *
educ:irain891       -29.4550    11.5242  -2.556 0.011829 *
educ:stckin891       38.2214    12.4722   3.065 0.002689 **
female1:age         -14.4274     7.8070  -1.848 0.067042 .
pyears:age           -0.8397     0.4294  -1.955 0.052839 .
choice1:finc501    -158.4909    66.8751  -2.370 0.019372 *
female1:finc1001    193.2202    93.7932   2.060 0.041537 *
age:finc1001         46.8591    13.9634   3.356 0.001058 **
pyears:finc751       -7.3850     4.7362  -1.559 0.121549
choice1:finc1001   -328.4064   101.4060  -3.239 0.001551 **
choice1:educ         24.3864    12.0994   2.016 0.046067 *
finc1001:irain891  -158.3423    98.3715  -1.610 0.110083
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for gaussian family taken to be 25652.67)

    Null deviance: 9318957  on 151  degrees of freedom
Residual deviance: 3103974  on 121  degrees of freedom
AIC: 2003.9

Number of Fisher Scoring iterations: 2
```

```
In [249]: lm.best2 <- glm(formula = wealth89 ~ pyears + choice + female + married + age + educ
              stckin89 + irain89 + married:finc75 + age:finc101 + female:finc101 + female:irai
              educ:stckin89 + female:age + pyears:age + choice:finc50 + female:finc100 + age:f
              choice:educ + finc100:irain89, family = gaussian, data = train)
          summary(lm.best2) # AIC: 2005
```

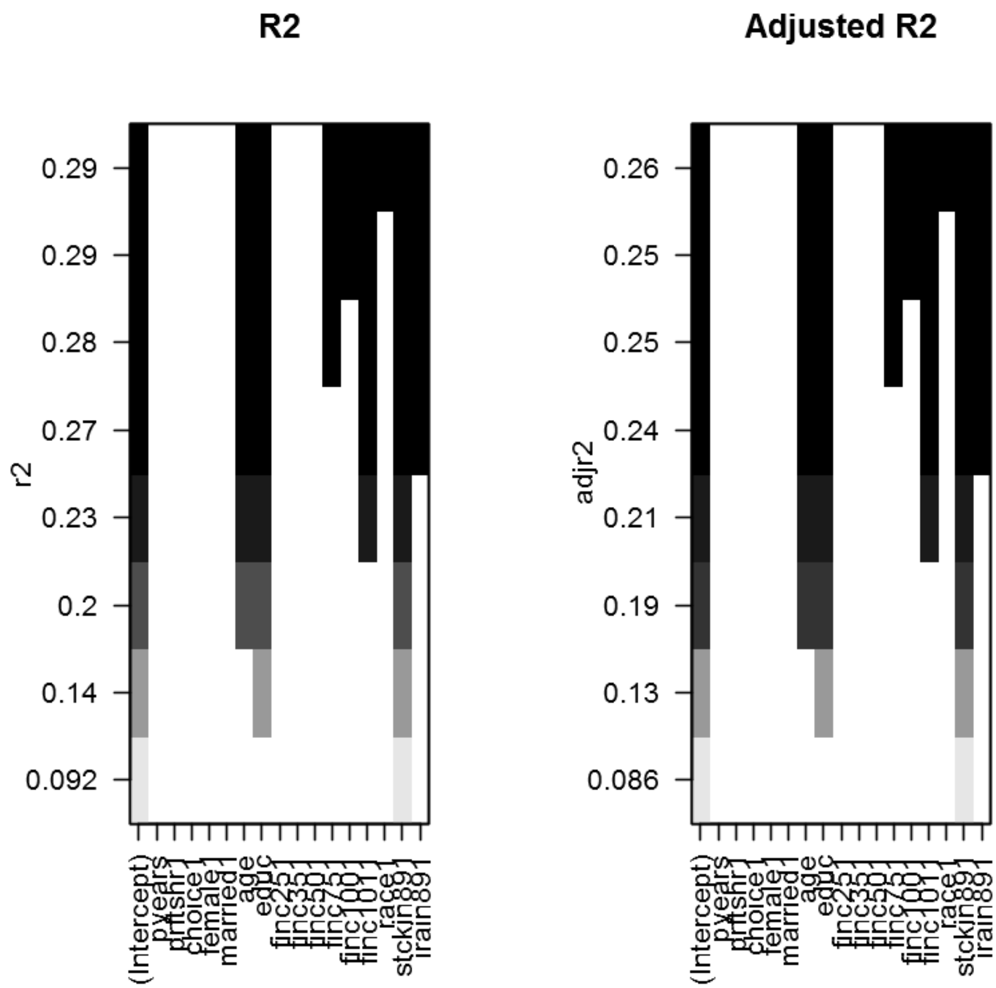
```
Call:
glm(formula = wealth89 ~ pyears + choice + female + married +
    age + educ + finc50 + finc75 + finc100 + finc101 + stckin89 +
    irain89 + married:finc75 + age:finc101 + female:finc101 +
    female:irain89 + pyears:finc50 + age:stckin89 + educ:irain89 +
    educ:stckin89 + female:age + pyears:age + choice:finc50 +
    female:finc100 + age:finc100 + pyears:finc75 + choice:finc100 +
    choice:educ + finc100:irain89, family = gaussian, data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-309.29   -68.18   -15.90    64.20   943.01

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        -843.4697   614.9481  -1.372 0.172702
pyears               49.7271    26.3942   1.884 0.061943 .
choice1            -294.4599   159.3137  -1.848 0.066980 .
female1             838.9570   483.6743   1.735 0.085346 .
married1             60.9021    38.7294   1.573 0.118423
age                  13.1785     9.6713   1.363 0.175507
educ                  1.4376    10.9126   0.132 0.895409
finc501              38.5126    71.2950   0.540 0.590053
finc751            1128.8213   182.1222   6.198 8.04e-09 ***
finc1001          -2600.8646   828.3577  -3.140 0.002121 **
finc1011          -6023.0753   956.2682  -6.299 4.94e-09 ***
stckin891         -1480.3054   555.6167  -2.664 0.008759 **
irain891            646.8110   164.4654   3.933 0.000140 ***
married1:finc751  -1029.4201   177.5618  -5.798 5.40e-08 ***
age:finc1011         95.4846    15.1223   6.314 4.58e-09 ***
female1:finc1011    709.4387   151.0817   4.696 7.02e-06 ***
```

```
female1:irain891   -202.9605    59.8488  -3.391 0.000938 ***
pyears:finc501        5.9580     3.4726   1.716 0.088752 .
age:stckin891        17.8365     8.3540   2.135 0.034753 *
educ:irain891       -32.4273    11.4409  -2.834 0.005376 **
educ:stckin891       38.5338    12.5477   3.071 0.002631 **
female1:age         -12.8637     7.7923  -1.651 0.101346
pyears:age           -0.8094     0.4316  -1.875 0.063152 .
choice1:finc501    -159.3892    67.2861  -2.369 0.019415 *
female1:finc1001    199.3274    94.2933   2.114 0.036559 *
age:finc1001         46.5178    14.0480   3.311 0.001222 **
pyears:finc751       -7.5582     4.7642  -1.586 0.115225
choice1:finc1001   -318.8968   101.8541  -3.131 0.002181 **
choice1:educ         25.3205    12.1597   2.082 0.039403 *
finc1001:irain891  -168.3136    98.7770  -1.704 0.090931 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 25970.78)

    Null deviance: 9318957  on 151  degrees of freedom
Residual deviance: 3168435  on 122  degrees of freedom
AIC: 2005

Number of Fisher Scoring iterations: 2
```

In [256]:
```r
# use regsubsets() to get the best model, which has the best R2 or ajusted-R2.
library(leaps)
lm.best3 <- regsubsets(wealth89~., data=train, nbest=1, method='exhaustive')
summary(lm.best3)
par(mfrow=c(1,2))
plot(lm.best3, scale='r2', main='R2')
plot(lm.best3, scale='adjr2', main='Adjusted R2')
```

```
Subset selection object
Call: regsubsets.formula(wealth89 ~ ., data = train, nbest = 1, method = "exhaustive")
16 Variables  (and intercept)
          Forced in Forced out
pyears        FALSE      FALSE
prftshr1      FALSE      FALSE
choice1       FALSE      FALSE
female1       FALSE      FALSE
married1      FALSE      FALSE
age           FALSE      FALSE
educ          FALSE      FALSE
finc251       FALSE      FALSE
finc351       FALSE      FALSE
```

```
finc501       FALSE      FALSE
finc751       FALSE      FALSE
finc1001      FALSE      FALSE
finc1011      FALSE      FALSE
race1         FALSE      FALSE
stckin891     FALSE      FALSE
irain891      FALSE      FALSE
1 subsets of each size up to 8
Selection Algorithm: exhaustive
         pyears prftshr1 choice1 female1 married1 age educ finc251 finc351
1  ( 1 ) " "    " "      " "     " "     " "      " " " "  " "     " "
2  ( 1 ) " "    " "      " "     " "     " "      " " "*"  " "     " "
3  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
4  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
5  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
6  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
7  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
8  ( 1 ) " "    " "      " "     " "     " "      "*" "*"  " "     " "
         finc501 finc751 finc1001 finc1011 race1 stckin891 irain891
1  ( 1 ) " "     " "     " "      " "      " "   "*"       " "
2  ( 1 ) " "     " "     " "      " "      " "   "*"       " "
3  ( 1 ) " "     " "     " "      " "      " "   "*"       " "
4  ( 1 ) " "     " "     " "      "*"      " "   "*"       " "
5  ( 1 ) " "     " "     " "      "*"      " "   "*"       "*"
6  ( 1 ) " "     "*"     " "      "*"      " "   "*"       "*"
7  ( 1 ) " "     "*"     "*"      "*"      " "   "*"       "*"
8  ( 1 ) " "     "*"     "*"      "*"      "*"   "*"       "*"
```




```
In [257]: # from the plot, we can get the best model is
          lm.best3 <- lm(wealth89~age+educ+finc75+finc100+finc101+race+stckin89+irain89, data=
          summary(lm.best3) # Multiple R-squared:0.2945,Adjusted R-squared:0.255
          # The p-value of finc75, finc100 and race is greater than 0.05, not significant. But
          AIC(lm.best3)# AIC: 2073.93
```

```
Call:
lm(formula = wealth89 ~ age + educ + finc75 + finc100 + finc101 +
    race + stckin89 + irain89, data = train)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-330.88 -125.45  -37.86   66.91 1190.52

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -916.920    290.428  -3.157  0.00194 **
age           13.515      4.395   3.075  0.00252 **
educ          14.771      7.439   1.986  0.04898 *
finc751      106.187     57.824   1.836  0.06838 .
finc1001      69.640     55.626   1.252  0.21264
finc1011     231.057     77.810   2.970  0.00350 **
race1        -59.121     56.818  -1.041  0.29985
stckin891     91.264     40.486   2.254  0.02570 *
irain891      91.428     37.496   2.438  0.01598 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 214.4 on 143 degrees of freedom
Multiple R-squared:  0.2945,Adjusted R-squared:  0.255
F-statistic: 7.462 on 8 and 143 DF,  p-value: 2.749e-08
```

2073.93207827551

In [258]:
```
# model remove finc100 and race
lm.best3 <- lm(wealth89~age+educ+finc75+finc100+finc101+stckin89+irain89, data=train)
summary(lm.best3) # Multiple R-squared:0.2892,Adjusted R-squared:0.2546
# The p-value of race is greater than 0.05, not significant, remove it.
AIC(lm.best3) # AIC: 2073.08
```

```
Call:
lm(formula = wealth89 ~ age + educ + finc75 + finc100 + finc101 +
    stckin89 + irain89, data = train)

Residuals:
    Min      1Q  Median      3Q     Max
-336.11 -127.37  -38.84   68.11 1193.75

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -987.442    282.491  -3.495 0.000629 ***
age           14.483      4.297   3.371 0.000963 ***
educ          14.787      7.441   1.987 0.048793 *
finc751      115.297     57.174   2.017 0.045594 *
finc1001      73.405     55.524   1.322 0.188249
finc1011     242.389     77.066   3.145 0.002017 **
stckin891     95.134     40.327   2.359 0.019663 *
```

```
irain891      93.650     37.446   2.501 0.013506 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 214.5 on 144 degrees of freedom
Multiple R-squared:  0.2892,Adjusted R-squared:  0.2546
F-statistic: 8.368 on 7 and 144 DF,  p-value: 1.463e-08
```

2073.07860745362

3.4 Predict And Evaluate

In [259]:
```
# predict wealth for train dataset with lm.best1 from stepAIC
lm.pred1.train <- predict(lm.best1,newdata=train[,-14],type="response")
# Calculate R2 for test dataset with lm.best1
(lm.best1.R2.train <- 1-sum((train$wealth89-lm.pred1.train)^2)/sum((train$wealth89-m

# predict wealth for train dataset with lm.best2 from regsubsets
lm.pred2.train <- predict(lm.best2,newdata=train[,-14],type="response")
# Calculate R2 for test dataset with lm.best2
(lm.best2.R2.train <- 1-sum((train$wealth89-lm.pred2.train)^2)/sum((train$wealth89-m

# predict wealth for train dataset with lm.best3 from regsubsets
lm.pred3.train <- predict(lm.best3,newdata=train[,-14],type="response")
# Calculate R2 for test dataset with lm.best2
(lm.best3.R2.train <- 1-sum((train$wealth89-lm.pred3.train)^2)/sum((train$wealth89-m

# predict wealth for test dataset with lm.best3 from stepAIC
lm.pred1.test <- predict(lm.best1,newdata=test[,-14],type="response")
# Calculate R2 for test dataset with lm.best3
(lm.best1.R2.test <- abs(1-sum((test$wealth89-lm.pred1.test)^2)/sum((test$wealth89-m

# predict wealth for test dataset with lm.best2 from regsubsets
lm.pred2.test <- predict(lm.best2,newdata=test[,-14],type="response")
# Calculate R2 for test dataset with lm.best1
(lm.best2.R2 <- abs(1-sum((test$wealth89-lm.pred2.test)^2)/sum((test$wealth89-mean(t

# predict wealth for test dataset with lm.best3 from regsubsets
lm.pred3.test <- predict(lm.best3,newdata=test[,-14],type="response")
# Calculate R2 for test dataset with lm.best1
(lm.best3.R2 <- abs(1-sum((test$wealth89-lm.pred3.test)^2)/sum((test$wealth89-mean(t
```

0.28052923018587
0.660001103653225
0.289157110912427
0.147110695248097

0.460305366490338
0.211240853365952

3.5 Interpretation

From the results, we can see that: 1. The AIC value of lm.best2(2005) is the best, which the AIC value of lm.best1 is 2072.9 and lm.best3 is 2073.08 in train dataset; 2. The prediction of both train and test dataset, lm.best2 has the best R-square value(train: 0.66, test: 0.46), which for lm.best1 is (train: 0.28, test: 0.147), lm.best3 is (train: 0.289, test: 0.211).

In conclusion, lm.best2 is the best model. which is: wealth89 = -843.4697 + 49.7271*pyears - 294.4599*choice + 838.9570*female + 60.9021*married + 13.1785*age + 1.4376*educ + 38.5126*finc50 + 1128.8213*finc75 - 2600.864*finc100 - 6023.0753*finc101 - 1480.3054*stckin89 + 646.811*irain89 - 1029.4201*married:finc75 + 95.4846*age:finc101 + 709.4387*female:finc101 - 202.9605*female:irain89 + 5.9580*pyears:finc50 + 17.8365*age:stckin89 - 32.4273*educ:irain89 + 38.5338*educ:stckin89 - 12.8637*female:age - 0.8094*pyears:age - 159.3892*choice:finc50 + 199.3274*female:finc100 + 46.5178*age:finc100 - 7.5582*pyears:finc75 - 318.8968*choice:finc100 + 25.3205*choice:educ - 168.3136*finc100:irain89

Even though some of the single variables are not significant, it interact with other variables, and the interaction is significant. So I keep them.

From the results, we can tell that all of these variables have complicated effects to the wealth. For example, when the pyears increasing, it will affect the interaction with finc101, age, finc75 and affect wealth directly. Some of them are positive effect, some of them are negative effect.

Besides, the best model, lm.best2, has the highest accuracy, but it's not enough. they accuracy is still low. It may need get other variables, or not fit linear model, but logistic model or KNN, and so on.

4 Problem 3.

(25 points) We are interested in finding the important predictors of online customers booking a room at a hotel, assess their adjusted effect sizes (in direction and magnitude) and use the best logistic regression model for interpretation and prediction. We want to analyze the Travel.txt dataset (available on Blackboard) that contains 20,000 observations on 26 variables (description of all variables is presented in the Data_Dictionary_Travel file available on Blackboard). Perform all necessary data analysis steps and write a section summarizing the findings.

4.1 Load data and Basic Analysis

In [2]:
```
# load data
travel_data <- read.table('Travel.txt',header=T)
dim(travel_data)
# omit rows that has NA
travel_data <- na.omit(travel_data)
dim(travel_data)
head(travel_data)
names(travel_data)
str(travel_data)
summary(travel_data)
```

1. 20000 2. 25

1. 19993 2. 25

date_time	user_location_region	user_location_city	user_location_latitude	user_location_lo
2015-07-13 13:22:00	CA	STOCKTON	37.983835	-121.30217
2015-10-21 16:19:00	WA	KIRKLAND	47.70401	-122.190513
2015-08-04 08:15:00	BZ	BELIZE	NULL	NULL
2015-06-02 20:23:00	MAN	MANDALUYONG	NULL	NULL
2015-09-18 21:37:00	CO	DENVER	39.7478	-104.9406725
2015-12-02 01:07:00	CA	SAN DIEGO	32.760625	-117.168396

1. 'date_time' 2. 'user_location_region' 3. 'user_location_city' 4. 'user_location_latitude' 5. 'user_location_longitude' 6. 'orig_destination_distance' 7. 'user_id' 8. 'is_mobile' 9. 'is_package' 10. 'channel' 11. 'srch_ci' 12. 'srch_co' 13. 'srch_adults_cnt' 14. 'srch_children_cnt' 15. 'srch_rm_cnt' 16. 'srch_destination_id' 17. 'hotel_country' 18. 'is_booking' 19. 'hotel_id' 20. 'prop_is_branded' 21. 'prop_starrating' 22. 'distance_band' 23. 'hist_price_band' 24. 'popularity_band' 25. 'cnt'

```
'data.frame':        19993 obs. of  25 variables:
 $ date_time                : Factor w/ 19525 levels "2015-01-01 08:22:00",..: 10614 16391 1204
 $ user_location_region     : Factor w/ 367 levels "***","?","00",..: 107 348 105 206 112 107 
 $ user_location_city       : Factor w/ 4315 levels "?","AACHEN","ABANCAY",..: 3720 1938 289 22
 $ user_location_latitude   : Factor w/ 3088 levels "-14.809929","-15.791410500000001",..: 1172
 $ user_location_longitude  : Factor w/ 3090 levels "-100.453205",..: 521 598 3090 3090 48 278
 $ orig_destination_distance: Factor w/ 13865 levels "0.0056","0.0286",..: 1460 11683 13865 138
 $ user_id                  : int  1332946867 1826431285 -1744380080 -805781454 64349796 32077
 $ is_mobile                : int  0 0 0 0 0 0 0 0 1 0 ...
 $ is_package               : int  1 0 1 0 0 0 0 1 1 0 ...
 $ channel                  : int  262 541 541 541 541 541 541 541 262 541 ...
 $ srch_ci                  : Factor w/ 625 levels "2015-01-01","2015-01-02",..: 445 295 223 1
 $ srch_co                  : Factor w/ 634 levels "2015-01-03","2015-01-04",..: 448 299 223 1
 $ srch_adults_cnt          : int  2 2 2 2 2 2 2 2 2 6 ...
 $ srch_children_cnt        : int  2 2 0 1 2 0 1 0 1 0 ...
 $ srch_rm_cnt              : int  1 1 1 1 1 1 1 1 1 2 ...
 $ srch_destination_id      : int  5581115 5527578 5525315 5581053 5527578 5526679 5525315 557
 $ hotel_country            : Factor w/ 140 levels "ANGOLA","ANGUILLA",..: 80 135 135 124 135 
 $ is_booking               : int  0 0 0 0 0 0 0 0 0 0 ...
 $ hotel_id                 : int  24630615 182496 758476 145417930 254413465 391591 373983 17
 $ prop_is_branded          : int  0 1 1 0 0 1 1 1 0 1 ...
 $ prop_starrating          : int  4 4 4 4 2 4 3 5 3 5 ...
 $ distance_band            : Factor w/ 5 levels "C","F","M","VC",..: 1 1 1 4 3 1 3 3 4 3 ...
 $ hist_price_band          : Factor w/ 5 levels "H","L","M","VH",..: 3 4 3 3 2 3 3 3 2 1 ...
 $ popularity_band          : Factor w/ 5 levels "H","L","M","VH",..: 4 4 4 4 3 4 4 4 1 4 ...
 $ cnt                      : int  1 1 3 1 1 1 3 3 1 1 ...
 - attr(*, "na.action")=Class 'omit'  Named int [1:7] 116 3885 8877 11095 13029 17566 18759
  .. ..- attr(*, "names")= chr [1:7] "116" "3885" "8877" "11095" ...
```

```
               date_time     user_location_region   user_location_city
 2015-03-13 23:15:00:    3   CA     : 2924        NEW YORK   :  366
 2015-05-27 19:37:00:    3   NY     : 1236        LOS ANGELES:  266
```

```
2015-06-09 13:11:00:    3   TX     : 1163        TORONTO    :  232
2015-06-10 17:31:00:    3   FL     : 1063        HOUSTON    :  229
2015-07-07 20:11:00:    3   ON     :  938        CHICAGO    :  215
2015-07-25 11:51:00:    3   WA     :  591        CALGARY    :  179
(Other)            :19975   (Other):12078        (Other)    :18506
user_location_latitude       user_location_longitude orig_destination_distance
NULL      : 4242       NULL               : 4242      NULL     : 4242
40.75512 :  366        -73.98300900000001:  366      227.5021:   34
34.059768:  266        -118.312427       :  266      0.6328  :   17
43.667179:  232        -79.390203        :  232      0.1175  :   16
29.769607:  229        -95.42647         :  229      342.2687:   16
41.89042 :  215        -87.62904         :  215      196.1892:   15
(Other)  :14443        (Other)           :14443      (Other) :15653
   user_id             is_mobile        is_package        channel
Min.   :-2.147e+09   Min.   :0.0000   Min.   :0.0000   Min.   :231.0
1st Qu.:-1.031e+09   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:293.0
Median : 2.384e+07   Median :0.0000   Median :0.0000   Median :510.0
Mean   : 1.763e+07   Mean   :0.2233   Mean   :0.1949   Mean   :418.6
3rd Qu.: 1.086e+09   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:541.0
Max.   : 2.147e+09   Max.   :1.0000   Max.   :1.0000   Max.   :541.0

      srch_ci           srch_co       srch_adults_cnt srch_children_cnt
2015-09-04:  134   2015-09-07:  152   Min.   :0.000   Min.   :0.0000
2015-07-03:  124   2015-07-05:  140   1st Qu.:2.000   1st Qu.:0.0000
2015-09-05:  121   2015-07-26:  124   Median :2.000   Median :0.0000
2015-08-14:  117   2015-08-09:  123   Mean   :2.056   Mean   :0.3105
2015-07-31:  114   2015-08-30:  122   3rd Qu.:2.000   3rd Qu.:0.0000
2015-08-07:  114   2016-01-02:  120   Max.   :9.000   Max.   :8.0000
(Other)   :19269   (Other)   :19212
 srch_rm_cnt    srch_destination_id                 hotel_country
Min.   :0.000   Min.   :      8152   UNITED STATES OF AMERICA:12009
1st Qu.:1.000   1st Qu.:   5527175   CANADA                  : 1141
Median :1.000   Median :   5627553   MEXICO                  : 1072
Mean   :1.077   Mean   :  67747386   ITALY                   :  537
3rd Qu.:1.000   3rd Qu.:187465121   UNITED KINGDOM          :  426
Max.   :8.000   Max.   :196871823   FRANCE                  :  376
                                     (Other)                 : 4432
  is_booking         hotel_id         prop_is_branded  prop_starrating
Min.   :0.00000   Min.   :      402   Min.   :0.0000   Min.   :0.000
1st Qu.:0.00000   1st Qu.:   725244   1st Qu.:0.0000   1st Qu.:3.000
Median :0.00000   Median : 21521873   Median :1.0000   Median :4.000
Mean   :0.08763   Mean   : 60268782   Mean   :0.6166   Mean   :3.528
3rd Qu.:0.00000   3rd Qu.: 76989646   3rd Qu.:1.0000   3rd Qu.:4.000
Max.   :1.00000   Max.   :410748015   Max.   :1.0000   Max.   :5.000

distance_band hist_price_band popularity_band      cnt
C :5129       H :4065         H :5972         Min.   : 1.000
F :2729       L :3870         L : 719         1st Qu.: 1.000
```

```
 M :7628       M :8076        M :5211        Median : 1.000
 VC:3155       VH:2108        VH:7969        Mean   : 1.421
 VF:1352       VL:1874        VL: 122        3rd Qu.: 1.000
                                             Max.   :38.000
```

```
In [3]: # change some variables to factors
        travel_data[,c(8,9,10,18,20)] <- lapply(travel_data[,c(8,9,10,18,20)], function(x) as.i
        # change date_time to month
        library(lubridate)
        travel_data$month <- as.factor(month(as.Date(travel_data$date_time)))
        # change check-in and check-out day to the number of days stay in the hotel.
        # if the days greater than 30 days or less than 0 days, change them to 0.
        travel_data$days <- as.integer(as.Date(travel_data$srch_co)-as.Date(travel_data$srch_ci
        travel_data$days <- ifelse(0<travel_data$days & travel_data$days<30 ,travel_data$days,0
        # change orig_destination_distance NULL to 0.
        travel_data$orig_destination_distance <- as.numeric(levels(travel_data$orig_destinatio
        #travel_data$orig_destination_distance[is.na(travel_data$orig_destination_distance)] <
        # remove redundent and non-important variables
        travel_data <- travel_data[,-c(1,4,5,7,11,12,16,19)]
        travel_data <- na.omit(travel_data)

        dim(travel_data)
        str(travel_data)
        summary(travel_data)
```

```
Warning message:
"package 'lubridate' was built under R version 3.3.3"
Attaching package: 'lubridate'

The following object is masked from 'package:base':

    date

Warning message in eval(expr, envir, enclos):
"NAs introduced by coercion"
```

1. 19986 2. 19

```
'data.frame':        19986 obs. of  19 variables:
 $ user_location_region     : Factor w/ 367 levels "***","?","00",..: 107 348 105 206 112 107 
 $ user_location_city       : Factor w/ 4315 levels "?","AACHEN","ABANCAY",..: 3720 1938 289 2
 $ orig_destination_distance: num  1243 686 0 0 950 ...
 $ is_mobile                : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 2 1 ...
 $ is_package               : Factor w/ 2 levels "0","1": 2 1 2 1 1 1 1 2 2 1 ...
 $ channel                  : Factor w/ 11 levels "231","262","293",..: 2 11 11 11 11 11 11 11
 $ srch_adults_cnt          : int  2 2 2 2 2 2 2 2 2 6 ...
 $ srch_children_cnt        : int  2 2 0 1 2 0 1 0 1 0 ...
```

```
$ srch_rm_cnt              : int  1 1 1 1 1 1 1 1 1 2 ...
$ hotel_country            : Factor w/ 140 levels "ANGOLA","ANGUILLA",..: 80 135 135 124 135
$ is_booking               : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
$ prop_is_branded          : Factor w/ 2 levels "0","1": 1 2 2 1 1 2 2 2 1 2 ...
$ prop_starrating          : int  4 4 4 4 2 4 3 5 3 5 ...
$ distance_band            : Factor w/ 5 levels "C","F","M","VC",..: 1 1 1 4 3 1 3 3 4 3 ...
$ hist_price_band          : Factor w/ 5 levels "H","L","M","VH",..: 3 4 3 3 2 3 3 3 2 1 ...
$ popularity_band          : Factor w/ 5 levels "H","L","M","VH",..: 4 4 4 4 3 4 4 4 1 4 ...
$ cnt                      : int  1 1 3 1 1 1 3 3 1 1 ...
$ month                    : Factor w/ 12 levels "1","2","3","4",..: 7 10 8 6 9 12 11 11 4 9
$ days                     : num  4 5 1 1 2 2 1 5 2 5 ...
- attr(*, "na.action")=Class 'omit'  Named int [1:7] 289 6822 7355 9678 12238 14162 17499
 .. ..- attr(*, "names")= chr [1:7] "290" "6824" "7357" "9681" ...
```

```
user_location_region   user_location_city orig_destination_distance is_mobile
CA     : 2921        NEW YORK   :  366    Min.   :    0.00          0:15527
NY     : 1236        LOS ANGELES:  266    1st Qu.:   26.58          1: 4459
TX     : 1163        TORONTO    :  232    Median :  422.30
FL     : 1063        HOUSTON    :  229    Mean   : 1330.88
ON     :  938        CHICAGO    :  215    3rd Qu.: 1767.37
WA     :  591        CALGARY    :  179    Max.   :11698.27
(Other):12074        (Other)    :18499
is_package    channel     srch_adults_cnt srch_children_cnt  srch_rm_cnt
0:16089    541    :7801   Min.   :0.000   Min.   :0.0000    Min.   :0.000
1: 3897    510    :3076   1st Qu.:2.000   1st Qu.:0.0000    1st Qu.:1.000
           231    :2612   Median :2.000   Median :0.0000    Median :1.000
           293    :2578   Mean   :2.056   Mean   :0.3106    Mean   :1.077
           262    :1796   3rd Qu.:2.000   3rd Qu.:0.0000    3rd Qu.:1.000
           324    :1276   Max.   :9.000   Max.   :8.0000    Max.   :8.000
           (Other): 847
                 hotel_country   is_booking prop_is_branded prop_starrating
UNITED STATES OF AMERICA:12003   0:18234    0: 7663         Min.   :0.000
CANADA                  : 1141   1: 1752    1:12323         1st Qu.:3.000
MEXICO                  : 1072                              Median :4.000
ITALY                   :  537                              Mean   :3.528
UNITED KINGDOM          :  426                              3rd Qu.:4.000
FRANCE                  :  376                              Max.   :5.000
(Other)                 : 4431
distance_band hist_price_band popularity_band      cnt              month
C :5129       H :4064         H :5970         Min.   : 1.000   7      :2148
F :2726       L :3869         L : 718         1st Qu.: 1.000   8      :2035
M :7626       M :8071         M :5208         Median : 1.000   6      :1974
VC:3155       VH:2108         VH:7968         Mean   : 1.421   5      :1750
VF:1350       VL:1874         VL: 122         3rd Qu.: 1.000   3      :1742
                                              Max.   :38.000   4      :1628
                                                               (Other):8709
     days
```

```
 Min.   : 0.000
 1st Qu.: 1.000
 Median : 2.000
 Mean   : 3.077
 3rd Qu.: 4.000
 Max.   :28.000
```

In [18]:
```
# basic analysis

# is_booking grouped by user_location_region,user_location_city
city<-aggregate(travel_data$is_booking, by=travel_data[,c(1,2)], FUN = length)
# is_booking grouped by channel
channel<-aggregate(travel_data$is_booking, by=list(travel_data$channel), FUN=length)
# is_booking grouped by is_mobile
mobile<-aggregate(travel_data$is_booking, by=list(travel_data$is_mobile), FUN=length)
# is_booking grouped by is_package
package<-aggregate(travel_data$is_booking, by=list(travel_data$is_package), FUN=lengt
# is_booking grouped by hotel_country
hcountry<-aggregate(travel_data$is_booking, by=list(travel_data$hotel_country), FUN=le
# is_booking grouped by prop_is_branded
brand<-aggregate(travel_data$is_booking, by=list(travel_data$prop_is_branded), FUN=le
# is_booking grouped by prop_starrating
stars<-aggregate(travel_data$is_booking, by=list(as.factor(travel_data$prop_starratin
# is_booking grouped by cnt
click<-aggregate(travel_data$is_booking, by=list(as.factor(travel_data$cnt)), FUN=len
# is_booking grouped by month
month<-aggregate(travel_data$is_booking, by=list(travel_data$month), FUN=length)
```

In [25]:
```
city[order(city$x,decreasing=T),]
```

	user_location_region	user_location_city	x
3095	NY	NEW YORK	366
2550	CA	LOS ANGELES	266
4469	ON	TORONTO	232
2002	TX	HOUSTON	229
835	IL	CHICAGO	215
645	AB	CALGARY	179
2797	FL	MIAMI	166
4029	WA	SEATTLE	156
3925	CA	SAN FRANCISCO	133
572	NY	BROOKLYN	130
4583	BC	VANCOUVER	128
3921	CA	SAN DIEGO	127
3928	CA	SAN JOSE	125
1297	AB	EDMONTON	124
1148	CO	DENVER	114
2916	QC	MONTREAL	113
1075	TX	DALLAS	107
2392	NV	LAS VEGAS	99
3473	AZ	PHOENIX	99
199	GA	ATLANTA	98
3468	PA	PHILADELPHIA	92
4685	DC	WASHINGTON	91
3301	FL	ORLANDO	87
3580	OR	PORTLAND	87
3325	ON	OTTAWA	78
1563	FL	FT LAUDERDALE	65
2997	BY	MUNICH	64
227	TX	AUSTIN	63
2849	MN	MINNEAPOLIS	63
395	BE	BERLIN	62
...	...	...	...
4922	MA	WRENTHAM	1
4923	NW	WUPPERTAL	1
4925	MN	WYKOFF	1
4927	PA	WYNNEWOOD	1
4930	CO	YAMPA	1
4931	SD	YANKTON	1
4932	M	YAVNE	1
4933	M	YEHUD	1
4934	NT	YELLOWKNIFE	1
4935	11	YEOKSAM-DONG	1
4936	11	YEONGEON-DONG	1
4938	NULL	YIGO	1
4939	16	YILDIRIM	1
4940	34	YILDIZ	1
4941	14	YOKOSUKA	1
4942	41	YONGIN	1
4945	ON	YORK	1
4947	YOR	YORK	1
4948	PA	YORK HAVEN	1
4953	NM	YOUNGSVILLE	1
4955	CA	YREKA	1

```
In [27]: channel[order(channel$x,decreasing=T),]
```

	Group.1	x
11	541	7801
10	510	3076
1	231	2612
3	293	2578
2	262	1796
4	324	1276
6	386	379
5	355	245
8	448	185
9	479	34
7	417	4

```
In [28]: mobile # more user not connect from mobile
```

Group.1	x
0	15527
1	4459

```
In [29]: package # more user not from package
```

Group.1	x
0	16089
1	3897

```
In [31]: hcountry[order(hcountry$x,decreasing=T),]
```

	Group.1	x
135	UNITED STATES OF AMERICA	12003
22	CANADA	1141
80	MEXICO	1072
62	ITALY	537
134	UNITED KINGDOM	426
42	FRANCE	376
114	SPAIN & CANARY ISLANDS	340
35	DOMINICAN REPUBLIC	275
45	GERMANY	223
126	THAILAND	193
64	JAPAN	175
133	UNITED ARAB EMIRATES	147
99	PUERTO RICO	138
47	GREECE	137
63	JAMAICA	136
129	TURKEY	125
88	NETHERLANDS	114
29	COSTA RICA	101
60	IRELAND	88
6	AUSTRALIA	78
26	CHINA	78
59	INDONESIA	78
96	PHILIPPINES	71
98	PORTUGAL	70
9	BAHAMAS	69
123	SWITZERLAND	66
55	HONG KONG	64
18	BRAZIL	63
27	COLOMBIA	62
113	SOUTH KOREA	56
...	...	...
17	BOSNIA AND HERZEGOVINA	2
38	ESTONIA	2
48	GRENADA	2
49	GUADELOUPE	2
53	HAITI	2
68	LAOS	2
71	LITHUANIA	2
74	MACEDONIA	2
78	MARTINIQUE	2
81	MOLDOVA	2
82	MONACO	2
103	RWANDA	2
121	STATE OF PALESTINE	2
139	ZAMBIA	2
1	ANGOLA	1
8	AZERBAIJAN	1
23	CAPE VERDE	1
44	GAMBIA	1
52	GUYANA	1
66	KAZAKHSTAN	1
83	MONTENEGRO	1

```
In [32]: brand # more book brand hotels
```

Group.1	x
0	7663
1	12323

```
In [34]: stars[order(stars$x,decreasing=T),] #4 and 3 stars hotel booked the most
```

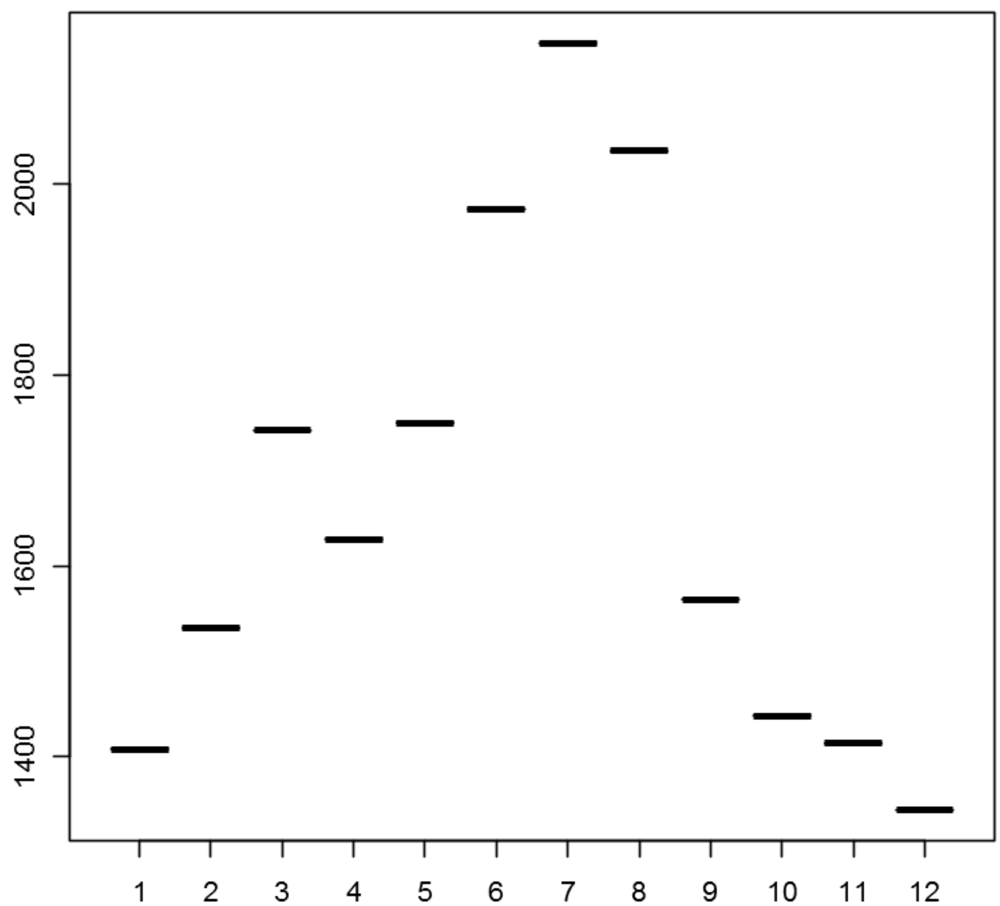
	Group.1	x
5	4	8222
4	3	6831
6	5	2623
3	2	1992
1	0	280
2	1	38

```
In [36]: click[order(click$x,decreasing=T),] # most of users click one or two times then booke
```

	Group.1	x
1	1	15717
2	2	2693
3	3	834
4	4	326
5	5	156
6	6	80
7	7	44
9	9	25
8	8	23
10	10	16
11	11	10
26	26	8
12	12	6
25	25	6
13	13	5
14	14	5
15	15	4
24	24	4
27	27	4
16	16	3
19	19	3
20	20	3
17	17	2
18	18	2
23	23	2
28	28	2
21	21	1
22	22	1
29	38	1

```
In [63]: month[order(month$x,decreasing=T),]
         plot(month$Group.1,month$x,type='l') # from may to august is the highest session book
```

	Group.1	x
7	7	2148
8	8	2035
6	6	1974
5	5	1750
3	3	1742
4	4	1628
9	9	1565
2	2	1535
10	10	1443
11	11	1414
1	1	1408
12	12	1344



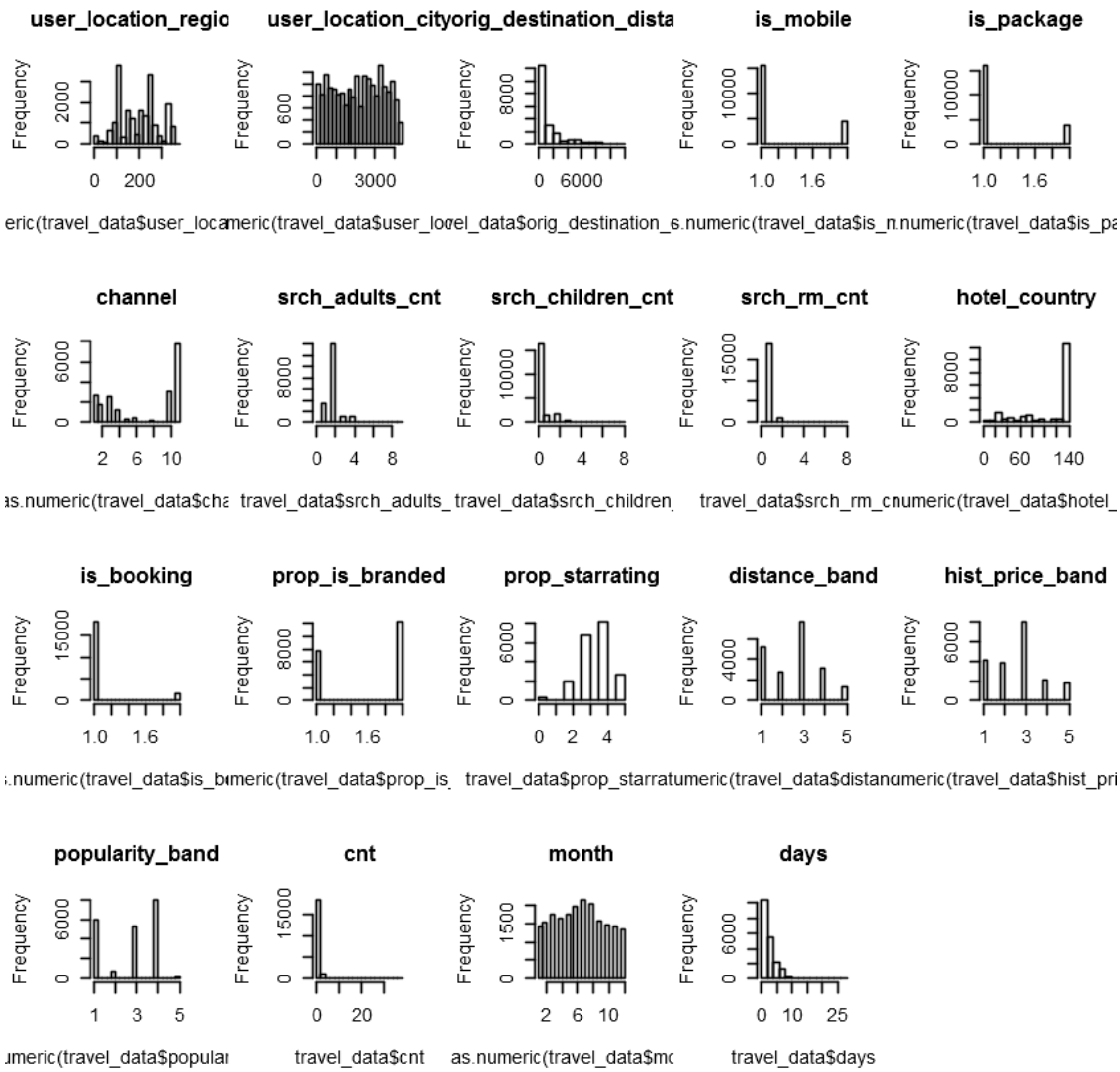
In [64]:
```
# change variables, user_location_region,user_location_city,hotel_country,distance_ba
travel_data$user_location_region <- as.factor(as.numeric(travel_data$user_location_reg
travel_data$user_location_city <- as.factor(as.numeric(travel_data$user_location_city)
travel_data$hotel_country <- as.factor(as.numeric(travel_data$hotel_country)-1)
travel_data$distance_band <- as.factor(as.numeric(travel_data$distance_band)-1)
travel_data$hist_price_band <- as.factor(as.numeric(travel_data$hist_price_band)-1)
travel_data$popularity_band <- as.factor(as.numeric(travel_data$popularity_band)-1)
summary(travel_data)
```

```
 user_location_region user_location_city orig_destination_distance is_mobile
 106    : 2921        2653   :  366      Min.   :    0.00          0:15527
 248    : 1236        2192   :  266      1st Qu.:   26.58          1: 4459
 327    : 1163        3887   :  232      Median :  422.30
 143    : 1063        1720   :  229      Mean   : 1330.88
 255    :  938        705    :  215      3rd Qu.: 1767.37
 347    :  591        538    :  179      Max.   :11698.27
 (Other):12074        (Other):18499
 is_package    channel     srch_adults_cnt srch_children_cnt  srch_rm_cnt
 0:16089    541    :7801   Min.   :0.000   Min.   :0.0000    Min.   :0.000
 1: 3897    510    :3076   1st Qu.:2.000   1st Qu.:0.0000    1st Qu.:1.000
            231    :2612   Median :2.000   Median :0.0000    Median :1.000
            293    :2578   Mean   :2.056   Mean   :0.3106    Mean   :1.077
            262    :1796   3rd Qu.:2.000   3rd Qu.:0.0000    3rd Qu.:1.000
            324    :1276   Max.   :9.000   Max.   :8.0000    Max.   :8.000
            (Other): 847
 hotel_country   is_booking prop_is_branded prop_starrating distance_band
 134    :12003   0:18234    0: 7663         Min.   :0.000   0:5129
 21     : 1141   1: 1752    1:12323         1st Qu.:3.000   1:2726
 79     : 1072                              Median :4.000   2:7626
 61     :  537                              Mean   :3.528   3:3155
 133    :  426                              3rd Qu.:4.000   4:1350
 41     :  376                              Max.   :5.000
 (Other): 4431
 hist_price_band popularity_band      cnt              month
 0:4064          0:5970          Min.   : 1.000   7      :2148
 1:3869          1: 718          1st Qu.: 1.000   8      :2035
 2:8071          2:5208          Median : 1.000   6      :1974
 3:2108          3:7968          Mean   : 1.421   5      :1750
 4:1874          4: 122          3rd Qu.: 1.000   3      :1742
                                 Max.   :38.000   4      :1628
                                                  (Other):8709
      days
 Min.   : 0.000
 1st Qu.: 1.000
 Median : 2.000
```

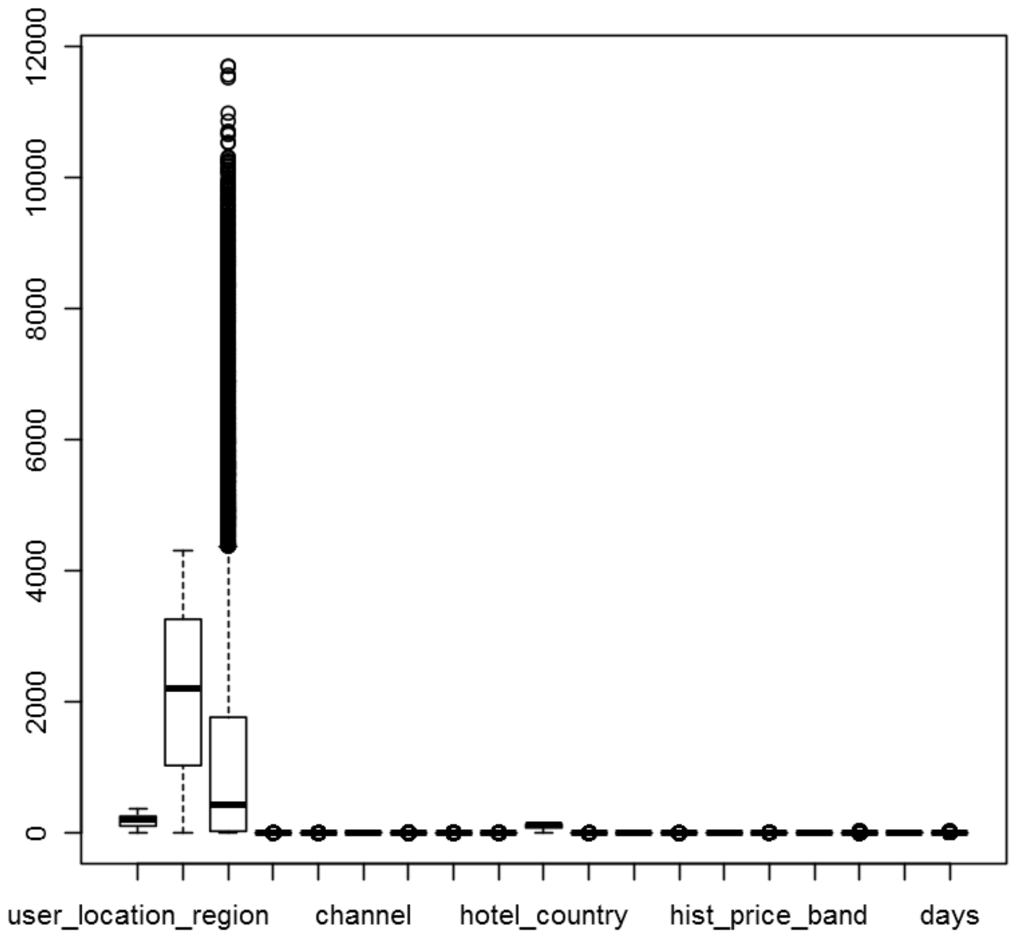
```
 Mean   : 3.077
 3rd Qu.: 4.000
 Max.   :28.000
```

```
In [65]: names(travel_data)
         par(mfrow=c(4,5))
         hist(as.numeric(travel_data$user_location_region),main='user_location_region')
         hist(as.numeric(travel_data$user_location_city),main='user_location_city')
         hist(travel_data$orig_destination_distance,main='orig_destination_distance')
         hist(as.numeric(travel_data$is_mobile),main='is_mobile')
         hist(as.numeric(travel_data$is_package),main='is_package')
         hist(as.numeric(travel_data$channel),main='channel')
         hist(travel_data$srch_adults_cnt,main='srch_adults_cnt')
         hist(travel_data$srch_children_cnt,main='srch_children_cnt')
         hist(travel_data$srch_rm_cnt,main='srch_rm_cnt')
         hist(as.numeric(travel_data$hotel_country),main='hotel_country')
         hist(as.numeric(travel_data$is_booking),main='is_booking')
         hist(as.numeric(travel_data$prop_is_branded),main='prop_is_branded')
         hist(travel_data$prop_starrating,main='prop_starrating')
         hist(as.numeric(travel_data$distance_band),main='distance_band')
         hist(as.numeric(travel_data$hist_price_band),main='hist_price_band')
         hist(as.numeric(travel_data$popularity_band),main='popularity_band')
         hist(travel_data$cnt,main='cnt')
         hist(as.numeric(travel_data$month),main='month')
         hist(travel_data$days,main='days')
```

1. 'user_location_region' 2. 'user_location_city' 3. 'orig_destination_distance' 4. 'is_mobile' 5. 'is_package' 6. 'channel' 7. 'srch_adults_cnt' 8. 'srch_children_cnt' 9. 'srch_rm_cnt' 10. 'hotel_country' 11. 'is_booking' 12. 'prop_is_branded' 13. 'prop_starrating' 14. 'distance_band' 15. 'hist_price_band' 16. 'popularity_band' 17. 'cnt' 18. 'month' 19. 'days'



In [66]:
```
# boxplot
par(mfrow=c(1,1))
boxplot(travel_data)
```



4.2 Split the dataset into training and testing dataset

```
In [67]: # train dataset: 80%, test dataset: 20%
         set.seed(6)
         (n <- dim(travel_data)[1])
         trainID <- sample(1:n,n*0.8)
         train <- travel_data[trainID,]
         test <- travel_data[-trainID,]
         dim(train)
         dim(test)
```

19986
1. 15988 2. 19
1. 3998 2. 19

4.3 Build Models

```
In [169]: # use stepAIC find the logit.best1 model
          #I runed about 2 hours, it still not get the results and my laptop shutdown automati
          # So remove user_location_region and user_location_city
          logit.fit1=glm(is_booking~.,family=binomial,data=train[,-c(1,2)])
          summary(logit.fit1) # AIC: 8637.5
```

```
Call:
glm(formula = is_booking ~ ., family = binomial, data = train[,
    -c(1, 2)])

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2697  -0.5164  -0.3523  -0.1161   3.3599

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.317e+01  3.956e+03  -0.003 0.997344
orig_destination_distance  5.083e-06  2.067e-05   0.246 0.805770
is_mobile1                -1.869e-01  7.357e-02  -2.541 0.011056 *
is_package1               -6.379e-01  1.083e-01  -5.892 3.81e-09 ***
channel262                -5.339e-02  1.240e-01  -0.431 0.666759
channel293                -3.680e-01  1.188e-01  -3.097 0.001954 **
channel324                 1.091e-01  1.345e-01   0.811 0.417313
channel355                 5.488e-01  2.214e-01   2.479 0.013174 *
channel386                 2.533e-01  2.207e-01   1.148 0.251157
channel417                -1.588e+01  2.282e+03  -0.007 0.994450
channel448                -1.244e+00  4.675e-01  -2.661 0.007782 **
channel479                 7.344e-01  5.214e-01   1.408 0.159028
channel510                -1.584e-02  1.226e-01  -0.129 0.897209
channel541                -3.787e-02  9.460e-02  -0.400 0.688893
srch_adults_cnt           -9.428e-02  3.998e-02  -2.359 0.018348 *
srch_children_cnt         -7.429e-02  4.252e-02  -1.747 0.080615 .
srch_rm_cnt                5.540e-02  8.488e-02   0.653 0.513949
hotel_country1             1.818e+01  3.956e+03   0.005 0.996333
hotel_country2             7.200e-01  4.100e+03   0.000 0.999860
hotel_country3             1.596e+01  3.956e+03   0.004 0.996781
hotel_country4             1.516e+01  3.956e+03   0.004 0.996942
hotel_country5             1.499e+01  3.956e+03   0.004 0.996977
hotel_country6             1.500e+01  3.956e+03   0.004 0.996975
hotel_country7            -5.777e-01  5.595e+03   0.000 0.999918
hotel_country8             1.382e+01  3.956e+03   0.003 0.997213
hotel_country9             6.170e-01  4.236e+03   0.000 0.999884
hotel_country10            5.142e-01  4.022e+03   0.000 0.999898
hotel_country11            1.459e+01  3.956e+03   0.004 0.997058
hotel_country12            1.512e+01  3.956e+03   0.004 0.996951
hotel_country13            1.551e+01  3.956e+03   0.004 0.996872
```

hotel_country14	1.707e+01	3.956e+03	0.004	0.996557
hotel_country15	-5.051e-02	4.845e+03	0.000	0.999992
hotel_country16	3.531e+01	5.595e+03	0.006	0.994965
hotel_country17	1.546e+01	3.956e+03	0.004	0.996882
hotel_country18	1.668e+01	3.956e+03	0.004	0.996636
hotel_country19	4.079e-01	4.517e+03	0.000	0.999928
hotel_country20	1.466e+01	3.956e+03	0.004	0.997044
hotel_country21	1.497e+01	3.956e+03	0.004	0.996981
hotel_country22	3.905e+01	5.595e+03	0.007	0.994431
hotel_country23	3.724e-01	4.092e+03	0.000	0.999927
hotel_country24	2.882e-01	4.030e+03	0.000	0.999943
hotel_country25	2.805e-02	3.983e+03	0.000	0.999994
hotel_country26	1.516e+01	3.956e+03	0.004	0.996943
hotel_country27	1.756e+01	3.956e+03	0.004	0.996459
hotel_country28	1.524e+01	3.956e+03	0.004	0.996926
hotel_country29	1.580e+01	3.956e+03	0.004	0.996813
hotel_country30	5.560e-01	4.148e+03	0.000	0.999893
hotel_country31	6.200e-01	4.211e+03	0.000	0.999883
hotel_country32	1.521e+01	3.956e+03	0.004	0.996932
hotel_country33	1.575e+01	3.956e+03	0.004	0.996824
hotel_country34	1.431e+01	3.956e+03	0.004	0.997114
hotel_country35	-2.702e-02	4.315e+03	0.000	0.999995
hotel_country36	1.569e+01	3.956e+03	0.004	0.996835
hotel_country37	7.385e-01	4.720e+03	0.000	0.999875
hotel_country38	1.736e+01	3.956e+03	0.004	0.996500
hotel_country39	1.578e+01	3.956e+03	0.004	0.996817
hotel_country40	-2.415e-01	4.137e+03	0.000	0.999953
hotel_country41	1.470e+01	3.956e+03	0.004	0.997034
hotel_country42	4.034e-01	4.075e+03	0.000	0.999921
hotel_country43	4.731e+00	5.595e+03	0.001	0.999325
hotel_country44	1.452e+01	3.956e+03	0.004	0.997072
hotel_country45	3.048e-01	4.515e+03	0.000	0.999946
hotel_country46	1.482e+01	3.956e+03	0.004	0.997011
hotel_country48	5.596e-01	4.838e+03	0.000	0.999908
hotel_country49	1.537e+00	4.551e+03	0.000	0.999731
hotel_country50	1.500e+01	3.956e+03	0.004	0.996974
hotel_country52	3.466e+00	5.595e+03	0.001	0.999506
hotel_country53	-9.789e-04	4.207e+03	0.000	1.000000
hotel_country54	1.480e+01	3.956e+03	0.004	0.997015
hotel_country55	1.420e+01	3.956e+03	0.004	0.997136
hotel_country56	1.529e+01	3.956e+03	0.004	0.996916
hotel_country57	1.380e+01	3.956e+03	0.003	0.997217
hotel_country58	1.545e+01	3.956e+03	0.004	0.996885
hotel_country59	1.454e+01	3.956e+03	0.004	0.997068
hotel_country60	-3.270e-01	4.019e+03	0.000	0.999935
hotel_country61	1.459e+01	3.956e+03	0.004	0.997058
hotel_country62	1.396e+01	3.956e+03	0.004	0.997185
hotel_country63	1.476e+01	3.956e+03	0.004	0.997023

hotel_country64	1.705e+01	3.956e+03	0.004	0.996561
hotel_country65	5.821e+00	5.595e+03	0.001	0.999170
hotel_country66	3.642e-01	4.255e+03	0.000	0.999932
hotel_country67	2.069e+00	5.595e+03	0.000	0.999705
hotel_country68	5.561e-01	4.366e+03	0.000	0.999898
hotel_country69	7.842e+00	5.595e+03	0.001	0.998882
hotel_country70	-1.931e-01	4.832e+03	0.000	0.999968
hotel_country71	-1.339e+00	5.595e+03	0.000	0.999809
hotel_country72	1.711e+01	3.956e+03	0.004	0.996550
hotel_country73	4.842e-02	4.836e+03	0.000	0.999992
hotel_country74	1.446e+01	3.956e+03	0.004	0.997084
hotel_country75	6.788e-01	4.003e+03	0.000	0.999865
hotel_country76	1.648e+01	3.956e+03	0.004	0.996675
hotel_country77	1.865e+00	4.615e+03	0.000	0.999678
hotel_country78	2.021e-01	4.178e+03	0.000	0.999961
hotel_country79	1.480e+01	3.956e+03	0.004	0.997016
hotel_country80	1.849e+01	3.956e+03	0.005	0.996271
hotel_country81	7.631e-01	5.595e+03	0.000	0.999891
hotel_country82	1.744e+00	5.595e+03	0.000	0.999751
hotel_country83	1.699e+00	4.051e+03	0.000	0.999665
hotel_country84	-5.531e-01	4.546e+03	0.000	0.999903
hotel_country85	4.885e+00	5.595e+03	0.001	0.999303
hotel_country86	1.694e+01	3.956e+03	0.004	0.996583
hotel_country87	1.479e+01	3.956e+03	0.004	0.997016
hotel_country88	1.525e+01	3.956e+03	0.004	0.996925
hotel_country89	1.542e+01	3.956e+03	0.004	0.996891
hotel_country90	1.499e+00	5.595e+03	0.000	0.999786
hotel_country91	1.492e+01	3.956e+03	0.004	0.996991
hotel_country92	8.844e-01	4.458e+03	0.000	0.999842
hotel_country93	1.469e+01	3.956e+03	0.004	0.997036
hotel_country94	1.517e+01	3.956e+03	0.004	0.996941
hotel_country95	1.362e+01	3.956e+03	0.003	0.997253
hotel_country96	1.514e+01	3.956e+03	0.004	0.996946
hotel_country97	1.575e+01	3.956e+03	0.004	0.996824
hotel_country98	1.443e+01	3.956e+03	0.004	0.997090
hotel_country99	2.046e-01	4.203e+03	0.000	0.999961
hotel_country100	-6.795e-02	4.388e+03	0.000	0.999988
hotel_country101	1.608e+01	3.956e+03	0.004	0.996757
hotel_country102	-1.830e-01	4.760e+03	0.000	0.999969
hotel_country103	1.582e+01	3.956e+03	0.004	0.996810
hotel_country104	7.537e+00	5.595e+03	0.001	0.998925
hotel_country105	9.656e+00	5.595e+03	0.002	0.998623
hotel_country106	2.834e-02	4.833e+03	0.000	0.999995
hotel_country107	1.423e+01	3.956e+03	0.004	0.997131
hotel_country108	9.334e-01	4.128e+03	0.000	0.999820
hotel_country109	-1.042e+00	5.595e+03	0.000	0.999851
hotel_country110	1.168e-01	5.595e+03	0.000	0.999983
hotel_country111	1.566e+01	3.956e+03	0.004	0.996841

```
hotel_country112           1.489e+01  3.956e+03   0.004 0.996998
hotel_country113           1.470e+01  3.956e+03   0.004 0.997035
hotel_country114          -3.721e-01  4.415e+03   0.000 0.999933
hotel_country116           7.163e-01  4.388e+03   0.000 0.999870
hotel_country117           9.963e-01  4.025e+03   0.000 0.999802
hotel_country118           5.714e-01  4.159e+03   0.000 0.999890
hotel_country119           2.866e+00  4.844e+03   0.001 0.999528
hotel_country120           4.716e-01  4.754e+03   0.000 0.999921
hotel_country121           1.489e+01  3.956e+03   0.004 0.996997
hotel_country122           1.440e+01  3.956e+03   0.004 0.997095
hotel_country123           1.535e+01  3.956e+03   0.004 0.996903
hotel_country124           1.176e+00  4.206e+03   0.000 0.999777
hotel_country125           1.504e+01  3.956e+03   0.004 0.996967
hotel_country126           1.198e-01  4.560e+03   0.000 0.999979
hotel_country127           3.152e-01  4.843e+03   0.000 0.999948
hotel_country128           1.498e+01  3.956e+03   0.004 0.996979
hotel_country129           7.257e-01  4.025e+03   0.000 0.999856
hotel_country130           1.460e+01  3.956e+03   0.004 0.997055
hotel_country131           5.444e-01  4.188e+03   0.000 0.999896
hotel_country132           1.555e+01  3.956e+03   0.004 0.996863
hotel_country133           1.521e+01  3.956e+03   0.004 0.996933
hotel_country134           1.509e+01  3.956e+03   0.004 0.996957
hotel_country135           4.651e-01  4.290e+03   0.000 0.999913
hotel_country136           4.834e-01  5.595e+03   0.000 0.999931
hotel_country137           1.484e+01  3.956e+03   0.004 0.997008
hotel_country138          -1.665e-01  4.844e+03   0.000 0.999973
prop_is_branded1           2.705e-01  6.728e-02   4.021 5.81e-05 ***
prop_starrating           -2.047e-01  4.087e-02  -5.010 5.44e-07 ***
distance_band1            -1.566e-01  1.032e-01  -1.518 0.129117
distance_band2             5.682e-02  7.435e-02   0.764 0.444770
distance_band3            -1.141e-01  9.400e-02  -1.213 0.224961
distance_band4             1.450e-01  1.238e-01   1.171 0.241475
hist_price_band1           1.237e-01  9.956e-02   1.242 0.214152
hist_price_band2           1.689e-02  8.373e-02   0.202 0.840123
hist_price_band3           1.710e-01  1.146e-01   1.493 0.135421
hist_price_band4          -2.398e-01  1.319e-01  -1.817 0.069142 .
popularity_band1          -8.256e-01  2.090e-01  -3.950 7.81e-05 ***
popularity_band2          -1.212e-01  8.101e-02  -1.496 0.134546
popularity_band3           2.775e-01  7.137e-02   3.889 0.000101 ***
popularity_band4          -7.086e-01  4.718e-01  -1.502 0.133146
cnt                       -2.709e+00  2.232e-01 -12.139  < 2e-16 ***
month2                     2.967e-02  1.578e-01   0.188 0.850892
month3                     1.446e-02  1.501e-01   0.096 0.923264
month4                    -6.771e-02  1.536e-01  -0.441 0.659379
month5                     8.374e-02  1.607e-01   0.521 0.602384
month6                     7.190e-02  1.606e-01   0.448 0.654326
month7                    -2.108e-01  1.650e-01  -1.277 0.201443
month8                    -1.345e-02  1.614e-01  -0.083 0.933597
```

```
month9                      1.217e-01  1.668e-01   0.730 0.465492
month10                     1.785e-01  1.689e-01   1.057 0.290540
month11                     7.827e-02  1.744e-01   0.449 0.653504
month12                     2.314e-01  1.721e-01   1.345 0.178604
days                       -1.808e-01  1.959e-02  -9.230  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9483.3  on 15987  degrees of freedom
Residual deviance: 8279.5  on 15809  degrees of freedom
AIC: 8637.5

Number of Fisher Scoring iterations: 16
```

In [168]:
```
# use stepAIC to get the best model
library(MASS)
logit.best1=stepAIC(logit.fit1,direction="both", trace = FALSE,data=train[,-c(1,2)])
summary(logit.best1) # AIC: 8507.1
# variables:srch_children_cnt is not significant, remove it.
```

```
Warning message:
"package 'MASS' was built under R version 3.3.3"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
```

```
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
```

```
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"


Call:
glm(formula = is_booking ~ is_mobile + is_package + channel +
    srch_adults_cnt + srch_children_cnt + prop_is_branded + prop_starrating +
    distance_band + hist_price_band + popularity_band + cnt +
    days, family = binomial, data = train[, -c(1, 2)])

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0124  -0.5192  -0.3660  -0.1389   3.5189

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)         1.807502   0.297697   6.072 1.27e-09 ***
is_mobile1         -0.176136   0.072336  -2.435 0.014893 *
is_package1        -0.700497   0.105022  -6.670 2.56e-11 ***
channel262         -0.052573   0.123091  -0.427 0.669303
channel293         -0.381748   0.117692  -3.244 0.001180 **
channel324          0.074915   0.132592   0.565 0.572069
channel355          0.561647   0.218070   2.576 0.010009 *
channel386          0.176829   0.217327   0.814 0.415842
channel417        -10.758173 187.377734  -0.057 0.954215
channel448         -1.238959   0.466169  -2.658 0.007867 **
channel479          0.788797   0.518156   1.522 0.127930
channel510         -0.076177   0.106833  -0.713 0.475813
channel541         -0.017783   0.088897  -0.200 0.841452
srch_adults_cnt    -0.081639   0.036589  -2.231 0.025665 *
srch_children_cnt  -0.068989   0.042111  -1.638 0.101371
prop_is_branded1    0.305120   0.063121   4.834 1.34e-06 ***
prop_starrating    -0.187794   0.037203  -5.048 4.47e-07 ***
distance_band1     -0.181276   0.101803  -1.781 0.074967 .
distance_band2      0.057788   0.073548   0.786 0.432027
distance_band3     -0.103343   0.093326  -1.107 0.268151
```

```
distance_band4      0.158380   0.121498   1.304 0.192386
hist_price_band1    0.121853   0.096986   1.256 0.208972
hist_price_band2    0.001405   0.082080   0.017 0.986343
hist_price_band3    0.115093   0.113105   1.018 0.308877
hist_price_band4   -0.210420   0.128388  -1.639 0.101227
popularity_band1   -0.795602   0.208055  -3.824 0.000131 ***
popularity_band2   -0.092947   0.079296  -1.172 0.241137
popularity_band3    0.275299   0.070514   3.904 9.45e-05 ***
popularity_band4   -0.723237   0.469936  -1.539 0.123801
cnt                -2.640927   0.216090 -12.221  < 2e-16 ***
days               -0.191243   0.018783 -10.181  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9483.3  on 15987  degrees of freedom
Residual deviance: 8445.1  on 15957  degrees of freedom
AIC: 8507.1

Number of Fisher Scoring iterations: 11
```

```
In [170]: # remove srch_children_cnt
          logit.best1 <- glm(formula = is_booking ~ is_mobile + is_package + channel +
              srch_adults_cnt  + prop_is_branded + prop_starrating +
              distance_band + hist_price_band + popularity_band + cnt +
              days, family = binomial, data = train[, -c(1, 2)])
          summary(logit.best1) # AIC: 8507.9
```

```
Warning message:
"glm.fit: fitted probabilities numerically 0 or 1 occurred"


Call:
glm(formula = is_booking ~ is_mobile + is_package + channel +
    srch_adults_cnt + prop_is_branded + prop_starrating + distance_band +
    hist_price_band + popularity_band + cnt + days, family = binomial,
    data = train[, -c(1, 2)])

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0273  -0.5192  -0.3663  -0.1396   3.5259

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)        1.804052   0.297813   6.058 1.38e-09 ***
```

```
is_mobile1        -0.175075   0.072318  -2.421  0.01548 *
is_package1       -0.697306   0.105028  -6.639 3.15e-11 ***
channel262        -0.051581   0.123065  -0.419  0.67512
channel293        -0.380355   0.117678  -3.232  0.00123 **
channel324         0.081007   0.132512   0.611  0.54099
channel355         0.557752   0.218065   2.558  0.01054 *
channel386         0.181859   0.217270   0.837  0.40258
channel417       -10.735697 187.379383  -0.057  0.95431
channel448        -1.228628   0.466082  -2.636  0.00839 **
channel479         0.805183   0.517946   1.555  0.12005
channel510        -0.075253   0.106810  -0.705  0.48109
channel541        -0.018351   0.088880  -0.206  0.83643
srch_adults_cnt   -0.088386   0.036429  -2.426  0.01525 *
prop_is_branded1   0.299662   0.063032   4.754 1.99e-06 ***
prop_starrating   -0.185572   0.037237  -4.984 6.24e-07 ***
distance_band1    -0.184832   0.101769  -1.816  0.06934 .
distance_band2     0.054592   0.073510   0.743  0.45769
distance_band3    -0.103447   0.093322  -1.109  0.26765
distance_band4     0.152857   0.121444   1.259  0.20815
hist_price_band1   0.124505   0.096965   1.284  0.19913
hist_price_band2   0.002048   0.082066   0.025  0.98009
hist_price_band3   0.113960   0.113099   1.008  0.31364
hist_price_band4  -0.204585   0.128326  -1.594  0.11088
popularity_band1  -0.795968   0.208006  -3.827  0.00013 ***
popularity_band2  -0.092916   0.079283  -1.172  0.24122
popularity_band3   0.274264   0.070496   3.890  0.00010 ***
popularity_band4  -0.717964   0.469819  -1.528  0.12647
cnt               -2.643898   0.216083 -12.236  < 2e-16 ***
days              -0.192581   0.018787 -10.251  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9483.3  on 15987  degrees of freedom
Residual deviance: 8447.9  on 15958  degrees of freedom
AIC: 8507.9

Number of Fisher Scoring iterations: 11
```

In [68]:
```
# Lasso regression
library(glmnet)
# create dummy variables
train.xmat <- model.matrix(is_booking~., data=train)[,-1]
test.xmat <- model.matrix(is_booking~., data=test)[,-1]
```

```
lasso.fit <- cv.glmnet(train.xmat, train$is_booking, alpha=1,family="binomial")
(lambda <- lasso.fit$lambda.min)  # 0.00472111648596817
```

```
Warning message:
"package 'glmnet' was built under R version 3.3.3"Loading required package: Matrix
Warning message:
"package 'Matrix' was built under R version 3.3.3"Loading required package: foreach
Warning message:
"package 'foreach' was built under R version 3.3.3"Loaded glmnet 2.0-13
```

0.00472111648596817

```
In [89]: # coeffecient
         ce <- coef(lasso.fit, s = "lambda.min")[which(coef(lasso.fit, s = "lambda.min") != 0)]
         # features that selected.
         fe <- colnames(train.xmat)[which(coef(lasso.fit, s = "lambda.min") != 0)]
         (n <- length(fe))
         # the fomula of this model
         fomula <- paste('log(pr(is_booking=1)/(1-pr(is_booking=1)))','=',ce[1],'*',fe[1])
         for (i in 2:n){
             fomula <- paste(fomula,'+',ce[i],'*',fe[i])
         }
         fomula
```

300

'log(pr(is_booking=1)/(1-pr(is_booking=1))) = -0.840175372380394 * user_location_region1 + 1.70064665164346 * user_location_region22 + 1.75037430767309 * user_location_region50 + 1.34757146927526 * user_location_region52 + 1.92621638805609 * user_location_region57 + 2.1966114724833 * user_location_region96 + 1.64498572322229 * user_location_region97 + 1.47093185065033 * user_location_region132 + 1.66825471520191 * user_location_region136 + 1.65327994128566 * user_location_region170 + 3.30869745225522 * user_location_region203 + 1.48393368588763 * user_location_region227 + 0.0776302603633484 * user_location_region230 + 0.379842165348271 * user_location_region262 + 2.3547967267212 * user_location_region266 + 1.72465244977372 * user_location_region279 + 0.183637427426377 * user_location_region282 + 2.36255454869789 * user_location_region286 + 1.61669538815595 * user_location_region296 + 1.73346642783372 * user_location_region323 + 1.71242619221405 * user_location_region345 + 2.66866636854355 * user_location_region357 + 1.72830370889992 * user_location_city23 + 1.91328689457381 * user_location_city25 + 1.82170701871056 * user_location_city43 + 1.34578747648795 * user_location_city62 + 0.171065000451806 * user_location_city145 + 0.769946669660274 * user_location_city154 + 1.46179218712341 * user_location_city161 + 1.91255710647529 * user_location_city188 + 1.57340288072404 * user_location_city190 + 2.66744055985309 * user_location_city202 + 0.625201658293635 * user_location_city210 + 1.20527540484896 * user_location_city225 + 0.00815313008617242 * user_location_city233 + 1.26156000929005 * user_location_city255 + 2.21353619712526 * user_location_city281 + 1.58240429249761 * user_location_city292 + 1.30318576331268 * user_location_city308 + 1.48776649780589 * user_location_city326 + 0.221991050010289 * user_location_city370 + 0.00897283018513941 * user_location_city396 + 2.08388660865866 * user_location_city399

+ 1.50810302908296 * user_location_city426 + 0.673150030556413 * user_location_city461 + 0.116011183583241 * user_location_city480 + 1.41594090800985 * user_location_city488 + 1.70405084239088 * user_location_city508 + 2.2956656986328 * user_location_city522 + 1.34608110011558 * user_location_city556 + 1.423711359596 * user_location_city559 + 1.42371135871114 * user_location_city573 + 0.20720038179701 * user_location_city595 + 0.618241346360835 * user_location_city607 + 0.0411031171108332 * user_location_city618 + 1.68226468836783 * user_location_city621 + 0.202959456053917 * user_location_city674 + 0.759334581797492 * user_location_city701 + 1.39265613284325 * user_location_city717 + 1.30446902568766 * user_location_city745 + 1.80742677118418 * user_location_city766 + 1.28519822850426 * user_location_city780 + 1.89229105223772 * user_location_city781 + 1.41594046593867 * user_location_city803 + 0.00764552989074245 * user_location_city807 + 1.56053485931506 * user_location_city811 + 1.51959680530411 * user_location_city819 + 1.37958746251617 * user_location_city823 + 2.1441767229971 * user_location_city824 + 1.35086935651256 * user_location_city860 + 0.0710943297740069 * user_location_city885 + 0.0466781230604732 * user_location_city903 + 1.43681538411199 * user_location_city907 + 0.130698050124748 * user_location_city910 + 0.271604556939314 * user_location_city911 + 1.94684193332267 * user_location_city915 + 1.93310636708276 * user_location_city952 + 1.58761010304687 * user_location_city965 + 1.20620315023935 * user_location_city975 + 1.10073170629572 * user_location_city977 + 1.6031510117803 * user_location_city980 + 1.69097798917911 * user_location_city988 + 1.42810151061056 * user_location_city1013 + 0.0173664772915361 * user_location_city1032 + 1.97783520565801 * user_location_city1060 + 1.69531316730153 * user_location_city1080 + 1.28899959933342 * user_location_city1086 + 0.0167581868446094 * user_location_city1101 + 1.5721576153883 * user_location_city1145 + 1.64260145085358 * user_location_city1154 + 0.705731557702907 * user_location_city1163 + 1.15517055215293 * user_location_city1170 + 1.54288955764277 * user_location_city1173 + 1.56438252651794 * user_location_city1180 + 0.412690526502291 * user_location_city1245 + 1.73717444923818 * user_location_city1257 + 2.38126002372614 * user_location_city1261 + 2.3787338292697 * user_location_city1291 + 1.45142075135272 * user_location_city1328 + 1.75502832813195 * user_location_city1335 + 2.70859467559473 * user_location_city1340 + 1.63818113374493 * user_location_city1360 + 1.61444253481964 * user_location_city1385 + 1.65524642533126 * user_location_city1412 + 1.64488800378399 * user_location_city1430 + 1.48776117301537 * user_location_city1433 + 1.32444037806607 * user_location_city1440 + 1.35478954587678 * user_location_city1442 + 1.30453579653865 * user_location_city1449 + 2.81148304322146 * user_location_city1478 + 3.06542496623121 * user_location_city1524 + 1.41593552700032 * user_location_city1525 + 0.141838688657448 * user_location_city1551 + 0.0180337853924606 * user_location_city1556 + 2.39285315223249 * user_location_city1564 + 1.45553683117395 * user_location_city1575 + 3.37109136045686 * user_location_city1594 + 1.42370603411847 * user_location_city1596 + 1.70754792902532 * user_location_city1597 + 1.74382248139009 * user_location_city1604 + 1.99578799236085 * user_location_city1607 + 1.87968702402652 * user_location_city1618 + 1.8484835534846 * user_location_city1626 + 0.195452375133235 * user_location_city1640 + 1.55038735070886 * user_location_city1698 + 0.0223783732009064 * user_location_city1720 + 2.83241933485316 * user_location_city1730 + 1.44099476551541 * user_location_city1771 + 1.67675235689419 * user_location_city1774 + 2.36442573933774 * user_location_city1800 + 0.0187045158538939 * user_location_city1804 + 1.06823492552214 * user_location_city1810 + 2.13849975397308 * user_location_city1819 + 0.685539856971392 * user_location_city1860 + 1.63841214275929 * user_location_city1874 + 1.75253611695693 * user_location_city1881 + 1.2933853385294 * user_location_city1894 + 1.29676516573578 * user_location_city1895 + 1.58760510419444 * user_location_city1908 +

1.76409180290395 * user_location_city1950 + 2.63653535284903 * user_location_city1956 + 1.52787846172861 * user_location_city1962 + 0.0223255307280383 * user_location_city1968 + 1.99064707662391 * user_location_city1979 + 2.18853880939357 * user_location_city1980 + 0.768887557596195 * user_location_city1984 + 2.01749304129641 * user_location_city1992 + 0.0140053569477208 * user_location_city2014 + 1.41593540858525 * user_location_city2034 + 1.35965079573736 * user_location_city2039 + 1.73614041442272 * user_location_city2044 + 1.51959174180391 * user_location_city2072 + 0.300559715124701 * user_location_city2078 + 1.43241959305018 * user_location_city2102 + 0.00347860498983713 * user_location_city2103 + 2.76174063262805 * user_location_city2114 + 2.25665304908358 * user_location_city2119 + 0.641235304492185 * user_location_city2121 + 0.535850561451285 * user_location_city2142 + 0.645058462797759 * user_location_city2145 + 1.71688377443262 * user_location_city2146 + 0.0280810470530736 * user_location_city2159 + 1.32859581144242 * user_location_city2180 + 0.0295914509071797 * user_location_city2205 + 1.44028055705431 * user_location_city2217 + 1.65771368540736 * user_location_city2219 + 2.49726189945697 * user_location_city2232 + 1.77160273732166 * user_location_city2241 + 1.59009697116719 * user_location_city2264 + 1.54906452921943 * user_location_city2300 + 1.36521074381114 * user_location_city2339 + 1.71282882770461 * user_location_city2344 + 1.36082007590793 * user_location_city2349 + 0.316885555520117 * user_location_city2372 + 1.57215246844488 * user_location_city2374 + -0.100530683153615 * user_location_city2405 + 2.23144801119273 * user_location_city2410 + 0.026486576985917 * user_location_city2417 + 1.34775152906851 * user_location_city2422 + 1.4402805430411 * user_location_city2460 + 0.677163400896585 * user_location_city2465 + 0.0356138983968096 * user_location_city2475 + 1.4867497564737 * user_location_city2513 + 1.74725545171385 * user_location_city2519 + 2.01954811849554 * user_location_city2531 + 1.50149434487202 * user_location_city2538 + 1.5919935394011 * user_location_city2577 + 1.61398477058887 * user_location_city2578 + 1.65771160686041 * user_location_city2601 + 0.641168791645382 * user_location_city2605 + 1.71912856848315 * user_location_city2607 + 1.26932268654331 * user_location_city2652 + 0.0323713575015137 * user_location_city2692 + 1.58714734022335 * user_location_city2703 + 0.428308556786427 * user_location_city2740 + 2.10560576684296 * user_location_city2750 + 1.59649526982019 * user_location_city2757 + 1.58769097934811 * user_location_city2759 + 0.00383717182970205 * user_location_city2799 + 1.44182943160601 * user_location_city2808 + 1.56793277815932 * user_location_city2833 + 1.24345450476768 * user_location_city2835 + 1.45553404302367 * user_location_city2880 + 0.137906917624647 * user_location_city2900 + 1.33997846365577 * user_location_city2909 + 1.53475305721225 * user_location_city2934 + 1.60540920670948 * user_location_city2936 + 0.0303729715924401 * user_location_city2965 + 1.58347371591598 * user_location_city2967 + 1.96942591317618 * user_location_city3001 + 1.57331940760417 * user_location_city3026 + 1.44348401842475 * user_location_city3038 + 0.502143637809217 * user_location_city3058 + 1.59365558190658 * user_location_city3069 + 1.51958862908655 * user_location_city3097 + 2.40590668375856 * user_location_city3114 + 1.41593227937248 * user_location_city3116 + 1.63840897727578 * user_location_city3132 + 1.53497881124773 * user_location_city3143 + 1.90738642508415 * user_location_city3166 + 0.136534080085431 * user_location_city3179 + 1.70499387252118 * user_location_city3244 + 1.39626266878445 * user_location_city3261 + 0.633666017607251 * user_location_city3269 + 0.291362620342793 * user_location_city3272 + 0.816417504131623 * user_location_city3284 + 1.32262598063512 * user_location_city3295 + 1.42370129699512 * user_location_city3301 + 0.98419094292816 * user_location_city3323 + 0.033826970645109 * user_location_city3339 + 1.44054848086501 * user_location_city3347 + 0.136698147263468 * user_location_city3349 + 0.607483822279127 * user_location_city3417 + 0.589289651161736 * user_location_city3423 + 1.8570528105965 * user_location_city3444

+ 1.59536984809037 * user_location_city3465 + 1.96752427660863 * user_location_city3477 + 1.67311340995418 * user_location_city3521 + 1.30115005436785 * user_location_city3523 + 0.0345032723031656 * user_location_city3527 + 1.903469155243 * user_location_city3565 + 1.47092109786354 * user_location_city3571 + 1.53497619488896 * user_location_city3592 + 1.46842906682882 * user_location_city3618 + 1.90106766605376 * user_location_city3631 + 1.53497618909107 * user_location_city3640 + 1.36520513604405 * user_location_city3658 + 1.47521094304858 * user_location_city3675 + 1.44195573618911 * user_location_city3680 + 2.42565685590999 * user_location_city3706 + 1.91813862366892 * user_location_city3721 + 1.61091072984783 * user_location_city3735 + 1.4603150668246 * user_location_city3767 + 0.420769294985221 * user_location_city3778 + 1.47092078669489 * user_location_city3795 + 2.56014162544353 * user_location_city3810 + 1.52603593853385 * user_location_city3826 + 0.0407796478970972 * user_location_city3835 + 1.6139808775342 * user_location_city3836 + 1.30114970925876 * user_location_city3868 + 1.42369980888356 * user_location_city3917 + 2.47712664823362 * user_location_city3945 + 1.70129168082216 * user_location_city3970 + 1.44520582988759 * user_location_city3978 + 0.0376645629145996 * user_location_city4001 + 1.69750535019174 * user_location_city4003 + 0.865496003488552 * user_location_city4048 + 1.40497338048667 * user_location_city4057 + 2.37745671716887 * user_location_city4081 + 1.57308983045378 * user_location_city4088 + 1.99809357128654 * user_location_city4097 + 1.37635508185944 * user_location_city4116 + 0.742650182802601 * user_location_city4149 + 1.56437605167919 * user_location_city4151 + 1.63723653095967 * user_location_city4183 + 1.62428689756246 * user_location_city4231 + 0.544013496634376 * user_location_city4285 + 1.4205142708612 * user_location_city4290 + 1.44027445616191 * user_location_city4295 + 2.55541530964901 * user_location_city4313 + -0.0318264177362581 * is_package1 + -0.448261666576075 * channel262 + -0.171683756628001 * channel324 + 0.140754753599276 * channel386 + -0.120724797287833 * channel479 + -0.0077713399951557 * srch_children_cnt + 0.570918745126533 * hotel_country2 + 2.3720265789286 * hotel_country17 + 0.46042139651852 * hotel_country23 + 1.22444921059706 * hotel_country39 + 0.110133301416886 * hotel_country135 + 0.135671015399629 * prop_starrating + -0.0640694520679812 * distance_band1 + -0.016576192445178 * distance_band2 + 0.027438314403372 * hist_price_band2 + -0.272422482034485 * popularity_band2 + 0.126954456011052 * popularity_band4 + -0.865817642323192 * month2 + -0.00907385931340032 * month8 + -0.148443004226004 * NA'

4.4 Predict and Evaluate

```
In [191]: logit.pred1.test <- predict(logit.best1,type='response',newdata=test[,-c(1,2,11)])
          min(logit.pred1.test)
          max(logit.pred1.test)
          unique(logit.pred1.test)
```

2.22044604925031e-16
0.403230212677895
1. 0.15371960423574 2. 0.196080876419912 3. 0.0587912600381845 4. 0.00125421510926031 5. 0.132423608432973 6. 0.0396861816807097 7. 0.0160537099137261 8. 0.0847520267639034 9. 0.0123355944721956 10. 0.0159211931161573 11. 0.048252591387093 12. 0.0524276569216523 13. 0.000978789203763779 14. 0.00353980290572955 15. 0.0922842074125851 16. 0.0100557268283705 17. 0.169473676177605 18. 0.184092604160856 19. 0.168159676901507 20. 0.215992760965985 21. 0.172603927891027 22. 0.132991969035463 23. 0.155312932566763 24. 0.000914429853907556 25. 0.121011833859889 26. 0.222547281985495 27. 0.0752964681038247

28. 0.209991489152388 29. 0.174190531281142 30. 0.0806490871639141 31. 0.171718482703863 32. 0.0972782012000671 33. 0.000585149792140442 34. 0.0933351670952805 35. 0.0307946419270644 36. 0.24602356793587 37. 0.167934865309499 38. 0.10581035802915 39. 0.00447868578787114 40. 0.0993502063858279 41. 0.113590164992191 42. 0.0813711026139276 43. 0.115975921696321 44. 0.00110292615075012 45. 0.0486543398493723 46. 0.0187029513896591 47. 0.0928758667997458 48. 0.168967517629418 49. 0.233269750531505 50. 0.0364007490517683 51. 0.080577681977002 52. 0.0854762450495025 53. 0.0845452409389507 54. 0.0860063595414643 55. 0.17884076406466 56. 0.0797517924713882 57. 0.113392775481006 58. 0.00933253231009517 59. 0.0532216702113453 60. 0.131425320548771 61. 0.0931619620401907 62. 0.0012107043792003 63. 0.00735543580904943 64. 0.000746142708521499 65. 0.0677873677508406 66. 0.141108409597164 67. 0.090492385236277 68. 0.170397210365865 69. 0.0399995702813921 70. 0.119235227846469 71. 0.169427326139594 72. 0.153173257616347 73. 0.10121019908798 74. 0.122693370940037 75. 0.0111569476371399 76. 0.0608284545859356 77. 0.00466334005690663 78. 0.0907877068515882 79. 0.0688095570008723 80. 0.194102632080644 81. 0.131799889599798 82. 0.0286470889609362 83. 0.00742202118598025 84. 0.0382044027109237 85. 0.14698478474695 86. 0.104422191258428 87. 0.101205815239725 88. 0.0380734271138187 89. 0.0584244627902995 90. 0.170807390636845 91. 0.0871202274866631 92. 0.0112132457596484 93. 0.131689238835481 94. 0.205101111285541 95. 0.0119893665744137 96. 0.0832486188146488 97. 0.214172369601674 98. 0.007297380728874 99. 0.151038486154062 100. 0.0682122211024601 101. 0.0415641722590144 102. 0.0635772705739719 103. 0.0466977851313856 104. 0.171477123331847 105. 5.54016911726286e-06 106. 0.0650871881255389 107. 0.221336970182166 108. 0.257320549851923 109. 0.00984437583652095 110. 0.0212535864841841 111. 0.0360764738462893 112. 0.00316459408819306 113. 0.00787364575925141 114. 0.0597456346766667 115. 0.108385921595767 116. 0.0545406575537666 117. 0.00813181437443863 118. 0.169271217518385 119. 0.0175420642916249 120. 0.174265945998708 121. 0.0559856636664471 122. 0.15778075892631 123. 0.138595764042007 124. 0.0445022333979972 125. 0.0540904172409825 126. 0.021981594997313 127. 0.014152489985729 128. 0.00980021976518355 129. 0.0651825456606913 130. 0.0014090684682003 131. 0.00331594166351338 132. 0.0779030439844852 133. 0.214724852089376 134. 0.111504646162052 135. 0.0956228750435141 136. 0.0927228993750739 137. 0.110192609831289 138. 0.0919752464407558 139. 0.00315210176352689 140. 0.0791245623169048 141. 0.207293385792985 142. 0.00661203435166702 143. 0.133983238213508 144. 0.0729847124737043 145. 0.198909762685845 146. 0.107320380975232 147. 0.128555093826857 148. 0.230503368267161 149. 0.209009949459428 150. 0.013072731246722 151. 0.187858417697366 152. 0.210665474853066 153. 0.0239484327596495 154. 0.0581063732626024 155. 0.1108267141831 156. 0.181936399664246 157. 0.084219924324744 158. 0.124722814098704 159. 0.0265107969666664 160. 0.0804260513419043 161. 0.00384011345816582 162. 0.0897087247977339 163. 0.0996413329102617 164. 0.232135868319775 165. 0.122647296074489 166. 0.0836895581597679 167. 6.96001112225765e-05 168. 0.0509400427077214 169. 0.047674162584635 170. 0.0750704178494485 171. 0.0740603359059777 172. 0.136772903460507 173. 0.129524201034465 174. 0.0349970777947528 175. 0.274958143621156 176. 0.0337975286926561 177. 0.143505492523669 178. 0.00572540709346485 179. 0.0331470720905459 180. 0.23748524112536 181. 0.145497388135788 182. 0.179332207828645 183. 0.0881742561864182 184. 3.38419428598802e-05 185. 0.084108883030288 186. 0.206651750261052 187. 0.0124106523426121 188. 0.0406540962763935 189. 0.012875487374671 190. 0.0037576270127154 191. 0.0284807088693756 192. 0.0230650064127305 193. 0.0235052620231478 194. 0.0785951292335289 195. 0.116368691726379 196. 0.00501362912129866 197. 0.116696443509907 198. 0.200469082960803 199. 0.146098944949433 200. 0.0451181257503151 201. 0.160784520521889 202. 0.096423390446227 203. 0.118911661458776 204. 0.000652366863043845

205. 0.0241685291133283 206. 0.100092919688574 207. 0.0158045645333918 208. 0.0841543245026593
209. 0.127249331910682 210. 0.220817512298232 211. 0.223477713870624 212. 0.103152358943157
213. 0.109067831941049 214. 0.0719188846297006 215. 0.0703216701664747 216. 0.125875957820417
217. 0.0182970719293007 218. 0.15193940118345 219. 0.014379662838432 220. 0.149911347102624
221. 0.256465145003667 222. 0.049497261005288 223. 0.0551398136499964 224. 0.0128032287303341
225. 0.192540572252555 226. 2.2548639737355e-06 227. 0.0646931458646559 228. 0.066784373834049
229. 0.115898507332731 230. 0.203854055235434 231. 0.0698090290515274 232. 0.19097677865253
233. 0.144592792249881 234. 0.150891729400015 235. 0.111750375858517 236. 0.140154593700496
237. 0.200710173435739 238. 0.077534366630925 239. 0.0706900889176725 240. 0.0701973405419703
241. 0.120683735261762 242. 0.0222012981857633 243. 0.013057527975489 244. 0.148452230806844
245. 0.0849553481068594 246. 0.0129383719961661 247. 0.169947168381451 248. 0.166098583580436
249. 0.158095470735296 250. 0.000230637205943885 251. 0.153595962594647
252. 0.035668950641861 253. 0.203087689329629 254. 0.191009532505445 255. 0.0771556754352243
256. 0.173633252325888 257. 0.112107562596121 258. 0.0367782836951366 259. 0.105206569279647
260. 0.144751597038378 261. 0.116485455950417 262. 0.0883983323300168 263. 0.0114866421953436
264. 0.00667011221537392 265. 0.0740362913542758 266. 0.00309085385943102
267. 0.0228510785146548 268. 0.113216925767586 269. 0.164119913634482 270. 0.0838405716505545
271. 0.177723148046036 272. 0.0636046653635321 273. 0.00396983299069414
274. 0.119124904784272 275. 0.0735045420050228 276. 0.00506713474601397
277. 0.130308024343726 278. 0.117619391712695 279. 0.00366551738696374 280. 0.012456784663556
281. 0.18504189262897 282. 0.184514457459003 283. 0.224073401463878 284. 0.046141218184639
285. 0.0373642382107676 286. 0.0112875747333549 287. 0.0535759477266519 288. 0.132577799057578
289. 0.0110211550909502 290. 0.0784297135137765 291. 0.0118339554565536
292. 0.00129208607319486 293. 0.139312077104084 294. 0.0194709051280914
295. 2.22044604925031e-16 296. 0.0105627673744813 297. 0.00835366485468225
298. 0.111953873641602 299. 0.103013175651467 300. 8.55099399579762e-12 301. 0.209390294958368
302. 0.190151878193501 303. 0.0161730805240864 304. 0.0999832760056608 305. 0.150740826952799
306. 0.198787174862336 307. 0.137155508731832 308. 0.112730986472569 309. 0.00239951835118765
310. 0.142230684197019 311. 0.169664861576105 312. 0.403230212677895 313. 3.19385702892313e-06
314. 0.0987108180753063 315. 0.0332838930831589 316. 0.116660712332557 317. 0.209501598912553
318. 0.0932921267519042 319. 0.124851869840113 320. 0.0998403756536528 321. 0.0940354115295324
322. 0.167765490700405 323. 0.0131936892838038 324. 0.171918831572113 325. 0.00150021343689652
326. 0.0669308549253867 327. 0.0896668268083468 328. 0.016370407342697 329. 0.258183055748139
330. 0.11916210239897 331. 0.00754538913833324 332. 0.101295773390849 333. 0.168746347303095
334. 0.0768745921727109 335. 0.34655056268858 336. 0.0484126836571637 337. 0.148707538875631
338. 0.000619112327950257 339. 0.194724936889914 340. 0.158824270263842 341. 0.156204772611291
342. 0.0277833439762578 343. 0.134620699038687 344. 0.107425165309816 345. 0.15403185362022
346. 0.0397907692346367 347. 0.104127500869404 348. 0.166943243407156 349. 0.194608089649545
350. 0.0166011737038407 351. 0.0841987245087531 352. 0.125492612060693 353. 0.193844580045803
354. 0.0642843109150916 355. 0.158219089613224 356. 0.0702302035987513 357. 0.135525951778934
358. 0.105214161065003 359. 0.0953386141054226 360. 0.150738710367833 361. 0.0316233342566758
362. 0.0457224159831273 363. 0.141382245434735 364. 0.00968796469417692
365. 0.0683749194977891 366. 0.050815524058144 367. 0.0558028087328515 368. 0.162610471488566
369. 0.0676870799704036 370. 0.198791915451122 371. 0.131870683119433 372. 0.0724704757778025
373. 0.0131764556552274 374. 0.00500301203862784 375. 0.0775659056528988 376. 0.11215362255976
377. 0.163441196968319 378. 0.269443569878569 379. 0.000660519372596781 380. 0.048119519044088
381. 0.0896771843486086 382. 0.0996191428120144 383. 0.0898107489984446 384. 0.102043365048618
385. 0.0658057442128303 386. 0.137586176952265 387. 0.0189372533098129 388. 0.138402167965278

389. 0.00240057772161755 390. 0.000475086829976257 391. 0.0451095545240999
392. 0.0275772315013723 393. 0.215091345063786 394. 0.139402817421036 395. 0.215101495630428
396. 0.0953030713152514 397. 0.209963226398932 398. 0.0621634387448944 399. 0.0807103269863647
400. 0.00770970931142549 401. 0.0407159919219521 402. 0.0730731007080289
403. 0.0988859591139116 404. 0.138941387265406 405. 0.135991687543543 406. 0.0117796311387097
407. 0.0801362795065813 408. 0.0220634623221032 409. 0.180407684493973 410. 0.0932548410685777
411. 0.0163171748287729 412. 0.012659243260282 413. 0.0828878244315267 414. 0.176120675157427
415. 0.0212746361916977 416. 0.129254511843987 417. 0.0795912061167902 418. 0.0676759888463939
419. 0.144835275658074 420. 0.00334156557452081 421. 0.0125915548961509 422. 0.100924753945494
423. 0.147473779231765 424. 0.0329366253510503 425. 0.154630803095305 426. 0.14846533998523
427. 0.160862608199382 428. 0.120872315180393 429. 0.00795450299697345 430. 0.12475097245239
431. 0.102301644241983 432. 0.0882331084306856 433. 0.0945243677230138 434. 0.137327748615396
435. 0.083032641177091 436. 0.158400895758571 437. 0.0694173384364956 438. 0.221481850429864
439. 0.075716092214268 440. 0.232054886010421 441. 0.0942268822153089 442. 0.0634688077942655
443. 0.0552395941583538 444. 0.0862508939500064 445. 0.0741812376022841 446. 0.171342297997518
447. 0.134949475610662 448. 0.0941758306097133 449. 0.242798864590013 450. 0.00399502328717075
451. 0.264915168065716 452. 0.102461680230877 453. 0.111229398055221 454. 0.000584393333818044
455. 0.137522666494995 456. 0.139369023961696 457. 3.91897203620831e-05 458. 0.192710692007951
459. 0.10568986493986 460. 0.104744303471957 461. 0.00495381401974821 462. 0.126812843298794
463. 0.00150705941630592 464. 0.114546389748839 465. 0.0789075338313958 466. 0.135603466787308
467. 0.000808051527128781 468. 0.164607770293291 469. 0.163201820446602
470. 0.0604161514086816 471. 0.00446330394370842 472. 0.109702457492259
473. 0.00544830466988421 474. 2.62073973183212e-06 475. 0.0374485772102404
476. 0.121778664880997 477. 0.159898030976839 478. 0.120322875497025 479. 0.103431375716528
480. 0.126557651017091 481. 0.148743567675392 482. 0.0803420518148978 483. 0.154298971258511
484. 0.154825667152561 485. 0.00663640337081087 486. 0.107425221722351 487. 0.0272629516283038
488. 0.196295393048131 489. 0.134670838742043 490. 0.11328901275031 491. 0.0361929075413587
492. 0.0185915966931782 493. 0.152576850771354 494. 0.140778472121283 495. 0.141196220432619
496. 0.232797957092721 497. 0.139108769297723 498. 0.0147523145554348 499. 0.0730118852451216
500. 0.124016584551343 501. 0.147308815231387 502. 0.104031406745934 503. 0.0700876899307217
504. 0.120944048117089 505. 0.160791061200455 506. 0.175140782467768 507. 0.123163732291743
508. 0.0548666111835705 509. 0.06264270505361 510. 0.0608021327587847 511. 0.171838713226303
512. 0.00343855398157177 513. 0.0894187716655077 514. 0.137583424400112
515. 0.000867924242638353 516. 0.146089470116499 517. 0.0605456932454604
518. 0.000217575320865782 519. 0.120965379972454 520. 0.0797277942799724
521. 0.069498492789436 522. 0.0348735287990418 523. 0.00992432118482336 524. 0.203097413368473
525. 0.145986730659754 526. 0.0622637623262042 527. 0.0683618215748599 528. 0.161968744124928
529. 0.0537028901553071 530. 0.072683179662585 531. 0.0451940543685547 532. 0.114896864784404
533. 0.130766277354244 534. 0.0491376596136096 535. 0.144979355722246 536. 0.184400644869165
537. 0.00041677412694907 538. 0.0408358692948073 539. 0.171906984663888 540. 0.108660293935841
541. 0.0115681694437582 542. 0.00999820411240574 543. 0.0145688616462109
544. 0.00107903319444799 545. 0.0846095703079682 546. 0.111342637723597
547. 0.0255122894502753 548. 0.0440996760663662 549. 0.214719833373436 550. 0.0639593503186678
551. 0.120263891744502 552. 0.147803081698953 553. 0.291634962869537 554. 0.126156550917987
555. 0.0118375862276798 556. 0.000884149782151814 557. 0.126122123002955
558. 0.100427239865812 559. 0.00145450262711761 560. 0.00589978631309946
561. 0.0898849380800375 562. 0.231867980157037 563. 0.0370180412733952 564. 0.0845768007447417
565. 0.0454481822711293 566. 0.0482858026919696 567. 0.0955822661961831 568. 0.216845850534697

569. 0.0744163450539446 570. 0.122259922629727 571. 0.112358306913838 572. 0.0887699859289368
573. 0.11333912290504 574. 0.113702894426106 575. 0.150595186126403 576. 0.129065905865417
577. 0.163376572018941 578. 0.0700203207864115 579. 0.118348756974583 580. 0.00502869569408864
581. 0.0136156000160011 582. 0.000193044695376205 583. 0.137622495137173
584. 0.012453955843793 585. 0.126148321906285 586. 0.046760023686477 587. 0.207579852538758
588. 0.123852918642953 589. 0.0621784765671752 590. 0.000582170270173388
591. 0.0158836687885279 592. 0.0101883531166305 593. 0.059323617423551
594. 0.105169518895304 595. 0.108919694824887 596. 0.0280660456675617 597. 0.0367041527080244
598. 3.04284894646984e-05 599. 1.6345406096366e-07 600. 2.99669937225509e-07
601. 0.0288616200258548 602. 0.161239333325993 603. 0.0932946226823287
604. 0.135418689401993 605. 0.0938394166005416 606. 0.0877518137694223 607. 0.169884984280277
608. 0.00658577382855624 609. 0.0152955955917815 610. 0.0931276394484805
611. 0.00313094906600631 612. 0.176299504580699 613. 0.0390431087527158 614. 0.205507008326116
615. 0.111681947957191 616. 0.0139226385003732 617. 0.0777431729371994 618. 0.13731596279386
619. 0.138058213150715 620. 0.11299442786886 621. 0.0840944823599567 622. 0.317959030400742
623. 0.0528013995443031 624. 0.0535680814598413 625. 0.144735481956051 626. 0.0327872231355097
627. 0.0948193474818056 628. 0.156443822021703 629. 0.031316832012206 630. 0.1613264450928
631. 0.036966632845203 632. 0.0127426254676949 633. 0.132948905890948 634. 0.271285762317356
635. 0.1726435807189 636. 0.000423266220604557 637. 0.136085002153692 638. 0.0520189906627012
639. 0.0114734384282404 640. 0.0958271364159731 641. 0.0627569760938794
642. 0.144026136845766 643. 0.000515068187724997 644. 0.10249128338864 645. 0.118666546689208
646. 0.0513069009287884 647. 0.0953299459433768 648. 3.56533761570731e-05
649. 0.020992780102919 650. 0.0306108576891061 651. 0.0818310924017707 652. 0.0382798371712531
653. 2.09177040356871e-05 654. 0.013876695727207 655. 0.00357232469512238
656. 0.147826726293402 657. 0.0802211407952965 658. 0.151657740719338 659. 0.0760432268554703
660. 0.0956261695842948 661. 0.0652776740645069 662. 3.35188624853819e-06
663. 0.191736799938421 664. 3.69804107826171e-07 665. 0.140131262415961
666. 0.0130783637907994 667. 0.113771826512286 668. 0.104129896433863 669. 0.171505728913722
670. 0.031841218665645 671. 0.03749066240277 672. 0.00627058763398354 673. 0.0879412975519536
674. 0.181886816943717 675. 0.081614409833868 676. 0.169488210602957 677. 0.0620024706497156
678. 2.2607429137993e-05 679. 0.181141804901116 680. 0.00682446462470257
681. 0.0965495718975011 682. 0.000401351776698035 683. 0.043451579456922
684. 0.0190799064577498 685. 0.235301134647014 686. 8.27615006642944e-05
687. 0.0502022696843928 688. 0.0847804472904651 689. 0.0128761252856115
690. 0.0984468044714703 691. 0.143777280168966 692. 0.10470419567787 693. 0.0378289247500545
694. 0.0119151133668571 695. 0.213458745843951 696. 0.157740748216457 697. 0.0113337053203791
698. 0.112332284291835 699. 0.00460873242731196 700. 0.122378325203197 701. 0.0566363028775853
702. 0.109991021779487 703. 0.219695343608889 704. 0.0350212508322018 705. 0.146030841355653
706. 0.195195951739893 707. 0.0732703174385258 708. 0.0524951304360902 709. 0.146829751785935
710. 0.128946203039365 711. 0.0581640042180195 712. 0.0596445846487258 713. 0.141965812080031
714. 0.0187490657756497 715. 0.0657604888472297 716. 3.41586148737104e-07
717. 0.107910125906839 718. 0.0803118662116847 719. 0.0123663206412478 720. 0.174047482238125
721. 0.166630590316408 722. 0.0931449543964386 723. 0.307490920253331 724. 0.154436252596836
725. 0.151018863872652 726. 0.143919248792373 727. 0.235221813922565 728. 0.205038517764891
729. 0.0482045220477544 730. 0.0276050744718808 731. 0.00558210340911453
732. 0.0762803226571412 733. 0.00033393011626967 734. 0.12652545253898 735. 0.0386708282943595
736. 7.35010631935433e-05 737. 0.0110716734842926 738. 0.114050779040177
739. 0.0139249799866787 740. 0.0216083819874609 741. 0.0540434510095264

742. 0.00411955368499832 743. 0.108524155189301 744. 0.101572437855838 745. 0.104293813310195
746. 0.0746481538352592 747. 0.0132119559934733 748. 0.131769169745335 749. 0.222878591912625
750. 0.00319659171449906 751. 0.194742121718993 752. 0.000431830602295682
753. 0.12548996362464 754. 0.135191155439645 755. 0.0949780924713149 756. 0.0707205440921681
757. 0.0891386333754999 758. 0.000859622007889783 759. 0.0136268716238092
760. 0.154384808324332 761. 0.040134654980931 762. 0.0108394732586442 763. 0.215579892150332
764. 0.156155348937379 765. 0.00962095164018971 766. 0.138214303435275 767. 0.18042077289461
768. 0.129740868605026 769. 0.186618744751259 770. 0.207630198545236 771. 0.000525143525619197
772. 0.101529727886117 773. 0.117597710973106 774. 0.14764492226636 775. 0.0898071535513632
776. 0.106974011693704 777. 0.0825166852119264 778. 0.00845585599532342 779. 0.189569017198322
780. 0.278183815781464 781. 0.00109834963738545 782. 0.0828342469686925
783. 0.101325119056926 784. 0.00353461756338652 785. 0.0213172083946244
786. 0.129306142256312 787. 0.0969111615292079 788. 0.00785859459580952
789. 0.0880056403339091 790. 0.146395442365986 791. 0.0999836695710014
792. 0.0178397589957737 793. 4.84621505211521e-05 794. 0.104501911367999
795. 0.0193976246593672 796. 0.0950664601449447 797. 0.0178021503677852 798. 0.246080342484797
799. 0.319236377040041 800. 0.0166498884105118 801. 0.117841335113694 802. 0.169559465863566
803. 0.00769190386345843 804. 0.0979096940799091 805. 0.000849595229885104
806. 0.0507022461391556 807. 0.0492588643418968 808. 0.119100945313765 809. 0.0460391933389986
810. 0.00983731438322747 811. 0.00690832304303177 812. 0.00735980164918437
813. 0.180470748567646 814. 0.000540238518871071 815. 0.0970526569452783
816. 0.262843507909009 817. 0.00983670589175138 818. 0.000597065392545441
819. 0.00441508501638957 820. 0.166019962737128 821. 0.0414959400827781
822. 0.0576966413859578 823. 0.155152347754089 824. 0.0187867898146209 825. 0.0775026547791743
826. 0.00626089953836683 827. 0.0359291349707541 828. 0.00746851648043018
829. 0.0324848871290005 830. 0.0244402794516164 831. 0.109559770905299 832. 0.0981902868585728
833. 0.149080386227762 834. 0.00814089284984939 835. 0.0484773238817139
836. 0.076390289135874 837. 0.000299572430146355 838. 0.161202926577248
839. 0.147709930405067 840. 0.158350587202048 841. 0.0187328626808724 842. 0.15510525364956
843. 0.0407497193561461 844. 0.126649122028601 845. 0.11125894568079 846. 0.000940033677884484
847. 0.0643738061740198 848. 0.125080462594239 849. 0.0184233928018674 850. 0.139383860671483
851. 0.026960888133076 852. 0.134035248703725 853. 0.0479874367748887 854. 0.111600869422401
855. 0.117832159225247 856. 0.209101989034781 857. 0.180295844998151 858. 0.146306837686015
859. 0.0308522618789272 860. 0.135666002691358 861. 0.217218427429423 862. 0.00854145908844683
863. 0.0492410823337223 864. 0.0367633722947952 865. 0.119664050824714 866. 0.128374692326677
867. 0.0988458293484911 868. 0.0506799746975039 869. 0.00885714561180809
870. 0.1001677622868 871. 0.173564754979024 872. 0.105397950030829 873. 0.093197813495189
874. 0.0671780409721019 875. 0.0597878482625573 876. 0.177175238183398 877. 0.011380155993094
878. 0.135598311206904 879. 0.195662193913572 880. 0.136513037428043 881. 0.0983420701781894
882. 0.03903381580886 883. 5.60062592493962e-05 884. 0.0567887257315536 885. 0.231565713576204
886. 0.000558777661070268 887. 0.00409026058901121 888. 0.111486621759412
889. 0.0205366528012435 890. 3.83101467287716e-07 891. 0.0137619890161097
892. 0.0786483249254133 893. 0.197731073005871 894. 0.0293249687083814 895. 0.0189728766728144
896. 0.093604671850797 897. 0.333260180315717 898. 0.183297930398864 899. 0.00764621154168975
900. 0.0259178984809143 901. 0.113121739366405 902. 0.183111716809115 903. 0.113588803418989
904. 0.00260162657313693 905. 0.103668066697539 906. 0.0282893524764251
907. 7.0045891872691e-05 908. 0.00269700511976082 909. 0.122501350839705
910. 0.17449371947683 911. 0.0324197975981447 912. 0.0193431289042107 913. 0.138363814784922

914. 0.0129568905788028 915. 0.0951301181017211 916. 0.0691047408391952
917. 0.000353623006841658 918. 0.105210269313933 919. 0.145450302378526 920. 0.186306832499302
921. 0.128622239903716 922. 0.0186343175339688 923. 0.136489965746634 924. 0.0413611848438707
925. 0.0179440992503659 926. 0.108578294992288 927. 0.202839645617638 928. 0.0900086601373457
929. 0.000506134521771633 930. 0.145321729001653 931. 0.152534560693994 932. 0.151216221536558
933. 0.00971576520441499 934. 0.185022757548804 935. 0.210810250929351 936. 0.111177609981674
937. 0.0386836344106913 938. 0.0827152757933178 939. 0.165130060059649 940. 0.090312555892125
941. 0.143471982550686 942. 0.111639226430733 943. 0.103703580039063 944. 0.000931589499321204
945. 0.179999873653442 946. 0.00970058343983165 947. 0.00655834858510683
948. 0.000237262920410365 949. 0.147593630044051 950. 0.0663404366880929
951. 0.129504373695659 952. 0.240011035302595 953. 0.149198453039297 954. 0.000142011765894594
955. 0.0932645972552192 956. 0.232647746797484 957. 0.000709637115755053
958. 0.0728067913193129 959. 1.673958392502e-05 960. 0.0754880483679393
961. 0.0659987687218527 962. 0.00713327161252501 963. 0.0604635546571655
964. 0.090346291232022 965. 0.124239294894437 966. 0.1141523560826 967. 0.00116869826927277
968. 0.172951090349792 969. 0.0312067682356253 970. 0.0194350011866108 971. 0.0485779675514124
972. 0.104755862728877 973. 0.00110856405374869 974. 0.0387958876442603
975. 0.0387460246675653 976. 0.0115760153218052 977. 0.0041273806482543 978. 0.144423794594224
979. 0.0271941836952932 980. 0.0273192903450221 981. 0.0607104274913421 982. 0.112770070550874
983. 0.000794216480746684 984. 0.103731361795981 985. 3.5809200215743e-06
986. 0.203031952461834 987. 0.103826899265684 988. 0.092280178351657 989. 0.0207086379683724
990. 0.101935608582606 991. 0.0170766152694551 992. 0.0252469721825242
993. 0.0967410892411956 994. 0.000214061831465556 995. 0.131068334349283
996. 0.06412588815012 997. 0.00808294869982698 998. 0.100177508034132 999. 0.116151464119848
1000. 0.0579944127113158 1001. 0.152715612643945 1002. 0.171109994302331
1003. 0.0682226719340101 1004. 0.225857436285342 1005. 0.0085204600863572
1006. 0.174976820787678 1007. 0.0882793672162389 1008. 0.145369027034485
1009. 0.120700625258698 1010. 0.00970798688053854 1011. 0.135843758188297
1012. 0.158168829588557 1013. 6.04866286537126e-05 1014. 0.0753285170662597
1015. 0.0614046600244929 1016. 0.128750621629162 1017. 0.0803651988151146
1018. 0.141922998284611 1019. 0.0129355348836614 1020. 0.010011203609217
1021. 0.0607341423863667 1022. 0.207907633545223 1023. 0.057946240603448
1024. 0.0518651630964553 1025. 0.158880861396577 1026. 0.031186962550685
1027. 0.237374036500508 1028. 0.225890671030508 1029. 0.00125324873249376
1030. 0.172126026630748 1031. 0.00279009007369945 1032. 0.0147687952884662
1033. 0.0611302793274277 1034. 0.00352120468398604 1035. 0.024996842408436
1036. 0.0961576042341142 1037. 0.0540677320413217 1038. 0.220131376878519
1039. 0.155393825821915 1040. 0.0279261275834619 1041. 0.177468363335265
1042. 0.000164872161605504 1043. 0.00777232649253145 1044. 0.0478892394988652
1045. 0.0287321934140445 1046. 0.179415250560754 1047. 0.217783189584735
1048. 0.0873499704309129 1049. 0.190543833215334 1050. 0.0657917570567914
1051. 0.14775654510513 1052. 0.139024242506272 1053. 8.52076544081975e-05
1054. 0.129076974006622 1055. 0.0921333888132815 1056. 0.146311628587494
1057. 0.114531840576043 1058. 0.0784270960141532 1059. 0.00558747491912923
1060. 0.16567181563592 1061. 3.98769807696929e-05 1062. 3.48851188929437e-06
1063. 0.127392838581308 1064. 0.15977014836744 1065. 0.0428780110849986
1066. 0.199125738126402 1067. 0.0489186879931324 1068. 0.0380604529713949
1069. 0.164285511940823 1070. 0.0676778654236718 1071. 0.17418392593138

1072.	0.0314876882313261	1073.	0.120393074919511	1074.	0.0413773892997302
1075.	0.0985004892971873	1076.	0.0978097278497872	1077.	0.180566764470675
1078.	0.0176910419494368	1079.	0.100279229495175	1080.	0.139201195472355
1081.	0.0449911144643921	1082.	0.130909338298615	1083.	0.0657335660658294
1084.	0.0664338652171066	1085.	0.153733482304547	1086.	0.178704796239889
1087.	0.159282611936712	1088.	0.0383996111044194	1089.	0.00211404662120566
1090.	0.12283755860781	1091.	0.229921038989328	1092.	0.00535862781937202
1093.	0.0737525409306605	1094.	0.10136650137223	1095.	0.0505544593841864
1096.	0.012385958431667	1097.	0.130388013572328	1098.	0.072528676431863
1099.	0.0117750181578316	1100.	0.200143529585166	1101.	0.0089695518337932
1102.	0.0607790699829764	1103.	0.0124699550913088	1104.	0.0243450275196819
1105.	0.114183367015325	1106.	0.0686808059180203	1107.	0.197482761325682
1108.	0.102497318236814	1109.	0.0857173728404029	1110.	0.148304258851873
1111.	0.224363182119184	1112.	0.014366766846482	1113.	0.188360866556142
1114.	0.209440964009147	1115.	0.120727033102703	1116.	0.131561463273581
1117.	0.164225047100783	1118.	0.173491306389761	1119.	0.0505227525392927
1120.	0.239911730908877	1121.	0.0112391448241914	1122.	0.0727751091757173
1123.	0.159762886475214	1124.	0.184556398050989	1125.	0.00800997431166877
1126.	0.000836858721967026	1127.	0.000139999570268056	1128.	0.123519565685412
1129.	0.0287163418830435	1130.	0.203849224643502	1131.	0.0377543036753963
1132.	0.107798550918935	1133.	0.106177420650346	1134.	0.176258524577216
1135.	0.00899203957567543	1136.	0.262253763733639	1137.	0.0284445051499776
1138.	8.34828728823783e-09	1139.	0.111988415286547	1140.	0.131964730017611
1141.	0.0742025055754725	1142.	0.154772294463865	1143.	0.13551453054529
1144.	0.10657813909867				
1145.	0.0981962673017042	1146.	0.00124668856469594	1147.	0.0357357992648928
1148.	0.0382040300803104	1149.	0.150035114220961	1150.	0.0463979571568682
1151.	0.0774287802473942	1152.	0.0400883213806618	1153.	0.0954630442822927
1154.	0.0178909162300876	1155.	0.0494585338003145	1156.	0.138769316097239
1157.	0.148607439548173	1158.	0.125030266953183	1159.	0.183315481201908
1160.	0.0751907994472322	1161.	0.00472194470963218	1162.	0.0241670809617926
1163.	0.116883059262984	1164.	0.025626181404429	1165.	0.0555593724878024
1166.	0.105679848633648	1167.	0.245252753400584	1168.	0.0150546874434395
1169.	0.0923061179867262	1170.	0.102130840949686	1171.	0.132589206501755
1172.	6.50803438840278e-05	1173.	0.169242908499025	1174.	0.0808011013494931
1175.	0.16736781245975	1176.	0.130889863910127	1177.	0.0826762485230843
1178.	0.0981315034504334	1179.	0.12862483861996	1180.	0.0112098279721235
1181.	0.102912793595758	1182.	0.128447635437371	1183.	0.0820106992853149
1184.	0.0694752711307875	1185.	0.121576365725199	1186.	0.013339981949161
1187.	0.0570777778645165	1188.	0.0986257653671166	1189.	0.0711377589927824
1190.	0.0025034537583858	1191.	0.000339286306805297	1192.	0.0927241625479984
1193.	0.172125919108315	1194.	0.051096753746434	1195.	0.0111026230370151
1196.	0.122829560347361	1197.	0.0247237786833291	1198.	0.0243108283206633
1199.	0.000512664496378473	1200.	0.0622627599492268	1201.	0.244451558542958
1202.	2.06016571384983e-07	1203.	0.0083697445172466	1204.	0.00527330264371481
1205.	0.078823159226133	1206.	0.163673690565643	1207.	0.084523467016572
1208.	0.0321297509753683	1209.	0.129923384697267	1210.	0.172918925650535
1211.	0.0572758218481614	1212.	0.0877289304114664	1213.	0.00118258772552936
1214.	0.138412169256811	1215.	0.0699141017338093	1216.	0.0514056186915155

1217.	0.0121806976788021	1218.	0.0992505979820652	1219.	0.0806413989443057
1220.	0.00622767557839998	1221.	0.113791454281177	1222.	0.0866169183353115
1223.	0.0772273440967068	1224.	0.0136195205105328	1225.	0.0408802992088813
1226.	0.0578824723384842	1227.	0.142741052503206	1228.	0.110897643537323
1229.	0.0793340392770821	1230.	0.0496513148820247	1231.	0.11279419932393
1232.	0.0691647808635185	1233.	0.000773241968398641	1234.	0.0100721784281321
1235.	0.0878080343179873	1236.	0.0231842680762728	1237.	0.0507247600476942
1238.	0.108426356724993	1239.	0.0447382118167655	1240.	0.141019063694366
1241.	0.0556580455518031	1242.	0.198743175009574	1243.	0.033347654499101
1244.	0.104431125874805	1245.	0.183848433925726	1246.	0.00342932097864332
1247.	0.016093166181939	1248.	0.193097053293724	1249.	0.0738749816109976
1250.	0.155004292002225	1251.	0.0727458436771442	1252.	0.0135836856374786
1253.	0.156834113812423	1254.	0.0567195572537112	1255.	0.0446275225676485
1256.	0.233319168414554	1257.	0.0125168183440267	1258.	0.10039417601213
1259.	0.0143271054357582	1260.	0.187976192898861	1261.	0.00417041852082415
1262.	0.0644372739738611	1263.	6.3717907820515e-05	1264.	0.0568529328086986
1265.	0.109976471401975	1266.	0.0244308923386422	1267.	0.000735076756684876
1268.	0.15680278529401	1269.	0.0175254854351839	1270.	0.0886007471574247
1271.	0.0528634048478134	1272.	0.0969831207791257	1273.	0.0155224416126103
1274.	0.0346368010059763	1275.	0.00502998807205259	1276.	0.0951286422950823
1277.	0.000302275856846815	1278.	0.222574663446053	1279.	0.207718129999074
1280.	0.103313652286422	1281.	0.0097255469953812	1282.	0.0246298731586631
1283.	0.151979439001627	1284.	0.0122102584470636	1285.	0.0544501451104681
1286.	0.0711596609329448	1287.	0.05342972890262	1288.	0.112603944402393
1289.	0.30421228502243				
1290.	0.0679994094658789	1291.	0.063631158722785	1292.	0.167589886328703
1293.	2.38059820187338e-05	1294.	0.0720419628474859	1295.	0.0376872960201066
1296.	0.159748234326948	1297.	3.53445815619286e-07	1298.	2.9767371254735e-05
1299.	0.264702322440057	1300.	6.43380457345513e-05	1301.	0.0373457193418742
1302.	0.0943707020440156	1303.	0.148651734609625	1304.	0.0380866456754133
1305.	0.0951466440235585	1306.	0.0184939248504688	1307.	0.116168896142626
1308.	0.0182492001040681	1309.	0.0902473574631276	1310.	0.000545310022802138
1311.	0.0873115921986897	1312.	0.0115235908016607	1313.	0.157198149398926
1314.	0.0659405091508552	1315.	0.0475388437336818	1316.	0.207723028351601
1317.	0.0162617881079926	1318.	0.00986908085194112	1319.	0.00860748432303347
1320.	0.137674993373802	1321.	0.0037785555496643	1322.	0.0946627640070944
1323.	0.000251636223216033	1324.	0.157350180654754	1325.	0.000100940187257189
1326.	0.0515937214524282	1327.	0.125752883095644	1328.	0.108932393901779
1329.	0.112145448992384	1330.	0.256295721505337	1331.	0.0148686385830036
1332.	0.061379255964718	1333.	0.0379779111541442	1334.	0.0168664411339473
1335.	0.0114488946834579	1336.	0.000469551982090917	1337.	0.0756747154663639
1338.	0.104343187056894	1339.	0.00553334704932668	1340.	0.000948018843528044
1341.	0.111691404082359	1342.	0.0950400841035781	1343.	0.0632656645936293
1344.	0.0657669974294955	1345.	0.0367052032264749	1346.	0.118371288118726
1347.	4.53538639812575e-05	1348.	0.170525001376082	1349.	0.168015768239547
1350.	0.0138660718404895	1351.	0.118582729733489	1352.	0.13845557992565
1353.	0.251543364513822	1354.	0.0290256370481766	1355.	0.00361189218688058
1356.	0.0324959510658707	1357.	0.0072334869029732	1358.	0.033526041082644
1359.	0.042979289977499	1360.	0.00147081266557148	1361.	0.0453822280199773

1362. 0.0833817711185703 1363. 0.239814296329196 1364. 0.0825209377882948

1365. 0.000185419085749329 1366. 0.193764156782769 1367. 0.105755626407709

1368. 8.71007138228385e-06 1369. 0.110957742781564 1370. 0.0905006910214361

1371. 0.0492479922876292 1372. 0.146826801454515 1373. 0.0129122370911208

1374. 0.139091416445804 1375. 0.00508126617242069 1376. 8.27938780098114e-10

1377. 0.000617841102357041 1378. 0.175422127078029 1379. 0.0214587371461381

1380. 0.00283568793138017 1381. 0.200276271598131 1382. 0.100126090567461

1383. 0.00841872256903765 1384. 0.00441609113830104 1385. 0.0685503525961745

1386. 0.108663572086276 1387. 0.164067953761274 1388. 0.0170743951786158

1389. 0.115527747229071 1390. 0.00458708246725263 1391. 0.11070602552228

1392. 0.0191140613678603 1393. 0.0728046125127814 1394. 0.0828310625960269

1395. 0.0111997667017609 1396. 0.149143504796104 1397. 0.00348028048501182

1398. 0.102718470243271 1399. 0.185339786507566 1400. 0.0500243565939711

1401. 0.0692226851436682 1402. 1.25675725690032e-05 1403. 0.255964730324385

1404. 0.185088049106881 1405. 0.148745376183409 1406. 0.0530051044072647

1407. 0.19270910681184 1408. 0.154021157493011 1409. 0.113488027721003 1410. 0.141808825255567

1411. 0.0892098712047958 1412. 0.0603908333088731 1413. 0.0670276658258329

1414. 0.00371444448566202 1415. 0.180095278442183 1416. 0.0129479144582469

1417. 6.15056664078445e-05 1418. 0.0888323146003592 1419. 0.0898342246300157

1420. 0.0736697925300103 1421. 0.154962659404059 1422. 0.0480318728660385

1423. 3.36268771071506e-06 1424. 0.0615516502540912 1425. 0.000932814915732095

1426. 0.0679353688625521 1427. 0.0737389442192565 1428. 0.0261701686790454

1429. 0.125071449445126 1430. 0.164351658111074 1431. 0.0163003231868838

1432. 0.0178306713815083 1433. 0.0921127545160149 1434. 0.0342969311554624

1435. 0.0432203638465231 1436. 0.0864194883703233 1437. 0.0908819100955922

1438. 0.0137627773578394 1439. 0.12262664091714 1440. 0.0124187372250347

1441. 4.093971893078e-05 1442. 0.150065685672459 1443. 0.139872222281567

1444. 0.0313360897712622 1445. 0.2127832319181 1446. 0.0050478002928799

1447. 0.0821326066549812 1448. 0.121272741665521 1449. 0.0674951175149792

1450. 0.182040606742825 1451. 0.0061863435574981 1452. 0.00465517211054286

1453. 0.000824021850065527 1454. 0.091854350168371 1455. 0.0221129661698628

1456. 0.145427060155232 1457. 0.012705241895925 1458. 0.105271436570213

1459. 0.0121894667821893 1460. 0.000726661535451343 1461. 0.00375879024293725

1462. 0.0870928318641079 1463. 0.223998137646244 1464. 0.00266762565115642

1465. 0.031061423588356 1466. 0.00542936040369784 1467. 0.149546584343158

1468. 0.0931553033390507 1469. 0.131654247382951 1470. 7.82916611143988e-09

1471. 0.0556892126629626 1472. 0.0713197237753369 1473. 0.164208245493208

1474. 0.0609000854116124 1475. 0.0283534699957809 1476. 0.100892154135499

1477. 4.44537455289705e-05 1478. 0.0765102108488191 1479. 0.0322954909499512

1480. 0.110726462441151 1481. 0.136421928029488 1482. 0.11467609037971 1483. 0.159716353162686

1484. 0.0615200636567943 1485. 0.144029642281795 1486. 0.131411740591883

1487. 0.0694991274668447 1488. 0.134351814831236 1489. 0.0559546126487537

1490. 0.181042169045108 1491. 0.00998356235832672 1492. 0.158029553389467

1493. 0.00814611265540103 1494. 0.00259183248653241 1495. 0.131218059344816

1496. 0.10568206638805 1497. 0.122756207946684 1498. 3.60527969670834e-05

1499. 0.182705427929933 1500. 0.0707718533691802 1501. 0.106856187437037

1502. 0.0446447224169152 1503. 0.147304620466172 1504. 0.197991241254983

1505. 0.0278166817439775 1506. 0.0586452002908439 1507. 0.0103925016737772

1508.	0.170980518071003	1509.	0.127071175253646	1510.	0.00134971418171135
1511.	0.00870174646979526	1512.	0.202916567838195	1513.	0.160163886860458
1514.	0.00581990884033794	1515.	0.212211852186092	1516.	0.0102404980056425
1517.	0.00141494348319055	1518.	0.18282716906707	1519.	0.00304432177999758
1520.	0.000540826488981254	1521.	0.0632698888856234	1522.	0.0461485611835041
1523.	0.0321060970581706	1524.	0.183511094930489	1525.	0.0817094767876343
1526.	0.186968152893255	1527.	0.00021246012436316	1528.	0.0329913492083257
1529.	0.0342846672303417	1530.	0.185705226334665	1531.	0.0246936137728781
1532.	0.107725500985455	1533.	0.0136229486817204	1534.	1.58266570331479e-08
1535.	0.100098439972503	1536.	0.142778685533512	1537.	0.0672715787385205
1538.	0.0599376460032566	1539.	0.00852083628896531	1540.	0.0946342157171158
1541.	0.20171662828553	1542.	0.15363246227661	1543.	0.131168712176624
1544.	0.0655277779918044	1545.	0.0279599680705669	1546.	0.034824759756351
1547.	0.113706157076145	1548.	0.0308819260924292	1549.	0.0972323199246849
1550.	0.140286598162167	1551.	0.0101019732435993	1552.	0.0497210943728619
1553.	0.15049917042381	1554.	0.0883419001121146	1555.	0.0748670723963691
1556.	0.0778282248883587	1557.	0.211453977785884	1558.	0.0868774151250764
1559.	0.0869811804523754	1560.	2.47979194294222e-08	1561.	0.0665610252457633
1562.	0.239418652206212	1563.	0.000552395553906833	1564.	0.127824341330999
1565.	0.145605248914023	1566.	0.0396604632839398	1567.	0.0374348335011292
1568.	6.76604909674647e-12	1569.	0.00593918167907409	1570.	0.0982845266660222
1571.	0.00646696138103839	1572.	0.146600394906571	1573.	0.126262981503952
1574.	0.0140765860645349	1575.	0.0920371698182883	1576.	0.0951016178847228
1577.	0.0130338582803274	1578.	0.187281762484381	1579.	0.0102948289786067
1580.	0.16205004174636	1581.	0.106705755333651	1582.	0.17670120452119
1583.	0.00747573690416112	1584.	0.161705200347452	1585.	0.0655471946228206
1586.	0.144630040627382	1587.	0.0780283948555917	1588.	0.137690325927255
1589.	0.00225994701596236	1590.	0.0947446742703318	1591.	0.00977219187841097
1592.	0.0716022780839124	1593.	0.0564324332438006	1594.	0.0641958959574638
1595.	0.120447646083311	1596.	0.20782054493177	1597.	0.0138934999956584
1598.	0.0591110593935515	1599.	0.0393447888365865	1600.	0.119946203485765
1601.	0.013011462735255	1602.	1.67787143842501e-06	1603.	0.132675133711245
1604.	0.162973717183523	1605.	0.130613796744463	1606.	0.166176739131164
1607.	0.120902215998004	1608.	0.08843442757033	1609.	0.00170343295708871
1610.	0.000945712915948555	1611.	0.0743802059584062	1612.	0.13065701044908
1613.	0.0144617734581749	1614.	0.0155744617912678	1615.	0.011412463308685
1616.	0.0367449850435847	1617.	0.0596415946532274	1618.	0.135947501272693
1619.	0.124695423936971	1620.	0.103176556586963	1621.	0.150307050889785
1622.	0.0210360218948854	1623.	0.00653380880236185	1624.	0.0710202660996014
1625.	0.0889577752153762	1626.	0.0767756463422442	1627.	0.0581125867938008
1628.	0.0280421423534446	1629.	0.146430616137386	1630.	0.000648180486194716
1631.	0.00611062667578122	1632.	0.0765137428751883	1633.	0.00015436609941662
1634.	0.120222087786551	1635.	0.00341813200504886	1636.	0.202355518334205
1637.	0.153386595811034	1638.	0.0798500534738814	1639.	0.000976361716515345
1640.	0.03582146155129	1641.	0.194214774722351	1642.	0.0257957898384873
1643.	0.187923578880281	1644.	0.095453045473273	1645.	0.213972003931628
1646.	0.113169205370623	1647.	0.0677170520959597	1648.	0.00046439393307622
1649.	0.0174341075358544	1650.	0.149583802316519	1651.	0.190334513596052
1652.	0.0162617633033135				

1653.	0.204882867608122	1654.	0.0751096091144597	1655.	0.148967049132418
1656.	0.0105056177841913	1657.	0.163004027385959	1658.	0.0445407902276409
1659.	0.0862774838380419	1660.	0.122067322885798	1661.	0.0938508436370889
1662.	0.0584582698301476	1663.	0.182376880730492	1664.	0.142323801801556
1665.	0.18534428059347	1666.	0.00192259709273875	1667.	0.170567244546628
1668.	0.0739886479660083	1669.	0.0897629725138334	1670.	0.0816283676873928
1671.	0.157139824993711	1672.	0.125571166741955	1673.	0.113396317814206
1674.	0.0813041976541826	1675.	0.0846503826944523	1676.	0.0070467363434923
1677.	0.0959912007045061	1678.	0.113226350266487	1679.	0.280324001331497
1680.	0.0889757669718873	1681.	0.0274827613884278	1682.	0.208123090017801
1683.	0.000684585330205395	1684.	0.169981369225567	1685.	0.0974417521267852
1686.	0.0349630001689848	1687.	3.87003942401511e-13	1688.	0.0316067074576847
1689.	0.100223826497602	1690.	0.148984056766858	1691.	1.40864940660216e-08
1692.	0.00882406135711528	1693.	0.13853046887225	1694.	4.44470011835545e-05
1695.	0.0617308944204692	1696.	0.132003704305012	1697.	0.0226254875209668
1698.	0.0839965782470218	1699.	0.0382514157162913	1700.	0.000409100487302531
1701.	0.0991427994446507	1702.	5.89223723751909e-05	1703.	0.116676487368843
1704.	0.0937033094901038	1705.	0.0648330603470971	1706.	0.139526406718759
1707.	0.160671990292954	1708.	0.0284338730706432	1709.	0.000231903345997424
1710.	0.00386447298276408	1711.	0.006644960117969	1712.	0.20287606241436
1713.	0.0743045510158566	1714.	0.0544559254373294	1715.	0.107246379153621
1716.	0.157463690018889	1717.	0.133786078210325	1718.	0.0941983145487716
1719.	0.0145229615656806	1720.	0.000403606399380612	1721.	0.033417863390823
1722.	0.2036089894874	1723.	0.111457610321221	1724.	0.0824599796152239	1725. 0.139377681073916
1726.	0.112769042448908	1727.	0.139843810424356	1728.	0.0617363814243149
1729.	0.0786949581796785	1730.	0.00062089547367089	1731.	0.24639548023091
1732.	0.15292503023955	1733.	0.112044692648861	1734.	0.125625102890383	1735. 0.120325262147985
1736.	0.16008049608777	1737.	0.0508992025514717	1738.	0.127334509855369
1739.	0.158760458951268	1740.	0.0698648338090574	1741.	0.0259217084981695
1742.	0.00252404780808777	1743.	0.144919672805312	1744.	0.183621696344786
1745.	0.0293260732386693	1746.	0.0871734177439353	1747.	0.000263333645196467
1748.	0.0861445172371452	1749.	0.116987848791447	1750.	0.0707360628128614
1751.	0.0280548047290561	1752.	0.124270731803618	1753.	0.0107589085797213
1754.	0.155521096270179	1755.	0.0849877588264371	1756.	0.150461692082959
1757.	0.0908600270971677	1758.	0.0985396034482392	1759.	0.0909911110633014
1760.	0.147419356109778	1761.	0.14090322264427	1762.	0.10541498999327	1763. 0.141573962915104
1764.	0.0465405058310627	1765.	0.132563656131802	1766.	0.0265548306024163
1767.	0.0588363780429617	1768.	0.104442535947384	1769.	0.0664725787476627
1770.	0.204619646969002	1771.	0.155223975653527	1772.	0.169384108931072
1773.	0.144895886202153	1774.	0.000155952134217432	1775.	0.132222360974284
1776.	0.104323308258524	1777.	0.0556556320574349	1778.	0.0576981584069611
1779.	0.0501898531643775	1780.	0.038948151420231	1781.	0.147690578165852
1782.	0.172455964324785	1783.	0.0388738993167444	1784.	0.064663590245256
1785.	0.0140498957984711	1786.	0.0652862023095525	1787.	0.172078261553505
1788.	0.0148189142461141	1789.	0.012456115897476	1790.	0.0564093736496353
1791.	0.162246901609688	1792.	0.0768180402030389	1793.	0.000283721223755443
1794.	0.0408309843444624	1795.	0.188529534706362	1796.	0.0517370786531792
1797.	0.0885850123871067	1798.	0.186276671626286	1799.	7.88386499394446e-09

1800.	0.113944632337592	1801.	0.00846682530130838	1802.	0.113654611450666
1803.	0.171939506381318	1804.	0.00607916474342513	1805.	0.0497716262521102
1806.	0.143614114415345	1807.	3.06304243096742e-06	1808.	0.156392462918597
1809.	0.0380259402017198	1810.	0.174957460108141	1811.	0.000472669619540218
1812.	0.092213511016066	1813.	0.167679020281901	1814.	0.0561007770725662
1815.	0.208013955903737	1816.	0.296963008508102	1817.	0.0999046742795149
1818.	0.0434952020651372	1819.	0.0168226255259575	1820.	0.0761179405473177
1821.	0.0505244047109074	1822.	0.00793835430932234	1823.	0.0308955853572147
1824.	0.134544375632828	1825.	0.146346309571879	1826.	0.0526698410674438
1827.	0.13635840347326	1828.	1.83190446909633e-06	1829.	0.0575753433252045
1830.	0.00196572068764672	1831.	0.157130799942758	1832.	0.00456031950711582
1833.	0.150129473566533	1834.	0.180045050311919	1835.	0.0521743589164171
1836.	0.189997748319802	1837.	0.0508662131790841	1838.	0.172231356458315
1839.	0.00895163764524531	1840.	0.105836385338509	1841.	0.0646817576712467
1842.	0.103201148141792	1843.	0.0858256087967283	1844.	0.164352388733172
1845.	0.00308455342547405	1846.	0.0418849851677167	1847.	0.0408825673918085
1848.	0.0168052143874944	1849.	5.76463535923863e-06	1850.	0.0526716062319791
1851.	0.0180600277160755	1852.	0.245184229260311	1853.	0.0130380782265129
1854.	0.070791509656699	1855.	0.0200869342342611	1856.	0.0334945320160273
1857.	0.102861185057805	1858.	0.00267667148825856	1859.	0.0174326643685015
1860.	0.153341657618351	1861.	0.0813944352865367	1862.	0.210325045299063
1863.	0.192142388743149	1864.	0.000251052434689456	1865.	0.0166975706659429
1866.	0.0344091264624252	1867.	0.0289355875438631	1868.	0.117398880865576
1869.	0.000691117400142847	1870.	0.00267832574938259	1871.	0.0157509950974128
1872.	0.0551790112268127	1873.	0.0921155969120467	1874.	0.03278298220252
1875.	0.0943612734484812	1876.	0.0714645989543595	1877.	0.0295720291535813
1878.	0.0703722733782346	1879.	0.17568708046654	1880.	0.204224383278196
1881.	0.056884191604777	1882.	0.0094400861791444	1883.	8.16483256614251e-05
1884.	0.225380668455301	1885.	0.112916388681769	1886.	0.166209497479783
1887.	0.09604373964326				
1888.	0.0382719558538369	1889.	0.0842557892160786	1890.	0.00682703123218679
1891.	0.0952348732259344	1892.	0.125194538044453	1893.	0.10157369070256
1894.	0.0105578525426795	1895.	0.0031074106711077	1896.	0.0132309667471959
1897.	0.233264427124008	1898.	0.0218336694655369	1899.	0.113012295546857
1900.	0.20012776789456	1901.	0.0712013263431809	1902.	0.0362678832677168
1903.	0.131823505519146	1904.	0.129672094223364	1905.	0.0364830563171185
1906.	0.0881330028966104	1907.	0.1746997458997	1908.	0.00741677167383772
1909.	0.213239911721279	1910.	0.119721689295474	1911.	0.0971557816259373
1912.	0.0601321853060189	1913.	0.00420643205459228	1914.	0.0407811260921976
1915.	0.147556962280561	1916.	0.129345879707842	1917.	0.000496635146706028
1918.	0.116556192875121	1919.	0.0715537624019247	1920.	0.0748677068619042
1921.	0.0280538214409791	1922.	0.0800797099841669	1923.	0.0387269051162929
1924.	0.0231870805478533	1925.	0.0344298123814417	1926.	0.107012694378017
1927.	0.138363487829075	1928.	0.145823927670908	1929.	0.00601678520235381
1930.	0.0354513045615338	1931.	0.0919311109867364	1932.	0.189081543732235
1933.	0.00918084088652839	1934.	0.0755531134067989	1935.	0.153679795080119
1936.	0.132799388497487	1937.	0.0900770110751134	1938.	0.183634043279155
1939.	0.245600836806576	1940.	0.115627338006666	1941.	0.0563913568974838
1942.	0.19801171839989	1943.	0.0912786497876031	1944.	0.154395656115755

1945.	0.000494621733431697	1946.	0.258392501407854	1947.	0.0473569585580544
1948.	0.0345966545827583	1949.	0.0101302443123233	1950.	0.196969751724474
1951.	0.11273939460234	1952.	0.0892699943645302	1953.	0.137533871855504
1954.	0.0569204787295437	1955.	0.0672417528339282	1956.	4.88618466229632e-05
1957.	0.00991781653332615	1958.	0.0441271531053299	1959.	0.0448032925319335
1960.	0.122176598402841	1961.	0.072509639533576	1962.	0.0647593663359836
1963.	0.212196220289366	1964.	0.0938782332085892	1965.	0.0460068622004617
1966.	0.0928953372959986	1967.	0.169547794754845	1968.	0.184693793576451
1969.	0.0251138003937164	1970.	0.173979314610897	1971.	0.202839870094801
1972.	2.42336567607985e-09	1973.	0.246343754584186	1974.	0.0924877789586386
1975.	0.0187658308655686	1976.	0.152108416116432	1977.	0.100001035075162
1978.	0.0125654369469623	1979.	0.223233596491197	1980.	0.153801275440318
1981.	0.00943375896955095	1982.	0.026740644991967	1983.	0.0690814552790354
1984.	0.169658395599468	1985.	0.125820726619689	1986.	0.0797879838858015
1987.	0.00774059783141221	1988.	0.021490396681675	1989.	0.115660412217383
1990.	0.109796303088228	1991.	0.120610174062552	1992.	0.0143384179014835
1993.	0.0120114297779907	1994.	0.101187055922996	1995.	0.00691555403862242
1996.	0.0744666375163023	1997.	0.00361915235298902	1998.	0.0138144367554715
1999.	0.236037819100698	2000.	0.15963514081986	2001.	0.0302433129258716
2002.	0.216825902710328	2003.	0.109976950702961	2004.	0.206694364457793
2005.	0.00100645377576108	2006.	0.000553551883221112	2007.	0.121647488369793
2008.	0.172368001696034	2009.	0.0903817017244458	2010.	0.00297602676552253
2011.	0.0529570000390401	2012.	0.00762508430940185	2013.	0.123270057533249
2014.	0.1303667917144	2015.	0.104118942950596	2016.	0.12049389406366
2017.	0.0574588756024366				
2018.	0.167320915593872	2019.	0.115926153062552	2020.	0.0778486998479957
2021.	0.122479258119387	2022.	0.0332635133501068	2023.	0.00829636011372459
2024.	0.0663255071307155	2025.	0.114428965881646	2026.	0.0268905112155449
2027.	0.127665044511015	2028.	0.112989381969438	2029.	0.234787089660255
2030.	0.0201236261034619	2031.	0.0203347528437258	2032.	0.0963937546694442
2033.	0.208282931568991	2034.	4.91280273397694e-11	2035.	0.0939273131428187
2036.	0.0740425409502402	2037.	0.13054375470451	2038.	0.0869244428327599
2039.	0.119926966062158	2040.	0.122549562200172	2041.	0.184401277784617
2042.	0.00404786440309004	2043.	0.133015478300071	2044.	0.0676040618262685
2045.	0.0261653372245778	2046.	0.115110337535456	2047.	0.0235118065152285
2048.	0.0950923684364553	2049.	0.154679282443822	2050.	0.0835425496385221
2051.	0.123922751217518	2052.	0.053694353621876	2053.	0.178386312068266
2054.	0.00158856331792128	2055.	0.134626259664957	2056.	0.214668458264576
2057.	0.0162876998664754	2058.	0.224354016905142	2059.	0.080151528987907
2060.	0.19237309346094	2061.	0.0973727506378264	2062.	0.0209601580996562
2063.	0.136576390531563	2064.	0.00994353832135102	2065.	0.0742207451305864
2066.	0.12292252908057	2067.	0.11364843592138	2068.	0.0511522974045018
2069.	0.0685893614992503	2070.	0.0233269862336909	2071.	0.0125143101370734
2072.	0.00743794148711392	2073.	0.0892242024297687	2074.	0.189651748017398
2075.	0.209362447935277	2076.	0.157834090049167	2077.	0.00272425275777694
2078.	0.127385837652603	2079.	0.0619004771484505	2080.	0.140256466683075
2081.	0.142912547756372	2082.	0.0145883664846646	2083.	0.0891197613595996
2084.	0.106548133172586	2085.	0.146148716953909	2086.	0.0161155120750872
2087.	0.112853429114185	2088.	0.000559524565965213	2089.	0.107521274109515

2090. 0.0128490982395019 2091. 0.118436038634169 2092. 0.0531239061903453
2093. 0.0837264346862186 2094. 2.61782868648051e-06 2095. 0.0500601446403614
2096. 0.0751230387375135 2097. 0.0785942175132563 2098. 0.0599375611814914
2099. 0.161322859550629 2100. 0.103037158465355 2101. 0.0786650806966946
2102. 0.0331861102131429 2103. 0.00934493732148749 2104. 0.0855358359286552
2105. 0.0340228661250603 2106. 0.0100592225043059 2107. 0.178156914628384
2108. 0.144473537357834 2109. 0.0153633019320705 2110. 0.0448549621672543
2111. 0.152450741948491 2112. 0.0903494770482844 2113. 0.00925358324308846
2114. 0.161702839500758 2115. 0.0704513912024027 2116. 0.04011258581037
2117. 0.155790320908205 2118. 0.114086509486132 2119. 0.000713792390321644
2120. 0.15951271385023 2121. 0.0504218452986799 2122. 0.17374815101461
2123. 0.00533148503730005 2124. 0.00555303407588569 2125. 0.0131293537158709
2126. 0.198877043349504 2127. 0.187319553658491 2128. 0.147678806114111
2129. 0.103305539341386 2130. 0.0586657875467009 2131. 0.000103402142491066
2132. 0.0319499139974831 2133. 0.118562666590372 2134. 0.174366939722095
2135. 0.108168038117711 2136. 0.00175790780644911 2137. 0.206479466765383
2138. 0.076367954189938 2139. 0.061771026553865 2140. 0.00625758917208333
2141. 0.0770554108856569 2142. 0.0103234730077748 2143. 0.00547740619517559
2144. 0.0987141647141552 2145. 0.128618104743693 2146. 2.95551464665256e-07
2147. 0.0126441982556171 2148. 0.136537769129521 2149. 0.182521323145919
2150. 0.00717440080792854 2151. 0.0508113920910543 2152. 0.0572195881417973 2153. 0.161546861399594
2154. 0.148025071133343 2155. 0.190516552146684 2156. 0.00571960067447774 2157. 0.109231017478242
2158. 0.183378888931859 2159. 0.0289407495344585 2160. 0.153562988713982 2161. 0.115938039529976
2162. 0.0253118215570338 2163. 0.120537202881094 2164. 0.0973631573722644 2165. 0.063925883449284
2166. 0.0611351986305681 2167. 0.165820064993869 2168. 0.115320008020139 2169. 0.100393600745545
2170. 0.185584781501827 2171. 0.00431380617410433 2172. 0.0609864990271236 2173. 0.05076190163095
2174. 0.00469243765408525 2175. 0.00451979716907037 2176. 0.126595869596715 2177. 0.199348054636492
2178. 0.000914627135614429 2179. 0.131992596436322 2180. 0.0523768216280667 2181. 0.00603354852421874
2182. 0.107436412492923 2183. 0.018871239218532 2184. 0.057480262788377 2185. 0.090962093544051
2186. 0.0541100746324851 2187. 0.0998236564414403 2188. 0.177693817518987 2189. 0.00884162579784305
2190. 0.11085779439135 2191. 0.0694676063380459 2192. 0.0617224239356116 2193. 0.0543447750102661
2194. 0.00447758141507031 2195. 0.00657391852321174 2196. 0.107009384922534 2197. 0.00623897912093164
2198. 0.0661197749045431 2199. 0.166769097514435 2200. 0.00637131697528712 2201. 0.113108907234555
2202. 0.139586819362468 2203. 0.184283861362404 2204. 5.64728292815881e-07 2205. 0.0300795095198048
2206. 3.32167377456345e-06 2207. 0.0280107421182174 2208. 0.0807098717630828 2209. 0.191743289966374
2210. 0.162871325561532 2211. 0.0775748304950686 2212. 0.0305834794871525 2213. 0.0842538925510479
2214. 0.0262944590369605 2215. 0.0422784054556861 2216. 0.180899890219425 2217. 0.0586354505983838
2218. 0.242116880623188 2219. 0.116798314349715 2220. 0.0381306819743764 2221. 0.123843607495106
2222. 0.166774638227076 2223. 0.106873650901512 2224. 0.0101580798151125 2225. 0.195448685010405
2226. 0.168701674497258 2227. 0.00582766321489818 2228. 0.0144559150169621 2229. 0.223484489888544
2230. 6.59327321375912e-11 2231. 0.0153968280050343 2232. 0.227363650663976 2233. 0.0772088637061584
2234. 0.0447205392210156 2235. 0.0538678499983455 2236. 0.00297255629932025 2237. 0.107929269044601
2238. 0.281150487090104 2239. 0.0602700681331482 2240. 0.0205927914436913 2241. 0.00571798434010258
2242. 0.179697717359287 2243. 0.177678427748965 2244. 0.0561435650541735 2245. 0.168649708521689
2246. 0.0110544964506955 2247. 0.109777961514844 2248. 0.0213427023761553 2249. 0.00373775928779147
2250. 0.134941491966256 2251. 0.0040148780410284 2252. 0.167570125115808 2253. 0.103825943155593
2254. 0.0692663086502873 2255. 0.106489366046264 2256. 0.102332969363247 2257. 0.0125481376882655
2258. 0.0103584248599393 2259. 0.0955293523370211 2260. 0.0701286010800738 2261. 0.000850493001087098

2262. 0.158093191373081 2263. 0.0514718936975529 2264. 0.0263652978247182 2265. 0.04152405731448
2266. 0.151013314644367 2267. 0.17945531008164 2268. 0.146136259635554 2269. 0.0690178361047847
2270. 0.141833122921751 2271. 0.137344186752286 2272. 0.109315432490377 2273. 0.212787903526467
2274. 0.0535404215297335 2275. 0.0988877609704313 2276. 0.0645693074422416 2277. 0.141641624038004
2278. 0.112385503348198 2279. 0.0855635402312091 2280. 5.26897571880902e-06 2281. 0.00090254318826728
2282. 0.0447913577371442 2283. 0.137128897358504 2284. 0.109597849796313 2285. 0.115527587406286
2286. 0.0868564564794695 2287. 0.0467145414589898 2288. 0.102083605389734 2289. 1.16692100139094e-07
2290. 0.145037807516382 2291. 0.0389451182067045 2292. 0.162408109982855 2293. 0.168937639681558
2294. 0.153120771672566 2295. 0.0312473757676187 2296. 1.06036203797158e-10 2297. 0.129668411618899
2298. 0.193144742840185 2299. 0.00988897386474876 2300. 0.0746521497509675 2301. 0.0068895003450868
2302. 0.207405488153807 2303. 0.145784821869341 2304. 0.0577840032765054 2305. 0.0953578204836309
2306. 0.0311370055493674 2307. 0.140884299540146 2308. 0.219636086727477 2309. 0.156967516247079
2310. 0.00699847449032727 2311. 0.0159904274730513 2312. 0.0400692964868436 2313. 0.106074834777696
2314. 0.0316605502996027 2315. 0.120797618663195 2316. 0.198424247884642 2317. 0.252403603552789
2318. 0.000406371611829233 2319. 0.00104900452067165 2320. 0.0957673016760459 2321. 0.12871646401574
2322. 0.124699244984951 2323. 0.00552811492672338 2324. 0.00107530496556949 2325. 0.1708892388299
2326. 0.0875394047520555 2327. 0.0172227277186707 2328. 0.0476485646497689 2329. 0.0993698454070259
2330. 0.0436936091115718 2331. 0.052438696260927 2332. 0.0219327799725175 2333. 1.76678223602015e-10
2334. 0.0073019718165103 2335. 0.0119115391738083 2336. 0.0581712970783863 2337. 0.0100011754640961
2338. 0.29422816044505 2339. 0.132113583449449 2340. 0.026461045823121 2341. 0.180735951056435
2342. 0.0574596329528419 2343. 0.102074091153775 2344. 0.0184318716096007 2345. 3.87375086141486e-05
2346. 0.00569952768366436 2347. 0.0870727991820519 2348. 0.106177436424824 2349. 0.230052625167679
2350. 0.1291737259501 2351. 0.0603960333827191 2352. 6.27602827867675e-05 2353. 0.0117501368120848
2354. 0.00701113589718357 2355. 0.0107426338167754 2356. 0.154509356118151 2357. 0.214745709216461
2358. 0.105003931542703 2359. 0.0312554080285544 2360. 0.0315008291146176 2361. 0.0494268577532148
2362. 0.172629517294549 2363. 0.0875312200151209 2364. 0.0470363890588493 2365. 0.0366090427499145
2366. 0.0196150635946819 2367. 0.00078405617437249 2368. 0.0788017142346984 2369. 0.0962976710087883
2370. 0.236776693000314 2371. 0.123634191329562 2372. 0.0787992296521226 2373. 0.005347720691577
2374. 0.120087638028793 2375. 0.121867477809817 2376. 0.136340142294926 2377. 0.0546736693373727
2378. 0.066194498224438 2379. 0.0102656526974985 2380. 0.000552867038952795 2381. 0.00706290396684968
2382. 0.0473927805584287 2383. 0.0140138542702226 2384. 0.0463464986573657 2385. 0.107065483568473
2386. 0.124686682692264 2387. 0.139240071949322 2388. 0.00295875183301111 2389. 0.104270617679438
2390. 0.000426597883631176 2391. 0.119753104319861 2392. 0.107074172923894 2393. 0.00841350273965282
2394. 0.105255486623319 2395. 0.123288600787847 2396. 0.144286730222094 2397. 0.121192140715668
2398. 0.12405525832244 2399. 8.13186632388629e-05 2400. 0.136431631591102 2401. 0.229641103585547
2402. 0.107025567899406 2403. 0.00170802693385548 2404. 0.140741365755814 2405. 0.184787928173565
2406. 0.0368058083985605 2407. 0.00244890418916002 2408. 0.00491815397523895 2409. 0.174100014443774
2410. 0.199874343901802 2411. 0.118180185925351 2412. 0.079019044777803 2413. 0.022235564689968
2414. 0.0813150910676644 2415. 0.0723144935640892 2416. 0.00981483910919838 2417. 0.00048617825425634
2418. 0.0840066124505324 2419. 0.088923616520595 2420. 0.0120140424098158 2421. 0.0816529404649864
2422. 0.0828458743670255 2423. 0.0325993181641505 2424. 0.179046419146637 2425. 0.0479926807680258
2426. 0.0384874247219786 2427. 0.00078263673116167 2428. 0.104291524684804 2429. 0.12910867380945
2430. 0.000706528861943197 2431. 0.165815043924884 2432. 0.0222852328386718 2433. 0.0235362482979218
2434. 0.0770693370154212 2435. 0.145022079302876 2436. 0.0595334442552797 2437. 0.109609851607633
2438. 0.068708336137737 2439. 0.155044381400222 2440. 0.0433101840794739 2441. 0.145460630379313
2442. 0.263637335378973 2443. 0.00156004864907354 2444. 0.0500908849856729 2445. 0.147401757315885
2446. 0.110546574586047 2447. 0.187577951736865 2448. 0.12469553697263 2449. 0.208888108836411
2450. 0.0119564388986783 2451. 0.126789024550512 2452. 0.0116263343161008 2453. 0.151056668536508

2454. 0.0997480978485484 2455. 0.00270890673340702 2456. 0.17742398664145 2457. 0.0842509315035317
2458. 0.0969991459306343 2459. 0.14175779527051 2460. 0.140860323630785 2461. 0.0516152699055991
2462. 0.0942354544425726 2463. 0.0574126277995748 2464. 0.0816119832224731 2465. 0.0827592629548196
2466. 0.146041348209056 2467. 0.0683920705612394 2468. 0.123875756405394 2469. 0.079345345273135
2470. 0.100322562568465 2471. 0.27160863015539 2472. 0.218621872600553 2473. 3.91457081391335e-07
2474. 0.135570971855978 2475. 0.0613492603665061 2476. 0.107345387494147 2477. 0.185943009476066
2478. 0.0583456409757082 2479. 0.0304100414752546 2480. 0.0350423229261663 2481. 0.000897702003080273
2482. 0.0660766333760471 2483. 0.190742255426811 2484. 0.175052520435186 2485. 0.0159145574180912
2486. 0.118437757678993 2487. 0.0924406828074542 2488. 0.0842724079065872 2489. 0.222193056637769
2490. 0.0688306136576809 2491. 0.0417351752237764 2492. 0.0438601103662558 2493. 0.109246176396361
2494. 0.130823167208236 2495. 0.170875688392819 2496. 0.0105608551824874 2497. 0.0932687373227975
2498. 0.00993799444689661 2499. 6.02819711210536e-06 2500. 0.0154643192489491 2501. 0.0126571397204969
2502. 0.160440851630559 2503. 0.121908075815979 2504. 0.142125397666553 2505. 0.0784585230898273
2506. 0.164269542885699 2507. 0.205897460200685 2508. 0.21285760836095 2509. 0.306239428929724
2510. 6.40280496323143e-05 2511. 0.0374314642642295 2512. 0.0488914610528243 2513. 0.0421126780071018
2514. 0.0957064942828503 2515. 0.123668791784242 2516. 0.0618733002960566 2517. 0.0743631873017671
2518. 0.132811621666005 2519. 0.00758111618494668 2520. 0.0973835566871471 2521. 0.168855898178878
2522. 0.0573370486312596 2523. 0.214653510418152 2524. 0.00924688563909623 2525. 0.10510767001501
2526. 0.0450434517544574 2527. 0.0696178342594341 2528. 0.167029380265151 2529. 0.156738434648813
2530. 0.0416562926317468 2531. 0.0719673653396506 2532. 0.0873654485654695 2533. 0.00881307605018706
2534. 0.0881396644931021 2535. 0.184354621033046 2536. 0.146449904299622 2537. 0.00114668316981826
2538. 0.00161438292625074 2539. 0.0842837922656591 2540. 0.0820166278820415 2541. 0.0146709720123434
2542. 0.0233875764744718 2543. 0.0258407894689934 2544. 0.219299596922768 2545. 0.13871663356197
2546. 0.0432926072518696 2547. 0.120256182773853 2548. 0.00142992115396337 2549. 0.0826461206380203
2550. 0.157925930877638 2551. 0.0986843732550759 2552. 0.0108051599559284 2553. 0.112040205872916
2554. 0.222299660958763 2555. 0.0587757167933933 2556. 0.00594892091507251 2557. 0.0875155361472494
2558. 0.0822369159051157 2559. 0.204602505561589 2560. 0.10975037907636 2561. 0.0955507992547762
2562. 0.175995883851739 2563. 0.105616699160312 2564. 0.198807148830523 2565. 0.013447597780491
2566. 0.000720787876013127 2567. 0.173362216508284 2568. 0.00671691518972328 2569. 0.000857941145518934
2570. 0.0377276231101524 2571. 0.0569525853246935 2572. 0.141710661893033 2573. 0.121949747346353
2574. 0.0314767703083785 2575. 0.0085294497714384 2576. 0.094037538780331 2577. 0.109691653473298
2578. 0.0721430461249635 2579. 0.0906795413032537 2580. 0.100815520055336 2581. 0.0999597437977451
2582. 0.012136046391969 2583. 0.0110905205996846 2584. 0.0935308322997084 2585. 0.177014838779177
2586. 0.0762342987918223 2587. 0.160948432271274 2588. 7.00381816085094e-11 2589. 0.00106783985878699
2590. 0.00396735931228104 2591. 0.0154024225430565 2592. 0.0985325465140829 2593. 7.02922619968254e-05
2594. 0.128513608597344 2595. 0.136526040376376 2596. 0.117463036172827 2597. 0.00848183072134702
2598. 0.122794381395839 2599. 0.107478408835202 2600. 0.190058175793068 2601. 0.0200956415080159
2602. 0.126448139963565 2603. 0.182422521253871 2604. 0.135973055532728 2605. 0.228727714849065
2606. 0.0085262976137831 2607. 0.00396180897940016 2608. 0.0345496744020766 2609. 0.0260784658483639
2610. 0.0145079481876027 2611. 0.0199794834717926 2612. 0.17791482731978 2613. 0.154097893362241
2614. 0.00232896386473836 2615. 0.173610822021449 2616. 0.0166442293705693 2617. 0.146752416442997
2618. 0.109445396013112 2619. 0.0873122732448953 2620. 1.99592735225555e-05 2621. 0.115297024127814
2622. 0.0329945060687294 2623. 0.0266228241479665 2624. 0.0835121753347581 2625. 0.0356006978063713
2626. 0.157434086322805 2627. 0.115559974455482 2628. 0.102241337947189 2629. 0.210989799680202
2630. 0.21863886827561 2631. 0.0348716804103538 2632. 0.115690288464182 2633. 0.013422653142937
2634. 0.173456371863878 2635. 0.0693992502822239 2636. 0.000228582614662773 2637. 0.146007513328849
2638. 0.0058674522275789 2639. 0.0500164928199405 2640. 0.154168940625562 2641. 0.166248073127013
2642. 0.018557680593043 2643. 2.9182247627101e-06 2644. 0.007079254461016 2645. 0.000316344478302914

2646. 0.15678403242018 2647. 0.0130624534209331 2648. 0.111238642255401 2649. 0.0608116708521276
2650. 0.139213445630944 2651. 0.034893436594774 2652. 0.0016457747385746 2653. 0.00412236066824863
2654. 0.139680633677788 2655. 0.0706852596030181 2656. 0.0675268353108367 2657. 0.140366062903889
2658. 0.0963792132424046 2659. 0.0348408704665349 2660. 0.0399038425465927 2661. 0.106377716809038
2662. 0.16910077406498 2663. 0.13999105531108 2664. 8.34039739501509e-09 2665. 0.104356858103482
2666. 0.0483974651794413 2667. 0.0569766430102384 2668. 0.0160877667930152 2669. 0.105339899604599
2670. 0.0115884452498457 2671. 0.0867606245859053 2672. 0.133226627947127 2673. 0.191872819736913
2674. 0.0896343676978527 2675. 0.0589092804442396 2676. 0.0195590561720365 2677. 0.000162336209077963
2678. 0.117831100062462 2679. 0.102979040146812 2680. 4.22293728795372e-06 2681. 0.0595843660531331
2682. 0.0259810440445556 2683. 0.109727668929267 2684. 0.183760102608986 2685. 0.0358796481583742
2686. 0.0608735378671701 2687. 0.0689625564148906 2688. 0.0497610475399749 2689. 0.190092795835165
2690. 0.119888653962669 2691. 0.122701848444794 2692. 0.0244563859871718 2693. 0.00429570218674683
2694. 0.063897726035773 2695. 0.00977091141353744 2696. 0.16623919061286 2697. 0.0769370574778609
2698. 0.201401206701931 2699. 0.00530371197167409 2700. 0.016410085058656 2701. 0.0536281805098338
2702. 0.163718893510189 2703. 0.145583371618402 2704. 0.00677921640516115 2705. 0.135077767378708
2706. 0.000939814990435236 2707. 0.0953810341032501 2708. 0.0184577544549734 2709. 0.0106997151051791
2710. 5.56464516295379e-05 2711. 0.00371494376771424 2712. 0.0980289281526271 2713. 0.0607366515590235
2714. 0.0211837966645882 2715. 0.20519580842889 2716. 0.130666570345972 2717. 0.120496881429932
2718. 0.0415745921798006 2719. 0.0623411006934673 2720. 0.0369820638614427 2721. 0.167557608056599
2722. 0.0156231347931585 2723. 0.0410922815652154 2724. 0.0645069695888926 2725. 0.000906797812747803
2726. 0.0795809442043367 2727. 0.13722628739223 2728. 0.130876164299725 2729. 0.0583581671069962
2730. 0.205455572045983 2731. 0.128294733788561 2732. 0.0530633738452395 2733. 0.00311605172007807
2734. 0.108099116629913 2735. 0.172392001029074 2736. 0.0748392860560036 2737. 0.23895483303345
2738. 0.00105862532955792 2739. 0.230003691591491 2740. 0.0128204004871592 2741. 0.174650092563796
2742. 0.0263624658796569 2743. 2.53346311610591e-07 2744. 0.156145264505829 2745. 0.0149738234680993
2746. 0.150544742172948 2747. 0.102650218015941 2748. 0.00126532323593807 2749. 0.0154117312647313
2750. 0.0631400369510895 2751. 0.0845419737454153 2752. 0.0107785541062219 2753. 0.0470978038454618
2754. 0.000378016341786755 2755. 0.137374733729314 2756. 0.0198098122390231 2757. 0.032194569843674
2758. 0.0189324776896974 2759. 0.149402860416451 2760. 0.0427291075604621 2761. 0.0933966317457822
2762. 0.0707769289896482 2763. 0.100968664589335 2764. 0.130361038815108 2765. 0.0418429353131679
2766. 0.0642354954801461 2767. 0.000778142882518895 2768. 0.20077692509563 2769. 0.0349489661042809
2770. 0.000806174624108468 2771. 0.0289171004836312 2772. 0.0228458047046508 2773. 0.122100187648025
2774. 0.000308078135757262 2775. 0.164802726048964 2776. 0.179257631115053 2777. 0.0336727489503189
2778. 0.0787750816850887 2779. 0.212480898837894 2780. 0.181782364675864 2781. 0.0799964009054388
2782. 0.0746457207033076 2783. 0.108338607299473 2784. 0.109868397300652 2785. 0.0174842345787892
2786. 0.00821316240420203 2787. 0.0281784938725619 2788. 0.021595836632838 2789. 0.0315751803764176
2790. 0.135286134531554 2791. 0.0742552416851352 2792. 0.12616966669372 2793. 0.0108031346536492
2794. 0.0601877611037471 2795. 0.00425391909970593 2796. 0.00697549467898261 2797. 0.0459156358898647
2798. 0.0857595219467924 2799. 0.15566290703359 2800. 0.0963817999934683 2801. 0.157648899073311
2802. 0.149722729444017 2803. 0.00926575620576598 2804. 0.134463840412303 2805. 0.0957943577338591
2806. 0.0119610045677668 2807. 0.0850609149801575 2808. 0.146370950077945 2809. 0.138042844341701
2810. 0.00997922804143974 2811. 0.00525606730551736 2812. 0.00757223139058055 2813. 0.125614137589772
2814. 0.157370048970189 2815. 0.0580308608988711 2816. 0.0930426292890962 2817. 0.00857469428150473
2818. 0.000173426558987174 2819. 0.0816637595107563 2820. 0.0707247852926806 2821. 0.134879590644664
2822. 0.0506255747890731 2823. 0.12911149786622 2824. 4.88386509775106e-05 2825. 0.130771234028436
2826. 0.0521348885064416 2827. 0.049706005901721 2828. 0.170963046779761 2829. 0.0901283643627014
2830. 0.000697685799781448 2831. 0.0208669882645036 2832. 0.0568426637058672 2833. 0.00265448950670851
2834. 0.000484906589705417 2835. 0.0354785110117647 2836. 0.13633771875161 2837. 0.0765416069560725

2838. 0.132414816094265 2839. 0.198066654612073 2840. 0.0187058404885577 2841. 0.0740909419612457
2842. 0.106793128204577 2843. 0.0216826047064952 2844. 0.0561906163399691 2845. 0.0422644616200932
2846. 0.0111796297402204 2847. 0.127653496627189 2848. 0.161755393477663 2849. 0.243175669802864
2850. 0.0571436045842981 2851. 0.00657945018735271 2852. 0.182820654287883 2853. 3.29206560878081e-06
2854. 0.188944736680086 2855. 0.00790194426995843 2856. 0.0801938803743938 2857. 0.136059986734696
2858. 0.0682141089399714 2859. 0.0194483229544856 2860. 0.0442037184166445 2861. 0.000520452575525023
2862. 0.113737369138417 2863. 0.209502368686199 2864. 0.0303171161440516 2865. 0.0217511359639566
2866. 0.113173176135089 2867. 0.0369015141955354 2868. 0.0145533763924146 2869. 0.151048926533855
2870. 0.0396105575890133 2871. 0.13686191697951 2872. 0.0798163448040715 2873. 0.069410121087559
2874. 0.12794454984896 2875. 0.107758114316174 2876. 0.0183187384571928 2877. 0.109182946051495
2878. 0.0102179804460857 2879. 0.162112027755896 2880. 0.111493697180986 2881. 0.0727839559131682
2882. 0.0922689687493263 2883. 0.101199583319762 2884. 0.000749073003976798 2885. 0.0516614757810986
2886. 0.126927741118342 2887. 0.23035350875251 2888. 0.0639637347368906 2889. 0.0848394098245754
2890. 0.0250589881559318 2891. 0.102124960977715 2892. 0.0724808810699657 2893. 0.0747277970026291
2894. 0.08500356454856 2895. 0.0835014612556034 2896. 0.0713163209772535 2897. 0.0891721037723914
2898. 0.0432523706071897 2899. 0.00835454856136903 2900. 0.1439009742994 2901. 0.171082023820745
2902. 0.177373725428947 2903. 9.56385019656876e-12 2904. 0.211743413640329 2905. 0.300342030888277
2906. 0.152370808672702 2907. 0.0761246556352015 2908. 0.0772703151491763 2909. 0.146433632518405
2910. 0.0850342695808435 2911. 0.0561832440076874 2912. 0.071156475654464 2913. 0.00799866078870632
2914. 0.134778890742054 2915. 0.110784900516541 2916. 0.0131079310298712 2917. 0.0171585827846037
2918. 0.156547009465701 2919. 0.151001060659791 2920. 0.0117851532931191 2921. 0.00158324270225253
2922. 0.00318695459245438 2923. 0.181877413105322 2924. 0.189484990055371 2925. 0.0599884085595337
2926. 0.0184174205443358 2927. 0.00378220939130089 2928. 0.00247276905621024 2929. 0.039480270995258
2930. 0.193416427754684 2931. 0.016344450033619 2932. 0.101378742453239 2933. 0.0772448674056148
2934. 0.0606816694963487 2935. 0.00041005037863028 2936. 0.0911826202835458 2937. 0.181714791524276
2938. 0.121350885071164 2939. 0.0678707112733639 2940. 0.0954734671534554 2941. 0.0979146026055801
2942. 0.138420528630454 2943. 0.000452562799935326 2944. 0.0115288915532865 2945. 0.000791839726924107
2946. 0.0938937590530475 2947. 0.000653390734211941 2948. 0.00660636445928139 2949. 0.114657870784488
2950. 0.114332450483322 2951. 0.0492364592637627 2952. 0.0725893355376732 2953. 0.073563445092482
2954. 0.107637663455667 2955. 0.136509851677423 2956. 0.00959512497084147 2957. 0.107529286726894
2958. 0.0492702310004323 2959. 0.212901431679695 2960. 0.131969049429538 2961. 0.190546159136779
2962. 0.0778921872593766 2963. 0.0322021912972106 2964. 0.0496085992037198 2965. 0.0840967748934554
2966. 4.1904686235638e-05 2967. 0.158106569087132 2968. 0.016183330428258 2969. 0.145590449363798
2970. 0.161734815756133 2971. 0.0502964860631248 2972. 0.00559907731977127 2973. 0.0866210502002119
2974. 0.136605450699605 2975. 7.4419305268412e-05 2976. 0.0437513512381827 2977. 0.0107477345455091
2978. 0.136746340884255 2979. 0.0517153264691958 2980. 0.0819918397938973 2981. 0.0169686934904167
2982. 0.267401063387758 2983. 0.0231756709924139 2984. 0.00643156449780511 2985. 0.00982976736715352
2986. 0.15480023622944 2987. 0.136809041023528 2988. 0.105909034495137 2989. 0.148734947661271
2990. 0.0133977724183076 2991. 0.0143728151435819 2992. 0.106645862496778 2993. 0.039465668602352
2994. 0.112375149133577 2995. 0.0569003723594877 2996. 0.00111727031498242 2997. 0.10493269795056
2998. 0.00109701361660765 2999. 1.32004281720622e-05 3000. 0.0649195718448203 3001. 0.0137925581324321
3002. 0.121926220053667 3003. 0.0200238752461641 3004. 0.167179842007643 3005. 0.0602119334503941
3006. 0.0814426927513533 3007. 2.02441481902591e-07 3008. 0.0113980899028889 3009. 0.0997977603286845
3010. 0.132479640142664 3011. 0.204344629608255 3012. 0.0164762544937708 3013. 0.133549647687781
3014. 0.00674004322641597 3015. 0.0962891294538905 3016. 0.00127309370959897 3017. 0.109556475469762
3018. 0.111027052505557 3019. 0.00186267570571738 3020. 0.0780318764751834 3021. 0.108825464648404
3022. 0.126268047468479 3023. 0.146761353505078 3024. 0.0534808185628119 3025. 0.0653294664633829
3026. 0.00489155321210438 3027. 0.015264772841204 3028. 0.0276728890515175 3029. 0.0625071486250638

3030. 0.0797452160362008 3031. 0.0878294435402578 3032. 0.00095206707722876 3033. 0.0169515243768339
3034. 0.187285257721044 3035. 0.0278102895712833 3036. 0.0829148958581506 3037. 0.00988719272411839
3038. 0.0507703662862541 3039. 0.0614247019288422 3040. 0.172064234602592 3041. 0.00121862508366219
3042. 0.00705338875174393 3043. 0.205280530799626 3044. 0.153672865914978 3045. 0.005266157231815
3046. 0.127979739633117 3047. 0.0161151100567943 3048. 0.0186164996803636 3049. 0.225738500716193
3050. 0.0324038512278819 3051. 0.118530625437387 3052. 0.157896185147392 3053. 0.102635889436239
3054. 0.193552220645457 3055. 0.0455788865479504 3056. 0.106665085974606 3057. 0.116543267823714
3058. 0.195113831657229 3059. 0.0946968098071895 3060. 0.0102740550079705 3061. 0.0928500857294563
3062. 0.022523871203326 3063. 0.210843250141371 3064. 0.0824697522572802 3065. 2.54836738572183e-05
3066. 0.0830924162620473 3067. 0.179656419649521 3068. 0.0790761454410727 3069. 8.98758826868135e-08
3070. 0.052833171011491 3071. 0.0728582319494709 3072. 0.0817365033305381 3073. 0.00284800045668045
3074. 0.164356194120346 3075. 0.00755153652506853 3076. 0.0292109086947217 3077. 0.0100318857693042
3078. 0.132369613900336 3079. 0.0703319975347744 3080. 0.00839182430501701 3081. 0.0794043643836296
3082. 0.0132802699359734 3083. 0.0328193292246213 3084. 0.00487319380882735 3085. 0.0598748765394953
3086. 0.0127550480397292 3087. 0.0135564549138163 3088. 0.144377564043659 3089. 0.000572439606178416
3090. 0.0429772168807395 3091. 6.36900777557421e-08 3092. 0.109687759560887 3093. 0.00933039584166491
3094. 0.00120220031130523 3095. 0.0440473948207553 3096. 0.150323515718056 3097. 0.15947200682506
3098. 0.0294535921921815 3099. 0.174198842535966 3100. 0.0359255272610912 3101. 0.148641311506309
3102. 0.0293315863074493 3103. 0.130300876901895 3104. 0.068538670712214 3105. 0.16767551505661
3106. 0.152164055876512 3107. 0.125481051662543 3108. 0.0625299381988447 3109. 0.142572731314506
3110. 0.0062492962044995 3111. 0.157474687033762 3112. 0.0128976519351016 3113. 0.138305869780107
3114. 0.0020830007906957 3115. 0.0559913180632609 3116. 0.170282937995016 3117. 0.000314412853748652
3118. 0.0999829530451933 3119. 0.0982631716581866 3120. 0.0922729726680896 3121. 0.00710338878100185
3122. 0.183969700764308 3123. 1.60093456894828e-11 3124. 0.0290506104737009 3125. 0.0585224825189197
3126. 0.0704372329032953 3127. 0.0231989559300087 3128. 0.0256425341599865 3129. 0.0360446891009065
3130. 0.0583700897303623 3131. 0.103621161455131 3132. 0.00603509509829172 3133. 0.0486717153423498
3134. 0.0731894341013574 3135. 0.08009888491668 3136. 0.0743103014113147 3137. 0.00359688207040047
3138. 0.0990086593901389 3139. 0.167044250891722 3140. 0.0635269769532335 3141. 0.000851817308503276
3142. 0.160489368403418 3143. 0.0935716539418327 3144. 0.135547367764234 3145. 0.125186608845625
3146. 0.00398375992879866 3147. 0.101299777597115 3148. 0.0234455737962026 3149. 4.13901553516412e-06
3150. 0.0200086707836681 3151. 0.0780484554940632 3152. 0.0445378969625043 3153. 0.119799671157598
3154. 0.132466700037725 3155. 0.122718354634395 3156. 0.0388188344145737 3157. 0.159249677375133
3158. 0.023902397216653 3159. 0.0957324231472256 3160. 0.096406337740335 3161. 0.108492900663687
3162. 0.131370627381785 3163. 0.0246257109541266 3164. 0.0409163208571779 3165. 0.0248830840805445
3166. 0.0134174428759339 3167. 0.057681326190584 3168. 0.264498177875594 3169. 0.0539015413441751
3170. 0.0995853714415209 3171. 0.131508230666321 3172. 0.0879936964746082 3173. 0.141890858786409
3174. 0.0515737030689871 3175. 0.00186013587957557 3176. 0.0442289883930824 3177. 6.73104878957647e-05
3178. 0.045068818961438 3179. 0.0161590313989212 3180. 0.027222406448011 3181. 0.0549558863875826
3182. 0.0709968851541591 3183. 0.111963735586042 3184. 0.0823137049663716 3185. 0.0346248210891539
3186. 0.0014448942928785 3187. 0.00640042842571717 3188. 0.113706523076192 3189. 0.00219461205364372
3190. 0.184051897186558 3191. 0.147189237123883 3192. 0.0603208812807786 3193. 1.50353017809333e-09
3194. 0.0491914701763855 3195. 0.0256042785520433 3196. 0.114806417423226 3197. 0.0386832180646132
3198. 0.0159529737079215 3199. 0.0243129451862758 3200. 0.1872628471315 3201. 0.0583423819344723
3202. 0.0481708048948533 3203. 0.0877892583509073 3204. 0.0842390267674619 3205. 0.00828592449146076
3206. 0.0707565678177245 3207. 0.192623666754579 3208. 0.0651344563754834 3209. 0.0396351489066338
3210. 0.00191102134476074 3211. 0.0227151235903877 3212. 0.00965513978434291 3213. 0.122414734411271
3214. 0.0886428328175278 3215. 0.137679026640766 3216. 0.175001641303576 3217. 0.0489894034704221
3218. 0.118626490607457 3219. 0.119218616304146 3220. 0.108095198267127 3221. 0.0318400516788726

3222. 0.00663638440210451 3223. 0.0272143104270798 3224. 0.0351536251103449 3225. 0.166037851026409
3226. 0.00494147882690052 3227. 0.129115364940625 3228. 0.0721298634839714 3229. 0.0169346809054269
3230. 0.151651985489368 3231. 3.56446718489747e-07 3232. 0.000111731695569878 3233. 0.0994271019363062
3234. 0.00225417601546575 3235. 0.105196936826741 3236. 0.193463164954972 3237. 0.164401125920919
3238. 0.0973653104851026 3239. 0.228820162780672 3240. 0.0954115267164349 3241. 0.125666830767467
3242. 0.0114661994064156 3243. 0.154107662511792 3244. 0.102006732609164 3245. 0.0137926688678199
3246. 0.00371583107338309 3247. 0.180636325865347 3248. 0.000741976965533273 3249. 2.26505906262193e-11
3250. 0.132457579104481 3251. 0.0120729715869926 3252. 0.0928208890901318 3253. 0.0939522354149114
3254. 0.267344037826434 3255. 0.170542365326946 3256. 0.000294771036902631 3257. 0.153661706005957
3258. 0.173588885990984 3259. 0.0997436179891315 3260. 0.112927611713325 3261. 0.0246042133026054
3262. 0.0552433480463306 3263. 0.0197411208502794 3264. 0.0180391009297682 3265. 0.12864124959958
3266. 0.208281251327344 3267. 0.0388217510877604 3268. 0.137975687303405 3269. 0.0482687919794951
3270. 0.112870883428186 3271. 0.0267760994574008 3272. 0.13640604938067 3273. 0.043358230933572
3274. 0.0825445368370642 3275. 0.0700599671879319 3276. 0.128379059988754 3277. 0.146557309762016
3278. 0.00180114616125953 3279. 0.134144962514419 3280. 0.142522353555549 3281. 0.00865911565753002
3282. 0.126500601267298 3283. 0.0947729243663857 3284. 0.0881770600478037 3285. 0.101833977928366
3286. 0.0328629453314035 3287. 0.10323636931389 3288. 0.133265917704624 3289. 6.47343719866613e-05
3290. 0.0250569827639237 3291. 0.00822036177275658 3292. 0.00349669878369358 3293. 0.101189244759619
3294. 0.146403631008481 3295. 0.0186182918085727 3296. 0.110181384976263 3297. 0.0892087483875167
3298. 0.057081223000053 3299. 0.171995122451687 3300. 0.00351718947194837 3301. 0.0748888717045272
3302. 0.0409573794150353 3303. 0.0226695502590034 3304. 0.110851012480437 3305. 0.0801220766705947
3306. 0.0709035307292345 3307. 0.15569211046483 3308. 0.0539272580799818 3309. 0.142532535042115
3310. 0.154994539844907 3311. 0.000182783445135936 3312. 0.0188783140076311 3313. 0.0116215429288148
3314. 1.01555165529793e-06 3315. 0.0253823860372041 3316. 0.0360951810281742 3317. 0.0309603886828725
3318. 0.0166107347794693 3319. 0.164506484939013 3320. 0.00570876102050007 3321. 0.158756483835788
3322. 0.0507548042894367 3323. 0.0533412982095643 3324. 0.0837462229685249 3325. 0.0211045008508437
3326. 0.0772809381162745 3327. 0.135558180445061 3328. 0.0575494098137446 3329. 2.97794170443113e-06
3330. 0.0876478276884936 3331. 0.0166675409211812 3332. 0.196373798478066 3333. 0.0551449174586385
3334. 0.19781606458331 3335. 0.0596394434559294 3336. 0.000816258682665114 3337. 0.037484420175577
3338. 0.0989740243124038 3339. 5.73020081230808e-06 3340. 0.0444260053046232 3341. 0.0208319713202322
3342. 0.0367794930623568 3343. 0.034056945022622 3344. 0.0438051612056514 3345. 0.000363994624231136
3346. 0.0556527371506961 3347. 0.0773879808558328 3348. 0.216939828049889 3349. 0.131133771112985
3350. 0.0554543802686153 3351. 0.0812221622449025 3352. 0.188024333787284 3353. 0.038753121255727
3354. 0.0605190978896749 3355. 0.200246209121514 3356. 0.0160485863621554 3357. 0.0606299474451515
3358. 0.117188553223517 3359. 0.00536564530388843 3360. 0.072982412942128 3361. 0.0549931159094654
3362. 0.0984883189155541 3363. 8.24394570197086e-06 3364. 0.0567398675102145 3365. 0.0627960162541854
3366. 0.116620455831179 3367. 0.161658344284529 3368. 0.0889416576248158 3369. 0.0951812559669205
3370. 0.170832333148812 3371. 1.99037099178605e-06 3372. 0.106070670516988 3373. 0.0762644785316578
3374. 0.147083596695214 3375. 0.0915956532863205 3376. 0.0346353120391996 3377. 0.00033255936309393
3378. 0.0907705037220251 3379. 0.11173670412789 3380. 0.165171361824162 3381. 0.010948000053903
3382. 0.146570151861353 3383. 0.00791343034450037 3384. 0.113720641768468 3385. 0.121874556116427
3386. 0.13538945872075 3387. 0.151262046363493 3388. 0.0612916608479257 3389. 0.151301341192216
3390. 0.107468240673718 3391. 0.0693185360407622 3392. 0.000233175868159437 3393. 0.0846573361891436
3394. 0.0523868712547162 3395. 0.124107202640184 3396. 0.145045093007381 3397. 0.168898854000392
3398. 0.165250062467283 3399. 0.0713083280910426 3400. 1.9830812223779e-12 3401. 0.282102560383813
3402. 0.0888773622368368 3403. 0.213956277592424 3404. 0.143201604978618 3405. 0.164411857575639
3406. 0.0194960997879346 3407. 0.15893789273665 3408. 0.0131414297264916 3409. 0.149966351518608
3410. 0.281016039704447 3411. 0.00325425040749086 3412. 0.00193229274036974 3413. 0.0795596878440596

3414. 0.0560275814030554 3415. 0.109838470478024 3416. 0.0246803459271205 3417. 0.0089356998516271
3418. 0.130268460383157 3419. 0.0742797013382912 3420. 0.108307520364214 3421. 0.132636734329931
3422. 0.0839003806121035 3423. 0.0954766055881712 3424. 5.70050229186318e-06 3425. 0.000478471454783512
3426. 3.13226817396444e-07 3427. 0.0722595850746881 3428. 0.127966455744009 3429. 0.0317020517667026
3430. 0.22072473359238 3431. 0.125930897923838 3432. 0.0307514705368292 3433. 0.002797624456339
3434. 0.0214612455946833 3435. 0.0479473656018151 3436. 0.0701848119178054 3437. 0.115121068267284
3438. 0.161456017794118 3439. 0.0603908667373027 3440. 0.0762727369427515 3441. 4.07406489699583e-05
3442. 0.115891587320809 3443. 0.000881249003350333 3444. 0.00143112833632565 3445. 0.0884550974293198
3446. 0.0144527268421997 3447. 0.115430054472231 3448. 0.100514627587079 3449. 0.0531642658517812
3450. 0.207658761426446 3451. 0.0628404861522389 3452. 0.0170184337709035 3453. 0.1119498074657
3454. 0.0666290539678073 3455. 0.135199620383822 3456. 0.0133905490325536 3457. 0.0725003096339942
3458. 0.0702314636154333 3459. 0.132323759448875 3460. 0.131420239706172 3461. 0.00766589254617829
3462. 0.1120380225293 3463. 0.193541349649673 3464. 0.0301128817534438 3465. 0.0165150930720599
3466. 0.0622683167769484 3467. 0.160811667190507 3468. 0.10768702773869 3469. 0.0154447172627932
3470. 2.69294232045684e-06 3471. 0.153624721169103 3472. 0.11291928839026 3473. 0.132011948210495
3474. 0.0996052776345528 3475. 0.110228915988427 3476. 0.123299919592454 3477. 0.0177988991267658
3478. 0.0531420369038825 3479. 0.13384174705464 3480. 0.117989342193648 3481. 0.115810521437997
3482. 0.0977752700618754 3483. 0.0703113496006928 3484. 0.00595657765896047 3485. 0.0569711485227335
3486. 0.0935682182612665 3487. 0.000190709157628346 3488. 0.065691255630001 3489. 0.15867416894576
3490. 0.000640608667893719 3491. 0.0593490390520557 3492. 0.139769551792485 3493. 0.155529901702058
3494. 0.135962308959702 3495. 0.137642626084755 3496. 0.0436246919087472 3497. 0.022329909907339
3498. 0.0118262440781171 3499. 0.0400558141774716 3500. 0.198880969513284 3501. 0.071890046272656
3502. 0.212552266942174 3503. 0.00897565897887979 3504. 0.146389093218431 3505. 0.0911358967548288
3506. 0.0311661538973198 3507. 0.121728401735662 3508. 0.0838225406426926 3509. 7.04363912358041e-05
3510. 0.0787151370268179 3511. 0.135855810951232 3512. 0.111395100788069 3513. 0.0375430636451105
3514. 0.0214451714135428 3515. 0.0849755235825456 3516. 0.19810135570553 3517. 0.214058587722386
3518. 0.0499201418086119 3519. 0.121795303186554 3520. 0.0479119848506843 3521. 0.00984847332139165
3522. 0.0154024930323064 3523. 0.189183494299134 3524. 0.16409229595037 3525. 0.0548505303990657
3526. 0.00472760359270932 3527. 0.0144589627919433 3528. 0.095377875153543 3529. 0.107130098931859
3530. 0.00646757632442479 3531. 0.0385506575052127 3532. 0.101989522387843 3533. 0.123398670513184
3534. 0.0445723604191613 3535. 0.00708380333656784 3536. 0.0199517854145515 3537. 0.0473299899339162
3538. 0.110305493682775 3539. 0.110983544841272 3540. 0.0428812896963631 3541. 0.0969331538426617
3542. 0.0693984871942778 3543. 0.112815414135567 3544. 0.00438944329029737 3545. 0.0518586265174763
3546. 0.000370370822394855 3547. 0.000291004098277332 3548. 1.64636976105961e-06 3549. 0.07200239806091
3550. 0.0795022012569347 3551. 0.110659127857515 3552. 0.00036906753478554 3553. 0.0590682764958098
3554. 0.151070097371163 3555. 0.0764438649610778 3556. 0.177204096431278 3557. 0.000754429002101991
3558. 0.101759524878005 3559. 0.128195221256176 3560. 0.0653044420789327 3561. 0.024851616577239
3562. 0.00407971718019351 3563. 0.221537608834213 3564. 0.216095381575936 3565. 0.129673370502063
3566. 0.0311701125743756 3567. 0.0806355555696028 3568. 0.0395165313844915 3569. 0.207222891942127
3570. 0.0011811664036589 3571. 0.20260070338569 3572. 0.0119399075073749 3573. 2.21884937929732e-08
3574. 0.146714025143062 3575. 0.0484304295550643 3576. 0.000215927096214122 3577. 2.64762634448856e-05
3578. 0.129133134684042 3579. 0.0742753721393637 3580. 0.0983916413541819 3581. 0.00353301239145605
3582. 0.217058699439566 3583. 0.0472917020341308 3584. 0.1031157405074 3585. 0.244998217119302
3586. 0.094765359693496 3587. 0.0554390510280213 3588. 0.0955127046077512 3589. 3.37457507593244e-06
3590. 0.128920703638444 3591. 0.123466807775555 3592. 0.0133386860716123 3593. 0.0955958553995933
3594. 0.0776445387975979 3595. 0.114704003331366 3596. 0.0268391148705145 3597. 0.188126572374704
3598. 0.0991315362476562 3599. 5.50543724082611e-05 3600. 0.089120004945207 3601. 0.176953778117839
3602. 0.1340387034496 3603. 0.072641137138791 3604. 4.78418780448209e-05 3605. 0.0247553734022383

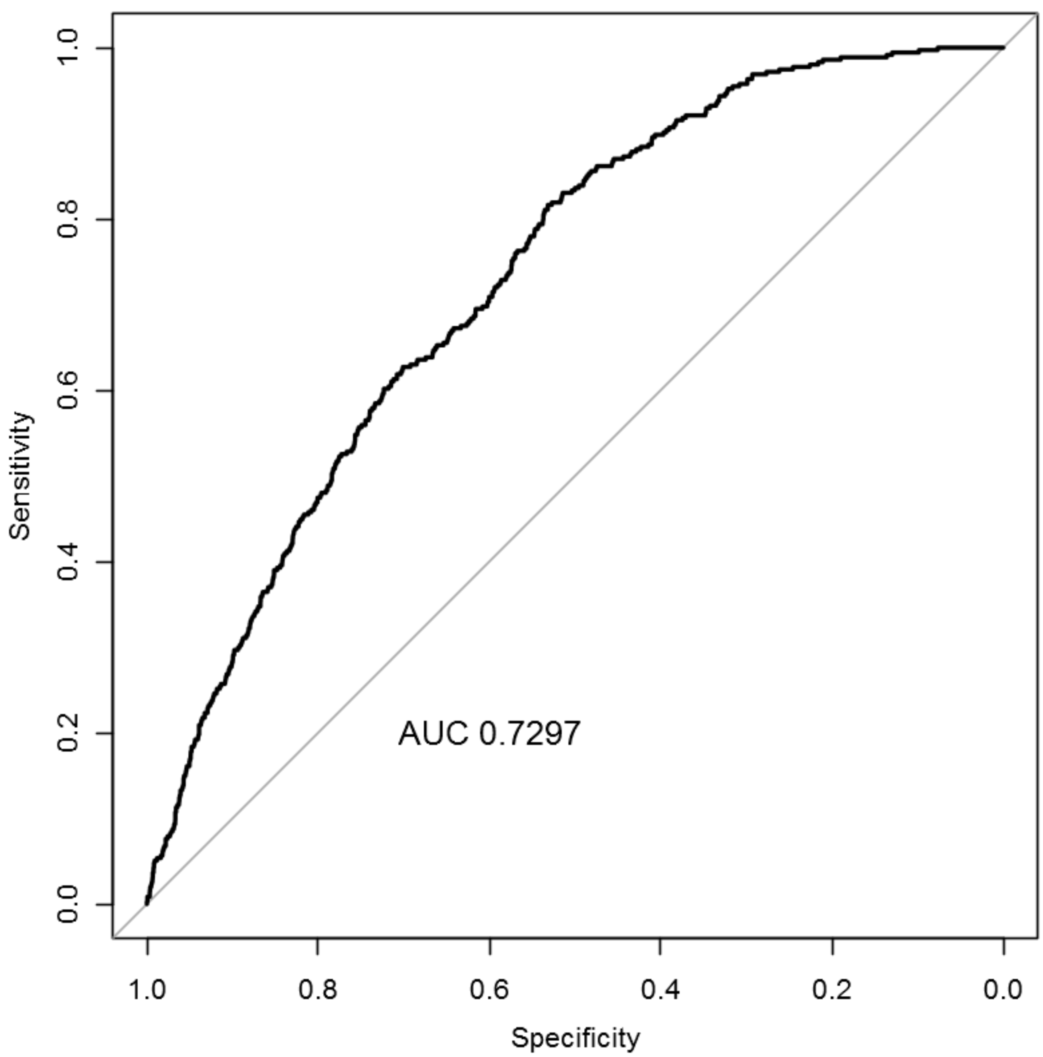
3606. 0.0547809085931345 3607. 0.0830033478597822 3608. 0.0581760508222719 3609. 0.0395244799188041 3610. 0.0983507427539001 3611. 0.0471833241217635 3612. 0.0857006778109695 3613. 0.0782580705324895 3614. 0.154288259965349 3615. 0.124898754434107 3616. 0.0562734467688876 3617. 1.25898383914651e-05 3618. 0.129026316893499 3619. 0.00622612228738438 3620. 0.154950367967547 3621. 0.102545353939842 3622. 0.0060653279412865 3623. 0.028344804449464 3624. 0.144489839289378 3625. 0.145636509486855 3626. 0.182183747393188 3627. 0.137461867563431 3628. 0.123037137930629 3629. 0.186241489509814 3630. 0.0103371094335435 3631. 0.0920040083599195 3632. 0.00801259827220425 3633. 0.0793321889304545 3634. 0.0429467578176204 3635. 0.032005589457977 3636. 0.135078554968512 3637. 0.00804155270586953 3638. 0.0695767903018929 3639. 0.0868902947310243 3640. 0.00833268081627328 3641. 0.201345523839111 3642. 0.00329757050572958 3643. 0.106235940809332 3644. 0.0578960543293345 3645. 0.178810919861844 3646. 0.109472708046718 3647. 0.111239802297677 3648. 0.0847437523148552 3649. 0.0881626561047546 3650. 0.0489093942237208 3651. 0.0555954849735504 3652. 0.125082224234951 3653. 0.217348101648833 3654. 3.77453843488917e-05 3655. 0.000434063029777052 3656. 0.0442051544275936 3657. 0.0111308593206375 3658. 0.050252710906739 3659. 0.17407651110159 3660. 0.00352629796725853 3661. 0.0691648552447488 3662. 0.0160344785405365 3663. 0.00632142503595582 3664. 0.00275189760388385 3665. 0.0118565066405449 3666. 0.120291886731497 3667. 0.0826096105272194 3668. 0.0597561641933947 3669. 0.0886511436189296 3670. 0.106636270169144 3671. 0.117523266197185 3672. 0.0854908988815136 3673. 0.119839252330057 3674. 0.13469253348233 3675. 0.146349861823319 3676. 0.129040629330503 3677. 0.0380723351480771 3678. 0.197126354390823 3679. 0.130531510322541 3680. 0.141573891414284 3681. 2.47072900643319e-06 3682. 5.05801986221827e-12 3683. 0.160874751230055 3684. 0.0130278052809087 3685. 0.156424519168853 3686. 0.0431706021385955 3687. 0.167725668156745 3688. 0.0329634277484425 3689. 0.0900625974774133 3690. 0.0472199799154547 3691. 0.106269354509136 3692. 0.0305399412689747 3693. 0.108076483952743 3694. 0.15057948116061 3695. 0.000400592828420038 3696. 0.262784213891133 3697. 0.0601342212530846 3698. 0.0789937675910654 3699. 0.00712999731904766 3700. 0.084931621211061 3701. 0.0798952913383591 3702. 9.85906478237219e-11 3703. 0.000390622151641622 3704. 0.000275961882423262 3705. 0.101681145508036 3706. 0.049337056828393 3707. 7.19353157874883e-12 3708. 0.00541230867531682 3709. 0.0429879641102183 3710. 0.116111973781688 3711. 0.104075126429935 3712. 0.0619168540896346 3713. 0.132613416239165 3714. 0.0110375266636061 3715. 0.146253021244301 3716. 0.123576467812992 3717. 0.0998821634795586 3718. 0.0587230470014263 3719. 0.0927243761791674 3720. 0.151224575341267 3721. 0.204304654413075 3722. 0.045720933041452 3723. 0.039150306221659 3724. 0.2002060834664 3725. 0.000743782085070824 3726. 0.121469493855542 3727. 0.224094907519568 3728. 0.157712851923218 3729. 0.0333291058153329 3730. 0.110120000834067 3731. 0.096985544790372 3732. 0.15962271314357 3733. 0.0106662813967996 3734. 0.113553660117426 3735. 0.025084694039649 3736. 0.03363891114251 3737. 0.193803338166788 3738. 0.00156726414778945 3739. 0.170444581382865 3740. 0.102370313544366 3741. 0.0955641357601637 3742. 0.110946007564115 3743. 0.147978386427318 3744. 0.139563154178447

In [202]:
```
# logit.best1 from stepAIC
logit.pred1.test <- predict(logit.best1,type='response',newdata=test[,-c(1,2,11)])
library(pROC)
# ROC
(roc1 <- roc(test$is_booking,logit.pred1.test,plot=TRUE))
par(cex=1.2)
text(0.6,0.2,"AUC 0.7297")
# Confusion Table and Accuracy
logit.pred1.test <- ifelse(logit.pred1.test > 0.35, 1, 0) #Best thereshold: 0.35
(confTable1 <- table(logit.pred1.test, test$is_booking))
(acc1 <- sum(diag(confTable1))/sum(confTable1)) #Accuracy: 0.911705852926463
```

```
Call:
roc.default(response = test$is_booking, predictor = logit.pred1.test,     plot = TRUE)

Data: logit.pred1.test in 3644 controls (test$is_booking 0) < 354 cases (test$is_booking 1).
Area under the curve: 0.7297
```

```
logit.pred1.test    0    1
               0 3644  353
               1    0    1
```

0.911705852926463

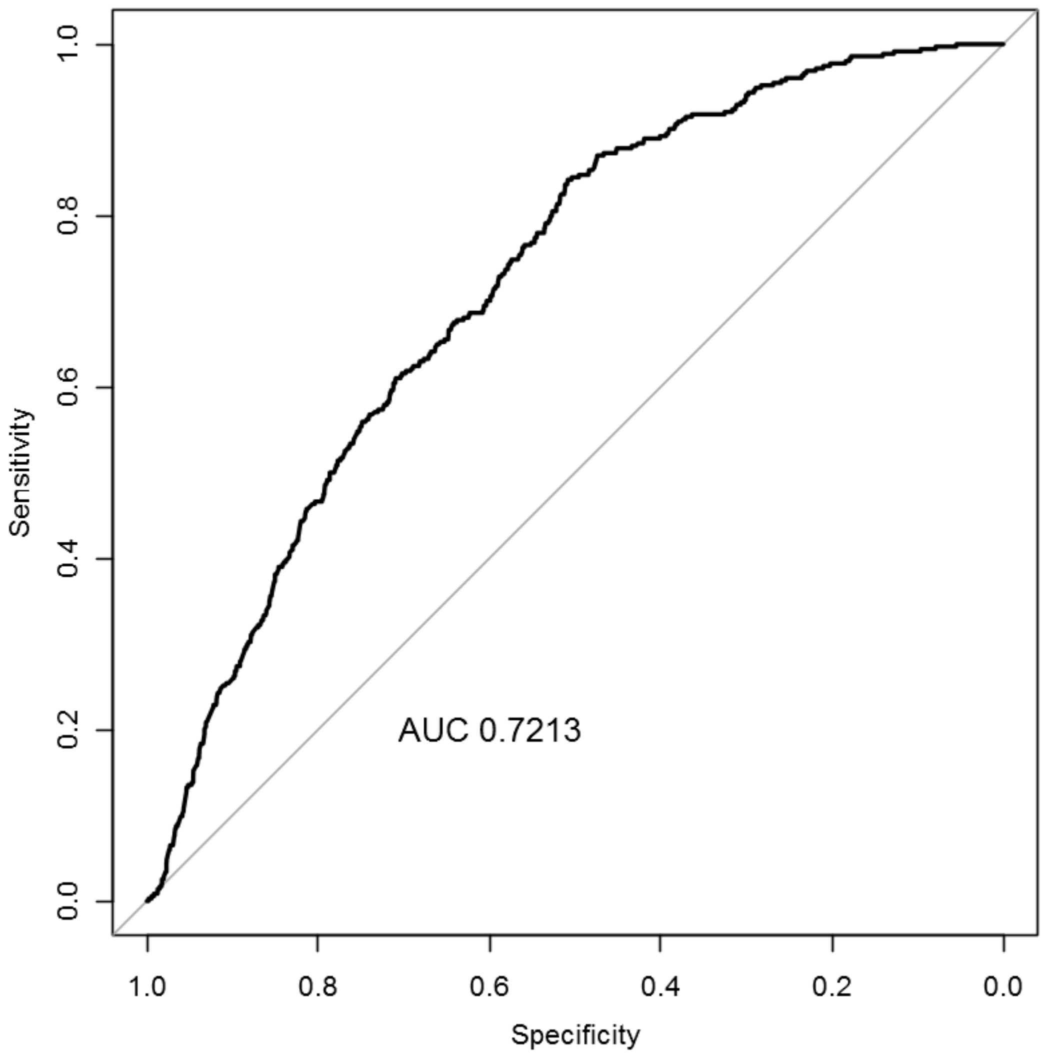


```
In [136]: library(pROC)
          lasso.pred <- predict(lasso.fit, s=lambda, newx=test.xmat)
          lasso.pred <- as.numeric(lasso.pred)
          # ROC
          (lasso.roc <- roc(test$is_booking,lasso.pred,plot=TRUE)) # AUC: 0.7213
          par(cex=1.2)
          text(0.6,0.2,"AUC 0.7213")
          # Confusion Table and Accuracy
          lasso.pred <- ifelse(lasso.pred > 0.3, 1, 0) #Best thereshold: 0.3
          (lasso.confTable <- table(lasso.pred, test$is_booking))
          (lasso.acc <- sum(diag(lasso.confTable))/sum(lasso.confTable)) #Accuracy: 0.91120560
```

```
Call:
roc.default(response = test$is_booking, predictor = lasso.pred,     plot = TRUE)

Data: lasso.pred in 3644 controls (test$is_booking 0) < 354 cases (test$is_booking 1).
Area under the curve: 0.7213
```

```
lasso.pred    0    1
         0 3643  354
         1    1    0
```

0.911205602801401



4.5 Interpretation

From the results, we can teel that logit.best1 is better than lasso.fit. the AUC logit.best1 vs lasso.fit is 0.7297 vs 0.7213 and the accuracy of logit.best1 is better than lasso.fit, which is 0.9117 vs 0.9112. So model logit.best1 is better than lasso.fit.

The best model is: log(pr(is_booking=1)/(1-pr(is_booking=1))) = 1.804052 - 0.175075*is_mobile1 - 0.697306*is_package1 - 0.051581*channel262 - 0.380355*channel293 + 0.081007*channel324 + 0.557752*channel355 + 0.181859*channel386 - 10.735697*channel417 - 1.228628*channel448 + 0.805183*channel479 - 0.075253*channel510 - 0.018351*channel541 - 0.088386*srch_adults_cnt + 0.299662*prop_is_branded1 - 0.185572*prop_starrating - 0.184832*distance_band1 + 0.054592*distance_band2 - 0.103447*distance_band3 + 0.152857*distance_band4 + 0.124505*hist_price_band1 + 0.002048*hist_price_band2 + 0.113960*hist_price_band3 - 0.204585*hist_price_band4 -

*0.795968*popularity_band1 - 0.092916*popularity_band2 + 0.274264*popularity_band3 - 0.717964*popularity_band4 - 2.643898*cnt - 0.192581*days

From the model, we can tell that: 1. is_mobile,is_package,srch_adults_cnt,prop_starrating, cnt, living days have negative effect to log of odds ratio of is_booking. 2. varibles: channel,distance_band,hist_price_band,popularity_band have both positive and negative effect to the log of odds ratio of is_booking.

5 Problem 4.

(25 points) We are interested in finding the important predictors of number of non-spinal bone fractures in women with low bone densities, assess their adjusted effect sizes (in direction and magnitude) and use the best Poisson regression model for interpretation and prediction. We want to analyze the FITglm2.txt dataset (available on Blackboard, use the command read.delim("…\FITglm2.txt",sep="")) that contains 6459 observations on 18 variables: alloc - id, ra_age –age in years, frx – indicator for spinal fractures, nosp – indicator for non-spinal fractures, numnosp – number of non-spinal fractures (outcome variable), trt01 – indicator for treatment, p3_weigh – weight over 100 pounds, htotbmd – bone mass density 1, nbmd – bone mass density 2, trialyrs – duration of follow-up, riskcat4 – risk category, tneck – bone density at the neck, bmd25 – indicator for osteoporosis (based on tneck values), hplac – indicator for high placebo dose, htrt – indicator for high dose treatment, lplac – indicator for low placebo dose, ltrt- indicator for lose dose treatment, rtgroup – risk group for falling. Perform all necessary data analysis steps and write a section summarizing the findings.

5.1 Load data and Basic Analysis

In [30]:
```
# load data
fitglm2_data <- read.table("FITglm2.txt",sep="\t",header=T)
dim(fitglm2_data)
# omit rows which has NA
fitglm2_data <- na.omit(fitglm2_data)
dim(fitglm2_data)
head(fitglm2_data)
```

1. 6459 2. 18

1. 6366 2. 18

alloc	ra_age	frx	nosp	numnosp	trt01	p3_weigh	htotbmd	nbmd	trialyrs	riskcat4
1	69	1	0	0	0	66.6	0.517	0.464	3.022587	1: HIGH RISK
2	76	0	0	0	1	69.0	0.583	0.518	3.003422	1: HIGH RISK
3	66	0	0	0	1	61.6	0.709	0.602	2.869268	0: LOW RISK
4	72	0	0	0	0	57.6	0.738	0.653	3.236140	0: LOW RISK
5	58	0	0	0	1	75.2	0.690	0.633	2.992471	0: LOW RISK
6	74	0	0	0	1	50.2	0.480	0.485	2.970568	0: LOW RISK

In [31]:
```
str(fitglm2_data)
summary(fitglm2_data)
```

```
'data.frame':        6366 obs. of  18 variables:
 $ alloc   : int  1 2 3 4 5 6 7 8 9 10 ...
```

```
$ ra_age  : int  69 76 66 72 58 74 66 63 62 69 ...
$ frx     : int  1 0 0 0 0 0 0 0 0 0 ...
$ nosp    : int  0 0 0 0 0 0 0 0 0 0 ...
$ numnosp : int  0 0 0 0 0 0 0 0 0 0 ...
$ trt01   : int  0 1 1 0 1 1 0 1 0 0 ...
$ p3_weigh: num  66.6 69 61.6 57.6 75.2 50.2 86.6 54 74 64.8 ...
$ htotbmd : num  0.517 0.583 0.709 0.738 0.69 0.48 0.816 0.538 0.726 0.514 ...
$ nbmd    : num  0.464 0.518 0.602 0.653 0.633 0.485 0.658 0.431 0.584 0.47 ...
$ trialyrs: num  3.02 3 2.87 3.24 2.99 ...
$ riskcat4: Factor w/ 3 levels "","0: LOW RISK",..: 3 3 2 2 2 2 2 2 2 2 ...
$ tneck   : num  -3.28 -2.83 -2.13 -1.71 -1.88 ...
$ bmd25   : int  1 1 0 0 0 1 0 1 0 1 ...
$ hplac   : int  1 0 0 0 0 0 0 0 0 0 ...
$ htrt    : int  0 1 0 0 0 0 0 0 0 0 ...
$ lplac   : int  0 0 0 1 0 0 1 0 1 1 ...
$ ltrt    : int  0 0 1 0 1 1 0 1 0 0 ...
$ rtgroup : Factor w/ 5 levels "","1:HIGH FALL RISK, PLACEBO GROUP",..: 2 3 5 4 5 5 4 5 4 4 .
- attr(*, "na.action")=Class 'omit'  Named int [1:93] 211 283 482 556 732 750 879 949 952 105
 .. ..- attr(*, "names")= chr [1:93] "211" "283" "482" "556" ...


    alloc          ra_age           frx              nosp
Min.   :   1   Min.   :54.00   Min.   :0.0000   Min.   :0.000
1st Qu.:2062   1st Qu.:64.00   1st Qu.:0.0000   1st Qu.:0.000
Median :3800   Median :68.00   Median :0.0000   Median :0.000
Mean   :3784   Mean   :68.12   Mean   :0.1407   Mean   :0.128
3rd Qu.:5583   3rd Qu.:73.00   3rd Qu.:0.0000   3rd Qu.:0.000
Max.   :7230   Max.   :81.00   Max.   :1.0000   Max.   :1.000
   numnosp          trt01           p3_weigh         htotbmd
Min.   :0.0000   Min.   :0.0000   Min.   : 36.30   Min.   :0.3700
1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.: 56.90   1st Qu.:0.6350
Median :0.0000   Median :1.0000   Median : 63.10   Median :0.6980
Mean   :0.1528   Mean   :0.5009   Mean   : 64.54   Mean   :0.6925
3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.: 70.80   3rd Qu.:0.7540
Max.   :4.0000   Max.   :1.0000   Max.   :124.60   Max.   :0.9860
     nbmd           trialyrs               riskcat4        tneck
Min.   :0.3370   Min.   :0.005476              :   0   Min.   :-4.342
1st Qu.:0.5420   1st Qu.:3.014374   0: LOW RISK :5239   1st Qu.:-2.633
Median :0.5900   Median :4.027379   1: HIGH RISK:1127   Median :-2.233
Mean   :0.5842   Mean   :3.795054                       Mean   :-2.282
3rd Qu.:0.6350   3rd Qu.:4.484600                       3rd Qu.:-1.858
Max.   :0.7830   Max.   :4.821355                       Max.   :-0.625
    bmd25            hplac              htrt              lplac
Min.   :0.0000   Min.   :0.00000   Min.   :0.00000   Min.   :0.0000
1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000
Median :0.0000   Median :0.00000   Median :0.00000   Median :0.0000
Mean   :0.3227   Mean   :0.08891   Mean   :0.08812   Mean   :0.4101
3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:1.0000
```

```
 Max.   :1.0000   Max.   :1.00000   Max.   :1.00000   Max.   :1.0000
     ltrt                                        rtgroup
 Min.   :0.0000                                      :   0
 1st Qu.:0.0000   1:HIGH FALL RISK, PLACEBO GROUP  : 566
 Median :0.0000   2:HIGH FALL RISK, TREATMENT GROUP: 561
 Mean   :0.4128   3:LOW FALL RISK, PLACEBO GROUP   :2611
 3rd Qu.:1.0000   4:LOW FALL RISK, TREATMENT GROUP :2628
 Max.   :1.0000
```

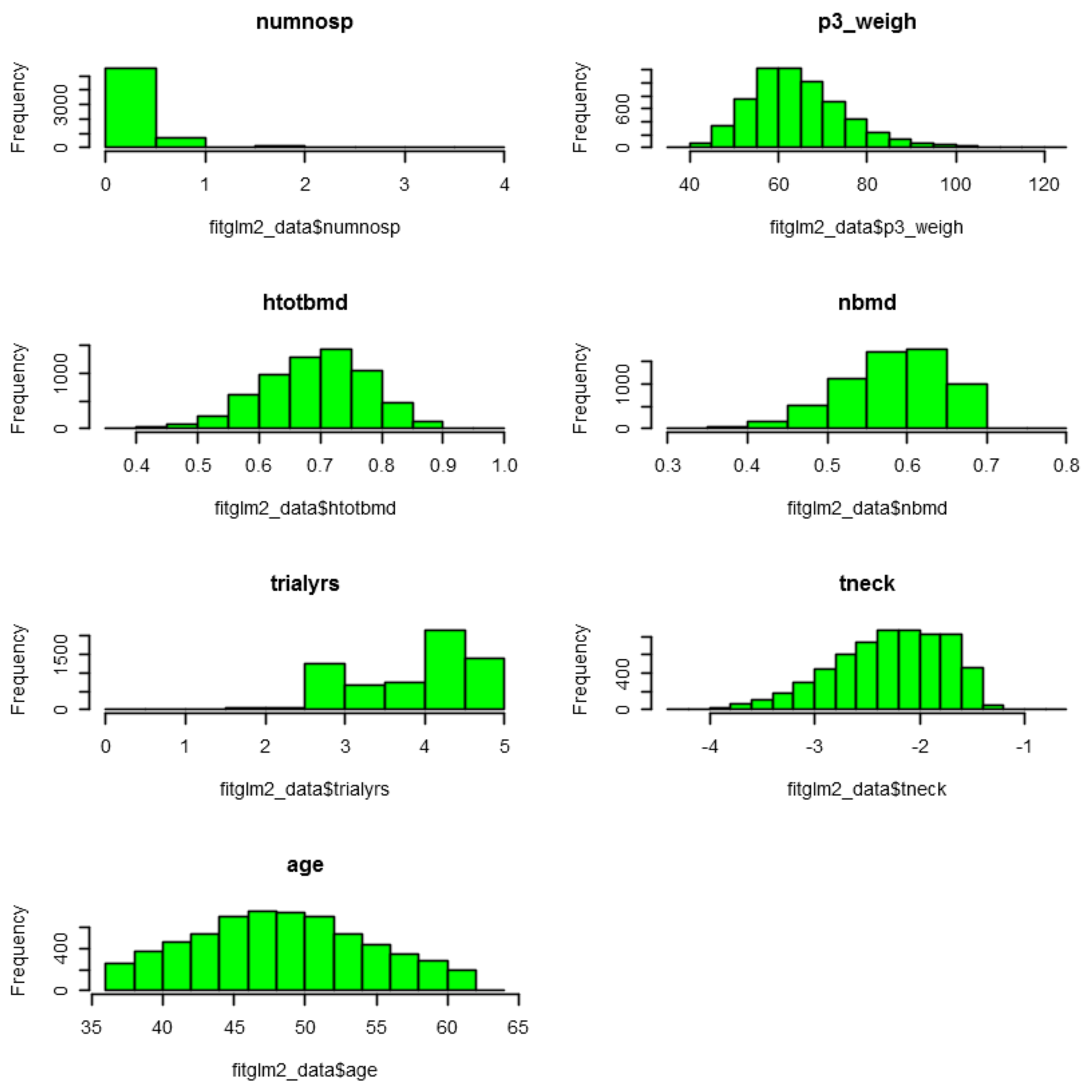
In [38]:
```
# change rg_age into age
fitglm2_data$age <- (117-fitglm2_data$ra_age)
# change indicator variables to factor
fitglm2_data[,c(3,4,6,13,14,15,16,17)]<-lapply(fitglm2_data[,c(3,4,6,13,14,15,16,17)]
# change riskcat4 and rtgroup to numeric factors.
fitglm2_data$riskcat4 <- as.factor(as.numeric(fitglm2_data$riskcat4)-2)
fitglm2_data$rtgroup <- as.factor(as.numeric(fitglm2_data$rtgroup)-2)
# remove alloc and ra_age
fitglm2_data <- fitglm2_data[,-c(1,2)]
dim(fitglm2_data)
summary(fitglm2_data)
```

1. 6366 2. 17

```
 frx      nosp        numnosp        trt01       p3_weigh         htotbmd
 0:5470   0:5551   Min.   :0.0000   0:3177   Min.   : 36.30   Min.   :0.3700
 1: 896   1: 815   1st Qu.:0.0000   1:3189   1st Qu.: 56.90   1st Qu.:0.6350
                   Median :0.0000            Median : 63.10   Median :0.6980
                   Mean   :0.1528            Mean   : 64.54   Mean   :0.6925
                   3rd Qu.:0.0000            3rd Qu.: 70.80   3rd Qu.:0.7540
                   Max.   :4.0000            Max.   :124.60   Max.   :0.9860
      nbmd           trialyrs          riskcat4     tneck          bmd25
 Min.   :0.3370   Min.   :0.005476   0:5239   Min.   :-4.342   0:4312
 1st Qu.:0.5420   1st Qu.:3.014374   1:1127   1st Qu.:-2.633   1:2054
 Median :0.5900   Median :4.027379            Median :-2.233
 Mean   :0.5842   Mean   :3.795054            Mean   :-2.282
 3rd Qu.:0.6350   3rd Qu.:4.484600            3rd Qu.:-1.858
 Max.   :0.7830   Max.   :4.821355            Max.   :-0.625
 hplac    htrt     lplac    ltrt     rtgroup       age
 0:5800   0:5805   0:3755   0:3738   0: 566   Min.   :36.00
 1: 566   1: 561   1:2611   1:2628   1: 561   1st Qu.:44.00
                                     2:2611   Median :49.00
                                     3:2628   Mean   :48.88
                                              3rd Qu.:53.00
                                              Max.   :63.00
```

In [41]:
```
par(mfrow=c(4,2))
hist(fitglm2_data$numnosp,col='green',main='numnosp')
```

```
hist(fitglm2_data$p3_weigh,col='green',main='p3_weigh')
hist(fitglm2_data$htotbmd,col='green',main='htotbmd')
hist(fitglm2_data$nbmd,col='green',main='nbmd')
hist(fitglm2_data$trialyrs,col='green',main='trialyrs')
hist(fitglm2_data$tneck,col='green',main='tneck')
hist(fitglm2_data$age,col='green',main='age')
```



5.2 Split the dataset into training and testing dataset

In [43]:
```
# train dataset: 80%, test dataset: 20%
(n <- dim(fitglm2_data)[1])
trainID <- sample(1:n,n*0.8)
train <- fitglm2_data[trainID,]
```

```
test <- fitglm2_data[-trainID,]
dim(train)
dim(test)
```

6366
1. 5092 2. 17
1. 1274 2. 17

5.3 Build Models

In [111]:
```
poreg.fit1 <- lm(log(numnosp+1)~., data=train)
# use stepAIC to get the best poisson regression model: poreg.best1.
library(MASS)
poreg.best1 <- stepAIC(poreg.fit1,data=train,trace=F)
summary(poreg.best1) #Multiple R-squared:0.9481,Adjusted R-squared:0.9481
# p3_weigh is not significant, remove it
```

```
Call:
lm(formula = log(numnosp + 1) ~ nosp + p3_weigh + bmd25 + hplac,
    data = train)

Residuals:
     Min       1Q   Median       3Q      Max
-0.08201 -0.00254  0.00024  0.00220  0.84796

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.022e-02  5.386e-03  -1.898   0.0578 .
nosp1        7.625e-01  2.513e-03 303.483   <2e-16 ***
p3_weigh     1.273e-04  7.938e-05   1.604   0.1088
bmd251       4.468e-03  1.868e-03   2.392   0.0168 *
hplac1       6.826e-03  2.921e-03   2.337   0.0195 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05974 on 5087 degrees of freedom
Multiple R-squared:  0.9481,Adjusted R-squared:  0.9481
F-statistic: 2.323e+04 on 4 and 5087 DF,  p-value: < 2.2e-16
```

In [117]:
```
# remove p3_weigh from model
poreg.best1 <- lm(formula = log(numnosp + 1) ~ nosp+bmd25+hplac, data = train)
summary(poreg.best1) # Multiple R-squared: 0.9481,Adjusted R-squared:0.948
```

```
Call:
lm(formula = log(numnosp + 1) ~ nosp + bmd25 + hplac, data = train)
```

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.07846 -0.00184  0.00176  0.00176  0.84852 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.001763   0.001087  -1.622   0.1048    
nosp1        0.762680   0.002511 303.701   <2e-16 ***
bmd251       0.003600   0.001788   2.013   0.0441 *  
hplac1       0.007088   0.002917   2.430   0.0151 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05975 on 5088 degrees of freedom
Multiple R-squared:  0.9481,Adjusted R-squared:  0.948 
F-statistic: 3.096e+04 on 3 and 5088 DF,  p-value: < 2.2e-16
```

```
In [116]: # add two way interaction
          poreg.best2 <- stepAIC(poreg.fit1, ~.^2, data=train,trace=F)
          summary(poreg.best2) #Multiple R-squared:0.949,Adjusted R-squared:0.9489
```

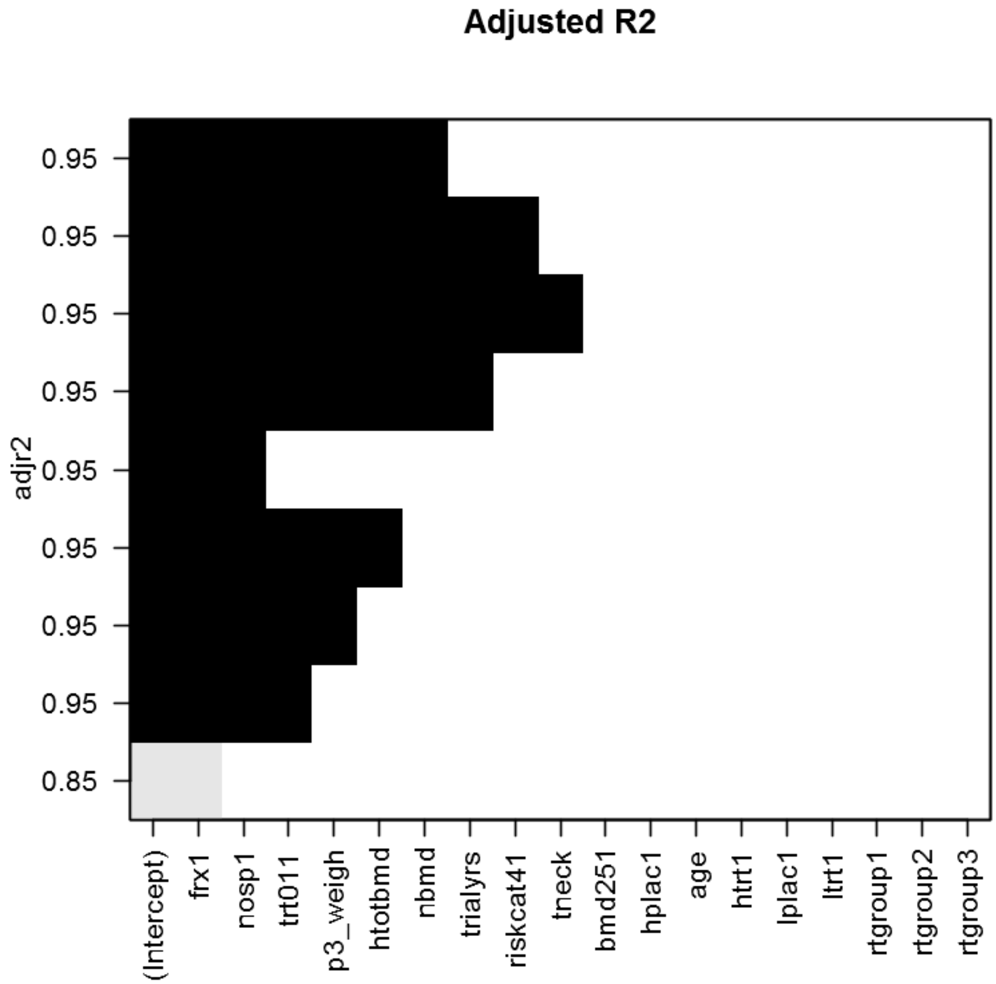
```
Call:
lm(formula = log(numnosp + 1) ~ nosp + trt01 + p3_weigh + bmd25 + 
    hplac + nosp:hplac + nosp:bmd25 + nosp:p3_weigh + trt01:p3_weigh + 
    p3_weigh:bmd25 + p3_weigh:hplac, data = train)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.15106 -0.00191 -0.00040  0.00082  0.85543 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      0.0127413  0.0090299   1.411   0.1583    
nosp1            0.6786481  0.0165184  41.084  < 2e-16 ***
trt011          -0.0185642  0.0105840  -1.754   0.0795 .  
p3_weigh        -0.0001963  0.0001365  -1.438   0.1506    
bmd251          -0.0221942  0.0112788  -1.968   0.0491 *  
hplac1           0.0249963  0.0170851   1.463   0.1435    
nosp1:hplac1     0.0455534  0.0081987   5.556 2.90e-08 ***
nosp1:bmd251     0.0292358  0.0052909   5.526 3.45e-08 ***
nosp1:p3_weigh   0.0010332  0.0002407   4.293 1.79e-05 ***
trt011:p3_weigh  0.0002907  0.0001622   1.792   0.0732 .  
p3_weigh:bmd251  0.0003605  0.0001807   1.995   0.0461 *  
p3_weigh:hplac1 -0.0003689  0.0002542  -1.451   0.1468    
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05926 on 5080 degrees of freedom
Multiple R-squared:  0.949,Adjusted R-squared:  0.9489
F-statistic:  8594 on 11 and 5080 DF,  p-value: < 2.2e-16
```

```
In [115]: # use regsubsets() to get the best model, which has the best R2 or ajusted-R2.
          library(leaps)
          poreg.best3 <- regsubsets(log(numnosp + 1)~., data=train, nbest=1, method='exhaustive
          summary(poreg.best3)
          plot(poreg.best3, scale='adjr2', main='Adjusted R2')
```

```
Warning message in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.in = force.in
"6  linear dependencies found"

Reordering variables and trying again:


Warning message in leaps.exhaustive(a, really.big):
"XHAUST returned error code -999"

Subset selection object
Call: regsubsets.formula(log(numnosp + 1) ~ ., data = train, nbest = 1,
    method = "exhaustive")
18 Variables  (and intercept)
          Forced in Forced out
frx1          FALSE      FALSE
nosp1         FALSE      FALSE
trt011        FALSE      FALSE
p3_weigh      FALSE      FALSE
htotbmd       FALSE      FALSE
nbmd          FALSE      FALSE
trialyrs      FALSE      FALSE
riskcat41     FALSE      FALSE
tneck         FALSE      FALSE
bmd251        FALSE      FALSE
hplac1        FALSE      FALSE
age           FALSE      FALSE
htrt1         FALSE      FALSE
lplac1        FALSE      FALSE
ltrt1         FALSE      FALSE
rtgroup1      FALSE      FALSE
rtgroup2      FALSE      FALSE
rtgroup3      FALSE      FALSE
1 subsets of each size up to 9
Selection Algorithm: exhaustive
```

```
         frx1 nosp1 trt011 p3_weigh htotbmd nbmd trialyrs riskcat41 tneck
1  ( 1 ) "*"  " "   " "    " "      " "     " "  " "      " "       " "
2  ( 1 ) "*"  "*"   " "    " "      " "     " "  " "      " "       " "
3  ( 1 ) "*"  "*"   "*"    " "      " "     " "  " "      " "       " "
4  ( 1 ) "*"  "*"   "*"    "*"      " "     " "  " "      " "       " "
5  ( 1 ) "*"  "*"   "*"    "*"      "*"     " "  " "      " "       " "
6  ( 1 ) "*"  "*"   "*"    "*"      "*"     "*"  " "      " "       " "
7  ( 1 ) "*"  "*"   "*"    "*"      "*"     "*"  "*"      " "       " "
8  ( 1 ) "*"  "*"   "*"    "*"      "*"     "*"  "*"      "*"       " "
9  ( 1 ) "*"  "*"   "*"    "*"      "*"     "*"  "*"      "*"       "*"
         bmd251 hplac1 htrt1 lplac1 ltrt1 rtgroup1 rtgroup2 rtgroup3 age
1  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
2  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
3  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
4  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
5  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
6  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
7  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
8  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
9  ( 1 ) " "    " "    " "   " "    " "   " "      " "      " "      " "
```



In [137]:
```
# from the plot, we get the poreg.best3
poreg.best3 <- lm(log(numnosp+1)~frx+nosp+trt01+p3_weigh+htotbmd+nbmd, data=train)
summary(poreg.best3)
#Multiple R-squared:0.948,Adjusted R-squared: 0.948
# variables frx, trt01, p3_weigh, htotbmd are not significant, remove them
```

```
Call:
lm(formula = log(numnosp + 1) ~ frx + nosp + trt01 + p3_weigh +
    htotbmd + nbmd, data = train)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-0.07643 -0.00219 -0.00026  0.00146  0.84822

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.864e-03  8.206e-03   1.080   0.2801
frx1        -1.098e-03  7.505e-03  -0.146   0.8837
nosp1        7.638e-01  7.788e-03  98.063   <2e-16 ***
trt011      -1.250e-03  1.677e-03  -0.745   0.4561
p3_weigh     1.309e-04  8.277e-05   1.582   0.1137
htotbmd      1.142e-02  1.548e-02   0.738   0.4607
nbmd        -4.199e-02  2.072e-02  -2.027   0.0427 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05978 on 5085 degrees of freedom
Multiple R-squared:  0.948,Adjusted R-squared:  0.948
F-statistic: 1.546e+04 on 6 and 5085 DF,  p-value: < 2.2e-16
```

```
In [140]: poreg.best3 <- lm(formula = log(numnosp + 1) ~  nosp + nbmd, data = train)
          summary(poreg.best3)
          #Multiple R-squared:0.948,Adjusted R-squared: 0.948
          # nbmd is not significant, remove it
```

```
Call:
lm(formula = log(numnosp + 1) ~ nosp + nbmd, data = train)

Residuals:
     Min       1Q   Median       3Q      Max
-0.07433 -0.00144 -0.00006  0.00106  0.84859

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.012851   0.007749   1.658   0.0973 .
nosp1        0.762830   0.002514 303.425   <2e-16 ***
nbmd        -0.021941   0.013141  -1.670   0.0950 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05978 on 5089 degrees of freedom
Multiple R-squared:  0.948,Adjusted R-squared:  0.948
F-statistic: 4.638e+04 on 2 and 5089 DF,  p-value: < 2.2e-16
```

```
In [141]: poreg.best3 <- lm(formula = log(numnosp + 1) ~  nosp, data = train)
```

```
summary(poreg.best3)
#Multiple R-squared:0.948,Adjusted R-squared: 0.948
```

```
Call:
lm(formula = log(numnosp + 1) ~ nosp, data = train)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.07002  0.00000  0.00000  0.00000  0.84627 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.979e-15  8.975e-04     0.0        1    
nosp1       7.632e-01  2.506e-03   304.5   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0598 on 5090 degrees of freedom
Multiple R-squared:  0.948,Adjusted R-squared:  0.948 
F-statistic: 9.273e+04 on 1 and 5090 DF,  p-value: < 2.2e-16
```

5.4 Predict and Evaluate

In [142]:
```
# evaluate models
pchisq(deviance(poreg.best1),df.residual(poreg.best1),lower.tail=FALSE) # p-value=1>
pchisq(deviance(poreg.best2),df.residual(poreg.best2),lower.tail=FALSE) # p-value=1>
pchisq(deviance(poreg.best3),df.residual(poreg.best3),lower.tail=FALSE) # p-value=1>


# predict and evaluate
# poreg.best1
# train
poreg.pred1.train <- predict(poreg.best1,type='response',data=train[,-3])
# unscale the predict value
poreg.pred1.train <- round(exp(poreg.pred1.train))-1
(confTable <- table(train$numnosp,poreg.pred1.train))
(acc1.train <- sum(diag(confTable))/sum(confTable)) # train accuracy: 0.97977

# test
poreg.pred1.test <- predict(poreg.best1,type='response',newdata=test[,-3])
# unscale the predict value
poreg.pred1.test <- round(exp(poreg.pred1.test))-1
(confTable <- table(test$numnosp,poreg.pred1.test))
(acc1.test <- sum(diag(confTable))/sum(confTable)) # test accuracy: 0.97252
```

```r
# poreg.best2
# train
poreg.pred2.train <- predict(poreg.best2,type='response',data=train[,-3])
# unscale the predict value
poreg.pred2.train <- round(exp(poreg.pred2.train))-1
(confTable <- table(train$numnosp,poreg.pred2.train))
(acc2.train <- sum(diag(confTable))/sum(confTable)) # train accuracy: 0.97977

# test
poreg.pred2.test <- predict(poreg.best2,type='response',newdata=test[,-3])
# unscale the predict value
poreg.pred2.test <- round(exp(poreg.pred2.test))-1
(confTable <- table(test$numnosp,poreg.pred2.test))
(acc2.test <- sum(diag(confTable))/sum(confTable)) # test accuracy: 0.97252


# poreg.best3
# train
poreg.pred3.train <- predict(poreg.best3,type='response',data=train[,-3])
# unscale the predict value
poreg.pred3.train <- round(exp(poreg.pred3.train))-1
(confTable <- table(train$numnosp,poreg.pred3.train))
(acc3.train <- sum(diag(confTable))/sum(confTable)) # train accuracy: 0.97977

# test
poreg.pred3.test <- predict(poreg.best3,type='response',newdata=test[,-3])
# unscale the predict value
poreg.pred3.test <- round(exp(poreg.pred3.test))-1
(confTable <- table(test$numnosp,poreg.pred3.test))
(acc3.test <- sum(diag(confTable))/sum(confTable)) # test accuracy: 0.97252
```

1
1
1

```
 poreg.pred1.train
     0    1
0 4439    0
1    0  550
2    0   90
3    0   12
4    0    1
```

0.979772191673213

```
 poreg.pred1.test
     0    1
```

```
0 1112    0
1    0  127
2    0   30
3    0    4
4    0    1
```

0.972527472527473

```
 poreg.pred2.train
     0    1
0 4439    0
1    0  550
2    0   90
3    0   12
4    0    1
```

0.979772191673213

```
 poreg.pred2.test
     0    1
0 1112    0
1    0  127
2    0   30
3    0    4
4    0    1
```

0.972527472527473

```
 poreg.pred3.train
     0    1
0 4439    0
1    0  550
2    0   90
3    0   12
4    0    1
```

0.979772191673213

```
 poreg.pred3.test
     0    1
0 1112    0
1    0  127
2    0   30
3    0    4
4    0    1
```

0.972527472527473

5.5 Interpretation

From the results, we can see all of these three best models get the same R-squared, 0.948, which is very good. After predict, they get the same train and test accuracy, 0.97977 and 0.9725, very high. As they get the same accuracy, I would choose poreg.best3 as the best model. Because it is the simplest.

The best modesl is: log(numnosp + 1) = 1.979e-15 + 0.7632*nosp

From the equation, if the bone is non-spinal bone, the number of non-spinal bone fractures in women with low bone densities will be 1, otherwise 0.

```
In [146]: exp(1.979e-15 + 0.7632)-1
```

1.14512966406385