

STA363Proj3 - Formal Report

Rickey huang

4/19/2021

```
# Library everything int this chunk
# Packages for ggplot2
library(rlang)
library(ggplot2)
# Packages for fitting and plotting trees
library(rpart)
library(rattle)
library(rpart.plot)
# Packages for fitting forests
# Run the next line of code if the "randomForest" package is not installed before
#install.packages("randomForest")
library(randomForest)

# Load the data
McDonalds <- read.csv("~/Desktop/2021Spring/STA-363/Projects/Project 3/STA363Proj3/McDonaldsProj3.csv")
# Explore the correlation among the numeric data
correlationMcDonalds <- subset(McDonalds, select = -Category)
correlationMcDonalds <- subset(correlationMcDonalds, select = -Item)
cor(correlationMcDonalds)
```

	Serving.Size	Calories	Calories.from.Fat
## Serving.Size	1.000000000	-0.09094113	-0.29532194
## Calories	-0.090941125	1.00000000	0.88650366
## Calories.from.Fat	-0.295321937	0.88650366	1.00000000
## Total.Fat	-0.293214800	0.88631148	0.99957642
## Total.Fat....Daily.Value.	-0.292950027	0.88598103	0.99965532
## Saturated.Fat	-0.117542091	0.86092735	0.87913547
## Saturated.Fat....Daily.Value.	-0.121311710	0.86254130	0.88122178
## Trans.Fat	0.005474563	0.52213706	0.43079165
## Cholesterol	-0.172995984	0.59107455	0.69717032
## Cholesterol....Daily.Value.	-0.170400416	0.58934014	0.69603085
## Sodium	-0.434273783	0.67029946	0.82733245
## Sodium....Daily.Value.	-0.433743207	0.67156454	0.82809644
## Carbohydrates	0.248288399	0.79280148	0.44424962
## Carbohydrates....Daily.Value.	0.249673406	0.79315513	0.44500559
## Dietary.Fiber	-0.367819224	0.51971165	0.57268268
## Dietary.Fiber....Daily.Value.	-0.347837992	0.52131200	0.56668588
## Sugars	0.519137438	0.30900026	-0.09813392
## Protein	-0.309875597	0.74832150	0.76669383
## Vitamin.A....Daily.Value.	-0.010435278	0.13298056	0.08121478
## Vitamin.C....Daily.Value.	-0.099537343	-0.08123161	-0.10582588
## Calcium....Daily.Value.	0.169118081	0.48638588	0.20517644
## Iron....Daily.Value.	-0.448768769	0.65195423	0.76904876

##	Total.Fat	Total.Fat....Daily.Value.
## Serving.Size	-0.29321480	-0.29295003
## Calories	0.88631148	0.88598103
## Calories.from.Fat	0.99957642	0.99965532
## Total.Fat	1.00000000	0.99970351
## Total.Fat....Daily.Value.	0.99970351	1.00000000
## Saturated.Fat	0.87906694	0.88003266
## Saturated.Fat....Daily.Value.	0.88115313	0.88212756
## Trans.Fat	0.42836696	0.43020715
## Cholesterol	0.69553485	0.69612472
## Cholesterol....Daily.Value.	0.69440054	0.69497352
## Sodium	0.82684166	0.82756832
## Sodium....Daily.Value.	0.82757003	0.82831891
## Carbohydrates	0.44379487	0.44305096
## Carbohydrates....Daily.Value.	0.44455421	0.44380207
## Dietary.Fiber	0.57229714	0.57209948
## Dietary.Fiber....Daily.Value.	0.56632398	0.56620727
## Sugars	-0.09830377	-0.09865072
## Protein	0.76647650	0.76662694
## Vitamin.A....Daily.Value.	0.07870723	0.07831506
## Vitamin.C....Daily.Value.	-0.10816524	-0.10821357
## Calcium....Daily.Value.	0.20737587	0.20655245
## Iron....Daily.Value.	0.76793913	0.76902668
##	Saturated.Fat	Saturated.Fat....Daily.Value.
## Serving.Size	-0.1175421	-0.12131171
## Calories	0.8609274	0.86254130
## Calories.from.Fat	0.8791355	0.88122178
## Total.Fat	0.8790669	0.88115313
## Total.Fat....Daily.Value.	0.8800327	0.88212756
## Saturated.Fat	1.0000000	0.99926132
## Saturated.Fat....Daily.Value.	0.9992613	1.00000000
## Trans.Fat	0.6140406	0.61359964
## Cholesterol	0.6219996	0.62434136
## Cholesterol....Daily.Value.	0.6210108	0.62333417
## Sodium	0.5693729	0.57372974
## Sodium....Daily.Value.	0.5707178	0.57509306
## Carbohydrates	0.5805911	0.58051132
## Carbohydrates....Daily.Value.	0.5812554	0.58102759
## Dietary.Fiber	0.3328876	0.33776513
## Dietary.Fiber....Daily.Value.	0.3282027	0.33270202
## Sugars	0.2110766	0.20947094
## Protein	0.5946765	0.59781260
## Vitamin.A....Daily.Value.	0.0715950	0.07211378
## Vitamin.C....Daily.Value.	-0.1846851	-0.18315769
## Calcium....Daily.Value.	0.4171685	0.41523970
## Iron....Daily.Value.	0.5698631	0.57227341
##	Trans.Fat	Cholesterol
## Serving.Size	0.005474563	-0.17299598
## Calories	0.522137059	0.59107455
## Calories.from.Fat	0.430791654	0.69717032
## Total.Fat	0.428366960	0.69553485
## Total.Fat....Daily.Value.	0.430207153	0.69612472
## Saturated.Fat	0.614040559	0.62199959
## Saturated.Fat....Daily.Value.	0.613599639	0.62434136

## Trans.Fat	1.000000000	0.24094114
## Cholesterol	0.240941136	1.00000000
## Cholesterol....Daily.Value.	0.238363200	0.99985282
## Sodium	0.159086996	0.61632594
## Sodium....Daily.Value.	0.159893624	0.61673888
## Carbohydrates	0.453800987	0.25356683
## Carbohydrates....Daily.Value.	0.453538005	0.25554319
## Dietary.Fiber	0.035653316	0.42063157
## Dietary.Fiber....Daily.Value.	0.039230478	0.42553348
## Sugars	0.344842837	-0.12795838
## Protein	0.375789975	0.55369719
## Vitamin.A....Daily.Value.	0.080380899	0.08640586
## Vitamin.C....Daily.Value.	-0.078914128	-0.08625158
## Calcium....Daily.Value.	0.393889160	0.14093620
## Iron....Daily.Value.	0.315746903	0.64874813
##	Cholesterol....Daily.Value.	Sodium
## Serving.Size	-0.17040042	-0.434273783
## Calories	0.58934014	0.670299458
## Calories.from.Fat	0.69603085	0.827332454
## Total.Fat	0.69440054	0.826841664
## Total.Fat....Daily.Value.	0.69497352	0.827568322
## Saturated.Fat	0.62101082	0.569372866
## Saturated.Fat....Daily.Value.	0.62333417	0.573729742
## Trans.Fat	0.23836320	0.159086996
## Cholesterol	0.99985282	0.616325936
## Cholesterol....Daily.Value.	1.00000000	0.614876354
## Sodium	0.61487635	1.000000000
## Sodium....Daily.Value.	0.61531045	0.999919583
## Carbohydrates	0.25169824	0.159709137
## Carbohydrates....Daily.Value.	0.25368716	0.160708849
## Dietary.Fiber	0.41983803	0.685677745
## Dietary.Fiber....Daily.Value.	0.42492960	0.681198789
## Sugars	-0.12885866	-0.431571075
## Protein	0.55258220	0.852094856
## Vitamin.A....Daily.Value.	0.08628407	0.100413202
## Vitamin.C....Daily.Value.	-0.08662792	-0.038090541
## Calcium....Daily.Value.	0.14133222	-0.008691481
## Iron....Daily.Value.	0.64684843	0.887985012
##	Sodium....Daily.Value.	Carbohydrates
## Serving.Size	-0.433743207	0.24828840
## Calories	0.671564543	0.79280148
## Calories.from.Fat	0.828096443	0.44424962
## Total.Fat	0.827570029	0.44379487
## Total.Fat....Daily.Value.	0.828318912	0.44305096
## Saturated.Fat	0.570717810	0.58059110
## Saturated.Fat....Daily.Value.	0.575093060	0.58051132
## Trans.Fat	0.159893624	0.45380099
## Cholesterol	0.616738880	0.25356683
## Cholesterol....Daily.Value.	0.615310445	0.25169824
## Sodium	0.999919583	0.15970914
## Sodium....Daily.Value.	1.000000000	0.16145545
## Carbohydrates	0.161455451	1.00000000
## Carbohydrates....Daily.Value.	0.162454784	0.99961372
## Dietary.Fiber	0.685168582	0.20308154

## Dietary.Fiber....Daily.Value.	0.680630759	0.20699635
## Sugars	-0.429883066	0.78299016
## Protein	0.852166203	0.32012479
## Vitamin.A....Daily.Value.	0.100619813	0.09029658
## Vitamin.C....Daily.Value.	-0.038278930	-0.03756545
## Calcium....Daily.Value.	-0.006639487	0.60513982
## Iron....Daily.Value.	0.887086466	0.19490016
##	Carbohydrates....Daily.Value.	Dietary.Fiber
## Serving.Size	0.24967341	-0.36781922
## Calories	0.79315513	0.51971165
## Calories.from.Fat	0.44500559	0.57268268
## Total.Fat	0.44455421	0.57229714
## Total.Fat....Daily.Value.	0.44380207	0.57209948
## Saturated.Fat	0.58125543	0.33288761
## Saturated.Fat....Daily.Value.	0.58102759	0.33776513
## Trans.Fat	0.45353801	0.03565332
## Cholesterol	0.25554319	0.42063157
## Cholesterol....Daily.Value.	0.25368716	0.41983803
## Sodium	0.16070885	0.68567774
## Sodium....Daily.Value.	0.16245478	0.68516858
## Carbohydrates	0.99961372	0.20308154
## Carbohydrates....Daily.Value.	1.00000000	0.20287429
## Dietary.Fiber	0.20287429	1.00000000
## Dietary.Fiber....Daily.Value.	0.20631999	0.98592990
## Sugars	0.78249293	-0.28940563
## Protein	0.32106228	0.63379901
## Vitamin.A....Daily.Value.	0.08975596	0.35201420
## Vitamin.C....Daily.Value.	-0.03825228	0.14144579
## Calcium....Daily.Value.	0.60539098	0.03755611
## Iron....Daily.Value.	0.19552194	0.73571180
##	Dietary.Fiber....Daily.Value.	Sugars
## Serving.Size	-0.34783799	0.51913744
## Calories	0.52131200	0.30900026
## Calories.from.Fat	0.56668588	-0.09813392
## Total.Fat	0.56632398	-0.09830377
## Total.Fat....Daily.Value.	0.56620727	-0.09865072
## Saturated.Fat	0.32820266	0.21107660
## Saturated.Fat....Daily.Value.	0.33270202	0.20947094
## Trans.Fat	0.03923048	0.34484284
## Cholesterol	0.42553348	-0.12795838
## Cholesterol....Daily.Value.	0.42492960	-0.12885866
## Sodium	0.68119879	-0.43157107
## Sodium....Daily.Value.	0.68063076	-0.42988307
## Carbohydrates	0.20699635	0.78299016
## Carbohydrates....Daily.Value.	0.20631999	0.78249293
## Dietary.Fiber	0.98592990	-0.28940563
## Dietary.Fiber....Daily.Value.	1.00000000	-0.28111307
## Sugars	-0.28111307	1.00000000
## Protein	0.65111688	-0.16974408
## Vitamin.A....Daily.Value.	0.37310143	0.04648691
## Vitamin.C....Daily.Value.	0.14964275	-0.06904391
## Calcium....Daily.Value.	0.06153190	0.59897973
## Iron....Daily.Value.	0.73304952	-0.36059418
##	Protein	Vitamin.A....Daily.Value.

## Serving.Size	-0.30987560	-0.01043528
## Calories	0.74832150	0.13298056
## Calories.from.Fat	0.76669383	0.08121478
## Total.Fat	0.76647650	0.07870723
## Total.Fat....Daily.Value.	0.76662694	0.07831506
## Saturated.Fat	0.59467647	0.07159500
## Saturated.Fat....Daily.Value.	0.59781260	0.07211378
## Trans.Fat	0.37578998	0.08038090
## Cholesterol	0.55369719	0.08640586
## Cholesterol....Daily.Value.	0.55258220	0.08628407
## Sodium	0.85209486	0.10041320
## Sodium....Daily.Value.	0.85216620	0.10061981
## Carbohydrates	0.32012479	0.09029658
## Carbohydrates....Daily.Value.	0.32106228	0.08975596
## Dietary.Fiber	0.63379901	0.35201420
## Dietary.Fiber....Daily.Value.	0.65111688	0.37310143
## Sugars	-0.16974408	0.04648691
## Protein	1.00000000	0.24893049
## Vitamin.A....Daily.Value.	0.24893049	1.00000000
## Vitamin.C....Daily.Value.	-0.05668039	0.06974444
## Calcium....Daily.Value.	0.37929861	0.17786694
## Iron....Daily.Value.	0.81772381	0.14328187
##	Vitamin.C....Daily.Value.	Calcium....Daily.Value.
## Serving.Size	-0.0995373433	0.169118081
## Calories	-0.0812316122	0.486385878
## Calories.from.Fat	-0.1058258798	0.205176438
## Total.Fat	-0.1081652432	0.207375870
## Total.Fat....Daily.Value.	-0.1082135675	0.206552446
## Saturated.Fat	-0.1846850538	0.417168501
## Saturated.Fat....Daily.Value.	-0.1831576882	0.415239702
## Trans.Fat	-0.0789141282	0.393889160
## Cholesterol	-0.0862515766	0.140936205
## Cholesterol....Daily.Value.	-0.0866279192	0.141332218
## Sodium	-0.0380905406	-0.008691481
## Sodium....Daily.Value.	-0.0382789296	-0.006639487
## Carbohydrates	-0.0375654517	0.605139816
## Carbohydrates....Daily.Value.	-0.0382522829	0.605390980
## Dietary.Fiber	0.1414457903	0.037556107
## Dietary.Fiber....Daily.Value.	0.1496427549	0.061531903
## Sugars	-0.0690439099	0.598979729
## Protein	-0.0566803872	0.379298610
## Vitamin.A....Daily.Value.	0.0697444385	0.177866943
## Vitamin.C....Daily.Value.	1.0000000000	-0.214921978
## Calcium....Daily.Value.	-0.2149219778	1.000000000
## Iron....Daily.Value.	-0.0005954675	0.040348640
##	Iron....Daily.Value.	
## Serving.Size	-0.4487687694	
## Calories	0.6519542293	
## Calories.from.Fat	0.7690487610	
## Total.Fat	0.7679391289	
## Total.Fat....Daily.Value.	0.7690266784	
## Saturated.Fat	0.5698630762	
## Saturated.Fat....Daily.Value.	0.5722734082	
## Trans.Fat	0.3157469025	

```
## Cholesterol 0.6487481269
## Cholesterol....Daily.Value. 0.6468484318
## Sodium 0.8879850117
## Sodium....Daily.Value. 0.8870864656
## Carbohydrates 0.1949001625
## Carbohydrates....Daily.Value. 0.1955219422
## Dietary.Fiber 0.7357118000
## Dietary.Fiber....Daily.Value. 0.7330495195
## Sugars -0.3605941782
## Protein 0.8177238099
## Vitamin.A....Daily.Value. 0.1432818680
## Vitamin.C....Daily.Value. -0.0005954675
## Calcium....Daily.Value. 0.0403486400
## Iron....Daily.Value. 1.0000000000
```

```
# Remove variables representing same thing
```

```
McDonalds <- subset(McDonalds, select = - Calories.from.Fat)
McDonalds <- subset(McDonalds, select = - Total.Fat....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Saturated.Fat....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Cholesterol....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Sodium....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Carbohydrates....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Dietary.Fiber....Daily.Value.)
McDonalds <- subset(McDonalds, select = - Item)
```

```
ggplot(McDonalds, aes(x = Calories)) +
  geom_histogram(aes(y=..density..), bins = 30, color = 'darkblue', fill = 'steelblue') +
  geom_vline(aes(xintercept = mean(Calories)), color = 'lightgreen', linetype = 'dashed') +
  geom_density(alpha=0.6, color = 'steelblue', fill = 'lightblue')
```

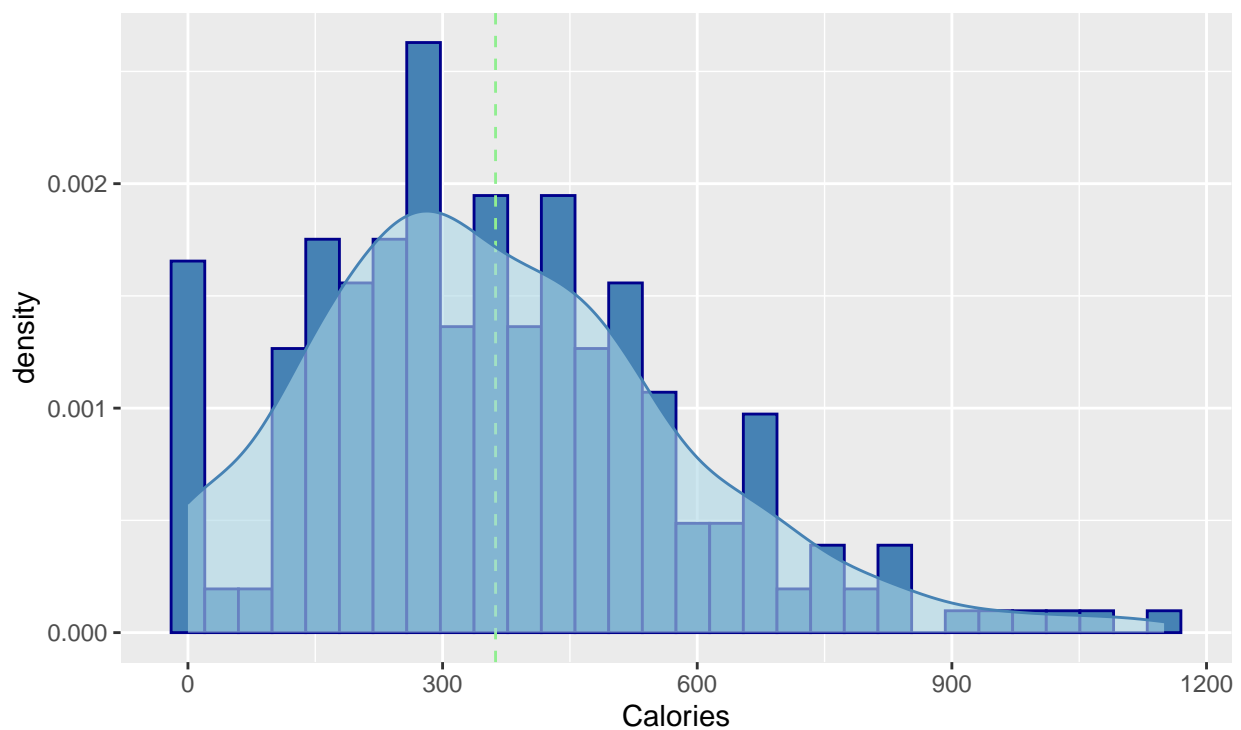


Figure 1: Histogram for Calories

```

mean(McDonalds$Calories)

## [1] 362.4324

# Set random seed
set.seed(2021)
# Train a tree model that includes every feature in the data set
fullTree1 <- rpart(Calories ~ ., method = "anova", data = McDonalds)
# Show the root node error for the full model
printcp(fullTree1)

##
## Regression tree:
## rpart(formula = Calories ~ ., data = McDonalds, method = "anova")
##
## Variables actually used in tree construction:
## [1] Carbohydrates Sodium          Sugars          Total.Fat
##
## Root node error: 12657818/259 = 48872
##
## n= 259
##
##      CP nsplit rel error  xerror    xstd
## 1 0.598805      0  1.000000  1.00395  0.101338
## 2 0.131256      1  0.401195  0.42713  0.051412
## 3 0.076137      2  0.269939  0.36359  0.045410
## 4 0.044482      3  0.193802  0.23799  0.025659
## 5 0.021568      4  0.149319  0.21965  0.021252
## 6 0.019259      5  0.127751  0.19660  0.020625
## 7 0.011656      6  0.108492  0.18006  0.020555
## 8 0.011175      7  0.096836  0.16502  0.019703
## 9 0.010000      8  0.085661  0.16342  0.019692

# Compute the test MSE for the full model
48872*0.1634231

## [1] 7986.814

# Output the cp table
knitr::kable(fullTree1$cptable, caption = "\\label{tab:cptableFullTree1}The cp table for the fullTree1")

```

Table 1: The cp table for the fullTree1

CP	nsplit	rel error	xerror	xstd
0.5988049	0	1.0000000	1.0039544	0.1013384
0.1312561	1	0.4011951	0.4271343	0.0514122
0.0761375	2	0.2699390	0.3635852	0.0454102
0.0444821	3	0.1938015	0.2379931	0.0256588
0.0215682	4	0.1493194	0.2196494	0.0212518
0.0192588	5	0.1277512	0.1965999	0.0206247
0.0116562	6	0.1084923	0.1800554	0.0205555
0.0111753	7	0.0968361	0.1650205	0.0197030
0.0100000	8	0.0856608	0.1634231	0.0196925

```
# Compute the test MSE for the pruned tree
48872*0.1650205
```

```
## [1] 8064.882
```

```
# Plot the relationship between xerror and the cp values
plotcp(fullTree1)
```

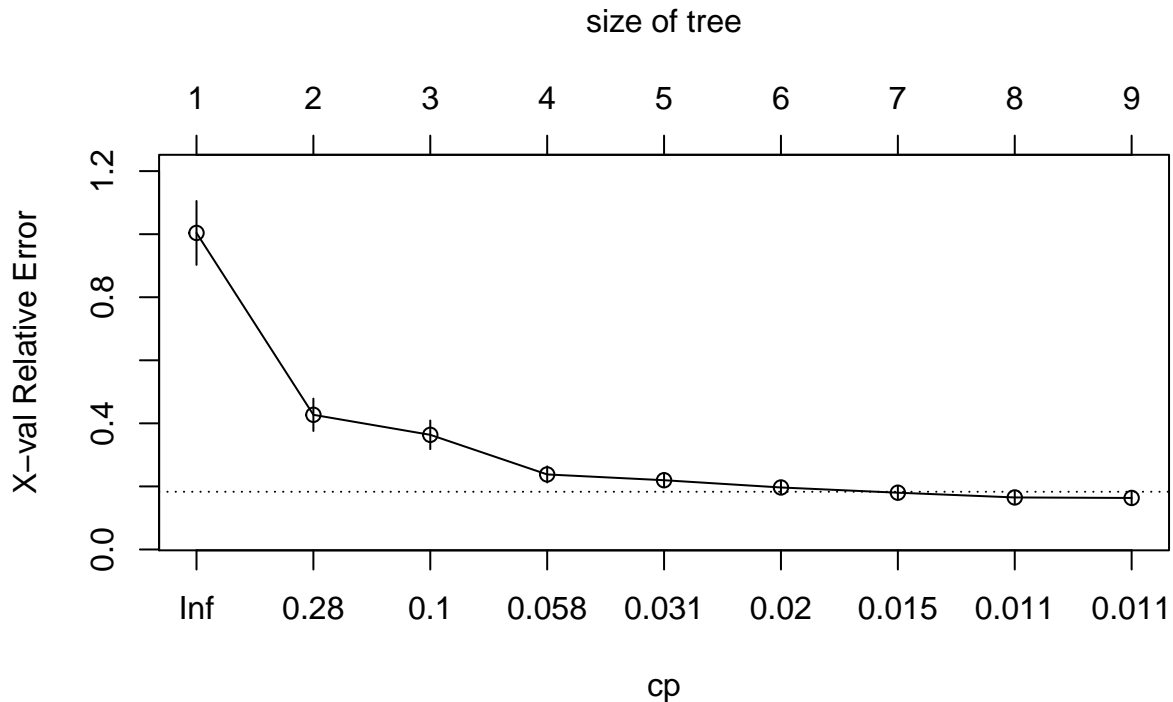


Figure 2: cp plot for fullTree1

```
# Pruning the fullTree1
Tree1 <- prune(fullTree1, cp = fullTree1$cptable[8,"CP"])
# Create a visualization
rpart.plot(Tree1, cex = 0.6)
```

```
# Compute the training RMSE
pred.Tree1 <- matrix(NA, nrow = nrow(McDonalds), ncol = 1)
pred.Tree1 <- predict(Tree1, newdata = McDonalds)
residual.Tree1 <- McDonalds$Calories - pred.Tree1
trainingRMSE <- sqrt((t(residual.Tree1)%*(residual.Tree1))/nrow(McDonalds))
trainingRMSE
```

```
## [1,] 68.79363
```

```
# Compute the training RMSE for the Tree1
sqrt(48872)
```

```
## [1] 221.0701
```

```
# Compute the test RMSE for the Tree1
sqrt(8064.882)
```

```
## [1] 89.80469
```

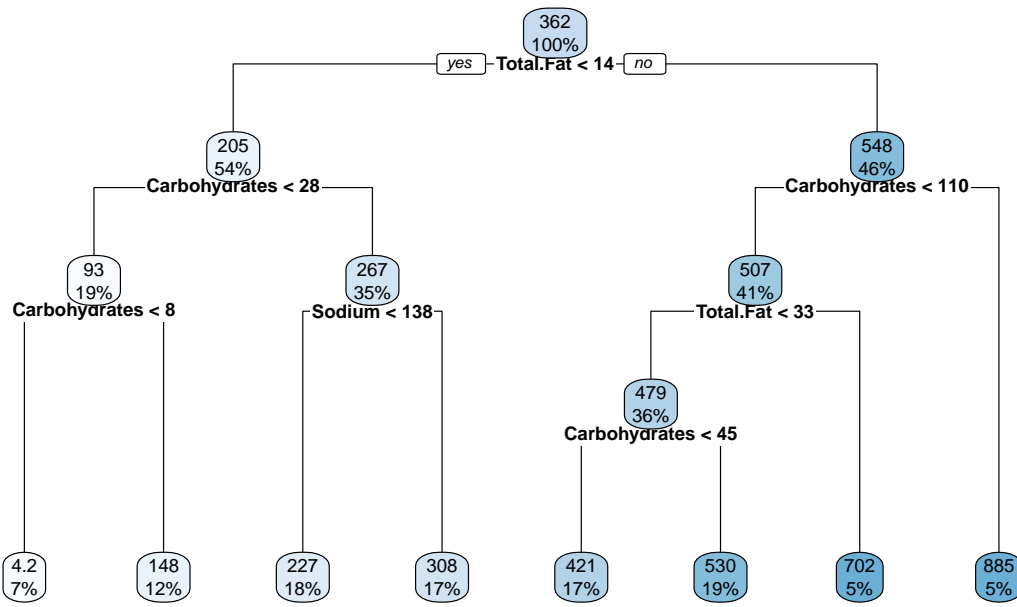



Figure 3: Visualization for Tree1

```
# Visualizing the distribution of the response variable
ggplot(McDonalds, aes(x = Category)) +
  geom_bar(width=0.7, color = 'darkblue', fill = 'steelblue') +
  geom_text(stat = 'count', aes(label=..count..), vjust=-0.3, size=3.5)

# Create a subset of Category data that would be served as the response variable
McDonaldsCat <- subset(McDonalds, Category == "Beef & Pork" | Category == "Breakfast" | Category == "Chicken")
# Set seed
set.seed(2021)
# train the fullTree2 model using all features in the data
fullTree2 <- rpart(Category ~ ., method = "class", data = McDonaldsCat)
# Show the RNE
printcp(fullTree2)

##
## Classification tree:
## rpart(formula = Category ~ ., data = McDonaldsCat, method = "class")
##
## Variables actually used in tree construction:
## [1] Protein      Saturated.Fat Sugars      Trans.Fat
##
## Root node error: 80/122 = 0.65574
##
## n= 122
##
##      CP nsplit rel error xerror      xstd
## 1 0.38750      0  1.0000 1.0000 0.065599
## 2 0.12917      1  0.6125 0.6125 0.067685
## 3 0.01250      4  0.2250 0.3625 0.058772
## 4 0.01000      5  0.2125 0.3875 0.060108
```

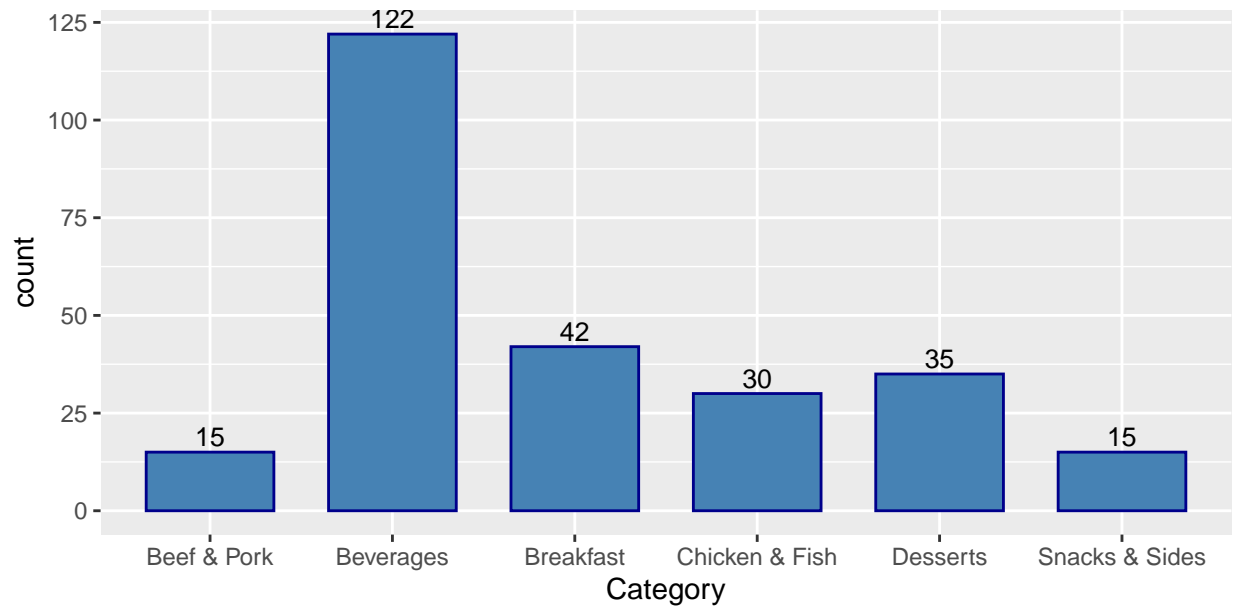


Figure 4: the Distribution of the Responce Variable Category

```
# Compute the test CER for the full model
0.65574*0.3875
```

```
## [1] 0.2540992
```

```
# Output the cp table
```

```
knitr::kable(fullTree2$cptable, caption = "\\label{tab:cptableFullTree2}The cp table for the fullTree2")
```

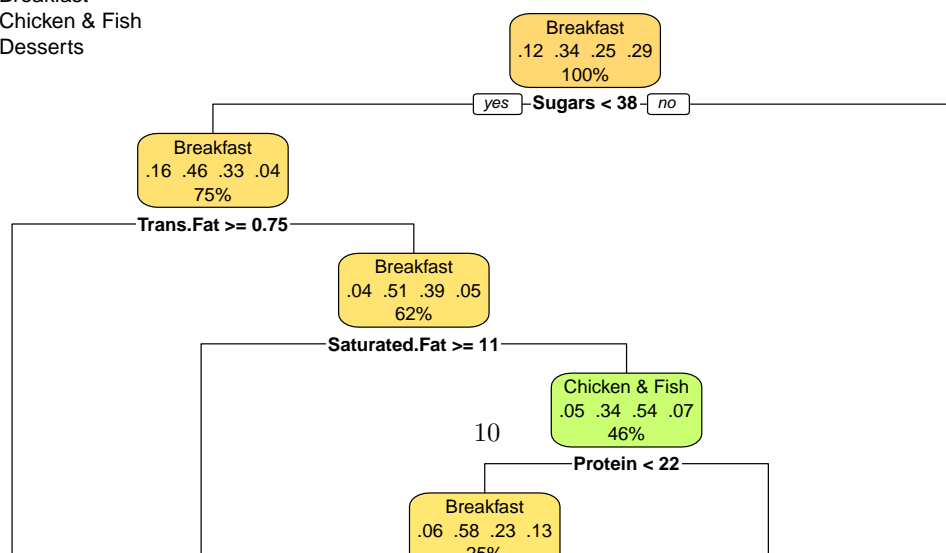
Table 2: The cp table for the fullTree2

CP	nsplit	rel error	xerror	xstd
0.3875000	0	1.0000	1.0000	0.0655994
0.1291667	1	0.6125	0.6125	0.0676846
0.0125000	4	0.2250	0.3625	0.0587720
0.0100000	5	0.2125	0.3875	0.0601079

```
# Visualization for the fullTree2 model
```

```
rpart.plot(fullTree2, box.palette = "RdYlGn")
```

■ Beef & Pork
■ Breakfast
■ Chicken & Fish
■ Desserts



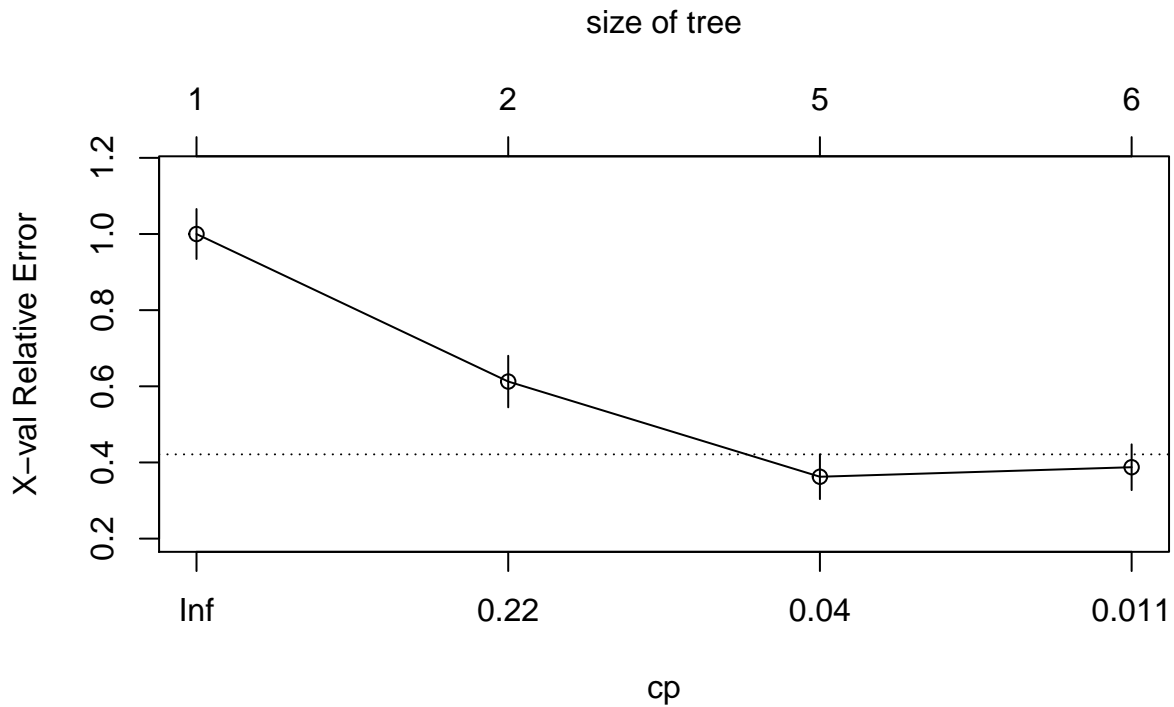


Figure 6: cp plot for fullTree2

```
# storing the optimal alpha in x
x <- fullTree2$cptable[which.min(fullTree2$cptable[, "xerror"]), "CP"]
# Pruning to get the Tree2 model
Tree2 <- prune(fullTree2, cp = x)
# Visualization 1 for the Tree2
rpart.plot(Tree2)

# Visualization 2 for the Tree2
prp(Tree2, box.palette = "RdYlGn")

# Compute the number of observations aren't predict correctly in the training data
122*0.65574

## [1] 80.00028

# Compute the test CER
0.65574*0.3625

## [1] 0.2377057

# Compute the numebr of observations aren't correctly predicted in the test data
122*0.23771

## [1] 29.00062

# Set seed
set.seed(2021)
# Training the Bagged Forest Model
BgForest <- randomForest(Calories ~ ., data = McDonalds, mtry = 15, importance = TRUE, ntree = 1000, cores = 4)
BgForest

##
```

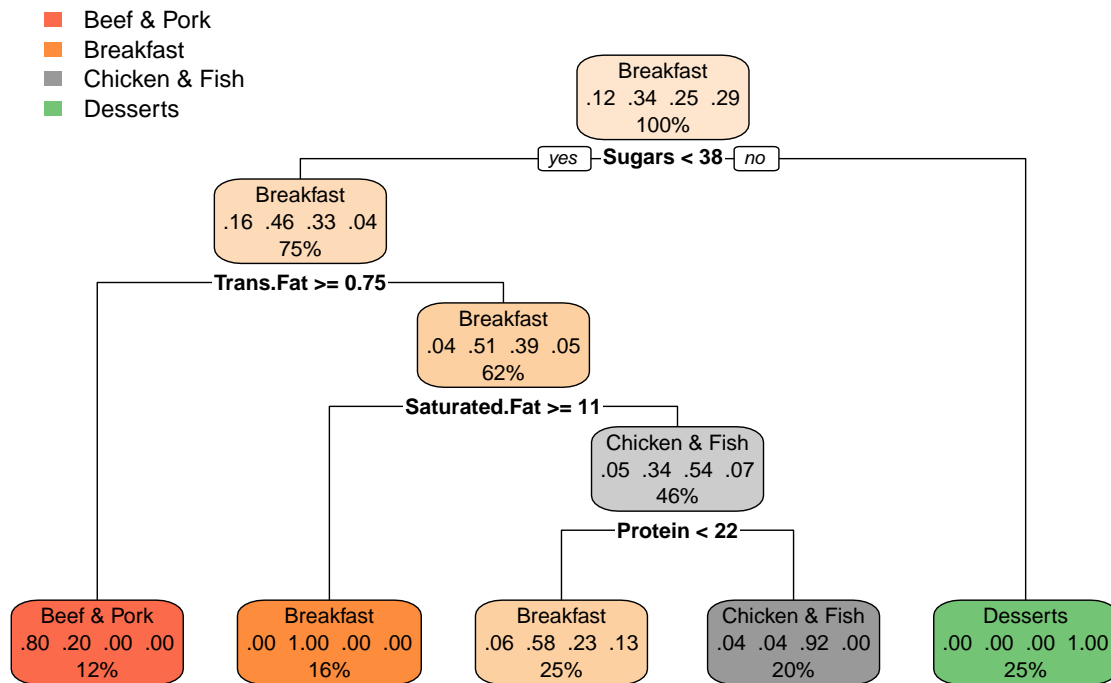


Figure 7: The Detailed Visualization for Tree2

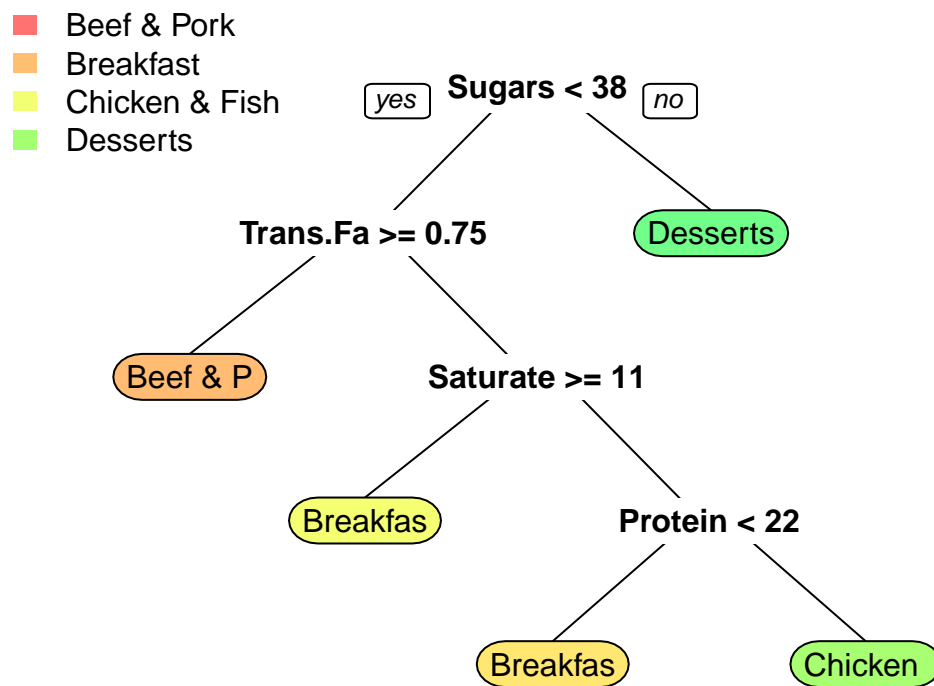


Figure 8: The Concise Visualization for Tree2

```
## Call:
## randomForest(formula = Calories ~ ., data = McDonalds, mtry = 15, importance = TRUE, ntree = 1000)
##           Type of random forest: regression
##           Number of trees: 1000
## No. of variables tried at each split: 15
##
##           Mean of squared residuals: 973.5829
##           % Var explained: 98.01
```

```
# Compute the OOB error estimate
predict.OOB.BgForest <- BgForest$predicted
# Compute the test RMSE for the BgForest
sqrt(mean((McDonalds$Calories - predict.OOB.BgForest)^2))
```

```
## [1] 31.20229
```

```
# Compute the improvement from the Tree1 Model
(89.80469 - 31.20229)/89.80469
```

```
## [1] 0.6525539
```

```
# Set seed
set.seed(2021)
# Train the Random Forest Model
RdForest <- randomForest(Calories ~ ., data = McDonalds, mtry = sqrt(15), importance = TRUE, ntree = 1000)
RdForest
```

```
##
## Call:
## randomForest(formula = Calories ~ ., data = McDonalds, mtry = sqrt(15), importance = TRUE, ntree = 1000)
##           Type of random forest: regression
##           Number of trees: 1000
## No. of variables tried at each split: 4
##
##           Mean of squared residuals: 1083.044
##           % Var explained: 97.78
```

```
# Compute the OOB error estimate
predict.OOB.RdForest <- RdForest$predicted
# Compute the test RMSE for the RdForest
sqrt(mean((McDonalds$Calories - predict.OOB.RdForest)^2))
```

```
## [1] 32.90963
```

```
# Compute the improvement from the Tree1 Model
(89.80469 - 32.90963)/89.80469
```

```
## [1] 0.6335422
```

```
# Compute the percent increase in OOB error estimate
importance(BgForest)[,1]
```

##	Category	Serving.Size	Total.Fat
##	7.701320	23.104697	102.321768
##	Saturated.Fat	Trans.Fat	Cholesterol
##	17.307040	5.082977	24.290918
##	Sodium	Carbohydrates	Dietary.Fiber
##	27.243796	86.031234	12.783145
##	Sugars	Protein	Vitamin.A....Daily.Value.

```
##                24.188490                16.902343                9.304025
## Vitamin.C....Daily.Value.  Calcium....Daily.Value.  Iron....Daily.Value.
##                6.971649                12.667627                11.390955

# Plot the importance
dotchart(importance(BgForest)[,1], xlab = "Percent Increase in the OOB Error Estimate")
```

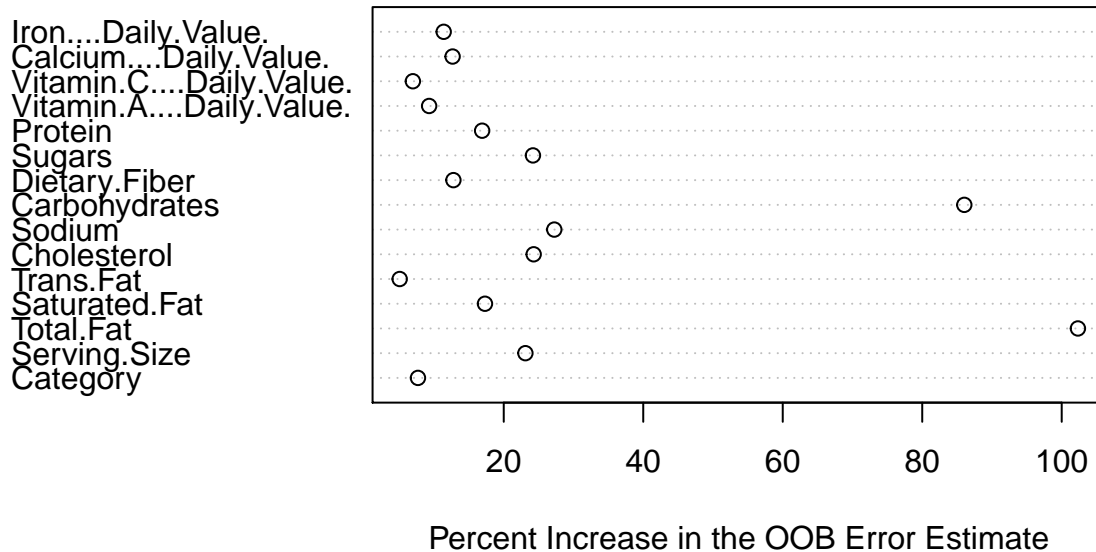


Figure 9: The Importance Plot for BgForest

```
# Compute the percent increase in OOB error estimate
importance(RdForest)[,1]

##                Category                Serving.Size                Total.Fat
##                9.807717                16.526972                32.778472
##                Saturated.Fat                Trans.Fat                Cholesterol
##                24.464427                9.194409                18.153171
##                Sodium                Carbohydrates                Dietary.Fiber
##                22.315483                37.492130                14.320768
##                Sugars                Protein Vitamin.A....Daily.Value.
##                30.560177                17.497024                12.185273
## Vitamin.C....Daily.Value.  Calcium....Daily.Value.  Iron....Daily.Value.
##                6.402707                16.026553                9.308504

# Plot the importance
dotchart(importance(RdForest)[,1], xlab = "Percent Increase in the OOB Error Estimate")

# Visualize the partial dependence between Calories and Carbonhydrates
partialPlot(BgForest, McDonalds, x.var = "Carbohydrates", ylab = "Calories")

# Visualize the partial dependence between Calories and Total.Fat
partialPlot(BgForest, McDonalds, x.var = "Total.Fat", ylab = "Calories")
```

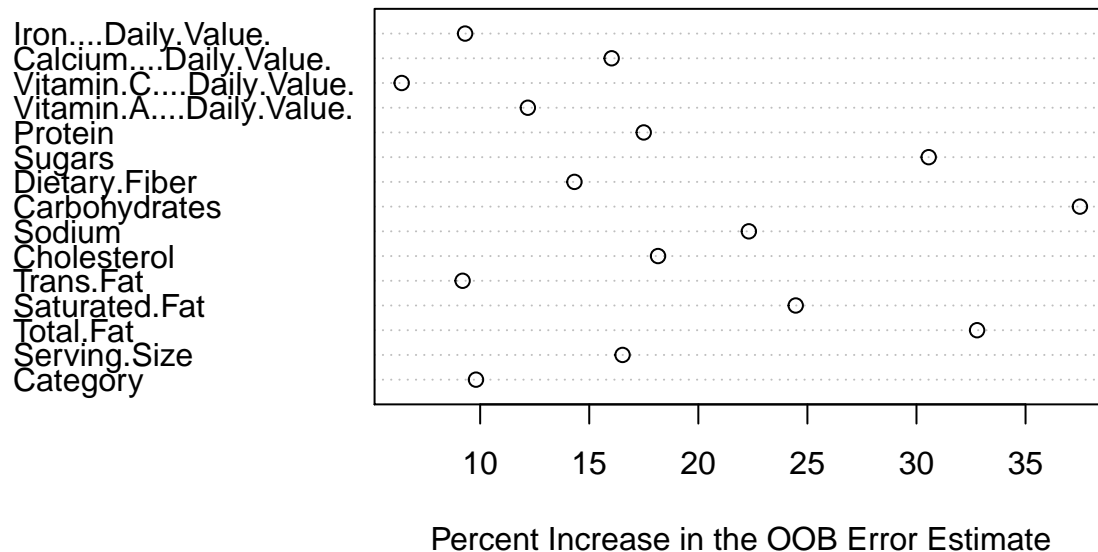


Figure 10: The Importance Plot for RdForest

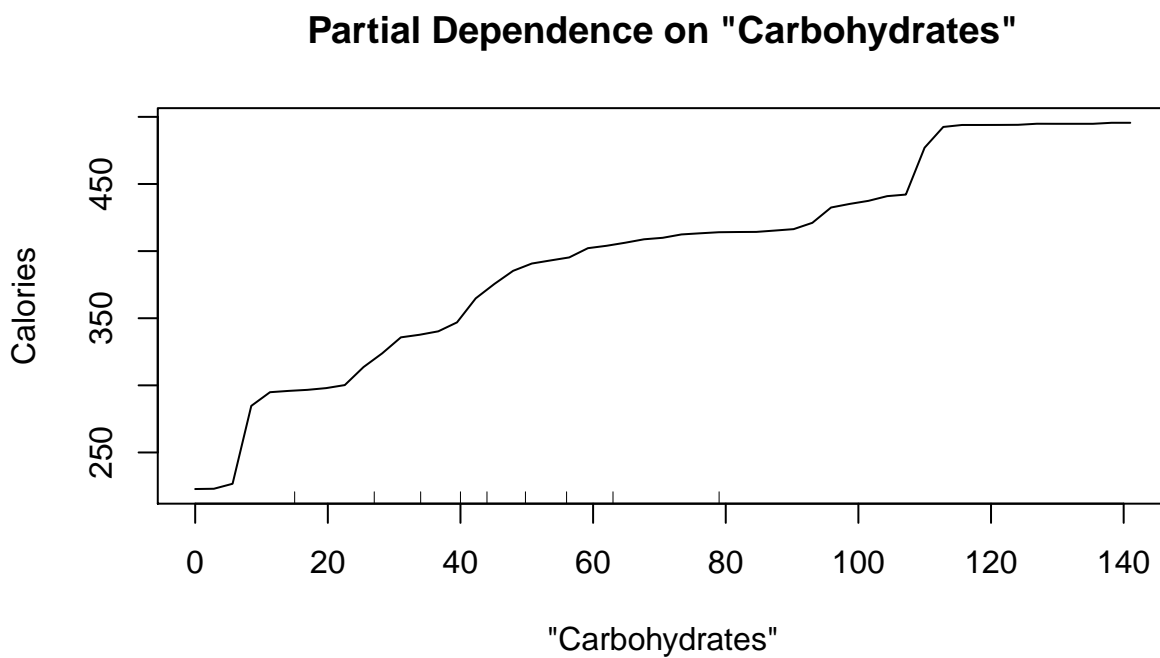


Figure 11: Partial Dependence on Carbohydrates

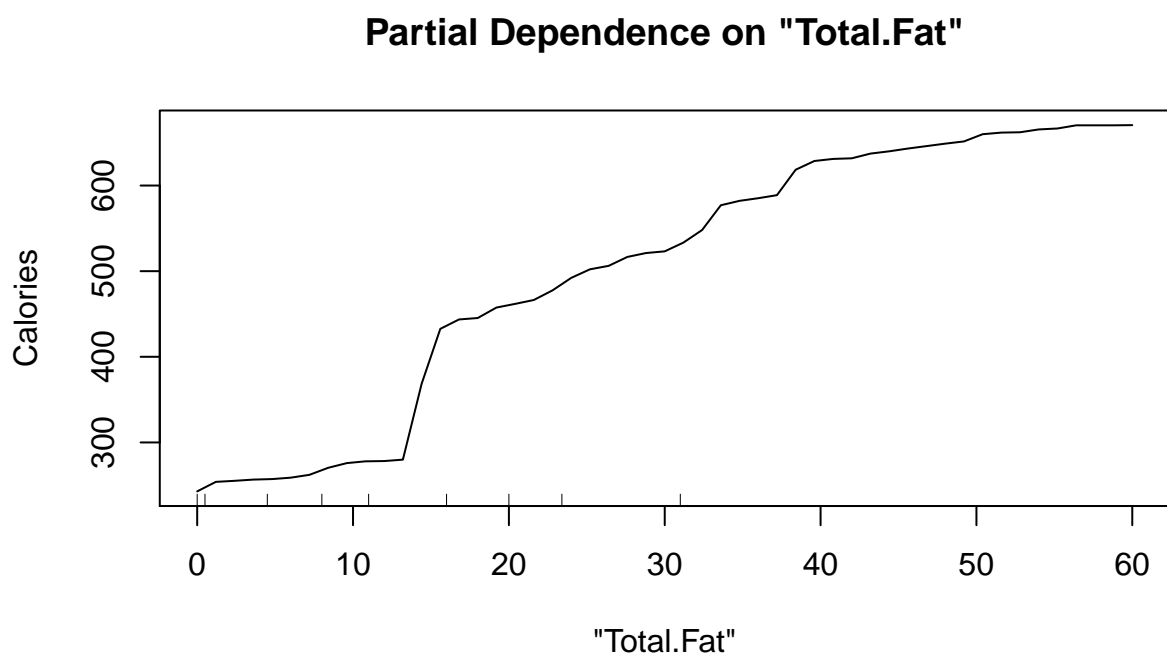


Figure 12: Partial Dependence on Total.Fat