Project -1.1

McDonalds Competition

Table of Contents

Case study:

McDonalds is an American fast food company with operations in more than 100 countries and revenues in excess of $ 20 billion. McDonald's offers a menu that includes fries, the Big Mac, chicken sandwiches, chicken nuggets, hamburgers, salads, wraps, desserts, soft drinks, and other beverages. However, to ensure that it connects with the international markets, McDonald's offers locally relevant food menus as well.

Here in your first flavor of a hackathon, we are providing you with dataset of the food menu at McDonald's and you are asked to leverage the skills you have learnt till now in analytics and R to answer the following questions and in return you can win amazon gift vouchers worth Rs. 500.

Questions:

1. Which food categories have the highest and lowest varieties ? (A visualization supporting your answer would be very much appreciated)
2. Which all variables have an outlier ?
3. Which variables have the Highest correlation. Plot them and tell the value.
4. Which category contributes to the maximum % of Cholesterol in a diet (% daily value)?
5. Which item contributes maximum to the Sodium intake?
6. Which 4 food items contains the most amount of Saturated Fat?

Solutions :

1. Which food categories have the highest and lowest varieties? (A visualization supporting your answer would be very much appreciated)

```
> Mcdonald<- read.csv("C:/Users/IBM_ADMIN/Documents/My Received Files/r file/R Programming/Mc
donald.csv", header= TRUE)
>
```

```
> str(Mcdonald)
'data.frame':   260 obs. of  24 variables:
 $ Category                 : Factor w/ 9 levels "Beef & Pork",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ Item                     : Factor w/ 260 levels "1% Low Fat Milk Jug",..: 76 77 228 229 230
245 12 11 14 13 ...
 $ Serving.Size             : Factor w/ 107 levels "1 carton (236 ml)",..: 55 54 42 69 69 83 63
72 65 73 ...
 $ Calories                 : int  300 250 370 450 400 430 460 520 410 470 ...
 $ Calories.from.Fat        : int  120 70 200 250 210 210 230 270 180 220 ...
 $ Total.Fat                : num  13 8 23 28 23 23 26 30 20 25 ...
 $ Total.Fat....Daily.Value.: int  20 12 35 43 35 36 40 47 32 38 ...
 $ Saturated.Fat            : num  5 3 8 10 8 9 13 14 11 12 ...
 $ Saturated.Fat....Daily.Value.: int  25 15 42 52 42 46 65 68 56 59 ...
 $ Trans.Fat                : num  0 0 0 0 1 0 0 0 0 ...
 $ Cholesterol              : int  260 25 45 285 50 300 250 250 35 35 ...
 $ Cholesterol....Daily.Value. : int  87 8 15 95 16 100 83 83 11 11 ...
 $ Sodium                   : int  750 770 780 860 880 960 1300 1410 1300 1420 ...
 $ Sodium....Daily.Value.   : int  31 32 33 36 37 40 54 59 54 59 ...
 $ Carbohydrates            : int  31 30 29 30 30 31 38 43 36 42 ...
 $ Carbohydrates....Daily.Value.: int  10 10 10 10 10 10 13 14 12 14 ...
 $ Dietary.Fiber            : int  4 4 4 4 4 4 2 3 2 3 ...
 $ Dietary.Fiber....Daily.Value.: int  17 17 17 17 17 18 7 12 7 12 ...
 $ Sugars                   : int  3 3 2 2 2 3 3 4 3 4 ...
 $ Protein                  : int  17 18 14 21 21 26 19 19 20 20 ...
 $ Vitamin.A....Daily.Value.: int  10 6 8 15 6 15 10 15 2 6 ...
 $ Vitamin.C....Daily.Value.: int  0 0 0 0 0 2 8 8 8 8 ...
 $ Calcium....Daily.Value.  : int  25 25 25 30 25 30 15 20 15 15 ...
 $ Iron....Daily.Value.     : int  15 8 10 15 10 20 15 20 10 15 ...
```

We found there is 260 observention and 24 variable and 9 levels in Category whereas 260 levels in Itema
and 107 levels are existed for Serving .Size.

Coffee & Tea is having the highetest (95)varities
Salads is having lowest (6) number of varities

Beef & Pork	Beverages	Breakfast	Chicken & Fish	Coffee & Tea
15	27	42	27	95

Desserts	Salads	Smoothies & Shakes	Snacks & Sides
7	6	28	13

```
> levels(Category)
[1] "Beef & Pork"       "Beverages"         "Breakfast"         "Chicken & Fish"
[5] "Coffee & Tea"      "Desserts"          "Salads"            "Smoothies & Shakes"
[9] "Snacks & Sides"
>
```

```
> summary(Mcdonald$Category)
       Beef & Pork          Beverages           Breakfast      Chicken & Fish         Coffee & Tea
                15                 27                  42                  27                   95
          Desserts             Salads  Smoothies & Shakes      Snacks & Sides
                 7                  6                  28                  13
>
```

```
> plot(Mcdonald$Category,col = "Blue", xlab ="Food Categries", main=" Maximum of Food Categories",ylim = c(0,10
0))
>
```

Comparision of Varities
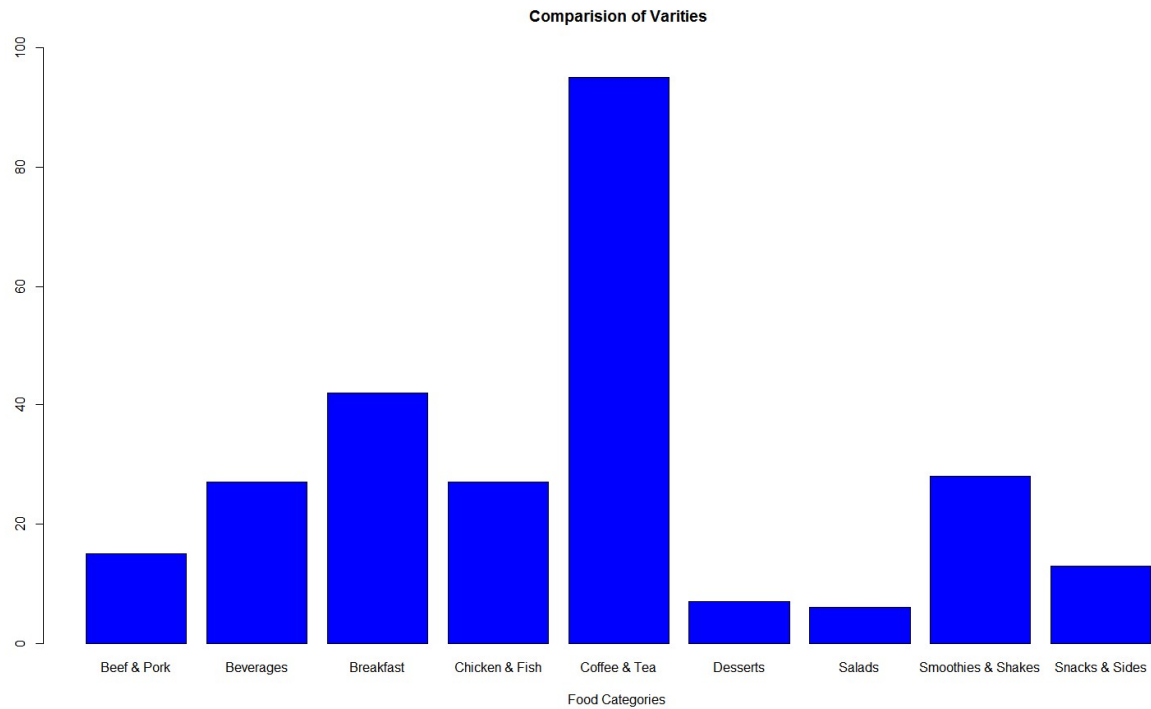
2. Which all variables have an outlier?

Following variables availed in the data set , Category , Item and Serving Size is not carrying any values of frequencies

```
> names(Mcdonald)
 [1] "Category"                        "Item"                     "Serving.Size"
 [4] "Calories"                        "Calories.from.Fat"        "Total.Fat"
 [7] "Total.Fat....Daily.Value."       "Saturated.Fat"            "Saturated.Fat....Daily.Value."
[10] "Trans.Fat"                       "Cholesterol"              "Cholesterol....Daily.Value."
[13] "Sodium"                          "Sodium....Daily.Value."   "Carbohydrates"
[16] "Carbohydrates....Daily.Value."   "Dietary.Fiber"            "Dietary.Fiber....Daily.Value."
[19] "Sugars"                          "Protein"                  "Vitamin.A....Daily.Value."
[22] "Vitamin.C....Daily.Value."       "Calcium....Daily.Value."  "Iron....Daily.Value."
> |
```

Calories
Calories.from.Fat
Total.Fat
Total.Fat....Daily.Value.
Saturated.Fat
Saturated.Fat....Daily.Value.
Trans.Fat
Cholesterol
Cholesterol....Daily.Value.
Sodium
Sodium....Daily.Value.
Carbohydrates

Carbohydrates....Daily.Value.
Dietary.Fiber
Dietary.Fiber....Daily.Value>
Sugars
Protein
Vitamin.A....Daily.Value.
Vitamin.C....Daily.Value.
Calcium....Daily.Value.
Iron....Daily.Value.

We have ploted here all variable and found these are the variable having outliers

```
> boxplot(Mcdonald,horizontal = TRUE,col=c("Red","Blue"),ylab ="Variables", main=" Box plot of Variables",
ylim = c(0,1000))
> boxplot(Cholesterol,Sodium,Carbohydrates,Sugars,Protein,Calories,Iron....Daily.Value.,horizontal = TRUE,
col=c("Red","Blue"),ylab ="Food Categories", main=" Box plot of varieties ",ylim = c(0,1000))
> par(mfrow= c(1,3))
> boxplot(Vitamin.C....Daily.Value.,col= "Red",xlab ="Vitamin.C....Daily.Value.", main=" Box plot of Varia
bles ",ylim = c(0,350))
> boxplot(Vitamin.A....Daily.Value. ,col= "Green",xlab ="Vitamin.A....Daily.Value. ", main=" Box plot of V
ariables ",ylim = c(0,350))
> boxplot(Cholesterol,col= "Blue",xlab ="Cholesterol", main=" Box plot of Variables ",ylim = c(0,350))
> |
```


Box plot of Variables
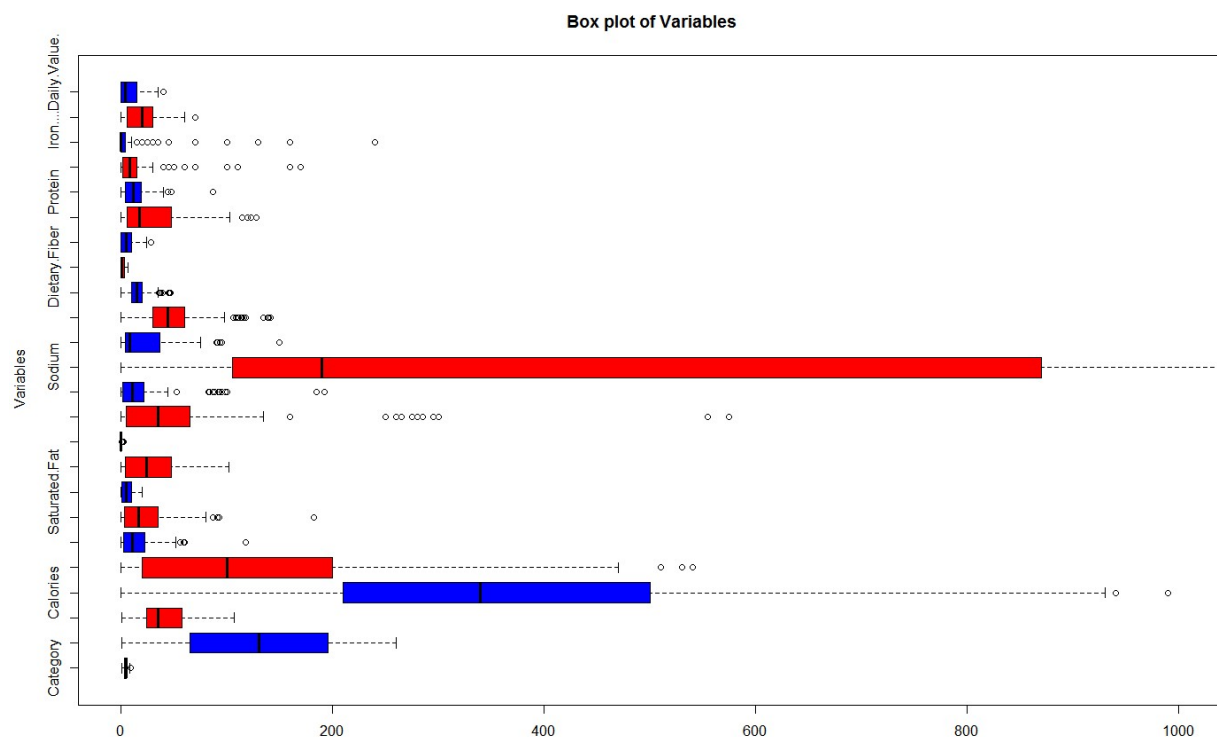
Below mentioned varibales are having large number of outliers

Box plot of Variables — Vitamin.C....Daily.Value.

Box plot of Variables — Vitamin.A....Daily.Value.

Box plot of Variables — Cholesterol
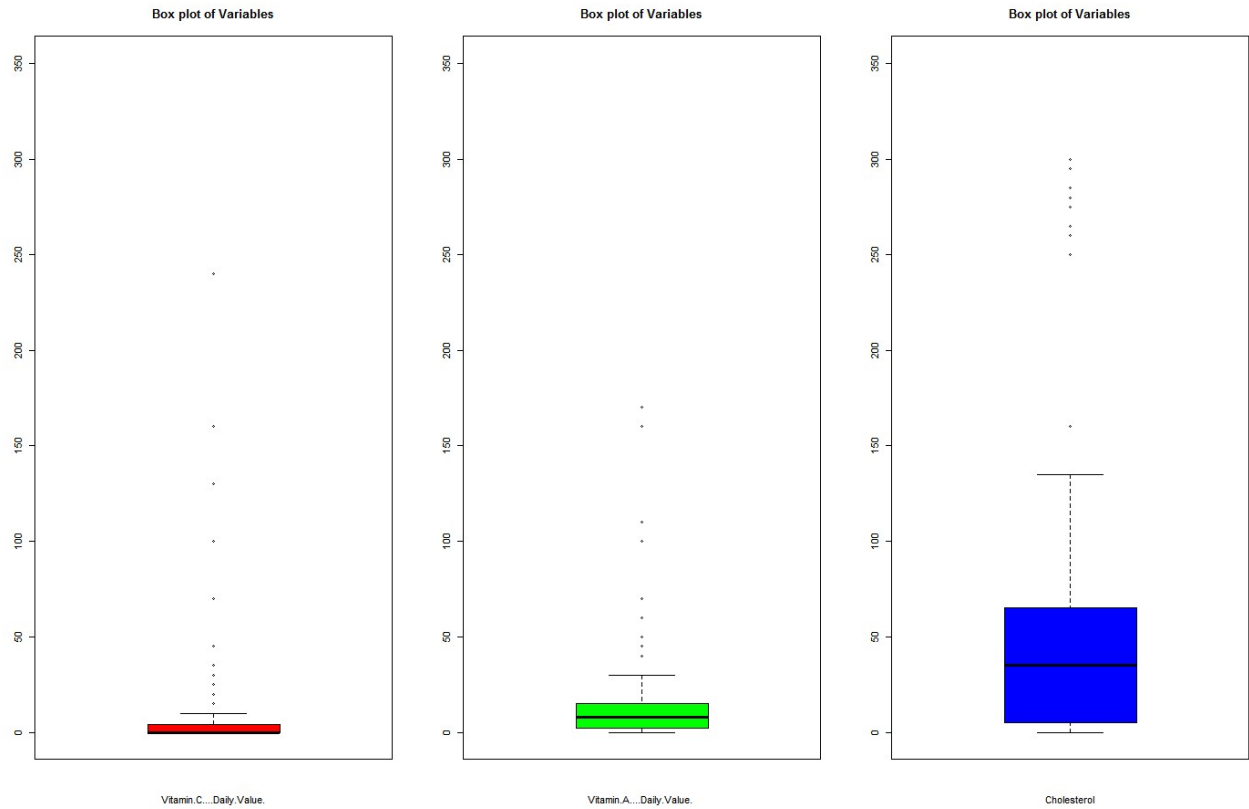
Category wise box plot
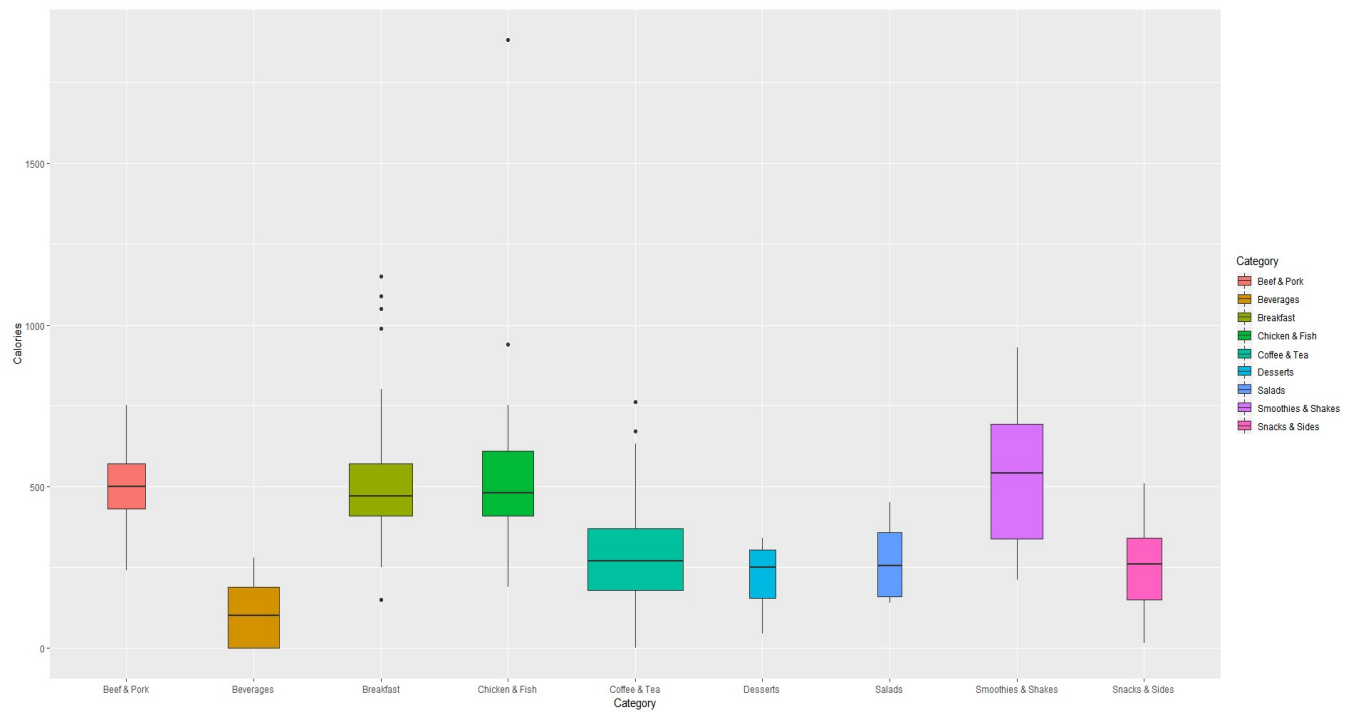
3. Which variables have the Highest correlation. Plot them and tell the value.

```
> cor(Mcdonald[,unlist(lapply(Mcdonald, is.numeric))])
                          Calories Calories.from.Fat    Total.Fat Total.Fat...Daily.Value. Saturated.Fat
Calories                1.00000000        0.90458780   0.90440916               0.90412255    0.84556364
Calories.from.Fat       0.90458780        1.00000000   0.99966350               0.99972526    0.84700759
Total.Fat               0.90440916        0.99966350   1.00000000               0.99976474    0.84670724
Total.Fat....Daily.Value. 0.90412255      0.99972526   0.99976474               1.00000000    0.84737925
Saturated.Fat           0.84556364        0.84700759   0.84670724               0.84737925    1.00000000
Saturated.Fat....Daily.Value. 0.84763077  0.84959196   0.84929279               0.84997301    0.99927862
Trans.Fat               0.52244092        0.43368614   0.43145290               0.43301574    0.62061061
Cholesterol             0.59639916        0.68216110   0.68054737               0.68093953    0.63121047
Cholesterol....Daily.Value. 0.59520769    0.68160700   0.67999970               0.68037809    0.63033406
Sodium                  0.71230869        0.84662441   0.84615842               0.84672792    0.58407526
Sodium....Daily.Value.  0.71341497        0.84727635   0.84678022               0.84736817    0.58532342
Carbohydrates           0.78153946        0.46167225   0.46121347               0.46051625    0.59126063
Carbohydrates....Daily.Value. 0.78124203  0.46146307   0.46100479               0.46029841    0.59174306
Dietary.Fiber           0.53889351        0.58127379   0.58083725               0.58059179    0.35181784
Dietary.Fiber....Daily.Value. 0.54001419  0.57562058   0.57520633               0.57503300    0.34715171
Sugars                  0.25959812       -0.11528469  -0.11544573              -0.11576122    0.19773362
Protein                 0.78784745        0.80791326   0.80777296               0.80792221    0.60302754
Vitamin.A....Daily.Value. 0.10884404      0.05673072   0.05443396               0.05403819    0.06497234
Vitamin.C....Daily.Value. -0.06874704    -0.08733073  -0.08935397              -0.08935258   -0.17967200
Calcium....Daily.Value. 0.42842646        0.16103409   0.16285952               0.16203088    0.40331077
```

We have prepared a new subset for principal numeric variable , the factor variables are taken out here.

We observed Saturated. Fat, Total . Fat ,Carbohydrates and protein are highly correlated

```
> Mcdonald_cor <- Mcdonald[,c("Calories","Protein","Calories.from.Fat","Sodium","Cholesterol","Sugars","Total.Fat","Saturate
d.Fat",            "Carbohydrates","Dietary.Fiber","Trans.Fat")]
> Mcdonald_cor
   Calories Protein Calories.from.Fat Sodium Cholesterol Sugars Total.Fat Saturated.Fat Carbohydrates Dietary.Fiber
1       300      17               120    750         260      3      13.0           5.0            31             4
2       250      18                70    770          25      3       8.0           3.0            30             4
3       370      14               200    780          45      2      23.0           8.0            29             4
4       450      21               250    860         285      2      28.0          10.0            30             4
5       400      21               210    880          50      2      23.0           8.0            30             4
6       430      26               210    960         300      3      23.0           9.0            31             4
7       460      19               230   1300         250      3      26.0          13.0            38             2
8       520      19               270   1410         250      4      30.0          14.0            43             3
9       410      20               180   1300          35      3      20.0          11.0            36             2
10      470      20               220   1420          35      4      25.0          12.0            42             3
11      430      11               240   1080          30      2      27.0          12.0            34             2
12      480      11               280   1190          30      3      31.0          13.0            39             3
```

```
> cors <- cor(subset(Mcdonald, select = c(Calories,Protein,Sodium,Cholesterol,Sugars,Total.Fat,Saturated.Fa
                              Carbohydrates,Dietary.Fiber )))
> cors
                Calories    Protein     Sodium Cholesterol     Sugars  Total.Fat Saturated.Fat
Calories       1.0000000  0.7878475  0.7123087   0.5963992  0.2595981  0.9044092     0.8455636
Protein        0.7878475  1.0000000  0.8698016   0.5615614 -0.1799396  0.8077730     0.6030275
Sodium         0.7123087  0.8698016  1.0000000   0.6243619 -0.4265355  0.8461584     0.5840753
Cholesterol    0.5963992  0.5615614  0.6243619   1.0000000 -0.1355183  0.6805474     0.6312105
Sugars         0.2595981 -0.1799396 -0.4265355  -0.1355183  1.0000000 -0.1154457     0.1977336
Total.Fat      0.9044092  0.8077730  0.8461584   0.6805474 -0.1154457  1.0000000     0.8467072
Saturated.Fat  0.8455636  0.6030275  0.5840753   0.6312105  0.1977336  0.8467072     1.0000000
Carbohydrates  0.7815395  0.3521222  0.2007956   0.2709775  0.7623621  0.4612135     0.5912606
Dietary.Fiber  0.5388935  0.6413448  0.6943895   0.4355748 -0.2951784  0.5808373     0.3518178
                Carbohydrates Dietary.Fiber
Calories            0.7815395     0.5388935
Protein             0.3521222     0.6413448
Sodium              0.2007956     0.6943895
Cholesterol         0.2709775     0.4355748
Sugars              0.7623621    -0.2951784
Total.Fat           0.4612135     0.5808373
Saturated.Fat       0.5912606     0.3518178
Carbohydrates       1.0000000     0.2245766
Dietary.Fiber       0.2245766     1.0000000
> |
```
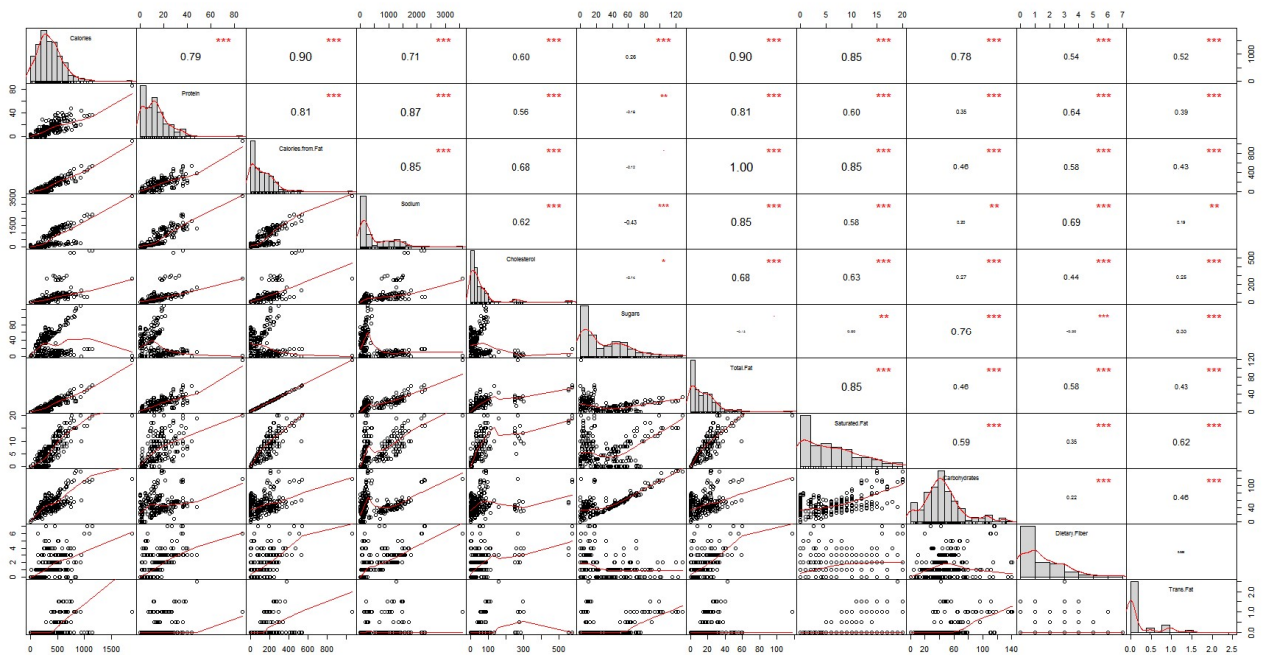
Analyzed principal variables correlation

```
> library("PerformanceAnalytics")
> chart.Correlation(Mcdonald_cor, histogram=TRUE, pch=19)
> |
```


- *The distribution of each variable is shown on the diagonal.*
- *On the bottom of the diagonal : the bivariate scatter plots with a fitted line are displayed*
- *On the top of the diagonal : the value of the correlation plus the significance level as stars*
- *Each significance level is associated to a symbol : p-values(0, 0.001, 0.01, 0.05, 0.1, 1) <=> symbols("***", "**", "*", ".",*
  " ")

4. Which category contributes to the maximum % of Cholesterol in a diet (% daily value)?

Created new subset of the data here with variable of daily cholesterol value only.

```
> Mcdonald2 <- Mcdonald[,c("Category","Cholesterol....Daily.Value.")]
> summary(Mcdonald2)
             Category Cholesterol....Daily.Value.
 Coffee & Tea    :95   Min.   :  0.00
 Breakfast       :42   1st Qu.:  2.00
 Smoothies & Shakes:28 Median : 11.00
 Beverages       :27   Mean   : 18.39
 Chicken & Fish  :27   3rd Qu.: 21.25
 Beef & Pork     :15   Max.   :192.00
 (Other)         :26
> Mcdonald2_new <- subset(Mcdonald2,Cholesterol....Daily.Value.>18)
> Mcdonald2_new
          Category Cholesterol....Daily.Value.
1        Breakfast                          87
4        Breakfast                          95
6        Breakfast                         100
7        Breakfast                          83
8        Breakfast                          83
13       Breakfast                          83
14       Breakfast                          83
19       Breakfast                          93
20       Breakfast                          84
```

We have checked the mean vlaue is 18

Hence we have ploted with more than of mean vlaue

```
255 Smoothies & Shakes                         20
> str(Mcdonald2_new)
'data.frame':  78 obs. of  2 variables:
 $ Category                   : Factor w/ 9 levels "Beef & Pork",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ Cholesterol....Daily.Value.: int  87 95 100 83 83 83 83 93 84 89 ...
> library(scales)
> library(ggplot2)
> myplot <- ggplot(Mcdonald2_new, aes(factor(Category))) +
+    geom_bar(aes(y = (..count..)/sum(..count..))) +
+    scale_y_continuous(labels=percent)
> myplot
> |
```

Breakfast is having maximum % of Cholesterol in a diet (% daily value)
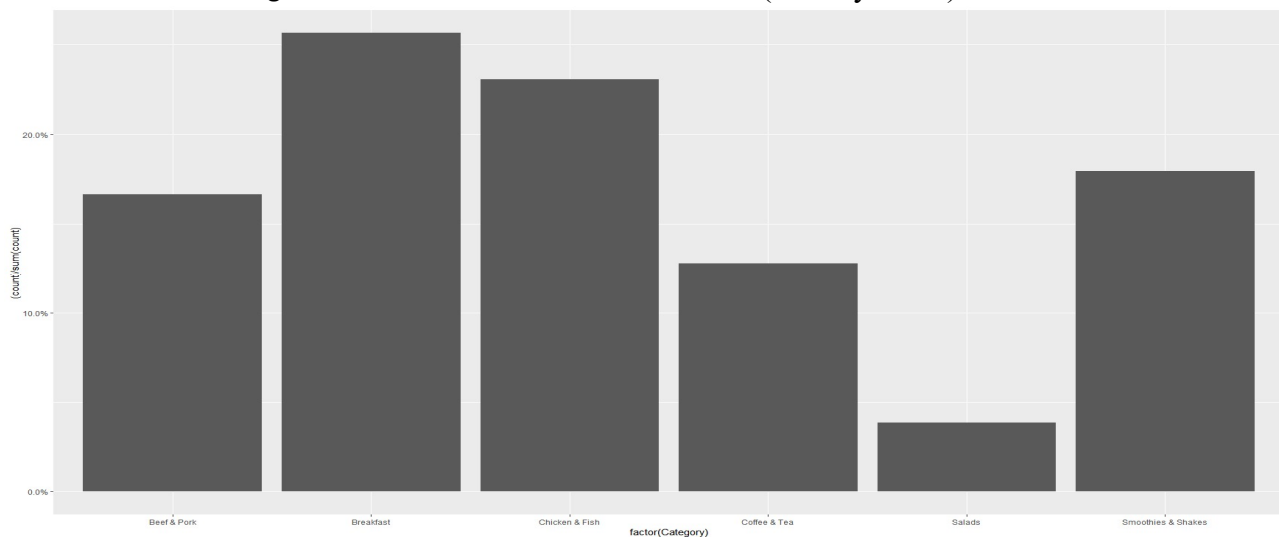
5. Which item contributes maximum to the Sodium intake?

We have created new subset of the data having only Items and Sodium

```
> names(Mcdonald)
 [1] "Category"                     "Item"                          "Serving.Size"
 [4] "Calories"                     "Calories.from.Fat"             "Total.Fat"
 [7] "Total.Fat....Daily.Value."    "Saturated.Fat"                 "Saturated.Fat....Daily.Value."
[10] "Trans.Fat"                    "Cholesterol"                   "Cholesterol....Daily.Value."
[13] "Sodium"                       "Sodium....Daily.Value."        "Carbohydrates"
[16] "Carbohydrates....Daily.Value." "Dietary.Fiber"                "Dietary.Fiber....Daily.Value."
[19] "Sugars"                       "Protein"                       "Vitamin.A....Daily.Value."
[22] "Vitamin.C....Daily.Value."    "Calcium....Daily.Value."       "Iron....Daily.Value."
> Mcdonald3 <- Mcdonald[,c("Item","Sodium")]
> summary(Mcdonald3)
                                     Item          Sodium
 1% Low Fat Milk Jug                  :  1    Min.   :   0.0
 Apple Slices                         :  1    1st Qu.: 107.5
 Bacon Buffalo Ranch McChicken        :  1    Median : 190.0
 Bacon Cheddar McChicken              :  1    Mean   : 495.8
 Bacon Clubhouse Burger               :  1    3rd Qu.: 865.0
 Bacon Clubhouse Crispy Chicken Sandwich:  1  Max.   :3600.0
 (Other)                              :254
>
```

```
> Mcdonald3_new <- subset(Mcdonald3,Sodium>1000)
> Mcdonald3_new
 > str(Mcdonald2_new)
'data.frame':    78 obs. of  2 variables:
 $ Category                 : Factor w/ 9 levels "Beef & Pork",..
 $ Cholesterol....Daily.Value.: int  87 95 100 83 83 83 83 93 84 89 ...

>
```

	Item	Sodium
7	Bacon, Egg & Cheese Biscuit (Regular Biscuit)	1300
8	Bacon, Egg & Cheese Biscuit (Large Biscuit)	1410
9	Bacon, Egg & Cheese Biscuit with Egg Whites (Regular Biscuit)	1300
10	Bacon, Egg & Cheese Biscuit with Egg Whites (Large Biscuit)	1420
11	Sausage Biscuit (Regular Biscuit)	1080
12	Sausage Biscuit (Large Biscuit)	1190
13	Sausage Biscuit with Egg (Regular Biscuit)	1170
14	Sausage Biscuit with Egg (Large Biscuit)	1280
15	Sausage Biscuit with Egg Whites (Regular Biscuit)	1180
16	Sausage Biscuit with Egg Whites (Large Biscuit)	1290
17	Southern Style Chicken Biscuit (Regular Biscuit)	1180
18	Southern Style Chicken Biscuit (Large Biscuit)	1290
19	Steak & Egg Biscuit (Regular Biscuit)	1470
20	Bacon, Egg & Cheese McGriddles	1250
21	Bacon, Egg & Cheese McGriddles with Egg Whites	1250
22	Sausage McGriddles	1030
23	Sausage, Egg & Cheese McGriddles	1320
24	Sausage, Egg & Cheese McGriddles with Egg Whites	1320
25	Bacon, Egg & Cheese Bagel	1480
26	Bacon, Egg & Cheese Bagel with Egg Whites	1480
27	Steak, Egg & Cheese Bagel	1510
28	Big Breakfast (Regular Biscuit)	1560
29	Big Breakfast (Large Biscuit)	1680
30	Big Breakfast with Egg Whites (Regular Biscuit)	1590
31	Big Breakfast with Egg Whites (Large Biscuit)	1700
32	Big Breakfast with Hotcakes (Regular Biscuit)	2150
33	Big Breakfast with Hotcakes (Large Biscuit)	2260
34	Big Breakfast with Hotcakes and Egg Whites (Regular Biscuit)	2170
35	Big Breakfast with Hotcakes and Egg Whites (Large Biscuit)	2290
44	Quarter Pounder with Cheese	1100
45	Quarter Pounder with Bacon & Cheese	1440

46	Quarter Pounder with Bacon Habanero Ranch	1180
48	Double Quarter Pounder with Cheese	1280
51	Double Cheeseburger	1040
52	Bacon Clubhouse Burger	1470
54	Bacon McDouble	1110
56	Jalapeño Double	1030
60	Premium Crispy Chicken Club Sandwich	1410
61	Premium Grilled Chicken Club Sandwich	1250
62	Premium Crispy Chicken Ranch BLT Sandwich	1400
63	Premium Grilled Chicken Ranch BLT Sandwich	1230
64	Bacon Clubhouse Crispy Chicken Sandwich	1720
65	Bacon Clubhouse Grilled Chicken Sandwich	1560
68	Bacon Cheddar McChicken	1260
69	Bacon Buffalo Ranch McChicken	1260
71	Premium McWrap Chicken & Bacon (Crispy Chicken)	1540
72	Premium McWrap Chicken & Bacon (Grilled Chicken)	1370
73	Premium McWrap Chicken & Ranch (Crispy Chicken)	1340
74	Premium McWrap Chicken & Ranch (Grilled Chicken)	1170
75	Premium McWrap Southwest Chicken (Crispy Chicken)	1480
76	Premium McWrap Southwest Chicken (Grilled Chicken)	1320
77	Premium McWrap Chicken Sweet Chili (Crispy Chicken)	1260
78	Premium McWrap Chicken Sweet Chili (Grilled Chicken)	1090
82	Chicken McNuggets (20 piece)	1800
83	Chicken McNuggets (40 piece)	3600

We found the above items are having maximum of Sodium intake.

Maximum of Sodium intake items
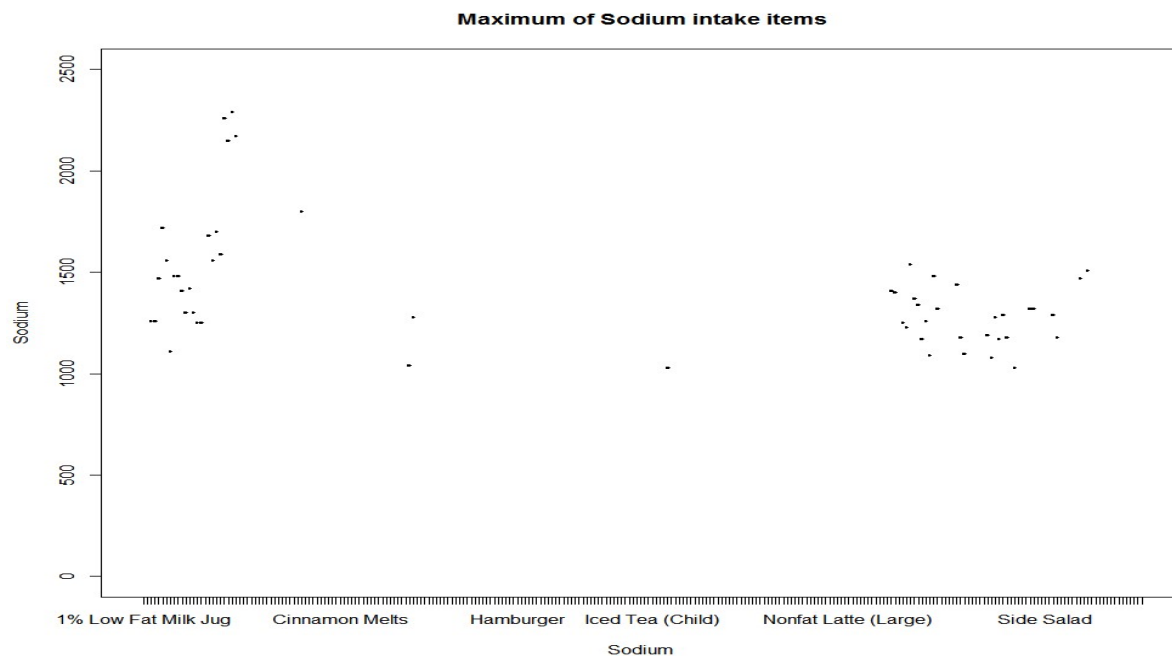
6. Which 4 food items contains the most amount of Saturated Fat?

We have prepared a new subset with Saturated Fat variable where we have considered all item which carries at least value of Saturated Fat. Hence this has been filtered out with greater than one

```
[ . cuoncu gctoption( max.pr int )       umritteu 199 rows ]
> Mcdonald4_new <- Mcdonald4[,c("Item","Saturated.Fat")]
> Mcdonald4_new
```

We found here total 194 observations

```
> str(Mcdonald4_new)
'data.frame':   194 obs. of  2 variables:
 $ Item         : Factor w/ 260 levels "1% Low Fat Milk Jug",..: 76 77 228 229 230 245 12 11 14 13 ...
 $ Saturated.Fat: num  5 3 8 10 8 9 13 14 11 12 ...
>
```

The maximum vlaues we found it is 20

```
ZUU                        MCI Tulily With RCCSC 3 I Cullut DuLLCI Cup3 (SHaCK)
> summary(Mcdonald4)
                                             Item          Saturated.Fat
 1% Low Fat Milk Jug                          :  1    Min.   : 1.500
 Bacon Buffalo Ranch McChicken                :  1    1st Qu.: 4.500
 Bacon Cheddar McChicken                      :  1    Median : 7.000
 Bacon Clubhouse Burger                       :  1    Mean   : 8.028
 Bacon Clubhouse Crispy Chicken Sandwich :  1    3rd Qu.:11.000
 Bacon Clubhouse Grilled Chicken Sandwich:  1    Max.   :20.000
 (Other)                                      :188
> Mcdonald_new4 <- subset(Mcdonald4,Saturated.Fat>19)
> Mcdonald_new4
                                          Item Saturated.Fat
33  Big Breakfast with Hotcakes (Large Biscuit)            20
83                  Chicken McNuggets (40 piece)            20
232              FrappÃ© Chocolate Chip (Large)            20
254      McFlurry with M&Mâ€™s Candies (Medium)            20
>
```

We have sorted out all Saturated . Fat values which is exceed with the value of 3rd quartile

```
> Mcdonald5_new <- subset(Mcdonald4_new,Saturated.Fat>11)
> Mcdonald5_new
                                                         Item Saturated.Fat
7                Bacon, Egg & Cheese Biscuit (Regular Biscuit)            13
8                  Bacon, Egg & Cheese Biscuit (Large Biscuit)            14
10   Bacon, Egg & Cheese Biscuit with Egg Whites (Large Biscuit)            12
11                    Sausage Biscuit (Regular Biscuit)            12
12                    Sausage Biscuit (Large Biscuit)            13
13         Sausage Biscuit with Egg (Regular Biscuit)            14
14           Sausage Biscuit with Egg (Large Biscuit)            15
15   Sausage Biscuit with Egg Whites (Regular Biscuit)            12
16     Sausage Biscuit with Egg Whites (Large Biscuit)            13
19             Steak & Egg Biscuit (Regular Biscuit)            16
23                      Sausage, Egg & Cheese McGriddles            12
```

```
27                                   Steak, Egg & Cheese Bagel              13
28                              Big Breakfast (Regular Biscuit)             17
29                                Big Breakfast (Large Biscuit)             18
30               Big Breakfast with Egg Whites (Regular Biscuit)            14
31                 Big Breakfast with Egg Whites (Large Biscuit)            14
32                Big Breakfast with Hotcakes (Regular Biscuit)             19
33                  Big Breakfast with Hotcakes (Large Biscuit)             20
34   Big Breakfast with Hotcakes and Egg Whites (Regular Biscuit)          16
35     Big Breakfast with Hotcakes and Egg Whites (Large Biscuit)          16
44                                  Quarter Pounder with Cheese             12
45                           Quarter Pounder with Bacon & Cheese           13
46                      Quarter Pounder with Bacon Habanero Ranch          13
48                           Double Quarter Pounder with Cheese            19
52                                     Bacon Clubhouse Burger              15
83                                  Chicken McNuggets (40 piece)           20
193                                      Hot Chocolate (Large)             12
224                                       FrappÃ© Mocha (Small)            12
225                                      FrappÃ© Mocha (Medium)            14
226                                       FrappÃ© Mocha (Large)            17
227                                     FrappÃ© Caramel (Small)            12
228                                    FrappÃ© Caramel (Medium)            15
229                                     FrappÃ© Caramel (Large)            17
230                              FrappÃ© Chocolate Chip (Small)            14
231                             FrappÃ© Chocolate Chip (Medium)            17
232                              FrappÃ© Chocolate Chip (Large)            20
243                                      Vanilla Shake (Medium)            12
244                                       Vanilla Shake (Large)            15
246                                   Strawberry Shake (Medium)            13
247                                    Strawberry Shake (Large)            15
249                                    Chocolate Shake (Medium)            12
250                                     Chocolate Shake (Large)            15
251                                     Shamrock Shake (Medium)            12
252                                      Shamrock Shake (Large)            15
253                         McFlurry with M&Mâ€™s Candies (Small)          14
254                        McFlurry with M&Mâ€™s Candies (Medium)          20
257                            McFlurry with Oreo Cookies (Medium)         12
259              McFlurry with Reese's Peanut Butter Cups (Medium)         15
```

>

These are the items which carry out most highest saturated . Fat

```
> Mcdonald5_new1 <- subset(Mcdonald4_new,Saturated.Fat>19)
> Mcdonald5_new1

                                     Item Saturated.Fat
33   Big Breakfast with Hotcakes (Large Biscuit)            20
83                 Chicken McNuggets (40 piece)            20
232              FrappÃ© Chocolate Chip (Large)            20
254        McFlurry with M&Mâ€™s Candies (Medium)          20
>
```

Appendix A – Source Code

```r
 1  #Mcdonald's Competition - Mini Hackathon#
 2  Mcdonald<- read.csv("C:/Users/IBM_ADMIN/Documents/My Received Files/r file/R Programming/Mcdonald.csv", head
 3  Mcdonald
 4  str(Mcdonald)
 5  head(Mcdonald)
 6  levels(Category)
 7  summary(Mcdonald)
 8  attach(Mcdonald)
 9  colSums(is.na(Mcdonald))
10  plot(Mcdonald$Category,col = "Blue", xlab ="Food Categories ", main=" Comparision of Varities ",ylim = c(0,
11  # Find out the outliers
12  names(Mcdonald)
13  boxplot(Mcdonald,horizontal = TRUE,col=c("Red","Blue"),ylab ="Variables", main=" Box plot of Variables",ylim
14  boxplot(Cholesterol,Sodium,Carbohydrates,Sugars,Protein,Calories,Iron....Daily.Value.,horizontal = TRUE,col=
15  par(mfrow= c(1,3))
16  boxplot(Vitamin.C....Daily.Value.,col= "Red",xlab ="Vitamin.C....Daily.Value.", main=" Box plot of Variables
17  boxplot(Vitamin.A....Daily.Value. ,col= "Green",xlab ="Vitamin.A....Daily.Value. ", main=" Box plot of Varia
18  boxplot(Cholesterol,col= "Blue",xlab ="Cholesterol", main=" Box plot of Variables ",ylim = c(0,350))
19  # Which variables have the Highest correlation. Plot them and tell the value.
20  cor(Mcdonald[,unlist(lapply(Mcdonald, is.numeric))])
21  library(corrplot)
22  Mcdonald_cor <- Mcdonald[,c("Calories","Protein","Calories.from.Fat","Sodium","Cholesterol","Sugars","Total.
23  Mcdonald_cor
24  library("PerformanceAnalytics")
25  chart.Correlation(Mcdonald_cor, histogram=TRUE, pch=19)
26  cors <- cor(subset(Mcdonald, select = c(Calories,Protein,Sodium,Cholesterol,Sugars,Total.Fat,Saturated.Fat,
27  cors
28
29  #Which category contributes to the maximum % of Cholesterol in a diet (% daily value)?
30  Mcdonald2 <- Mcdonald[,c("Category","Cholesterol....Daily.Value.")]
31  summary(Mcdonald2)
32  Mcdonald2_new <- subset(Mcdonald2,Cholesterol....Daily.Value.>18)
33  Mcdonald2_new
34  str(Mcdonald2_new)
35  library(scales)
36  library(ggplot2)
37  myplot <- ggplot(Mcdonald2_new, aes(factor(Category))) +
38    geom_bar(aes(y = (..count..)/sum(..count..))) +
39    scale_y_continuous(labels=percent)
40  myplot
41  #.  Which item contributes maximum to the Sodium intake?
42  names(Mcdonald)
43  Mcdonald3 <- Mcdonald[,c("Item","Sodium")]
44  summary(Mcdonald3)
45  Mcdonald3_new <- subset(Mcdonald3,Sodium>1000)
46  Mcdonald3_new
47  plot(Mcdonald3_new,col = "Blue", xlab ="Sodium ", main=" Maximum of Sodium intake items ",ylim = c(0,2500))
48  str(Mcdonald2_new)
49
50  #Which 4 food items contains the most amount of Saturated Fat?
51  Mcdonald4 <- subset(Mcdonald,Saturated.Fat > 1)
52  Mcdonald4
53  Mcdonald4_new <- Mcdonald4[,c("Item","Saturated.Fat")]
54  Mcdonald4_new
55  str(Mcdonald4_new)
56  summary(Mcdonald4_new)
57  Mcdonald5_new <- subset(Mcdonald4_new,Saturated.Fat>11)
58  Mcdonald5_new
59  Mcdonald5_new1 <- subset(Mcdonald4_new,Saturated.Fat>19)
60  Mcdonald5_new1
```

Mcdonald.R