Assignment 3-2

File Name: Assign32_LastName.doc

Using R, please conduct the following tasks. Create your report in a Word document, and upload it to Blackboard. In your report, provide the source code of R, the screen capture of the output, and some descriptions of the outputs.

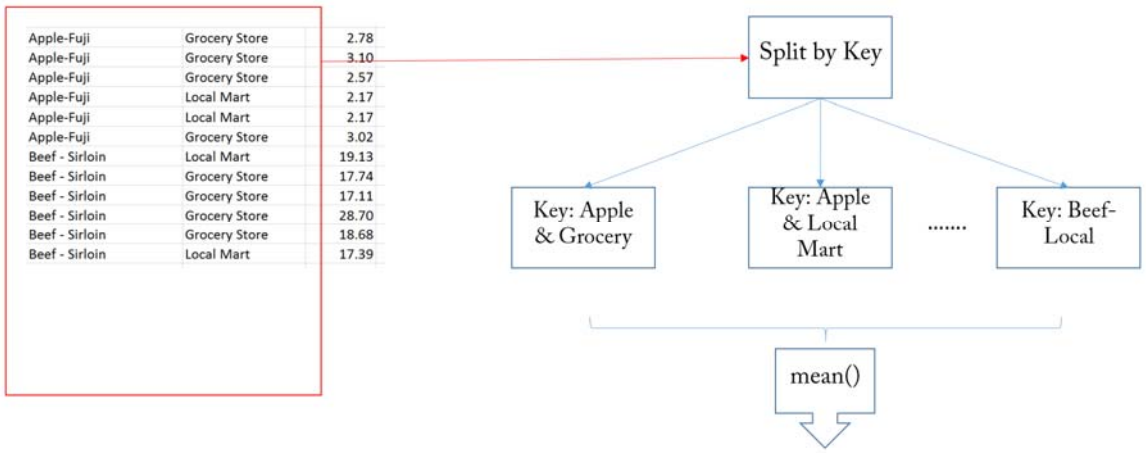
Instructions:

1. Download the data file named "M2.csv" and read the data into your project. Name the data as "market" and explore the data.

```
   P_SEQ M_SEQ M_NAME A_SEQ          A_NAME    A_UNIT Price P_YEAR_MONTH M_TYPE_CODE M_TYPE_NAME
1 319572   199     SC   305      Apple-Fuji         1  2.17       12-Jul           1  Local Mart
2 319573   199     SC   306           Pear          1  2.61       12-Jul           1  Local Mart
3 319574   199     SC   307 Chiness cabbage        1  4.35       12-Jul           1  Local Mart
4 319575   199     SC   308          Radish        1  1.30       12-Jul           1  Local Mart
5 319576   199     SC   309          Onions    6 lbs  1.74       12-Jul           1  Local Mart
6 319577   199     SC    23         Lettuce 3.5 ounce  1.04       12-Jul           1  Local Mart
  M_GU_CODE City
1    140000   JG
2    140000   JG
3    140000   JG
4    140000   JG
5    140000   JG
6    140000   JG
```

*P_SEQ: id, M_SEQ: Store Number, M_NAME: Mart type, A_SEQ: item number, A_NAME: item name, A_UNIT: unit, Price: price, M_TYPE_CODE: mart type code

2. Apply the "Split-Apply-Combine" approach to the below figure using "tapply," "by," and "ddply" functions.



The output will look like the figure below when tapply( ) is applied:

```
                A_NAME
M_TYPE_NAME      Apple Apple-Fuji Beef - Sirloin  Chicken Chiness cabbage   Cod Cucumber   Egg
  Grocery Store 2.150   3.433636        20.8925 7.764167        1.774167 5.030   0.4825 2.420
  Local Mart    2.045   2.372500        17.5980 5.111000        2.904000 2.667   0.3960 1.906
                A_NAME
M_TYPE_NAME      Frozen pollack  Lettuce Mussel   Onions     Pear  Pollack     Pork   Radish  squash
  Grocery Store           0.86 1.195833 2.1775 2.783333 4.313333 1.310000 15.46083 1.661667 1.04125
  Local Mart              4.26 0.690000 2.3030 2.075000 2.646000 2.173333 11.55200 1.517000 0.63625
                A_NAME
M_TYPE_NAME      Yellow corbina
  Grocery Store       3.491818
  Local Mart          2.732000
```

by( ) produces the following output:

```
M_TYPE_NAME: Grocery Store
A_NAME: Apple
[1] 2.15
------------------------------------------------------------------------------
M_TYPE_NAME: Local Mart
A_NAME: Apple
[1] 2.045
------------------------------------------------------------------------------
M_TYPE_NAME: Grocery Store
A_NAME: Apple-Fuji
[1] 3.433636
------------------------------------------------------------------------------
M_TYPE_NAME: Local Mart
A_NAME: Apple-Fuji
[1] 2.3725
------------------------------------------------------------------------------
```

Using the ddply() function, summarize the data like the figure below:

```
##   "M_TYPE_NAME" "A_NAME"             P_SEQ          M_SEQ       M_NAME
## 1   M_TYPE_NAME   A_NAME Min.   :319572   Min.   :199   SC      :1
## 2   M_TYPE_NAME   A_NAME 1st Qu.:319572   1st Qu.:199   Chung   :0
## 3   M_TYPE_NAME   A_NAME Median :319572   Median :199   Doggok  :0
## 4   M_TYPE_NAME   A_NAME Mean   :319572   Mean   :199   Dow     :0
## 5   M_TYPE_NAME   A_NAME 3rd Qu.:319572   3rd Qu.:199   Emart   :0
## 6   M_TYPE_NAME   A_NAME Max.   :319572   Max.   :199   Emart Y:0
## 7   M_TYPE_NAME   A_NAME              <NA>           <NA> (Other):0
##           A_SEQ               A_NAME       A_UNIT          Price
## 1 Min.   :305   Apple-Fuji      :1   1       :1   Min.   :2.17
## 2 1st Qu.:305   Apple           :0   1 box   :0   1st Qu.:2.17
## 3 Median :305   Beef - Sirloin  :0   1.3 lbs :0   Median :2.17
## 4 Mean   :305   Chicken         :0   10 lbs  :0   Mean   :2.17
## 5 3rd Qu.:305   Chiness cabbage:0   11 lbs  :0   3rd Qu.:2.17
## 6 Max.   :305   Cod             :0   12      :0   Max.   :2.17
## 7          <NA> (Other)         :0   (Other) :0              <NA>
```

3. Install data.table package (library (data.table). Transform the current data to data table format and name the new data as "market.dt" using data.table() function. If you successfully transformed, the following output will be displayed.

```
   M_NAME
1:     SC
```

4. Using the "sqldf()" function, display the output below. An example of the sql statement is ("select . . . . , avg(Price) as avg from . . . . group by . . . ").

```
      M_TYPE_NAME          A_NAME       avg
1  Grocery Store           Apple  2.150000
2  Grocery Store      Apple-Fuji  3.433636
3  Grocery Store  Beef - Sirloin 20.892500
4  Grocery Store         Chicken  7.764167
5  Grocery Store Chiness cabbage  1.774167
6  Grocery Store             Cod  5.030000
7  Grocery Store        Cucumber  0.482500
8  Grocery Store             Egg  2.420000
9  Grocery Store  Frozen pollack  0.860000
10 Grocery Store         Lettuce  1.195833
```