Pricing of First-Exposure Items

A) In Assignment 6, you estimated the actual demand of the previously sold items. Using the estimated actual demand and the other available features in the dataset of this assignment, determine the optimal price for the first-exposure items available in the "First-exposures.csv" file. The objective of the optimization model should be to maximize the total revenue of these items and the demand should be estimated using the regression tree approach. To determine the optimal price of each item, consider {$25.99, $30.99, $35.99, $40.99} and {$45.99, $50,99} as the set of possible prices for items in event A and event B, respectively.

Answer A) From the 'Assignment6_Data.csv' dataset, there are 1,000 obs. of 29 variables. At first, we have imported the data and then converted the categorical data into numerical data. We have used the estimated actual demand and the other available features in the dataset and have found the stockout_time as given below:

```
> stockout_time
  [1] 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22
 [47] 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18
 [93] 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22
[139] 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16
[185] 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22 21 15 22 17 18 16 22 20 14 22 22 22 22
[231] 21 15 22
```

We have applied the *'Elbow Method'* to find and plot the optimal number of clusters as shown below:

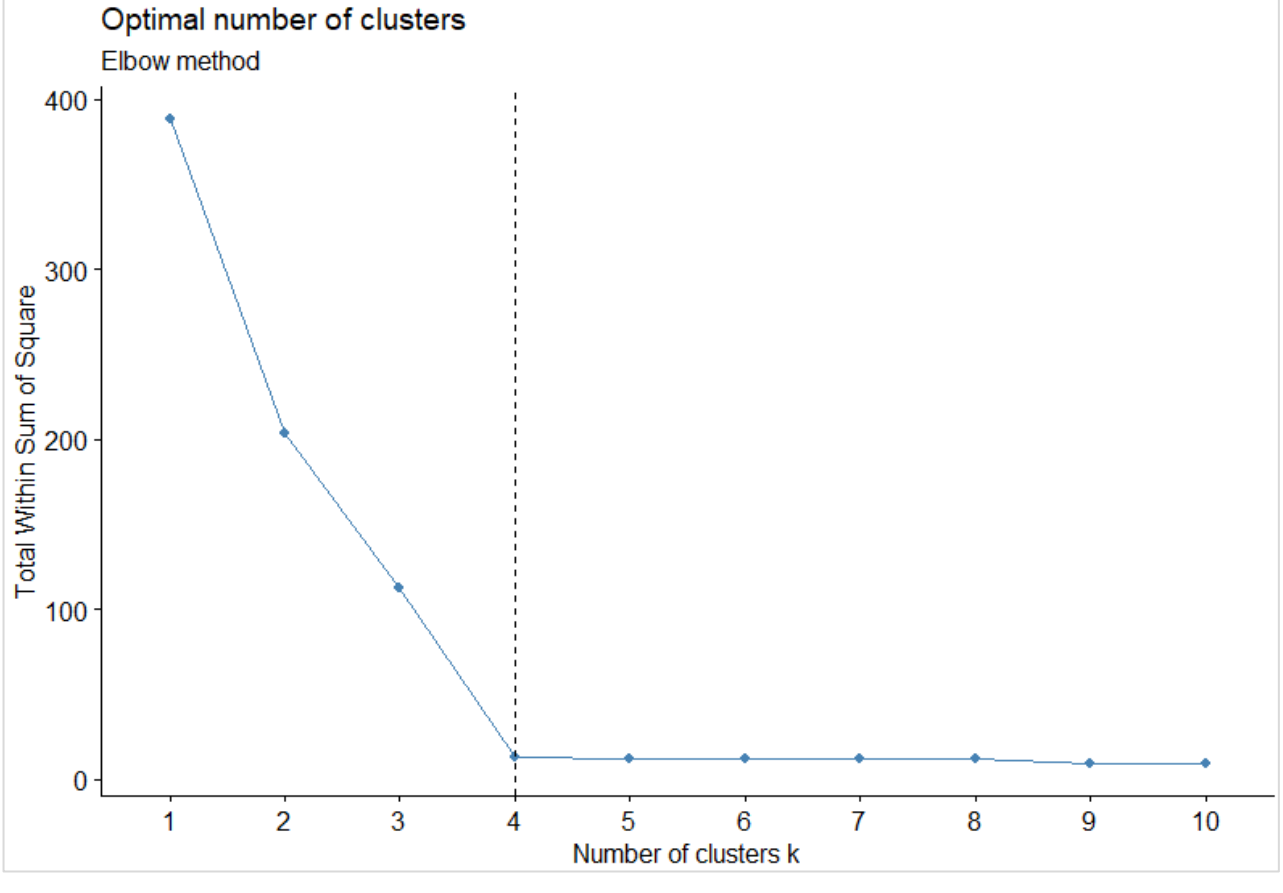

Figure 01: *Number of clusters using K-Means*

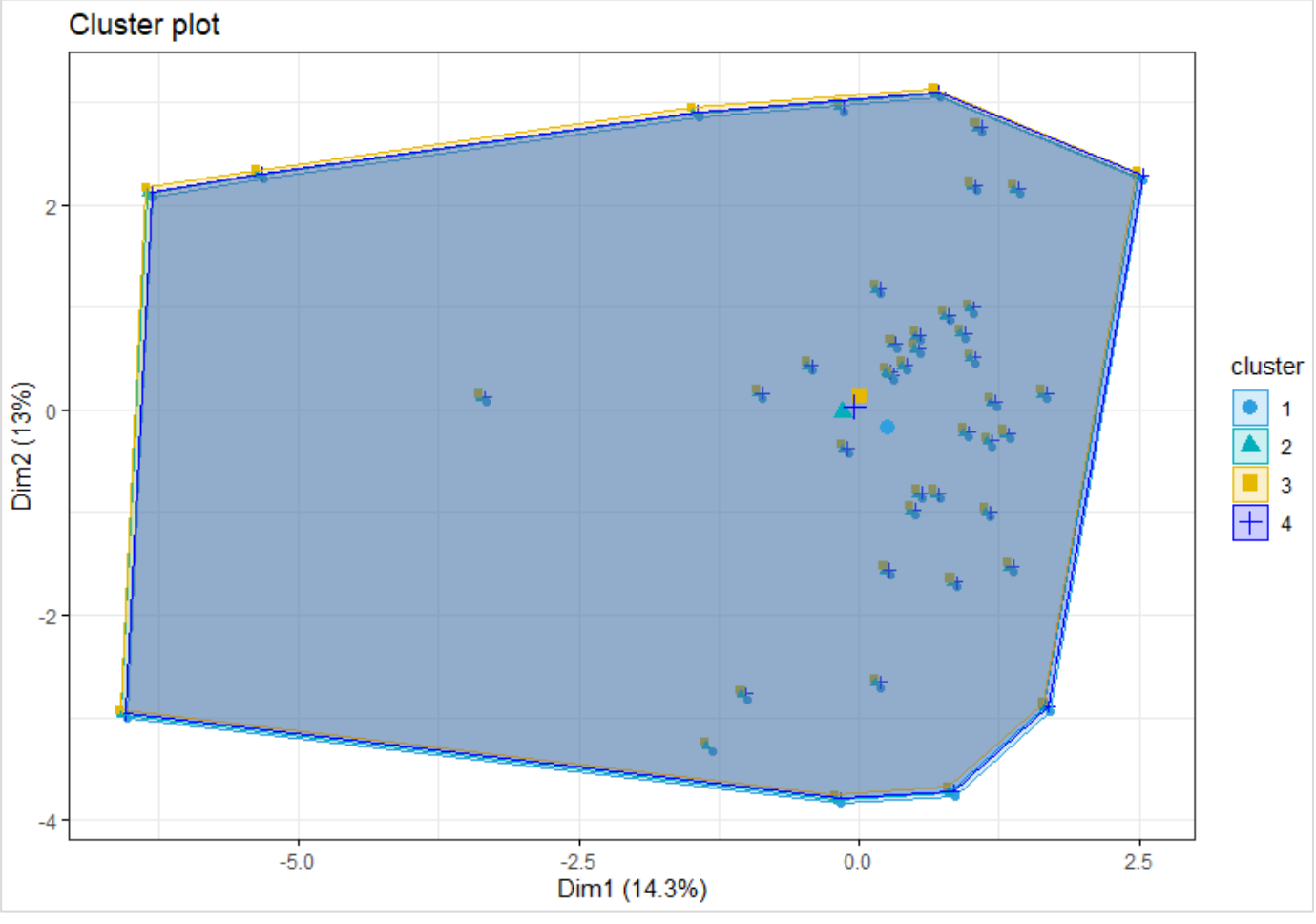

Figure 02: *Cluster plots of not_stockouts*

Then we have generated the output of the 'Stockouts' and 'Not_Stockouts' in CSV format and this has been submitted with this assignment. The clusters center output is given below:

```
> clusters$center
  Dep_category Part_of_day    hour.1    hour.2     hour.3     hour.4     hour.5     hour.6     hour.7     hour.8     hour.9
1            1           0 0.2568421 0.1385167 0.07392344 0.06622010 0.05440191 0.04344498 0.04043062 0.03856459 0.02287081
2            1           1 0.2557769 0.1375697 0.06832669 0.07083665 0.05430279 0.04462151 0.03924303 0.03984064 0.02382470
3            0           1 0.2668902 0.1382927 0.07347561 0.06951220 0.05378049 0.04256098 0.04134146 0.04140244 0.02560976
4            0           0 0.2622378 0.1329371 0.07146853 0.07405594 0.05041958 0.04020979 0.04020979 0.04069930 0.02692308
     hour.10    hour.11    hour.12    hour.13    hour.14     hour.15    hour.16     hour.17     hour.18     hour.19
1 0.02727273 0.02588517 0.02368421 0.02100478 0.01363636 0.008995215 0.01105263 0.006889952 0.004928230 0.004114833
2 0.02685259 0.02601594 0.02179283 0.02043825 0.01310757 0.009681275 0.01258964 0.007370518 0.004780876 0.004501992
3 0.02603659 0.02554878 0.02128049 0.02134146 0.01207317 0.008475610 0.01085366 0.005670732 0.003475610 0.002743902
4 0.02615385 0.02650350 0.02181818 0.02041958 0.01188811 0.009720280 0.00965035 0.007832168 0.004055944 0.004195804
     hour.20     hour.21     hour.22     hour.23     hour.24
1 0.005358852 0.008373206 0.006985646 0.004784689 0.006076555
2 0.005896414 0.008286853 0.008007968 0.004223108 0.006653386
3 0.004451220 0.006646341 0.006341463 0.003353659 0.006585366
4 0.006293706 0.008251748 0.007132867 0.004405594 0.006503497
```

We have used the ANOVA method in regression tree approach, thus getting the output as shown below:
The output of the regression tree is as following:

```
> regression_tree
n= 1000

node), split, n, deviance, yval
      * denotes terminal node

 1) root 1000 1522537.000  90.34239
   2) Price>=59.95 609  535513.500  80.24741
     4) Branded< 0.5 139   39087.820  67.05189 *
     5) Branded>=0.5 470  465064.900  84.14992
      10) Price>=164.9 78   24983.970  69.52501 *
      11) Price< 164.9 392  420078.000  87.05997 *
   3) Price< 59.95 391  828296.600 106.06580
     6) Num_Branded_Events12< 7.5 371  512832.500 102.82610
      12) Num_Branded_Events12< 0.5 107  128193.000  88.99956
        24) Brand_MSRP_Index>=0.575 84   33364.150  80.28134 *
        25) Brand_MSRP_Index< 0.575 23   65126.470 120.84000
          50) ConcurrentEvents>=2.5 11    5819.847  81.16605 *
          51) ConcurrentEvents< 2.5 12   26120.970 157.20780 *
      13) Num_Branded_Events12>=0.5 264  355893.300 108.43000 *
     7) Num_Branded_Events12>=7.5 20  239340.300 166.16150
      14) Event_Length< 2.5 12    9724.686 108.80040 *
      15) Event_Length>=2.5 8  130906.800 252.20310 *
```

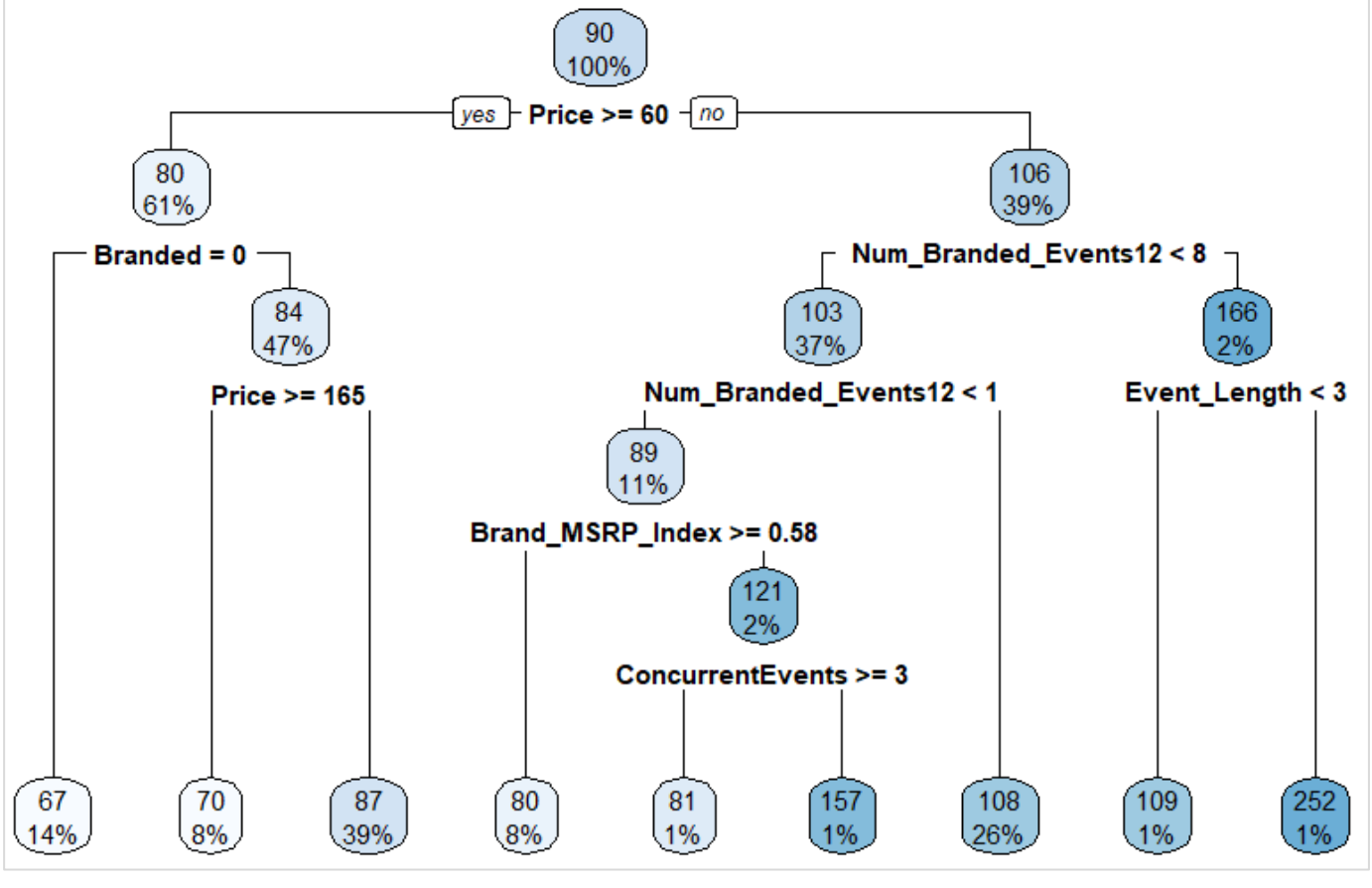

Figure 03: *Output using the ANOVA method using regression tree approach*

For Event A – we have put the set of possible prices for items and thus found the optimal prices, which has shown below:

```
> D1
 [1] 108.43004 108.43004 157.20781  87.05997  87.05997  87.05997  87.05997  87.05997  87.05997  87.05997  87.05997
[12]  87.05997
> Final_solutions_A
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    1    0    0    0    1    0    0    0    1     0     0     0
 [2,]    0    1    0    0    1    0    0    0    1     0     0     0
 [3,]    0    0    1    0    1    0    0    0    1     0     0     0
 [4,]    0    0    1    0    0    1    0    0    1     0     0     0
 [5,]    0    0    1    0    0    0    1    0    1     0     0     0
 [6,]    0    0    1    0    1    0    0    0    0     0     0     1
 [7,]    0    0    1    0    0    0    0    1    0     1     0     0
 [8,]    0    0    1    0    0    0    1    0    0     0     0     1
 [9,]    0    0    1    0    0    0    0    1    0     0     0     1
[10,]    0    0    0    1    0    0    0    1    0     0     0     1
> Final_objectives_A
 [1]  7343.474  7885.624 10183.286 10618.586 11053.886 11489.186 11924.486 12359.786 12795.086 10705.765
> Final_solutions_A[match(max(Final_objectives_A),Final_objectives_A),]
 [1] 0 0 1 0 0 0 0 1 0 0 0 1
```

For Event B – we have put the set of possible prices for items and thus found the optimal prices, which has shown below:

```
> D2
[1] 108.43004 108.43004  87.05997  87.05997
> Final_solutions_B
     [,1] [,2] [,3] [,4]
[1,]    1    0    1    0
[2,]    0    1    1    0
[3,]    0    1    0    1
> Final_objectives_B
[1] 8990.586 9532.736 9968.036
> Final_solutions_B[match(max(Final_objectives_B),Final_objectives_B),]
[1] 0 1 0 1
```

Bibliographic Reference:

Creed J. (2016) *Optimal Pricing and Predicting Demand* Retrieved from https://rpubs.com/jcreed/opplastersand

Monroe J. T. (2021) *Business Analytics with R* Retrieved from https://bookdown.org/jeffreytmonroe/business_analytics_with_r7/marketing.html

Laurinec P. August 22, 2017 *Using regression trees for forecasting double-seasonal time series with trend in R* Retrieved from https://petolau.github.io/Regression-trees-for-forecasting-time-series-in-R