1. Task Description

Hello,

Blackwell Electronics' board of directors is considering acquiring Electronidex, a start-up electronics online retailer. The board of directors has asked us to help them better understand the clientele that Electronidex currently is serving and if it would be an optimal partnership.

They need our help to identify purchasing patterns that will provide insight into Electronidex's clientele. Attached is a CSV file that contains one month's (30 days' worth) of Electronidexes online transactions and a file containing all the electronics that they currently sell. Due to their lack of funding, Electronidex is only able to pull data on the items that customers purchased per their transactions.

I would like you to to use R to conduct a market basket analysis. You will be discovering any interesting relationships (or associations) between customer's transactions and the item(s) they've purchased. These associations can then be used to drive sales-oriented initiatives such as recommender systems like the ones used by Amazon and other eCommerce sites.

To help Blackwell's board of directors form a clearer picture of Electronidex's customer buying patterns, please consider the following questions while you're performing your analysis:

- Are there any interesting patterns or item relationships within Electronidex's transactions?
- Would Blackwell benefit from selling any of Electronidex's items?
- In your opinion, should Blackwell acquire Electronidex?
- If Blackwell does acquire Electronidex, do you have any recommendations for Blackwell? (Ex: cross-selling items, sale promotions, should they remove items, etc.)

Once you've completed your market basket analysis, please put together a formal business report in Word. Thank you in advance!

Best,

Danielle Sherman
Chief Technology Officer
Blackwell Electronics
www.blackwellelectronics.com

2. Task Solution

Two data sets (ElectronidexItems2017.xlsx and ElectronidexItems2017.pdf) were provided by the CTO to help Blackwell's board of directors to get a clear picture of Electronidex's customer buying patterns, the final deliverables are:

- Are there any interesting patterns or item relationships within Electronidex's transactions?
- Would Blackwell benefit from selling any of Electronidex's items?
- In your opinion, should Blackwell acquire Electronidex?
- If Blackwell does acquire Electronidex, do you have any recommendations for Blackwell? (Ex: cross-selling items, sale promotions, should they remove items, etc.)

In order to solve this task, the list below, are the general steps taken to approach the solution.

2.1 Steps Sequence

Step 1 – Exploratory Analysis

The exploratory analysis give us ideas of the level of analysis and data cleaning requirements. Two levels were defined: Categories (18) and Products (125). The data cleaning didn't required efforts, a product rearrangement within two categories was conducted. Two items were identified as software ("Computer Game" and "Microsoft Office Home and Student 2016"). As a result, a new category was created and these two items were placed on it. The product "iPhone Charger Cable" was moved from computer cords to accessories. The product rearrangement is listed in figure 1.

Categories	Qty of Products
Monitors	10
Laptops	10
Computer Headphones	10
Computer Mice	10
Mouse and Keyboard Combo	9
Computer Cords	9
Speakers	9
Keyboard	9
Desktop	9
Active Headphones	6
Printers	5
Printer Ink	5
Smart Home Devices	5
Computer Tablets	5
Computer Stands	5
External Hardrives	5
Accessories	4
Total	**125**

Categories	Qty of Products
Monitors	10
Laptops	10
Computer Headphones	10
Computer Mice	10
Mouse and Keyboard Combo	9
Computer Cords	8
Speakers	9
Keyboard	9
Desktop	9
Active Headphones	6
Printers	5
Printer Ink	5
Smart Home Devices	5
Computer Tablets	5
Computer Stands	5
External Hardrives	5
Accessories	3
Software	2
Total	**125**

Figure 1 – Categories rearrangement

Step 2 – Create frequency plots to determine the best rules to discover product associations

Step 3 – Rules creation and analysis

Step 4 – Final recommendations

3. Results

The market basket analysis was conducted with the transaction file, a frequency charts (Figure 2) were develop showing some patterns in regards the quantity of transactions per products.

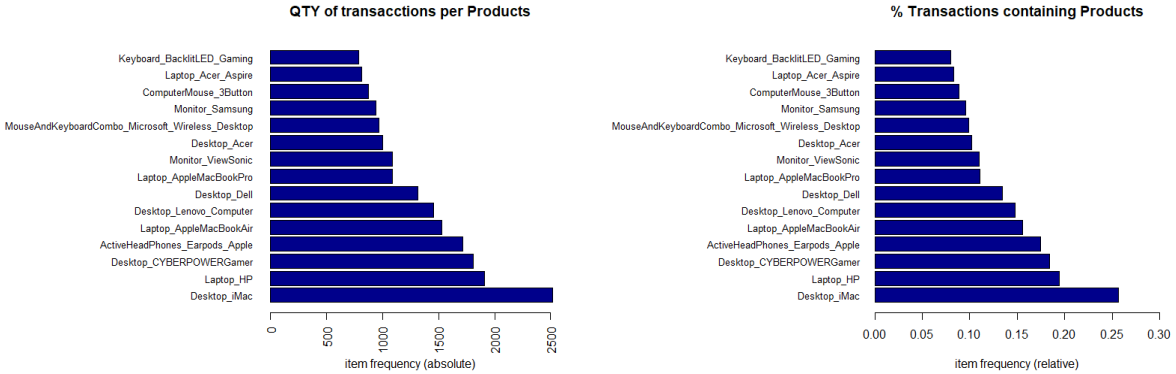


Figure 2 – Item frequency for product categories in terms of transactions.

In order to clearly identify the most frequent items, two levels of analysis were created (Level I - Categories and Level II - Products).

Level I – Analysis per Category

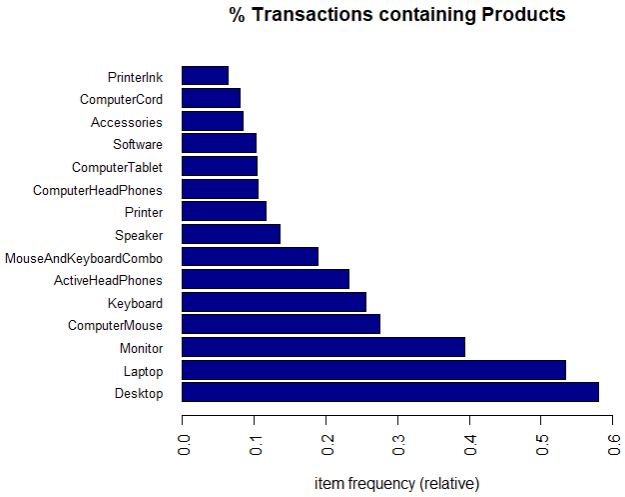


Figure 3 – Analysis per Category

Figure 3 shows that Desktops, Laptops and Monitors are the most frequent categories in terms of transactions. So, we can clearly identify main products and complementary products, since customers that first buys a computer tends to buy a mouse and not the opposite. Therefore, Desktops and Laptops will be considered “main products” and the other categories will be considered as “complementary products”.

Figure 4 contains a matrix with the categories designated as “main” and “complementary” products, the products considered “complementary” were included as recommended products after a client purchasing a main product.

Categories	Main / Complementary
Desktop	Main Product
Laptops	Main Product
Computer Headphones	Complementary Product
Computer Mice	Complementary Product
Mouse and Keyboard Combo	Complementary Product
Computer Cords	Complementary Product
Speakers	Complementary Product
Keyboard	Complementary Product
Monitors	Complementary Product
Active Headphones	Complementary Product
Printers	Complementary Product
Printer Ink	Complementary Product
Smart Home Devices	Complementary Product
Computer Tablets	Complementary Product
Computer Stands	Complementary Product
External Hardrives	Complementary Product
Accessories	Complementary Product
Software	Complementary Product

Figure 4 – Main Product and Complementary Products

Rule # 1 – Productcategory: Based on the statements described above, rule # 1 (Figure 5) was created subtracting the Laptops and Desktops categories from the rhs side of the rule, with support: 0.025 and confidence: 0.1

```
> inspect(sort(Productcategory.rules, by='confidence', decreasing = T)[1:10])
     lhs                    rhs                  support    confidence lift     count
[1]  {Desktop,Laptop}    => {Monitor}            0.18254856 0.5689382  1.447069 1795
[2]  {Desktop}           => {Monitor}            0.27865351 0.4803647  1.221786 2740
[3]  {Laptop}            => {Monitor}            0.24051663 0.4499619  1.144458 2365
[4]  {Desktop,Laptop}    => {ComputerMouse}      0.12244483 0.3816165  1.390306 1204
[5]  {Desktop,Laptop}    => {Keyboard}           0.11471575 0.3575277  1.398954 1128
[6]  {Desktop}           => {ComputerMouse}      0.18793857 0.3239832  1.180336 1848
[7]  {Desktop}           => {Keyboard}           0.17827723 0.3073282  1.202530 1753
[8]  {Laptop}            => {ComputerMouse}      0.16332757 0.3055556  1.113200 1606
[9]  {Desktop,Laptop}    => {ActiveHeadPhones}   0.09396929 0.2928685  1.260296  924
[10] {Laptop}            => {Keyboard}           0.14736093 0.2756849  1.078715 1449
```

Figure 5 –Top 10 categories for Rule # 1 - Productcategory

The top 10 rules show customers that first buys Desktops and / or Laptops tends to buy monitors, computer mouse, keyboard and active head phones as complementary products.

Level II – Analysis per Products (items)

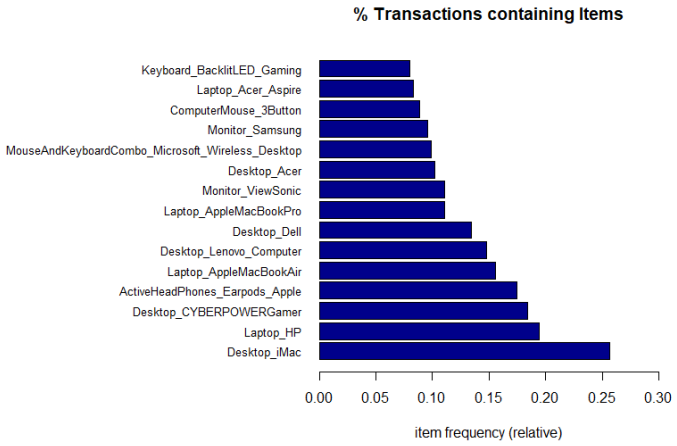


Figure 6 – Analysis per Product

Figure 6 shows the item frequency per product, for example Desktop iMac was the most transacted product.

Rule # 2 – Product Type: This rule was created to subtract the Laptops and Desktops categories from the rhs side of the rule, using support: 0.001 and confidence: 0.1

```
> inspect(sort(ProductType.rules, by='confidence', decreasing = T)[1:10])
     lhs                              rhs                     support confidence     lift count
[1]  {Laptop_Acer_Aspire,
      Laptop_ChromebookASUS,
      Desktop_Dell,
      Laptop_HP}                   => {Monitor_ViewSonic} 0.001220380  0.8571429 7.768005    12
[2]  {Laptop_ChromebookASUS,
      Desktop_CYBERPOWERGamer,
      Desktop_Dell,
      Laptop_HP}                   => {Monitor_ViewSonic} 0.001728872  0.7391304 6.698497    17
[3]  {Laptop_Acer_Aspire,
      Desktop_Acer,
      Desktop_Dell,
      Laptop_HP,
      Desktop_iMac}                => {Monitor_ViewSonic} 0.001322079  0.7222222 6.545264    13
[4]  {Laptop_Acer_Aspire,
      Laptop_ChromebookASUS,
      Desktop_Dell,
      Desktop_iMac}                => {Monitor_ViewSonic} 0.001016984  0.7142857 6.473338    10
[5]  {Laptop_Acer_Aspire,
      Laptop_ChromebookASUS,
      Laptop_HP,
      Desktop_iMac}                => {Monitor_ViewSonic} 0.001627174  0.6956522 6.304468    16
[6]  {Laptop_ChromebookASUS,
      Desktop_CYBERPOWERGamer,
      Desktop_Dell,
      Laptop_HP,
      Desktop_iMac}                => {Monitor_ViewSonic} 0.001118682  0.6875000 6.230588    11
[7]  {Laptop_Acer_Aspire,
      Laptop_ChromebookASUS,
      Laptop_HP}                   => {Monitor_ViewSonic} 0.002237364  0.6470588 5.864082    22
[8]  {Laptop_AppleMacBookAir,
      Laptop_ChromebookASUS,
      Desktop_Dell}                => {Monitor_ViewSonic} 0.001118682  0.6470588 5.864082    11
[9]  {Laptop_ChromebookASUS,
      Desktop_CYBERPOWERGamer,
      Desktop_Dell,
      Desktop_iMac}                => {Monitor_ViewSonic} 0.001423777  0.6363636 5.767155    14
[10] {Desktop_Acer,
      Laptop_HP,
      Desktop_iMac,
      Desktop_Intel}               => {Monitor_ViewSonic} 0.001220380  0.6315789 5.723793    12
```

Figure 7 – Rule # 2 – Product Type

Figure 7 shows that Monitor ViewSonic is the most frequent item when Desktop and Laptops are bought together or separately. In order to get in-depth understanding of complementary items, three additional rules were created to identity the top 10 items when Desktop iMac, Laptop HP and Desktop Cyberpower Gamer are bought.

Rule # 3 – iMac Association Rule: This rule was created to find out which items are frequently bought with iMac Desktops. Laptops and Desktops categories were removed from the rhs side of the rule and lhs = iMac Desktop, using support: 0.001 and confidence: 0.1.

```
> inspect(sort(imac.association.rules, by='confidence', decreasing = T)[1:10])
     lhs               rhs                                        support    confidence lift      count
[1]  {Desktop_iMac} => {Monitor_ViewSonic}                        0.04942540 0.1929337  1.7484950 486
[2]  {Desktop_iMac} => {ActiveHeadPhones_Earpods_Apple}           0.04027255 0.1572052  0.9013406 396
[3]  {Desktop_iMac} => {ComputerMouse_3Button}                    0.03335706 0.1302104  1.4632673 328
[4]  {Desktop_iMac} => {Keyboard_AppleMagic}                      0.03234008 0.1262406  1.7607426 318
[5]  {Desktop_iMac} => {Software_MicrosoftOffice_HomeandStudent2016} 0.03101800 0.1210798  1.8204551 305
[6]  {Desktop_iMac} => {Monitor_Samsung}                          0.03091630 0.1206828  1.2610777 304
[7]  {Desktop_iMac} => {Monitor_ASUS2}                            0.02806875 0.1095673  1.9001325 276
[8]  {Desktop_iMac} => {Monitor_ASUS}                             0.02766195 0.1079794  1.9481854 272
[9]  {Desktop_iMac} => {Keyboard_BacklitLED_Gaming}               0.02756026 0.1075824  1.3475892 271
[10] {Desktop_iMac} => {Monitor_HP}                               0.02684837 0.1048035  1.4494132 264
```

Figure 8 – Rule # 3 – iMac association rule

Figure 8 shows that Monitors from different brands, active headphones, computer mouse and MS software are the type of products that are bought together with an iMac Desktop.

Rule # 4 – HP Laptop Association Rule: This rule was created to find out which items are frequently bought with HP Laptops. Desktops and Laptops categories were removed from the rhs side of the rule and lhs = HP Laptop, using support: 0.001 and confidence: 0.1.

```
> inspect(sort(hplaptop.association.rules, by='confidence', decreasing = T)[1:10])
     lhs             rhs                                                  support    confidence lift      count
[1]  {Laptop_HP} => {Monitor_ViewSonic}                                   0.04800163 0.2472499  2.2407447 472
[2]  {Laptop_HP} => {ActiveHeadPhones_Earpods_Apple}                      0.03274687 0.1686747  0.9671011 322
[3]  {Laptop_HP} => {Keyboard_AppleMagic}                                 0.02888233 0.1487690  2.0749581 284
[4]  {Laptop_HP} => {Monitor_Samsung}                                     0.02756026 0.1419591  1.4834051 271
[5]  {Laptop_HP} => {Software_MicrosoftOffice_HomeandStudent2016}         0.02339062 0.1204819  1.8114660 230
[6]  {Laptop_HP} => {MouseAndKeyboardCombo_Microsoft_Wireless_Desktop}    0.02318723 0.1194343  1.2119681 228
[7]  {Laptop_HP} => {ComputerMouse_3Button}                               0.02267873 0.1168151  1.3127346 223
[8]  {Laptop_HP} => {Monitor_LG}                                          0.02166175 0.1115767  1.9349808 213
[9]  {Laptop_HP} => {Monitor_HP}                                          0.02105156 0.1084337  1.4996187 207
[10] {Laptop_HP} => {Monitor_ASUS}                                        0.02023797 0.1042431  1.8807743 199
```

Figure 9 – Rule # 4 – HP Laptop association rule

Figure 9 shows that Monitors from different brands, active headphones, computer mouse and MS software are the type of products that are bought together with a HP Laptop.

Rule # 5 – Cyberpower Gamer Desktop Association Rule: This rule was created to find out which items are frequently bought with Cyberpower Desktops. Laptops and Desktops categories were removed from the rhs side of the rule and lhs = Desktop_CYBERPOWERGamer, using support: 0.001 and confidence: 0.1.

```
> inspect(sort(cyberpower.association.rules, by='confidence', decreasing = T)[1:10])
     lhs                          rhs                                                  support    confidence lift     count
[1]  {Desktop_CYBERPOWERGamer} => {ActiveHeadPhones_Earpods_Apple}                    0.03834028 0.20840243 1.194881 377
[2]  {Desktop_CYBERPOWERGamer} => {Monitor_Samsung}                                   0.03101800 0.16860144 1.761804 305
[3]  {Desktop_CYBERPOWERGamer} => {Monitor_ViewSonic}                                 0.02430591 0.13211719 1.197335 239
[4]  {Desktop_CYBERPOWERGamer} => {ComputerMouse_3Button}                             0.02094986 0.11387507 1.279695 206
[5]  {Desktop_CYBERPOWERGamer} => {Keyboard_BacklitLED_Gaming}                        0.01972948 0.10724157 1.343320 194
[6]  {Desktop_CYBERPOWERGamer} => {MouseAndKeyboardCombo_Microsoft_Wireless_Desktop}  0.01952609 0.10613599 1.077023 192
[7]  {Desktop_CYBERPOWERGamer} => {Monitor_Acer}                                      0.01922099 0.10447761 1.771256 189
[8]  {Desktop_CYBERPOWERGamer} => {Software_MicrosoftOffice_HomeandStudent2016}       0.01637344 0.08899945 1.338122 161
[9]  {Desktop_CYBERPOWERGamer} => {Accessories_MousePad_Belkin}                       0.01474626 0.08015478 1.368337 145
[10] {Desktop_CYBERPOWERGamer} => {Keyboard_AppleMagic}                               0.01464456 0.07960199 1.110250 144
```

Figure 10 – Rule # 5 – Cyberpower Gamer desktop association rule

Figure 10 shows that active headphones, monitors, computer mouse and keyboards are the type of products that are bought together with Cyberpower Gamer desktops.

To finalize this result section, an analysis of the transaction table is made, based on the outcomes of the transaction summary listed in figure 11.

```
> summary(transactions_table)
transactions as itemMatrix in sparse format with
 9835 rows (elements/itemsets/transactions) and
 125 columns (items) and a density of 0.03506172

most frequent items:
                  iMac               HP Laptop CYBERPOWER Gamer Desktop            Apple Earpods        Apple MacBook Air                  (Other)
                  2519                                  1909                                  1715                                  1530                    33622

element (itemset/transaction) length distribution:
sizes
   0    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19   20   21   22   23   25   26   27   29   30
   2 2163 1647 1294 1021  856  646  540  439  353  247  171  119   77   72   56   41   26   20   10   10   10    5    3    1    1    3    1    1

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.000   2.000   3.000   4.383   6.000  30.000

includes extended item information - examples:
                        labels
1 1TB Portable External Hard Drive
2 2TB Portable External Hard Drive
3              3-Button Mouse
```

Figure 11 – Transaction table summary

There are 9835 transactions and 125 items. Using density, we can estimate how many items were purchased (9835 x 125 x 0.035061) ~ 43,103 items.

In terms of number of items per transaction we found that 62% of the transactions contains between 1 and 4 items. The split of this 62% is: 22% of transactions contains 1 item, 17% of transactions contains 2 items, 13% of transaction contains 3 items and 10% of transaction contains 4 items. An image to identify this trend is showed in figure 12.

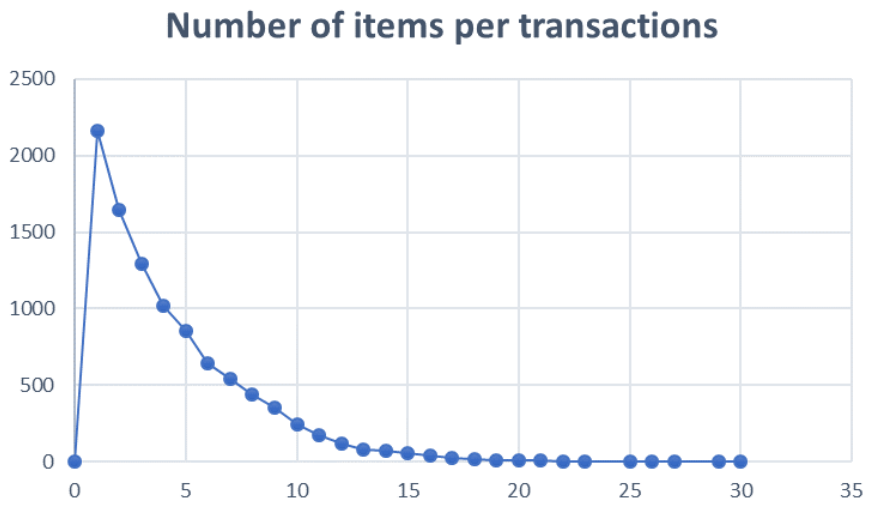
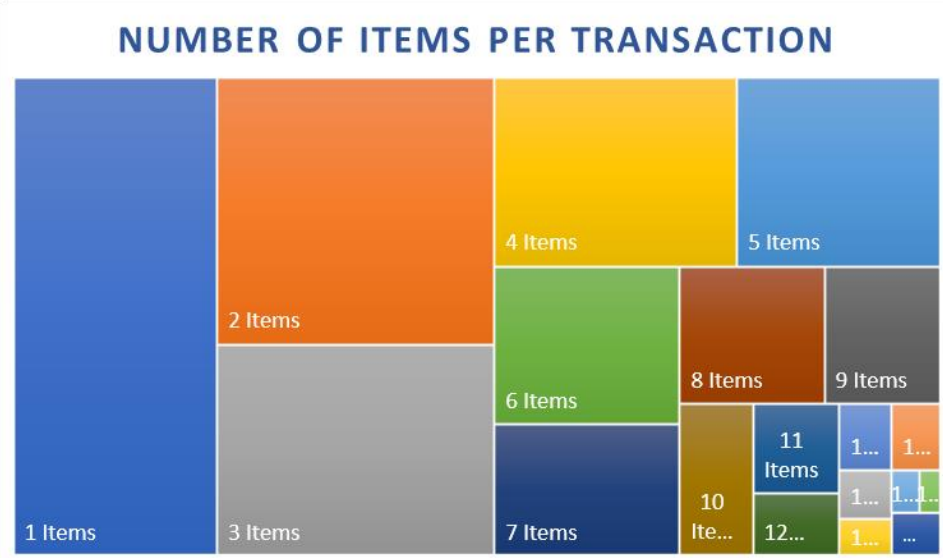


Figure 12 – Number of items per transactions

4. Results Discussion

There are interesting patterns within Electronidex's transactions such as:

- 62% of transactions containing between 1 – 4 items.
- 31% of transactions containing between 5 – 10 items.



This transaction pattern clearly shows that two or more type of customer exists, such as individuals and corporate customers.

Generally, the main products are Desktops (iMac and CyberGamer) and Laptops (HP) and the complementary products are Monitors (Viewsonic, Samsung), Apple Earpods and computer mouse.

However, we can drill down the top products to more specific "complementary products" if we analyze rule # 3, 4 and 5 separately.

5. Recommendations

The recommendations for this analysis are:

Disclaimer: Since this is one-month data, the results of this analysis can be biased. Additional months are required to determine consistent sales patterns.

1- Blackwell can be beneficed of this purchase, since Eletronidex sales are oriented to high end products (Desktop, Laptops), while Blackwell products are oriented to low end products (monitors, warranties, peripheral). As a result, Electronidex's portfolio is complementary to Blackwell's portfolio.
2- Blackwell may acquire Electronidex to completement his product offerings and to enhance the client base.
3- Identify complementary products offered by Blackwell that are duplicated in Electronidex and remove those.
4- Rationalize the product portfolio of Desktops and Laptops
5- Rationalize the vendor base and increase the negotiation power to reduce the cost and increase gross margin per product

6. R scripts

Provided in a notebook.