K.J. SOMAIYA INSTITUTE OF MANAGEMENT STUDIES & RESEARCH

PGDM (RM) – 2016-18 III TRIM

Retail Finance & Analytics

MS Excel is required to answer few/all of the questions

Date : 05/04/2017   Marks : 50   Time : 3 Hours

Please attempt all questions

Section-A – 10 Marks (7+3)

1) Following is the observed sales against allocated space of detergent powder.

Space(inches)	100	200	300	400	500	600
Sales(Units)	228	236	256	278	285	301

   a) Draw a scatter diagram to illustrate the above information - 1
   b) Model the relationship between space and sales of this product assuming linearity - 3
   c) Interpret the gradient/slope of this line -1
   d) Interpret the intercept of this line on Y axis -1
   e) Estimate the sales when the space allocated is 650 inches -1
   f) What could go wrong if we try to use the same model to determine sales when space allocated is 2400 inches? -1

2) Write the cost function of logistic regression and how it behaves when Y=1 and Y = 0

Section – B (10 Marks) (6+4)

1) Following is the part of the transactional data of a retail store for a glass cleaner: The variable names have their usual meaning as per the header row. Date format: m/dd/yy

a) Assuming 5th April 2017 is the date of analysis, and customers inactive for last 30 days are at a risk of churn, determine which customers are at the risk of churn. Who is your most valuable customer? - 5

Quantity	UnitPrice	CustomerID	dateofpurchase
6	2.55	17850	2/6/17
6	3.39	17850	3/1/17
6	4.25	17850	2/27/17
6	4.25	13047	2/16/17
3	4.95	13047	3/29/17
3	4.95	13047	3/27/17
3	4.95	13047	3/25/17
3	5.95	13047	1/19/17
24	3.75	12583	1/5/17
24	3.75	12583	2/5/17
12	3.75	12583	3/7/17
24	1.65	12583	3/22/17
24	2.95	12583	3/12/17
12	3.75	12583	4/1/17
6	4.95	17850	2/3/17
6	1.06	17850	3/18/17
6	1.06	17850	1/21/17
6	1.06	17850	3/3/17
32	10.95	15100	1/19/17
6	2.55	17850	3/17/17
6	3.39	17850	3/12/17
8	2.75	17850	2/25/17
6	4.95	17850	1/1/17
6	3.39	17850	2/13/17
2	7.65	17850	2/20/17
6	4.25	17850	3/26/17
48	3.45	15291	1/12/17
64	2.55	15291	2/10/17
6	1.85	17850	3/22/17
6	1.85	17850	4/1/17

2) Following is the transactional data for customers in a book shop

Customer	Book A	Book B	Book C	Book D
1	1	1	0	0
2	0	1	0	0
3	1	1	0	0
4	1	0	0	1
5	0	0	0	0
6	1	1	0	0
7	1	0	1	0
8	0	0	1	0
9	1	1	0	0
10	1	1	0	0
11	0	1	0	0
12	0	1	0	0
13	1	1	0	0
14	0	0	0	0
15	0	0	0	0

a) Define support, confidence and Lift. Compute the same for the rule **Book A > Book B**

Section – C (10 Marks) (1*5 + 5)

a) Answer the following questions briefly:
1) Mention two uses of Neural Network in retail business.
2) What is TFIDF?
3) Describe the K-Means clustering algorithm step by step
4) What is reinforced learning? Provide an example
5) Provide an example of stemming.

b) Compute **cosine similarity** between the following sentences:

1) Priyanka is the brand ambassador of cow brand ghee
2) A cow is hit by an ambassador with Priyanka inside it.

Section-D (10 marks) (5+2+3)

a) Following the sales versus price for the newly launched cookie by a local company

Sales	Price.Cookies
18	4.74
20	4.81
23	4.36
23	4.29
23	4.17
23	4.66
24	4.73
25	4.11
26	4.21
26	4.25
26	4.62
26	4.53
27	4.44
28	4.19
28	4.37
29	4.29
29	4.57
30	4.21
30	4.77
31	4
31	4.31
33	4.34
34	4.05
35	4.73
38	4.07
39	4.75
41	4
44	4.15
44	4.34
46	4.15

1) Determine price elasticity of demand for this cookie
2) Explain the sign of price elasticity (why positive or negative)?

3) Draw a constant price elasticity price response function and explain it.

Section-E (10 marks) (3+2+2+2+1)

1) One candidate issues the following command to perform linear regression in R

   ```
   Result <- lm(formula) Sales ~ Price.Cookies, data = sales.data)
   ```

   Where Sales in the independent variable and Price.Cookies in the dependent variable in the dataset sales.data. It produces an error. Identify the error and correct it.

2) Which function would you issue in R to locate the current working directory? Which is the function for setting up the working directory?

3) Write down the command you would use to import the dataset **sales.csv** from the location **"C:\Users\yourname\Desktop"** in R

4) Which functions are used in R to view the beginning and end lines of a dataset?

5) Which function is used to display the column names of a dataset?