Reading

1. Refresh your understanding on probability, conditional probability, Testing Hypothesis, Different types of continuous distributions Chi-squared, Student t, Normal etc)
2. Refresh on Linear Regression concepts, multi linear regression

Python

1. Given sales dataset and perform the following operations on it,

a. print the name of columns

b. print 10 rows and 10 columns in the data set

c. plot any graph to represent the total profit.

d. print the item\_type with total cost grater than 1000000

2. Write a python program to create a 2D array and perform the following operations on it,

a. print the dimension of the array

b. update the array by changing the dimension to 3D

c. add 5 to every element, subtract 2 from each element, multiply each element by 5

3. Write a python program to represent marks of 4 students (roll number 1-4) in 4 subjects,

a. find the highest mark of each student

b. find the topper

c. print the marks of first two roll numbers

d. print the marks of all students in ascending order.

e. update the details by adding marks of roll number 5 and 6

4. Create a dictionary Exam\_result with name, score, no\_of\_attempts and qualify, perform the following,

a. create another dictionary label with inputs a, b, c, d, e…..etc and display a DataFrame from exam\_result dictionary data which has the index labels.

b. print the first 4 rows of the data frame

c. print name and qualify from this data frame

d. print number of score in between 20- 35 and also print the sum of attempts by each students

5. Consider the following sales data,

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Account number | Name | Sku | Category | Quantity | Unit price | Ext price |
| 0 | 852659 | Carroll PLC | QC-82856 | Belt | 13 | 44.48 | 578.24 |
| 1 | 654987 | Heidenreich-Bosco | MJ-58694 | Shoes | 19 | 53.62 | 1018.78 |
| 2 | 258369 | Kerluke | AS-58546 | Shirt | 12 | 24.16 | 289.92 |
| 3 | 741852 | Waters-Walker | AS-46987 | Shirt | 5 | 82.68 | 413.40 |
| 4 | 693471 | Waelchi-Fahey | AS-36987 | Shirt | 18 | 99.64 | 1793.52 |

1. Summarise the sales table for quick analysis.
2. Print summary statistics for unit price
3. Print the data type used for each fields
4. Create and print a data frame called customers, which only contains the name and ext price