21BLC1342
Devanshu Agarwal

__/__/____

DA-I

Q-1 State and explain different types of Data Analytics.

Ans - There are four types of Analytics
1) Descriptive
2) Diagnostic
3) Predictive
4) Prescriptive

-> Descriptive Analytics -
It provides a defiction or "sumary view" of figures in an understandable format. This towards preparing data for further analysis (i.e. data preparation)
Eg - Data visualisation through line plots, graphs, pie-charts.

-> Diagnostic Analytics -
Digs deeper into an issue so that they can arrive at the source of a problem. tools both descriptive and diagnostic analytics go parallel.
Eg - Data discovery, correlation

-> Predictive Analytics -
Forecasts trends based on the current predicts the probability of an event happening in future on estimating the accurate time will happen. Uses model constructed past data to predict the future on as impact of one variable.
Eg - linear regression

-> Prescriptive Analytics
Indicates the best course of action. In other words, it tells what decision to make to optimize the outcome. Goals quality improvement, service enhancement, cost reduction, productivity increment, etc.
eg - Optimization algorithms (Python), decision analysis

Q-2 Write 'R' syntax for various mathematical operations?
Ans - # Arithmetic operations
Addition, subtraction, multiplication and division

-> 4 * 2
-> 4 + 2
-> 4 - 2
-> 4 %% 2
-> 4 / 2

Power
5 ^ 4

Order of calculation
1 + 2 * 4
(1+2) * 4
4 / 2 * 3

Functions for calculations
225 ^ 0.5 -> Gives the square root
sqrt(225) -> Gives the square root
abs(-15) -> Gives the absolute value
round(3.14102, 2) Gives 3.14 rounded value

__/__/____

Q-3 Give 'R' example for various data cleaning and manipulation operations?

Ans - For data cleaning suppose we have data for 'air quality'.

```
data = airquality
mean(airquality$ozone)
mean(airquality$wind)
mean(airquality$Solar.R, narm = TRUE)
mean(airquality$ozone, narm = TRUE)
Summary(airquality)
new_df = airquality
new_df$Ozone = ifelse(is.na(Newdf$ozone),
                      median(New_df$Ozone,
                             na.rm = TRUE),
                      New_df$Ozone)

New_df$Solar.R = if else(is.na(New_df$Solar.R),
                         median(New_df$Solar.R,
                                na.rm = TRUE),
                         New_df$Solar.R)

Summary(New_df)
head(New_df)
```

For data manipulation
Here we will first create a data frame and then rename it.

```
# Creating a Data Frame
df = data.frame(row1=0:2, row 2 = 3:5, row 3 = 6:8)
print("original data frame")
print(df)
print("Modified data frame")

# Renaming data frame
rename(df, c("row 1" = "one", "row 2" = "two",
             "row 3" = "three"))

# Adding empty column
dt[["col 4"]] <- 0

# Assigns a value NA to the data frame column 5
df $"Col 5" = NA

# Updating value of column added

df[["col 5"]] = df((("col 1"])) * dt [["col 2"]]
print("Modified Data frame")
print(df)
```

__/__/____

Q-4 Write the 'R' code for data imputation.

Ans- For imputation:

```
# create a dataframe
data = data.frame(marks1 = c(NA, 27, NA, 49, 75),
                  marks2 = c(81, 14, NA, 61, 12),
                  marks3 = c(78.5, 19.325, NA, 28, 48.0))

# impute manually
data$marks1[is.na(data$marks1)] = mean(data$marks1,
                                       na.rm = T)
```

Q-5 Explain a use study on healthy care for data analytics life cycles.

Ans- Let us take an example of life cycle of data analytics for health care, with the objective of providing required medicines for patients of 'cancer'

Stage 1: Objective:
To provide medicines to the patients of cancer.

Stage 2: Understanding the data:
We need the data of the hospital database. We need to analyse the number of patients of cancer.

Stage 3: Data cleaning & data transformation:
We will remove all the unnecessary data such as Doctor details in this case.

__/__/____

Stage 4: Data Enhancement:
The number of patients of cancer will indicate how much medicine will be required that means adding value to data.

Stage 5: Data Analytics:
Seeing the number of patients and demands of medice, we can predict the amount of medicine required.

Stage 6: Data Visualization:
Now at the end we can visualize the data and predictions using graphs in 'matlab'.

Q-6 Illustrate a case study for business data analytics?

Ans- Descreptive Analytics:
It has been observed that ice cream parlours are opened in and near VIT during summers.

Diagnostic Analytics:
It is because of the obvious reasons that during summer demand of ice-cream increases. And VIT has a large no. of students.

Predictive Analytics:
Next year also, around 2-3 new ice-cream parlours will open

Prescriptive Analytics
There are many local ice-cream parlours but there is no 'Baskin Robbins' parlour. So, if someone opens it, the chances of success is high.