

**Perform the data classification using classification algorithm using R/Python.**

```
rainfall <- c(234,543,56,876,89,987,566,987,242,554,78.2,567)
rainfall.timeseries <- ts(rainfall,start = c(2012,1),frequency = 12)
print(rainfall.timeseries)
png(file="rainfall.png")
plot(rainfall.timeseries)
dev.off()
```

**Perform the data clustering using clustering algorithm using R/Python.**

```
newiris <- iris
newiris$Species <- NULL
(kc <- kmeans(newiris,3))
table (iris$Species,kc$cluster)
#plot(newiris[c("Sepal.Length","Sepal.Width")],col=kc$cluster)
```

**Perform the Linear regression on the given data warehouse data using R/Python.**

```
x <- c(151, 174, 138, 186, 128, 136, 179, 163, 152, 131)
y <- c(63, 81, 56, 91, 47, 57, 76, 72, 62, 48)
relation <- lm(y~x)
print(summary(relation))
x <- c(151, 174, 138, 186, 128, 136, 179, 163, 152, 131)
y <- c(63, 81, 56, 91, 47, 57, 76, 72, 62, 48)
relation <- lm(y~x)
a <- data.frame(x=170)
result <- predict(relation,a)
print(result)
x <- c(151, 174, 138, 186, 128, 136, 179, 163, 152, 131)
y <- c(63, 81, 56, 91, 47, 57, 76, 72, 62, 48)
relation <- lm(y~x)
png(file="linearregression.png")
```

```
plot(x, y, col="blue", main="Height & Weight Regression",
abline(lm(y~x)),cex=1.3,pch=16,xlab="Weight in kg",ylab="Height in cm" )
dev.off()
```

```
plot(x, y, col="blue", main="Height & Weight Regression",
abline(lm(y~x)),cex=1.3,pch=16,xlab="Weight in kg",ylab="Height in cm" )
```

**Perform the logistic regression on the given data warehouse data using R/Python.**

```
input <- mtcars[, c("am","cyl","hp","wt")]
print(head(input))
input <- mtcars[, c("am","cyl","hp","wt")]
am.data= glm(formula = am ~ cyl + hp + wt,data =input, family =binomial)
print(summary(am.data))
```

Write a Python program to read data from a CSV file, perform simple data analysis, and generate basic insights. (Use Pandas is a Python library). B1

```
import pandas as pd

#Loading data into a DataFrame
data_frame=pd.read_csv(r'C:\Users\Vipin Kumar Gupta\Downloads\Financial Sample.csv')

#displaying first five rows
display(data_frame.head())

#displaying last five rows
display(data_frame.tail())

print(list(data_frame.columns))

data_frame.info()

data_frame.describe()
```

Perform data visualization using Python on any sales data.

```
import pandas as pd

from matplotlib import pyplot as

plt.rcParams["figure.figsize"] = [8.00, 4.50] plt.rcParams["figure.autolayout"] = True
columns = ["Item_name", "Qty_sold"]
```

```
df = pd.read_csv(r'C:\Users\LENOVO\OneDrive\Desktop\Sales.csv', usecols=columns)
print("Contents in csv file:", df)

plt.plot(df.Item_name, df.Qty_sold)

plt.show()
```

**<https://services.odata.org/V3/Northwind/Northwind.svc/>**