

## EXPERIMENT-3

### PERFORM CORRELATION ANALYSIS AND NORMALIZATION USING R-TOOL

Name: S.G.DEVSACHIN

Reg.No: 192111088

Subject: CSA1672, Data warehouse and data mining

#### CORRELATION ANALYSIS

OUTPUT:

```
R Console
5  57    0
6  80    2
7  51    1
8  53    3
9  49    1
10 39    2
11 57    3
12 58    1
13 69    0
14 33    2
15 55    1
16 45    4
17 35    2
18 25    0
19 69    1
> chisq.test(diabetes1)

Pearson's Chi-squared test

data:  diabetes1
X-squared = 19.943, df = 18, p-value = 0.3361

Warning message:
In chisq.test(diabetes1) : Chi-squared approximation may be incorrect
> |
```

#### NORMALIZATION:

##### 1) MIN-MAX NORMALIZATION

OUTPUT:

```
> A<-c(diabetes1$age)
> mean<-mean(A)
> minimum<-min(diabetes1$age)
> maximum<-max(diabetes1$age)
> minmax<-(A-minimum)/(maximum-minimum)
> minmax
[1] 0.2727273 0.5636364 0.8181818 0.6363636 0.5818182 1.0000000 0.4727273
[8] 0.5090909 0.4363636 0.2545455 0.5818182 0.6000000 0.8000000 0.1454545
[15] 0.5454545 0.3636364 0.1818182 0.0000000 0.8000000
> |
```

## 2) Z SCORE NORMALIZATION

### OUTPUT

```
> A<-c(diabetes$age)
> mean<-mean(A)
> std<-sd(A)
> zscore<-(A-mean)/std
> zscore
Error: object 'zscore' not found
> zscore
[1] -0.89639817  0.23432815  1.22371368  0.51700973  0.30499855  1.93041763 -0.11902382  0.02231697 -0.26036461 -0.96706856  0.30499855  0.37566894  1.15304329 -1.39109093  0.16365776
[16] -0.54304619 -1.24975014 -1.95645410  1.15304329
> |
```

## 3) DECIMAL SCALING NORMALIZATION

### OUTPUT

```
> decimalscaling=(A/100)
> decimalscaling
[1] 0.40 0.56 0.70 0.60 0.57 0.80 0.51 0.53 0.49 0.39 0.57 0.58 0.69 0.33 0.55 0.45 0.35 0.25 0.69
> |
```