

1. Find the set of cars that have the highest similarity with each other.

Answer: The set of cars with highest similarity is 0.8521429 or 85%

2. Find the set of cars that have the highest dissimilarity with each other.

Answer: The set of cars with highest dissimilarity is 0.1000 or 100%

3. Find the two categorical attributes that have the highest positive/negative correlation. Draw a scatter plot of these two attributes. Analyze the correlation using the following significance levels:

0.01, 0.05, 0.001

Answer: High Similarity between Car\_Value and Persons

Correlation (0.01):

Pearson's product-moment correlation

```
data: DF_Corr$Persons and DF_Corr$Class_value
t = -13.04, df = 1726, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
99 percent confidence interval:
 -0.3548260 -0.2420182
sample estimates:
           cor
-0.2994683
```

Correlation(0.05):

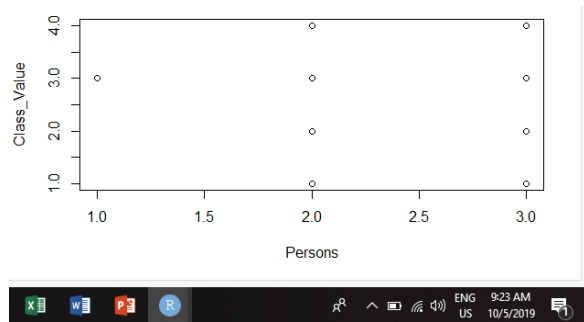
Pearson's product-moment correlation

```
data: DF_Corr$Persons and DF_Corr$Class_value
t = -13.04, df = 1726, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.341797 -0.255927
sample estimates:
           cor
-0.2994683
```

Correlation(0.10):

Pearson's product-moment correlation

```
data: DF_Corr$Persons and DF_Corr$Class_value
t = -13.04, df = 1726, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
90 percent confidence interval:
 -0.3350791 -0.2630031
sample estimates:
           cor
-0.2994683
```



4. Present the dissimilarity matrix for the very good cars.

Answer: The set of cars with highest dissimilarity is 0.7142857 or 71.43%