Data Analytics Lab 6

Data Imputation using kNN method

```
rm(list=ls())
```

```
setwd("your data directory")
```

Read the data and check the content.

```
Data <- read.csv("HS.csv", header=TRUE, sep= ",")
```

```
Data
```

```
str(Data)
```

```
attach(Data)
```

See how the data look like:

```
head(Data)
```

See where you have some missing values

```
summary(Data)
```

Install the package VIM:

```
install.packages("VIM")
```

```
library(VIM)
```

Use other columns of your data for data imputation of the missing values.

```
Impute1 <- kNN(Data, variable= "Genre", k=5)
```

See the result

```
summary(Impute1)
```

You can use kNN data imputation for continuous variables as well.

```
Impute2 <- kNN(Data, variable= c("Audience..score..", "Profitability"), k=6)
```

```
summary(Impute2)
```

You can use kNN data imputation for integer variables as well.

```
Impute3 <- kNN(Data, variable= "Rotten.Tomatoes..", k=6)
```

```
summary(Impute3)
```