# Data Preprocessing Template

# Importing the dataset

dataset = read.csv('Data.csv')

# Splitting the dataset into the Training set and Test set

# install.packages('caTools')

library(caTools)

set.seed(123)

split = sample.split(dataset$DependentVariable, SplitRatio = 0.8)

training\_set = subset(dataset, split == TRUE)

test\_set = subset(dataset, split == FALSE)

# Feature Scaling

# training\_set = scale(training\_set)

# test\_set = scale(test\_set)

# Data Preprocessing

# Importing the dataset

dataset = read.csv('Data.csv')

# Taking care of missing data

dataset$Age = ifelse(is.na(dataset$Age),

ave(dataset$Age, FUN = function(x) mean(x, na.rm = TRUE)),

dataset$Age)

dataset$Salary = ifelse(is.na(dataset$Salary),

ave(dataset$Salary, FUN = function(x) mean(x, na.rm = TRUE)),

dataset$Salary)

# Data Preprocessing

# Importing the dataset

dataset = read.csv('Data.csv')

# Taking care of missing data

dataset$Age = ifelse(is.na(dataset$Age),

avg(dataset$Age, FUN = function(x) mean(x, na.rm = TRUE)),

dataset$Age)

dataset$Salary = ifelse(is.na(dataset$Salary),

ave(dataset$Salary, FUN = function(x) mean(x, na.rm = TRUE)),

dataset$Salary)

# Encoding categorical data

dataset$Country = factor(dataset$Country,

levels = c('France', 'Spain', 'Germany'),

labels = c(1, 2, 3))

dataset$Purchased = factor(dataset$Purchased,

levels = c('No', 'Yes'),

labels = c(0, 1))