

## PROJECT: GERMAN CREDIT DATA

### Credit Risk Assessment

\*\*\*\*\*

### Analysis Exploratory

## Análise Exploratória de Dados

\*\*\*\*\*

### 1. Working Directory

## Configurando o diretório de trabalho

```
setwd("C:/Users/Utilizador/repos/Formacao_cientista_de_dados/big_data_analytics_R_microsoft_azure_machine_learning")
getwd()
```

**Observação:** Este código contém comandos para filtrar e transformar os dados de aluguel de bikes,

Este código foi criado para executar tanto no Azure, quanto no RStudio;

Para executar no Azure, altere o valor da variavel Azure para TRUE;

Ou seja, se o valor for FALSE, o código sera executado no RStudio;

### 2. Data Collection and Transformation

## Variável que controla a execução do script

```
Azure <- FALSE
```

```
2.1 Data Collect  if(Azure){ source("src/ClassTools.R") Credit <- maml.mapInputPort(1) }
```

### 3. Plots

## Plots usando ggplot2

```
library(ggplot2) lapply(colNames2, function(x){ if(is.factor(Credit[,x])) { ggplot(Credit, aes_string(x)) +  
geom_bar() + facet_grid(. ~ CreditStatus) + ggtitle(paste("Total de Credito Bom/Ruim por",x))}})
```

## Plots CreditStatus vs CheckingAcctStat

```
lapply(colNames2, function(x){ if(is.factor(Credit[,x]) & x != "CheckingAcctStat") { ggplot(Credit,  
aes(CheckingAcctStat)) + geom_bar() + facet_grid(paste(x, " ~ CreditStatus"))+ ggtitle(paste("Total de  
Credito Bom/Ruim CheckingAcctStat e",x)) }})
```