

PROJECT: BIKESHARE CAPITAL RENTAL FORECAST

Forecasting Bicycle Rental Demand

Exploratory Analysis: BoxPlots

Este código contém comandos para análise de variáveis usando Box-Plots

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;

Variável que controla a execução do script

```
Azure <- FALSE  
if(Azure){ source("src/Tools.R") Bikes <- maml.mapInputPort(1) Bikes$dteday <- set.asPOSIXct(Bikes)  
}else{ bikes <- bikes }
```

Convertendo a variável dayWeek para fator ordenado e plotando em ordem de tempo

```
bikesdayWeek <- fact.conv(bikesdayWeek)
```

Demanda de bikes x potenciais variáveis preditoras

```
labels <- list("Boxplots - Demanda de Bikes por Hora", "Boxplots - Demanda de Bikes por Estação",  
"Boxplots - Demanda de Bikes por Dia Útil", "Boxplots - Demanda de Bikes por Dia da Semana")  
xAxis <- list("hr", "weathersit", "isWorking", "dayWeek")
```

Função para criar os boxplots

```
plot.bboxes <- function(X, label){ ggplot(bikes, aes_string(x = X, y = "cnt", group = X)) + geom_boxplot(  
  ) + ggtitle(label) + theme(text = element_text(size = 18)) }  
Map(plot.bboxes, xAxis, labels)
```

Gera saída no Azure ML

```
if(Azure) maml.mapOutputPort('bikes')
```