

Ciência de Dados com Software R

OBJETIVO

Capacitar o aluno a utilizar a ferramenta R Studio com a linguagem R e pacotes como tidyverse, para importar, tratar, limpar, organizar e manipular massas consideráveis de dados afim de retirar informações relevantes para organizações e exibir tais informações por meio de gráficos e relatórios feitos com R Markdown. O R tem sido uma opção muito utilizada para ciência de dados pela sua simplicidade em realizar esta tarefa se comparado a outras linguagens de programação, além de ser possível representar os dados de maneira eficiente. O cientista de dados em R estará apto a trabalhar em empresas que demandem a obtenção de informações para fins de tomada de decisão sobre as informações coletadas. Ciência de dados é, certamente, uma das carreiras mais promissoras atualmente.

PRÉ-REQUISITOS

Lógica de programação

METODOLOGIA

As aulas são compostas por módulos, onde será passado para os alunos todo o conhecimento teórico para a realização de exercícios práticos. Cada aluno terá um computador para realizar os exercícios e criar todo ambiente de programação de de alta performance e com tolerância a falhas.

O curso é totalmente presencial e disponibiliza o material virtual ou impresso.

VANTAGENS TREINAR

- Grade curricular baseada na vivencia em programação;
- Aulas práticas;
- Instrutores com anos de experiência em docência.

Carga Horária: 40 horas (5 Dias / 10 Noites)

CONTEÚDO PROGRAMÁTICO

- O que é o software R?
- Instalação
- R e Rstudio
- Instalação de pacotes
- Programação Imperativa
 - Entrada e saída
 - Estrutura condicional
 - Estrutura de repetição
- Tipos de dados
- Funções
- Vetores, Listas e Data Frames
- Operações vetorizadas
- Manipulação de Strings
- Expressões Regulares
- Manipulação de Data e Hora
- Pacote Tidyverse
- Importação de dados
- Datasets presentes no próprio R
- Tibbles
- Valores NA
- Variáveis e observações
- Fatores
- Revisão de estatística básica
 - Média, quartis, variância, covariância
 - Correlação de variáveis
 - Variáveis categóricas e contínuas
- Visualização de dados por meio de gráficos
- A semântica para plotagem de gráficos em R
- Pipes
- Limpeza de dados
 - Spreading e Gathering
 - Separating e Pull
 - Unite

- Manipulação de dados
 - filter()
 - arrange()
 - select()
 - mutate()
 - summarize()
 - outras de funções de manipulação de dados
- Manipulação de Dados Relacionais
 - Chaves únicas
 - Chaves compostas
 - Joins
- Programação funcional
- Funções de mapeamento
- R Markdown