

Introdução

Transcrição

[00:00] Eu vou mostrar para vocês como é que faz uma análise de dados, utilizando o RStudio. A gente vai aprender a desenvolver todas as etapas de análise de um problema prático, utilizando o R, mas antes de entrar exatamente no problema, nós precisamos saber quais são as principais etapas desse projeto, que a gente vai desenvolver.

[00:28] Então, as principais etapas que a gente vai estar abordando nesse curso e utilizando o RStudio, se passa exatamente por esse quadro aqui, que é um fluxo do projeto que a gente vai estar executando. Esse projeto, a gente vai utilizar uma base como experimento.

[00:51] Imagina que você está numa empresa e você é da área de ciência de dados, recebe uma base da área de planejamento estratégico, aonde você tem nessa base de dados diferentes ofertas, imóveis sobre aluguéis, então você precisa explorar essa base e dar, fornecer resultados, fornecer (insight), justamente para a área de planejamento estratégico.

[01:19] A gente vai utilizar, mesmo utilizando o RStudio como ferramenta principal e a gente vai seguir esse processo de análise, aonde a gente vai passar mostrando como é que a gente importa uma base, para dentro do R, as principais bibliotecas que a gente vai estar utilizando, que a principal biblioteca é a tidy, dentro do tidyverse, que seria o universo de dados.

[01:44] Tem uma parte bastante extensiva de transformação, visualização e modelagem. Nesse curso, pelo tempo, como ele é um curso introdutório, também a gente não vai entrar no detalhe da modelagem, mas a gente vai deixar... preparar, a gente vai explicar exatamente.

[02:01] A gente vai executar toda a parte de processo de manusear a base de dados, transformar os dados na parte de visualização também e a parte de comunicação, mostrando sempre em tabelas como é que fica. Então, esse fluxo, ele vai seguir a seguinte... só, entrando mais no detalhe.

[02:20] Ele tem essa primeira parte de importação, como a gente vai justamente mostrar como é que... as principais bibliotecas que a gente utiliza lá dentro do RStudio e que servem, o que? Para importar as bases para dentro do R, essas importações podem ser uma base txt., Excel, tabulado.

[02:36] Então tem o entendimento inicial da base de dados também, que de importado e importar a base, a gente vai ter esse entendimento olhando a estrutura, variáveis, quais são os principais comandos que a gente pode estar utilizando.

[02:50] Depois que a gente tem esse entendimento da base de dados, o próximo passo vai ser justamente a arrumação dos dados que passam pela biblioteca tidy, que é a preparação dos dados para as próximas etapas. Então, é onde você tem um pouco mais de atenção sobre os campos, que tipo de campo você está trabalhando.

[03:11] Se esses dados, eles estão completos, não estão completos, se existem valores faltantes e assim por diante. E aí, nessa parte de transformação, é realmente no projeto de análise de dados, é aonde se despende mais tempo, que aí, você tem a parte de transformação, visualização e modelagem.

[03:33] Aqui a gente está tratando exatamente numa introdução de preparação e transformação dessa base de dados que a gente vai estar trabalhando. Essa parte de transformação, é realmente se tem que ter muita atenção, sobre o entendimento dos dados.

[03:51] E aí, na sequência, a gente tem a parte da... simplesmente da comunicação, onde você vai comunicar em forma de relatório a análise dos dados que você fez exatamente. Então, esse curso, ele mostra muito fortemente como é que se faz... ele recebe uma base de dados na mão e você consegue sair do outro lado com uma visualização, um relatório, com indicadores.

[04:19] Então, mostrar exatamente como é que a gente faz, cria novos campos, como é que a gente relaciona tabelas, diferentes tipos de tabelas dentro do RStudio. Então, essa que é a ideia, exatamente desse curso, de tratamento de dados, de análise de dados, utilizando o RStudio, ok?