

04

Etapas do Projeto de Análise

Transcrição

[00:00] Depois de ter preparado o ambiente de dados, com o RStudio instalado já, a gente vai começar a nossa análise de dados, mas antes de começar a nossa análise de dados, eu vou explicar quais as etapas que a gente vai utilizar nesse projeto de análise.

[00:16] Só lembrando, qual que é o nosso desafio? Nós somos um cientista de dados de uma empresa, a gente recebe da área de planejamento e estratégia uma base de dados de aluguel de imóveis, onde nós precisamos tirar insumos analíticos para a tomada de decisão dessa área de estratégia de dados.

[00:40] Então, antes de a gente começar a análise de fato, eu vou mostrar para vocês quais são as principais etapas que a gente vai estar passando por esse projeto. Então, as etapas do projeto que a gente vai estar desenvolvendo, basicamente são o quê? A gente vai passar para [XXIninteligívelXX], importar os dados para o R.

[01:01] Depois, a gente tem uma biblioteca, que a gente vai ter que instalar dentro do R, carregar, que chama Tidy, que seria uma parte de arrumar os dados. Depois tem uma parte de transformação, visualização e comunicação. Então, eu vou explicar cada uma delas, a primeira seria Import.

[01:18] A (parte) de importação, quando a gente... a gente precisa importar a base lá para dentro do R e tem o entendimento inicial da base de dados junto à estrutura dela, às variáveis, quais os tipos de variáveis que a gente está carregando, quantas linhas... com as observações.

[01:38] Essa parte de Import, a gente vai fazer e vai carregar para uma base que trabalha dentro do R. A gente não vai precisar ficar... O RStudio tem uma funcionalidade de gerar relatório, tabelas, gráficos, então a gente vai fazer praticamente toda a nossa análise dentro do R.

[01:57] Não vai precisar ficar copiando e colando para o Excel, volta... Então, a gente consegue fazer isso tudo pelo R Notebook. A parte de Tidy, a gente vai usar uma biblioteca bastante específica que está dentro de uma biblioteca chamada tidyverse, depois a gente vai fazer essa instalação também.

[02:16] Essa biblioteca empacota várias outras, então a gente não precisa ficar instalando uma a uma, isso facilita bastante a arrumação dos dados, a preparação dos dados para as próximas etapas. A próxima etapa, talvez seja a etapa mais importante, que é a etapa de transformação, visualização e modelagem, aonde a gente precisa, de fato, ter mais atenção para os dados e entendimento dos dados também.

[02:42] Então, a gente vai receber... ao longo do projeto, a gente vai receber alguns insumos da área de planejamento quanto ao tipo de imóvel que a empresa pretende alugar, uma faixa de renda que ela queira alugar, número de quartos, número de... se tem vaga ou se não tem vaga, o bairro que ela quer procurar.

[03:03] Então, a gente vai fazer essas seleções, esses filtros, esses cruzamentos, dentro dessa etapa de entendimento. A parte de modelagem, de fato, a gente não vai gerar, a gente não vai abordar nesse curso, que acaba sendo um curso mais introdutório, mas a gente estaria capacitado para fazer depois... exatamente, a gente já tivesse transformado, entendido, criado novas variáveis.

[03:31] Então, essa parte de transformação é bastante importante. Por último, a gente tem a parte de comunicação, a parte de comunicação do relatório, dos resultados, quando a gente está dentro do R, utilizando o R Notebook, que é uma

ferramenta dentro do RStudio, a gente consegue fazer num único documento, toda essa parte, tanto de documentação, quanto comunicação, já com as tabelas, os filtros.

[03:55] E também, automaticamente todos os códigos que a gente gerar no relatório, eles vão sendo salvos e guardados num único documento. Então, fica muito fácil de depois você publicar isso daí ou algum outro analista checar o que que você fez.

[04:12] Então, o próprio código fica num mesmo documento, o código out put, toda a parte de documentação fica no mesmo documento, que facilita bastante o entendimento dos processos de análise que forem executados ao longo do projeto. Por enquanto é isso, no próximo vídeo, então, a gente começa a fazer essa parte de importação no R. Até mais, valeu.