Exercício de Analise de dados com R

Planejamento do Código em R

#Situação fictícia

ESTRUTURA DO TRABALHO E ENTREGA

1. business problema
2. business context
3. outpult

1.PROBLEMA DE NEGOCIO: A Organização Internacional do Trabalho, agência da ONU destinada a cuidar de questões trabalhistas, contratou a ONG *Trabalho Brasil* para realizar uma visualização de dados e produzir um índice de correlação de Pearson que apontem a associação linear entre anos de estudo/escolaridade e renda do trabalho no Brasil, em 2020, no primeiro trimestre.

**Context**

1.1 Empresas: ONG Trabalho Brasil

1.2 O que a empresa faz: presta serviços para a OIT

1.3 A ONG Trabalhos Brasil contratou o analista de Dados Alan Silva para fazer uma associação linear entre escolaridade e renda no Brasil no primeiro trimestre de 2020.

**Output**

2. Planejamento da solução

2.1 Planejamento do Produto final

1. O que vou entregar?

Resposta: Gráfico animado correlacionando escolaridade e renda

3. Planejamento do Processo

**#Primeira parte: baixar e organizar pacotes, em seguida importar BD e transformar em objeto**

1. Abrir um *New Projet* no R e salvar no diretório de trabalho escolhido, depois baixar o banco de dados para o diretório de trabalho, carregar os pacotes necessários, importar a base de dados para o R, conhecer a base de dados, tratar a base de dados, fazer analise univariadas (estatísticas, gráficas e tabelas), fazer analise bivariadas e multivariadas (correlações);
2. Onde estão os dados a serem trabalhados ?

Resposta: Banco de Dados pnadc\_2020\_t1.fs

1. Quais os pacotes necessários?

Resposta: rio (importar a base de dados), tidyverse (manipulação de bases de dados e gráficos ) weights (realizar coisas com PESOS AMOSTRAIS, algo fundamental quando usamos bases de dados oriundas de amostras por conglomerado, como a PNAD contínua )

1. Como importar a base de dados no R?

Resposta: Use a função import para importar a base de dados dentro do R. Salve em um objeto e não esqueça de verificar se está no envinroment.

**#Primeira parte feita com sucesso.**

**#Segunda parte: tratar a BD**

1. Leia o dicionário de variáveis em anexo, ele vai te dizer que as variáveis de renda do trabalho é a VD4019 . A variável de anos de estudo é a VD3005. E V1027 é a de pesos amostrais. Não esqueça dela!
2. Conheça seu banco de dados com a função str(). Passe o objeto do banco de dados dentro dela e dê run. No console, você vai ver todas as variáveis e seus tipos, de uma vez só. Faça todas as verificações que precisar nas variáveis de interesse.