

O objetivo desta lista é criar um mini-projeto, o qual contém todas as camadas (Camada de dados, camada de lógica, e camada do Usuário). O projeto consiste em fazer a análise dos dados da COVID-19.

Questão 1 (*Criando um projeto*)*(2 pontos)*

Crie um projeto no R-Studio com a seguinte estrutura:

1. Um diretório chamado: “Database”
2. Um diretório chamado: “Graficos”
3. Baixe a base de dados do site <https://ourworldindata.org/coronavirus-source-data>
4. Salve a base no diretório “Database” com o nome “owid-covid-data.xlsx”

(Uma versão antiga da base pode ser obtida no e-class)

Caso você não esteja utilizando o R-Studio apenas organize seu diretório de maneira a seguir a estrutura sugerida.

Questão 2 (*Camada de dados(Data Loader)*)*(2 pontos)*

Crie um script chamado “DataLoader.r”, o script deve realizar as seguintes tarefas:

1. Abrir o banco de dados de nome “owid-covid-data.xlsx” que está no diretório “Database”
2. A coluna data deve ser convertida para o tipo data. (garantindo que a coluna é do tipo “Date”)
3. O script deve filtrar a base e mostrar apenas os seguintes países: “Brazil”, “United States”, “Mexico”, “Germany”, “France”, “United Kingdom”
4. O script deve selecionar apenas as variáveis: **location**, **date**, **total_cases**, **new_cases**

Questão 3 (*Função*)*(2 pontos)*

Crie uma função que recebe um vetor com datas e determina o dia da semana bem como faz uma contagem das semanas.

1. A função deve retornar um data.frame com as seguintes colunas: date, weekday, week. Exemplo:

date	weekday	week
2020-01-01	4	1
2020-01-02	5	1
2020-01-03	6	1
2020-01-04	7	1
2020-01-05	1	2
2020-01-06	2	2
2020-01-07	3	2
2020-01-08	4	2
2020-01-09	5	2
2020-01-10	6	2
2020-01-11	7	2
2020-01-12	1	3

Questão 4 (*Camada lógica*)*(2 pontos)*

Crie um script chamado “main.r”, o script deve fazer as seguintes tarefas:

1. Carregar a função criada anteriormente.
2. Carregar o script DataLoader.r
3. Utilizar a função criada para determinar qual a semana associada a cada registro na base “covid_data”. (dica utilizar *join*)

4. Agrupar os dados por semana para cada pais selecionado. (dica: utilize *group-by* e *summarise*)
5. excluir a informação da ultima semana pois essa pode estar incompleta.

Questão 5 (*Camada do usuário*)*(2 pontos)*

Crie uma rotina que cria 2 gráficos e salva eles no diretório “Graficos”

1. Gráfico 1: utilizando os dados agrupados, crie um gráfico de linha do total de mortes por covid-19, com uma serie para cada pais selecionado. (eixo x: Semana, eixo y: Total de mortes).
2. Gráfico 2: utilizando os dados agrupados, crie um gráfico de linha das novas de mortes por covid-19, com uma serie para cada pais selecionado. (eixo x: Semana, eixo y: Novas mortes)
3. Salve os dois gráficos no diretório “Graficos”. (Configurações recomendadas: `scale=1`, `units = “in”`, `dpi = 300`, `width = 10.4`, `height = 5.85`)