

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE INFORMÁTICA
DEPARTAMENTO DE COMPUTAÇÃO CIENTÍFICA

FUNDAMENTOS MATEMÁTICOS E ESTATÍSTICOS PARA CIÊNCIA DE DADOS

PROJETO FINAL

DESCRIÇÃO DA PROPOSTA: Tivemos no nosso curso uma introdução à linguagem `Python`, abordando também bibliotecas como `sympy`, `numpy` e `matplotlib`. Vimos também como podemos realizar análises de dados utilizando a biblioteca `pandas`, análise gráfica com as bibliotecas `matplotlib` e `seaborn`, construção de gráficos interativos com a biblioteca `plotly`. No decorrer do nosso curso, nós utilizamos o banco de dados oficial da Covid-19 do Brasil para ilustrar o uso de várias ferramentas de análise de dados em `Python`, principalmente as ferramentas gráficas. Com estas análises conseguimos realizar comparações buscando obter respostas para possíveis questionamentos oriundos da análise da doença Covid-19 em âmbito nacional. No entanto, seria interessante utilizar estas mesmas ferramentas com o intuito de entender o que ocorre com respeito à Covid-19 em âmbito internacional. Mais precisamente, nos aprofundarmos acerca de um tema que tem gerado debate em torno da eficiência do isolamento social. De fato, está sendo comum apresentar os dados de mortalidade da Suécia como um exemplo de que não há necessidade do isolamento social. Com o intuito de analisarmos com mais profundidade, neste projeto propomos a comparação de dados de mortalidade da Suécia com dados de seus vizinhos escandinavos. Vamos tomar com base para comparação a Noruega e a Dinamarca que foram países que adotaram o isolamento social. Os dados sobre Covid-19 nestes países podem ser encontrados em:

- <https://www.populationpyramid.net/> - Dados de pirâmides etárias para baixar, somando sexos masculino e feminino;
- <https://tinyurl.com/y4ksyohz> - Dados diários de mortalidade por país;
- <https://tinyurl.com/y6xc9f5y> - Tabela 3.1 clicando em Dødsfald, na coluna “Dødsfald (%)”, contém o total de óbitos por Covid-19 por faixa-etária na Dinamarca;
- <https://tinyurl.com/yyqrulng> - A tabela na seção “COVID-19-associated deaths, by age and sex” contém o total de óbitos por Covid-19 por faixa-etária na Noruega;
- <https://tinyurl.com/y9u5g89j> - Contém o total de óbitos por Covid-19 por faixa-etária na Suécia.

O trabalho de comparar a epidemia em países diferentes é bastante difícil (ver, por exemplo, <https://www.bbc.com/news/52311014>). Assim a única possibilidade de realizarmos uma análise minimamente confiável, é comparando países bastante similares, com pirâmides etárias semelhantes, sistemas de saúde semelhantes (ver <https://tinyurl.com/yxnstykb> e <https://tinyurl.com/y5h573fs>), etc. Mesmo assim, é possível haver diferenças entre as formas em que os diferentes países confirmam mortes por Covid-19.

Por fim, existe também uma especulação de que na Suécia houve uma concentração de óbitos em pessoas bastante idosas que moravam em asilos e afins (ver <https://tinyurl.com/ybcgpxwx> e <https://tinyurl.com/yc7gdcem>). Para verificar esta hipótese, devemos comparar a distribuição de óbitos ao longo das diferentes faixas-etárias nos diferentes países e verificar se realmente houve uma concentração desproporcionalmente maior nas faixas mais idosas na Suécia em comparação com os outros países.

O QUE DEVE SER FEITO: Escreva um texto apresentando o problema, os indicadores que serão utilizados, a metodologia a ser utilizada, as dificuldades e o que foi feito para tentar mitigá-las.

Em seguida, implemente em **Python** códigos para:

- Elaborar gráficos de barras para comparar as pirâmides etárias entre os países em questão;
- Elaborar gráficos de barras para comparar os seguintes indicadores sócio-econômicos entre os países em questão: IDH, densidade populacional, PIB per capita;
- Elaborar gráficos de linhas comparando a curva de mortes nos países em questão colocando como primeiro dia o dia em que¹: (a) houve um total de 10 mortes confirmadas por Covid-19; (b) houve um total de 50 mortes confirmadas por Covid-19;
- Elaborar um gráfico comparando as distribuições dos números de mortes ao longo das diferentes faixas-etárias nos diferentes países.

DESAFIO 1: Acrescentar outros países nórdicos na comparação.

DESAFIO 2: Acrescentar na comparação o impacto do Covid-19 no PIB destes países e verificar se o fato da Suécia não ter realizado o isolamento social reduziu o impacto na economia, ou se o impacto na economia foi similar ao dos demais países nórdicos considerados na análise que realizaram o isolamento social.

COMO DEVE SER APRESENTADO: A apresentação deverá ser feita em **formato de artigo científico**. Com **cabeçalho** contendo título, autor e filiação. Seguido de um **resumo**. Deve conter uma **introdução** com breve abordagem sobre o objeto de estudo. Deve conter também o objetivo e a relevância do estudo proposto. Em seguida, deve conter uma seção de **desenvolvimento**, apresentando a **fundamentação teórica, a metodologia, os resultados e discussão**. É nesta parte que são apresentados os gráficos e as tabelas. Deve conter uma seção de **Conclusões** ou **considerações finais**, que deve responder a questões da pesquisa relacionadas aos objetivos e hipóteses. Além disso, pode conter recomendações e sugestões para trabalhos futuros. Deve conter uma seção com as **referências** utilizadas. E por fim, o **apêndice**, com os códigos implementados no **Python**.

DATA DE ENTREGA: O projeto deve ser entregue até o dia 19/08/2014 às 23:59h.

¹Seguindo o Report 13 do Imperial College, as mortes que ocorrem no início da epidemia são de infecções de fora do país, e portanto podem viesar a análise, desta forma, recomenda-se considerar a análise a partir do primeiro dia em que há pelo menos 10 mortes acumuladas