**Trabalho de INF1026**

Integrantes do grupo: Cladio Alagão (1820293) e Daniel Stulberg Huf (1920468)

Turma 33C – prof. Claudia Ferlin

* Fonte de Dados:

1. WORLD BANK GROUP (org.). ***World Development Indicators: Structure of output***. Ultima atualização em 28 mai. 2020. Disponível em: <http://wdi.worldbank.org/table/4.2>. Acesso em: 8 jun. 2020.
2. UNITED NATIONS DEVELOPMENT PROGRAMME. Human Development Reports (org.). ***Table 1. Human Development Index and its components***. Disponível em: <http://hdr.undp.org/en/composite/HDI>. Acesso em: 8 jun. 2020.
3. WORLD ATLAS. ***Complete List of Country & Dialing Codes***. Disponível em: < <https://www.worldatlas.com/aatlas/ctycodes.htm>>. Acesso em: 8 jun. 2020.

* Descrição dos dados

O arquivo “dadosMundo.xlsx” armazena alguns dados socioeconômicos de diversos países do mundo, e possui 4 planilhas:

* A planilha PIB\_2018 armazena dados referentes ao PIB (Produto Interno Bruto) calculado para alguns países no ano de 2018. As linhas contêm as informações de cada país e as colunas representam, respectivamente, o PIB do país em bilhões de dólares, as porcentagens do PIB em relação à agricultura, indústria, manufatura e serviços e, por último, o código do país;
* A planilha PIB\_2010 tem como linhas, os mesmos países selecionados e como colunas, o PIB dos países calculado em 2010, em bilhões de dólares, e o código de cada país.
* A planilha IDH armazena dados referentes aos cálculos mais recentes do IDH (Índice de Desenvolvimento Humano). As linhas contêm as informações de cada país e as colunas representam, respectivamente, o IDH, a expectativa de vida, média de anos de estudo e PIB per capita de cada país em dólar.
* A planilha Cont tem como linhas os mesmos países selecionados e como coluna os respectivos continentes a que pertencem.
* Utilizando os conceitos de DataFrame da biblioteca Pandas, responda as seguintes perguntas a respeito do arquivo em questão:

1. Crie e exiba:

* DataFrame dfPib2018 a partir da planilha PIB\_2018, que tem como índice a lista de países e como colunas os dados de cada país.
* DataFrame dfPib2010 a partir da planilha PIB\_2010, que tem como índice a lista de países e como colunas os dados de cada país.
* DataFrame dfIdh a partir da planilha IDH, que tem como índice a lista de países e como colunas os dados de cada país.
* DataFrame dfCont a partir da planilha Cont, que tem como índice a lista de países e como coluna os continentes a que pertencem.

1. Renomeie a coluna “PIB” de dfPib2018 para “PIB2018” e a coluna “PIB” de dfPib2010 para “PIB2010”. Exiba os DataFrames alterados.
2. Exiba, através de um gráfico de pizza, a relação percentual entre os 4 setores da economia em relação ao PIB de 2018 do Brasil (note que Brasil está escrito com a letra z).
3. Qual é o número de países listados por continente? Exiba a tabela de frequência.
4. Crie o DataFrame dfPibIdh, que possui como índice a lista de países e como colunas os PIB´s totais de 2018 e de 2010, o código dos países e todas as colunas de dfIdh. Nomeie o seu índice para “País” e exiba-o.

Trabalhando com dfPibIdh:

1. Alguns códigos de países de dfPibIdh estão escritos em letra minúscula e com caracteres aleatórios. Altere esses códigos para as três primeiras letras do país correspondente escritas em maiúsculo. Exiba o DataFrame alterado.
2. Substitua os valores ausentes do PIB de 2010 e de 2018 e pelo valor mínimo de cada um desses anos. Exiba.
3. Exiba o resumo estatístico do PIB calculado em 2010, um país com o maior PIB do ano e um país com o menor PIB do ano.
4. Quantos países tiveram um decrescimento de seu PIB de 2010 para 2018?
5. Adicione a coluna “Continente”, com os continentes a que cada país pertence. Depois, substitua os valores ausentes das colunas de IDH, expectativa de vida, anos de estudo e PIB per capita pelos seus respectivos valores médios do continente de cada país. Exiba o DataFrame atualizado.
6. Exiba em ordem decrescente os 10 países com maiores PIB´s per capita (com seus respectivos valores) e em ordem crescente os 10 países com menores PIB´s per capita (com seus respectivos valores).
7. Exiba um resumo das medidas de sumarização dos anos de estudo dos países por continente, isto é, Máximo, Mínimo, Médio e Mediano. Tais medidas devem estar escritas exatamente dessa maneira.
8. Quais são os países cuja expectativa de vida possui, no máximo, 1 ano de diferença da média?
9. Qual a relação gráfica (scatter) entre o PIB per capita e o IDH dos países de dfPibIdh?
10. Divida os valores de IDH em 5 categorias: de 0.350 a 0.555 inclusive (baixo), de 0.555 a 0.700 inclusive (medio), de 0.700 a 0.800 inclusive (alto) e de 0.800 a 1.00 inclusive (muito alto). Mude os valores da coluna de IDH pelas suas respectivas categorias e exiba o DataFrame alterado.
11. Exiba a tabela de frequência das categorias de IDH e um gráfico de barras representando-a.
12. Exiba o código dos países com o PIB de 2018 acima da média e categoria de IDH que não seja a mais alta.
13. Exiba o PIB per capita máximo, mínimo, médio, mediano e a soma por continente e por faixa de IDH.
14. Quais são as quantidades de países por:

* Faixa de IDH x continente?
* Faixa de IDH x expectativa de vida acima (ou igual) e abaixo da média
* Faixa de IDH x Continente/expectativa de vida acima (ou igual) e abaixo da média

1. Quais são os maiores PIB´s de 2018 por faixa de IDH x continente?
2. Qual é o continente mais comum por cada faixa de IDH? Considere que pode haver repetição.