## **Resumo**

* O desafio proposto trata sobre o transtorno depressivo, um problema multifatorial que pode ter várias possíveis combinações de situações como causa.
* Os dados disponíveis são resultados da NHNES (National Health and Nutrition Examination Survey), realizada anualmente nos EUA para avaliar a **saúde e nutrição** de adultos e crianças;
* Seu desafio é responder as seguintes perguntas:  
  + Qual o perfil de indivíduos (adultos maiores de 18 anos) com sintomas depressivos nos EUA no período de 2005-2006?
  + Hábitos saudáveis de alimentação e atividade física estão associados a menores índices de depressão nesta população?

## **Bases de dados**

Abaixo você tem acesso aos bancos de dados apresentados pela expert:

<https://drive.google.com/file/d/1RtMi9PdKOnVASGP_2285sPcUDUnS9cEH/view?usp=sharing>

https://drive.google.com/file/d/1pXHL4aUGgQshj6DegIQQcmXLYxkQkPIn/view?usp=sharing

**Dicionários das bases**

**DEMO\_PHQ.csv**: banco de dados contendo 5334 observações de adultos pesquisados no NHANES 2005-2006:https://drive.google.com/file/d/1fiqIsCDDrCGFLujKKV5yRnu2pwTG2P-0/view?usp=sharing

**PAG\_HEI.csv**: banco de dados contendo 9424 observações de crianças e adultos pesquisados no NHANES 2005-2006:https://drive.google.com/file/d/1B0tXh0EvfW2nWH2XXVEgWt\_61MWx1UXo/view?usp=sharing

**Sugestão**

* **TRATAMENTO DOS DADOS**
  + Faça a leitura do banco e os tratamentos que achar necessários nos dados;
  + Combine os dois bancos e lide com possíveis dados faltantes. O esperado é um banco final com 5334 informações;
  + Monte o score que resume a informação do questionário PHQ-9. Esse score será uma soma de todas as perguntas envolvendo depressao
* **ANÁLISE EXPLORATÓRIA (EDA) UNIVARIADA**
  + Identificar corretamente quais variáveis são qualitativas e quais são quantitativas;
  + Para medidas quantitativas, podemos usar medidas de posição e dispersão e para gráficos como histogramas ou de densidade;
  + Para medidas qualitativas, podemos olhar para as frequências absolutas e relativas, ou olhar para um gráfico de barras;
  + Avalie dados faltantes, reagrupe ou os re-codifique.
* **ANÁLISE EXPLORATÓRIA (EDA) BIVARIADA**
  + Para avaliar relação entre duas variáveis numéricas use um gráfico de dispersão;
  + Para avaliar relação entre duas variáveis categóricas use um gráfico de barras;
  + Para avaliar relação entre variáveis numéricas e variáveis categóricas use um gráfico boxplot;
  + Avalie a variável dependente versus todas as variáveis independentes graficamente
* **TESTES DE HIPÓTESES**
  + Para avaliar a relação entre duas variáveis numéricas use teste de correlação de Pearson;
  + Para avaliar a associação entre duas variáveis categóricas use Teste qui-quadrado de independência;
  + Para comparar as médias de dois grupos independentes use Teste-t independente;
  + Para comparar as médias de mais de dois grupos independentes use Teste F (ANOVA).
* **ANÁLISE CRÍTICA DOS RESULTADOS**
  + Pense criticamente sobre os resultados que você observou
  + Quais vieses podem existir nos dados e quais as limitações das análises?
  + O que de fato podemos inferir pensando em efeitos causais?

Se acharem que cabe, usem também regressões!