



## LABORATÓRIO 1

### Relação da concentração de glicose no plasma com características diversas

O Instituto Nacional de Diabetes e de Doenças Digestivas e Renais dos EUA conduziram um estudo com 768 mulheres da tribo Pina, que residem próximo a Phoenix. As seguintes características foram coletadas: número de gestações [**pregnat**], concentração de glicose no plasma (obtido duas horas depois da realização de um teste de tolerância a glicose) [**glucose**], pressão sanguínea diastólica (mmHg) [**diastolic**], largura do tríceps (mm) [**triceps**], nível de insulina ( $\mu\text{U}/\text{ml}$ ) [**insulin**], índice de massa corpórea ( $\text{kg}/\text{m}^2$ ) [**bmi**], nível de função diabética [**diabetes**], idade em anos [**age**] e um teste para avaliação de sinais de diabetes (0 = negativo e 1 = positivo) [**teste**]. Na base de dados as características estão rotuladas em inglês conforme as indicações entre colchetes.

Com base nos dados disponíveis no arquivo “Dados\_Lab01.csv” pede-se:

1. Sumarização dos dados do estudo (por exemplo, apresente medidas de tendência central e de variabilidade, construa histogramas).
2. Verificação de observações não usuais. Note que algumas variáveis (características) apresentam valores iguais a zero em situações em que o valor “zero” não poderia existir. No presente estudo, os investigadores representaram os valores faltantes (“*missing*”) pelo código “zero” (0). Descarte as observações não usuais e refaça as análises do item anterior. Alguma mudança foi observada em relação aos resultados anteriores? Quais?
3. Relação da variável resposta (**diabetes**) com as demais variáveis. Construa diagramas de dispersão para avaliação das relações.
4. Comparação do **teste** com as variáveis quantitativas. Construa boxplots para comparar os níveis de glicose e de insulina, a pressão diastólica, o tríceps, o bmi, a idade e o nível de função diabética entre aqueles que apresentaram resultados do teste positivo e negativo.

### Observações:

- Apresente as conclusões sobre os resultados dessa análise preliminar em forma de relatório (arquivo PDF). O arquivo deve ter texto corrido, sem inclusão de *outputs*.
- O *script* utilizado também deverá ser enviado pelo AVA Moodle.