

Pesquisa sobre Partos

Os itens seguintes referem-se aos dados contidos no arquivo de nome *partos.csv* (<http://bessegato.sites.uol.com.br/UFJF/partos.csv>). Esse arquivo contém os dados de uma pesquisa sobre partos realizados em maternidade da Grande BH, vinculada ao SUS, no período de janeiro a dezembro de 2007. O banco de dados é apresentado em nove colunas representando as seguintes variáveis de interesse:

<i>coluna 1:</i>	identificação do paciente
<i>coluna 2:</i>	sexo do neonato: 1 = feminino. 2 = masculino.
<i>coluna 3:</i>	peso do neonato (em g).
<i>coluna 4:</i>	comprimento do neonato (em cm).
<i>coluna 5:</i>	quantidade de nascimentos: 1 = único 2 = gêmeos 3 = trigêmeos.
<i>coluna 6:</i>	idade da mãe: 1 = de 15 a 19 anos 2 = de 20 a 24 anos 3 = de 25 a 29 anos 4 = de 30 a 34 anos 5 = de 35 a 39 anos 6 = de 40 a 44 anos.
<i>coluna 7:</i>	quantidade de gestações anteriores da mãe do neonato
<i>coluna 8:</i>	duração da gestação (em semanas).
<i>coluna 9:</i>	tipo de parto: 1 = normal. 2 = cesária 3 = forceps.

NA indica dado não disponível (*missing data*).

1. Escolha quatro variáveis dentre as colunas 2 a 9. Classifique-as, faça o gráfico da distribuição e a tabela de frequência para cada uma delas. Analise as variáveis quanto ao seu formato, posição, dispersão, pontos discrepantes (atípicos ou *outliers*) e padrões de agrupamentos que julgue relevantes. Comente sobre os resultados encontrados.
2. O senso comum indica que os recém-nascidos de cesarianas têm mais peso que os nascidos de parto normal. Para os dados dessa amostra, o que você diria a respeito? Justifique sua resposta baseando-se em gráficos e tabelas de frequência.
3. Obtenha as medidas de posição e de variabilidade para as variáveis: *peso* e *comprimento*. Comente os resultados obtidos.

4. Repita o item (3) para cada tipo de parto (*tipopart*). Compare as respostas obtidas e comente os resultados obtidos.
5. Analisando a literatura, classifique os nascimentos em dois grupos, prematuros e gestação normal. Denote essa nova variável por *parto* e construa uma tabela de dupla entrada entre *parto* e *tipopart*. Você diria que as duas variáveis estão relacionadas de alguma forma? Justifique resumidamente a razão de sua resposta.
6. Considere os valores da variável *idade* em três grupos: *jovem* com até 24 anos (inclusive), *adulta* para indivíduos com idades entre 25 e 34 anos (inclusive) e *sênior* para maiores de 34 anos. Construa uma tabela de dupla entrada para estudar o comportamento desses grupos em relação ao peso e ao comprimento dos recém-nascidos, tirando as conclusões pertinentes. Comente os resultados obtidos.
7. Escolhendo-se uma parturiente ao acaso, qual a probabilidade de que ela tenha tido gestações prévias, dado ter mais de 34 anos? E ter acima de 34 anos, dado que não tenha tido gestações prévias? Utilize tabelas de dupla entrada para apoiar sua resposta.
8. Considere a variável comprimento de recém-nascidos de mães com pelo menos uma gestação prévia.
 - a. Obtenha o histograma e algumas medidas descritivas. Justifique suas escolhas e comente;
 - b. Você diria que os dados são simétricos? Identifique possíveis *outliers*. Qual a percentagem de observações compreendidas no intervalo entre a média mais ou menos 1 desvio-padrão? E, no intervalo entre a média mais ou menos 2 desvios-padrão? E, no intervalo entre a média mais ou menos 3 desvios-padrão?
9. Deseja-se verificar a influência da duração de gestação no peso do recém nascido. Considere baixo peso ao nascer os valores menores que 2.500 g.
 - a. Suponha que sejam selecionados apenas aqueles nascimentos que você considerou prematuros. Construa um gráfico de dispersão para duração de gestação e peso do recém-nascido. O que pode ser dito?
 - b. Considere agora, o restante dos nascimentos. Construa um gráfico de dispersão para duração de gestação e peso do recém-nascido. Compare com o gráfico obtido no item (a). Comente os resultados obtidos.
 - c. Nos dois casos, ajuste as retas de regressão. Interprete os coeficientes angulares e os interceptos obtidos. Há explicação clínica para sua resposta?
10. Que dados adicionais deveriam ser coletados para aprimorar seu entendimento sobre o banco de dados? Que tipo de estudo você propõe para estudar pelo menos uma hipótese que você levanta com sua análise exploratória deste banco de dados?