

Para este lab, iremos usar dados de deputados federais da câmara, referentes ao ano de 2011. Os dados distribuídos no arquivo dados-deputados.csv, contendo:

- informações de nome, estado e partido de cada deputado;
- informações dos gastos em 2011 para cada deputado;
- informações de presença dos deputados nas sessões de 2011;
- lista dos partidos com sua posição (Governo, Oposição ou Independente)

Os dados originais foram extraídos de <http://www.excelencias.org.br/>.

Questões

Todas as questões desse lab requerem um script em R e um documento PDF com a sua análise referente ao que foi pedido na questão.

1. Implemente um script R que plota dois gráficos:

- Um gráfico png com os intervalos de confiança das médias dos gastos anuais dos deputados de cada uma das cinco regiões do Brasil. Observe o gráfico gerado e compare os gastos dos deputados de cada região. É possível concluir algo interessante com base nos intervalos de confiança? Existem intervalos maiores que outros? A que você atribui isso?
- Um gráfico png com os intervalos de confiança das proporções de presenças dos deputados em sessões durante o ano, sabendo que o número total de sessões que ocorreram no ano foi 202 e considerando o número médio de presenças contabilizadas por região. Observe o gráfico gerado compare as assiduidades dos deputados de cada região. É possível identificar algo interessante com base nos intervalos de confiança? É possível identificar alguma relação entre gastos e assiduidade? Comente.

Para gerar os gráficos acima, considere nível de significância de 5%.

2. Use o mesmo script anterior para gerar os mesmos gráficos da questão 2, porém considerando o nível de significância de 25%. Compare os gráficos gerados na questão 2 e nesta questão. O que aconteceu com os seus intervalos de confiança? ao aumentar o nível de significância estatística? E a sua análise dos intervalos, mudou? É possível identificar alguma relação entre gastos e assiduidade? Comente.

3. Gere um gráfico com os intervalos de confiança da média dos gastos anuais dos deputados por estado. Você consegue identificar com 90% de confiança, um estado cujos deputados gastam, em média, mais que os deputados dos outros estados? Existem intervalos com tamanhos maiores que outros? Analise os

dados para tentar descobrir porque a precisão de alguns intervalos está menor do que a de outros.

4. Gere um gráfico com os intervalos de confiança da proporção de presenças dos deputados em sessões por estado. Você consegue identificar com 90% de confiança, um estado cujos deputados são, em média, mais assíduos que os deputados dos outros estados? Existem intervalos com tamanhos maiores que outros? Analise os dados para tentar descobrir porque a precisão de alguns intervalos está menor do que a de outros.

5. Encontre o estado A do Brasil onde os deputados federais são, em média, mais assíduos. Encontre também o estado B do Brasil onde os deputados federais são, em média, menos assíduos. É possível afirmar com 95% de confiança que o deputado menos assíduo de A falta menos sessões que o deputado mais assíduo de B?

6. Encontre o estado A do Brasil onde os deputados federais gastam mais em média. Encontre também o estado B do Brasil onde os deputados federais gastam menos, em média. É possível afirmar com 95% de confiança que o deputado menos gastador de A gasta mais que o deputado mais gastador de B?

7. Use sua criatividade para analisar os dados dos deputados como você desejar. Relate suas descobertas informando o seu raciocínio para chegar em cada conclusão.