

Exercício sobre Processo de Geração de Dados (PGD)

ABJ

Orientações gerais

- Formem grupos de 5 alunos. Todas as questões deverão ser respondidas nestes grupos;
- Os grupos terão 2 horas para responder às questões;
- Depois das discussões em grupo, retornaremos à sala para discutir as questões conjuntamente;
- Lembrem de anotar as respostas às questões em um documento compartilhado entre vocês.

Perguntas

Classes e assuntos

1) Investigando o processo de geração de dados dos assuntos (e classes) processuais. Vamos usar, para esse exercício, uma [base de dados que está no Drive](#).

Essa planilha contém a contagem de todos a contagem de todos os assuntos extraídos do Esaj do TJAM, entre os anos de 2014 e 2019. Poderíamos fazer uma planilha igual para as classes, pois o raciocínio será o mesmo.

Os assuntos estão padronizados, graças às [Tabelas Processuais Unificadas](#). As TPUs são uma iniciativa do CNJ para padronizar a classificação dos processos judiciais. Existem TPUs para classes, assuntos e movimentações processuais.

a) Veja a imagem abaixo, construída a partir da tabela:

Assunto - Nome3	Assunto - Nome4
Ministério Público	Afastamento
	Aposentadoria
	Processo Disciplinar / Si...
	Promoção
	Remoção
	Remuneração
	-

Nesta imagem, estamos vendo o ramo de assuntos ligados ao Ministério Público. Neste caso, o Assunto - Nome4 é o nível de assunto mais específico que esse ramo consegue chegar. Qual é o assunto com que foram classificados os processos que possuem “-” nesta coluna?

Os casos com “-” na coluna de “Assunto - Nome4” foram classificados como “Ministério Público” apenas.

b) Por que um caso acaba sendo classificado como “-” e não como “Remuneração”, por exemplo?

Isso acontece porque quem faz a classificação dos assuntos é quem protocola a petição inicial. Neste caso, há duas situações:

1. Um caso “-” pode ser um caso de “Remuneração”, mas que foi classificado apenas como “Ministério Público”
2. Um caso “-” pode ter sido classificado dessa forma pois se trata de um assunto ligado ao Ministério Público, mas que não se liga a nenhum dos assuntos específicos que estão no nível 4.

c) Com base no que foi discutido no item anterior, discuta possíveis problemas no gráfico abaixo. Podemos afirmar com certeza que “Remuneração” é o assunto mais frequente a respeito do Ministério Público? O que significam os casos de cinza?

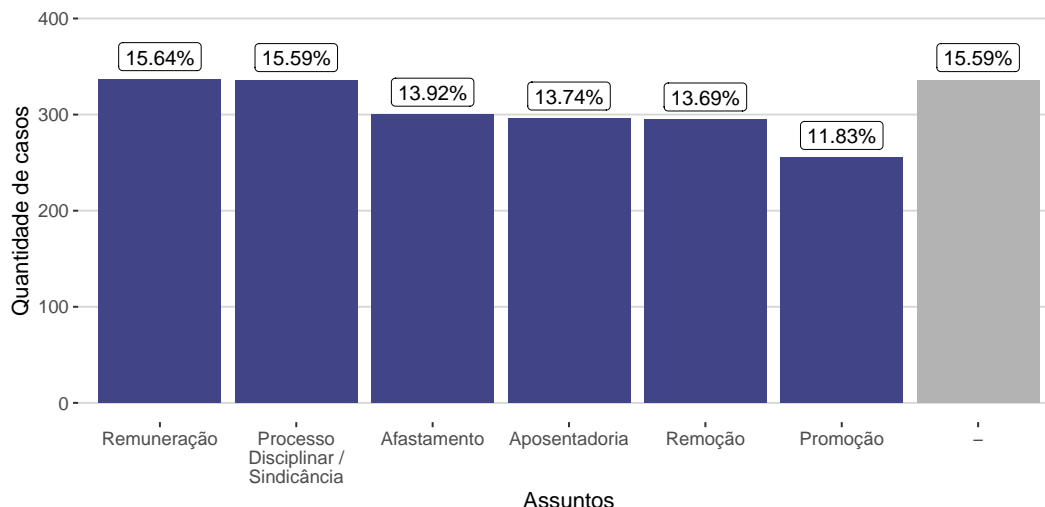


Figura 1: Distribuição de sub-assuntos dentro de DIREITO ADMINISTRATIVO E OUTRAS MATÉRIA DE DIREITO PÚBLICO > Agentes Políticos > Ministério Público

Nós não sabemos como os casos marcados com “-” se distribuem. Para analisar os dados, temos de partir de uma premissa de que os casos marcados com “-” foram marcados assim de forma aleatória. Podemos nos fazer algumas perguntas para nos fazer questionar essa aleatoriedade.

Será que os casos marcados com “-” são aqueles casos que se enquadram em mais de um assunto como a combinação de dois ou mais assuntos (por exemplo, remuneração + promoção) e, justamente por não se enquadrar em um único assunto, eles foram marcados com “-”? Se este for o caso, será que o fato de esses casos representarem a combinação de mais de um assunto não indica que esses casos são mais complexos que os demais, que tratam de um único tema? E por serem casos mais complexos, será que eles não demoram mais tempo para serem resolvidos, ou são casos com menos consenso entre os juízes e, portanto, são casos com mais recursos?

Uma outra série de perguntas possíveis é: será que todos os casos que estão em “-” deveriam ser classificados como “Promoção”, sendo, portanto, “Promoção” o assunto mais frequente?

Será que existe algum motivo para que casos que de fato sejam de “Promoção” sejam classificados de forma mais genérica do que os outros assuntos processuais? Por exemplo, será que SEMPRE que os promotores se deparam com um caso de “Remuneração”, eles registram no sistema do tribunal como o assunto “Remuneração”, mas quando o caso é de “Promoção”, eles registram no sistema do tribunal ou com o assunto “Promoção” ou com o assunto “Ministério Público”?

d) Um pesquisador deseja listar *todos* os processos relacionados à Remuneração dos promotores públicos. Com base na planilha deste exercício e nas discussões dos itens anteriores, quais assuntos você recomendaria a ele utilizar para filtrar a sua busca?

Uma primeira resposta seria dizer que basta procurar pelo assunto “Remuneração”. Entretanto, o enunciado diz que o pesquisador deseja listar *todos* os processos relacionados à remuneração dos promotores. Nesse caso, o que acontece com os processos que foram mal classificados? O pesquisador estaria deliberadamente excluindo esses processos da análise. E teria problema excluir de forma sistemática os processos mal classificados? Novamente, aqui há um problema do que podemos assumir sobre o caso. Será que os casos mal classificados possuem características específicas, ou eles são aleatórios? Se for o caso de aleatoriedade, não tem problema o pesquisador deixar de fora esses casos (apesar de ser da sua vontade pessoal listar *todos* os processos). Mas se não tiver como o pesquisador defender a aleatoriedade da má-classificação, então talvez seja melhor ele tentar pegar os processos mal classificados. Neste caso, ele precisaria procurar pelos assuntos “Remuneração” e “Ministério Público”.

Se ele quiser ser mais abrangente ainda, ele deveria pegar também os assuntos “Agentes Políticos” e “DIREITO ADMINISTRATIVO E OUTRAS MATÉRIAS DE DIREITO PÚBLICO”. Mas se ele pegar esses assuntos tão genéricos, ele vai acabar pegando muito lixo na pesquisa dele, fazendo com que muitos processos que estão fora do escopo dele entrem na análise. Aqui, então, entra uma discussão importante a respeito da parcimônia x precisão. O que está em jogo aqui é saber sopesar a simplicidade da pesquisa com a sua precisão. Escolher assuntos tão genéricos vai fazer com que tenha muito lixo na lista de processos. Se isso não for retirado da base, então a análise não vai mais ser sobre processos de Remuneração do Ministério Público, mas sim, uma análise sobre qualquer tema sobre o Ministério Público, ou sobre qualquer outro Agente Político, ou ainda sobre qualquer assunto de Direito Administrativo. Neste caso, talvez seja melhor sacrificar alguns processos que foram mal classificados, a fim de se preservar o escopo da base de dados.

Valor da causa

2) Investigando o processo de geração de dados do valor da causa

O valor da ação é uma informação que pode ser obtida nas informações de capa dos processos. No Esaj, essa informação aparece disposta da seguinte maneira:

←

Visualizar autos

Peticionar

1021742-63.2020.8.26.0196

Extinto

Tramitação prioritária

Classe

Assunto

Foro

Vara

Juiz

Procedimento Comum Cível

Interpretação / Revisão de Contrato

Foro de Franca

1ª Vara Cível

João Sartori Pires

Distribuição

Controle

Área

Valor da ação

Recolher

14/08/2020 às 11:03 - Livre

2020/001493

Cível

R\$ 3.373,20

Sabendo que essa informação estava disponível nas capas dos processos, uma pesquisadora extraiu essa informação de 100 processos escolhidos aleatoriamente. A partir disso, foi encontrado um valor médio da ação de R\$ 17654.1, conforme a figura abaixo.

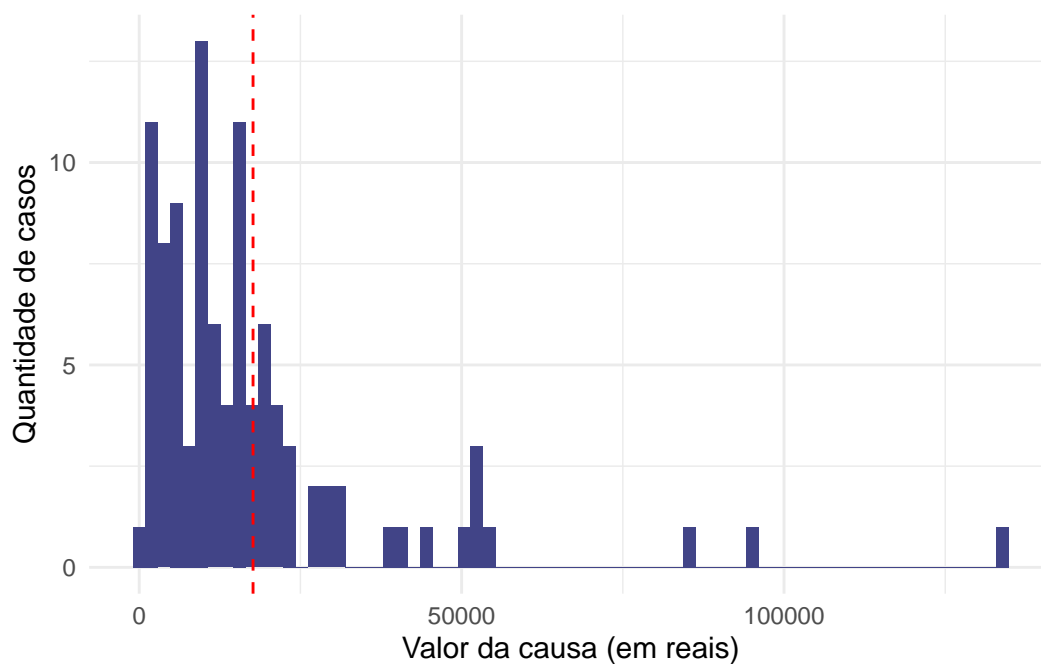


Figura 2: Valor da causa em ações genéricas

A partir disso, responda (para algumas questões será necessário consultar o CPC):

a) O valor da ação corresponde ao valor efetivamente pedido? E se o valor pedido na inicial for impugnado, a informação da capa irá refletir o valor originalmente pedido, o valor contestado, ou novo valor arbitrado pelo juiz?

O valor da ação corresponde ao pedido inicial, e não ao valor contestado, tampouco ao valor arbitrado pelo juiz. O valor arbitrado pelo juiz é o valor da condenação

b) O valor da ação pode ser corrigido? Ele pode estar desatualizado?

O valor da ação pode não estar sempre corrigido. Não existe um sistema dos TJs que garanta a atualização dos valores.

c) Segundo art. 292 do CPC, quais aspectos podem influenciar na definição do valor da causa?

Se houver pedidos cumulativos, pelo art. 292, VI, o valor da ação será a soma de todos os pedidos.

Se há vários pedidos alternativos, pelo art. 292, VII, o valor da causa refletirá o pedido de maior valor.

Se houver pedidos subsidiários, pelo art. 292, VIII, o valor da ação refletirá o valor do pedido principal.

d) Supondo que a coleta dos dados foi realizada de forma adequada, como você avalia a utilização do valor da causa para aproximar o valor pedido? Discuta se podemos falar que o valor médio da ação encontrado reflete o valor pedido. Esse valor pode estar subestimado ou sobrestimado?

O valor da causa é um valor complicado de ser utilizado para representar o valor real do pedido. A depender do caso, ele pode ser um valor mais ou menos consistente. Por exemplo, em ações de alimentos, por causa da lei, sabe-se que os valores da causa refletem a soma de 12 meses da prestação. Neste caso, então, é possível ter mais certeza sobre o que significa o valor da ação.

O valor da ação poderá estar subestimado caso existam pedidos subsidiários, ou no caso de o período para se pagar as prestações for superior a 1 ano.

E o valor da ação poderá estar sobrestimado caso existam pedidos alternativos, se o juiz conceder o pedido de menos valor. Mesmo neste caso, o valor que irá aparecer na capa do processo é o valor do pedido do maior pedido.

Datas

3) Investigando o processo de geração de dados da data de liminares

Ao se estudar a Justiça Federal do Ceará, um pesquisador criou um robô para extrair todas as movimentações processuais de todos os processos de seu interesse. Ele queria calcular o tempo até as decisões de tutela antecipada dos processos. Para tanto, ele precisava de duas datas: (a) a distribuição do processo e (b) a data da decisão liminar.

Para pegar a data de distribuição, ele pegou a data da primeira movimentação registrada no sistema da JFCE, independente de qual fosse a descrição da movimentação.

Para pegar a data da decisão de tutela antecipada, o pesquisador buscou pegar as datas das movimentações que continham em sua descrição as palavras “liminar” ou “antecipação de tutela”, por serem as decisões de tutela antecipada que ele observou na TPU de movimentações.

a) Quais problemas poderiam aparecer ao pegar sempre a data da primeira movimentação como data de distribuição?

Muitas vezes, a primeira data de movimentação é uma data de migração de sistema ou uma data escrita incorretamente (por exemplo, do ano de 1899). Além disso, as datas podem não estar ordenadas na tabela de movimentações.

b) Em alguns processos havia duas ou mais movimentações de antecipação da tutela. Qual data o pesquisador deve considerar? A data da primeira movimentação ou da última? Ajude o pesquisador a decidir qual data escolher, indicando as vantagens e desvantagens de cada alternativa.

Existem alguns problemas que o pesquisador deve pensar:

1. Se a petição inicial trazer mais de um pedido liminar, esses pedidos podem ser julgados em tempos diferentes. A primeira decisão liminar pode julgar apenas uma parte dos pedidos, enquanto a última decisão liminar pode julgar o resto dos pedidos. Neste caso, escolher a primeira data, significaria escolher o tempo para se ter *alguma* resposta judicial para o problema; enquanto escolher a última data significaria escolher o tempo para se ter uma satisfação completa dos pedidos liminares.
2. Se a petição inicial trouxe apenas um pedido liminar e, depois, em outra peça, a parte autora demanda um novo pedido liminar, podemos presumir que a primeira decisão liminar tenha julgado os pedidos da petição inicial, enquanto a última decisão liminar tenha julgado o último pedido. Neste caso, escolher a última data seria um problema muito grande, pois a pesquisa estaria calculando o tempo entre a distribuição do pedido (que normalmente corresponde à petição inicial) e a decisão que julga um pedido que NÃO foi realizado neste momento, sobrestimando demais o tempo de decisão.

Dessa forma, parece ser mais acertado escolher sempre a data da primeira liminar.

Interpretação de gráficos

4) Exercício sobre boxplot

Os exercícios sobre boxplot dizem respeito aos processos administrativos sancionadores do Mercado de Capitais. Basicamente, a autarquia federal responsável por fiscalizar as infrações no mercado financeiro é a Comissão de Valores Mobiliários (CVM), tais como *insider trading*, abuso de poder, falta com os deveres de lealdade, etc. O Colegiado da CVM realiza o julgamento desses acusados para estabelecer as penalidades do mercado financeiro. As penalidades podem ser de três tipos: multas, restrições de exercício de cargo (inabilitação, suspensão e proibição) e advertências. Para todas as aplicações de penalidade, é possível entrar com recursos. Mas esses recursos são destinados, não à CVM, mas ao Conselho de Recursos do Sistema Financeiro Nacional (CRSFN). O CRSFN pode reformar as condenações aplicadas à CVM.

Feita essa introdução ao processo administrativo sancionador no Mercado de Capitais, podemos entrar na comparação dos valores das penalidades de multas e de restrições. As multas são sempre impostas de forma pecuniária e, portanto, podem ser medidas pelo seu valor em reais; já as restrições são sempre impostas com alguma medida temporal (por exemplo, 1 ano de inabilitação para o cargo de diretor) e podem, portanto, ser medidas em anos. Assim, vemos dois boxplots a seguir nas Figuras 3 e 4, comparando os valores de multas e de restrições na CVM e no CRSFN, respectivamente.

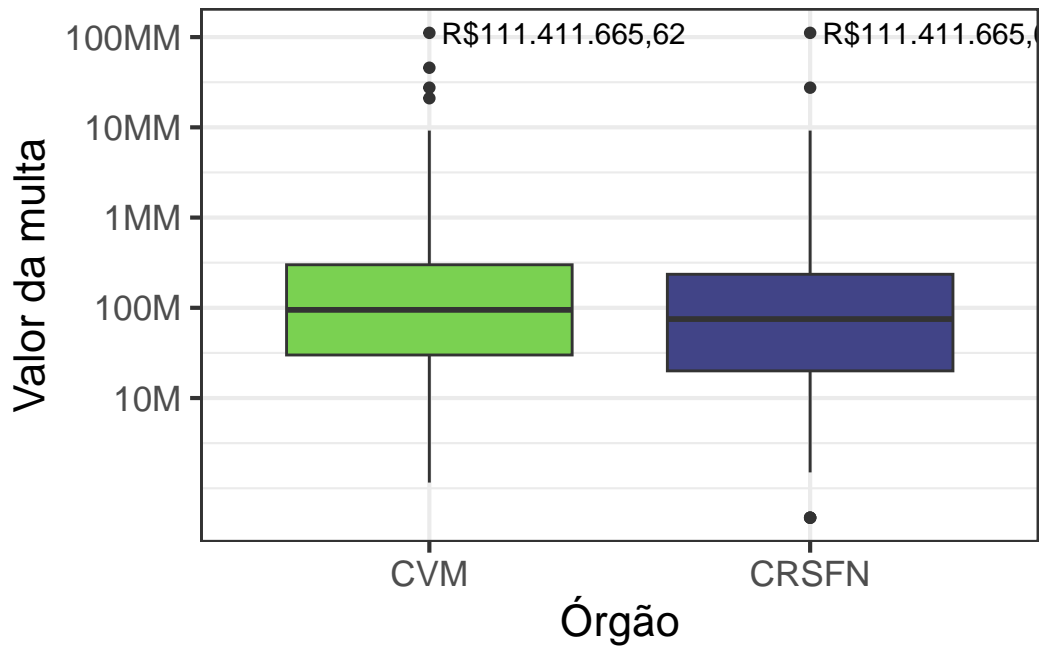


Figura 3: Comparação do valor das multas entre CVM e no CRSFN.

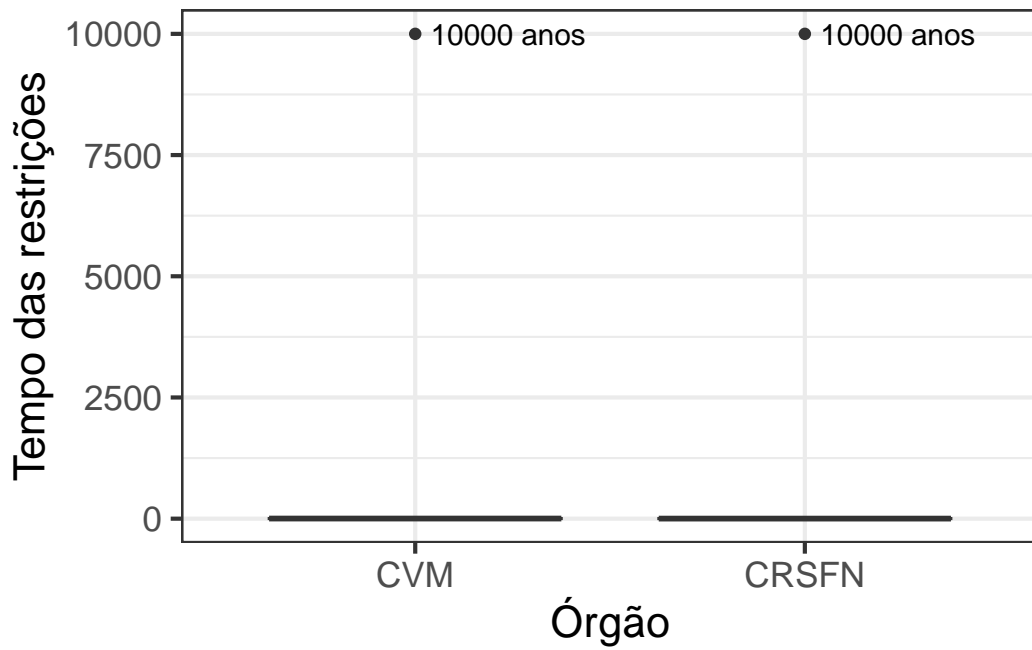


Figura 4: Comparação do tempo das restrições entre CVM e no CRSFN.

Com base nesses gráficos, responda:

a) Olhando a mediana, que 50% dos dados, qual dos dois órgãos aplicam multas mais altas? E penas de restrições?

As multas, tanto na CVM, como no CRSFN, são muito parecidas. Não é possível falar que alguma multa é maior que a outra. Quanto às restrições, podemos afirmar que a CVM geralmente aplica sanções mais severas.

b) Qual dos dois órgãos possui maior variabilidade quanto à aplicação de multas? E quanto às restrições? Quais elementos sustentam essas conclusões?

A CVM possui maior variabilidade quanto às penalidades de multa. Vemos isso pela quantidade de possíveis outliers que existem nela, em comparação com o CRSFN.

Já na comparação das penas de restrição, podemos dizer que é o CRSFN quem possui penas mais dispersas. Podemos observar isso pelo tamanho da caixa central do boxplot. A caixa azul, do CRSFN, é muito maior que a caixa da CVM. O tamanho da caixa representa o IQR e o IQR impacta diretamente também no tamanho dos bigodes. Vemos que os bigodes do CRSFN são maiores também.

c) Considerando tanto a mediana quanto a variabilidade conjuntamente, o que se pode dizer sobre a comparação entre os valores de multas e restrições nos órgãos?

Olhando tanto para a mediana quanto para a variabilidade, a comparação é indecidível tanto para multas quanto para restrições. No caso das multas, a mediana é muito parecida, já no caso das restrições, a variabilidade das restrições no CRSFN é muito alta.