

JURIMETRIA:

Uma aplicação para a base de dados do TJ-MG

Islaine Soneman Fernandes de Oliveira¹

Filipe Jaeger Zabala²

“Nada é impossível, algumas coisas só são menos improváveis do que outras.”

~ JONATHAN WINTERS

“Aquilo que os homens, de fato, querem não é o conhecimento, mas a certeza.”

~ BERTRAND RUSSELL

Resumo

Este artigo aborda a literatura disponível sobre Jurimetria como instrumento a permitir motivação para aplicações práticas e teóricas do tema, bem como considerações na busca da melhor resposta ao litígio, na obtenção de bases de dados públicos e na sua previsibilidade como forma de fortalecer a segurança jurídica.

Palavras-chaves: Jurimetria. Estatística.

1. Introdução

As Ciências Sociais, Exatas, Humanas e Jurídicas estão, a cada ano, mais inter-relacionadas. Direito e Economia, Estatística e Biologia, Direito e Estatística são alguns exemplos cujas pesquisas e trocas de saberes estão sendo cada vez mais conhecidos e estudados. Com a utilização do computador estes aprendizados se ampliaram.

O advogado de Minesota Lee Loevigner [1949] cunha o termo *jurimetrics* onde apresenta a jurisprudência como alvo de interesse e torna o uso do Direito mais previsível. Assim, pela primeira vez une Direito, Estatística e Métodos Computacionais. Conforme Zabala

¹ Bacharel em Economia pela Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre/RS, Brasil. E-mail: islaine.soneman@gmail.com

² Professor da FAMAT-PUCRS e sócio da ZN Consultoria Estatística. E-mail: filipe.zabala@gmail.com

e Silveira (2014, p.76), “à luz das ideias de Loevigner, define-se Jurimetria como a *aplicação de métodos quantitativos no Direito*”.

No Brasil, quando este tema começou a ser difundido, houve muitas dúvidas com relação a sua empregabilidade junto aos principais usuários, como, advogados, legisladores e decisores. Segundo, Amoldi (2010, p. 94):

As concentrações de estudos nessa área refletem uma preocupação social com o aperfeiçoamento do Poder Judiciário. As autoridades judiciárias, desde os tribunais estaduais, passando pela mais alta corte de justiça do Brasil e o Conselho Nacional de Justiça estão cientes da necessidade de estudos quantitativos-empíricos para aperfeiçoar a sua gestão e iniciaram esforços na promoção de pesquisas desse gênero, estabelecendo metas para serem cumpridas desde os magistrados de primeira instância até os Tribunais Superiores.

A análise das informações organizadas em banco de dados públicos faz da Jurimetria um instrumento fundamental para elucidar a situação sócio-econômica vigente. Por conseguinte, a elaboração legislativa e a gestão pública, juntamente com o uso dos avançados meios da tecnologia da informática, fazem-nos pensar em uma jurisdição mais célere e com efetividade na prestação de seus serviços. A Constituição Federativa do Brasil, em seu artigo 5º, inciso LXXVIII já trata da duração razoável do processo como uma de suas garantias.

Desse modo, Zabala e Silveira (2014, p. 77) salientam que “com o suporte de peritos, os legisladores podem se valer da análise dos fatos descritos nos autos, bem como da própria jurisprudência como forma de embasar a criação e manutenção legislativa contemporânea”.

Sob outro aspecto, a Jurimetria busca manter a autoridade e autonomia de quem decide as questões judiciais. Ou melhor, o decisor jurídico pode fazer uso de uma ferramenta isenta como suporte aos seus argumentos. Assim, “tal embasamento pode ou não ser considerado pelo magistrado, ficando a seu critério a utilização e contestação dos métodos utilizados. Desta maneira, a informação disponível agrega-se à opinião do julgador de forma intuitiva, sem jamais substituí-la”. (ZABALA e SILVEIRA, 2014, p.79).

A advocacia pode fazer uso de modelos estatísticos a partir de informações e experiências disponíveis e aplicá-las jurimetricamente a fim de obter estatísticas que forneçam as medidas de evidência nos processos. Ter os resultados com relativa antecedência e eficiência para dar suporte à argumentação pelos usuários do Direito na instrução probatória é a terceira forma de olhar a Jurimetria.

Na mesma linha de raciocínio, Zabala e Silveira (2014, p.) afirmam que os modelos estatísticos permitem fornecer subsídios mais sólidos aos clientes, sendo uma segurança jurídica para o próprio advogado ou escritório. A celeridade com que as decisões judiciais

necessitam ocorrer revela que a segurança jurídica dá espaço para uma jurisdição mais efetiva. Assim sendo, a exposição suficiente dos argumentos que norteiam o julgamento do decisor é que se fundamenta a mister segurança jurídica a despeito da estabilidade e previsibilidade, tendo a Jurimetria como ponte para a previsibilidade das relações sociais e jurídicas.

2. Metodologia

Ao se elaborar um estudo estatístico, a metodologia a ser utilizada pelo pesquisador necessita de alguns atributos, o que a torna desafiante. Uma pesquisa precisa de uma determinada população ou amostra, que evidencie representando de maneira fiel os dados e características daquela, um método para obter os dados e questionamentos que tornem as informações coletadas passíveis de serem analisados.

Nesse sentido, a ciência adequada para estudar de forma empírica um universo de eventos e fatos é a Estatística. A Estatística utiliza métodos para coletar (recolher os dados), analisar, sistematizar (organizar e resumir estes dados) e interpretar as informações obtidas.

Segundo Escotet (2007 apud SERRA, 2013), Estatística é a técnica que computa, numera, mede fatos relacionados aos elementos de uma amostra ou população; coordena e classifica os dados obtidos com objetivo de determinar suas causas, consequências e tendências, e se divide em estatística descritiva e estatística inferencial.

2.1 Análise Exploratória de Dados

A análise descritiva ou análise exploratória de dados é aquela que faz uso de medida-resumo para descrever fatos. Esta técnica vem com o intuito de simplificar as informações e pode utilizar os dados provenientes tanto da população como da amostra. Entende-se por população (ou universo) um grupo de portadores das características de interesse. A amostra é um subgrupo desta população. Nesse sentido, a população pode ser finita ou infinita. Se olharmos o conjunto de dados em um determinado tempo, será finita. Por exemplo: a população formada pela quantidade total de prendedores fabricados em um determinado turno do dia é finita, já ao verificarmos todos os resultados em sucessivos lances de uma moeda (cara ou coroa) é uma população infinita.

Assim, a análise exploratória de dados é uma ferramenta para compreender problemas de inferência estatística, e no caso de possuímos uma população, serve como resumo dos dados. Estas características são chamadas de variáveis.

2.2 Inferência Estatística

A estatística inferencial é a área mais importante da Estatística. Consiste na obtenção de resultados que possam ser projetados para toda a população a partir de uma amostra da mesma. Conforme Spiegel (1993, p. 2):

se uma amostra é representativa de uma população, conclusões importantes sobre a população podem ser inferidas de sua análise. [...] Como esta inferência não pode ser absolutamente certa, a linguagem da probabilidade é muitas vezes usada, no estabelecimento das conclusões.

Desse modo, os problemas de inferência estatística se encontram em três grandes grupos: problemas que estimam pontualmente os parâmetros da população, intervalos de confiança e os testes de hipótese. Estes verificam através de um método se os dados amostrais trazem evidências que fundamentam ou não uma hipótese formulada sobre o universo. Os testes de hipóteses podem ser paramétricos ou não-paramétricos.

2.1 Análise Textual

A análise textual é um tipo específico de análise de dados. Com esta técnica é possível analisar textos, documentos, relatórios, entre outros. Também permite descrever um material produzido por um produtor, como também pode comparar produções diferentes em função de variáveis específicas que permitem descrever quem escreveu o texto.

Existem diferentes formas de análises textuais indo desde o cálculo de frequência de palavras até classificação hierárquica descendente. Assim, o pesquisador pode identificar a quantidade de palavras, sua frequência média, a coocorrência entre as palavras, entre outros aspectos.

Segundo Zabala (2014, p.5):

Pode-se considerar as informações processuais dentro deste paradigma, dada a grande quantidade de texto gerada diariamente no judiciário. Pode-se avaliar o estilo de cada julgador, sendo uma possibilidade de tornar mais tangível o *merecimento*, carente de uma definição formal.

2.1.1 Wordcloud

A Nuvem de Palavras ou *Wordcloud* é uma ferramenta que agrupa palavras e as organiza graficamente em função da frequência dos termos e sua proximidade. É uma análise textual mais simples, porém interessante graficamente.

2.2 Modelagem ARIMA

O modelo Auto Regressive Integrated Moving Average – ARIMA(p,d,q)–fornece uma descrição parcimoniosa de um processo estocástico estacionário em termos de dois polinômios, descrito primeiramente por Whittle (1951) e popularizado posteriormente por Box & Jenkins (1971). O modelo geral segue a forma

$$y_t = c + \sum_{i=1}^p \phi y_{t-i} + \sum_{i=0}^q \theta \varepsilon_{t-i} ,$$
$$\varepsilon \sim N(0, \sigma_\varepsilon^2).$$

Através deste modelo pode-se modelar a média de uma série temporal, permitindo realizar previsões e o estudo do comportamento dos dados. Para realizar o ajuste, utiliza-se a metodologia de Box & Jenkins, que faz uma avaliação dos gráficos de autocorrelação e autocorrelação parcial, que fornecem indícios sobre a ordem do modelo (parâmetros p,d e q), conforme Anexo II.

3. Resultados

As bases de dados foram tratadas com a planilha eletrônica *Excel for Mac 2011* (versão 14.4.7). A análise dos dados foi feita com o pacote estatístico *R* versão 3.1.1 (2014-07-10) (R Core Team [2014]); assim, todos os códigos apresentados neste texto estão em linguagem R.

3.1 Base de dados

A base de dados foi fornecida pelo TJ-MG (TJMG-SEPLAG-CEINFO) em 05/12/2014, reunindo processos até 26/06/2014. Possui ao todo 5.080.270 linhas (processos³) e 11 colunas (variáveis). As colunas estão na Tabela 1.

Tabela 1

Descrição	Nome da variável	Tipo
Ano da Distribuição	<code>ano.distribuicao</code>	numeric
Data Distribuição	<code>data.distribuicao</code>	date
Comarca	<code>comarca</code>	string
Vara	<code>vara</code>	string
Número do Processo	<code>numero.do.processo</code>	string
Julgado (s/n)	<code>julgado</code>	binary
Data Último Julgamento	<code>data.ultimo.julgamento</code>	date
Última Movimentação	<code>ultima.movimentacao</code>	string
Data Última Movimentação	<code>data.ultima.movimentacao</code>	date
Classe	<code>classe</code>	string
Assunto	<code>assunto</code>	string

3.1.1 Ano de distribuição

O processo "008006006004-5" possui `ano.distribuicao` igual a 2206, e há 153 processos com `ano.distribuicao` inferior a 1950, listados no Anexo I.

A Figura 1 apresenta um resumo do volume anual dos processos, excluindo-se aqueles $153 + 1 = 154$ processos com datas fora do escopo de interesse. Nota-se um

³Para garantir que o número de linhas equivale ao de processos, contou-se o número de processos distintos, i.e., `length(numero.do.processo) = length(unique(numero.do.processo))`.

crescimento exponencial no volume anual, considerando que para 2014 os dados vão até 26 de junho.

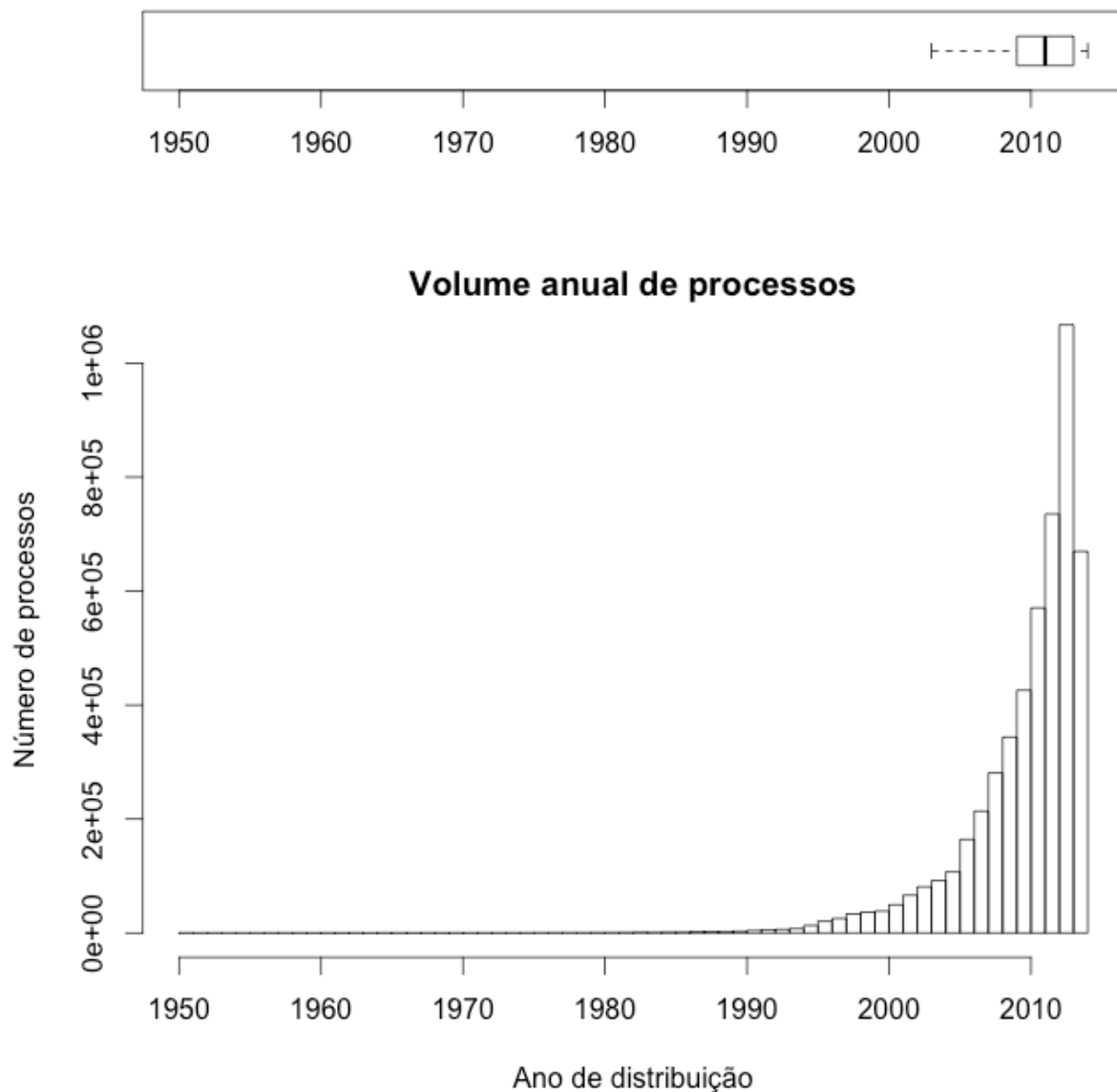


Figura 1

Utilizando o modelo⁴ ARIMA(1,2,0) obteve-se as projeções do número de processos esperados para os próximos 10 anos, apresentadas na Tabela 2 e Figuras 2 e 3. O volume de processos previsto para 2014 é de 1.5 MM⁵, chegando a 5.3 MM em 2020 e 7.4 MM em 2023.

⁴ Detalhes no Anexo III.

⁵MM = milhões.

Projeção do volume anual de processos

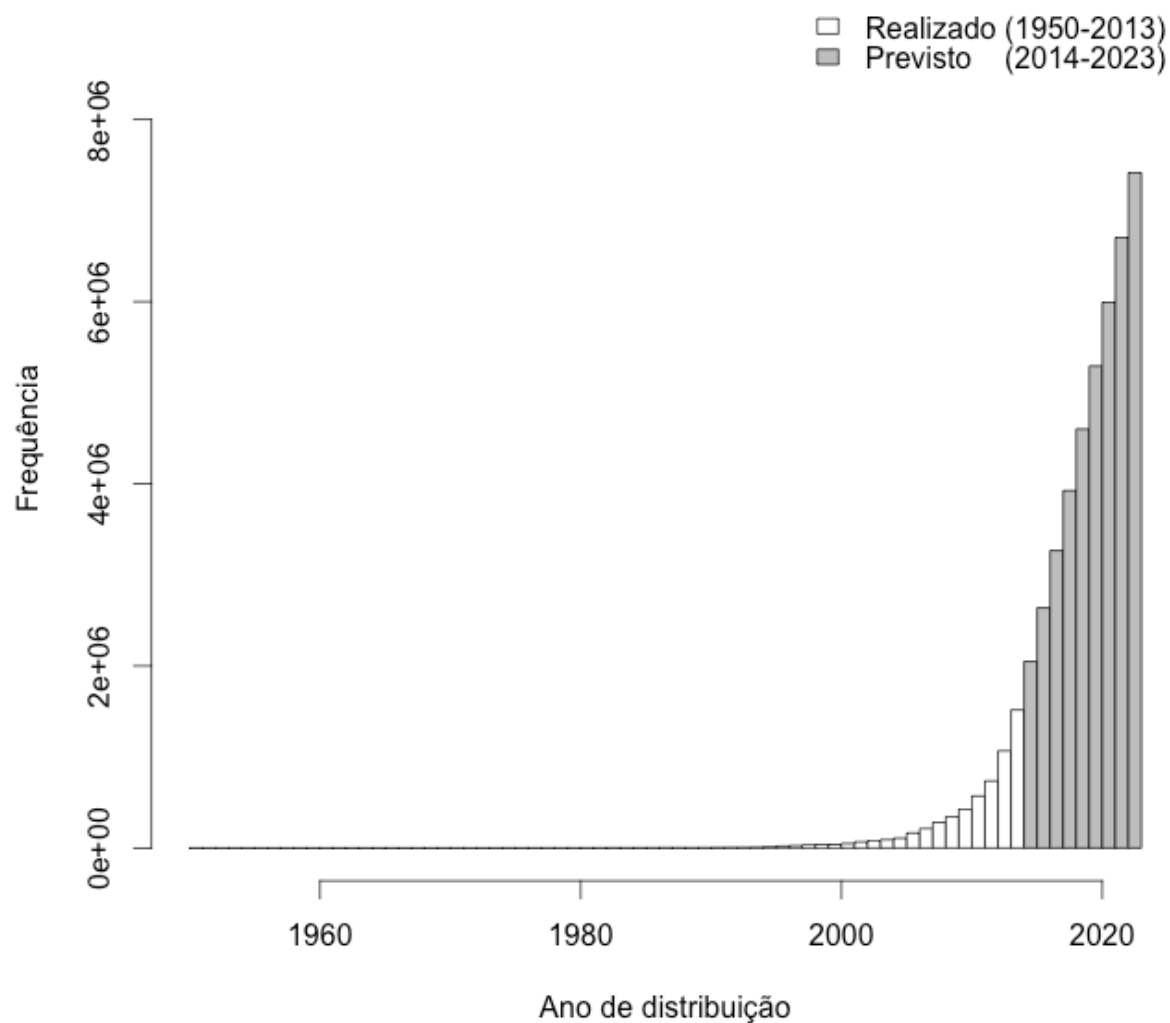


Figura 2

Ano	Projeção	Lo 80	Hi 80	Lo 95	Hi 95
2014	1516620	1488159	1545081	1473093	1560147
2015	2048121	1966158	2130083	1922770	2173472
2016	2637176	2475607	2798745	2390078	2884275
2017	3266557	3000486	3532629	2859636	3673478
2018	3924192	3530317	4318068	3321812	4526573
2019	4601624	4058301	5144946	3770684	5432564
2020	5292925	4580094	6005756	4202744	6383106
2021	5993945	5092987	6894902	4616049	7371841
2022	6701774	5595362	7808186	5009662	8393885
2023	7414373	6086316	8742431	5383285	9445462

Visualizando o gráfico em linha, a Figura 3 apresenta as bandas de confiança para a projeção.

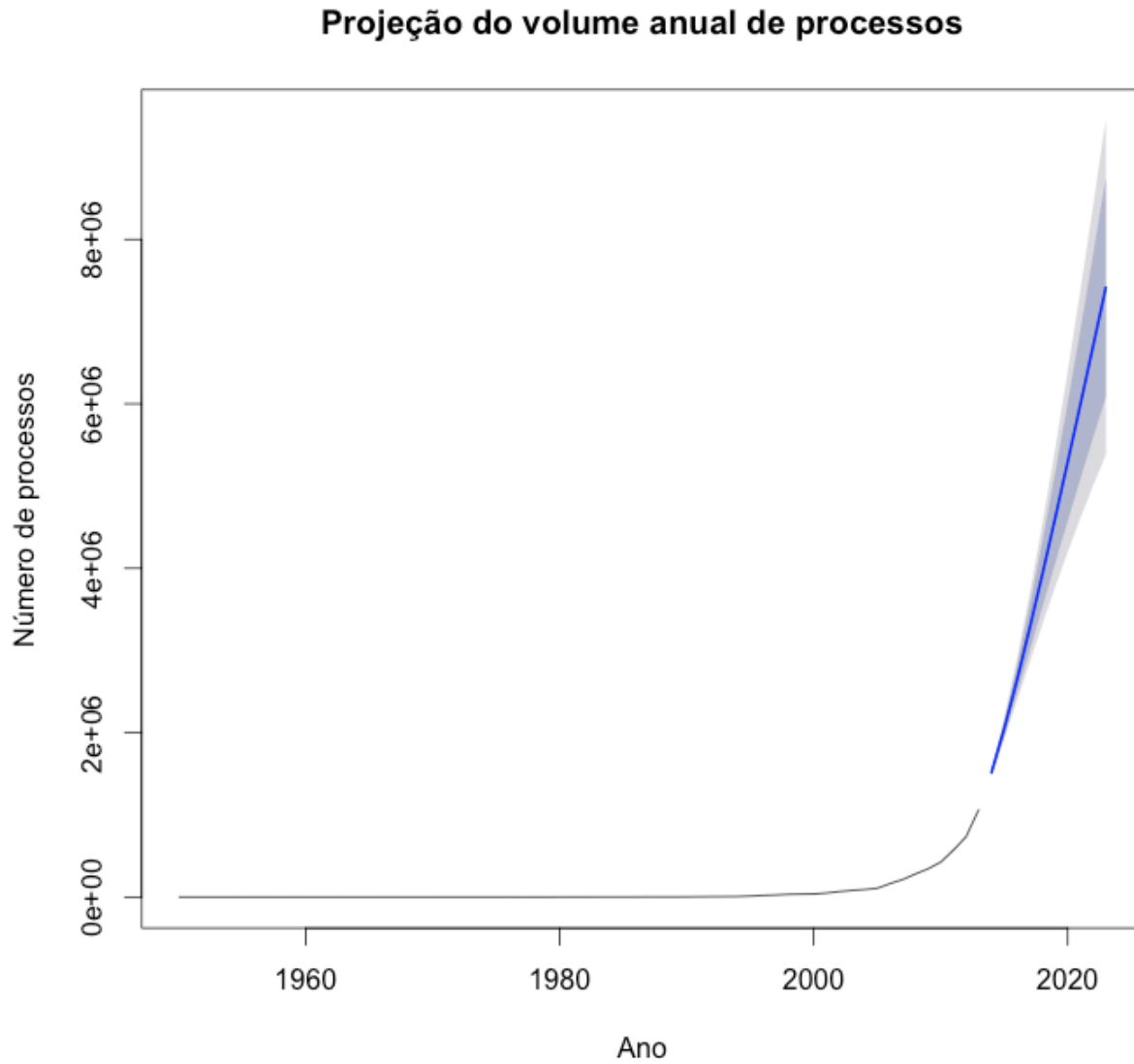


Figura 3

3.1.2 Comarca

Para avaliar o impacto das comarcas no volume de processos, foi gerado um gráfico de Pareto. Este gráfico indica as comarcas no eixo horizontal (não listadas), ordenadas pela frequência. A frequência de cada comarca no eixo vertical esquerdo e o percentual acumulado no eixo vertical direito.

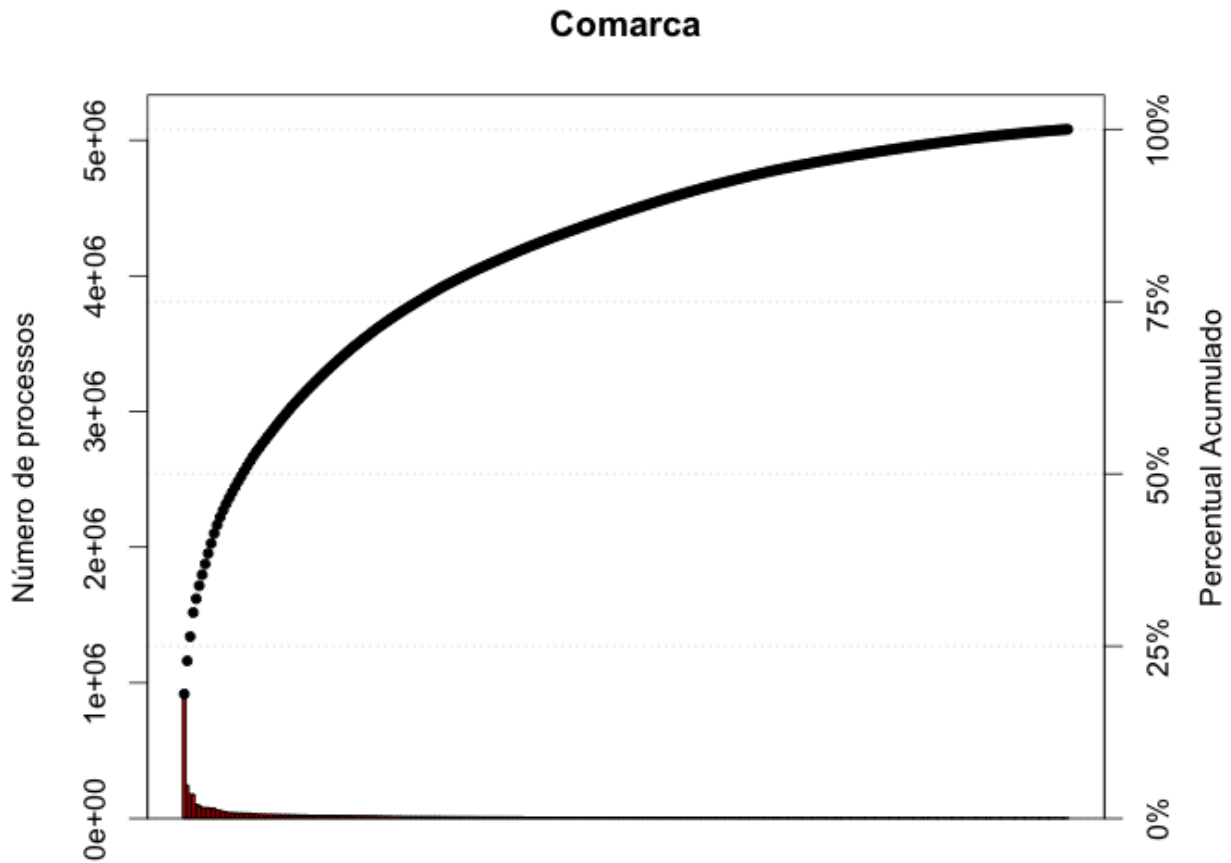


Figura 4

Abaixo estão listadas as 10 comarcas com maior volume de processos, que representam juntas 39.9% dos processos do TJ-MG.

	Processos	Cum.Freq.	Percentage	Cum.Perc.
BELO HORIZONTE	917667	917667	18.063351	18.06335
UBERLÂNDIA	243990	1161657	4.802697	22.86605
CONTAGEM	178806	1340463	3.519616	26.38566
JUIZ DE FORA	177622	1518085	3.496310	29.88197
UBERABA	101847	1619932	2.004756	31.88673
MONTES CLAROS	96157	1716089	1.892754	33.77948
SETE LAGOAS	81026	1797115	1.594915	35.37440
GOVERNADOR VALADARES	78784	1875899	1.550784	36.92518
SANTA LUZIA	77911	1953810	1.533600	38.45878
BETIM	74091	2027901	1.458407	39.91719

3.1.3 Vara

Para avaliar o impacto das varas, foi gerado um gráfico de Pareto. Este gráfico indica a comarca no eixo horizontal (não listadas); a frequência ordenada de cada vara no eixo vertical esquerdo e o percentual acumulado no eixo vertical direito.

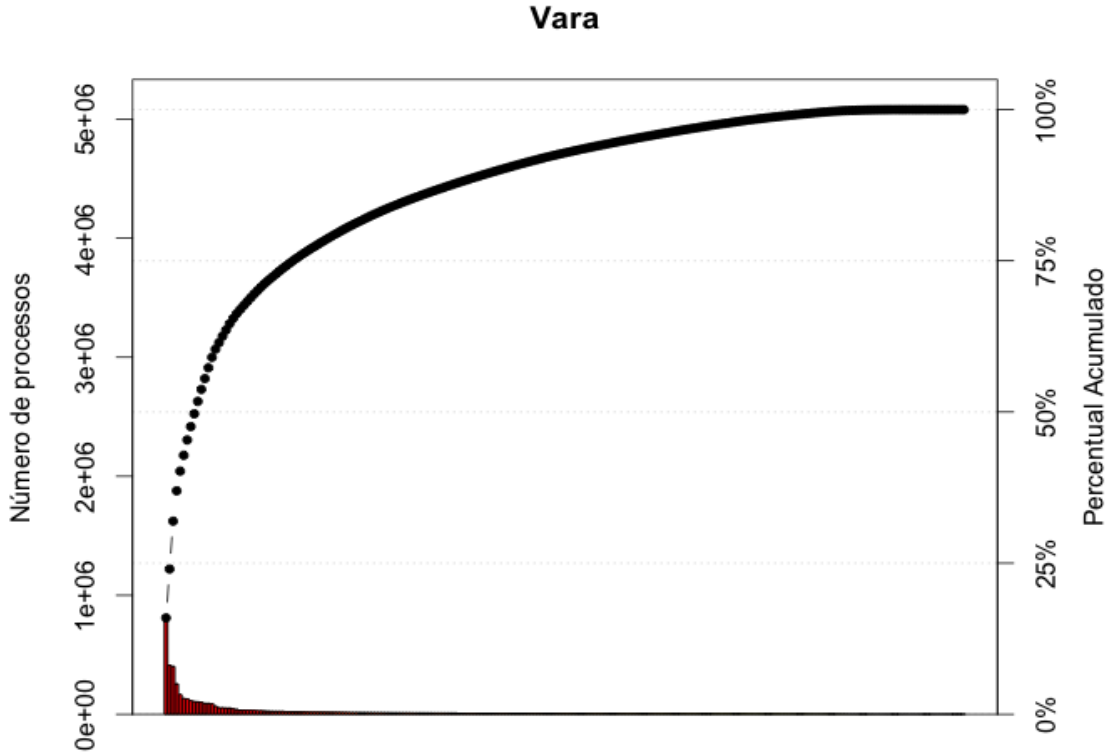


Figura 5

Abaixo estão listadas as 10 varas com maior volume de processos, que representam juntas 51.8% dos processos do TJ-MG.

	Processos	Cum.Freq.	Percentage	Cum.Perc.
SECRETARIA DO JUÍZO	809821	809821	15.940511	15.94051
2ª VARA CÍVEL	410509	1220330	8.080456	24.02097
1ª VARA CÍVEL	402147	1622477	7.915859	31.93683
EXECUÇÕES CRIMINAIS	255273	1877750	5.024792	36.96162
3ª VARA CÍVEL	165240	2042990	3.252583	40.21420
INFÂNCIA E JUVENTUDE	133203	2176193	2.621967	42.83617
VARA CRIMINAL	128726	2304919	2.533842	45.37001
4ª VARA CÍVEL	112801	2417720	2.220374	47.59038
1ª VARA CRIMINAL	107049	2524769	2.107152	49.69754
2ª CÍVEL, CRIME E JIJ	104332	2629101	2.053670	51.75121

3.1.4 Julgado

Ao todo há 74.2% (3,770,020) de processos não julgados, contra 25.8% (1,310,250) de processos julgados. Na Figura 6 está o histograma apresentando as quantidades de processos anuais segmentado pelos processos julgados e não julgados.

Julgado	Frequência	Percentual
Sim	1,310,250	25.8%
Não	3,770,020	74.2%
Total	5,080,270	100.0%

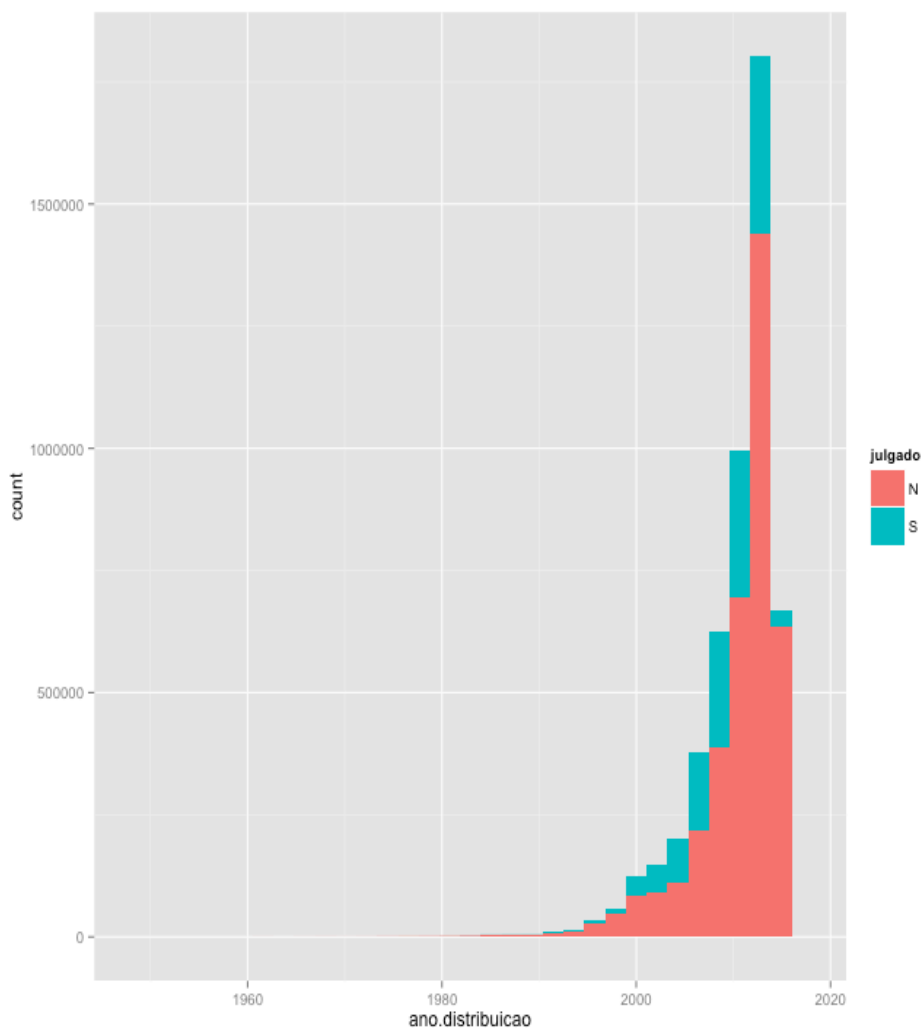
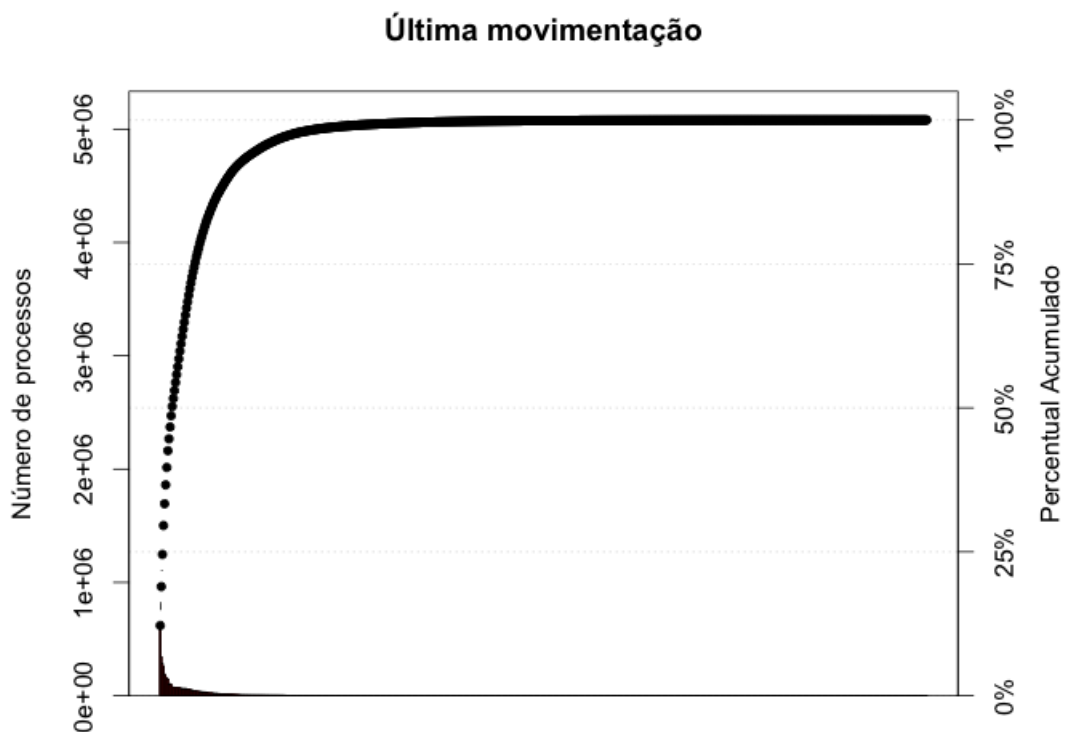


Figura 6

3.1.4 Última Movimentação

Para avaliar o impacto das últimas movimentações foi gerado um gráfico de Pareto. Este gráfico indica a comarca no eixo horizontal (não listadas); a frequência ordenada de cada classe de movimentação no eixo vertical esquerdo e o percentual acumulado no eixo vertical



direito.

Figura 7

Abaixo estão listadas as 10 categorias de movimentação com maior número de processos, que representam juntas 46.7% do total de processos do TJ-MG.

	Processos	Cum.Freq.	Percentage	Cum.Perc.
(0212-1) CONCLUSO PARA DESPACHO/DECISÃO	619400	619400	12.192265	12.19227
(2241-8) PROFERIDO DESPACHO -	343530	962930	6.762042	18.95431
(0436-6) REMETIDOS AUTOS DELEGACIA	284209	1247139	5.594368	24.54868
(2254-1) ATO ORDINATÓRIO	255165	1502304	5.022666	29.57134
(2064-4) SUSPENSO P/DECISÃO JUDICIAL	191579	1693883	3.771040	33.34238
(1058-7) SENTENCIADO EM EXECUÇÃO PENAL	166928	1860811	3.285810	36.62819
(0172-7) REMETIDOS AUTOS TJMG	153895	2014706	3.029268	39.65746
(0681-7) RECEBIDOS OS AUTOS	147337	2162043	2.900181	42.55764
(2205-3) JUNTADA PET (OUTRAS)	104918	2266961	2.065205	44.62284
(0150-3) CONCLUSOS PARA JULGAMENTO	104860	2371821	2.064064	46.68691

3.1.5 Classe

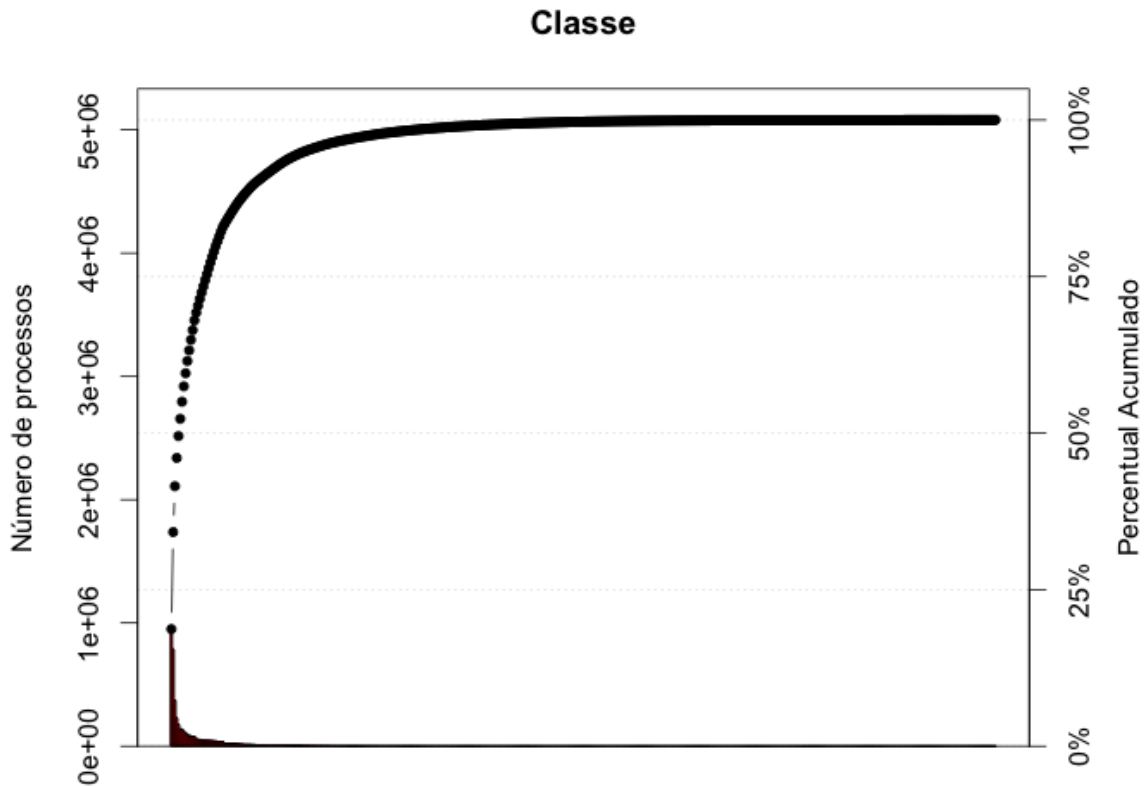


Figura 9

Abaixo estão listadas as 10 classes com maior volume de processos, que representam juntas 61.5% dos processos do TJ-MG.

	Processos	Cum.Freq.	Percentage	Cum.Perc.
(281-6) EXECUÇÃO FISCAL	950365	950365	18.706978	18.70698
(317-8) PROCEDIMENTO ORDINÁRIO	786719	1737084	15.485771	34.19275
(418-4) INQUÉRITO POLICIAL	372071	2109155	7.323843	41.51659
(424-2) AÇÃO PENAL-PROC ORDINÁRIO	229686	2338841	4.521138	46.03773
(854-0) EXECUÇÃO DA PENA	177047	2515888	3.484992	49.52272
(240-2) CUMPRIMENTO DE SENTENÇA	139856	2655744	2.752925	52.27565
(354-1) CRIME C/ PATRIMÔNIO	138701	2794445	2.730190	55.00584
(182-6) EXECUÇÃO TÍTULO EXTRAJUDI	125131	2919576	2.463078	57.46891
(388-9) CARTA PRECATÓRIA	106630	3026206	2.098904	59.56782
(254-3) BUSCA/APREENSÃO DL 911/69	99171	3125377	1.952081	61.51990

3.1.6 Assunto

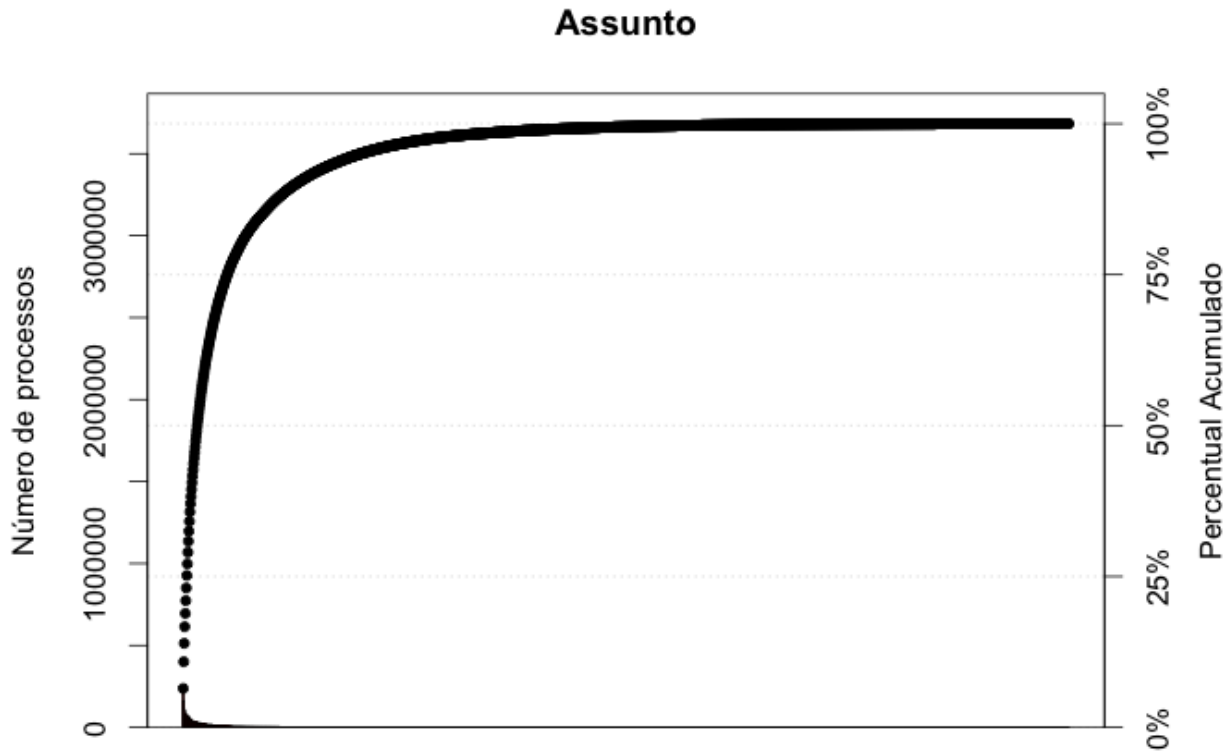


Figura 10

Abaixo estão listadas os 10 assuntos com maior volume de processos, que representam juntas 29.0% dos processos do TJ-MG.

	Processos	Cum.Freq.	Percent.	Cum.Perc.
(6017) Dívida Ativa	237507	237507	6.448352	6.448352
(5952) IPTU	163024	400531	4.426127	10.874479
(7691) Inadimplemento	112463	512994	3.053388	13.927867
(3416) Furto	101508	614502	2.755958	16.683826
(7770) Interpretação/Revisão de Contrato	80983	695485	2.198701	18.882527
(3402) Ameaça	77941	773426	2.116110	20.998637
(5560) Decorrente de Violência Doméstica	77218	850644	2.096481	23.095118
(7687) Inventário e Partilha	76451	927095	2.075657	25.170774
(3608) Tráfico de Drogas e Condutas Afins	71020	998115	1.928204	27.098979
(9582) Alienação Fiduciária	70406	1068521	1.911534	29.010513

3.1.7 Tempo até o julgamento

A duração do processo foi calculada de duas formas distintas:

- para os processos onde o julgamento já ocorreu, foi feito
`data.ultimo.julgamento - data.distribuicao,`
- para os processos onde o julgamento ainda não ocorreu, foi feito
`26/06/2014 - data.distribuicao.`

A tabela X resume os tempos de espera dos processos.

Julgado	Mínimo	Q1	Mediana	Média	Q3	Máximo
Sim	0.067	25.23	50.47	61.27	85.2	780.10
Não	0.033	9.30	25.40	43.58	57.17	783.50
Total	0.033	11.93	31.50	48.14	66.57	783.50

As Figuras 11 e 12 apresentam os histogramas dos tempos de espera dos processos, consolidado e quebrado por julgado (S/N), respectivamente.

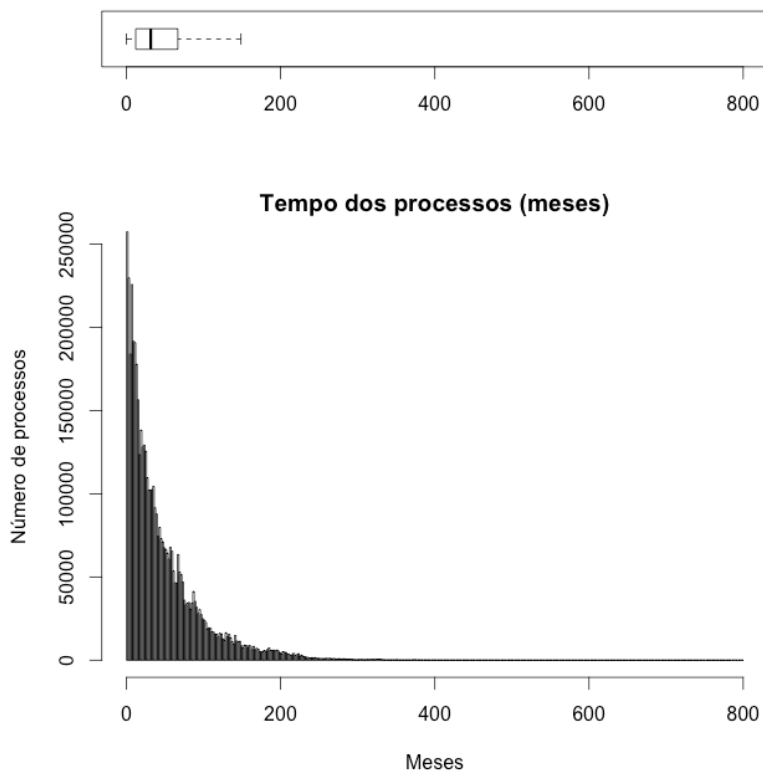


Figura 11

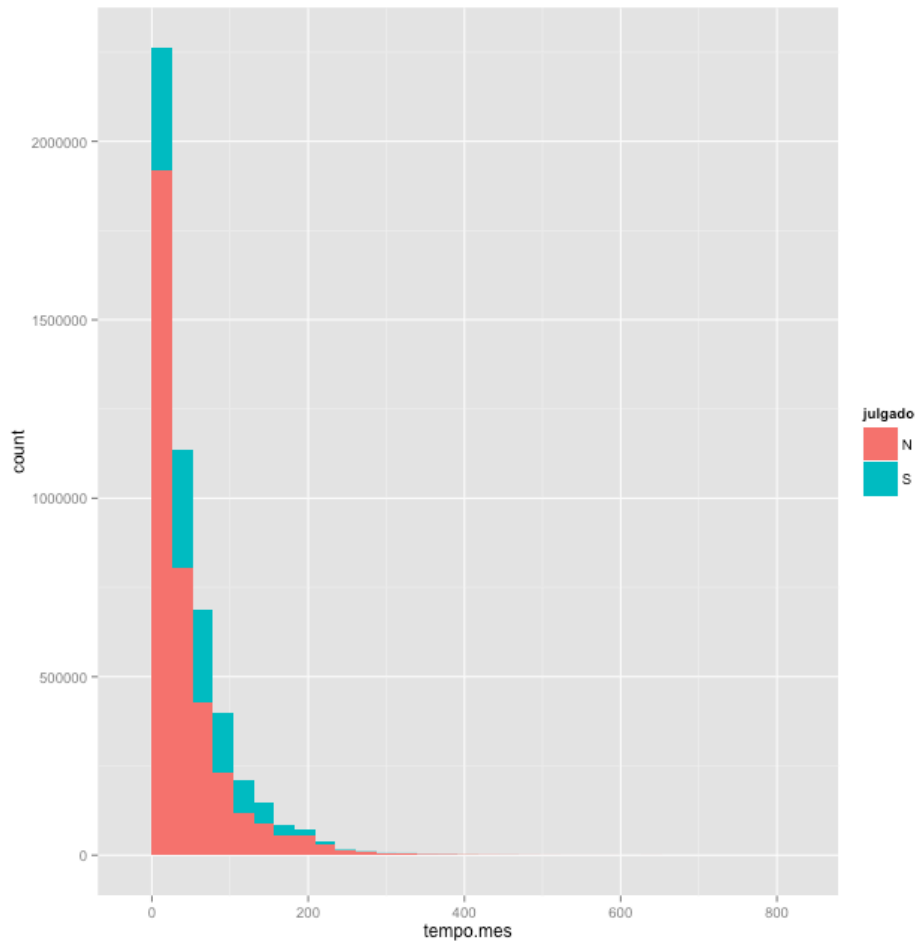


Figura 12

4. Conclusão

Compreendemos que as pesquisas acadêmicas interdisciplinares fundamentam a integração de saberes e que, a Jurimetria é um dos elementos de integração do Direito e Estatística. O esforço simultâneo em coletar dados de base pública e especializar profissionais no uso de ferramental técnico disponível urge.

Foi apresentado um cenário sobre a situação atual do Tribunal de Justiça de Minas Gerais e de alguns conceitos da metodologia adotada, bem como sobre análise textual.

Diante do que foi visto, a questão essencial reside também na responsabilidade do Poder Público contribuir para o desenvolvimento socioeconômico e democrático ao tornar

acessível o conteúdo de suas bases de dados públicas. Isso pode ser claramente exemplificado através do Ofício encaminhado pelo Tribunal de Justiça do Distrito Federal e dos Territórios, no anexo IV.

Enfim, a efetividade dos Tribunais Judiciais poderia impactar positivamente na previsibilidade das decisões judiciais.

REFERÊNCIAS

AMOLDI, Paulo Roberto Colombo. **Novas Perspectivas para o Direito Concursal Brasileiro com os Estudos Interdisciplinares da Economia e da Estatística**. Revista Estudos Jurídicos UNESP, Franca, A. 14 n.19, p. 01-404, 2010.

CORRAR, Luiz J., PAULO, Edilson e DIAS FILHO, José Maria (coordenadores). **Análise Multivariada**: para os cursos de administração, ciências contábeis e economia. FIPECAFI - Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras. 1ª ed. São Paulo: Atlas, 2012.

FÁVERO, Luiz Paulo [et al.]. **Análise de Dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Elsevier, 2009.

SERRA, Márcia Milena Pivatto. **Como Utilizar Elementos da Estatística Descritiva na Jurimetria**. Revista Eletrônica do Curso de Direito das Faculdades OPET. Curitiba:PR. Ano IV, nº 10, p. 156-169, jun-dez 2013.

SPIEGEL, Murray R.. **Estatística**. 3ª Ed. Tradução e revisão técnica Pedro Consentino. Série Schaum. São Paulo: Pearson Makron Books, 1993.

ZABALA, Filipe Jaeger, SILVEIRA, Fabiano Feijó. **Jurimetria: estatística aplicada ao direito**. Revista Direito e Liberdade, Natal, v. 16, n. 1, p. 73-86, jan./abr. 2014. Quadrimestral.

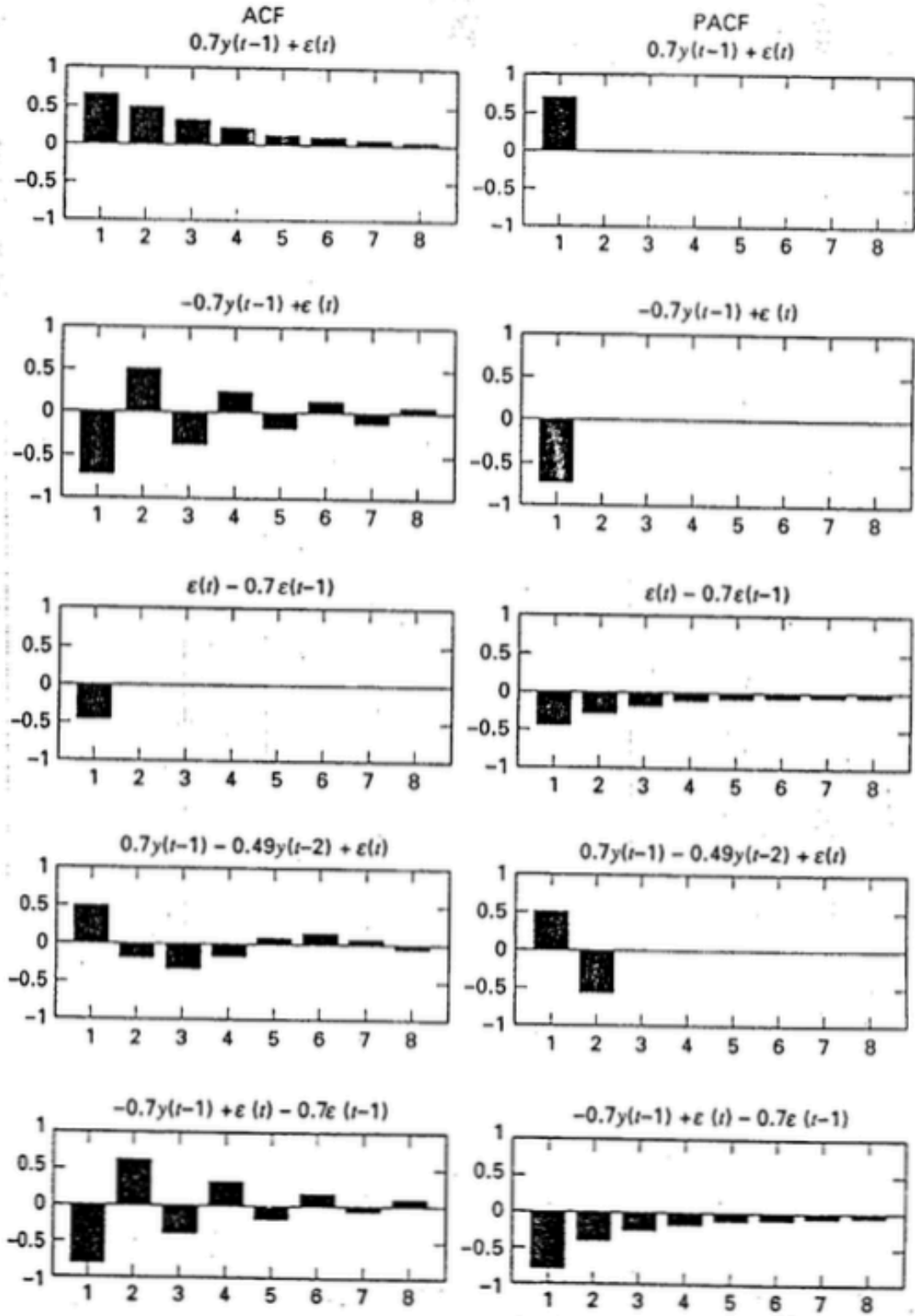
ZABALA, Filipe Jaeger. **Jurimetria: O Direito como Aplicação Estatística**. Boletim ISBrA. Volume 7, nº 1, Julho 2014, p. 9-12.

ANEXOS

Anexo I - Processos com ano.distribuicao inferior a 1950

[1] "001114000305-1" "001609099155-1" "001612008422-9" "001501002783-5"
[5] "003510011298-2" "004006044192-6" "077804006399-3" "005112000467-9"
[9] "005113001147-4" "005613018191-2" "005613013350-9" "005613012798-0"
[13] "002491832875-8" "002485226734-3" "002414012202-9" "002485323139-7"
[17] "002485239160-6" "002412309610-9" "002414012302-7" "002411325419-7"
[21] "002406198006-6" "002485269444-7" "002407451196-5" "002485269445-4"
[25] "007309046961-7" "008614000899-5" "008613002663-5" "008614001592-5"
[29] "008613003928-1" "008613002653-6" "010907009779-4" "010914000683-3"
[33] "011604001848-7" "013306033448-8" "013314001288-0" "018314002545-7"
[37] "018307131876-4" "077513002362-2" "019114001376-1" "019114001347-2"
[41] "019114001377-9" "019114001368-8" "020905044844-5" "020913003766-3"
[45] "020914002141-8" "020911002277-6" "020914002147-5" "020914002144-2"
[49] "020914003813-1" "020914002146-7" "020914002145-9" "021613003695-9"
[53] "021608054950-6" "022307221458-6" "023913001701-9" "023912001417-4"
[57] "024910001941-6" "026107048511-3" "026196002421-2" "010595020249-6"
[61] "010513019925-7" "028712006454-1" "031306203558-6" "031704041092-8"
[65] "031714000808-5" "031910002299-9" "031913003422-0" "033509013692-0"
[69] "033514001194-1" "033805034517-6" "035807013856-7" "014505258401-1"
[73] "014594006905-0" "014596025962-3" "014593008794-8" "037707011469-1"
[77] "038213015067-7" "038410083251-6" "038414005603-5" "038613000876-9"
[81] "039214000189-3" "040014000726-3" "040012000665-7" "040012000402-5"
[85] "040012000403-3" "040012000400-9" "040013000212-6" "040703003557-7"
[89] "043314006255-8" "043113004395-0" "018813000285-3" "047312002789-0"
[93] "047313002680-9" "047106065102-6" "047114001787-5" "047907124930-0"
[97] "047907130600-1" "048013008911-7" "048012017168-5" "048107066886-0"
[101] "048113007136-0" "048613001399-9" "048710006426-9" "049113001681-0"
[105] "021013000587-4" "021011007759-6" "021013000884-5" "051214003230-5"
[109] "051512009667-9" "051512009678-6" "051714001077-1" "052111023156-5"
[113] "054013000181-6" "054014001171-4" "055608015757-2" "055814000730-0"
[117] "055813000289-9" "056711004170-2" "056795006491-3" "056711004171-0"
[121] "057209021898-1" "057214000571-9" "024594003121-5" "024500004959-4"
[125] "024505073112-5" "024505067889-6" "060714003809-4" "061012000851-7"
[129] "061113002278-7" "061111001099-2" "061111001098-4" "062502020652-4"
[133] "064206000828-0" "064707071856-2" "068606182122-5" "068606184594-3"
[137] "070199011086-1" "070101025684-3" "070299019770-0" "070201040774-1"
[141] "070297009285-5" "070297018047-8" "070203077999-6" "070201045537-7"
[145] "070203089648-5" "070203097277-3" "071810001421-3" "072013006777-3"
[149] "072014002203-2" "071311010076-3" "071314004988-1" "071314004566-5"
[153] "070809031434-3"

Anexo II – Gráficos de ACF e PACF teóricos
 Retirado de Enders (1994), p. 79.



Anexo III - Modelo ARIMA(1,2,0)

Definição do modelo

$$Y_t = 0.7007 Y_{t-1} + \varepsilon_t,$$

$$\varepsilon_t \sim N(0,493196904)$$

e

$$Y_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2},$$

Análise dos resíduos

Na Figura 10, os gráficos de autocorrelação (FAC) e autocorrelação parcial (FACP) sugerem resíduos estacionários, dentro das bandas de confiança.

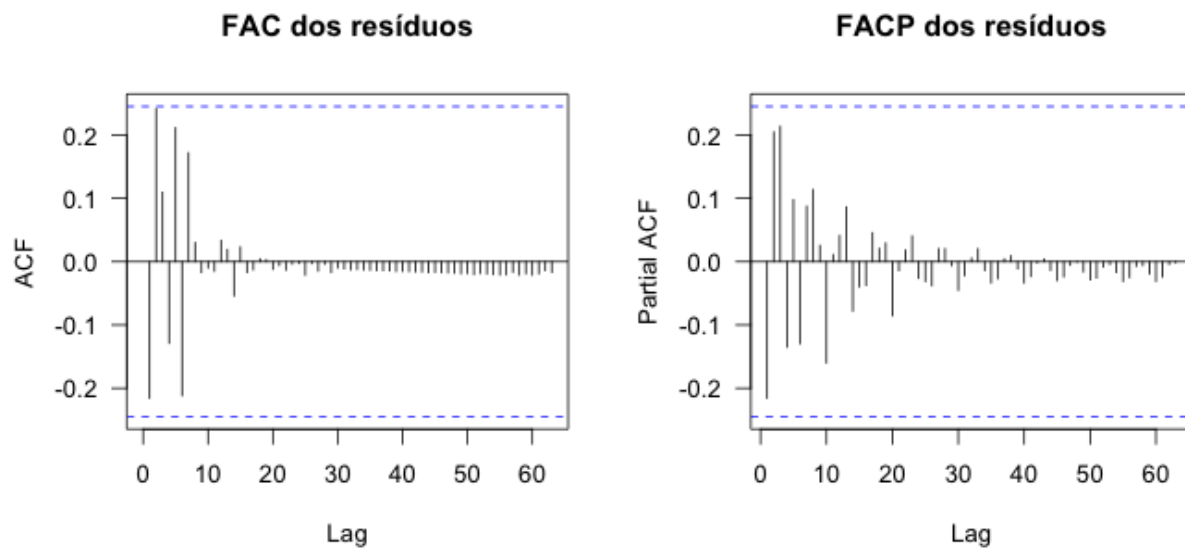


Figura 12

Anexo IV - Código R da análise

```
###
## TJ-MG
#

##
# Removendo objetos

rm(list=ls(all=TRUE))

##
# Instalando e lendo pacotes necessarios

# packs <- c('tseries', 'lmtest', 'forecast', 'tm', 'wordcloud',
#'qcc', 'mosaic')
# install.packages(pack, dep=T)

# Atualizando pacotes
update.packages()

# Chamando pacotes
library(tseries)
library(lmtest)
library(forecast) # auto.arima, forecast
library(qcc) # pareto.chart
library(mosaic) #dotPlot
library(lattice) #dotplot
library(plotrix) # dotplot.mtb
library(TeachingDemos) #dots
library(ggplot2) #ggplot

##
# Criando funções

cleanCorpus<- function(corpus)
{
corpus.tmp<- tm_map(corpus, stripWhitespace) # retirando espaços em branco
corpus.tmp<- tm_map(corpus.tmp, content_transformer(tolower)) # transformando
em minúsculas
corpus.tmp<- tm_map(corpus.tmp, removePunctuation) # retirando pontuação
corpus.tmp<- tm_map(corpus.tmp, removeNumbers) # retirando números
corpus.tmp<- tm_map(corpus.tmp, removeWords, stopwords('portuguese'))
corpus.tmp<- tm_map(corpus.tmp, removeWords, stopwords('english'))
myStopwords<- c(stopwords('portuguese'), "nao", "pag")
corpus.tmp<- tm_map(corpus.tmp, removeWords, myStopwords)
return(corpus.tmp)
}
```

```

##
# Lendo as bases de dados
#####

# Set options
options(stringsAsFactors=F)

tjmg<- read.table('~/base_tj-mg_26-06-2014.txt', head = F, sep = '\\t')
dim(tjmg)
length(unique(tjmg$numero.do.processo)) # processos distintos
head(tjmg)
colnames(tjmg) <- c('ano.distribuicao', 'data.distribuicao', 'comarca',
'vara',
                    'numero.do.processo', 'julgado',
                    'data.ultimo.julgamento',
                    'ultima.movimentacao', 'data.ultima.movimentacao', 'classe',
'assunto')
head(tjmg)

dell <- c(which(tjmg$ano.distribuicao< 1950),
which.max(tjmg$ano.distribuicao))
tjmg<- tjmg[-dell,]
dim(tjmg)
head(tjmg)

##
# Análise exploratória e inferencial

## Ano da Distribuição
summary(tjmg$ano.distribuicao)
boxplot(tjmg$ano.distribuicao)

# maiores
tjmg$numero.do.processo[which.max(tjmg$ano.distribuicao)]
tjmg$ano.distribuicao[which.max(tjmg$ano.distribuicao)]

nf<- layout(mat = matrix(c(1,2),2,1, byrow=TRUE), height = c(1,3))
par(mar=c(5.1, 4.1, 1.1, 2.1))
boxplot(tjmg$ano.distribuicao, horizontal=TRUE, outline=FALSE, ylim =
c(1950,2014))
hist(tjmg$ano.distribuicao, main = "Volume anual de processos", breaks =
1950:2014,
ylab = "Número de processos", xlab = "Ano de distribuição", xlim =
c(1950,2014))

# série temporal

```



```

( ad.2014 <- ts(data = table(tjmg$ano.distribuicao), start = c(1950, 1),
frequency = 1) )
ad.2014
( ad.2013 <- ts(data = ad.2014[-length(ad.2014)], start = c(1950, 1),
frequency = 1) )
ad.2013
par(mfrow = c(1,1), mar = c(5, 4, 4, 2) + 0.1) # parametros gráficos
originais
ts.plot(ad.2014)
ts.plot(ad.2013)

# ajustando o modelo
fit<- auto.arima(ad.2013)
fit
forecast(fit)
forecast(fit)$method
plot(forecast(fit), main = "Projeção do volume anual de processos",
xlab = 'Ano', ylab = 'Número de processos')
forecast(fit)$mean
# table(rep(2014:2023, forecast(fit)$mean)) # demora!
par(mfrow = c(2,2))
fit.acf<- acf(fit$res, lag.max=length(fit$res), plot = F)
fit.acf$acf[1] <- NA
plot(fit.acf, main = 'FAC dos resíduos', las = 1)
pacf(fit$res, main = 'FACP dos resíduos', lag.max=length(fit$res), las = 1)

# apresentando as projeções
base.prev<- c(rep(2014:2023, forecast(fit)$mean), tjmg$ano.distribuicao[
c(which(tjmg$ano.distribuicao == 2014))])
# table(base.prev)
cores <- c(rep('white', 2013-1950+1), rep('gray', 2023-2014+1))

par(mfrow = c(1,1), mar = c(5, 4, 4, 2) + 0.1) # parametros gráficos
originais
hist(base.prev, main = "Projeção do volume anual de processos", breaks =
1950:2023, col = cores,
ylab = "Número de processos", xlab = "Ano de distribuição", xlim =
c(1950,2023), ylim = c(0,9000000) )
legend("topright", legend=c("Realizado (1950-2013)", "Previsto (2014-
2023)"), fill= c('white', 'gray') , bty="n")

## Data distribuição
class(tjmg$data.distribuicao)
head(as.Date(tjmg$data.distribuicao, "%d/%m/%y"))
boxplot(as.Date(tjmg$data.distribuicao, "%d/%m/%y"))
data.distr<- head(as.Date(tjmg$data.distribuicao, "%d/%m/%y"))
summary(data.distr)
ano <- as.numeric(format(data.distr, "%Y"))

```

```

summary(data.distr[which(data.distr>as.Date('26/06/14', '%d/%m/%y'))])

data.distr[ano > 2014] <- data.distr[ano > 2014] - 100 * 365.25
head(data.distr)
summary(data.distr)

## Comarca
table(tjmg$comarca)
(pc.comarca<- pareto.chart(table(tjmg$comarca), xaxt = 'n', main = "Comarca",
ylab = "Número de processos", ylab2 = "Percentual Acumulado") )
head(pc.comarca, 10)

## Vara
table(tjmg$vara)
(pc.vara<- pareto.chart(table(tjmg$vara), xaxt = 'n', main = "Vara",
ylab = "Número de processos", ylab2 = "Percentual Acumulado") )
head(pc.vara, 10)

#####
## Julgado
table(tjmg$julgado)
round(prop.table(table(tjmg$julgado)),3)
by(tjmg$julgado, tjmg$ano.distribuicao, table) # por ano.distribuicao

par(mfrow = c(1,1), mar = c(5, 4, 4, 2) + 0.1) # parametros gráficos
originais

# histograma do ggplot2
with(tjmg, table(ano.distribuicao, julgado))
ggplot(tjmg, aes (x = ano.distribuicao, fill = julgado)) + geom_bar()

#####
## Data ultimo julgamento
class(tjmg$data.ultimo.julgamento)
head(as.Date(tjmg$data.ultimo.julgamento[1:1000], "%d/%m/%y"), 100)

## Última movimentação
table(tjmg$ultima.movimentacao)
(pc.ultima.movimentacao<- pareto.chart(table(tjmg$ultima.movimentacao), xaxt
= 'n',
main = "Última movimentação",
ylab = "Número de processos", ylab2 = "Percentual Acumulado") )

```

```

head(pc.ultima.movimentacao, 10)

# wordcloud
# write(t(tjmg$ultima.movimentacao),
#       '~/ultima-movimentacao.txt',
#       ncolumns = 1)
txt<- '~'
(um <- Corpus(DirSource(txt, encoding = "UTF-8"), readerControl =
list(language = "lat")))
cl.um <- cleanCorpus(um[7])
inspect(cl.um)
wordcloud(cl.um, max.words=100, random.order=T, rot.per=0.35,
use.r.layout=FALSE, colors=brewer.pal(8, "Dark2"))

## Data última movimentação
class(Data.Última.Movimentação)
head(as.Date(Data.Última.Movimentação, "%d/%m/%y"))

## Classe
table(tjmg$classe)
(pc.classe<- pareto.chart(table(tjmg$classe), xaxt = 'n',
main = "Classe",
ylab = "Número de processos", ylab2 = "Percentual Acumulado" )
head(pc.classe, 10)

## Assunto
table(tjmg$assunto)
del2 <- which(tjmg$assunto != 'Nenhum assunto cadastrado')
assunto <- table(tjmg$assunto[del2])
(pc.assunto<- pareto.chart(assunto, xaxt = 'n',
main = "Assunto",
ylab = "Número de processos", ylab2 = "Percentual Acumulado" )
head(pc.assunto, 10)

## Tempo até o julgamento
class(tjmg$data.ultimo.julgamento)
head(as.Date(tjmg$data.ultimo.julgamento, "%d/%m/%y"), 30)
par(mfrow = c(1,1), mar = c(5, 4, 4, 2) + 0.1) # parametros gráficos
originais
boxplot(as.Date(tjmg$data.ultimo.julgamento, "%d/%m/%y"))

class(tjmg$data.distribuicao)
head(as.Date(tjmg$data.distribuicao, "%d/%m/%y"), 30)
head(tjmg$data.distribuicao, 30)

ifelse(tjmg$julgado == 'S',
tempo.dias<- as.Date(tjmg$data.ultimo.julgamento, "%d/%m/%y") - data.distr,
tempo.dias<- as.Date('26/06/14', "%d/%m/%y") - data.distr)

```

```

summary(tempo.dias)
head(tempo.dias,100); class(tempo.dias); length(tempo.dias)
summary(as.numeric(tempo.dias))
tempo.mes<- as.numeric(tempo.dias/30)
head(tempo.mes,100); class(tempo.mes); length(tempo.mes)
summary(tempo.mes)
by(tempo.mes, tjmg$julgado, summary)

nf<- layout(mat = matrix(c(1,2),2,1, byrow=TRUE), height = c(1,3))
par(mar=c(5.1, 4.1, 1.1, 2.1))
boxplot(tempo.mes, horizontal=TRUE, outline=FALSE, ylim = c(0,800))
hist(tempo.mes, main = "Tempo dos processos (meses)", breaks = seq(0, 800,
2),
ylab = "Número de processos", xlab = "Meses")

# quebrando por julgado
with(tjmg, table(tempo.mes, julgado))
ggplot(tjmg, aes (x = tempo.mes, fill = julgado)) + geom_bar()

```

Anexo IV – Ofício do Tribunal de Justiça do Distrito Federal e dos Territórios



TJDFT

Poder Judiciário da União

TRIBUNAL DE JUSTIÇA DO DISTRITO FEDERAL E DOS TERRITÓRIOS

COSIST/SGC

Coordenação de Projetos e Sistemas de 1ª Instância – COSIST

Secretaria-Geral da Corregedoria



PA nº 22.031/2014

À Ouvidoria-Geral,

Em atenção à solicitação esta Coordenação informa que o sistema informatizado da primeira instância dispõe de diversos relatórios estatísticos.

Contudo, os relatórios exibem dados quantitativos que requerem a indicação de parâmetros restritivos, tais como competência, o tipo de classe processual, período de verificação, entre outros.

No caso em questão o pedido não limita os dados desejados, e indica apenas a extração da base de dados públicos, o que não é viável tendo em vista a quantidade de registros daquela, e a ausência de acesso desta Coordenação a tais registros para extração.

Além disso, não há integração entre os sistemas informatizados da primeira e da segunda instância, de modo que a verificação de algumas informações relacionadas no presente exigirá a verificação manual de processos.

Diante do exposto, esta Coordenação informa a impossibilidade de atendimento da solicitação.

COSIST, em 9 de dezembro de 2014.

Kátia S. de S. Prates
Kátia Santana de Souza Prates

Coordenadora de Projetos e Sistemas de Primeira Instância