

Workshop on Legal Digital Intelligence
20/01/2023

Topological Data Analysis + Súmulas Vinculantes

Lucas Reck, Henrique Hennes, Jean Roberto Ponciano, Jorge Poco,
Raphaël Tinarrage

Beatriz Sabdin Chagas, Carla Marcondes Damian, Ana Clara Macedo
Jaccoud, Pedro Burlini de Oliveira

As súmulas vinculantes surgiram com a reforma do judiciário de 2004.

Têm como objetivo:

- garantir a isonomia, seguindo a lógica do *common law*,
- aumentar a eficiência do poder judiciário, no sentido de deixar de originar processos com demandas idênticas.

Muitas críticas afirmam que estes objetivos não estão sendo atingidos.



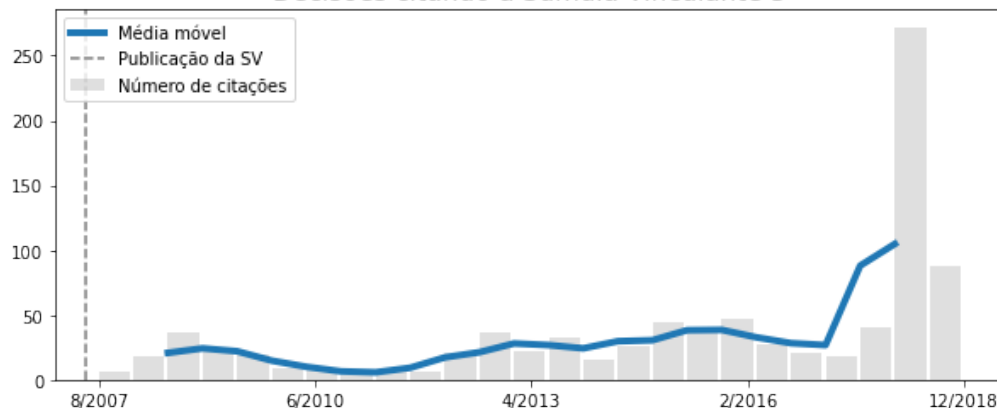
Nosso conjunto de dados consiste em 150'00 decisões do STF, entre 1988 e 2018.

Existem 58 SV. Pretendemos estudar apenas as 10 mais citadas: 3, 4, 10, 11, 14, 17, 20, 26, 33, 37.

Curvas de citações das SV

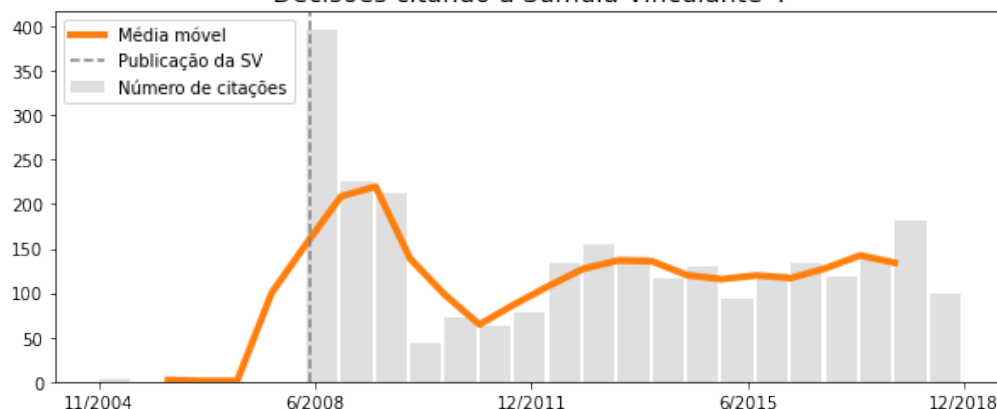
3/17 (1/3)

Decisões citando a Súmula Vinculante 3



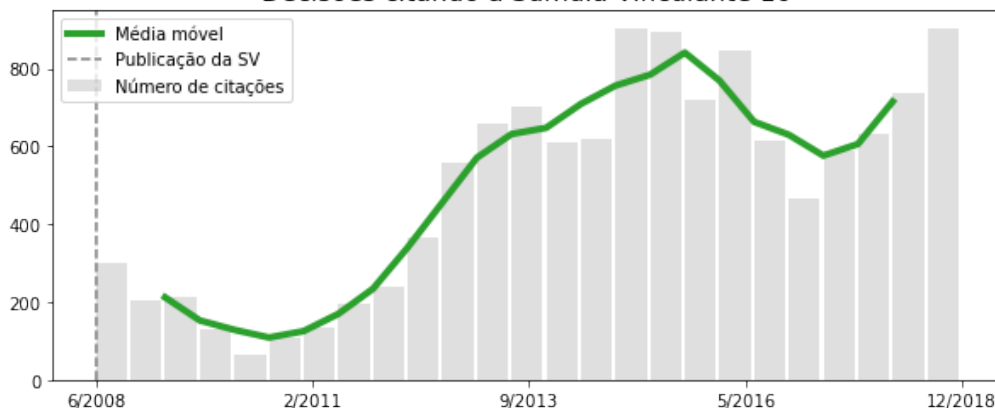
Nos processos perante o Tribunal de Contas da União asseguram-se o contraditório e a ampla defesa quando da decisão puder resultar anulação ou revogação de ato administrativo que beneficie o interessado, excetuada a apreciação da legalidade do ato de concessão inicial de aposentadoria, reforma e pensão.

Decisões citando a Súmula Vinculante 4



Salvo nos casos previstos na Constituição, o salário mínimo não pode ser usado como indexador de base de cálculo de vantagem de servidor público ou de empregado, nem ser substituído por decisão judicial.

Decisões citando a Súmula Vinculante 10

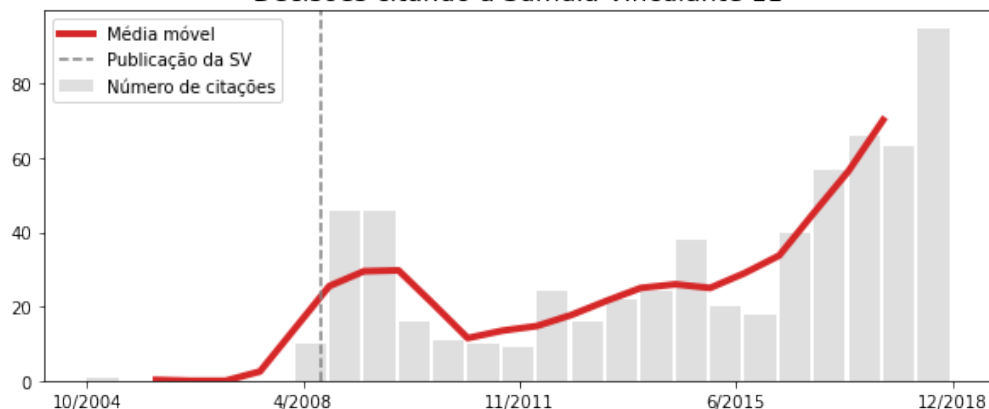


Viola a cláusula de reserva de plenário (CF, artigo 97) a decisão de órgão fracionário de Tribunal que embora não declare expressamente a inconstitucionalidade de lei ou ato normativo do poder público, afasta sua incidência, no todo ou em parte.

Curvas de citações das SV

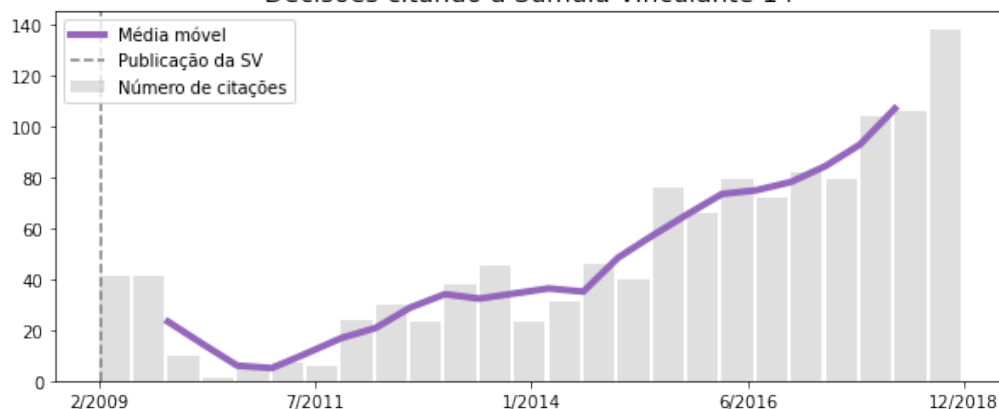
3/17 (2/3)

Decisões citando a Súmula Vinculante 11



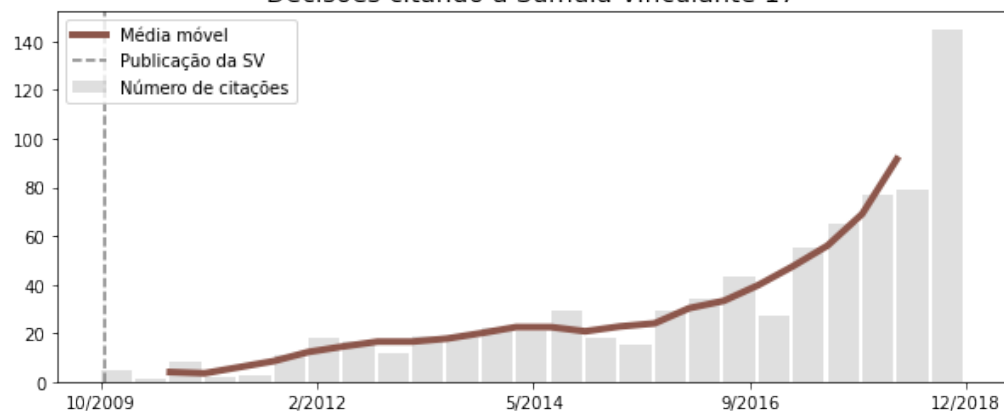
Só é lícito o uso de algemas em casos de resistência e de fundado receio de fuga ou de perigo à integridade física própria ou alheia, por parte do preso ou de terceiros, justificada a excepcionalidade por escrito, sob pena de responsabilidade disciplinar, civil e penal do agente ou da autoridade e de nulidade da prisão ou do ato processual a que se refere, sem prejuízo da responsabilidade civil do Estado.

Decisões citando a Súmula Vinculante 14



É direito do defensor, no interesse do representado, ter acesso amplo aos elementos de prova que, já documentados em procedimento investigatório realizado por órgão com competência de polícia judiciária, digam respeito ao exercício do direito de defesa.

Decisões citando a Súmula Vinculante 17

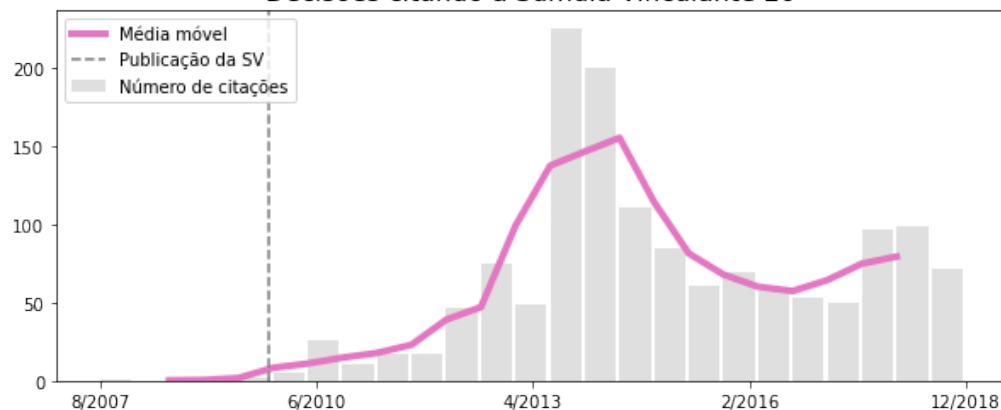


Durante o período previsto no parágrafo 1º do artigo 100 da Constituição, não incidem juros de mora sobre os precatórios que nele sejam pagos.

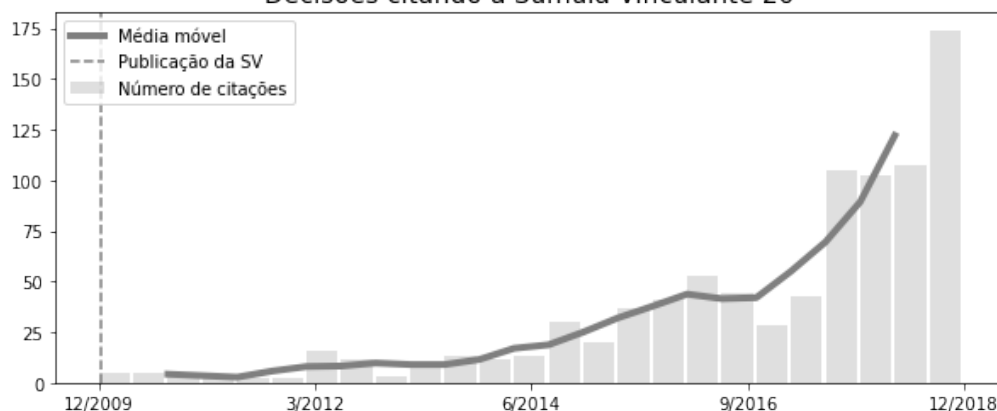
Curvas de citações das SV

3/17 (3/3)

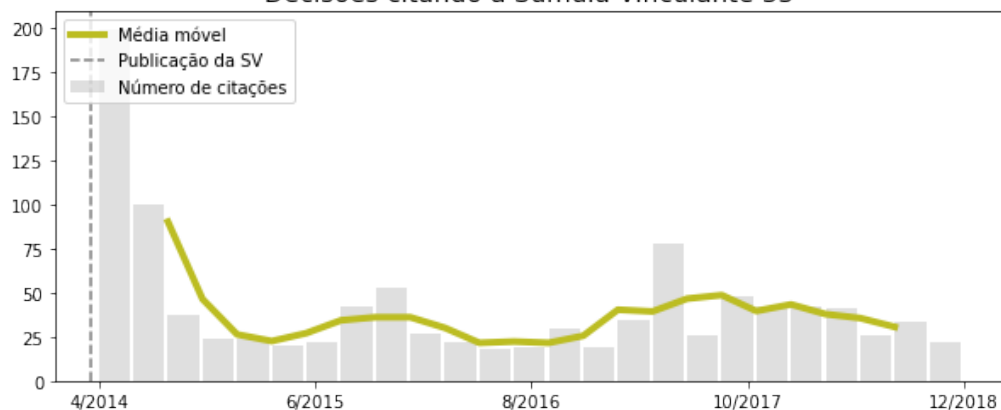
Decisões citando a Súmula Vinculante 20



Decisões citando a Súmula Vinculante 26



Decisões citando a Súmula Vinculante 33



A Gratificação de Desempenho de Atividade Técnico-Administrativa - GDATA, instituída pela Lei nº 10.404/2002, deve ser deferida aos inativos nos valores correspondentes a 37,5 (trinta e sete vírgula cinco) pontos no período de fevereiro a maio de 2002 e, nos termos do artigo 5º, parágrafo único, da Lei nº 10.404/2002, no período de junho de 2002 até a conclusão dos efeitos do último ciclo de avaliação a que se refere o artigo 1º da Medida Provisória no 198/2004, a partir da qual passa a ser de 60 (sessenta) pontos.

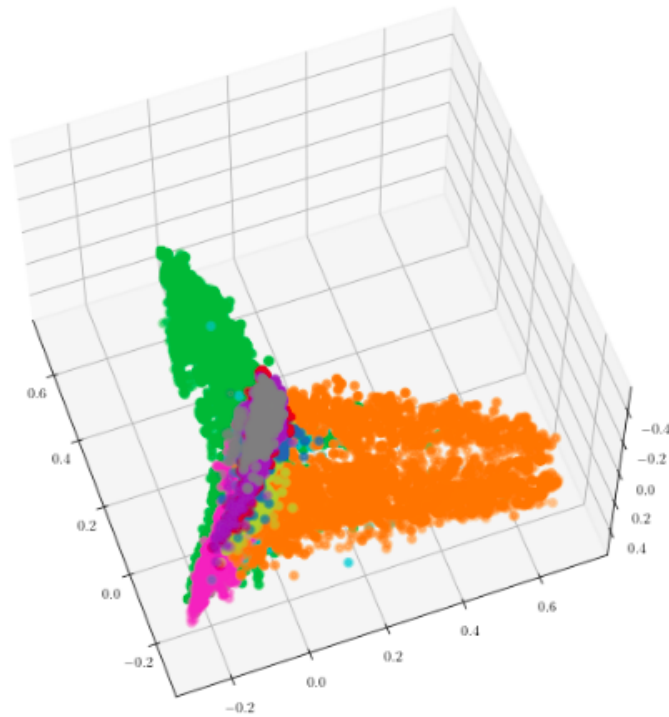
Para efeito de progressão de regime no cumprimento de pena por crime hediondo, ou equiparado, o juízo da execução observará a inconstitucionalidade do art. 2º da Lei n. 8.072, de 25 de julho de 1990, sem prejuízo de avaliar se o condenado preenche, ou não, os requisitos objetivos e subjetivos do benefício, podendo determinar, para tal fim, de modo fundamentado, a realização de exame criminológico.

Aplicam-se ao servidor público, no que couber, as regras do regime geral da previdência social sobre aposentadoria especial de que trata o artigo 40, § 4º, inciso III da Constituição Federal, até a edição de lei complementar específica.

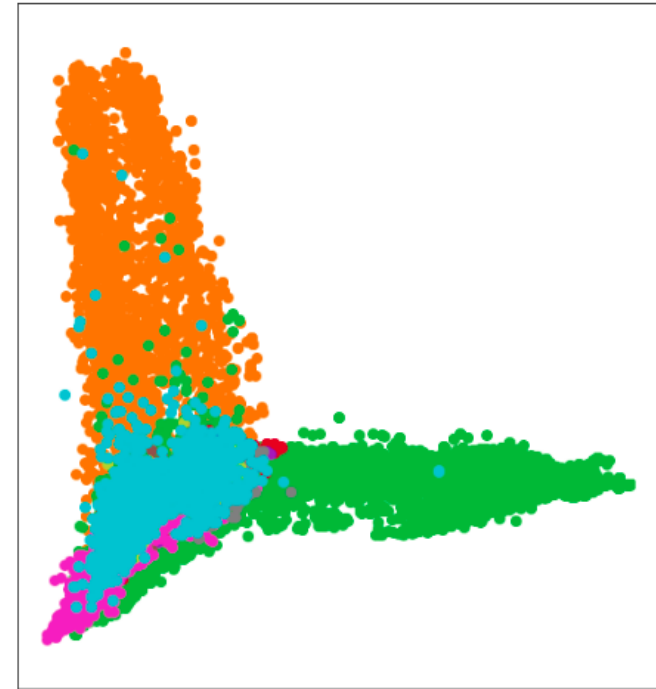
I - Análise exploratória com o TDA

II - Análise empírica da SV 26

Aplicamos TF-IDF a cada documento.



PCA no \mathbb{R}^3



PCA no \mathbb{R}^2

SV:



3



4



10



11



14



17



20



26



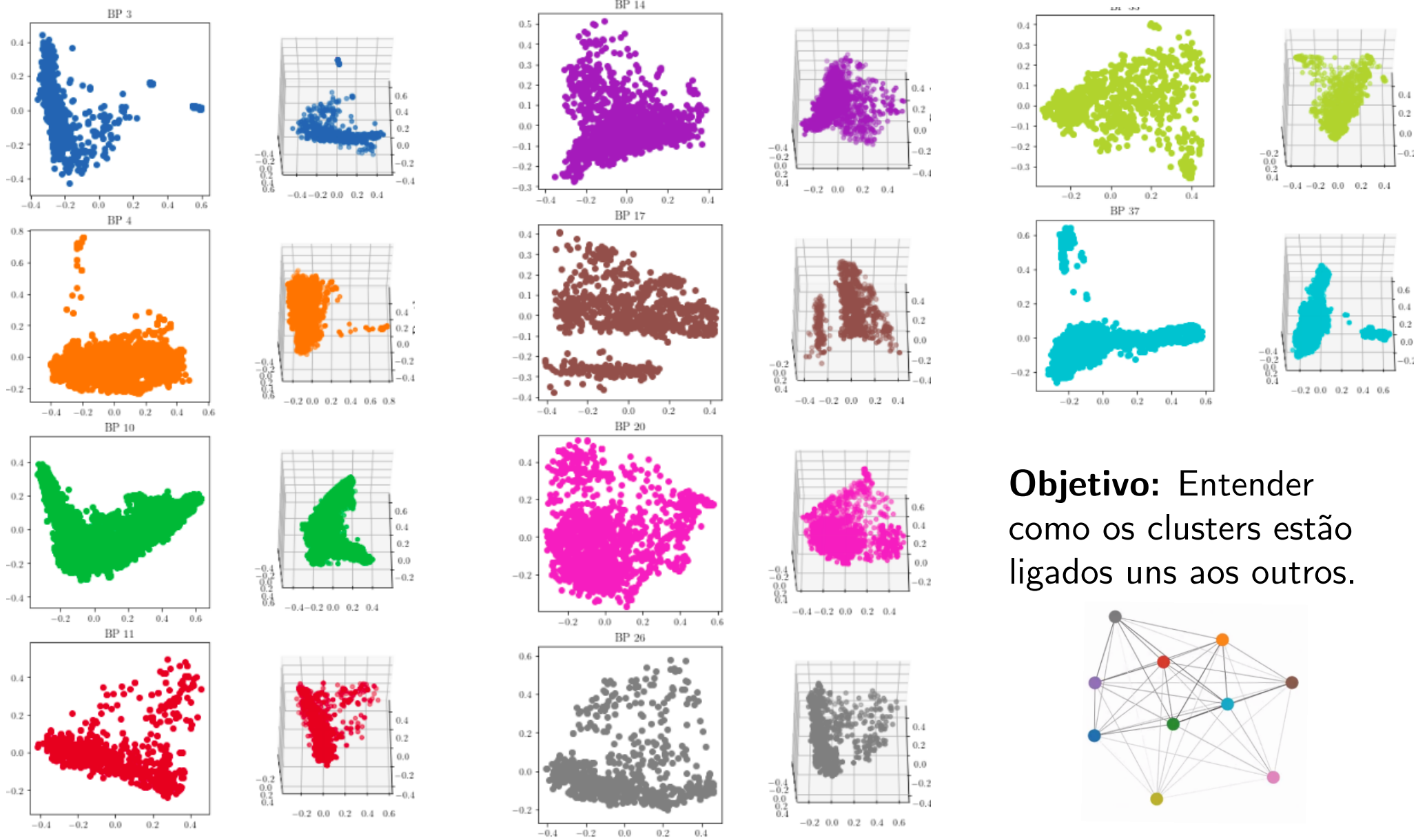
33



37

TF-IDF

5/17 (2/3)



Objetivo: Entender como os clusters estão ligados uns aos outros.



SV:



Não só a distância euclidiana pode ser interessante. Por exemplo, a distância *coseno* parece ser particularmente pertinente na NLP.

$$d_{\cos}(x, y) = 1 - \cos(x, y) = 1 - \frac{x \cdot y}{|x||y|}$$

comparação euclidiano / cosseno

SV:



3



4



10



11



14



17



20



26



33



37

Números de Betti

6/17 (1/9)


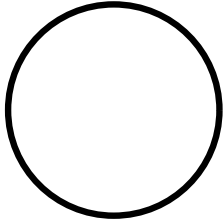
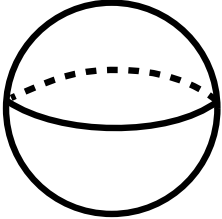
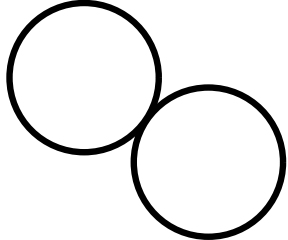
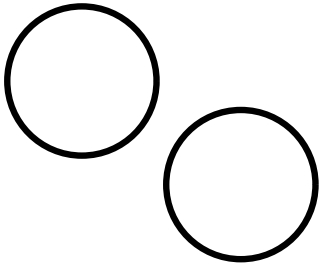
Para todo espaço topológico X , podemos definir uma sequência de números inteiros

$$\beta_0(X), \beta_1(X), \beta_2(X), \beta_3(X), \dots$$

chamados de **números de Betti**.


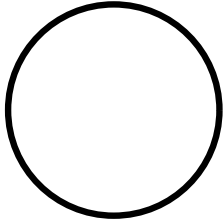
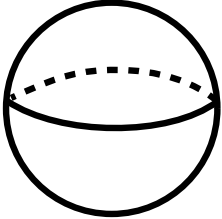
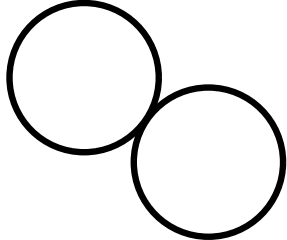
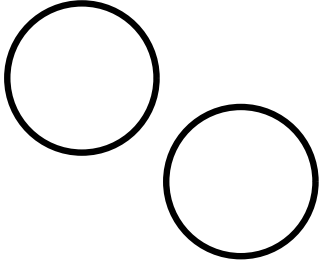
Construção dos números Betti: baseada na teoria da homologia. O grupo de homologia (no $\mathbb{Z}/2\mathbb{Z}$) $H_i(X)$ pode ser escrito como

$$H_i(X) = (\mathbb{Z}/2\mathbb{Z})^{\beta_i(X)}.$$

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0


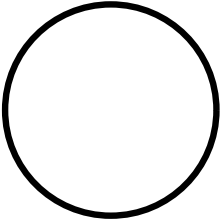
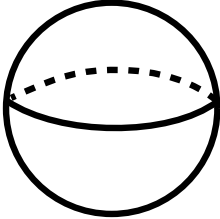
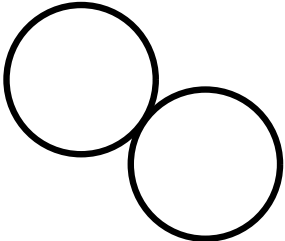
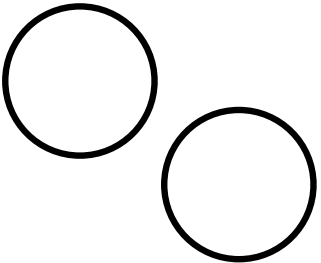
Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0


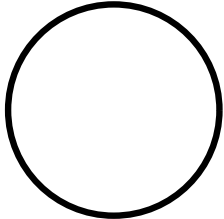
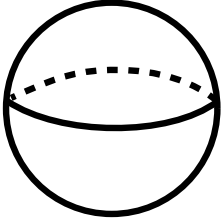
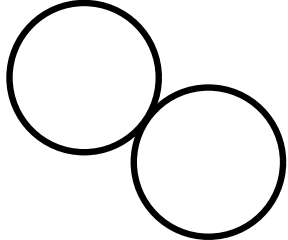
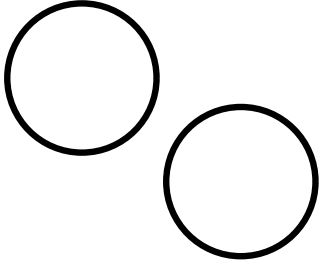
Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0


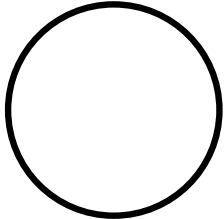
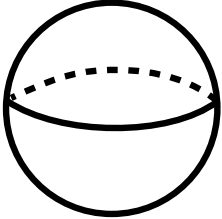
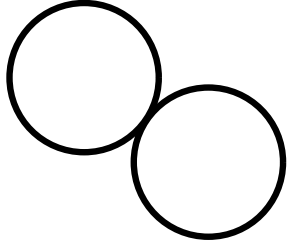
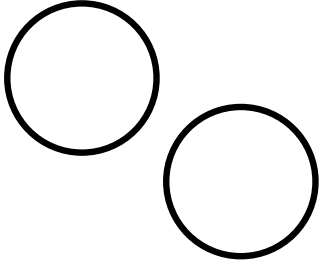
Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

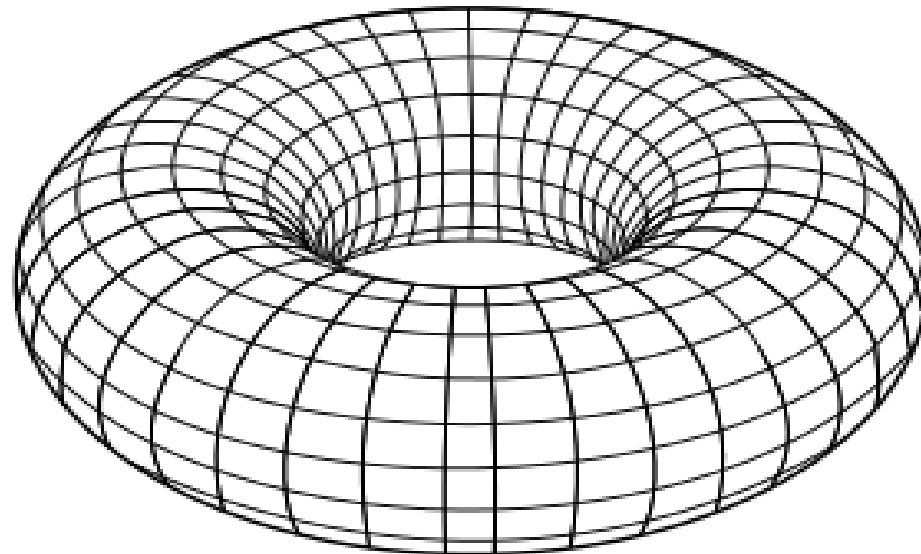
X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

Exemplo: Números de Betti do toro:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \dots$$

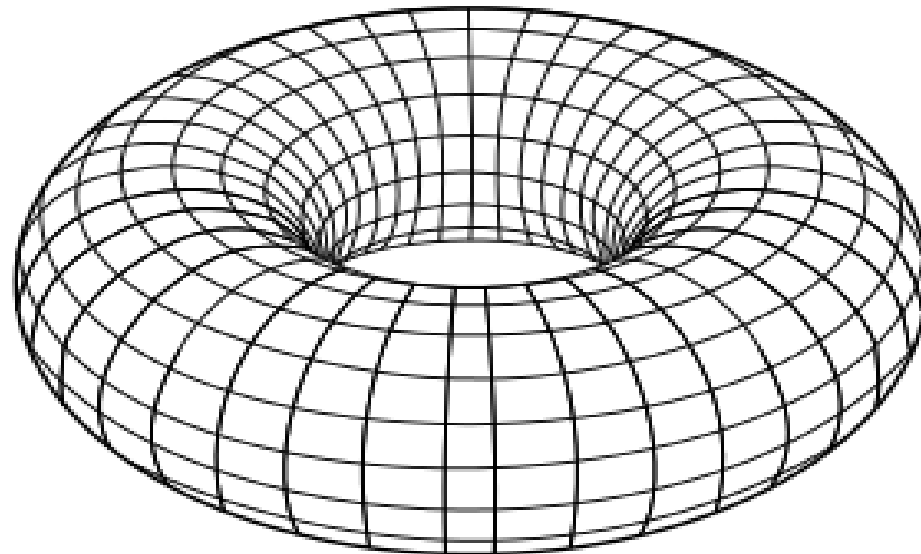


Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

Exemplo: Números de Betti do toro:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \dots$$

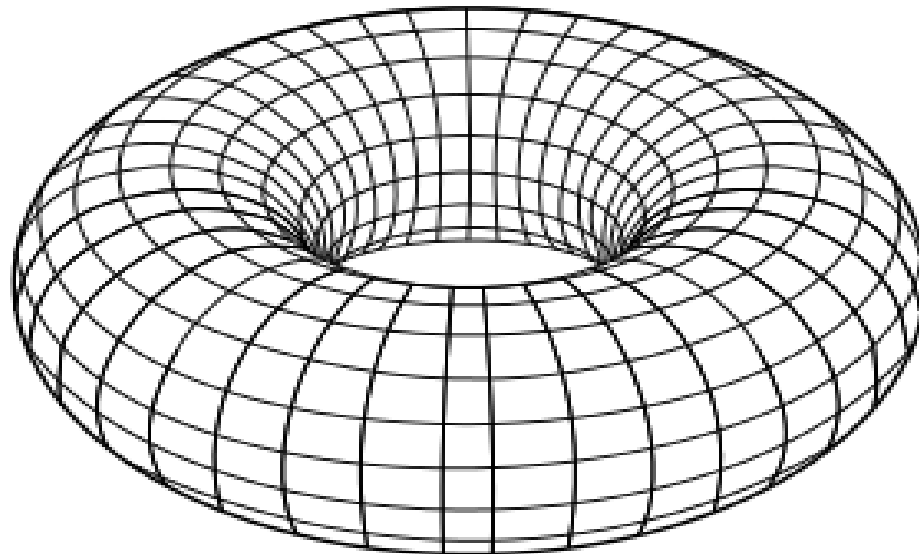


Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

Exemplo: Números de Betti do toro:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \dots$$

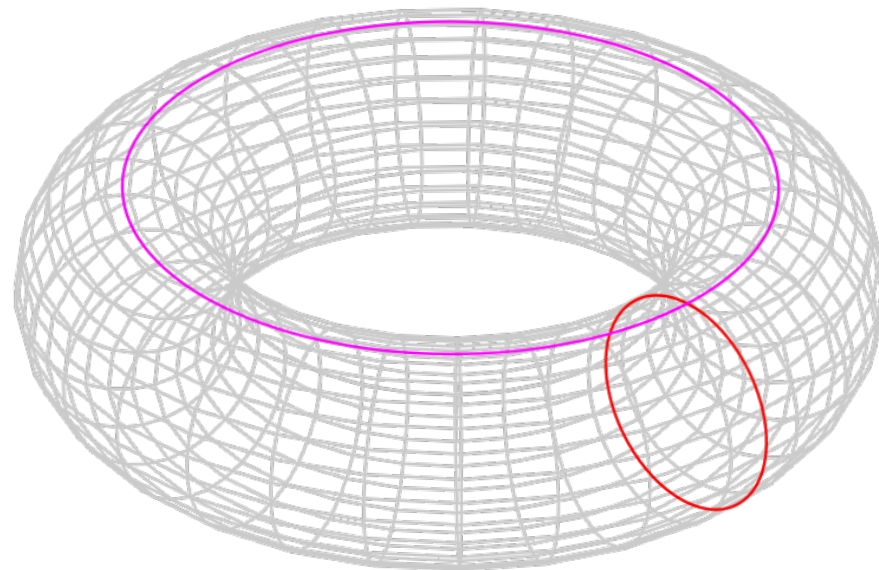


Interpretação:

- $\beta_0(X)$ é o número de componentes conexas de X
- $\beta_1(X)$ é o número de 'buracos' em X
- $\beta_2(X)$ é o número de 'vazios' em X
- ...

Exemplo: Números de Betti do toro:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \dots$$

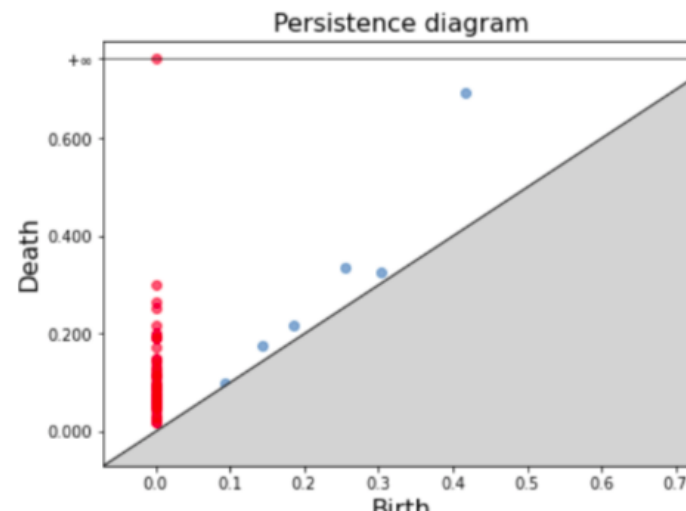


A **homologia persistente** (PH) é usada para estimar a homologia dos conjuntos de dados.

Filtração de Čech



Output: diagrama de persistência

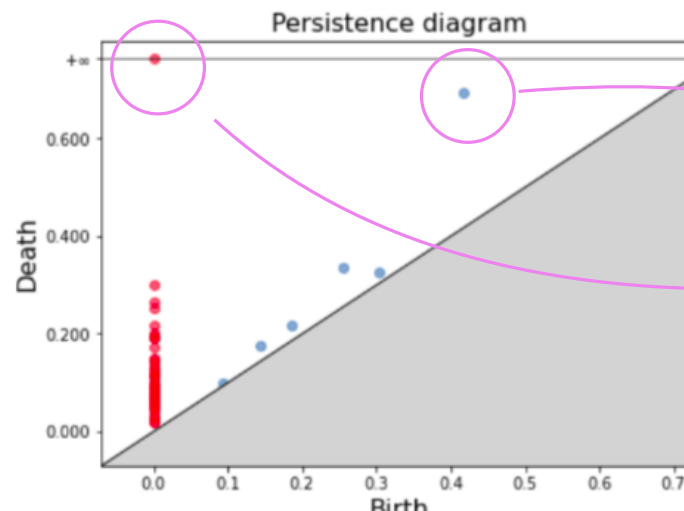


A **homologia persistente** (PH) é usada para estimar a homologia dos conjuntos de dados.

Filtração de Čech



Output: diagrama de persistência



$$\beta_1(X) = 1$$

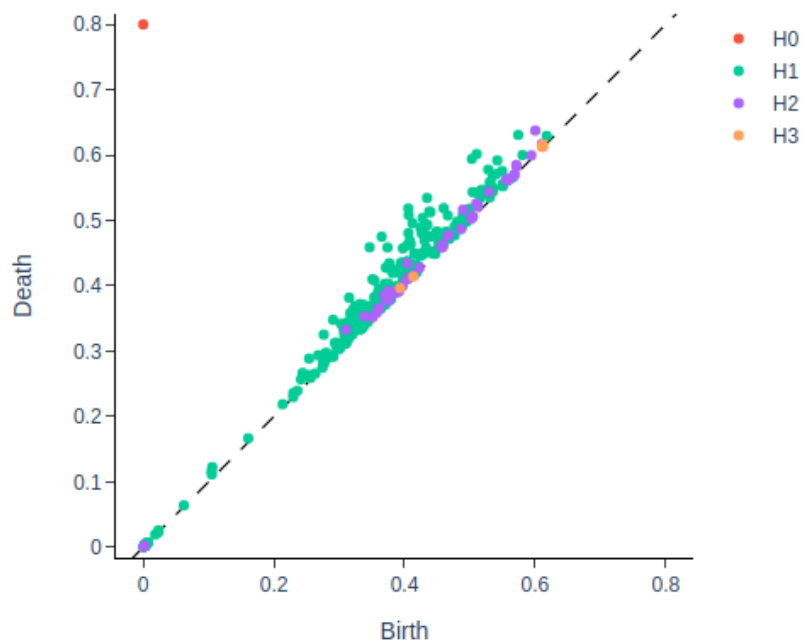
$$\beta_0(X) = 1$$

O espaço é um círculo

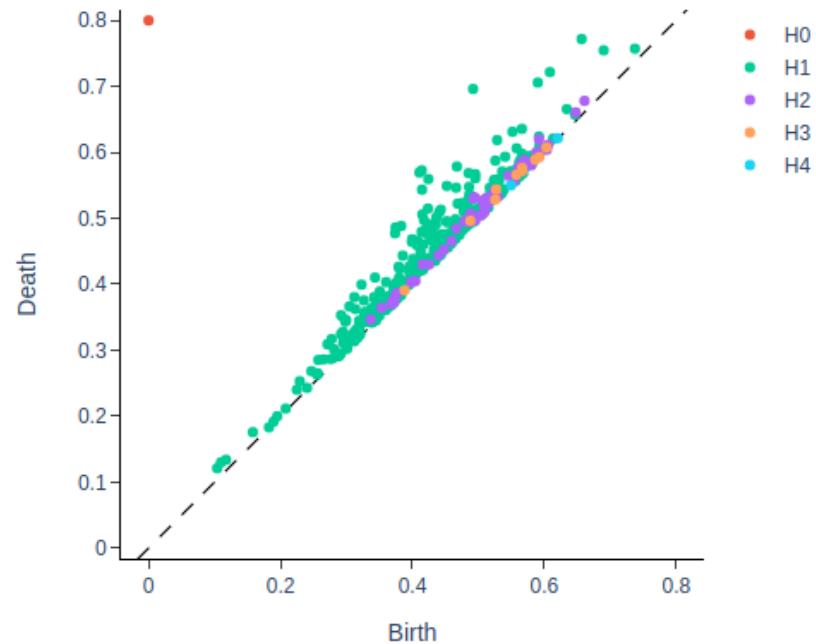
Código de barras do módulo de persistência associado com a filtração de Čech : H_0 em vermelho et H_1 em verde.



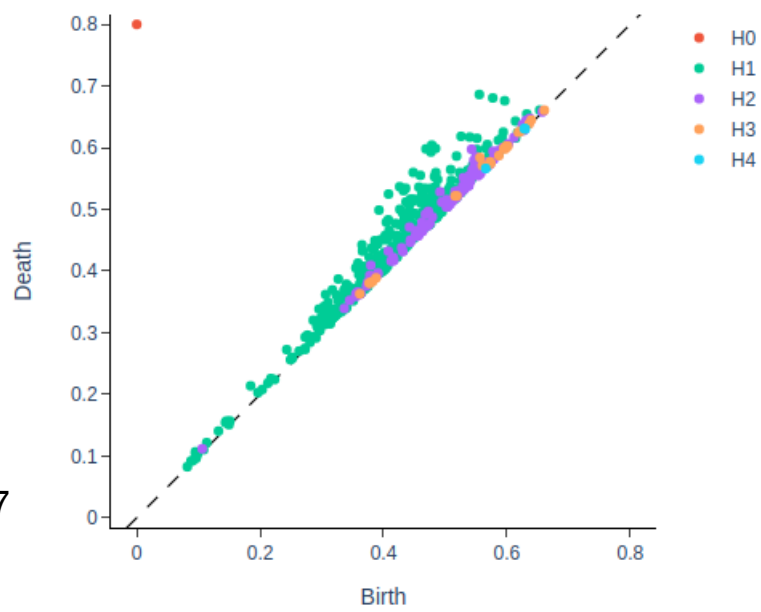
SV 3



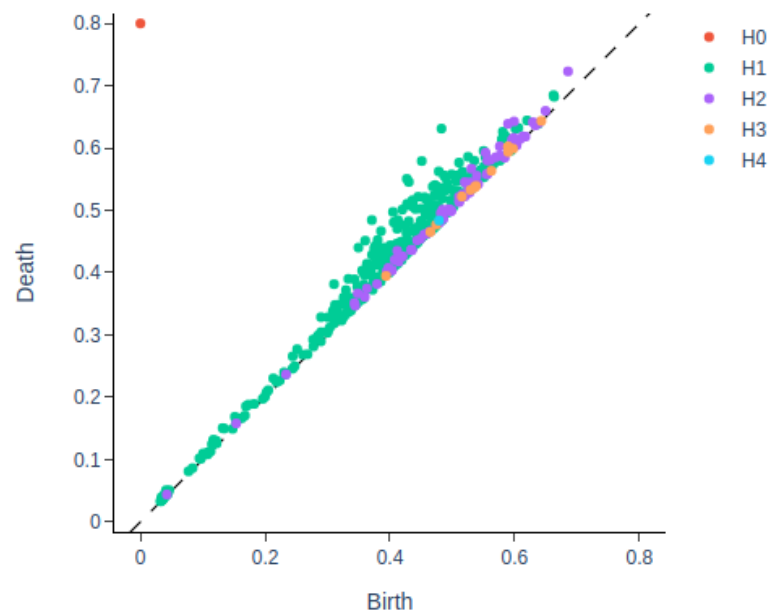
SV 11



SV 17



SV 26

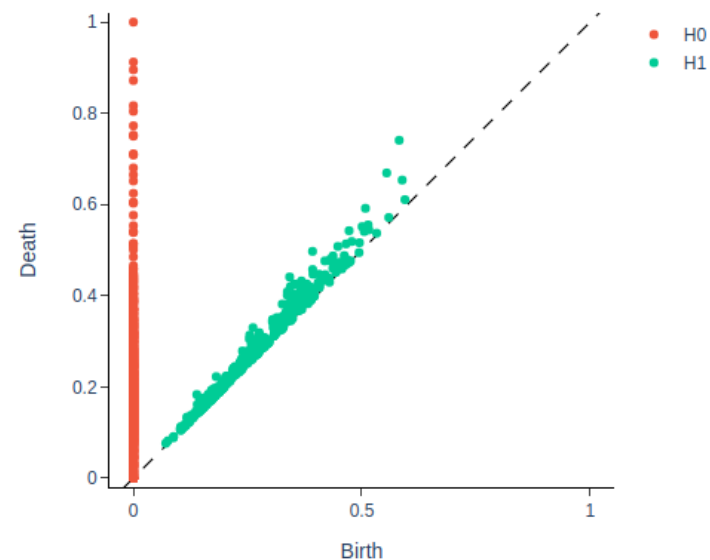
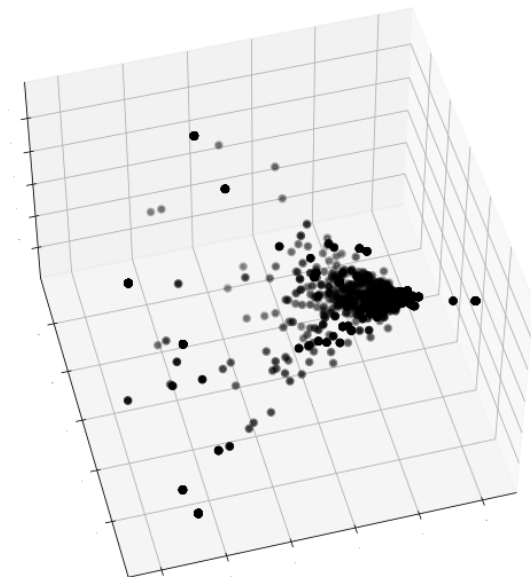


São extraídas somente as frases conjugadas na primeira pessoa do singular.

'Sem prejuízo de novo exame por ocasião do julgamento de mérito, considero ausentes os requisitos que ensejariam a concessão da antecipação de tutela pleiteada.'

'O que faço para fins de adequação do acórdão recorrido aos termos fixados pela Súmula Vinculante 20.'

'Portanto, eu sempre decidi na mesma linha do Ministro Cezar Peluso, no sentido de não conhecer, de considerar a matéria inconstitucional, mas se o Plenário entender que, dada a importância, o impacto da questão com relação à Administração, então talvez convenha que nós ultrapássemos essa questão do conhecimento e adentremos no âmago do tema.'



Para cada SV, selecionamos algumas palavras-chave pertinentes.

SV 3: ['contas', 'anterior', 'tcu', 'defeso', 'pensão', 'universidade', 'contraditório', 'união', 'amplo', 'atar', 'paradigma']

SV 4: ['fepasa', 'salário', 'insalubridade', 'mínimo', 'adicional', 'ipatinga', 'cálculo']

SV 10: ['inconstitucionalidade', 'reservar', 'afastar', 'adc', 'declarar', 'plenário', 'lei', 'art', 'cláusula', 'fracionário']

SV 11: ['réu', 'júri', 'eletrônico', 'agente', 'audiência', 'justificar', 'prisão', 'algemar', 'nulidade']

SV 14: ['acesso', 'defeso', 'interceptação', 'ministério', 'colaboração', 'eleitoral', 'inquérito', 'advogar', 'criminal', 'polícia', 'investigação', 'penal', 'procedimento', 'cópia']

SV 17: ['período', 'juros', 'juro', 'precatório', 'moratório', 'morar']

SV 20: ['gdass', 'turma', 'gdatem', 'ativos', 'gdata', 'utilização', 'individual', 'avaliação', 'desempenho', 'prêmio', 'inativos', 'recursal', 'gratificação']

SV 26: ['corte', 'processual', 'departamento', 'regime', 'hc', 'execuções', 'criminais', 'criminológico', 'realização', 'progressão', 'interpretação', 'reclamação']

SV 33: ['previdência', 'especial', 'injunção', 'mi', 'averbação', 'aposentadoria']

SV 37: ['acre', 'aumentar', 'estadual', 'judiciário', 'vencimento', 'leis', 'equiparação', 'reajuste', 'isonomia', 'rondônia', 'função', 'maranhão', 'remuneratório', 'carreira', 'remuneração', 'salarial', 'amapá', 'magistratura', 'anual', 'magistrado', 'proposto', 'municipal', 'legislação']

Para cada SV, selecionamos algumas palavras-chave pertinentes.

Súmula vinculante 33

Aplicam-se ao servidor público, no que couber, as regras do regime geral da **previdência** social sobre **aposentadoria especial** de que trata o artigo 40, § 4º, inciso III da Constituição Federal, até a edição de lei complementar específica.



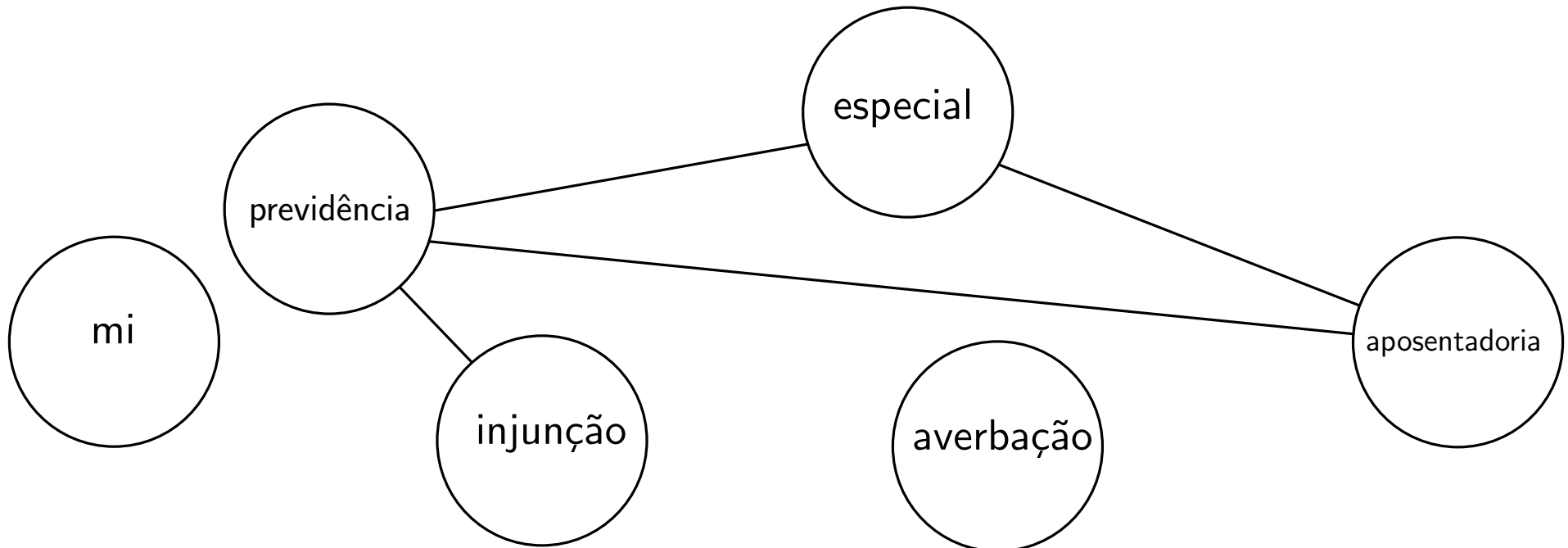
SV 33: ['previdência', 'especial', 'injunção', 'mi', 'averbação', 'aposentadoria']

SV 37: ['acre', 'aumentar', 'estadual', 'judiciário', 'vencimento', 'leis', 'equiparação', 'reajuste', 'isonomia', 'rondônia', 'função', 'maranhão', 'remuneratório', 'carreira', 'remuneração', 'salarial', 'amapá', 'magistratura', 'anual', 'magistrado', 'proposto', 'municipal', 'legislação']

A fim de avaliar a complexidade sintáctica de uma SV, construímos um grafo

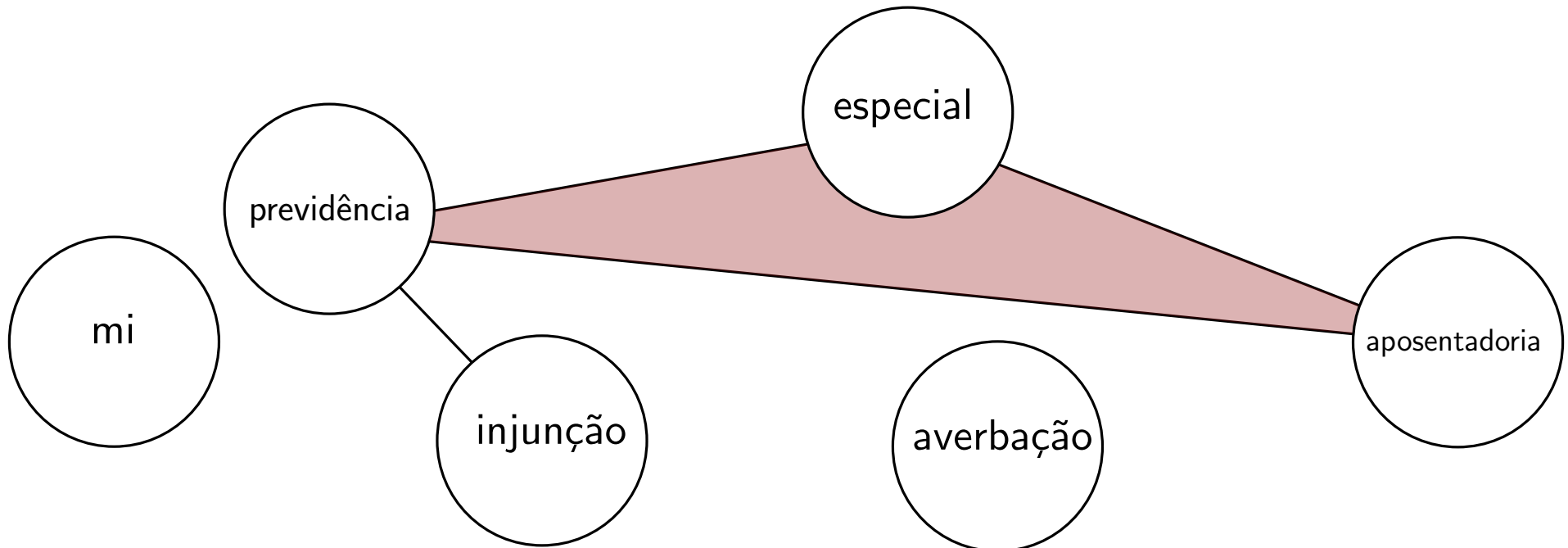
- cujos nós são as palavras-chave selecionadas
- e onde temos uma aresta $[word_1, word_2]$ se existir um documento que contenha ambas as palavras.

SV 33: ['previdência', 'especial', 'injunção', 'mi', 'averbação', 'aposentadoria']



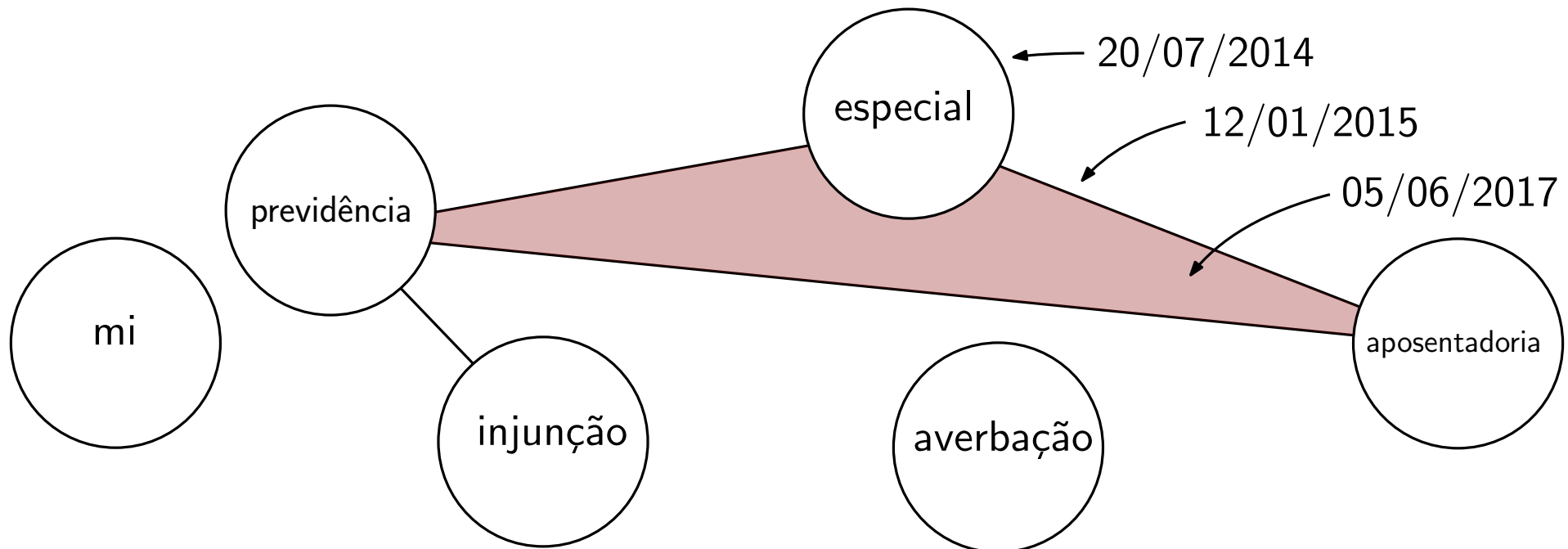
- A fim de avaliar a complexidade sintáctica de uma SV, construímos um ~~grafo~~ **complexo simplicial**
- cujos nós são as palavras-chave selecionadas
 - e onde temos uma aresta $[word_1, word_2]$ se existir um documento que contenha ambas as palavras.
 - e onde temos um triângulo $[word_1, word_2, word_3]$ se houver um documento que contenha as três palavras

SV 33: ['previdência', 'especial', 'injunção', 'mi', 'averbação', 'aposentadoria']



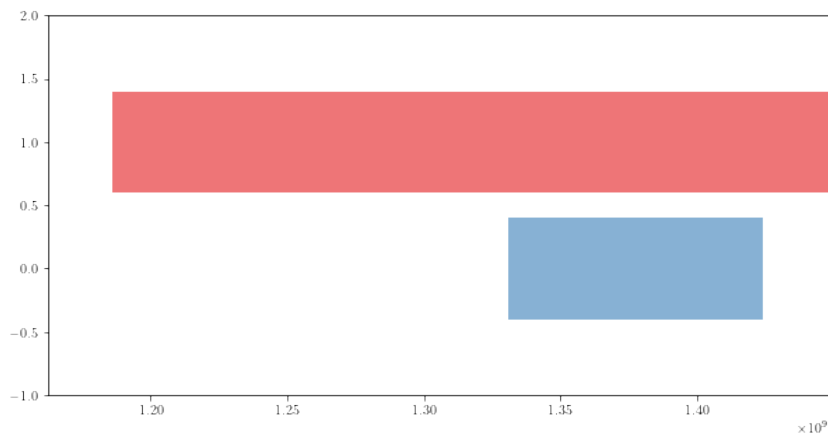
- A fim de avaliar a complexidade sintáctica de uma SV, construímos um ~~grafo~~ **complexo simplicial**
- cujos nós são as palavras-chave selecionadas
 - e onde temos uma aresta $[word_1, word_2]$ se existir um documento que contenha ambas as palavras.
 - e onde temos um triângulo $[word_1, word_2, word_3]$ se houver um documento que contenha as três palavras
 - também dotamos as simplices com um **valor de filtragem**: a primeira vez que elas aparecem

SV 33: ['previdência', 'especial', 'injunção', 'mi', 'averbação', 'aposentadoria']

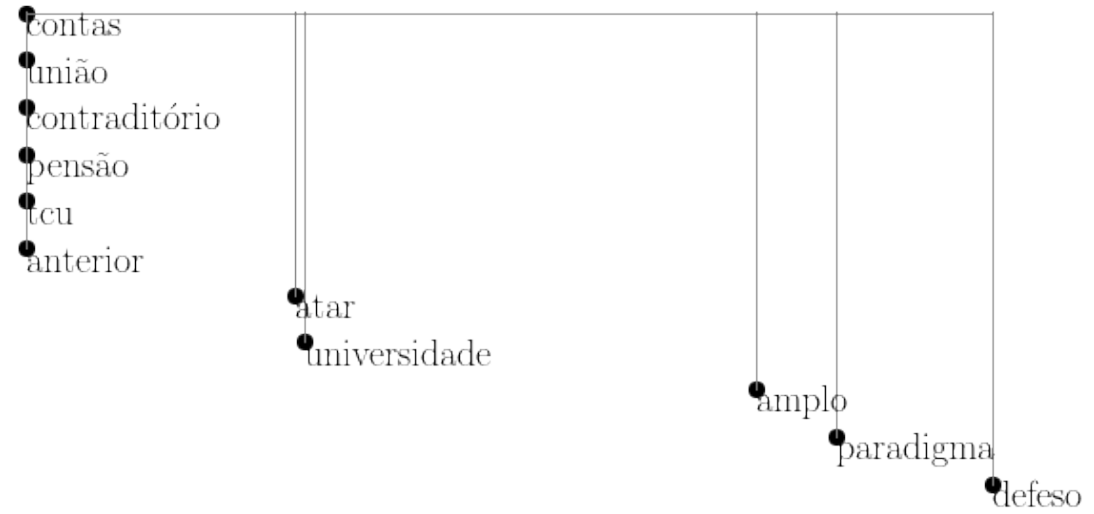


SV 3 : 'contas', 'anterior', 'tcu', 'defeso', 'pensão', 'universidade', 'contraditório', 'união', 'amplo', 'atar', 'paradigma'

Persistence diagram



Merge tree

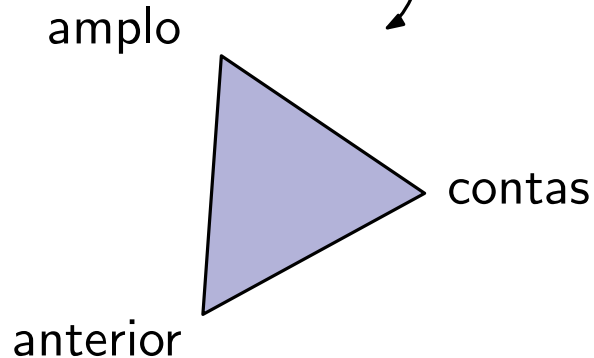
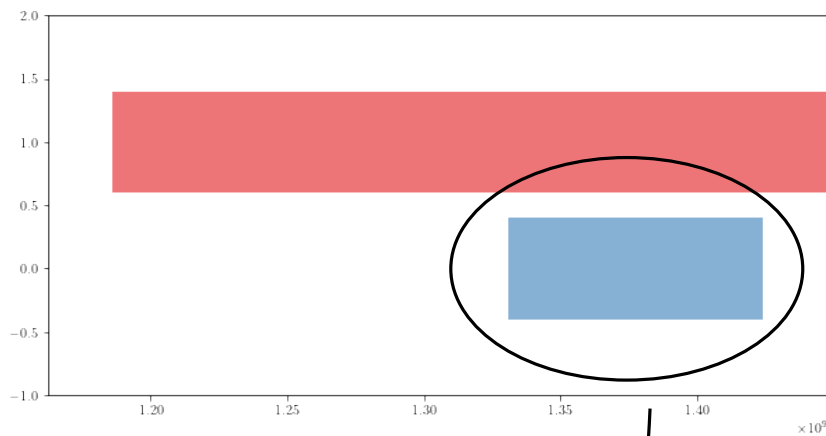


Exemplos

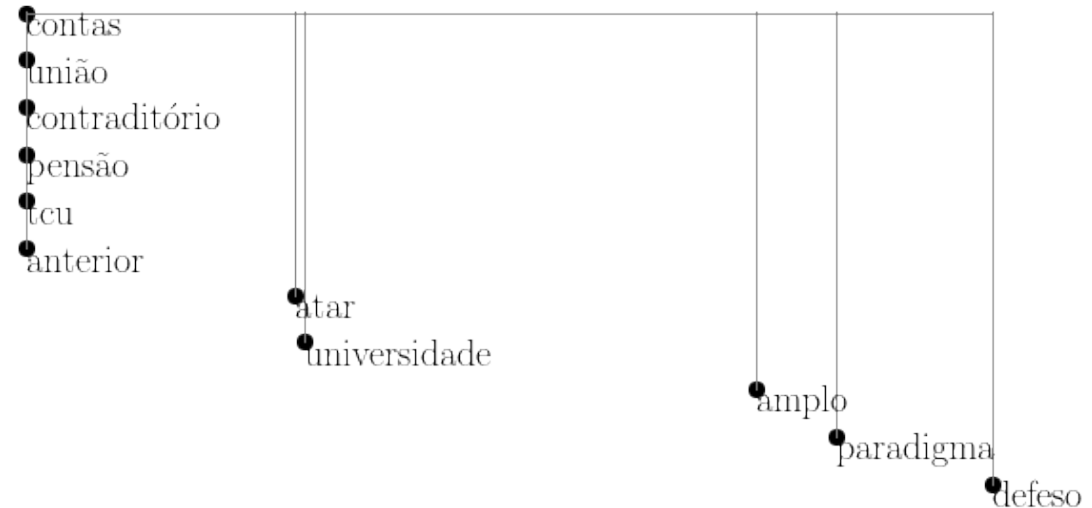
12/17 (2/3)

SV 3 : 'contas', 'anterior', 'tcu', 'defeso', 'pensão', 'universidade', 'contraditório', 'união', 'amplo', 'atar', 'paradigma'

Persistence diagram



Merge tree



Filtration values

(anterior): 1186005600.0
(contas): 1186005600.0
(anterior, contas): 1186005600.0
(amplo): 1228950000.0
(amplo, contas): 1228950000.0
(amplo, anterior): 1330902000.0
(amplo, anterior, contas): 1424300400.0

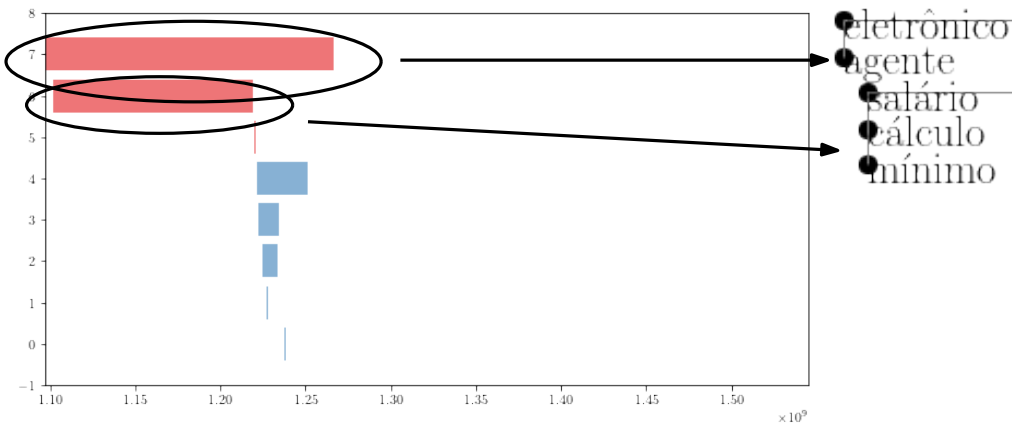
Exemplos

12/17 (3/3)

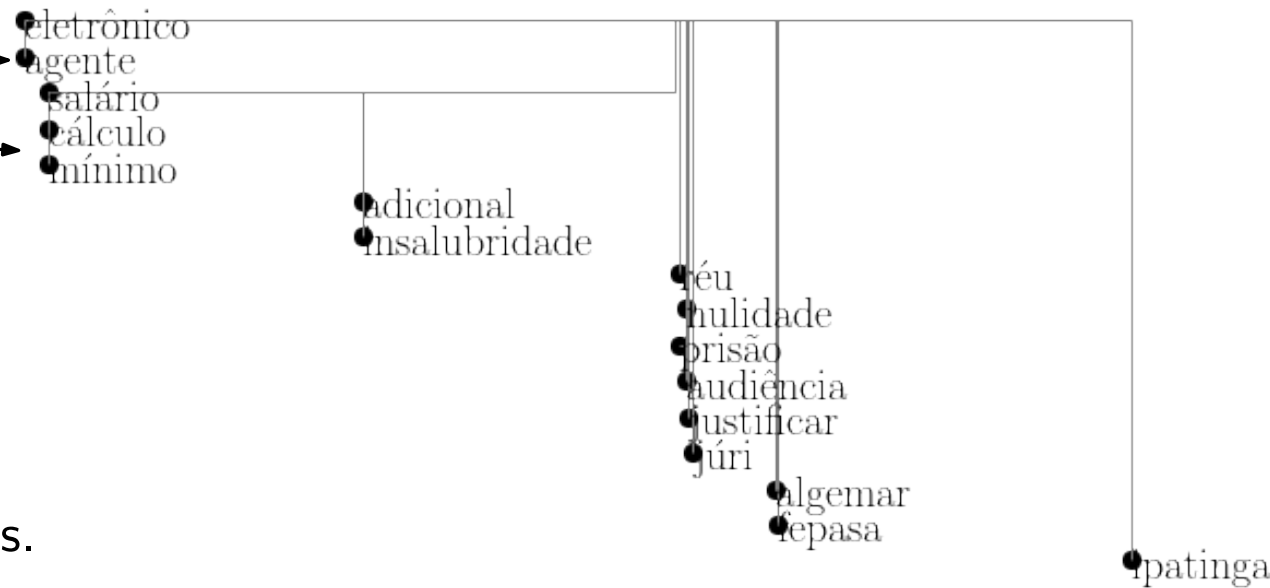
SV 4: 'fepasa', 'salário', 'insalubridade', 'mínimo', 'adicional', 'ipatinga', 'cálculo'

BP 11: 'réu', 'júri', 'eletrônico', 'agente', 'audiência', 'justificar', 'prisão', 'algemar', 'nulidade'

Persistence diagram



Merge tree



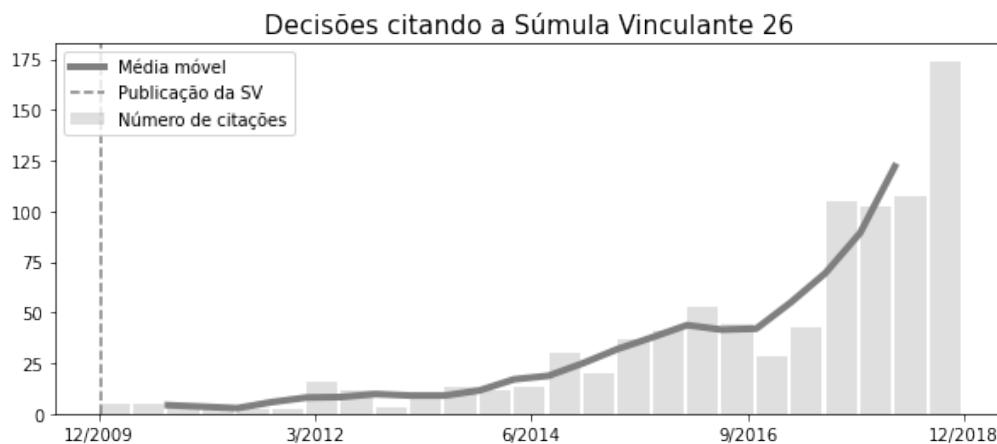
We see two connected components.

I - Análise exploratória com o TDA

II - Análise empírica da SV 26

[*O impacto da Súmula Vinculante 26 na diminuição de demanda similar no STF: uma análise quantitativa por modelos de machine learning*, with B. S. Chagas and C. M. Damian]

[*Progressão de regime em crimes hediondos no Supremo Tribunal Federal: uma análise empírica pela Súmula Vinculante 26*, with A. C. M. Jaccoud and Pedro B. de Oliveira]



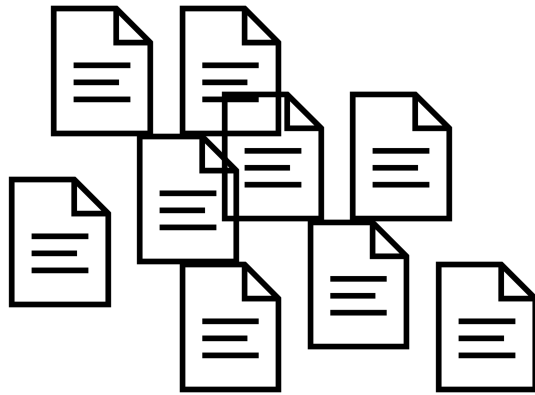
Para efeito de progressão de regime no cumprimento de pena por crime hediondo, ou equiparado, o juízo da execução observará a inconstitucionalidade do art. 2º da Lei n. 8.072, de 25 de julho de 1990, sem prejuízo de avaliar se o condenado preenche, ou não, os requisitos objetivos e subjetivos do benefício, podendo determinar, para tal fim, **de modo fundamentado, a realização de exame criminológico.**

1 - Estimção automática de documentos: detectar documentos, antes de 2009, que tratam do assunto da SV 26.

2 - Análise jurídica manual: Ler os documentos que citam a SV 26, e entender porque há um aumento.

Estimação automática de documentos 15/17 (1/2)

Problema: Quais são as decisões a que a SV 26 poderia ter se aplicado?

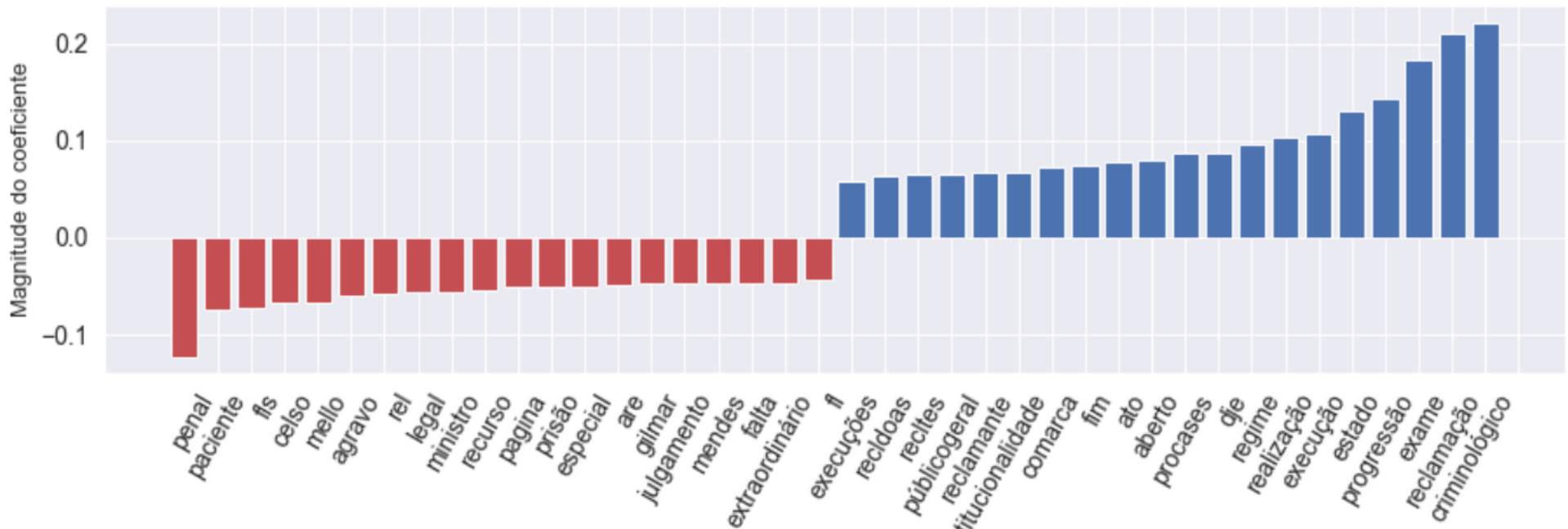


Decisões do STF



SV 26

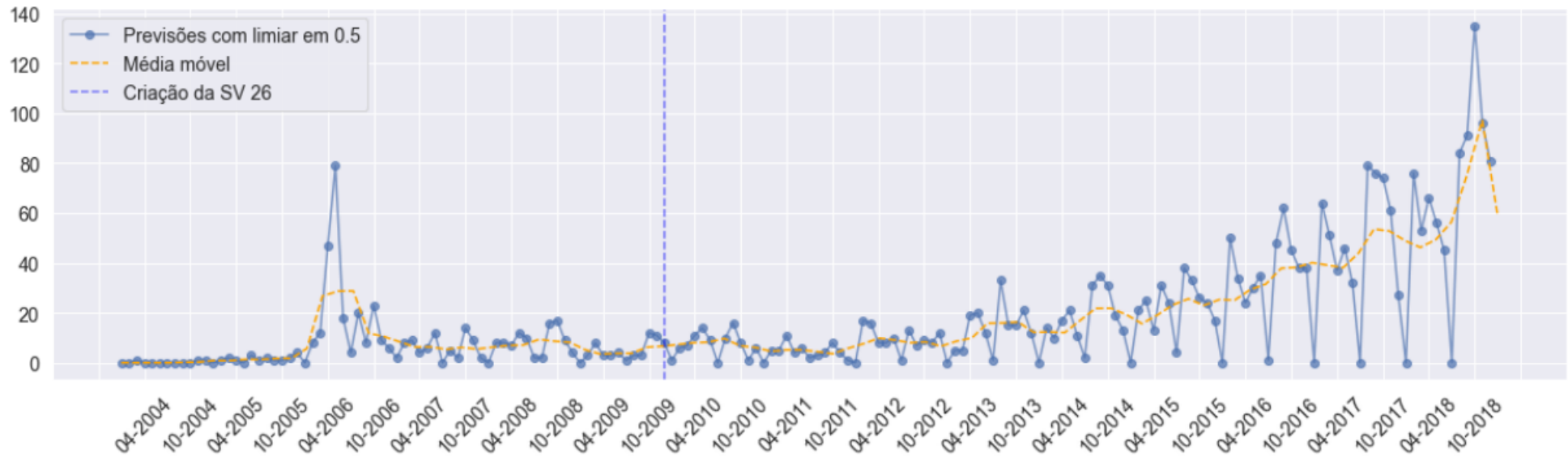
Imersão TF-IDF (in $\mathbb{R}^{11'573}$) + regressão logística (erro de validação cruzada: 0.98).



Estimação automática de documentos ^{15/17} (2/2)

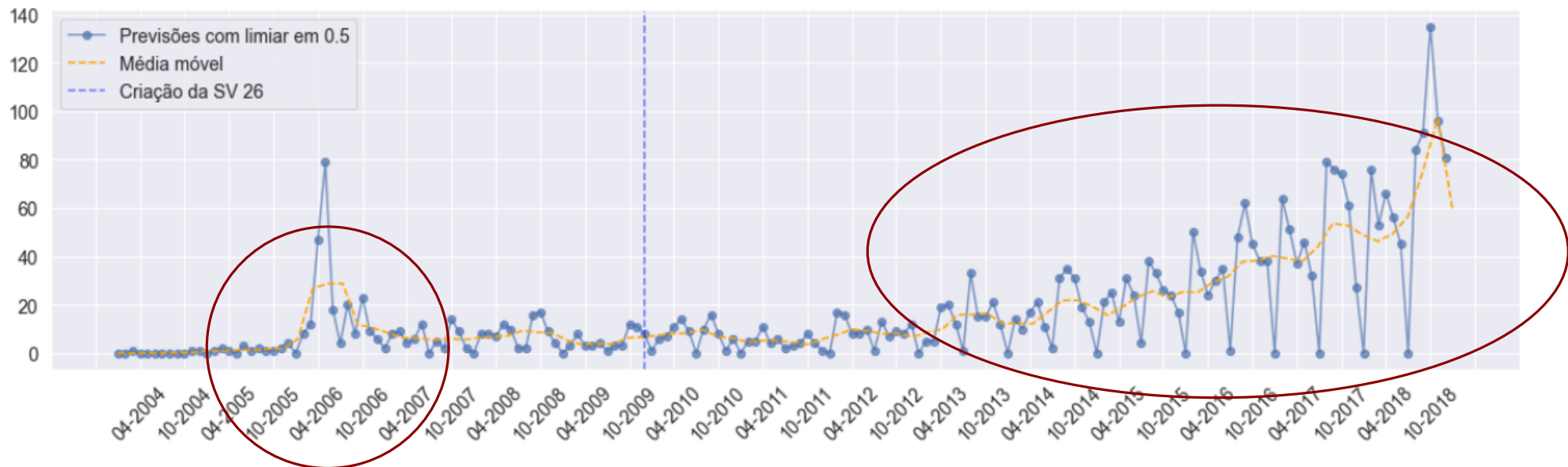
Resultado:

Número absoluto de documentos detectados por mês



Resultado:

Número absoluto de documentos detectados por mês

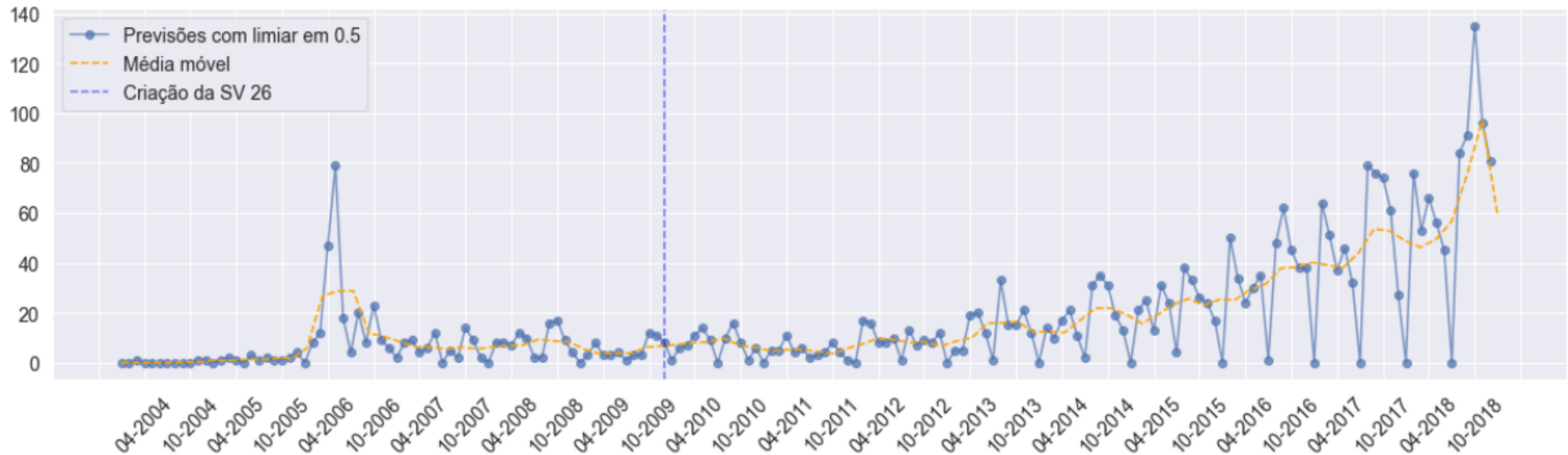


demandas sobre a possibilidade de progressão do regime

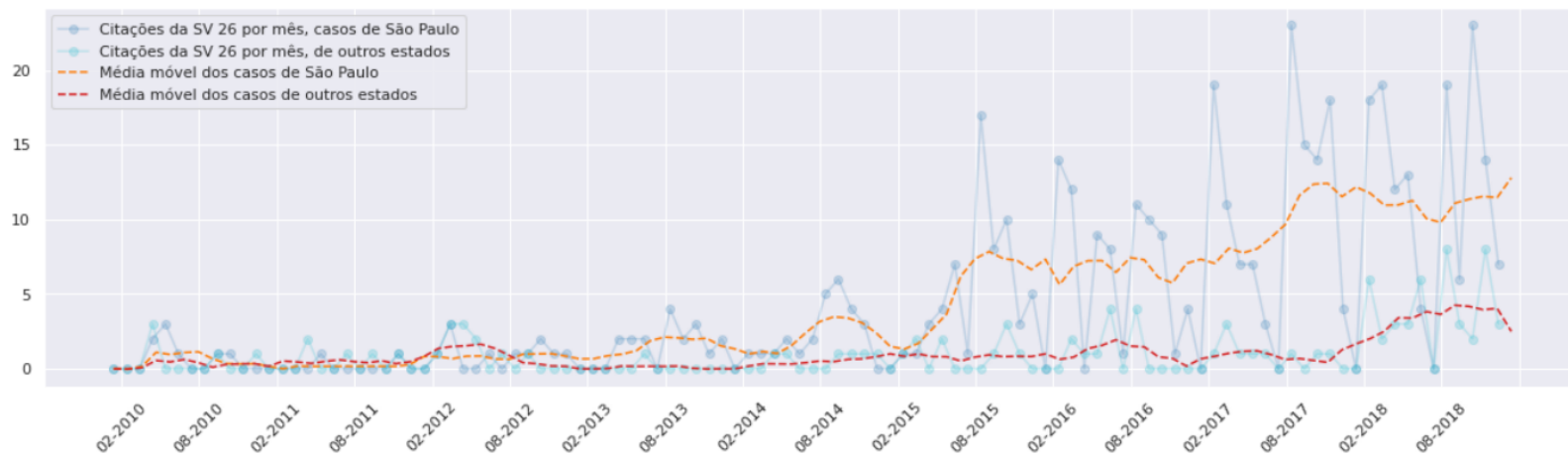
demandas sobre a fundamentação do pedido de exame criminológico

Resultado:

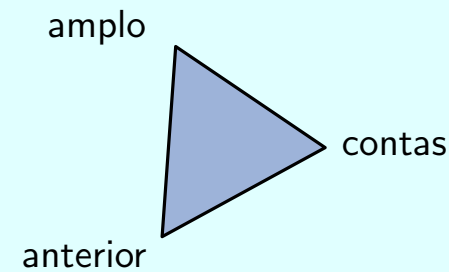
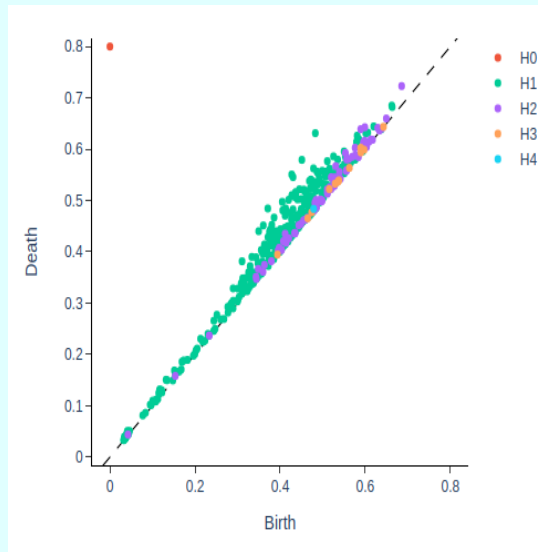
Número absoluto de documentos detectados por mês



- A súmula teve um impacto inicial na diminuição de reclamações.
- Depois, grande quantidade de reclamações a respeito da fundamentação do exame criminológico, provavelmente impulsionada pela Defensoria Pública de São Paulo.



Primeiras construções de métodos de análise topológica para documentos legais.



Análise da eficácia das SVs com TF-IDF + regressão linear.

