

Análise de dados de produtividade de Órgão do Judiciário utilizando aprendizado de máquina não supervisionado

Caique Ferreira Borges¹, Marcelo Martins Pinto¹

¹Instituto Metr pole Digital (IMD)
Universidade Federal do Rio Grande do Norte (UFRN)
CEP: 59078-900 – Natal – RN – Brazil

c.ferreira321@gmail.com, pintomarc@gmail.com

Abstract. *This article presents an initial study on the productivity of civil servants of a brazilian judiciary body, aiming to define an average productivity for civil servants, for later measurement of the activities carried out by those who work in a telecommuting regime. To this end, we carried out a study based on data collected from January 2019 to December 2022, with documents attached to the processes, proposing the use of unsupervised machine learning techniques to create clusters indicative of productivity, as presented.*

Resumo. *Este artigo apresenta estudo inicial sobre produtividade dos servidores de um  rg o do judici rio brasileiro visando a defini  o de uma produtividade m dia para os servidores, para posterior medi  o das atividades realizadas pelos que laboram em regime de teletrabalho. Para tal, fizemos um estudo baseado em dados coletados no per odo de janeiro de 2019 a dezembro de 2022, com os documentos juntados aos processos, propondo a utiliza  o de t cnicas de aprendizado de m quina n o supervisionado para cria  o dos agrupamentos indicativos de produtividade, conforme apresentamos.*

1. Introdu  o

O Judici rio brasileiro, notadamente ap s a implanta  o do Sistema de Processo Judicial Eletr nico (PJe), vem procurando aprimorar a **gest o de pessoas**, focando na motiva  o, comprometimento, clima organizacional e qualidade de vida dos servidores. Em raz o disto, o Conselho Nacional de Justi a (CNJ) vem aprimorando cada vez mais a legisla  o para viabilizar o regime de teletrabalho aos servidores, e os demais  rg os do judici rio v m seguindo e aprimorando cada vez mais essa legisla  o.

Uma boa parte dos  rg os do judici rio adotou a pr tica de definir que os serventu rios em teletrabalho tem como meta produzir mais do que aqueles que laboram presencialmente, definindo metas percentuais superiores ou apenas “*prodizir mais*”.

Ressalte-se que n o era pr tica medir o volume de trabalho realizado, bem como indicar se o servidor   produtivo ou n o, o que se tornou realidade com a op  o pela modalidade de teletrabalho, principalmente ampliada em raz o da Pandemia da COVID-19, deflagrada no Brasil em meados de mar o de 2020.

As estat sticas de produtividade do judici rio, na mesma linha, passaram a utilizar a Base Nacional de Dados do Poder Judici rio (DATAJUD). Contudo, os dados estat sticos s o colhidos at  o n vel de magistrado, sem contabilizar a produtividade dos servidores em geral.

Assim, entendemos que é necessária a realização do presente estudo, visando obter algumas possíveis respostas para a questão referente à jurimetria relativa aos servidores do judiciário, não levantada até então pelo CNJ.

Este trabalho visa, com base nos dados de produtividade coletados, analisar, tratar e, utilizando as técnicas de Aprendizagem Não Supervisionada, identificar o melhor algoritmo a ser utilizado para classificar e identificar os possíveis agrupamentos, de forma a possibilitar rotulagem e classificação do conjunto de dados utilizado, visando identificar o nível de produtividade de cada grupo.

2. Conceitos relacionados

O conjunto de dados utilizado neste estudo refere-se às atividades realizadas pelos servidores na área jurídica de um órgão da justiça no Rio Grande do Norte, durante o período de janeiro de 2019 a dezembro de 2022. Os seguintes critérios foram considerados na seleção dos dados: chave de identificação do serventuário, data da operação, a indicação se a atividade ocorreu durante a pandemia, o nome do órgão do processo, cidade da unidade de lotação do serventuário, quantidade diárias de expedientes criados e anexados aos processos para cada tipo de documento, sexo do serventuário, unidade de lotação do serventuário, papel desempenhado pelo serventuário atua na unidade, a idade do serventuário e o tempo de serviço do serventuário.

```
RangeIndex: 143187 entries, 0 to 143186
Data columns (total 23 columns):
#   Column              Non-Null Count  Dtype
---  -
0   id_chave             143187 non-null  int64
1   dt_doc               143187 non-null  object
2   fl_pandemia          143187 non-null  object
3   ds_orgao             143187 non-null  object
4   ds_cidade            143187 non-null  object
5   despachos            143187 non-null  int64
6   decisao              143187 non-null  int64
7   sentenca             143187 non-null  int64
8   intimacao            143187 non-null  int64
9   mandado              143187 non-null  int64
10  notificacao          143187 non-null  int64
11  certidao             143187 non-null  int64
12  ataudaudiencia       143187 non-null  int64
13  docdiverso           143187 non-null  int64
14  alvara               143187 non-null  int64
15  calculos             143187 non-null  int64
16  peticaopdf            143187 non-null  int64
17  cartaprecatoria      143187 non-null  int64
18  fl_sexo              143187 non-null  object
19  un_lotacao           143187 non-null  object
20  ds_cargo              121738 non-null  object
21  nu_idade_dias         143187 non-null  int64
22  nu_tempo_svc_dias     143187 non-null  int64
```

Figura 1. Estrutura dos dados disponibilizados

Jurimetria - é a quantificação das atividade por meio da análise de dados gerados e, a partir dela, determinar padrões;

Dados categóricos - são dados qualitativos, em formato não numérico, que precisam ser tratados, representados na figura 1, como os tipos

Normalização de dados - é um processo importante e tem como objetivo transformar os dados para uma escala comum, a fim de melhorar o desempenho dos algoritmos

de aprendizado de máquina e garantir que as características tenham impacto equilibrado na análise. No trabalho em tela, foi utilizada a normalização Min-Max, em que os dados são reescalados para um intervalo específico, entre 0 e 1.

3. Trabalhos relacionados

O Conselho Nacional de Justiça (CNJ) e os demais órgãos do judiciário vem buscando cada vez mais aprimorar as pesquisas para melhoria da produtividade no judiciário, entretanto, praticamente todos esses estudos se concentram em analisar o que foi produzido e identificar as lacunas a serem preenchidas para o órgão atingir as metas traçadas. Não há estudos que granularizam a produtividade até o nível dos serventuários em geral, sendo a menor unidade de medição, os dados relacionados à produtividade dos magistrados, que não são foco do presente estudo.

Dessa forma, não foram localizados outros trabalhos com o foco proposto na presente publicação.

4. Metodologia dos experimentos

Para facilitar o estudo, dividimos o trabalho em quatro etapas: Tratamento dos dados, Normalização dos dados, Aplicação das técnicas de aprendizado não supervisionado e Análise dos dados obtidos.

4.1. Tratamento dos dados

Esta etapa visa a correção de dados quanto à existência de eventuais imperfeições e inadequações que podem afetar os resultados obtidos.

Assim, tão logo identificado a existência de dados em branco, foi realizada a análise para saber se as linhas poderiam ser descartadas ou corrigidas, sendo atribuídos os valores cabíveis, por exemplo: Atribuir o valor ASSISTENTE ao atributo ds_cargo, onde seu valor fosse.

Utilizando a metodologia aprendida durante o curso e, considerando o volume de dados inicialmente existente, foi realizada a redução de dados por meio do agrupamento das informações mais relevantes em apenas alguns elementos chave, tais como: id_chave, fl_pandemia, ds_orgao, ano_mes e ds_cargo.

Além disso, foram excluídas colunas que possuíam dados duplicados, assim como aquelas que não tinham impacto na análise dos dados em si, como o código de identificação do servidor que elaborou o expediente. Por outro lado, foram acrescentadas outras colunas mais relevantes para a análise.

4.2. Normalização dos dados

Após a conclusão da etapa de ajustes no dataset, os 143187 registros em 23 colunas foram convertidos em 13068 registros em 21 colunas, contudo, devido a existência de muitas colunas com valores categóricos, foi necessária a transformação dos dados, de forma a converter os valores nominais em numéricos e, em seguida, normalizar todos esses valores numéricos, de forma a eliminar os ruídos, deixando-os aptos para mineração.

Após esta etapa, os dados ficaram distribuídos em 13068 registros, com 62 colunas, conforme disposto na figura 2.

RangeIndex: 13068 entries, 0 to 13067
Data columns (total 62 columns):

#	Column	Non-Null Count	Dtype
0	despachos	13068 non-null	float64
1	decisao	13068 non-null	float64
2	sentenca	13068 non-null	float64
3	intimacao	13068 non-null	float64
4	mandado	13068 non-null	float64
5	notificacao	13068 non-null	float64
6	certidao	13068 non-null	float64
7	ataaudiencia	13068 non-null	float64
8	docdiverso	13068 non-null	float64
9	alvara	13068 non-null	float64
10	calculos	13068 non-null	float64
11	peticaopdf	13068 non-null	float64
12	cartaprecatoria	13068 non-null	float64
13	nu_idade_dias	13068 non-null	float64
14	nu_tempo_svc_dias	13068 non-null	float64
15	nu_dias_periodo	13068 non-null	float64
16	fl_pandemia_N	13068 non-null	uint8
17	fl_pandemia_S	13068 non-null	uint8
18	ds_orgao_Cejusc-MOS	13068 non-null	uint8
19	ds_orgao_Cejusc-NAT	13068 non-null	uint8
20	ds_orgao_DINT	13068 non-null	uint8
21	ds_orgao_DIPRE	13068 non-null	uint8
22	ds_orgao_PAAZN	13068 non-null	uint8
23	ds_orgao_VTASSU	13068 non-null	uint8
24	ds_orgao_VTCAICO	13068 non-null	uint8
25	ds_orgao_VTCMIRIM	13068 non-null	uint8
26	ds_orgao_VTCNOVOS	13068 non-null	uint8
27	ds_orgao_VTGOIAN	13068 non-null	uint8
28	ds_orgao_VTM1	13068 non-null	uint8
29	ds_orgao_VTM2	13068 non-null	uint8
30	ds_orgao_VTM3	13068 non-null	uint8
31	ds_orgao_VTM4	13068 non-null	uint8
32	ds_orgao_VTMAU2	13068 non-null	uint8
33	ds_orgao_VTN1	13068 non-null	uint8
34	ds_orgao_VTN10	13068 non-null	uint8
35	ds_orgao_VTN11	13068 non-null	uint8
36	ds_orgao_VTN12	13068 non-null	uint8
37	ds_orgao_VTN13	13068 non-null	uint8
38	ds_orgao_VTN2	13068 non-null	uint8
39	ds_orgao_VTN3	13068 non-null	uint8
40	ds_orgao_VTN4	13068 non-null	uint8
41	ds_orgao_VTN5	13068 non-null	uint8
42	ds_orgao_VTN6	13068 non-null	uint8
43	ds_orgao_VTN7	13068 non-null	uint8
44	ds_orgao_VTN8	13068 non-null	uint8
45	ds_orgao_VTN9	13068 non-null	uint8
46	ds_orgao_VTPFERROS	13068 non-null	uint8
47	ds_cargo_ANALISTA JUDICIARIO	13068 non-null	uint8
48	ds_cargo_ASSISTENTE	13068 non-null	uint8
49	ds_cargo_ESTAGIARIO NIVEL SUP	13068 non-null	uint8
50	ds_cargo_TECNICO JUDICIARIO	13068 non-null	uint8
51	ds_cidade_Assu	13068 non-null	uint8
52	ds_cidade_Caicó	13068 non-null	uint8
53	ds_cidade_CearA-Mi	13068 non-null	uint8
54	ds_cidade_Currais	13068 non-null	uint8
55	ds_cidade_Goianinh	13068 non-null	uint8
56	ds_cidade_Macau	13068 non-null	uint8
57	ds_cidade_Mossoró	13068 non-null	uint8
58	ds_cidade_Natal	13068 non-null	uint8
59	ds_cidade_Pau dos	13068 non-null	uint8
60	fl_sexo_F	13068 non-null	uint8
61	fl_sexo_M	13068 non-null	uint8

Figura 2. Estrutura dos dados normalizados

4.3. Aplicação das técnicas de aprendizado não supervisionado

A ideia do projeto é utilizar aprendizado de máquina não supervisionado através de modelos descritivos de agrupamento, uma vez que não há classe associada a cada instância

dos dados analisados.

Dessa forma, a aplicação dos modelos possibilita a associação dos objetos por similaridade, onde cada grupo tem elevada homogeneidade interna elevada heterogeneidade externa.

Levando-se em conta os principais algoritmos, consideramos o uso do **kMeans**, com agrupamento baseados em centroide, do **Expectation-Maximization (EM)** com uso de modelos de mistura Gaussiana (GMM), que realiza o agrupamento com base em distribuição; e do **Hierárquico**, que realiza agrupamentos hierárquicos aglomerativos.

Aos algoritmos de agrupamento indicados, foi considerada a aplicação de duas métricas, visando auxiliar a escolha do melhor algoritmo a ser utilizado:

1. **Davies-Bouldin** é uma métrica de avaliação de clustering que mede a qualidade da divisão dos dados em grupos ou clusters. Essa métrica busca encontrar clusters que sejam bem separados uns dos outros e com alta coesão interna. Quanto menor o valor do índice de Davies-Bouldin, melhor é a qualidade do clustering. Valores mais próximos de zero indicam clusters bem separados e coesos.
2. **Silhouette** é uma medida de avaliação de clustering que quantifica a qualidade e a coesão dos clusters encontrados em um conjunto de dados. Essa métrica fornece uma pontuação para cada ponto de dados com base em quão bem ele se encaixa em seu próprio cluster em comparação com outros clusters. Uma pontuação Silhouette alta indica que o ponto de dados está bem colocado em seu próprio cluster, com uma distância média menor para pontos do mesmo cluster em comparação com outros grupos.

Em ambos os casos, a aplicação dos algoritmos e métricas foi feita para geração de 2 a 20 grupos para avaliação e escolha da melhor opção.

A partir da análise dos gráficos comparativos, a formação de Elbow (cotovelo) foi considerada como o ponto onde ocorre uma mudança na taxa de diminuição da soma das distâncias intragrupos ou variância explicada, indicando uma possível estimativa para o número ideal de clusters.

4.3.1. Métrica DB aplicada a kMeans, EM e HA

Para o caso dessa figura 3, observa-se que houve formação do Elbow em DB_EM e DB_HAav, sendo mais evidente a formação do cotovelo no DB_HAav, com 4 grupos, conforme abaixo.

Clusters	DB_kMeans	DB_HAav	DB_EM
2	2,295	2,622	2,337
3	2,096	2,352	2,348
4	2,115	1,869	1,976
5	2,084	2,042	2,06
6	2,07	1,885	1,977
7	1,946	1,833	1,887
8	1,978	1,716	1,935
9	1,828	1,662	2,245
10	1,822	1,64	1,833
11	1,75	1,662	1,761
12	1,785	1,702	1,967
13	1,73	1,679	1,933
14	1,692	1,676	1,756
15	1,682	1,712	1,664
16	1,62	1,676	1,635
17	1,559	1,647	1,873
18	1,534	1,636	1,747
19	1,528	1,595	1,916
20	1,526	1,578	1,705

Tabela 1. Valores para kMeans, HAav e EM, com Davis Bouldin (DB).

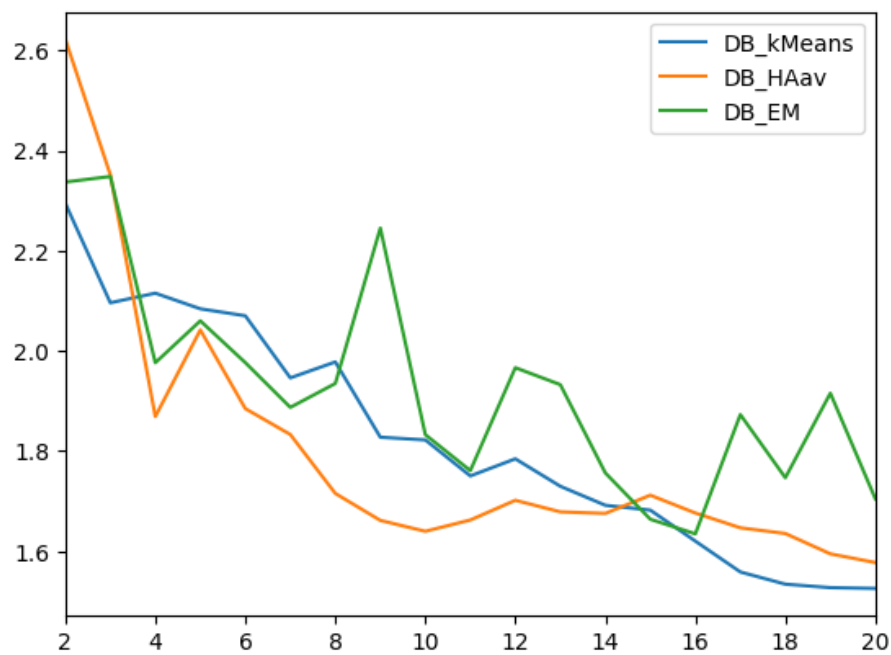


Figura 3. Comparativo usando a métrica DB

4.3.2. Métrica Silhouette aplicada a kMeans, EM e HA

Neste caso - figura 4, verifica-se que há formação do cotovelo em SIL_kMeans e SIL_EM, sendo mais evidente no SIL_kMeans, com 4 grupos, pois há uma maior suavização após o cotovelo, de acordo como abaixo.

Clusters	Sil_kMeans	Sil_HAav	Sil_EM
2	0,156	0,088	0,053
3	0,146	0,102	0,146
4	0,174	0,109	0,07
5	0,154	0,106	0,098
6	0,164	0,115	0,143
7	0,162	0,12	0,117
8	0,17	0,137	0,152
9	0,178	0,143	0,171
10	0,185	0,154	0,16
11	0,197	0,16	0,185
12	0,203	0,168	0,17
13	0,209	0,168	0,164
14	0,215	0,171	0,19
15	0,226	0,18	0,212
16	0,228	0,187	0,202
17	0,229	0,186	0,224
18	0,234	0,188	0,226
19	0,239	0,188	0,199
20	0,246	0,19	0,211

Tabela 2. Valores para kMeans, HAav e EM, com Silhouette.

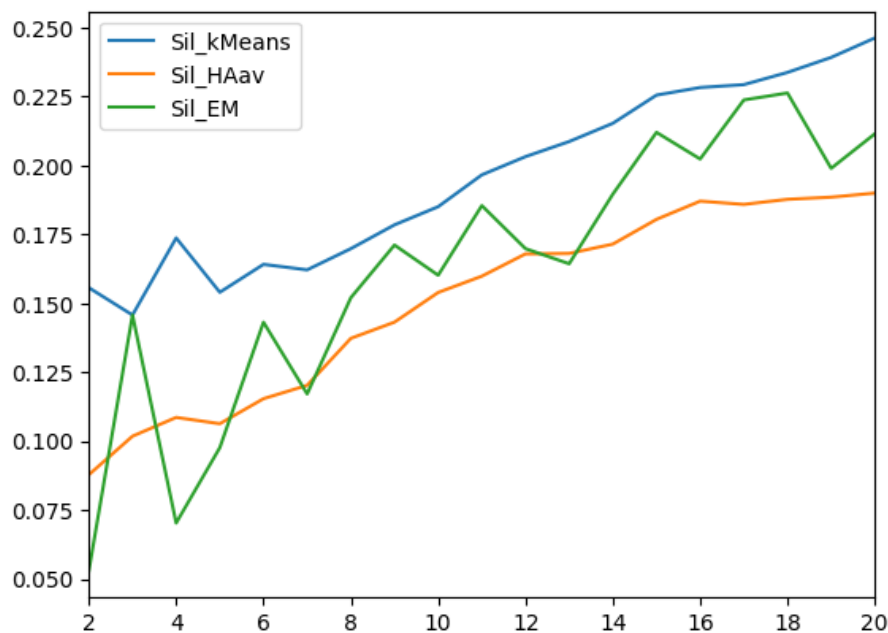


Figura 4. Comparativo usando a métrica Silhouette

4.4. Análise dos dados

Com base na análise dos gráficos e verificando possível variação positiva e negativa do quantitativo de grupos, foram selecionados DB_HAav e Sil_kMeans para realizar o agrupamento dos dados. Foram gerados de 2 a 6 grupos utilizando esses algoritmos, permitindo explorar diferentes configurações e identificar a estrutura dos dados de forma mais precisa.

Foi aplicado ao conjunto de dados não normalizados os agrupamentos indicados e, a partir deles, foi elaborado um conjunto de *dashboards* utilizando o Microsoft Power BI, de forma a facilitar a visualização dos resultados e auxiliar na escolha do melhor método e algoritmo a ser utilizado para o agrupamento desejado. *Clique aqui para ter acesso aos dashboards.*

Feita a análise dos painéis, constatou-se que ainda havia um certo nível de ruído nos dados, pois as classes que tinham o atributo `ds_cargo` com valor "Estagiário" não faziam todas as atividades possíveis, além disso o atributo `ds_orgao` não agregava nenhum valor às informações de agrupamento. Com isso, foi rodado o pipeline completo para ajustar dados, excluindo-se os estagiários e suprimindo as informações dos órgãos no dataset. Dessa forma os dados normalizados figuraram com 10689 registros e 31 colunas e como resultado, os gráficos abaixo foram gerados para fins de comparação.

1. Segundo comparativo de kMeans, HAav e EM, utilizando DB (figura 5)

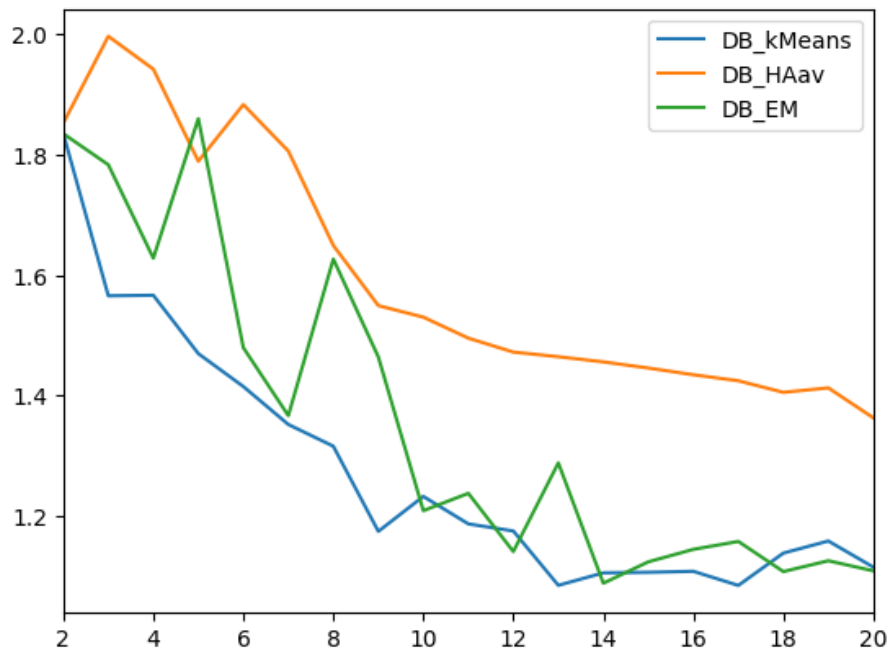


Figura 5. Segunda comparação usando a métrica DB

2. Segundo comparativo de kMeans, HAav e EM, utilizando Silhoutte (figura 6)

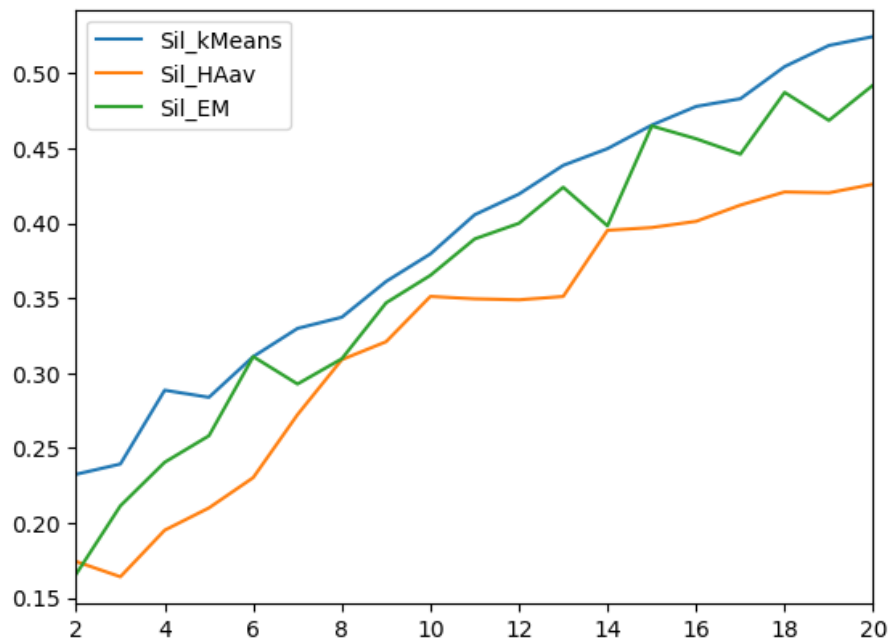


Figura 6. Segunda comparação usando a métrica Silhouette

Dessa 2ª execução do pipeline verificamos que as melhores opções selecionadas foram DB_EM e Sil.kMeans, ambos com 4 grupos. Da mesma forma que no pipeline anterior, também foram gerados de 2 a 6 grupos utilizando esses algoritmos e populados os datasets com a informação dos clusters. Os dados obtidos foram coletados e estão apresentados no mesmo painel citado acima.

5. Resultados

Considerando a questão de produtividades estar muito associada ao conjunto de tarefas realizadas, que por sua vez estão transversalmente vinculadas aos cargos que executam essas atividades, aguardávamos que a distribuição dos grupos contemplasse em cada um, membros de cargos distintos e com diferença significativa entre as quantidades de observações, que seriam a indicação de grupos de melhor ou pior desempenho, em termos de produtividade.

Observamos na análise dos dashboards uma grande disparidade entre os tamanhos dos grupos e a composição deles, que continham algumas vezes só uma parte dos cargos ou outros. Houve a formação de grupos que não possuem determinados cargos associados, com quantidade muito grande em relação aos demais grupos identificados.

De todos os painéis analisados, o que entendemos ter chamado mais a atenção foi o gerado a partir do **kMeans com 4 grupos**, pois apresenta uma distribuição de membros uniforme, com quantidade de elementos equilibrada, contendo servidores de todos os cargos e diferenças no volume de trabalho proporcionais.

Assim, pudemos abstrair a possibilidade de mapear os grupos segundo a seguinte critério:

1. muito produtivo - Grupos 2
2. produtivo - Grupo 1
3. pouco produtivo - Grupo 4
4. improdutivo - Grupo 3

6. Conclusão

Foi visto que a aplicação das técnicas de aprendizado de máquina não supervisionado obteve resultados satisfatórios quanto ao conjunto de dados analisado.

Considerando que a análise realizada se aplicou apenas à área judicial, sem contemplar as atividades administrativas, que também são realizadas pelos tribunais como suporte à área finalística, tais como controle de pessoal, processos de contratação, dentre outras rotinas. Entendemos que o mesmo pipeline pode ser aplicado a um conjunto de dados relacionados às atividades administrativas, obtendo resultados semelhantes.

É válido destacar que, assim como não há dados referentes às atividades administrativas, o conjunto de dados utilizado não fornece informações sobre a complexidade de cada uma das atividades. Portanto, na conclusão, os resultados são considerados como tendo o mesmo peso, independentemente do seu grau de dificuldade.

Para o algoritmo escolhido, os resultados obtidos foram coerentes, apresentando grupos bem definidos, claramente separados e com formas geométricas semelhantes. O conjunto de dados utilizado, juntamente com a abordagem de análise da produtividade mensal, permitiu a identificação de características que possibilitaram a criação de grupos distintos entre si, com variação nos níveis de produtividade, desde muito produtivo, produtivo, pouco produtivo até improdutivo.