



**UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE CIÊNCIAS SOCIAIS
CURSO DE CIÊNCIAS SOCIAIS /
BACHARELADO**

MAYRES LANE PEQUENO DOS SANTOS SILVA

**ÍNDICE H5 DO GOOGLE ACADÊMICO E
WEBQUALIS CAPES NA PRODUÇÃO
DA SOCIOLOGIA BRASILEIRA**

**Maceió/AL
2022**



**UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE CIÊNCIAS SOCIAIS
CURSO DE CIÊNCIAS SOCIAIS /
BACHARELADO**

MAYRES LANE PEQUENO DOS SANTOS SILVA

**ÍNDICE H5 DO GOOGLE ACADÊMICO E
WEBQUALIS CAPES NA PRODUÇÃO
DA SOCIOLOGIA BRASILEIRA**

Trabalho de Conclusão de Curso em forma de artigo científico apresentado como requisito parcial para a obtenção do título de Bacharel em Ciências Sociais pelo Instituto de Ciências Sociais da Universidade Federal de Alagoas, sob orientação do Profa. Dra. Marina Melo.

**Maceió/AL
2022**



UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE CIÊNCIAS SOCIAIS
CURSO DE CIÊNCIAS SOCIAIS

ATA DE AVALIAÇÃO DE TRABALHO DE CONCLUSÃO DE CURSO

Aos 11 (ONZE) dias do mês de FEVREIRO do ano de 2022, às 9h horas compareceu perante a banca Examinadora o(a) aluno(a) MAYRES LANGE PESVENO DOS SANTOS SILVA autor(a) do Trabalho de Conclusão de Curso - TCC intitulado Índice H5 do Google Acadêmico e Webometrics Capes na Produção da Sociologia Brasileira sendo a Banca Examinadora constituída pelos professores: MARINA FÉLIX DE MELO (orientador/a), JOSE ALEXAN- DRE DA SILVA JÚNIOR E RANULFO PARANHOS DOS SANTOS FILHO que atribuíram respectivamente as seguintes notas: 1º examinador 9,0 (NOVE), 2º examinador 9,0 (NOVE), 3º examinador 9,0 (NOVE), cuja média aritmética é 9,0 (NOVE), tendo a referida banca considerado(a) aprovado(a) e apto(a) para a Colação de Grau de BACHARELADO em Ciências Sociais.

F por estar conforme, eu _____

_____ técnico do Instituto de Ciências Sociais lavrei a presente ata que vai assinada por mim, pelos membros da banca e pelo Diretor do Instituto de Ciências Sociais.

1º Examinador(a):

Marina Félix de Melo

2º Examinador(a):

Documento assinado digitalmente
gov.br Jose Alexandre da Silva Junior
Data: 11/02/2022 11:42:25-0300
Verifique em <https://verificador.it6.br>

3º Examinador(a):

Ranulfo Paranhos

Diretor(a) do Instituto de Ciências Sociais

Coordenação do Curso de Ciências Sociais

“Um pouco de Ciência nos afasta de Deus.

Muito, nos aproxima.”

BLAISE PASCAL

AGRADECIMENTOS

Agradeço a todas as pessoas que compõem o Instituto de Ciências Sociais da Universidade Federal de Alagoas. Especialmente à minha orientadora Dra. Marina Félix de Melo e ao grupo CNPQ produção acadêmica científica e tecnológica. Minha gratidão também ao Grupo de Pesquisa Cidadania e Políticas Públicas na pessoa do Dr. Ranulfo Paranhos que a partir de metodologia das Ciências Sociais me fez ter a certeza de que eu poderia continuar na graduação. Agradeço especialmente ao professor Dr. José Alexandre da Silva Júnior que possui minha admiração e respeito por seu compromisso com a docência e a pesquisa científica.

À minha família, especialmente minha mãe que é uma mulher forte e pode ver a primeira pessoa de sua família se graduando. Aos meus amigos e amigas de turma da graduação: Laura Gusmão, Pedro Vitorino e especialmente Quemuel Baruque que se tornou meu irmão de jornada. Aos pesquisadores do Grupo de Pesquisa Denisson Santos, Willber Nascimento e Breno Avelino. Aos amigos que sempre foram um refúgio durante as horas e horas de universidade: Laura Melo, Edvonaldo Horácio, Thaís Galvão, Débora Moreira, Pedro Paz, Valentina Souza, Nathalie Nayara, Larissa Jessica. Aos Grupos formados por cristãos na universidade: Cru Campus e Cristãos na Ciência. Os diálogos que pude ter com vocês foram importantes para ressaltar em mim o desejo e continuar pesquisando e explorando fenômenos sociais.

Por fim, ao Deus trino, criador de todas as coisas.

ÍNDICE H5 DO GOOGLE ACADÊMICO E WEBQUALIS CAPES NA PRODUÇÃO DA SOCIOLOGIA BRASILEIRA

Resumo: Qual será a diferença nas qualificações das revistas após a suposta extinção do sistema qualis e adoção do sistema h5 google para avaliação dos periódicos? O presente artigo objetiva demonstrar as diferenças das avaliações de periódicos científicos examinados pelo sistema Webqualis da Capes e pelo sistema h5 do google na área da sociologia. No Brasil, o qualis Capes é um dos principais fatores que orienta e estrutura o processo da produção acadêmica e científica, sobretudo, dos programas de pós-graduação. Na presente investigação, analisamos 2.019 periódicos qualificados pela plataforma do Webqualis da Capes da área de sociologia, classificados de 2016 a 2020. Estatisticamente, com o auxílio de um banco de dados construído no software R, utilizamos um modelo de análise de conglomerados para avaliar a diferença entre os dois sistemas. Os resultados apontaram 4 clusters ótimos, onde o cluster com score h5 mais alto não continha a totalidade de revistas de qualis A1 da Capes, que se distribuíam nos clusters 1 a 3. A pesquisa indicou que na área de avaliação da sociologia nem sempre publicar em revistas de qualis A1 tende a ser parte do score de citações mais altas no sistema h5 do google.

Palavras-Chave: Índice H5; Qualis Capes; Análise de Clusters; Sociologia.

H5 INDEX AND WEBQUALIS CAPES IN ACADEMIC PRODUCTION ON BRAZILIAN SOCIOLOGY

Abstract: What will be the difference in magazines after the supposed extinction of the qualis system and the adoption of the h5 google? This paper aims to demonstrate the differences between evaluations of scientific journals examined by Webqualis Capes and by google h5 system in sociology area. In Brazil, the Capes qualis is one of the main factors that guides and structures the process of academic and scientific production, especially in graduate programs. In this investigation, we analyzed 2.019 journals qualified by the Webqualis platform in the sociology, classified from 2016 to 2020. Statistically, using a database built in the R software, we had a cluster analysis model to assess the difference between the two systems. The results showed 4 optimal clusters, where the cluster with the highest h5 score did not contain all the journals with A1 qualis from Capes, which were distributed in clusters 1 to 3. The research indicated that in the area of sociology evaluation, not always publishing in A1 qualis journals tend to be part of the highest citation score in google's h5 system.

Keywords: H5 Index; Qualis Capes; Cluster Analysis; Sociology.

1. Introdução

Artigos científicos desempenham um papel importantíssimo na produção acadêmica. Comumente pesquisadores são incentivados a publicarem suas pesquisas em revistas bem avaliadas para alcançarem nível de internacionalização (BORNMANN & WILLIAMS, 2017). Medidas de impacto científico são cada vez mais utilizadas para identificar se pesquisadores e trabalhos acadêmicos podem receber promoções científicas, avaliações de bolsas e avaliações de departamentos universitários e centros de pesquisa (DOGAN *et al*, 2016). A validação de uma revista científica dá crédito, reconhecimento e visibilidade em publicações. Uma revista bem classificada por uma agência acaba atraindo um maior número de leitores e autores e logicamente, nesse processo, a revista tenderá a ser melhor avaliada. Tradicionalmente, na academia tem-se utilizado o fator de impacto de um periódico para sua hierarquização, isto é, dividem-se as citações e número de artigos publicados em determinado período.

Em 2005, o índice H começou a ser utilizado para quantificar a produtividade de cientistas com base em seu registro de publicações (HIRSCH, 2005). Esse tipo de métrica pode ajudar a identificar um número significativo de citações, como também serem úteis na divulgação de impactos intelectuais e internacionais mais amplos. (MEHO & YANG, 2007; DOGAN *et al* 2016; DHAMDHERE, 2018). O *Google Scholar Metrics* (GSM) tem diferentes índices para revistas científicas, o h5 e i10. Neste artigo escolhemos trabalhar com o índice h5 que considera apenas o número de citações dos artigos publicados nos últimos 5 anos. Uma vantagem disso é a de que o Google Scholar possui acesso livre e está disponível gratuitamente perante outros índices, sejam a *Web of Science* ou o *Scopus*, que são computados de maneira precisa, mas não acessíveis a maior parte das revistas no Brasil.

Com o aumento da produção científica no País, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), fundação vinculada ao Ministério da Educação (MEC), criou a avaliação Qualis-Periódicos - uma forma de classificar revistas onde os pesquisadores brasileiros publicam artigos, estabelecendo padrões de qualidade nas publicações científicas. Atualmente é o sistema Qualis que mais incita discussões devido a sua visibilidade nacional. O Qualis é um conjunto de métodos utilizados pela CAPES para a estratificação da qualidade da produção científica dos programas de pós-graduação do Brasil. A estratificação da produção é realizada por meio da análise da qualidade dos veículos de divulgação dos periódicos científicos, um índice que varia desde o estrato C até o A1. (CAPES, 2013)

A atribuição dos estratos indicativos de qualidade vão de A1 (mais elevado), A2; B1; B2; B3; B4; B5 a C – com peso zero. Toda avaliação é feita de acordo com os critérios de cada áreas do conhecimento, especificados na documentação de área relativa ao último triênio e quadriênio da Capes (2010 a 2012; 2013 a 2016) disponíveis no site da Instituição.¹ O mesmo periódico pode ser classificado em mais de uma área de conhecimento e receber diferentes avaliações, expressando assim a pertinência do conteúdo que é veiculado.

Souza (2021) se propõe a dar visibilidade à cronologia sobre a mudança no Qualis Periódico no quadriênio 2017-2020. Em julho de 2020 publicado pela diretoria de avaliação da CAPES publica um relatório técnico reconhecendo a dificuldade da avaliação de periódicos em humanidades². O mesmo recomendava o uso de índices bibliométricos como tentativa de sanar o problema exposto pela carta-manifesto de coordenadores de programas de pós-graduação em humanidades. Em Outubro de 2020 durante uma palestra promovida pela CAPES são apresentadas em seminário ponderações acerca do índice h, via Google Scholar, que poderá ser adotado por parte das humanidades.³

Neste contexto, perguntamos: qual será a diferença nas qualificações das revistas após a suposta extinção do sistema qualis e adoção do sistema h5 google para avaliação dos periódicos? Propomos utilizar análise de cluster para perceber grupamentos das revistas que são avaliadas pelo comitê da Capes na área de sociologia. Essa técnica permite que os objetos do grupos sejam observados de forma similar tanto quanto possível e, de forma análoga, que os grupos de clusters sejam os mais diferentes o possível em sua constituição (KASSAMBARA, 2017), tal agregação é feita baseada na distância euclidiana (proximidade) dessas observações em relação aos outros observados (HAIR *et al.* 2009).

Dentre os diversos métodos de clusterização que existem, utilizamos nesse artigo a Clusterização Particionada, utilizando a técnica *K-means*. Tal decisão metodológica ocorre para se verificar os diferentes qualis em citações pareadas. Assim, poderemos classificar objetos em diferentes grupos e cada um deve conter objetos semelhantes por meio da distância estatística mensurada entre eles. Existem outros métodos que permitem mensurar a distância euclidiana entre os objetos, neste caso escolhemos adotar o método de partição, pois, nos fornece o quão homogêneos são os elementos nos grupos (BUSSAB *et al.*, 1990). Para esse grupamento,

¹<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>. Acessado em 06 de Fevereiro de 2022.

² <https://www.gov.br/capes/pt-br/centrais-de-conteudo/23072020-dav-aph-pdf>. Acessado em 06 de Fevereiro de 2022.

³ <https://www.youtube.com/watch?v=QsuHiUsZdgQ&list=PLOGKhHkjqQoIvwX1KDfi5CCSoZrehR0p>. Acessado em 06 de Fevereiro de 2022.

utilizaremos as variáveis H5.Google que indexa h dos artigos publicados nos últimos cinco anos. Trata-se do maior número h de uma publicação, em que h artigos publicados de 2016 a 2020 tenham sido citados no mínimo h vezes cada. Da mesma forma, descrevemos a mediana do H5 que chamamos de H5.Mediana.Google, variável que consiste na média de citações para os artigos que compõem seu índice $h5$.

A fim de identificarmos quais seriam as principais implicações caso a adoção do índice $h5$ do google fosse utilizada em substituição do qualis CAPES, subdividimos a análise desse fenômeno nesse artigo da seguinte forma: a primeira parte diz respeito à descrição da técnica de recolha de informações; logo após, definição e normalização das variáveis no banco de dados; a terceira parte apresenta a análise dos dados; a quarta discute os resultados e possíveis questionamentos não exaustivos acerca do índice apresentado aqui.

2. Técnica de Recolha de Informações

A presente investigação conta a análise de 2.019 periódicos avaliados pela plataforma do WebQualis da Capes de Sociologia⁴. A variáveis aferidas foram: 1. Nome do Periódico; 2. ISSN; 3. Nota do Qualis Capes (A1, A2, B1, B2, B3, B4, B5 e C); 4. Peso Equivalente à nota Qualis de acordo com o Documento de Área (2016) do Setor de Sociologia da Capes (A1: 100, A2: 85, B1: 70, B2: 60, B3: 40, B4: 30, B5: 10 e C: 0); 5. Qualis Superior (A1, A2 e B1) ou Inferior (B2, B3, B4, B5 e C); 6. Valor do Índice H5 do Google Acadêmico e; 7. Mediana do Índice H5 do Google Acadêmico.

Para o processo da técnica de recolha de informações desta pesquisa, os dados foram coletados de forma manual, um a um, da página web que indica o índice H5 Google Acadêmico⁵, tabuladas e analisadas no softawre de código aberto *Rstudio*, por meios dos seguintes pacotes: *dplyr* (manipulação), *ggplot2* (visualização), *magrittr* (operador) e *factoextra* (análise exploratória de dados). Optamos por não coletar os dados por técnica *web scrapping*. Esta seria, grosso modo, a criação de uma função para extrair rapidamente informações do site em que constam as informações dos índices H5 de cada revista. Esta técnica auxilia na exportação de dados para planilhas e demandaria a criação de uma função no *software R* para a extração de dados em xml, sendo esta função uma maneira bastante interessante de se realizar

⁴<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>. Acessado em 06 de Fevereiro de 2022. É importante ressaltar que não são periódicos de Sociologia, mas sim, avaliados pela Sociologia.

⁵ https://scholar.google.com/citations?view_op=top_venues&hl=pt-BR&vq=en. Acessado em 06 de Fevereiro de 2022

a coleta automatizada de informações. Entretanto, ao realizarmos a análise exploratória da coleta manual dos dados na página web do índice H5, percebemos alguns problemas na forma como estão disponíveis as informações. Não há no site uma padronização na disposição da busca pelas informações, o que poderia nos levar bastante tempo para adaptar a coleta automatizada. Os principais problemas técnicos na recolha dos dados e que nos fizeram optar por uma recolha manual foram:

1. Na lista do WebQualis Capes, algumas revistas têm uma nomenclatura que precisa ser traduzida para outro idioma para ser localizada na página web de informações do índice H5;
2. A página web de informações do índice H5, na maioria dos casos, não considera sinais como parênteses ou colchetes, muitos utilizados na página do WebQualis;
3. Algumas revistas estão cadastradas com sinais de “&”, outras com “and”, outras com “e”, sem uma padronização;
4. Acentos muitas vezes não são reconhecidos pelo sistema e por vezes são reconhecidos a ponto da revista só ser localizada na página web de informações do índice H5 se estiver acentuada conforme a página a tem cadastrada;
5. Alguns periódicos são localizados com a palavra “Revista” ou “Journal” antes do nome das revistas, já outros não. O lançamento da palavra “revista” ou “journal” é determinante na localização do periódico. Usar as palavras “revista” ou “journal” pode fazer com que o periódico não seja localizado, como não usa-las pode derivar no mesmo tipo de erro;
6. Alguns periódicos são identificados na página do WebQualis Capes como “Journal” e na página do Google H5 como “Revista”, e vice-versa, e a troca destes termos impede a localização do periódico no sistema;
7. No sistema H5, por vezes, encontramos o mesmo nome de revista grafado duas vezes, com índices diferentes para letras maiúsculas e minúsculas, indicando dois periódicos avaliados distintamente. Como não há diferenciação das revistas por ISSN, o usuário não tem como localizar se a avaliação dupla trata de um mesmo periódico;
8. Periódicos que possuem homônimos, especialmente homógrafos, ou seja, periódicos com o mesmo nome e que o índice H5 só reconhece uma nomenclatura, como se existisse apenas um periódico naquele título. Não sabemos ao certo como estes dados são computados na página do índice H5 posta a pouca divulgação destes ajustes

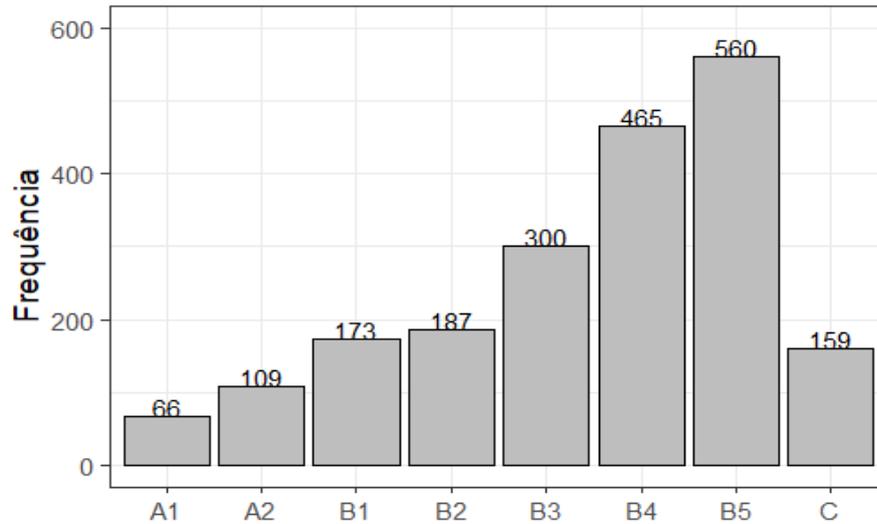
metodológicos do sistema google. Todavia, a preocupação da nossa coleta com estes dados foi a de isolar com um *missing* específico (97) estes casos por não sabermos se as citações das duas ou mais revistas se confundem, ou se somam-se, ao índice apresentado no site. Alguns destes casos são resolvidos pelo H5 com a apresentação da origem (universidade, cidade, país) da Revista, outros, não.

Problemas que sistematizamos podem ser um empecilho para o usuário, que pode não encontrar o periódico e supor que este não exista na base de dados do índice H5 do google acadêmico. Embora no site do índice H5 google não exista a possibilidade de consulta do ISSN da Revista, como esta informação está contida no site do WebQualis Capes, utilizamos o ISSN por diversas vezes para localizar os sites dos periódicos e obtermos informações detalhadas da origem da revista para localizarmos quais revistas do H5 são devidamente correspondidas no WebQualis. Após a busca pelos 2.025 periódicos, realizamos uma verificação de todos os casos de *missing*, não localizados na página do H5 google, a partir dos pontos explicados acima para termos uma margem mais segura de que os casos omissos apresentados nesta pesquisa significam, de fato, que os periódicos considerados pelo WebQualis Capes não estão na base de dados do índice H5 (supostamente por ainda não terem citações localizadas pelo algoritmo do fator de impacto do google acadêmico).

3. Descrição das Variáveis

Conforme o gráfico de barras abaixo, temos as 2.019 revistas divididas nos seguintes estratos do WebQualis Capes: A1 (66 Periódicos); A2 (109); B1 (173); B2 (187); B3 (300); B4 (465); B5 (560) e; C (159). Posto que a Capes considera como revistas de estrato superior os periódicos avaliados como A1, A2 e B1, temos 348 revistas (17,5%) avaliadas neste estrato e 1.671 (82,5%) avaliadas como revistas de estrato inferior pelo WebQualis Capes.

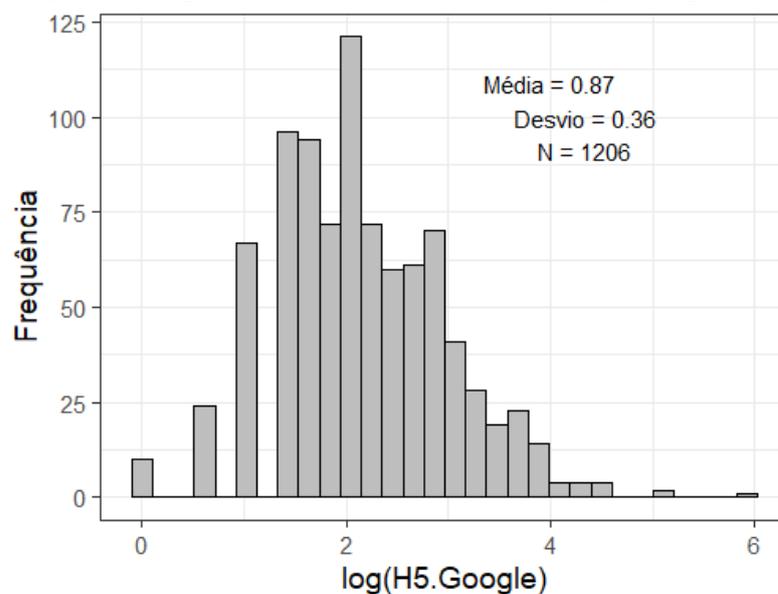
Figura 1 - Nota WebQualis Capes (A1 a C)



Fonte: Elaboração própria

A variável dependente da nossa pesquisa, o índice H5 do google acadêmico, é uma variável que possui originariamente uma distribuição bastante assimétrica⁶. Buscamos alguns tipos de correção, exponencial, por raiz quadrada e por logaritmo na base 10 (LG10). A variável que demonstrou padrões mais simétricos foi a criada a partir de correção logarítmica na base 10 (LG10)⁷. Utilizaremos esta variável para a análise dos dados, independentemente da utilização de testes paramétricos. A distribuição pode ser visualizada na Figura 2.

Figura 2 - Distribuição da Frequência do Índice H5 com Correção Logarítmica



⁶ O valor de sua medida de assimetria dividida pelo seu erro padrão é de 128 e da sua curtosis dividida pelo seu erro padrão de 1.002. Estes valores estão, devido à grande presença de *outliers*, casos distantes da curva, muito afastados dos valores de -2 a 2, referenciais para os padrões de simetria baseados na curva normal estatística.

⁷ O valor de sua medida de assimetria dividida pelo seu erro padrão é de 7,26 e da sua curtosis dividida pelo seu erro padrão é de 2,37, com a exclusão inicial de 6 casos de *outliers*.

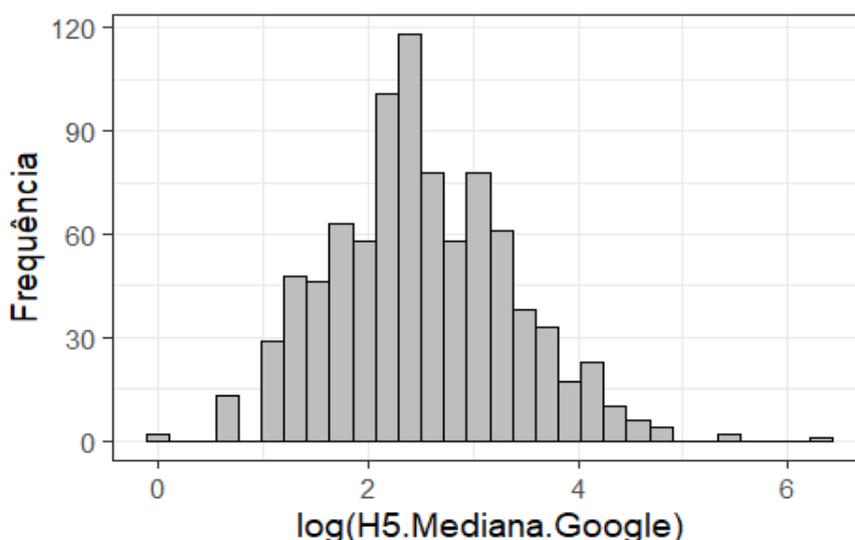
Fonte: Elaboração própria

A variável H5.Google (LG10) conta com uma quantidade expressiva de casos omissos. De 2.025 revistas consultadas na base do google, 1.206 (60% dos periódicos) têm algum valor de índice registrado, porém, 819 (40%) não possuem qualquer registro, indicando que ou há problemas técnicos para este registro ou estamos a lidar com revistas que não possuem fator de impacto registrado nos últimos 5 anos.

Em termos brutos, isto é, a considerar a variável H5.Google sem a correção por LG10, verificamos que a revista que possui o maior índice H5 da nossa amostra é uma revista que registra o valor máximo de um índice de 376, sendo esta uma revista de notoriedade máxima na comunidade acadêmica internacional. A média do índice H5 da amostra de revistas avaliadas no Qualis Capes da Sociologia é de 11,75, sendo o desvio padrão de 18,3, demonstrando a grande assimetria da distribuição. Nota-se que 40% das revistas concentram índices que variam entre 1 e 9, ou seja, com fatores de impacto relativamente baixos em relação à distribuição completa. A moda da distribuição, com 151 casos (7,5%) é composta por periódicos que recebem o índice H5 de valor 4.

Observamos também na distribuição a variável H5.Mediana.Google, como indicada na Figura 3, que de forma similar também possui uma quantidade expressiva de casos omissos com distribuição assimétrica. A média da Mediana H5 corresponde a 18,54 sendo o desvio padrão de 22,00 reforçando a distribuição assimétrica.

Figura 3- Distribuição da Frequência da Mediana H5 com Correção Logarítmica



Fonte: Elaboração própria

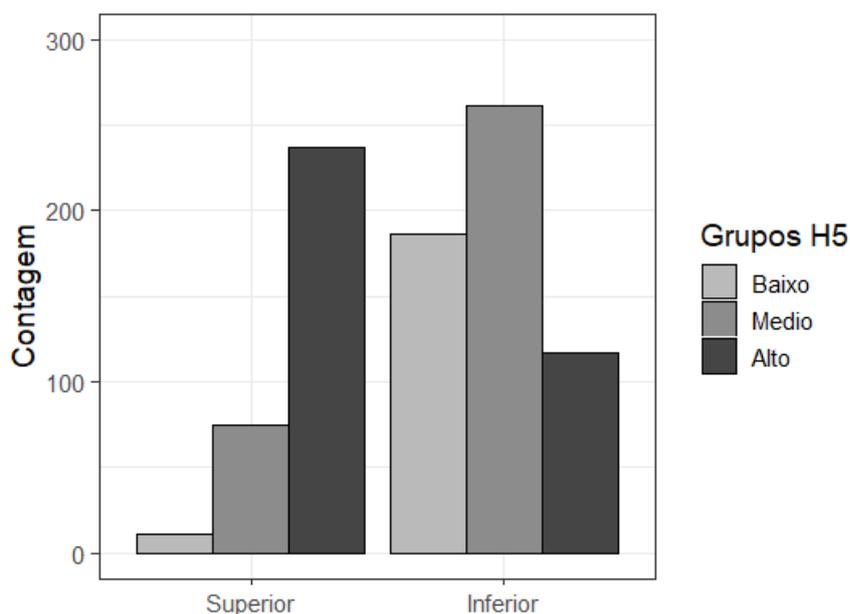
4. Análise dos Dados

Realizamos uma redução de mensuração da variável Índice H5 Google, transformando a variável originalmente escalar em variável ordinal, dividida em 3 estratos a partir das suas frequências observadas. Geramos 3 grupos, sendo estes: 1. revistas que têm baixo índice H5, variando este índice de 1 a 4 (29% da amostra); 2. revistas com índice H5 mediano, variando de 5 a 9 (38% da amostra) e; 3. revistas de índice H5 elevado, acima de 10 (33% da amostra), variando de 10 a 100, posto que retiramos da análise os 6 casos de *outliers* com valores acima de 100. As 819 revistas que não constam na base de informação do índice H5 do google acadêmico não entram na análise, tampouco são representadas como valor 0 na variável ordinal criada.

A partir do teste não paramétrico de associação de *Spearman Rho*, encontramos um coeficiente de correlação de 0,533, com significância estatística ($p=0,00$) indicando uma associação positiva substancial entre o nível do qualis capes e o índice H5 google. Semelhantemente, na verificação da associação entre o índice H5 google e o fato das revistas serem avaliadas como qualis superior ou inferior pela Capes, encontramos um coeficiente de associação (*V de Crámer*) de -0,525, com significância estatística ($p=0,00$), indicando uma associação substancial entre as duas variáveis. Ou seja, não é tão aleatória a relação quanto poderia se supor ao encontrarmos revistas com alto qualis com fator de impacto baixo e vice-versa.

A Figura 4 mostra, as revistas de qualis superior (A1 a B1) e as revistas avaliadas em qualis inferior Capes (B2 a C) de acordo com o índice H5 do google acadêmico:

Figura 4 - Qualis Superior e Inferior de acordo com Grupos do Índice H5



Fonte: Elaboração própria

Aplicamos um teste de diferença entre duas médias, *T Student*, para verificação de diferença estatisticamente significativa entre as médias do índice H5 de periódicos avaliados como superiores e inferiores pelo Qualis Capes. A média do índice H5 (Log10) para as revistas de qualis superior é de 1,17 e para as revistas de qualis inferior é de 0,76. Assumindo igualdade de variância a partir dos resultados do teste de *Levene*, verificamos que a diferença entre estas duas médias (0,41) é estatisticamente significativa ($p=0,00$), indicando que, de fato, revistas de qualis superior e de qualis inferior possuem médias do índice H5 com diferenças na realidade que não são fruto do mero acaso. Este resultado, embora aparentemente tautológico posto que revistas melhor qualificadas pelo WebQualis Capes tendem a ser revistas de maior visibilidade e fator de impacto no mundo acadêmico, não se replica quando separamos a análise das médias do índice H5 pelos diferentes estratos do WebQualis.

A partir de uma análise de variância (ANOVA), ratificada por análise de clusters de agrupamento hierárquico, que será detalhada a partir da Figura 6, observamos as seguintes médias do Quadro 1 do índice H5 (LG10), divididas pelos estratos do WebQualis Capes:

Quadro 1 - Média H5 (LG10) de acordo com Estrato Qualis

A1	1,26
A2	1,27
B1	1,07
B2	0,87
B3	0,77
B4	0,75
B5	0,69
C	0,68

O ANOVA aplicado demonstra homogeneidade de variância a partir do teste de *Levene* analisado e tem um modelo como um todo significativo ($p=0,00$).

A considerar as diferenças de médias do índice H5 de cada estrato em relação a outros, encontramos como resultado 4 grupos de homogeneidade, apresentados abaixo.

Quadro 02: Grupos de Homogeneidade das Médias H5 por Estratos do Qualis

Tukey HSD ^{a,b}		Subconjunto para alfa = 0.05			
Nota WebQualis Capes (A1 a C)	N	1	2	3	4

C	38	,6874			
B5	233	,6947			
B4	271	,7534	,7534		
B3	191	,7796	,7796		
B2	150		,8756		
B1	150			1,0724	
A1	61				1,2606
A2	106				1,2750
Sig.		,366	,076	1,000	1,000

Os 4 grupos apresentados nas colunas à direita demonstram que embora as médias destes grupos tenham valores diferentes, estas são, na prática, descritoras de uma mesma realidade, posta que são médias que não apresentam diferenças estatisticamente significantes entre si.

Grupo 01: Revistas dos estratos A1 e A2

Grupo 02: Revistas do estrato B1

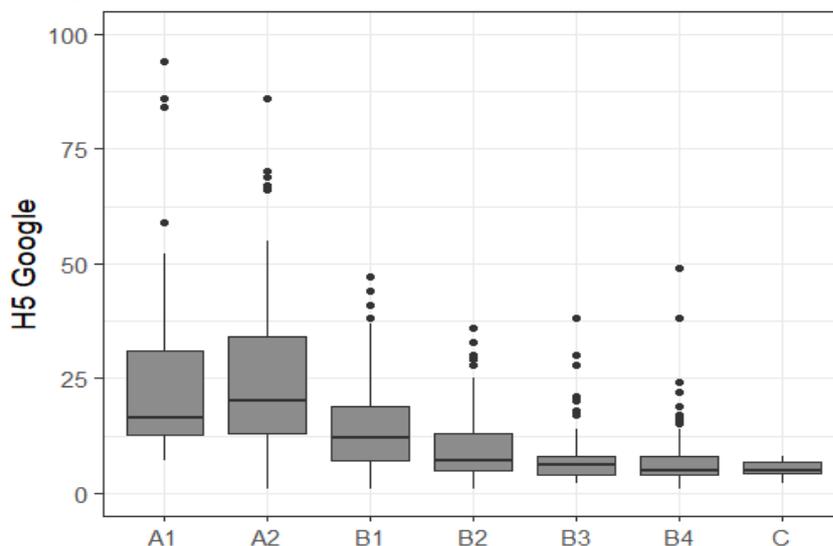
Grupo 03: Revistas dos estratos B2, B3 e B4

Grupo 04: Revistas dos estratos B3, B4, B5 e C

Em outros termos, se separássemos 4 grupos com características semelhantes a respeito do índice H5 e que fossem diferentes entre si, estes seriam os 4 grupos apresentados, e não os 8 estratos existentes no modelo do WebQualis⁸.

⁸ Realizamos também uma análise de clusters que tem por finalidade agrupar dados em grupos com características em comum, unindo as informações por distâncias métricas através de um agrupamento hierárquico. O resultado aferido pelo dendograma da análise de clusters confirma o resultado apresentado pelos grupos de homogeneidade da análise de variância desta pesquisa.

Figura 5 - Distribuição de Citações do Índice H5 por qualis

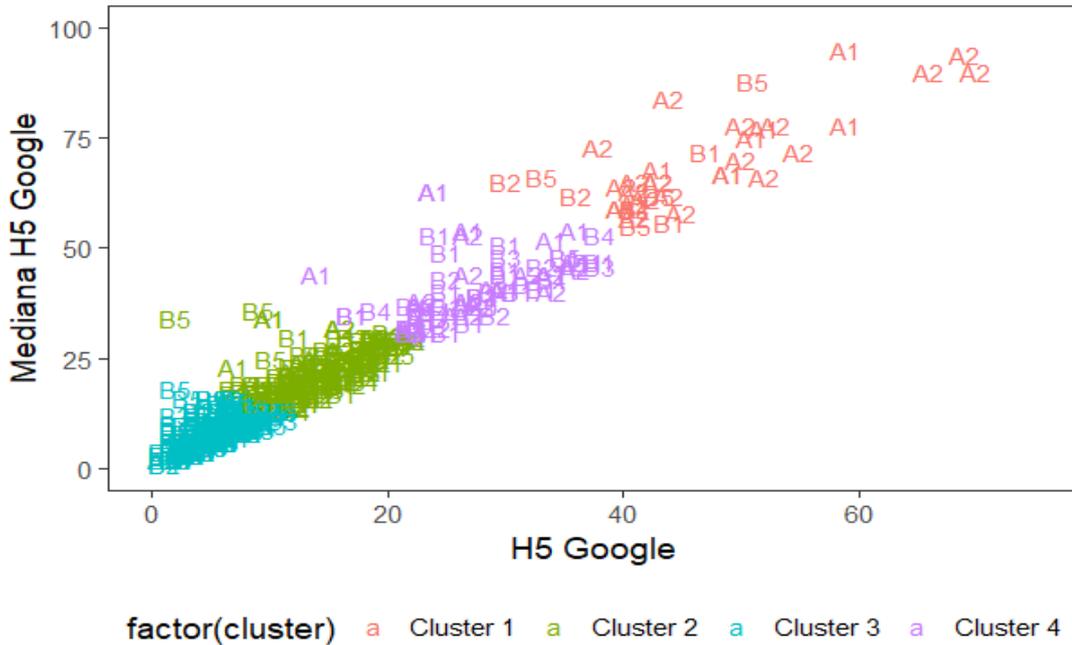


Fonte: Elaboração própria

A Figura 5 nos mostra por meio do bloxplot comparativo as distribuições de citações do índice H5 por qualis. Percebemos que as distribuições dos grupos que compõem o qualis superior possuem o segundo quartil com medianas mais altas que os do grupo de qualis inferior. Tal fenômeno é interessante, sobretudo, tendo em vista que o qualis A2 tem a maior distribuição, inclusive maior do que o qualis A1. O Qualis B1 diferente dos demais é o que mais se aproxima de uma distribuição simétrica.

Com o intuito de ratificar os testes paramétricos realizados acima decidimos utilizar ambas as variáveis, H5.Google e H5.Mediana.Google, a fim de cruzar tais métricas e observar a similaridade entre as revistas por meio dos clusters gerados. Essa similaridade é determinada através do coeficiente de similaridade e correspondência entre os itens.

Figura 6 - Grupamento por Clusters

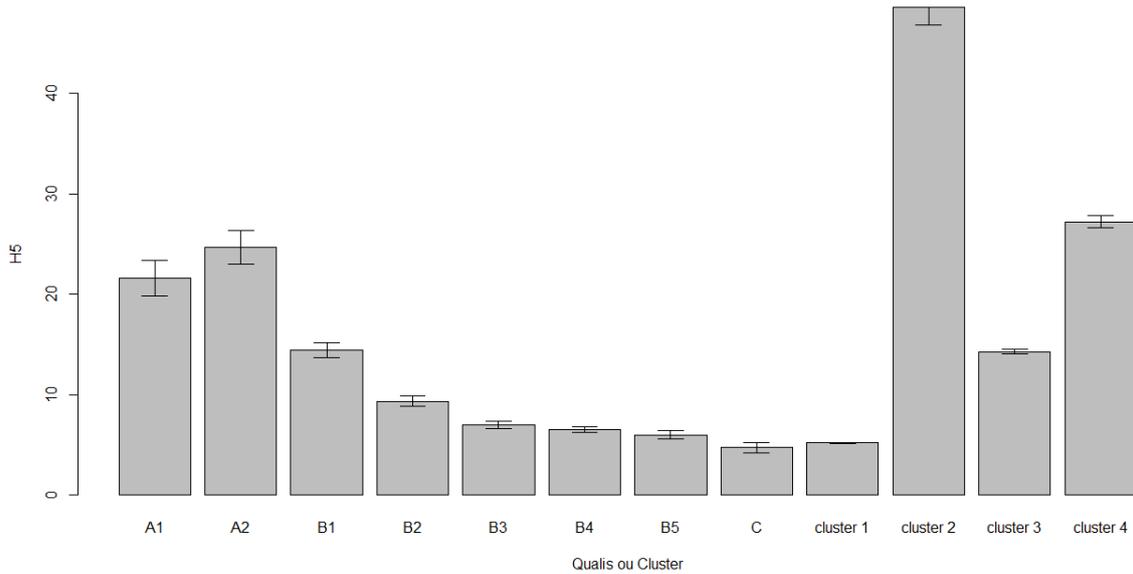


Fonte: Elaboração própria

O resultado descrito na Figura 6 são os grupamentos gerados. Podemos observar que o cluster de número 2, que é o grupamento com scores de citações mais altas, agrupa em sua maioria (83%) revistas Qualis A2 (57,1%) e A1 (21,4%). Já o cluster de número 4, que agrupa o segundo grupo com maiores scores, tem em sua maioria (62,1%) revistas de qualis A2 (26,8%) e B1(35,3%).

Podemos ver na Figura 6 que apesar do cluster 1 ser o menos homogêneo e possuir a maior variação, tanto do H5 Google quanto a mediana do Google, ele é o que tem a maior definição de quais revistas são diferentes entre si neste cluster. Já as revistas que o compõem estão em um cenário de citação mais distante das demais, diferentemente dos clusters 3 e 4, que têm uma maior homogeneidade, porém, parte dos seus indivíduos poderiam ser facilmente permutados entre si.

Figura 7 - Distribuição Anova Qualis / Clusters



Fonte: Elaboração própria

Por meio do modelo ANOVA descrito na Figura 7 podemos perceber que o modelo de estratificações de revistas em diferentes níveis de qualis acaba sendo pouco homogêneo quando analisamos o número citações. O modelo indica que quando comparamos os grupamentos com os diferentes qualis percebemos uma baixa paridade, na qual apenas os estratos inferiores têm paridade com seus grupamentos. É visível, aliás, que os grupos de clusters formados se diferem entre si quanto ao índice H5, confirmando a distância relativa que precisa existir entre a formação dos grupos. Vemos que o cluster 2 acaba agregando maior índice de distribuição dos casos.

5. Conclusão

Objetivamos com os dados e informações aqui apresentados realizar os seguintes movimentos: (a) descrever como se comporta o H5 do Google quanto às revistas avaliadas pelo comitê CAPES de Sociologia; (b) verificar como esse índice se comporta sendo analisado diante de outra métrica, o H5.Mediana.Google. Afim de replicação da análise aqui discutida os dados e scripts estarão disponíveis publicamente.⁹

Assumimos os limites técnicos e metodológicos do presente artigo. Quando a CAPES se refere a índices bibliométricos cita os principais de três bases: *Scopus (CiteScore)*, *Web of Science (Fator de Impacto)* e *Google Scholar (índice H5)*. Nossa análise referiu-se somente a

⁹ H5Google <https://osf.io/dxm9b/> via @OSFramework

um desses índices, mas ressaltamos que é necessário considerar a relevância que uma pesquisa posterior poderia ter a partir da coleta dos dados das demais bases. Outra limitação decorre dos valores *missing* encontrados em nossa análise. Cerca de 40% das revistas que não possuíam valores no h5 provavelmente são revistas abaixo do qualis C. Então, algumas perguntas surgem a partir disso: (a) será que tais revistas realmente não possuem visibilidade pelo índice h5? Logicamente teremos limitações ao compreender realmente como revistas C se comportam dentro da análise. Como 40% das revistas podem não ser consideradas pelo H5 Google, isso limita e direciona ainda mais publicações para revistas que se mantêm em Qualis superior enquanto outras revistas podem não receber boas publicações com o passar do tempo. Tal ação pode gerar invisibilidade dessas revistas que antes no meio científico recebiam algum conceito e com a adoção do índice H5 passarão a não receber mais.

Essa análise assume o pressuposto de que quanto maior o número de citações de uma revista melhor avaliada ela é, o que não necessariamente pode ser tomado como garantia. Suponhamos que existem revistas especializadas em um ou dois temas. Uma dada revista pode obter um H5 elevado entre essa temática devido ao fato de autores citarem uns aos outros entre si. Será que revistas mais especializadas tendem a ser mais citadas?

Relembramos que de acordo com nossa análise descritiva, as publicações mais citadas são de revistas de Qualis superior. Mas, dentre os variados Qualis CAPES, as revistas que se enquadram no Qualis A2 são mais citadas que as demais, o que é completamente inesperado de acordo com o pressuposto de avaliação convencional. O esperado seria que revistas que se encontram no Qualis A1 deveriam ser as mais citadas.

Foi possível perceber na clusterização que revistas de conceitos que não são elevados também publicam bons artigos. Como interpretação do resultado dos 4 grupos de semelhanças formados na análise de cluster, encaramos que é possível ratificar com a decisão da CAPES de qualificar estratos como superior e inferior. Quando verificada a homogeneidade que aparece nos grupos, esta parece corroborar com tal lógica, pois foram encontrados dois grupos homogêneos para qualis inferior e dois grupos homogêneos para qualis superior.

Caso o H5 Google seja adotado para avaliações de produção, especialmente para programas de pós-graduação em sociologia, a lógica dos qualis não será completamente modificada. Em verdade, essa lógica é reforçada pelo H5 Google. Isso não quer dizer que o sistema não precisa ser revisado pois, como observados na formação de clusters, existem grupos com índice H5 e mediana H5 elevados mesmo tendo revistas de qualis inferior. O contrário também é verdadeiro, existem clusters formados com a maioria das revistas com índice H5 e mediana baixas, mas que contêm revistas de qualis superior.

Algumas ponderações são válidas a partir dessa análise. A tomada de decisão para a mudança de avaliação de periódicos, qualificando revistas a partir do H5 Google, não diz respeito somente a uma nova qualificação de revistas, mas à política classificatória hierarquizante gerada por meio do Qualis. Tal política sugere influenciar diretamente no fluxo de publicações e, conseqüentemente, influenciará no cotidiano de seus autores e em suas expectativas. A hierarquia do Qualis culmina diretamente nos processos de avaliação aos quais os pesquisadores poderão ser submetidos para a captação de investimentos por meio de agências de fomento. Quanto melhor situada na hierarquia do Qualis, maior o poder de atração para influenciar na captação de investimentos.

7. Referências

BORNMANN, Lutz; WILLIAMS, Richard. Can the journal impact factor be used as a criterion for the selection of junior researchers? A large-scale empirical study based on ResearcherID data. **Journal of Informetrics**, v. 11, n. 3, p. 788-799, 2017.

BUSSAB, W.O.; MIAZAK, E.S.; ANDRADE, D.F. Introdução à Análise de Agrupamentos. 9º Simpósio Brasileiro de Probabilidade e Estatística. São Paulo: IME – USP, 1990

DHAMDHARE, S. N. Cumulative citations index, h-index and i10-index (research metrics) of an educational institute: A case study. **International Journal of Library and Information Science**, v. 10, n. 1, p. 1-9, 2018.

DOGAN, Guleda; SENCAN, Ipek; TONTA, Yasar. Does dirty data affect google scholar citations? **Proceedings of the Association for Information Science and Technology**, v. 53, n. 1, p. 1-4, 2016.

HAIR, Joseph F., WILLIAN C. Black, BARRY J. Babin, e RONALD L. Análise Multivariada de Dados. 6a ed. São Paulo: Bookman. Tatham, 2009.

HIRSCH, Jorge E. An index to quantify an individual's scientific research output. **Proceedings of the National academy of Sciences**, v. 102, n. 46, p. 16569-16572, 2005.

KASSAMBARA, Alboukadel. Practical Guide To Cluster Analysis in R. 1 ed. USA: STHDA, 2017.

MEHO, Lokman I.; YANG, Kiduk. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. **Journal of the american society for information science and technology**, v. 58, n. 13, p. 2105-2125, 2007.

SOUZA, Marcos Eduardo. "Cronologia da discussão sobre a mudança no Qualis Periódico no quadriênio 2017-2020", <https://doi.org/10.7910/DVN/5UNJXH>, **Harvard Dataverse**, V4, UNF:6:I/9JbqvIa46mI0ktLUEZXg== [fileUNF], 2021.

SANTOS, Paulo Jorge dos. Novo Qualis - Live com Prof. Paulo Jorge dos Santos (membro do GT CAPES que elaborou a proposta). CPGSS UFPR. 14 de ago. de 2020. Disponível em: https://youtu.be/zTeT_k-wLqw. Acesso em: 06 fev. 2022