

Analizando dados públicos: um estudo sobre as compras de diárias e passagens no contexto pandêmico

Leonardo Marçal*

Resumo

A pergunta de pesquisa que envolve este trabalho de análise busca levantar a problemática à respeito do uso de diárias e passagens concedidas no ano de 2020 ao servidores público, levando em consideração o contexto pandêmico vivenciado e realizando um parâmetro com o ano de 2019. Desta forma buscamos identificar qual o valor gasto por cada órgão, valores gastos por cidade e a quantidade de viagens realizadas por mês. De forma metodológica utilizaremos o modelo *KDD – Knowledge Discovery in Databases (KDD)* é um processo de Descoberta de Conhecimento em Banco de Dados é utilizado na criação de projetos de ciência de dados. A hipótese levantada é que, com o crescimento da pandemia e a impossibilidade de viagens e aberturas de órgãos públicos em algumas regiões do país, é possível que houve economia a respeito, porém o fato necessita ser avaliado da melhor forma possível. Logo, justifica-se a realização da pesquisa em uma curiosidade acadêmica sobre o trabalho com dados no setor público e suas aplicações com dados verídicos e de confiança.

Palavras-chaves: análise de dados. biblioteconomia orientada à dados. dados abertos.

Introdução

Propositalmente, iniciamos uma análise de dados diante de um problema a ser resolvido, ou uma pergunta referente, que requer uma análise dos dados e das informações existentes acerca da temática tratada. A pergunta de pesquisa que envolve este trabalho de análise busca levantar a problemática à respeito do uso de diárias e passagens concedidas no ano de 2020 ao servidores público.

Podemos dizer que *Data Analytics*, também conhecida como Análise de Dados é o processo que torna possível a transformação de dados e informações em conhecimento para

*UFPE, Departamento de Ciência da Informação. E-mail: leonardo.marcal@ufpe.br

um propósito específico. Tão importante quando compreender que a análise de dados pode ocorrer sem o uso de computadores. Pois a maior relevância é buscar conseguir realizar a transformação uma informação em conhecimento.

Quando absorvemos informações, reações são desencadeadas em nossa mente. De acordo com a psicologia temos um mapa cognitivo, do qual se modifica na medida que aprendemos. Unido com nossas experiências esse mapa é traçado, e o utilizamos com a finalidade de organizar nossas vidas e procurar informação. É através de nossas experiências diárias que aprendemos com o surgimento do novo.

Levando em consideração o contexto pandêmico vivenciado e realizando um parâmetro com o ano de 2019. Desta forma buscamos identificar qual o valor gasto por cada órgão, valores gastos por cidade e a quantidade de viagens realizadas por mês. De forma metodológica utilizaremos o modelo *KDD – Knowledge Discovery in Databases (KDD)* é um processo de Descoberta de Conhecimento em Banco de Dados é utilizado na criação de projetos de ciência de dados. A hipótese levantada é que, com o crescimento da pandemia e a impossibilidade de viagens e aberturas de órgão públicos em algumas regiões do país, é possível que houve economia a respeito. Logo, justifica-se a realização da pesquisa em uma curiosidade acadêmica sobre o trabalho com dados no setor público e suas aplicações com dados verídicos e de confiança.

Dados versus Informação

A informação é um dado concreto do produto de conhecimento seja qual ele for. É através do contato com a informação que o sujeito desenvolverá conhecimento, porém o sujeito deve estar disposto a observar os dados e produzir novos conhecimentos estando de acordo com sua habilidade de percepção e experiência. Desenvolvendo então um conjunto de ação e reação promovendo sua aprendizagem e demonstrando os objetivos até então propostos na informação. Uma informação deve ter objetivos, se oposto for será apenas um conhecimento supérfluo, subjetivo e sem propósito. Devemos propor ação e reação nos sujeitos, contribuindo assim para a resolução dos grandes problemas em nossa sociedade.

É importante salientar os conceitos de Ciência da Informação e Teoria da Informação, a primeira estuda os processos de armazenamento e recuperação de informação. A segunda estuda a transmissão de informação por meio de sinais. As definições de informação são altamente influenciadas pela teoria da informação, a teoria propõe medir a quantidade de informação, sendo uma área especializada nas ciências sociais e biológicas. Seguindo tais constatações é possível definir a informação como: um fato, um reforço do que já se conhece (produto), a liberdade de escolha ao selecionar uma mensagem, a matéria-prima do qual se obtém o conhecimento, o que gera aprendizagem e ação nos sujeitos, a redução de incertezas, transforma hipóteses em veridades ([MACGARRY, 1999](#)).

A todo momento absorvemos informações, porém quase nada fica armazenado. Somente é extraído o que é relevante para nós, nosso cérebro classifica e seleciona o que é óbvio, o que surtirá ação, reação e aprendizagem para nossas vidas. Ignoramos certas informações, apreendemos o resto e agimos. É a tomada de decisão, esses são os objetivos dos eventos.

O que é evento? O evento é o que pode ser adaptado, além de modificar o espaço ele modifica o tempo da ação. São estímulos que promovem essa habilidade de percepção e geram o aprendizado. Podem ser divididos em duas categorias internas e externas. A primeira está associada às ações involuntárias, um espirro por exemplo; a segunda refere-se

a uma ação voluntária e tem objetivo. Somente após a definição desses termos é que é possível organizar os eventos da experiência, sejam eles internos ou externos, espaço e tempo. Pois como foi dito são modificáveis e adaptáveis ([MACGARRY, 1999](#)).

Em suma, a desordem (entropia) é a aversão de informação. O mundo está desordenado e nossa função é colocá-lo em ordem. Localizar um evento no tempo e no espaço é o primeiro passo essencial para sua organização. Desta forma a ordenação cria uma base de seleção e classificação dos conhecimentos e das informações, conforme analisamos no início desta análise, seja qual for a forma que se encontra o produto. O tempo é a experiência principal e gera uma transformação independente do contexto histórico em que se aplica. Logo, o tempo é uma experiência a ser ordenada e transformada em informação.

Analizando Dados Públicos

É notável que cada vez mais instituições públicas, estão tomando decisões embasada nos dados, com o crescimento informacional de seus bancos de dados, eles conseguem detectar anomalias, criar e monitorar indicadores e propor e implementar melhorias em seus processos de gestão.

É importante salientar que o Brasil é um dos grandes percussores de dados abertos no mundo, estes dados também denotados de dados interligados, podem ser dados governamentais como gastos diversos e investimentos, dados climáticos, dados de pesquisas censitárias, informações geográficas como dados de georreferenciamento, mapas, endereços porém necessitam que estejam disponíveis gratuitamente, com permissão e licenciados para uso, cópia, download, análise, reuso e acessíveis em via software livre ([OLIVEIRA; SILVA, 2016](#)).

Desta forma no Brasil, a Lei de Acesso à Informação (LAI) dispõe que os órgãos públicos devem promover a divulgação de informações de interesse coletivo ou geral na internet. A lei também define as hipóteses de sigilo e de informações pessoais, que são as exceções à regra de que os dados devem ser abertos.

Metodologia

Para melhor abrangência da pesquisa e sucesso sua concepção, seguimos o modelo *KDD – Knowledge Discovery in Databases (KDD)* é um processo de Descoberta de Conhecimento em Banco de Dados é utilizado na criação de projetos de ciência de dados.

Segundo Goebel ([DESCOBERTA...](#)), o termo KDD é usado para representar o processo de tornar dados de baixo nível em conhecimento de alto nível, enquanto mineração de dados pode ser definida como a extração de padrões ou modelos de dados observados.

Compreendendo que para realizar a análise de dados, deve-se primeiro determinar quais os tipos de dados disponíveis e quais dados serão necessários para realizar as análises necessária. Em seguida, são definidas as fontes de dados que serão usadas para coletar ou adquirir dados, consequentemente, será determinado o tipo de análise a ser feito e a ferramenta a ser utilizada.

Os dados obtidos através da plataforma de dados abertos do Portal da Transparência do Governo Federal¹, extraídos da base de dados de Viagens a Serviço, com a temporalidade de dois anos, sendo 2019 e 2020. A ferramenta utilizada para realizar as etapas do KDD foi o Rstudio com uso próprio de Linguagem de Programação R. Sendo criado um *script*

para a preparação e processamento, assim como a criação do modelo preditivo e produção da visualização.

Comunicação dos Resultados

Visando encontrar respostas das perguntas que envolviam o nosso problema de pesquisa, geramos através dos dados as seguintes visualizações, que estão organizadas de acordo com as perguntas objetivadas na introdução.

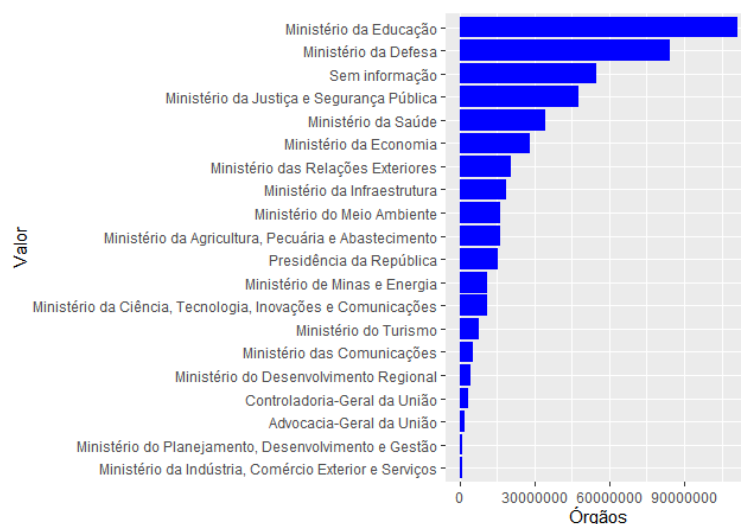


Figura 1 – Órgãos que mais gastaram no ano de 2019

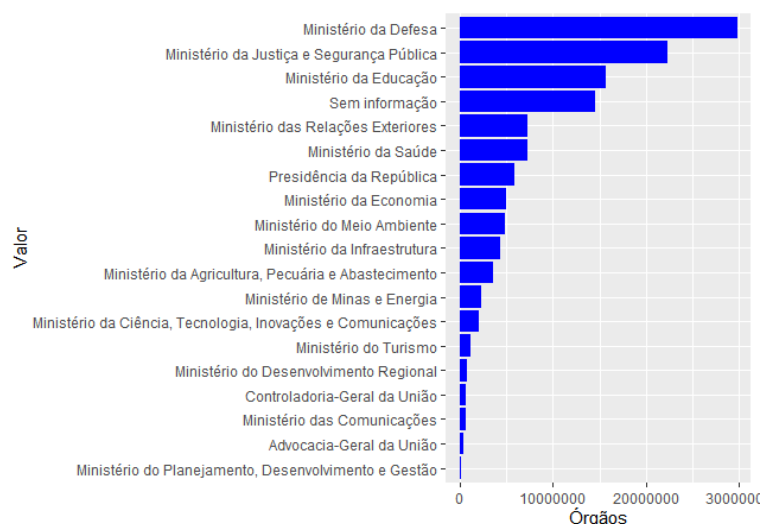


Figura 2 – Órgãos que mais gastaram no ano de 2020

Para salientar no ano de 2019 foram realizadas cerca de 820.936 viagens, a maioria realizada por servidores do Ministério da Educação, grande parte do público participaram de eventos educacionais e científicos, o que totalizou o valor de R\$ 1.249.052.823,88. Entretanto no ano de 2020 foram realizadas cerca de 306.000 viagens, grande parte pelos

servidores do Ministério da Defesa, contextualiza o período pandêmico ocasionado pela crise do coronavírus, com isso ou uma redução no valor de R\$ 519.533.322,48.

Nossa hipótese estava correta, houve uma economia de R\$ 729.519.501,4 nos cofres públicos e no orçamento destinado à viagens e serviços. É interessante notar que ao tentarmos extrair dados dos locais de viagens pelo Ministério, o dados estavam com a seguinte denominação: "informações protegidas por sigilo" conforme a Figura 4. Desta forma não sabemos ao certo em que foram utilizadas estas viagens e em que parâmetros elas se enquadram neste eixo.

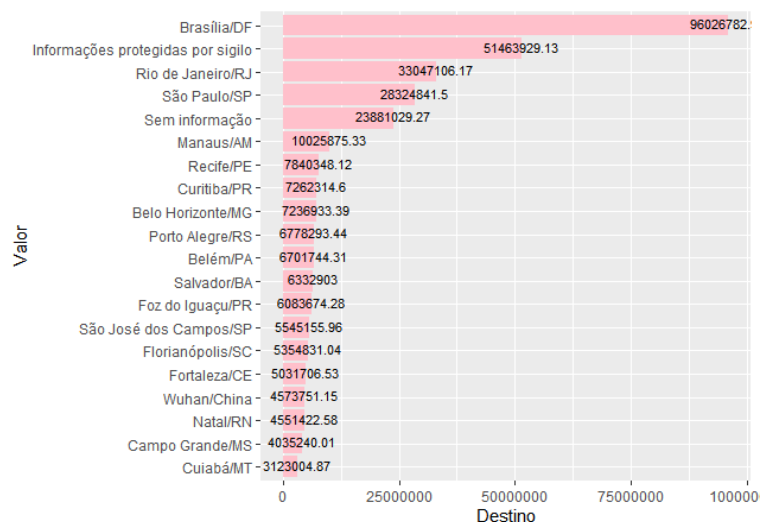


Figura 3 – Cidades de destino no ano de 2019

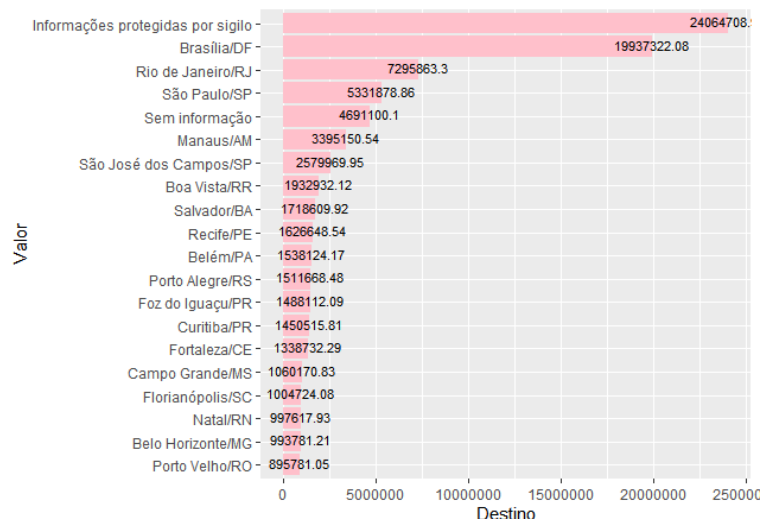


Figura 4 – Cidades de destino no ano de 2020

Por fim realizamos a análise de dispersão, que consiste em averiguar os resultados das quantidades de viagens realizadas e o valor, com a temporalidade de ano e mês que a sucede. Com isto podemos verificar em quais meses foram realizadas mais viagens. Conformes as figuras 5 e 6.

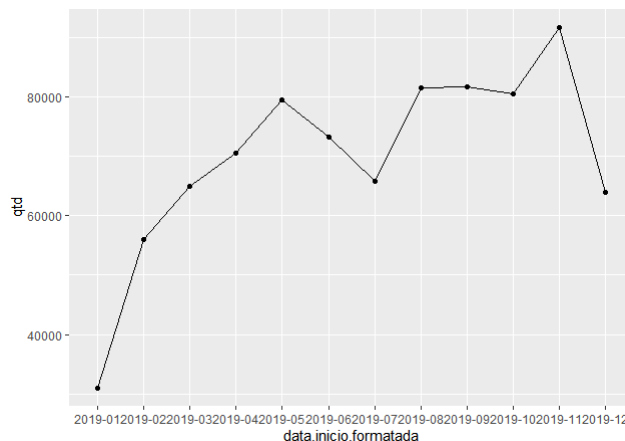


Figura 5 – Dispersão das médias dos valores gastos no ano de 2019

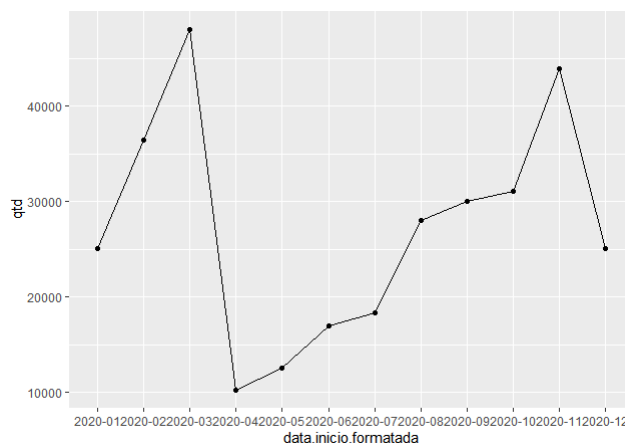


Figura 6 – Dispersão das médias dos valores gastos no ano de 2019

Conclusão

Com toda esta pesquisa e trabalho de laboratório realizado, compreendemos que trazer respostas é gerar dados. A informação é um dado! Entretanto, para que a resposta seja um dado ela precisa ser constituída em fatos. O fato é um aspecto conveniente e inteligível de um evento, ou seja, a forma como um evento se apresenta para a mente que o contempla. Através de toda formulação do fato é possível dar forma e estrutura a resposta (dado/informação). Logo, trouxemos respostas para as nossas perguntas, a necessidade de informação, reduz as incertezas este fluxo informacional onde aqui se enraizou de forma interdisciplinar, nas áreas de biblioteconomia, estatística, computação e por fim a ciência da informação com sua ênfase no dinâmico, no acesso à informação e a transferência da informação. Podemos dizer que nesta transferência ocorreu a modificação da consciência, algo que é pregado por alguns estudiosos da informação como (BARRETO, 2002) salienta em sua pesquisa "A condição da informação". Faz jus todo o universo de pesquisa aqui comentado e debatido além da possibilidade de abranger esta pesquisa, tratando e processando estes dados para maior compilamento e mineração, como tirar petróleo da vastidão marítima.

Referências

BARRETO, A. d. A. A condição da informação. *São Paulo em Perspectiva*, SciELO Brasil, v. 16, n. 3, p. 67–74, 2002. Citado na página 6.

DESCOBERTA de conhecimento e mineração de dados. In: PROCEDIMENTOS da oitava conferência internacional ACM SIGKDD sobre descoberta de conhecimento e mineração de dados, 23-26 de julho de 2002, Edmonton, Alberta, Canadá. [S.l.: s.n.]. Citado na página 3.

MACGARRY, K. *O contexto dinâmico da informação*. [S.l.]: Brasília: Briquet de Lemos, 1999. Citado 2 vezes nas páginas 2 e 3.

OLIVEIRA, A. C. S. de; SILVA, E. M. da. Ciência aberta: dimensões para um novo fazer científico. *Informação & Informação*, v. 21, n. 2, p. 5–39, 2016. Citado na página 3.