



Técnicas de visualização da informação aplicadas à produção de petróleo no Brasil

Cássio Antonio Andrade¹, Claudio Marcio Cassela Inacio Junior²

ICMC-USP

Afonso Paiva³

ICMC-USP

1 Introdução

Nos últimos tempos, com o advento a exploração do pré-sal, o Brasil passou a figurar entre os principais produtores de petróleo no cenário mundial [5]. Tal descoberta foi severamente importante, uma vez que esses poços possuem o potencial de produção dez vezes maior que os poços em pós-sal. Atualmente, o Brasil produz cerca de 3 milhões de barris por dia (Mbdp), sendo que o pré-sal representa mais do que 50% da produção total nacional [5].

Diversas técnicas de ciência de dados vêm sendo publicadas e desenvolvidas nos últimos anos no contexto petrolífero. No cenário nacional, Leal e Moura [3] realizaram uma mineração desses datasets e aplicaram técnicas de data analytics como forma de avaliar os dados. Em diferente abordagem, de Góes [2] utilizou a base de dados da ANP dos poços de petróleo para a criação de um banco de dados não-relacional que se conecta com um dashboard de um software de Business Intelligence (BI) com scatterplots georeferenciados, representações de Tagclouds, entre outros. Já no âmbito internacional, Cao et al. [1] propuseram uma diferente abordagem com gráficos de dispersão para uma avaliação mais abrangente dos dados de perfilagem de diferentes poços em diferentes camadas permitindo a visualização de dados heterogêneos multidimensionais. Plaisant et al. [6] utilizam as técnicas Treemap e SpaceTree para visualização de complexas estruturas geológicas em espaços tridimensionais. Por fim, é possível citar uma plataforma que utiliza técnicas multidimensionais de visualização combinada com métodos estatísticos para identificar, classificar, validar e interpretar propriedades geológicas [4].

Neste estudo, foi abordada uma análise completa dos datasets de produção de petróleo por poço, disponibilizados pela ANP, que vai desde a aquisição até a visualização de atributos. Dentre

¹andrade.cassio@usp.br

²claudio.inacio@usp.br

³apneto@icmc.usp.br

as ferramentas abordadas, optou-se pelo uso das técnicas de visualização scatterplots, mapas coropleto, séries temporais e *treemap*.

2 Metodologia

Os dados sobre os volumes produzidos mensalmente em cada poço, disponibilizados pela ANP, encontram-se em base mensal no formato de separação por vírgula (csv). No estudo em questão foram avaliados os dados de produção de janeiro de 2021 a maio de 2022. Para aquisição dos dados utilizou-se a biblioteca *requests*, que permite a aquisição dos dados através do canal de dados abertos da ANP. Como os dados estão disponibilizados no formato compactado (zip) foram aplicadas as funções disponíveis nas bibliotecas *zipfile* e *io* para descompactação dos arquivos e salvamento local. Os dados em csv foram processados, transformados e analisados no formato de dataframe utilizando as funções do pacote *pandas*.

Os dados são disponibilizados de acordo com o tipo de extração de petróleo: mar, terra e pré-sal. Sendo assim, os três conjuntos de dados foram agrupados em um único dataset e foi adicionada uma nova coluna que especifica a extração correspondente, resultando em uma base inicial com mais de 180 mil observações e 49 atributos. Além dos dados de produção, foi utilizada uma base com características de cada poço para utilização das coordenadas geográficas a fim de ilustrar a produção de cada poço em sua localização real conforme sua latitude e longitude, também fornecida no site da ANP.

Para cada visualização, foi preparada uma base específica por meio de junções das bases de dados, agregações, tratamento dos valores ausentes, entre outras técnicas de pré-processamento e exploração de dados.

3 Resultados

A técnica scatterplot relaciona as coordenadas geográficas latitude e longitude, o tipo de extração (mar ou pré-sal), definido pela cor do marcador, e a produção do petróleo mensal de cada poço, definida pelo tamanho do marcador. Já no mapa coropleto, a produção mensal total de petróleo brasileiro é representada pela coloração no interior dos limites territoriais do Brasil, sendo que quanto mais escuro o tom de azul, maior a produção de petróleo no país no mês em análise. A junção das duas técnicas em diferentes períodos de tempo, como ilustrado na Figura 1, permitiu relacionar diversos atributos e entender a evolução da produção de petróleo no Brasil.



Figura 1: Scatterplot e mapa coropleto.

Para análise da qualidade do petróleo, também foram utilizados scatterplots. Na Figura 2, a relação entre a quantidade produzida e a qualidade do petróleo, representado pelo atributo grau API, é apresentada. Verifica-se que há vários poços do tipo terra, mas a produção é inferior quando comparada aos tipos mar e pré-sal. É possível notar que o petróleo brasileiro possui um grau API, em média, entre 20 a 30, que caracteriza um petróleo do tipo pesado a médio. Entretanto, também nota-se a ocorrência de petróleos mais leves em alguns poços, com um grau API acima de 33. Também é possível verificar o comportamento da série de dados ao longo do tempo.

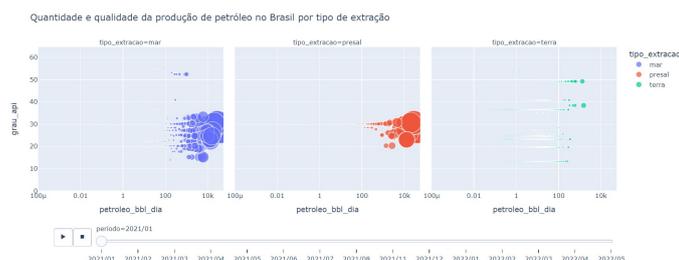


Figura 2: Análise da qualidade com scatterplot.

A Figura 3 representa a produção de petróleo em milhões de barris por dia por tipo de extração. Pela escala, é possível notar que o nível de produção via mar e pré-sal possuem magnitudes superiores à produção em terra. Além disso, nota-se que devido ao atual processo de desinvestimento de alguns poços de petróleo pela maior empresa petrolífera brasileira (Petrobras), os poços em terra estão em declínio de produção desde o início de 2021.



Figura 3: Time Series da produção de petróleo por tipo de extração.

A magnitude da produção por categorias é exibida na Figura 4a, ficando evidente a predominância da produção do estado Rio de Janeiro. Por fim, na Figura 4b, verifica-se a maior relevância da produção via mar, mas com importante participação do pré-sal, que tende a produzir cada vez mais nos próximos anos.

Outras visualizações e análises podem ser acessadas em <https://sites.google.com/usp.br/petrovis>.

4 Conclusões

Verificou-se uma curva de crescimento da produção de petróleo ao longo dos meses do período observado (jan/2021 a mai/2022), com uma alta concentração no litoral sudeste do país. Também



Figura 4: Treemap da produção de petróleo.

foi observado que o tipo de extração em terra, apesar de ter mais poços, representa uma pequena proporção na produção total quando comparado aos tipos mar e pré-sal. Avaliando a qualidade do petróleo, o tipo pré-sal apresentou a menor variabilidade com grau API em torno do valor 30. Entretanto, a média nacional está no intervalo entre 20 a 30, que classifica o petróleo como pesado a médio.

A queda observada da produção via terra pode estar relacionada ao processo de desinvestimento dos poços operados pela Petrobras. Sendo assim, possivelmente as produções em mar e pré-sal devem ser incentivadas no futuro próximo, já que as bacias de Santos e Campos possuem a predominância da produção no país, representadas pelos estados de RJ e SP.

Referências

- [1] M. Cao, Z. Gao, Y. Yuan, Z. Yan, and Y. Zhang, A Visualization and Analysis Method by Multi-Dimensional Crossplots from Multi-Well Heterogeneous Data, *Energies*, 15(7), 2575, 2022.
- [2] G. L. G. de Góes, Estruturação de Banco de Dados Não-Relacional para as informações de poços de petróleo das bacias sedimentares continentais brasileiras, Monografia (Bacharel em Engenharia de Petróleo), Escola Politécnica da Universidade de São Paulo, 2021.
- [3] A. B. Leal and T. R. da Silva Moura, Data analytics applied to the analysis of petroleum production in Brazil, *Brazilian Applied Science Review*, v. 5, n. 2, p. 818-835, 2021.
- [4] A. E. Mostafa, J. Cevolani, E. V. Brazil, E. Sharlin and M. C. Sousa, PetroVis: exploratory visualization for petrographic characterization. *IEEE VISWeek 2013, Poster and Extended Abstract, Electronic Conference Proceedings*, 2013.
- [5] E. Petersohn. Pre-Salt Super Play: Leading Brazil into the World's Top 5 Oil Suppliers, *AAPG Latin America And Caribbean Region Geoscience Technology Workshop*, 2019.
- [6] C. Plaisant, G. Chintalapani, C. Lukehart, D. Schiro and J. Ryan, Using Visualization Tools to Gain Insight Into Your Data, *SPE Annual Technical Conference and Exhibition, OnePetro* 2003