

## **Pipeline de dados para automação do cálculo da estimativa do Produto Interno Bruto do estado de Alagoas**

### **Clerivaldo Holanda dos Santos Junior**

Secretaria de Estado do Planejamento, Gestão e Patrimônio de Alagoas, Brasil

E-mail: clerivaldo.santos@seplag.al.gov.br

### **Juliana Carla da Silva Santos**

Secretaria de Estado do Planejamento, Gestão e Patrimônio de Alagoas, Brasil

E-mail: juliana.santos@seplag.al.gov.br

### **Roberson Leite Silva Junior**

Secretaria de Estado do Planejamento, Gestão e Patrimônio de Alagoas, Brasil

E-mail: roberon.leite@seplag.al.gov.br

### **Márcia Núbia Barbosa Lopes**

Secretaria de Estado do Planejamento, Gestão e Patrimônio de Alagoas, Brasil

E-mail: marcianubia.lopes@seplag.al.gov.br

## **Resumo**

O Produto Interno Bruto (PIB) é um indicador que visa mensurar os bens e serviços finais produzidos em determinado período de um estado ou país. Ele desempenha um papel crucial na avaliação do desempenho econômico de uma região, proporcionando informações essenciais para o planejamento e execução de políticas públicas. Sua divulgação é realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), que publica as Contas Regionais, porém com defasagem de dois anos, fornecendo informações detalhadas sobre a economia em diferentes áreas do país. O estado de Alagoas por meio da Superintendência de Informações e Cenário (SINC), da Secretaria de Estado do Planejamento, Gestão e Patrimônio (SEPLAG) realiza um esforço para antecipar o resultado do PIB e assim tomar decisões mais assertivas de acordo com a conjuntura econômica do estado. Porém, a metodologia de cálculo utilizada demanda muito tempo, esforço e dedicação, o que pode ocasionar em erros operacionais devido à necessidade de alimentação manual de diversas fontes de dados. Nesse contexto, a automação do cálculo da estimativa do PIB de Alagoas torna-se uma solução eficaz para reduzir esses possíveis erros operacionais. E com a implantação de um pipeline de dados, é possível o mapeamento das fontes relevantes de informação, coletar e alimentar automaticamente um banco de dados, permitindo uma abordagem mais ágil e precisa na obtenção dos dados necessários para o cálculo do PIB, e exibindo-as em um painel de visualização do *software Power BI* com o auxílio de linguagem de programação em *Python*, melhorando a forma em que as informações são apresentadas, possibilitando à equipe técnica se concentrar em tarefas mais estratégicas devido ao ganho de tempo. Como trabalho futuro, propõe-se a construção de uma plataforma para visualização de dados atualizados por meio de APIs ligadas ao banco de dados, melhorando a visualização da estimativa do PIB e eliminando o processo de automação de atualização no *Power BI* chegando próximo a uma solução em nuvem. Este estudo apresenta uma abordagem eficaz para a automação do cálculo da estimativa do PIB e demonstra como a ciência de dados pode ser aplicada na administração pública,

proporcionando uma melhor governança de dados e tomadas de decisões mais informadas e estratégicas.

**Palavras-chave:** Produto Interno Bruto; Pipeline de dados; Ciência de dados.

## 1. Introdução

Considerado por muitos como um dos principais termômetros da economia de um território, o Produto Interno Bruto (PIB) de uma localidade nada mais é do que a representação da totalidade dos valores monetários de bens e serviços finais produzidos em um determinado país, estado ou cidade, sendo usualmente calculado em um período de um ano. Os bens e serviços finais que compõem o PIB são avaliados levando em conta o preço de venda ao consumidor. Portanto, essa medida também incorpora os impostos incidentes sobre os produtos comercializados. (IBGE, s.d.)

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), para o cálculo do PIB são utilizadas diversas fontes de dados, como o resultado de pesquisas realizadas pelo próprio IBGE como também de fontes externas que mostram informações acerca dos setores da agropecuária, da indústria, e de serviços de uma forma geral, além das informações provenientes de impostos. Algumas das fontes são o Balanço de pagamentos, do Banco Central, Índice de Preços ao Produtor Amplo (IPA), da Fundação Getúlio Vargas (FGV), Produção Agrícola Municipal (PAM), Pesquisa Anual de Comércio (PAC), Pesquisa Anual de Serviços (PAS), Pesquisa Industrial Anual – Empresa (PIA-Empresa), do IBGE, além de outras fontes utilizadas.

Com isso, é possível realizar diversas análises a partir do resultado do PIB de um território, podendo ser feita uma comparação temporal de uma mesma localidade para observar se há um crescimento econômico ao longo da série histórica, comparações com outras localidades parecidas, a fim de analisar se o crescimento ou decréscimo encontrado para determinado território está acima ou abaixo da média dos outros locais, ou até mesmo analisar como seria a distribuição do PIB pela população se cada habitante obtivesse frações iguais das riquezas de determinado lugar, sendo denominado por PIB per capita, entre outros estudos técnicos. No entanto, o PIB não reflete fatores importantes como qualidade de vida ou distribuição de renda, mas ajuda a entender o contexto econômico que determinada localidade se encontra em um período predefinido. (IBGE, s.d.)

Dessa forma, o resultado encontrado para o Produto Interno Bruto mostra-se como um indicador eficaz para o planejamento e elaboração de políticas públicas, pois é a partir do conhecimento do contexto atual do país, estado ou cidade que é possível realizar atividades com mais assertividade. Porém, a disponibilização dessas informações pelo IBGE é feita com defasagem de dois anos. Ou seja, quando é tomada alguma decisão relacionada a algum acontecimento com base no resultado do PIB, este pode nem existir mais ou já estar controlado de uma forma menos invasiva.

Segundo Rebelo (2022), a previsão do Produto Interno Bruto (PIB) desempenha um papel fundamental na implementação de políticas econômicas, com especial ênfase nas políticas monetárias e orçamentárias. O PIB trimestral, por sua vez, é um indicador crucial que revela a tendência de curto prazo no desempenho da economia. Nessa linha o IPECE

(2023), assevera que esse indicador assume uma relevância significativa ao assinalar a propensão ao crescimento ou contração da economia.

A corroborar com a assertiva acima, o IPARDES (2018) discorre sobre a importância de a estimativa trimestral do PIB para acompanhamento da conjuntura, visto que alguns entes nacionais calculam e elaboram o Produto Interno Bruto Trimestral, que apresenta a evolução da economia com pouca defasagem de tempo, contribuindo para a tomada de decisões de governo, empresas e investidores no curto prazo. Os dados anuais das Contas Nacionais fornecem a estrutura de ponderação para o PIB trimestral, existindo um processo integrado entre o PIB anual e trimestral. Quando da divulgação do PIB anual, os resultados das contas trimestrais são ajustados aos resultados anuais. É importante, no entanto, ressaltar que as informações e resultados apresentados são preliminares e, portanto, estão sujeitos a retificações.

A compreensão do PIB e suas projeções têm um impacto importante no direcionamento das políticas econômicas, visto que fornece uma visão importante sobre o atual estado e direção futura da economia. Nesse contexto, a precisão dessas previsões é crucial para a tomada de decisões informadas pelos responsáveis pelas políticas públicas e pelos agentes econômicos.

E é a partir da necessidade de antecipar este resultado que a tentativa de calcular o PIB do estado de Alagoas mostra-se eficiente, pois com a análise do contexto atual e sua projeção para o ano é possível realizar ações para controle ou até mesmo incentivos na economia da região. Na Tabela 1 é possível observar o esforço do corpo técnico em calcular a estimativa de forma que seja aproximada ao valor apresentado pelo IBGE, mesmo com a defasagem de dois anos, ao longo de toda série histórica.

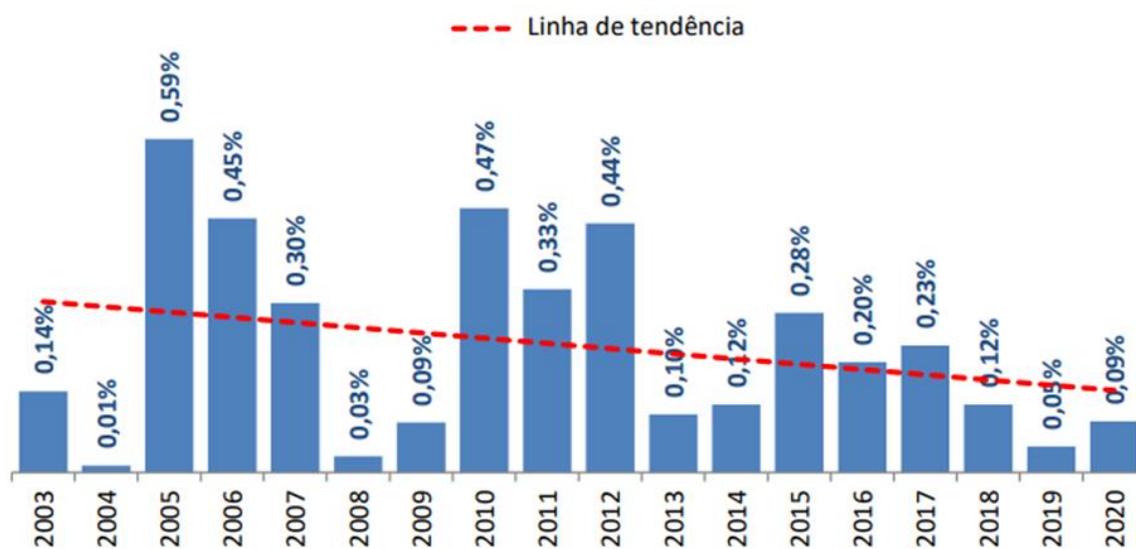
**Tabela 1 –** Comparativo da estimativa do PIB e o resultado das Contas Regionais (CR)

<b>ANO</b>	<b>SEPLAG</b>	<b>IBGE</b>
2003	-1,29%	-1,15%
2004	6,04%	6,03%
2005	2,97%	3,56%
2006	2,39%	2,84%
2007	4,94%	5,24%
2008	6,84%	6,81%
2009	0,87%	0,96%
2010	4,87%	5,34%
2011	4,33%	4,65%
2012	1,60%	2,05%
2013	0,49%	0,38%
2014	4,65%	4,77%
2015	-2,59%	-2,88%
2016	-1,15%	-1,35%
2017	3,56%	3,33%
2018	0,99%	1,11%
2019	1,90%	1,95%
2020	-4,32%	-4,23%

Fonte: IBGE/CONAC – SEPLAG/SINC

Com a Tabela 1, é possível observar que os resultados apresentados pelo órgão, mesmo que de maneira manual, conseguem alcançar resultados bem aproximados em relação aos que são apresentados pelo IBGE e que essa diferença vem diminuindo ao longo dos anos. Essa tendência de diminuição da diferença pode ser mais bem vista na Figura 1 a seguir.

**Figura 1** – Diferença em pontos percentuais entre os resultados da estimativa SEPLAG e CR considerando o valor absoluto do erro



Fonte: IBGE/CONAC – SEPLAG/SINC

Para que o cálculo da estimativa do PIB seja realizado, faz-se necessário a compreensão de como o Produto Interno Bruto é calculado além de quais fontes de dados são utilizadas oficialmente pelo IBGE, sendo subdividido em três setores, da agropecuária, da indústria, de serviço e os impostos, e a partir daí realiza-se o mapeamento de quais fontes de dados podem ser utilizadas, devido a algumas particularidades do estado de Alagoas, e então coletar e armazenar os dados obtidos em um banco de dados para que posteriormente sejam realizados os cálculos necessários para alcançar o resultado do PIB naquele período, seja trimestral, semestral ou anual. Com o resultado calculado, a apresentação deste resultado para a alta gestão faz com que sejam sinalizados possíveis ações, que possibilite a tomada de decisões baseadas em dados de maneira mais assertiva.

Todo o passo a passo descrito anteriormente, pode ser mais bem visualizado como um ciclo de vida dos dados, onde ele passa pelas etapas de coleta; armazenamento; análise e utilização; monitoramento e manutenção, além da visualização e comunicação da informação e sua posterior ação a partir dos dados obtidos ao final do ciclo. E, a partir de um processo bem estruturado, é possível obter resultados bem satisfatórios, desde que haja um monitoramento da qualidade desses dados e seus possíveis ajustes.

Apesar de entender o ciclo de vida de dados, a complexidade das etapas e o custo de operação por tempo de execução, se as atividades forem realizadas de maneira manual podem se tornar extremamente cansativas ocasionando em alguns erros operacionais, diminuindo a qualidade dos dados em si. Em contrapartida, um processo automatizado como o de um *pipeline* de dados, quando for possível sua utilização, reduz a ocorrência de erros

humanos além de realizar os cálculos de forma precisa e consistente garantindo a qualidade dos dados em seus diversos níveis de detalhamento possíveis.

Por meio da contribuição de um *pipeline* de dados para automação de etapas mais trabalhosas como obtenção, armazenamento de dados e seus respectivos cálculos e da contribuição humana para processos mais analíticos e decisórios, como as etapas de análise, monitoramento, comunicação e ação, consegue-se alcançar um nível alto de qualidade dos dados aliado a melhores decisões tomadas com base nas informações obtidas, sendo assim um ciclo de vida mais eficiente.

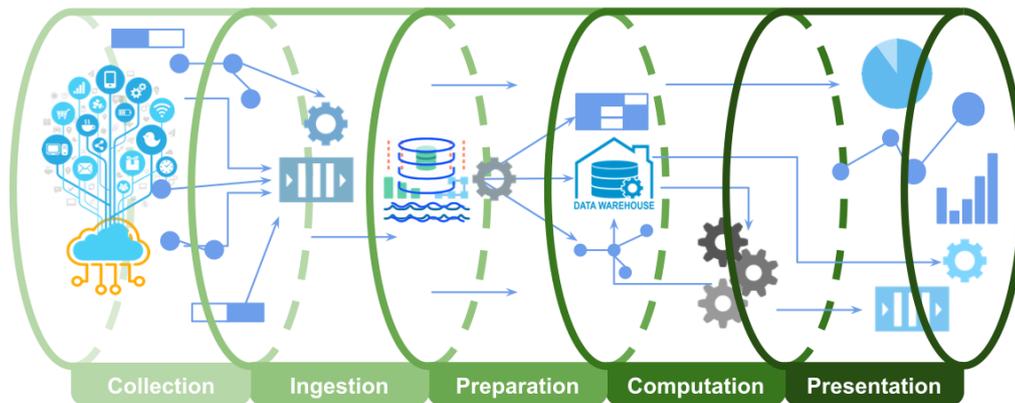
Sendo assim, este artigo é classificado com uma pesquisa quantitativa, pois Lakatos e Marconi (1985, apud CORDEIRO, 2014) afirmam que tal tipo de pesquisa ocorre com tudo que pode ser mensurável, transformando números em informações que podem transmitir um significado que podem ser analisados e também classificados. Além disso, a pesquisa se enquadra também como descritiva. De acordo com Sampieri, Collado e Lúcio (1998, apud CORDEIRO, 2014), este tipo de pesquisa tenta mensurar e comunicar como um acontecimento ocorre e seus respectivos impactos. Dessa forma, pelo fato de a pesquisa utilizar dados que podem ser mensuráveis, e por meio de uma modelagem estatística para então transmitir informações para uma tomada de decisão mais assertiva faz com que seja classificada nesse formato.

Com isso, os objetivos acerca desse artigo mostram-se importantes para descrever o processo de automação do cálculo da estimativa do PIB do estado de Alagoas através da alimentação de um banco de dados local e também da execução de cálculos para alcançar o resultado final deste indicador. Eliminando assim, possíveis erros operacionais de qualquer natureza, como erro de digitação ou a não atenção a detalhes como data de atualização de outros períodos, através do fluxo de automação elaborado. Conseguindo diminuir o grande esforço para manter esse fluxo atualizado manualmente, eliminando a necessidade da execução de tarefas repetitivas pelo corpo técnico para que se tenha uma dedicação maior aos aspectos econômicos, alcançando resultados com uma melhor qualidade além de decisões mais assertivas.

### 1.1. *Pipeline* de Dados

Segundo a Amazon Web Services (AWS), um *pipeline* de dados refere-se a uma sequência de etapas de processamento projetadas para preparar dados corporativos para análise. As organizações lidam com um vasto volume de dados provenientes de diversas fontes. No entanto, os dados brutos por si só são inúteis; eles precisam ser movidos, organizados, filtrados, reformados e analisados para se obter *insights* relevantes para que seja possível uma tomada de decisão. Um *pipeline* de dados abrange uma variedade de tecnologias que são utilizadas para verificar, resumir e identificar padrões nos dados, a fim de embasar as decisões de negócio. O fluxograma de um *pipeline* de dados pode ser mais bem visto na Figura 2 abaixo.

Figura 2 – Fluxograma de um pipeline de dados



Fonte: Artigo no LinkedIn “Pipeline de dados – O encanamento na sua empresa é de água ou de esgoto?”<sup>1</sup> (2020)

Na fase da coleta de dados é que são definidas as funções de cada dado e qual a fonte que disponibiliza tal informação. A partir disso, é feita a extração e captura e uma posterior validação e limpeza das informações obtidas para que sejam armazenadas somente as informações que sejam necessárias de acordo com o mapeamento das funções dos dados feitos anteriormente. Após isso, o ciclo segue com o armazenamento dos dados em um sistema de gerenciamento de banco de dados que suporta o tipo de informação utilizada, além da sua transformação, caso seja necessário, além da integração com outras bases de dados coletadas no passo anterior.

Com todos os dados necessários armazenados, segue-se o ciclo com a análise dos dados. É aqui que os dados são explorados mais detalhadamente, além da aplicação de técnicas de análise de dados necessárias para a identificação de padrões e *insights* significativos. Concomitante a isso, tem-se o passo de monitoramento, para que seja possível analisar a qualidade dos dados, e uma possível atualização e correção caso seja observada a necessidade de acordo com o próprio dado obtido ou mudança do comportamento a ser analisado.

Com os resultados obtidos na etapa de análise, segue-se com a etapa de visualização e comunicação. É nessa etapa que os resultados são mais bem transmitidos de uma forma simples e objetiva de maneira que quem tem o poder de decisão veja de forma clara e concisa as informações necessárias para auxiliar de forma mais assertiva na tomada de decisões, seja por meio de relatórios ou painéis interativos.

Sendo assim, os *pipelines* de dados proporcionam uma série de benefícios, com destaque para o processamento eficiente e a automação de tarefas. Ao automatizar as etapas de transformação e carga de dados, a equipe de dados pode concentrar-se na descoberta de *insights* de negócios valiosos, ao invés de realizar tarefas repetitivas. Além disso, esses *pipelines* permitem processar os dados brutos de forma mais rápida, garantindo

<sup>1</sup> Disponível em: <https://www.linkedin.com/pulse/pipeline-de-dados-o-encanamento-na-sua-empresa-%C3%A9-%C3%A1gua-lobes/>. Acesso em jul. de 2023.

que não percam seu valor ao longo do tempo. Essa abordagem eficiente e automatizada disponibilizada pelo *pipeline* de dados facilita a análise e contribui para uma tomada de decisão mais informada e precisa (AWS, s.d.).

## 2. Metodologia Aplicada

A metodologia utilizada seguiu a partir da aplicação das etapas do ciclo de vida dos dados e da automação de algumas de suas etapas de forma estratégica para que houvesse um equilíbrio entre ganho de tempo com a automação de etapas repetitivas e desgastantes além de usar a capacidade analítica da equipe técnica da Secretaria de Estado do Planejamento, Gestão e Patrimônio de Alagoas (SEPLAG) para poder analisar e melhorar as informações utilizadas no cálculo da estimativa do PIB do estado, garantindo assim a qualidade dos dados e do resultado final.

### 2.1. Pipeline de dados

Sabendo-se que o ciclo é iniciado com a etapa de coleta, mais precisamente com o entendimento do problema, foi aqui que começou o mapeamento das fontes de dados com as atividades de cada setor que tivessem uma relevância maior na economia do estado de Alagoas. Após levantar ideias acerca de evidências, foram listadas fontes de dados que mostrassem dados sobre a produção agropecuária, como alguns dados públicos disponibilizados pelo IBGE, para o setor agropecuário.

Para o setor industrial e de serviços, por não existir pesquisas que disponibilizassem os dados para o estado por diversos motivos, seja pelo custo de realização destas serem relativamente caras, ou pelo quantitativo do estado não ser tão relevante quanto ao dos estados vizinhos. Então, foram buscadas fontes de dados que conseguissem evidenciar o desenvolvimento destes setores de uma forma mais aproximada.

Para isso, foram utilizados dados de pessoal ocupado em alguns setores relevantes economicamente para o estado, obtidos através dos dados disponibilizados pelo Cadastro Geral de Empregados e Desempregados (CAGED) e também pelo Cadastro Central de Empresas (CEMPRE). Além disso, o consumo de água, energia e gás se mostraram como identificadores de padrões que podiam influenciar no resultado da estimativa do PIB, pois se há uma atividade industrial ou de algum outro serviço em crescimento, o consumo desses recursos também demonstraram um aumento, além de também conseguir demonstrar em um cenário de decréscimo. Então, buscou-se fontes de dados como os da Empresa de Pesquisa Energética (EPE), como também dados internos da Companhia de Saneamento de Alagoas (CASAL), e também da BRK Ambiental, por conta da privatização de algumas atividades da CASAL.

Para dados relacionados à saúde, a fonte de dados do Departamento de Informática do Sistema Único de Saúde (DataSUS) mostrou-se eficaz para a identificação de tendências quanto à quantidade de procedimentos eletivos e hospitalares, que após uma análise do corpo técnico foi possível observar que interferiam no cálculo da estimativa. Além disso, fontes como a Receita Federal, Banco Central e do Conselho Nacional de Política Fazendária (CONFAZ), contribuíram para a aquisição de dados relacionados aos impostos

do estado. Outras fontes foram necessárias para algumas outras particularidades da região, fazendo-se uso de fontes públicas e privadas para alcançar o resultado final.

Após o mapeamento de todas as fontes, antes mesmo de iniciar a coleta de dados, tentou-se também mapear a periodicidade da disponibilização das informações na fonte para que tivesse um melhor acompanhamento e não ter um custo operacional desnecessário relacionado à programação, como em casos em que havia a necessidade de verificação dia a dia da disponibilidade da informação pela fonte. Em alguns casos foi possível determinar a periodicidade, em outros foram necessários estabelecer uma linha de corte de forma empírica devido à correção e atualização dos dados que já tinham sido disponibilizados, a fim de que não prejudicasse o cálculo da estimativa.

Após todo o processo supracitado, deu-se início com a coleta de dados propriamente dita. Foram analisadas quais as fontes de dados era possível a automação de uma forma simples para a criação de rotinas realizadas em linguagem de programação em *Python* para a obtenção de dados facilitadas por meio de uma Interface de Programação de Aplicações (API) da própria fonte de dados, ou quando não existe essa ferramenta, o uso de técnicas de *web scraping* e *web crawling* para a simulação de rotinas como se fosse uma pessoa a obter os dados que por vezes estão em uma página de internet ou até mesmo no *download* de PDFs. Além disso, para fontes de dados internas, não há como obter as informações através de automação, devido a sigilosidade dos dados, porém há poucas informações provenientes desse tipo de fonte, fazendo com que seja necessária a alimentação manual de uma planilha denominada parâmetros, as quais são lidas no momento de cálculo da estimativa.

Há também o problema relacionado às atualizações de dados, como exemplo citamos o IBGE como uma das fontes que informa a data de atualização. Para casos onde não há essa informação, a definição da linha de corte temporal para comparação torna-se uma alternativa muito eficiente, para verificar se houve ou não atualização da base.

Com todas as informações desejadas obtidas, fez-se necessária a transformação por meio da estruturação desses dados, que vieram em diversos formatos, tais como planilhas e PDFs, como também a extração de algumas informações relevantes para o cálculo da estimativa do Produto Interno Bruto. Esta etapa foi feita também com o auxílio de linguagem de programação em *Python*, devido à alta disponibilidade de bibliotecas para que sejam possíveis essas transformações de uma forma mais simplificada.

Após todo o processo de transformação com a estruturação dos dados, foi feita a inserção dos dados em um sistema de gerenciamento de banco de dados local nas dependências da própria SEPLAG, e que, por conta das informações que são armazenadas, necessitam de um nível de segurança maior, sendo assim, foi a melhor decisão a ser tomada para esse tipo de armazenamento. Além disso, o formato do banco de dados foi seguido como um *Data Warehouse*, pela necessidade de armazenamento e utilização somente de dados estruturados para a realização desse projeto. Todas as informações armazenadas seguiram técnicas de organização para que no futuro não houvesse dificuldades nas consultas das mesmas, subdividindo as informações entre os setores e fontes de dados em *Schemas* e Tabelas para que houvesse uma fácil identificação posterior.

Após todas as informações terem sido armazenadas, seguiu-se o ciclo com a etapa de análise de dados com a aplicação dos cálculos. Esta etapa também foi realizada com linguagem de programação em *Python*, e o passo a passo a ser seguido foi reproduzir os cálculos que existiam em uma planilha, mas que era manipulada completamente de forma

manual e que apresentava uma suscetibilidade maior para erros por conta de atualização de dados e replicação de cálculos devido à alta quantidade de informações.

Ao iniciar o processo de cálculo, fez-se necessário a consulta ao banco de dados e ao manual, além das informações sigilosas da SINC para a atribuição de pesos referentes a cada atividade. Esse processo se deu para cada setor, quais sejam: agropecuária, indústria e serviços, além de impostos. Ademais a subsetor do comércio, visto sua singularidade, necessitou sua desagregação do setor de serviços. De onde sucedem cinco processos de cálculo independentes, incluindo também os impostos, o qual posteriormente comporá o cálculo da estimativa do produto interno bruto do estado de Alagoas.

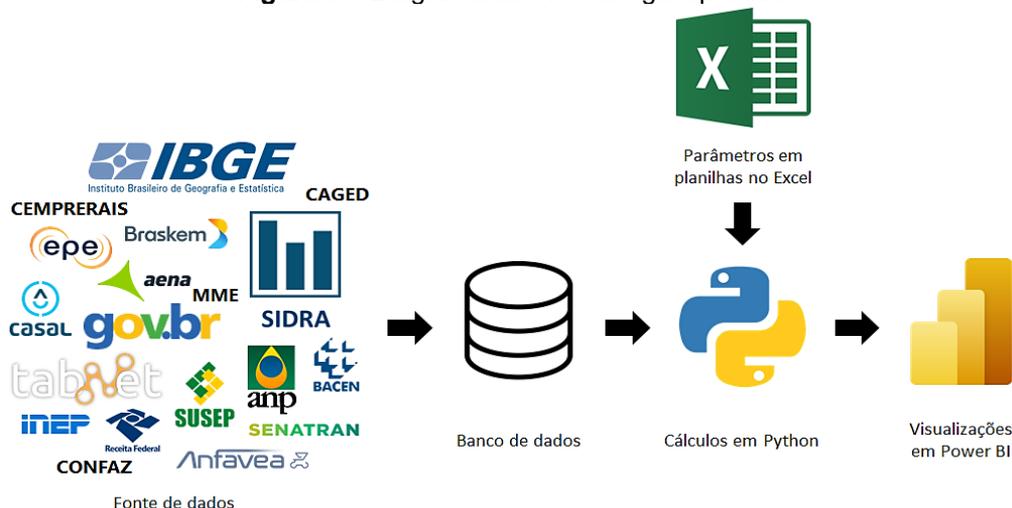
Com todos os cálculos realizados, a visualização de cada etapa de cálculo faz-se importante para obter um nível de observação mais detalhado e como consequência poder analisar melhor que atividade específica está apresentando tal comportamento. Toda a etapa de visualização dos cálculos para cada setor foi realizada com o auxílio do *software* de visualização de dados *Power BI*, com cada página possuindo uma visualização mais detalhada de cada etapa do cálculo e também para outros períodos de tempo, tais como anual, trimestral, semestral e acumulado de 12 meses. Ao final, há um painel interativo centralizador para o acompanhamento do resultado da estimativa do PIB final.

A etapa de comunicação das informações é feita por meio da análise do cálculo da estimativa para determinado período e disponibilizado em boletins conjunturais na plataforma Alagoas em Dados e Informações, a demonstrar o cenário atual da economia estadual. Posteriormente, há a publicação da estimativa do Produto Interno Bruto final antecipando o resultado em aproximadamente um ano da publicação oficial do IBGE. Apesar disso, por conta da disponibilidade dos dados na fonte serem menores que um ano, é possível antecipar até mesmo o PIB do ano corrente, sendo esta informação interna e que pode auxiliar a uma tomada de decisão por parte da alta gestão.

Após a publicação oficial do resultado pelo IBGE, é realizada a verificação dos cálculos para garantir a precisão das informações obtidas e se estão dentro do intervalo de confiança predefinido na publicação oficial do órgão. Em alguns casos, há a necessidade de usar um fator de ajuste tendo em vista que o Produto Interno Bruto das outras Unidades da Federação somado ao de Alagoas resulte no PIB do Brasil, o que faz com que possam resultar em divergências no valor estimado entre uma publicação e outra, daí a importância de determinar um intervalo de possíveis valores e não um valor em específico.

Com isso, todo o processo da metodologia que foi aplicada pode ser mais bem visualizado no diagrama da Figura 3 a seguir.

Figura 3 – Diagrama da metodologia aplicada



### 3. Resultados e Discussões

Com a metodologia aplicada para a automação do cálculo da estimativa do Produto Interno Bruto, foi possível alcançar o mesmo resultado ao que era feito de forma manual pelo corpo técnico da SINC/SEPLAG somente para o setor de atividade agropecuária. A planilha de cálculo utilizada que demonstrava o passo a passo do cálculo foi replicada em toda sua série histórica e suas respectivas projeções, sendo a estimativa do valor adicionado da agropecuária de 2022 e também de 2023, devido a disponibilização dos dados pelas fontes utilizadas.

Figura 4 – Painel interativo na página de acompanhamento dos produtos do setor agropecuário

PIB VBP AGROPECUÁRIA																	
PAM E PPM																	
Ano	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomate
2021	73145	5138	24436	7900345	112404	40494	1321236	17419082	99612	13403	21691	138990	652835	508652	81642	39591	8990
2020	69646	5276	16072	8823684	112394	42027	1290414	15293563	97994	10735	20725	140088	615297	481553	61097	35456	7799
2019	71598	5316	20177	8532961	104048	38013	1233394	18702251	92395	9802	19401	142324	603808	384152	54122	33481	8720
2018	64672	3696	19841	8243360	104910	32560	1247619	16738990	77498	8257	5411	173764	587299	378203	25421	32395	7543

LSPA E PESQUISAS TRIMESTRAIS																	
Ano	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomate
2023	50819	4688	23862	140596	120494	40450	1254938	18747990	108188	27328	16980	113223	6913	583763	145675	1846	8578
2022	76951	4607	16971	140596	111691	44014	1254938	16870542	107017	9512	12323	132157	6638	641383	58368	1821	9303
2021	76012	4999	24388	140596	115308	41239	1176405	17990900	100200	16464	21492	137049	5865	527181	70490	1750	8825
2020	74084	5323	16020	140596	109361	42267	1106961	15241840	99113	10681	19973	140673	5417	535365	61283	1847	7432

VARIAÇÃO PERCENTUAL LSPA																	
Ano	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomate
2023	0,66	1,02	1,41	1,00	1,08	0,92	1,00	1,11	1,01	2,87	1,38	0,86	1,04	0,91	2,50	1,01	0,92
2022	1,01	0,92	0,70	1,00	0,97	1,07	1,07	0,94	1,07	0,58	0,57	0,96	1,13	1,22	0,83	1,04	1,05
2021	1,05	0,97	1,52	0,90	1,00	0,96	1,02	1,14	1,02	1,25	1,05	0,99	1,06	1,06	1,34	1,12	1,15
2020	0,97	0,99	0,80	1,03	1,08	1,11	1,05	0,82	1,06	1,10	1,07	0,98	1,02	1,25	1,13	1,06	0,89

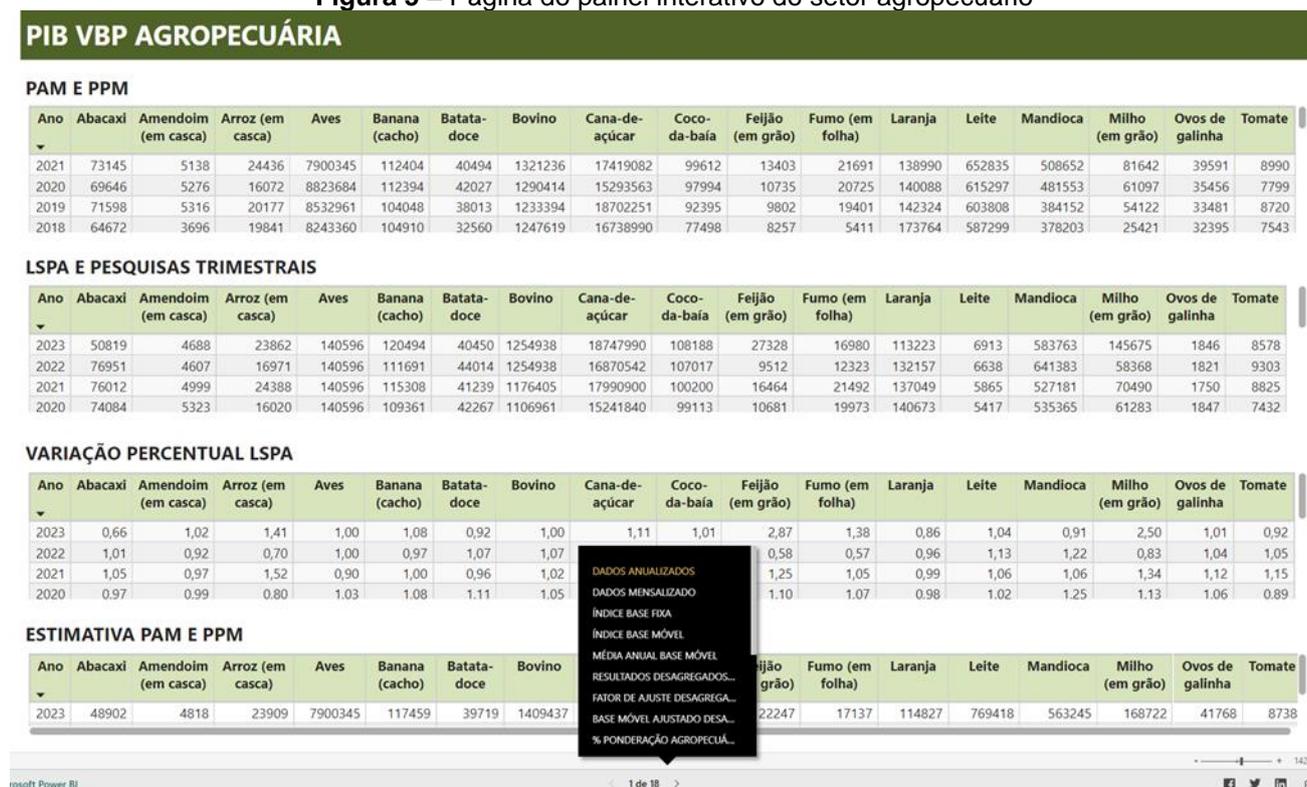
  

ESTIMATIVA PAM E PPM																	
Ano	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomate
2023	48902	4818	23909	7900345	117459	39719	1409437	18152109	107553	22247	17137	114827	769418	563245	168722	41768	8738

Fonte: Autores (2023)

Com a Figura 4, é possível observar que as informações aparecem de forma simples e precisa. Na Figura 5, observamos que elas se apresentam num total de 18 páginas contando a etapa de cálculo da estimativa do setor agropecuário, fazendo com que exista a possibilidade de uma análise mais detalhada de cada etapa de cálculo.

Figura 5 – Página do painel interativo do setor agropecuário



Fonte: Autores (2023)

Além disso, com a construção do painel, é possível observar o desempenho em diversos períodos determinados pela equipe, como anual, semestral, trimestral e acumulado de 12 meses. Todas essas formas de visualização fazem com que haja uma análise melhor do que somente a apresentada pelo IBGE, que além de ter a defasagem de dois anos na divulgação, seu resultado só é apresentado no final do ano.

E é sabendo como está o desempenho do Valor Adicionado, nesse caso, somente do setor agropecuário, ao longo do trimestre ou semestre, por exemplo, que faz com que possam ser tomadas decisões mais assertivas devido ao melhor conhecimento do contexto econômico do estado nesse aspecto.

Figura 6 – Abas do painel interativo com a variação percentual anual e acumulado de 12 meses

## ACUMULADO ANUAL

TABELA COM OS RESULTADOS DO ACUMULADO ANUAL EM RELAÇÃO AO MESMO PERÍODO DO ANO ANTERIOR  
VALOR DE PRODUÇÃO

### VARIÇÃO PERCENTUAL

Ano	Trimestre	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomat
2023	1	-33,96	1,76	40,60	0,00	7,88	-8,10	0,00	11,13	1,09	187,30	37,79	-14,33	4,14	-8,98	149,58	1,41	-7,7
2023	2	-33,96	1,76	40,60	0,00	7,88	-8,10	0,00	11,13	1,09	187,30	37,79	-14,33	4,14	-8,98	149,58	1,41	-7,7
2023	3	-33,96	1,76	40,60	0,00	7,88	-8,10	0,00	11,13	1,09	187,30	37,79	-14,33	4,14	-8,98	149,58	1,41	-7,7
2023	4	-33,96	1,76	40,60	0,00	7,88	-8,10	0,00	11,13	1,09	187,30	37,79	-14,33	4,14	-8,98	149,58	1,41	-7,7
2022	1	1,24	-7,84	-30,41	0,00	-3,14	6,73	6,68	-6,23	6,80	-42,23	-42,66	-3,57	13,18	21,66	-17,20	4,03	5,4
2022	2	1,24	-7,84	-30,41	0,00	-3,14	6,73	6,68	-6,23	6,80	-42,23	-42,66	-3,57	13,18	21,66	-17,20	4,03	5,4
2022	3	1,24	-7,84	-30,41	0,00	-3,14	6,73	6,68	-6,23	6,80	-42,23	-42,66	-3,57	13,18	21,66	-17,20	4,03	5,4
2022	4	1,24	-7,84	-30,41	0,00	-3,14	6,73	6,68	-6,23	6,80	-42,23	-42,66	-3,57	13,18	21,66	-17,20	4,03	5,4
2021	1	5,02	-2,62	52,04	-10,46	0,01	-3,65	2,39	13,90	1,65	24,85	4,66	-0,78	6,10	5,63	33,63	11,66	15,2
2021	2	5,02	-2,62	52,04	-10,46	0,01	-3,65	2,39	13,90	1,65	24,85	4,66	-0,78	6,10	5,63	33,63	11,66	15,2

## 12 MESES

TABELA COM OS RESULTADOS DO ACUMULADO DOS 12 MESES EM RELAÇÃO AO MESMO PERÍODO DO ANO ANTERIOR  
VALOR DE PRODUÇÃO

### VARIÇÃO PERCENTUAL

Ano	Trimestre	Abacaxi	Amendoim (em casca)	Arroz (em casca)	Aves	Banana (cacho)	Batata-doce	Bovino	Cana-de-açúcar	Coco-da-baía	Feijão (em grão)	Fumo (em folha)	Laranja	Leite	Mandioca	Milho (em grão)	Ovos de galinha	Tomat
2023	1	-7,64	-5,59	-17,04	0,00	-0,45	2,84	4,92	-2,09	5,30	-5,16	-29,75	-6,19	10,70	12,82	18,88	3,36	1,5
2023	2	-16,47	-3,24	-1,27	0,00	2,28	-0,93	3,23	2,17	3,85	41,82	-13,34	-8,85	8,38	4,84	58,35	2,70	-1,2
2023	3	-25,24	-0,79	17,60	0,00	5,06	-4,57	1,59	6,58	2,45	103,32	8,21	-11,56	6,19	-2,39	101,71	2,05	-4,6
2023	4	-33,96	1,76	40,60	0,00	7,88	-8,10	0,00	11,13	1,09	187,30	37,79	-14,33	4,14	-8,98	149,58	1,41	-7,7
2022	1	4,04	-3,90	24,31	-8,06	-0,78	-1,13	3,48	8,36	2,96	5,14	-7,58	-1,48	7,95	9,80	17,97	9,59	12,5
2022	2	3,08	-5,19	2,30	-5,52	-1,56	1,44	4,56	3,18	4,25	-12,39	-19,54	-2,17	9,74	13,86	4,56	7,64	9,5
2022	3	2,15	-6,51	-15,59	-2,84	-2,35	4,06	5,62	-1,67	5,53	-28,09	-31,23	-2,87	11,48	17,82	-7,05	5,79	7,6
2022	4	1,24	-7,84	-30,41	0,00	-3,14	6,73	6,68	-6,23	6,80	-42,23	-42,66	-3,57	13,18	21,66	-17,20	4,03	5,4
2021	1	-0,83	-1,22	-5,16	-0,15	5,90	6,73	4,05	-11,35	5,35	13,62	4,28	-1,38	2,97	28,20	18,56	7,40	-4,6
2021	2	1,10	-1,68	11,75	-3,64	3,86	3,10	3,48	-3,77	4,08	17,53	4,41	-1,19	4,02	19,46	23,88	8,86	1,6

Fonte: Autores (2023)

Como a metodologia de automação do cálculo ainda está sendo aplicada é esperado que os setores da indústria, de serviços, comércio e os impostos, além do resultado final da estimativa do PIB do estado, tenham os seus resultados a convergir para o verificado para o setor agropecuário na replicação do cálculo das planilhas utilizadas. A não realização da replicação para estes setores se dá pelo fato de que o método de automação do cálculo apresentado é uma proposta recente. Ela, todavia, se mostra promissora tendo em vista o resultado alcançado até o momento, de tal maneira na implantação do método automatizado como a aproximação dos resultados apresentados pela SEPLAG e os divulgados pelo IBGE.

Por ainda estar em fase de implantação, a partir do momento em que o resultado do cálculo do Produto Interno Bruto para o estado de Alagoas esperado for alcançado com o método de automação por meio de um *pipeline* de dados, o acompanhamento se faz necessário para que seja observado se existe ou não a ocorrência de divergências entre os dois métodos. Sendo assim, na medida em que se observar a convergência dos valores calculados de maneira usual com os resultados provenientes do método de automação, tenderá a uma função mais voltada para verificação dos valores de ambas as abordagens.

Com a descontinuação do método anterior, já com a replicação do cálculo automatizado realizada e monitorada, também é esperado que sejam alcançado níveis maiores de detalhamento. Caso seja observado pela equipe que há a necessidade desse

acompanhamento, tendo em vista que as novas atribuições serão mais focadas em etapas mais analíticas, como para a revisão metodológica do cálculo da estimativa do Produto Interno Bruto em si, além da possibilidade de uma melhor visualização, seja por meio de gráficos ou outras formas observadas, além da possibilidade de tomadas de decisão mais assertivas com base nos dados apresentados em cada etapa pelo painel interativo.

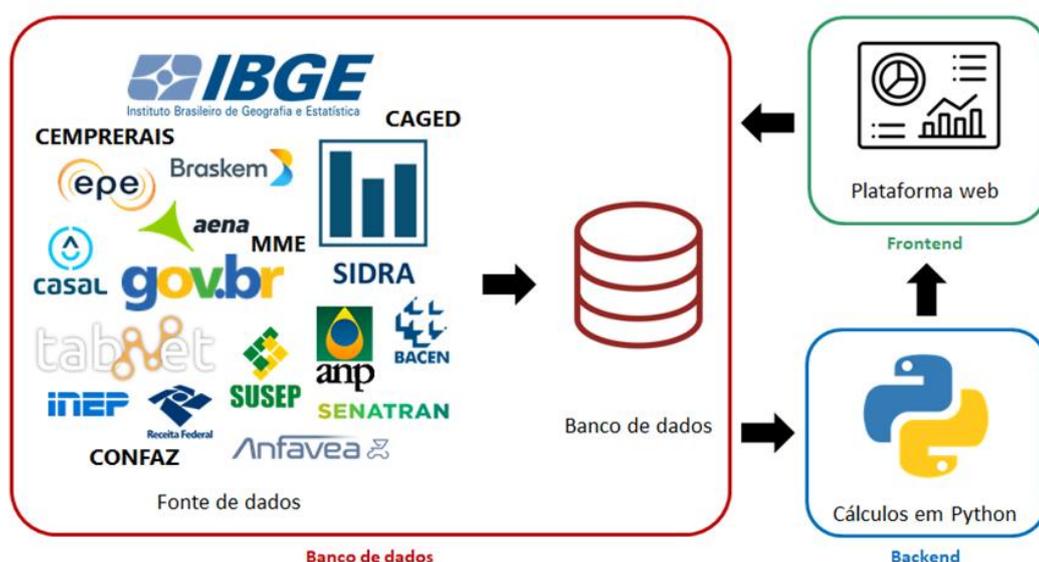
Atualmente, ainda há a necessidade de manipulação dos dados de forma manual em alguns momentos pontuais como na atribuição de pesos, além do fator de ajuste para alguns produtos e atividades, seja pelo fato da sigiliosidade dos dados ou pela forma em que são adquiridos através das fontes de dados. Há outro ponto a ser observado, como a aquisição de dados em algumas fontes que não permitem que tenham processos de automatização quanto ao período, por não ser possível saber qual a periodicidade de sua disponibilização, fazendo com que seja necessária a execução manual do código quando for percebido que houve uma atualização e assim alimentar o banco de dados utilizado.

Outro momento que exige uma atividade manual é a atualização dos dados no painel interativo do *Power BI*, onde é necessária a ativação de um botão para realização dos cálculos e a respectiva visualização de cada etapa. É considerada como uma atividade de pouco impacto de realização, mas pode ser avaliada como um gargalo ao travar o fluxo de todo o processo que deveria ocorrer de maneira fluída. Apesar disso, o ganho de tempo obtido através da automação de tarefas repetitivas citadas pela equipe técnica apresenta grande relevância podendo concentrar-se em atividades mais analíticas.

Em longo prazo, com a continuação do trabalho, poderá ser desenvolvida uma plataforma com o *backend* acessando os dados do banco de dados, e o *frontend* fazendo a parte visual por meio de *dashboards* e alimentação de dados manuais e também informações sigilosas diretamente no banco de dados, sendo ambas as comunicações entre *backend* e *frontend* realizadas através de uma API.

Seguindo a sugestão supracitada, o fluxo de automação ficaria como o diagrama abaixo.

Figura 7 – Diagrama com o fluxo proposto



Fonte: Autores (2023)

Com isso, os processos de coleta e alimentação do banco de dados automatizados, as consultas e cálculos feitos com o auxílio de linguagem de programação em *Python*, geram os dados para que sejam mais bem vistos em uma plataforma web, sendo que para esta plataforma deve-se ter a possibilidade de alimentação do banco de dados com os dados que de qualquer forma devem ser inseridos manualmente, tendo assim um produto mais perto de uma solução em nuvem e também com todos os objetivos deste artigo alcançados.

#### 4. Considerações Finais

Tendo em vista o resultado alcançado, utilizando como comparação o setor da agropecuária entre o método manual por meio de planilhas e o método automatizado por meio do *pipeline* de dados, foi possível alcançar o mesmo resultado final para ambos, tendo sido observado um ganho de tempo muito grande do segundo método em relação ao primeiro, levando-se em consideração que não houve a necessidade de atualização dos dados na base e também dos seus respectivos ajustes de cálculos. Com isso, é esperado que para os outros setores, indústria, serviço, comércio e impostos, além do resultado final da estimativa do Produto Interno Bruto alagoano, sejam alcançados com o mesmo sucesso.

Desta forma, o *pipeline* de dados mostra-se como uma alternativa eficaz para a redução de esforço em tarefas repetitivas que demandam muita mão de obra. Como consequência, o ganho de tempo da equipe técnica pode ser mais bem utilizado para adequação do método caso haja mudança na base de dados ou na possibilidade de acréscimo de outras informações que sejam necessárias, sabendo-se que o esforço aumentaria pouco com a inclusão de mais níveis de detalhamento no cálculo da estimativa do PIB do estado. Além disso, a utilização de painéis interativos do *Power BI* permite que haja uma melhor visualização dos dados ao invés da utilização de planilhas para uma posterior manipulação e criação de gráficos, garantindo economia de tempo também nessa etapa.

Em contrapartida ao sucesso esperado do método aplicado, algumas limitações por parte da obtenção de dados em algumas fontes que não disponibiliza seus dados de uma forma mais integrada, como por meio de APIs, faz com que ainda haja um esforço manual neste método. Porém, não diminui o objetivo alcançado na redução do esforço da equipe nas tarefas repetitivas.

Como melhorias contínuas o próximo passo é o desenvolvimento de uma plataforma que já faça a extração de dados e o armazenamento em um banco de dados com os cálculos da estimativa no *backend*, e *dashboards* de forma interativa no *frontend* fazendo com que o método proposto se aproxime de uma solução em nuvem, acoplado ainda mais benefícios, como a eliminação de uma planilha de parâmetros que poderá ser alimentada no *frontend* da plataforma e armazenada no próprio banco de dados, o que fará ser possível acessar, editar e visualizar de qualquer lugar aonde quem tem o poder de tomar decisões estiver.

## Referências

IBGE. Produto Interno Bruto - PIB, [s.d.]. Disponível em: <https://www.ibge.gov.br/explica/pib.php>. Acesso em: 14 jul. 2023.

IPECE, PIB Trimestral - Ano 2023. Disponível em [https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2023/06/APRESENTACAO\\_PIB\\_1o\\_TRIM2023.pdf](https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2023/06/APRESENTACAO_PIB_1o_TRIM2023.pdf). Acesso em: 27 jul. 2023.

IPARDES. Contas Regionais Trimestrais do Paraná: Metodologia, 2018. Disponível em [https://www.ipardes.pr.gov.br/sites/ipardes/arquivos\\_restritos/files/documento/2022-01/METODOLOGIA\\_PIB\\_2018.pdf](https://www.ipardes.pr.gov.br/sites/ipardes/arquivos_restritos/files/documento/2022-01/METODOLOGIA_PIB_2018.pdf) . Acesso em: 27 jul. 2023.

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da Ciência da Informação. **Informação & Informação**, [S. l.], v. 21, n. 2, p. 116–142, 2016. DOI: 10.5433/1981-8920.2016v21n2p116. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/27940>. Acesso em: 17 jul. 2023.

AWS. O que é um pipeline de dados? - Explicação sobre pipeline de dados - AWS, [s.d.]. Disponível em: <https://aws.amazon.com/pt/what-is/data-pipeline/>. Acesso em: 19 jul. 2023.

LOPES, J. M.. Pipeline de Dados - O encanamento na sua empresa é de água ou de esgoto?. **LinkedIn**, 2020. Disponível em: <https://www.linkedin.com/pulse/pipeline-de-dados-o-encanamento-na-sua-empresa-%C3%A9-%C3%A1gua-lobes/>. Acesso em: 19 jul. 2023.

REBELO, M. A. C. Interpretabilidade de modelos Machine Learning para a previsão macroeconómica. 2022. Tese de Doutorado. Instituto Superior de Economia e Gestão.

SANTOS JUNIOR, C. H.; SANTOS, J. C. S.; EMERY, T. M. R. L.; LOPES, M. N. B.; SILVA JUNIOR, R. L.. Intervalo de Confiança para a Estimativa do Produto Interno Bruto (PIB) do Estado de Alagoas. **Alagoas em Dados e Informações**, 2023. Disponível em: <https://dados.al.gov.br/catalogo/dataset/notas-tecnicas-estudos/resource/f0d8efa1-8ee5-4eab-b5ad-608e98bf4a00>. Acesso em: 19 jul. 2023.

CORDEIRO, R. A.; SANCHES, P. L. B.; CAVALCANTE, K. de O.; PEIXOTO, A. F.; LEITE, J. C. de L. Quantitative Research in Finance: An Analysis of Statistics Techniques Used by Articles Published in Qualified Scientific Journals from 2007 to 2009. **Revista de Administração da UFSM**, [S. l.], v. 7, n. 1, p. 117–134, 2014. DOI: 10.5902/198346595851. Disponível em: <https://periodicos.ufsm.br/reaufsm/article/view/5851>. Acesso em: 21 jul. 2023.