

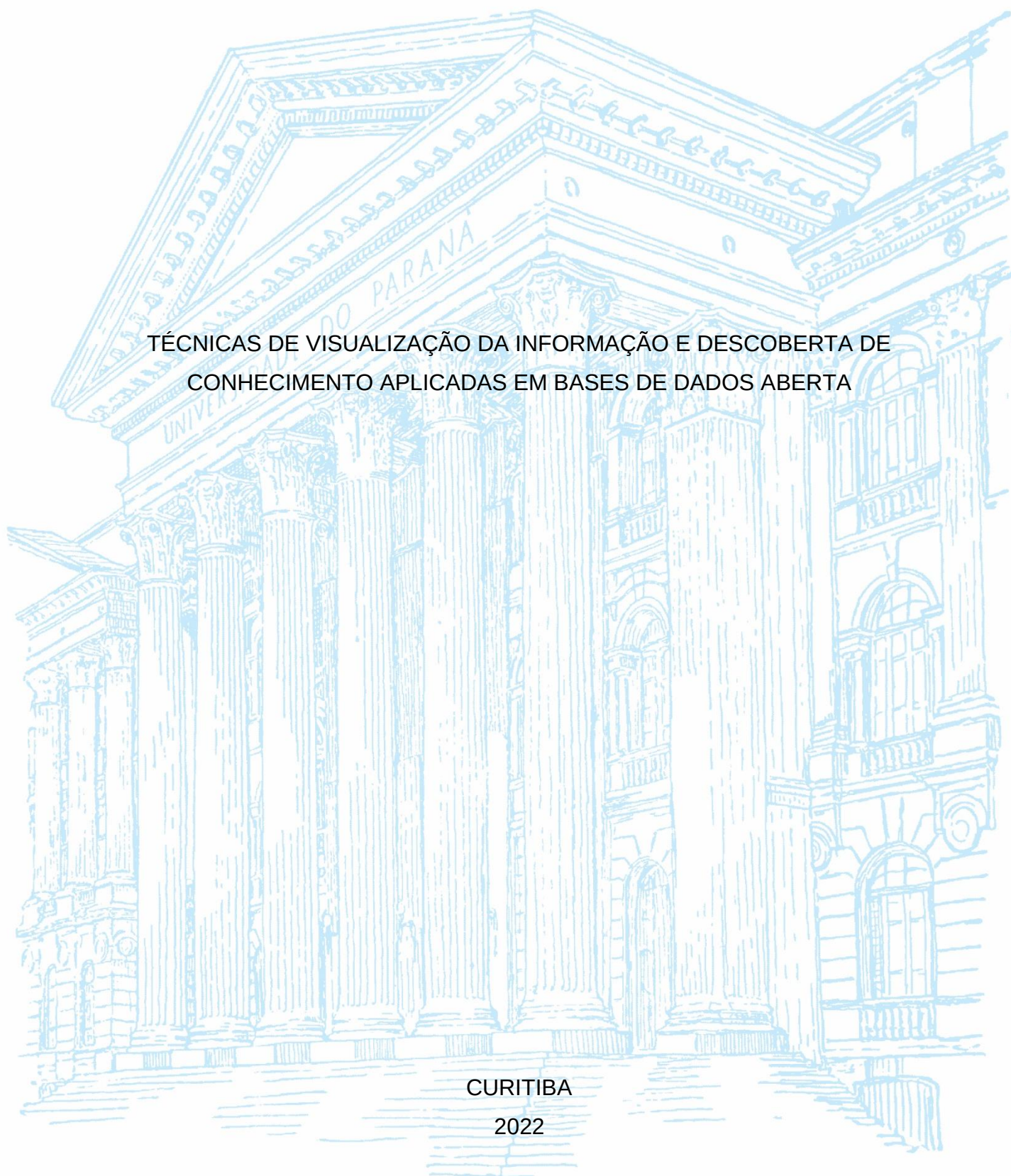
UNIVERSIDADE FEDERAL DO PARANÁ

ALLAN CHIAVELI

TÉCNICAS DE VISUALIZAÇÃO DA INFORMAÇÃO E DESCOBERTA DE
CONHECIMENTO APLICADAS EM BASES DE DADOS ABERTA

CURITIBA

2022



ALLAN CHIAVELI

TÉCNICAS DE VISUALIZAÇÃO DA INFORMAÇÃO E DESCOBERTA DE
CONHECIMENTO APLICADAS EM BASES DE DADOS ABERTA

Trabalho de Conclusão de Curso apresentado
ao curso de graduação em Gestão da
informação, Setor e Ciências Sociais Aplicadas,
Universidade Federal do Paraná, como requisito
parcial à obtenção do título de Bacharel em
Gestão da Informação.

Orientador: Prof. Dr. José Marcelo Almeida
Prado Cestari

CURITIBA
2022

RESUMO

Este estudo visa contribuir para a visualização de dados e informações provenientes de dados abertos, demonstrar como o tratamento adequado pode transformar dados brutos em conhecimentos relevantes. A pesquisa, feita em cinco capitais, nos últimos cinco anos, se baseia em dados abertos sobre a gestão pública brasileira. Através de agências públicas nacionais que pesquisam, registram e compartilham nos meios digitais dados como: o preço da gasolina e do gás de cozinha praticados em todo Brasil, o preço médio de uma cesta básica em diferentes capitais brasileiras e também o salário mínimo vigente no país, e através de correlações entre esses dados, demonstrar como é possível obter possíveis novos conhecimentos sobre estudos socioeconômicos brasileiros de interesse público.

Palavras-chave: Dados abertos. Visualização da informação. Descoberta de conhecimento

ABSTRACT

This study aims to contribute to the visualization of data and information from open data, to demonstrate how proper treatment can transform raw data into relevant knowledge. The survey, carried out in five capitals in the last five years, is based on open data on Brazilian public management. Through national public agencies that research, record and share in digital media data such as: the price of gasoline and cooking gas practiced throughout Brazil, the average price of a basic food basket in different Brazilian capitals and also the minimum wage in force in the country. , and through correlations between these data, demonstrates how it is possible to obtain possible new knowledge about Brazilian socioeconomic studies of public interest.

Keywords: Open data. Information visualization. knowledge discovery

SUMÁRIO

1 INTRODUÇÃO	1
1.1 PROBLEMATIZAÇÃO.....	1
1.2 OBJETIVOS.....	2
1.2.1 Objetivo Geral	3
1.2.2 Objetivos Específicos.....	3
2 REFERENCIAL TEÓRICO	4
2.1 DADOS, INFORMAÇÃO E CONHECIMENTO	4
2.2 DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (DCBD)	5
2.2.1 Etapas do Processo de Descoberta de Conhecimento	6
2.4 VISUALIZAÇÃO DA INFORMAÇÃO	7
2.5 MINERAÇÃO VISUAL DE DADOS.....	8
3 APLICAÇÃO	10
3.1 COLETA DE DADOS	10
3.2 TRATAMENTO DOS DADOS	13
4 RESULTADOS	21
5 CONSIDERAÇÕES FINAIS.....	31
REFERÊNCIAS	33

1 INTRODUÇÃO

Os avanços tecnológicos, conquistados pelo ser humano, ocasionalmente atingem marcos históricos que dão início a novos ciclos de transformações na sociedade. Em meio a essas transformações, o ser humano cresceu de forma exponencial o volume de informações ao seu redor e isso repercute na difusão de um mundo digital, um meio de comunicação instrumentalizado pela informática e pela internet.

Devido ao aparato tecnológico que acompanha essa transformação, uma característica notória é a enorme capacidade de armazenamento de dados, informações e conhecimentos nos meios digitais. “A crescente informatização das empresas, órgãos públicos e demais organizações levou a um vertiginoso aumento no volume de dados produzidos por estas instituições. Tal situação conduziu a visualização da informação a um elevado patamar de importância no mundo corporativo” (RODRIGUES; BRAGA, 2017).

Nesse novo cenário informacional surge o gestor da informação, que tem o papel de encontrar formas inteligentes de tratar e agregar valor a esses dados, transformando em informações relevantes que permitem extrair conhecimentos valiosos. Dentre essas técnicas são utilizadas as de visualização da informação, que implicam no tratamento visual das informações para sua melhor compreensão.

“Uma das áreas que sustenta o mesmo objetivo é a Visualização da Informação (VI). A VI busca facilitar e aprimorar o entendimento de situações, por meio da utilização de técnicas e recursam que permitem uma representação visual de um conjunto de dados analisado.” (SILVA, 2007)

1.1 PROBLEMATIZAÇÃO

Atualmente no Brasil existem diversas bases de dados, disponibilizadas publicamente ao cidadão, graças às políticas de dados abertos e políticas de transparência vigentes na atual legislação brasileira. Os dados estão disponíveis para acesso e para download em diversos portais de instituições públicas, mas em especial no Portal Brasileiro de Dados Abertos, um repositório nacional que possui

um acervo 11.478 conjuntos de bases, provenientes de 208 diferentes agências públicas de todo o Brasil, com temas referentes a todas as áreas da administração pública (agricultura, transporte, saúde, segurança, ciência etc.).

Uma das instituições que depositam bases de dados no Portal Brasileiro de Dados Abertos é a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis - ANP, que dentre outras atividades, realiza uma pesquisa rotineira sobre os preços praticados em cada posto de combustível registrado no Brasil, que hoje chegam a cerca de 120 mil revendedoras em todo país, além das revendedoras de combustível, a pesquisa também é feita com as revendedoras de gás de cozinha.

Outra instituição que realiza pesquisas importantes no âmbito nacional é o Instituto de Pesquisa Econômica Aplicada - IPEA, uma fundação pública federal vinculada ao Ministério da Economia que mantém disponível em seu portal, o histórico de pesquisas macro e microeconômicas feitas no país, que dão suporte técnico às ações governamentais. Dentre esses registros existe o histórico do preço médio de uma cesta básica, estimado de acordo com pesquisas feitas nas capitais de todo o país, mensalmente. Além do preço médio da cesta básica, o IPEA mantém registrado o histórico do salário mínimo vigente no país, tanto o nominal quanto o necessário. Segundo o próprio instituto o salário nominal é aquele decretado pelo governo federal, e o necessário, uma estimativa do que o salário mínimo deveria ser considerando os valores da inflação acerca das necessidades de moradia, alimentação, educação, saúde, lazer, vestuário, higiene, transporte e previdência social.

Informações sobre os preços do combustível, das cestas básicas e dos salários mínimos praticados no país são importantes e de grande interesse público, como podemos utilizar técnicas de visualização da informação e descoberta de conhecimento e aplicá-las em bases de dados abertas para tornar as informações geopolíticas mais acessíveis à população e motivar possíveis *insights* e novos conhecimentos sobre temas relevantes para sociedade?

1.2 OBJETIVOS

Neste tópico é apresentado os objetivos desejados com o desenvolvimento do trabalho, que estão divididos em objetivo geral e específicos.

1.2.1 Objetivo Geral

O objetivo geral deste trabalho consiste em proporcionar condições para que, através da aplicação de técnicas de visualização da informação, conhecimentos possam ser gerados, a partir de bancos de dados abertos, isso deve auxiliar na percepção dos dados e permitir a descoberta de padrões e *insights* e que possam ser de interesse público.

1.2.2 Objetivos Específicos

Os objetivos específicos do trabalho são:

- Selecionar as bases de dados disponíveis nos portais governamentais.
- Identificar e organizar os dados indicadores sobre o tema proposto
- Tratar os dados para relaciona-los e comparar valores.
- Tratar visualmente os dados para torná-los acessíveis e de fácil entendimento.
- Discutir possíveis deduções de conhecimentos e insights.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os conteúdos utilizados para a fundamentação teórica do trabalho, abordando de forma objetiva o tema e seus conceitos gerais.

2.1 DADOS, INFORMAÇÃO E CONHECIMENTO

É importante para compreensão do trabalho, esclarecer as definições de dado, informação e conhecimento, e suas diferenças, pois embora sejam aparentemente semelhantes, são distintas e não podem ser confundidas.

O termo 'dado' aparece muito na literatura da área de Ciência da Informação e de Informática. É definido por Miranda (1999) como um “conjunto de registros qualitativos ou quantitativos conhecido que organizado, agrupado, categorizado e padronizado adequadamente transforma-se em informação”. Da mesma maneira, Miranda (1999) conceitua informação como sendo “dados organizados de modo significativo, sendo subsídio útil à tomada de decisão”. Sendo assim, podemos dizer que a informação está em um nível hierárquico mais alto, constituída por um conjunto de dados que foram agrupados e relacionados ganhando um significado maior.

Por serem simples registros, os dados, por si só, não possuem grande significado ou valor, a menos que lhes seja atribuído por meio de contextualização, categorização, cálculo, correção ou condensação, por exemplo. “Ao acrescentar-lhes significado ou valor, os dados tornam-se informação.” (DAVENPORT; PRUSAK, 1998).

O conhecimento, por sua vez, é algo mais complexo, resultante de uma mistura de elementos.

“É derivado da informação, do mesmo modo que a informação deriva de dados. Assim o conhecimento é obtido através de indivíduos ou grupo de indivíduos que detém algum conhecimento e o transmite através de meios estruturados ou contatos de pessoa a pessoa.” (DAVENPORT; PRUSAK, 1998).

Conhecimento é definido como “a informação valiosa da mente humana e inclui reflexão, síntese e contexto; a sua compreensão e transferência é difícil e este é um dos grandes desafios organizacionais atualmente” (DAVENPORT; PRUSAK, 1998). O conhecimento tem origem ao agregarmos valor às informações. O quadro a seguir traz as principais características de dado, informação e conhecimento segundo Davenport.

QUADRO 1 - DADOS, INFORMAÇÃO E CONHECIMENTO

DADOS	INFORMAÇÃO	CONHECIMENTO
Simple observações sobre o estado do mundo	Dados dotados de relevância e propósito	Informação valiosa da mente humana. Inclui reflexão, síntese e contexto
• Facilmente estruturado • Facilmente obtido por máquinas • Frequentemente Quantificado • Facilmente transferível	• Requer unidade de análise • Exige consenso em relação ao significado • Exige a mediação humana	• De difícil estruturação • De difícil captura em máquinas • Frequentemente tácito • De difícil transferência

FONTE: Davenport, Prusak - 1998 - p.18 (adaptado)

2.2 DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (DCBD)

Em meio a um enorme volume de dados gerados diariamente pelas instituições, em especial aquelas de alcance nacional e internacional, incide a necessidade de estruturação dos dados para que seja possível extrair conhecimento. O processo de descoberta do conhecimento (DCBC), conhecido também como KDD (*Knowledge Discovery from Data Base*) é um processo que auxilia neste processo, ele é definido como o “processo iterativo para identificar nos dados novos padrões que sejam válidos, potencialmente úteis e interpretáveis” (FAYYAD; PIATETSKY-SHAPIO; SMITH, 1996). Os mesmos autores ainda explicam o processo de DCBD como “um processo, de várias etapas, não trivial, interativo e iterativo, para a identificação de padrões compreensíveis, válidos, novos, e potencialmente úteis a partir de grande conjunto de dados” (FAYYAD;

PIATETSKY-SHAPIRO; SMITH, 1996). Ou seja, o processo de DCBC permite que valor seja acrescido no conjunto de dados de acordo com os objetivos do processo. Para implementar o processo de DCBD faz-se necessário ter objetivos claros e definidos, saber quais resultados deseja-se atingir e ter conhecimento da regra de negócio e da aplicação, mesmo que se trate apenas de um estudo acadêmico.

2.2.1 Etapas do Processo de Descoberta de Conhecimento

“O processo de descoberta de conhecimento envolve várias etapas complexas, como a preparação dos dados, a busca por padrões, a avaliação e o refinamento do conhecimento” (HALMENSCHLAGER, 2002). Para o método tradicional pesquisado para este trabalho, o processo de descoberta de conhecimento possui cinco etapas, seleção dos dados, pré-processamento, transformação, mineração e interpretação, que são executadas com o objetivo de identificar padrões nos dados.

“A etapa de seleção se desenvolve através do entendimento das regras do negócio ou in e aplicação para mensurar o objetivo final do cliente” (FAYYAD; PIATETSKYSHAPIRO; SMYTH, 1996). Fazendo uma adaptação para os objetivos do trabalho podemos entender o “negócio” como as instituições públicas pesquisadas e o “cliente” como o interesse público. Nesta etapa também são definidos quais bases de dados e quais os dados específicos que deverão ser utilizados para descoberta de conhecimento.

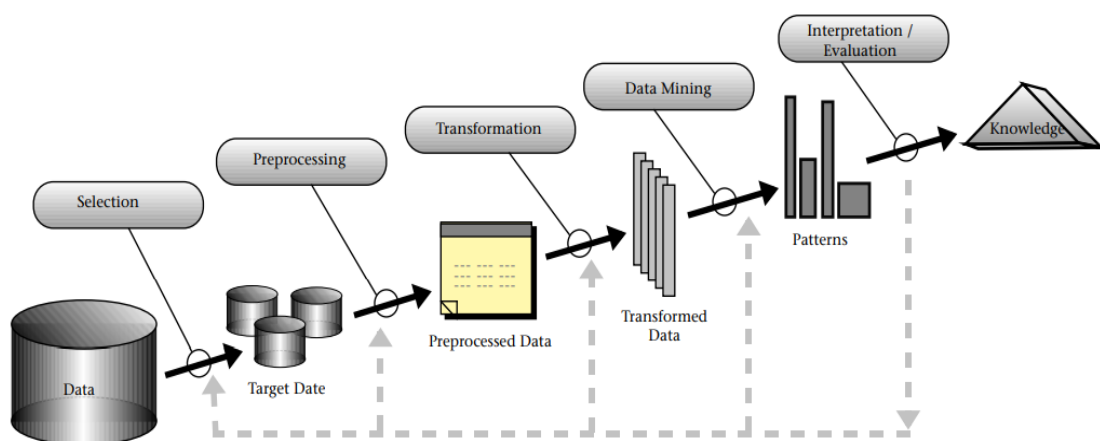
A etapa de pré-processamento, conforme Fonseca (2013), “é um processo de grande importância para remoção de dados fora do domínio, com ausência de valores e valores que não condizem com o atributo.” Deste modo, os dados tornam-se mais confiáveis no processo de mineração de dados. Nessa etapa é feita a “limpeza” dos dados selecionados, a remoção de *outliers* e dados corrompidos, isso é, dados que fogem demasiado da média observada e dados que não condizem com o metadado observado, causados por possíveis erros de digitação ou má formatação da base.

Na transformação os dados são convertidos em um modelo para que possam ser submetidos à etapa posterior. A mineração de dados trata da aplicação de técnicas, implementadas por meio de algoritmos, capazes de receber, como

entrada, um conjunto de fatos ocorridos no mundo real e devolver, como saída, um padrão de comportamento (SILVA; MARQUES; BOSCARIOLI, 2006).

Após esse tratamento temos a fase de visualização e interpretação dos dados, foco deste trabalho, nesta fase “deve-se interpretar o conhecimento apresentado, verificando a relevância (ou não) na obtenção dos padrões” (GÓES; STEINER, 2016). Na Figura 1 são ilustradas as cinco etapas do processo para melhor entendimento:

FIGURA 1 – DESCOBERTA DO CONHECIMENTO



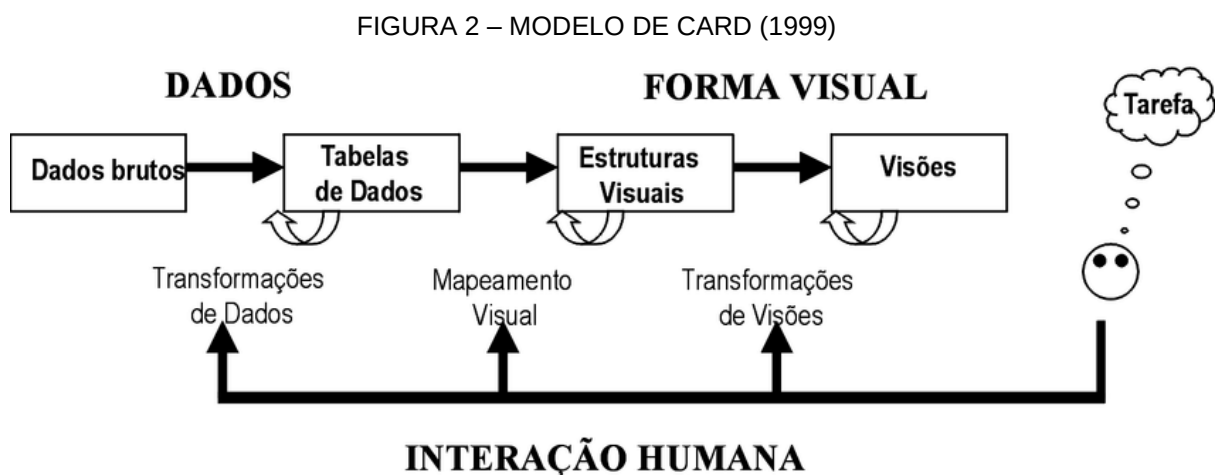
Fonte: Fayyad; Piatetsky-Shapiro; Smyth, 1996

2.4 VISUALIZAÇÃO DA INFORMAÇÃO

Coletar e armazenar dados se tornou algo quase ‘trivial’ e quando é necessário trabalhar com esses grandes volumes de dados, complexos e de diferentes origens, uma dificuldade que se encontra é a sobrecarga de dados e de informações. Uma abordagem para contornar essa dificuldade e selecionar informações relevantes é utilizar técnicas de visualização da informação, para facilitar a descoberta de novos conhecimentos e melhor compreensão desses conhecimentos.

“Visualização da informação é uma área da Ciência que tem por objetivo o estudo das principais formas de representações gráficas para apresentação de informações, a fim de contribuir para o entendimento delas, bem como ajudar a percepção do usuário a fim de deduzir novos conhecimentos baseados no que está sendo apresentado. É uma ciência que combina aspectos de computação gráfica, interação humano-computador, cartografia e mineração de dados” (Freitas 2001).

Card (1999) propõem um modelo para a criação de estruturas de Visualização da Informação, como mostra a Figura 2:



Fonte: Card (1999)

De acordo com o modelo proposto por Card (1983), o sistema perceptivo entra em funcionamento quando o usuário observa a estrutura visual, recebe informações através da visão, inicia o processamento cognitivo que permite a interpretação da estrutura gráfica no seu cérebro e, também, a dedução de novos conhecimentos. As informações apresentadas são transformadas a partir de dados brutos armazenados, por exemplo, em um banco de dados, posteriormente refinados e organizados. Por fim, são criadas imagens gráficas que possibilitam uma análise interativa das informações.

2.5 MINERAÇÃO VISUAL DE DADOS

A aplicação de técnicas de mineração de dados e visualização de informação é referenciada na literatura como mineração visual de dados ou Visual Data Mining (VDM).

“A Mineração Visual de Dados aproxima o usuário/pesquisador e o processo de descoberta de conhecimento em termos de técnicas de visualização eficientes, capacidade de interação e transferência de conhecimento. A visualização de informação não é capaz de substituir as técnicas convencionais de Mineração de Dados, mas as duas técnicas unidas podem potencializar enormemente a exploração de informação, observando que a utilização intercalada pode causar penalidades relativas às deficiências e limitações de cada uma.” (WONG, 1999).

De acordo com BARIONI (2002), a visualização e mineração de dados são processos que podem ser integrados das seguintes formas:

- Dados armazenados em banco de dados podem ser visualizados sob diferentes níveis de abstração, podendo ser utilizadas diferentes combinações de atributos. As formas visuais utilizadas podem ser cubos 3D, curvas, superfícies, grafos ligados e outras.
- Os resultados da mineração de dados podem ser visualizados por meio de formas visuais.
- Visualizar as etapas do KDD de forma que o usuário possa acompanhar o processo desde a extração dos dados até a apresentação do resultado.
- Mineração de dados visual (VDM – Visual Data Mining): Ferramentas de visualização de informação podem ser utilizadas tanto para extrair conhecimentos quanto para a análise dos resultados obtidos com a aplicação de alguma técnica de mineração de dados. Em casos que envolvem grande volume de dados, o usuário pode selecionar porções da base de dados de interesse utilizando técnicas de visualização de informação, diminuindo assim a árdua tarefa exercida no entendimento dos resultados de MD para grande volume de dados.

3 APLICAÇÃO

Nesta seção serão aplicadas as técnicas apresentadas, com base nos objetivos determinados pelo trabalho e os conceitos teóricos apresentados.

3.1 COLETA DE DADOS

A amostra escolhida para pesquisa foi referente aos dados dos últimos cinco anos (2017 a 2021) em cinco capitais, Belém-PA, Curitiba-PR, Goiânia-GO, Salvador-BA e São Paulo-SP.

Os dados sobre os preços da gasolina e do gás de cozinha, praticados no Brasil, foram coletados por meio do Portal Brasileiro de Dados Abertos, onde a Agência Nacional do Petróleo - ANP, deposita várias bases de dados, dentre elas, uma pesquisa realizada semestralmente em todas as revendedoras de gás de cozinha e revendedoras de combustível (postos de gasolina) registradas no país. Essa pesquisa é disponibilizada no formato CSV (*Character Separated Values*), um arquivo em formato de tabela com valores separados por vírgula. A figura a seguir ilustra como a pesquisa é realizada no portal.

FIGURA 3 – PORTAL BRASILEIRO DE DADOS ABERTOS



FONTE: Portal Brasileiro de Dados Abertos (2022)

No portal estão disponíveis as pesquisas feitas de todos os semestres desde 2004, para essa pesquisa foram selecionadas como amostra, as pesquisas dos de 2017 a 2021, isso corresponde a dez bases de dados. As bases podem ser baixadas do portal e lidas pelo aplicativo de criação de tabelas *Excel*, elas contêm dados sobre cada revendedora de gasolina, localização e os preços dos combustíveis. A figura a seguir ilustra como os dados são visualizados dentro da planilha.

FIGURA 4 – DADOS ANP

Regiao - Sigla																
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Regiao - Sigla	Estado - Sigla	Municipio	Revenda	CNPJ da R	Nome da R	Numero R	Complem	Bairro	Cep	Produto	Data da Colet	Valor de V	Valor de C	Unidade d	Bandeira	
2	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	GASOLINA	04/01/2017	4,059		R\$ / litro	BRANCA		
3	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	ETANOL	04/01/2017	3,495		R\$ / litro	BRANCA		
4	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	DIESEL	04/01/2017	3,249		R\$ / litro	BRANCA		
5	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	DIESEL S10	04/01/2017	3,319		R\$ / litro	BRANCA		
6	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	345		SANTA RC 68440-000	GASOLINA	04/01/2017	4,19		36,181 R\$ / litro	BRANCA		
7	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	345		SANTA RC 68440-000	DIESEL	04/01/2017	3,46		28,907 R\$ / litro	BRANCA		
8	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	1306		68440-000	GASOLINA	04/01/2017	4,22		36,252 R\$ / litro	BRANCA		
9	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	1306		68440-000	DIESEL S10	04/01/2017	3,43		29,114 R\$ / litro	BRANCA		
10	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	GASOLINA	04/01/2017	4,3		R\$ / litro	IPIRANGA		
11	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	ETANOL	04/01/2017	3,85		R\$ / litro	IPIRANGA		
12	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	DIESEL	04/01/2017	3,41		R\$ / litro	IPIRANGA		
13	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	DIESEL S10	04/01/2017	3,51		R\$ / litro	IPIRANGA		
14	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	GASOLINA	11/01/2017	4,059		R\$ / litro	BRANCA		
15	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	ETANOL	11/01/2017	3,495		R\$ / litro	BRANCA		
16	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	DIESEL	11/01/2017	3,249		R\$ / litro	BRANCA		
17	N	PA	ABAETETU P CARVAL	14.140.90	AVENIDA	1610		SANTA RC 68440-000	DIESEL S10	11/01/2017	3,319		R\$ / litro	BRANCA		
18	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	345		SANTA RC 68440-000	GASOLINA	11/01/2017	4,19		36,181 R\$ / litro	BRANCA		
19	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	345		SANTA RC 68440-000	DIESEL	11/01/2017	3,56		28,907 R\$ / litro	BRANCA		
20	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	1306		68440-000	GASOLINA	11/01/2017	4,27		36,252 R\$ / litro	BRANCA		
21	N	PA	ABAETETU E CARVALI	04.780.56	RUA JUSTI	1306		68440-000	DIESEL S10	11/01/2017	3,43		29,114 R\$ / litro	BRANCA		
22	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	GASOLINA	11/01/2017	4,3		R\$ / litro	IPIRANGA		
23	N	PA	ABAETETU E M C CON	11.171.19	RODOVIA	2940		CRISTO RE 68440-000	ETANOL	11/01/2017	3,85		R\$ / litro	IPIRANGA		

FONTE: O Autor (2022)

Além dos dados sobre os preços do gás e do combustível, coletados no Portal Brasileiro de Dados Abertos, foram selecionados também dados sobre o valor estimado da cesta básica no país e para comparação foram coletados os dados sobre o salário mínimo vigente no país, que é atualizado todos os ano, levando em consideração que alguns estados do Brasil possuem um valor diferente do valor federal, ou seja, tem seu valor estipulado pelo governo do estado. Essa prática está prevista por lei e pode ocorrer contanto que o valor estipulado pelos estados não esteja abaixo do federal. Os dados sobre a cesta básica e sobre o salário mínimo federal estão disponíveis no portal do Instituto de Pesquisa Econômica Aplicada - IPEA, e as pesquisas são realizadas pelo Departamento Intersindical de Estatísticas e Estudos Socioeconômicos - DIEESE.

Dentro do portal do IPEA é possível pesquisar sobre dados macroeconômicos e regionais, e dentro dos regionais é possível encontrar as pesquisas realizadas mensalmente em quase todas as capitais brasileiras sobre o preço médio de uma cesta básica. O portal também informa a metodologia utilizada na pesquisa, que obedece a estrutura de produtos e quantidades estipuladas por decreto federal, que varia de acordo com cada região do país.

Dentre as capitais da amostra, apenas São Paulo e Curitiba possuem o valor de salário mínimo regional, enquanto as outras capitais seguem o valor vigente federal. Os dados do salário de Curitiba foram coletados no portal da Federação dos Trabalhadores Rurais Agricultores Familiares do Estado do Paraná - FETAEP, que disponibiliza PDF 's dos decretos de lei que atualizam o piso salarial ano a ano no estado do Paraná. E para os dados de São Paulo, foi feita uma pesquisa em sites de notícia sobre o piso salarial vigente dos últimos cinco anos. Após a coleta os dados foram anexados em planilhas. A figura a seguir mostra como a pesquisa é feita no portal do IPEA.

FIGURA 5 - IPEADATA

The screenshot shows the IPEADATA portal interface. At the top, there is a search bar with the text 'insira a palavra' and a search icon. Below the search bar, there are tabs for 'regional', 'macroeconômico', 'ipeadata', and 'social'. The 'ipeadata' tab is selected. On the left side, there is a sidebar with a menu containing 'Temas', 'Fontes', 'Nível geográfico', and 'Ajuda'. The main content area displays the following information:

Cesta básica - gasto mensal
 Fonte: Departamento Intersindical de Estatística e Estudos Sócio-Econômicos (Dieesse)
 Frequência: Mensal de 1959.01 até 2022.03
 Unidade: R\$

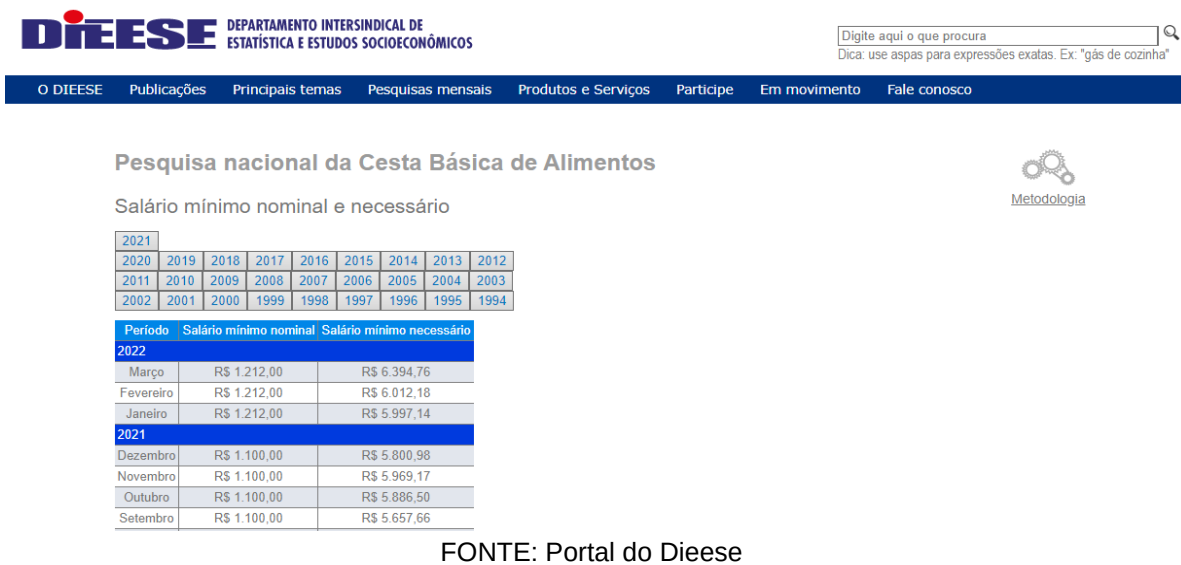
Configuração

Escolha o nível geográfico:	Municípios	Para acessar as séries selecionadas, escolha o nível geográfico de seu interesse que pode ser Brasil, estados, municípios, áreas comparáveis, regiões metropolitanas etc.
Escolha a abrangência:	Brasil	Escolha uma área de abrangência geográfica relevante que pode ser todo Brasil, uma região ou estado específico, ou uma região administrativa como Amazônia Legal, SUDENE, municípios que participam do Programa Fome Zero, entre outros.
Início:	1 / 2022	Escolha o período inicial de interesse.
Fim:	12 / 2022	Escolha o período final de interesse.

FONTE: Portal do IPEAdاتا

A figura a seguir é um recorte do site da DIESSE, onde estão também disponíveis os dados sobre o piso salarial federal.

FIGURA 5 - DIEESE



DIEESE DEPARTAMENTO INTERSINDICAL DE ESTATÍSTICA E ESTUDOS SOCIOECONÔMICOS

Busca: Digite aqui o que procura
Dica: use aspas para expressões exatas. Ex: "gás de cozinha"

O DIEESE | Publicações | Principais temas | Pesquisas mensais | Produtos e Serviços | Participe | Em movimento | Fale conosco

Pesquisa nacional da Cesta Básica de Alimentos

Salário mínimo nominal e necessário

Metodologia

Período	Salário mínimo nominal	Salário mínimo necessário
2022		
Março	R\$ 1.212,00	R\$ 6.394,76
Fevereiro	R\$ 1.212,00	R\$ 6.012,18
Janeiro	R\$ 1.212,00	R\$ 5.997,14
2021		
Dezembro	R\$ 1.100,00	R\$ 5.800,98
Novembro	R\$ 1.100,00	R\$ 5.969,17
Outubro	R\$ 1.100,00	R\$ 5.886,50
Setembro	R\$ 1.100,00	R\$ 5.657,66

FONTE: Portal do Dieese

3.2 TRATAMENTO DOS DADOS

Para uma análise mais eficaz foi preciso organizar os dados de acordo com os objetivos da pesquisa, utilizando para isso o programa de criação de planilhas Excel. Primeiramente foi necessário unir as dez tabelas de dados sobre os preços da gasolina e as dez tabelas sobre os preços do gás para que fosse possível observar a evolução dos preços ao longo dos cinco anos, pois cada tabela continha os dados de apenas um semestre.

Depois que todos os semestres de pesquisa foram reunidos em uma única planilha Excel, foi preciso limpar os dados que estavam em excesso, pois as tabelas originais possuíam dados desnecessários para a pesquisa, colunas e linhas foram excluídas em uma versão mais “limpa” da tabela. Na tabela de preços da gasolina e na tabela de preço do gás, foram descartadas as cidades que não faziam parte da amostra e também as colunas de dados que não interessam para a pesquisa, como o CEP ou o CNPJ das revendedoras.

Foi preciso também normalizar e corrigir alguns valores da coluna de preço, pois no caso da gasolina, que são utilizados três casas decimais para os centavos,

algumas células da tabela não estavam considerando a vírgula como um separador decimal mas sim um indicador de milhar, por exemplo, uma célula com o preço da gasolina a R\$3,155 não estava considerando a vírgula como separador entre reais e centavos e considerava o valor como “R\$3.150,00”. Para evitar esses valores enganosos, foi preciso localizar esses valores e digitá-los novamente, eliminando possíveis erros nos cálculos. Ao final dessa normalização, que foi feita utilizando apenas o Excel, foi obtida uma única planilha com o histórico de preços dos cinco anos da amostra. A seguir a versão “normalizada” dos dados da Gasolina:

FIGURA 6 – TABELA ANP

	A	B	C	D	E	F	G
	Município	Revenda	Produto	Data da Coleta	Valor de Venda	Bandeira	
2	BELEM	POSTO EXEMPLO COMERCIO LTDA	GASOLINA	03/01/2017	3.929	IPIRANGA	
5	BELEM	GP PETROLEO COMERCIO DE COMBUSTIVEIS LTDA	GASOLINA	03/01/2017	3.899	SABBÁ	
8	BELEM	SAITO COMERCIO DE COMBUSTIVEIS LTDA	GASOLINA	03/01/2017	3.899	SABBÁ	
11	BELEM	AUTO POSTO BELEM LTDA	GASOLINA	03/01/2017	3.899	SABBÁ	
13	BELEM	IRMAOS TEIXEIRA LTDA	GASOLINA	03/01/2017	3.999	IPIRANGA	
16	BELEM	AUTO POSTO BICOLOR LTDA	GASOLINA	03/01/2017	3.899	IPIRANGA	
19	BELEM	AUTO POSTO VERDÃO LTDA	GASOLINA	03/01/2017	3.999	PETROBRAS DISTRIBUIDORA S.A.	
22	BELEM	SUPER POSTO TRIANGULO LTDA	GASOLINA	03/01/2017	3.999	IPIRANGA	
25	BELEM	POSTO QUATRO LTDA	GASOLINA	03/01/2017	3.879	PETROBRAS DISTRIBUIDORA S.A.	
27	BELEM	POSTO SAO DOMINGOS LTDA	GASOLINA	03/01/2017	3.999	IPIRANGA	
29	BELEM	AUTO POSTO SENADOR LTDA	GASOLINA	03/01/2017	3.999	IPIRANGA	
31	BELEM	ABASTECEDORA BRASILIA LTDA	GASOLINA	03/01/2017	3.899	SABBÁ	
33	BELEM	IRMAOS TEIXEIRA LTDA	GASOLINA	03/01/2017	4.029	IPIRANGA	
36	BELEM	POSTO HUMAITA LTDA	GASOLINA	03/01/2017	3.939	PETROBRAS DISTRIBUIDORA S.A.	
39	BELEM	SUPER POSTO VIEIRA LTDA	GASOLINA	03/01/2017	3.789	BRANCA	
42	BELEM	ALBINO F. SANTOS & CIA. LTDA.	GASOLINA	11/01/2017	3.99	IPIRANGA	
45	BELEM	GP PETROLEO COMERCIO DE COMBUSTIVEIS LTDA	GASOLINA	11/01/2017	3.899	SABBÁ	
48	BELEM	SAITO COMERCIO DE COMBUSTIVEIS LTDA	GASOLINA	11/01/2017	3.899	SABBÁ	
51	BELEM	AUTO POSTO BELEM LTDA	GASOLINA	11/01/2017	3.899	SABBÁ	
53	BELEM	IRMAOS TEIXEIRA LTDA	GASOLINA	11/01/2017	3.999	IPIRANGA	
56	BELEM	A.M.C AGUILERA COMERCIO DE COMBUSTIVEIS LTDA	GASOLINA	11/01/2017	3.999	RAIZEN	
58	BELEM	AUTO POSTO VERDÃO LTDA	GASOLINA	11/01/2017	3.999	PETROBRAS DISTRIBUIDORA S.A.	

FONTE: O Autor (2022)

Os dados coletados no portal do IPEA sobre a cesta básica foram pesquisados cidade por cidade, no período de 2017 a 2021, como a pesquisa é realizada mensalmente, os dados foram transferidos para uma tabela da mesma forma utilizando planilhas do Excel. A seguir um recorte da planilha com os dados da Cesta Básica produzida.

FIGURA 7 – TABELA CESTA BASICA

	A	B	C	D	E	F	G	H	I	J	K
1	Cidade	Mês-Ano	Preço								
2	Belém	jan-17	R\$ 406,40								
3	Belém	fev-17	R\$ 395,57								
4	Belém	mar-17	R\$ 394,21								
5	Belém	abr-17	R\$ 398,12								
6	Belém	mai-17	R\$ 402,76								
7	Belém	jun-17	R\$ 393,01								
8	Belém	jul-17	R\$ 388,67								
9	Belém	ago-17	R\$ 375,82								
10	Belém	set-17	R\$ 369,89								
11	Belém	out-17	R\$ 367,03								
12	Belém	nov-17	R\$ 358,74								
13	Belém	dez-17	R\$ 356,67								
14	Belém	jan-18	R\$ 366,99								
15	Belém	fev-18	R\$ 379,36								
16	Belém	mar-18	R\$ 367,06								
17	Belém	abr-18	R\$ 366,19								
18	Belém	mai-18	R\$ 367,56								
19	Belém	jun-18	R\$ 381,65								
20	Belém	jul-18	R\$ 361,11								
21	Belém	ago-18	R\$ 360,30								
22	Belém	set-18	R\$ 359,51								
23	Belém	out-18	R\$ 361,70								

FONTE: O Autor (2022)

Utilizando também o Excel, os dados sobre o salário mínimo foram inseridos em uma planilha. Os dados sobre o salário mínimo variam anualmente, isso é, eles sofrem alteração apenas uma vez no ano, contudo, foram colocados em uma tabela separados mês a mês para que posteriormente fosse mais fácil de compará-los com a variação da cesta básica, que é atualizada mensalmente, mesmo que os valores se repetissem ao longo de 12 meses. A seguir a tabela do salário mínimo, agora com as médias semestrais.

FIGURA 8 – TABELA SALÁRIO MÍNIMO (MÉDIA)

Mês/Ano											
Mês/Ano	Belém	Curitiba	Goiânia	Salvador	São Paulo						
jan/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
fev/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
mar/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
abr/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
mai/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
jun/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
jul/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
ago/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
set/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
out/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
nov/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
dez/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
jan/18	R\$ 954,00	R\$ 1.269,40	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
fev/18	R\$ 954,00	R\$ 1.269,40	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
mar/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
abr/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
mai/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
jun/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
jul/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
ago/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
set/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
out/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						

FONTE: O Autor (2022)

De acordo com a metodologia de pesquisa apresentada em seu portal, a DIEESE estima o valor da cesta básica calculando a média dos preços dos produtos através de uma pesquisa em vários mercados das cidades, com isso, a soma das médias dos produtos que compõem a cesta resulta no valor da cesta apresentada no portal.

No caso da pesquisa feita pela Agência Nacional do Petróleo, sobre os preços do gás de cozinha e da gasolina, ela contém os valores individuais de cada revendedora da cidade, para que seja possível comparar o preços do gás e da gasolina com os demais valores da pesquisa, foi calculada a média dos preços em cada cidade por semestre de acordo com a pesquisa da ANP, utilizando o próprio Excel. E da mesma maneira que o salário mínimo, as médias foram colocadas numa tabela mês a mês, para facilitar a comparação com as demais bases da pesquisa.

FIGURA 9 – TABELA GASOLINA (MÉDIAS)

	A	B	C	D	E	F	G	H	I
1	Data	Belém	Curitiba	Goiânia	Salvador	São Paulo			
2	jan-17	R\$ 3,887	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
3	fev-17	R\$ 3,887	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
4	mar-17	R\$ 3,887	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
5	abr-17	R\$ 3,724	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
6	mai-17	R\$ 3,724	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
7	jun-17	R\$ 3,724	R\$ 3,397	R\$ 3,775	R\$ 3,645	R\$ 3,441			
8	jul-17	R\$ 3,699	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
9	ago-17	R\$ 3,699	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
10	set-17	R\$ 3,699	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
11	out-17	R\$ 3,908	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
12	nov-17	R\$ 3,908	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
13	dez-17	R\$ 3,908	R\$ 3,724	R\$ 4,125	R\$ 3,827	R\$ 3,611			
14	jan-18	R\$ 4,026	R\$ 4,070	R\$ 4,512	R\$ 4,292	R\$ 4,025			
15	fev-18	R\$ 4,026	R\$ 4,070	R\$ 4,512	R\$ 4,292	R\$ 4,025			
16	mar-18	R\$ 4,026	R\$ 4,070	R\$ 4,512	R\$ 4,292	R\$ 4,025			
17	abr-18	R\$ 4,026	R\$ 4,070	R\$ 4,512	R\$ 4,292	R\$ 4,025			

FONTE: O Autor (2017-2021)

Os dados sobre os preços do gás e da gasolina das cinco cidades foram organizados de forma que ficaram juntos em uma única tabela, assim como os dados sobre as cestas básicas. Da mesma forma, os dados sobre os salários mínimos das cinco cidades também foram unidos em uma única tabela. Isso permite que a evolução dos valores seja comparada entre as cidades de maneira mais fácil.

FIGURA 10 – TABELA COMPARATIVA

	A	B	C	D	E	F	G	H	I	J	K	L
1	Mês/Ano	Belém	Curitiba	Goiânia	Salvador	São Paulo						
2	jan/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
3	fev/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
4	mar/17	R\$ 937,00	R\$ 1.190,20	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
5	abr/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
6	mai/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
7	jun/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
8	jul/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
9	ago/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
10	set/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
11	out/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
12	nov/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
13	dez/17	R\$ 937,00	R\$ 1.269,40	R\$ 937,00	R\$ 937,00	R\$ 1.000,00						
14	jan/18	R\$ 954,00	R\$ 1.269,40	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
15	fev/18	R\$ 954,00	R\$ 1.269,40	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
16	mar/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
17	abr/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
18	mai/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
19	jun/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
20	jul/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
21	ago/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
22	set/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						
23	out/18	R\$ 954,00	R\$ 1.293,60	R\$ 954,00	R\$ 954,00	R\$ 1.108,00						

FONTE: O Autor (2022)

Uma vez que as tabelas foram organizadas desta forma, é possível relacionar os dados de maneira mais fácil pois estão tabulados de maneira semelhante. É possível comparar a evolução dos preços dos produtos ao lado da evolução do salário mínimo para entender quais as discrepâncias que existem entre a inflação dos produtos e os ajustes do piso salarial. Uma maneira de esclarecer isso também, é calculando a porcentagem que os produtos representam no salário mínimo, ou seja, qual a porcentagem que um trabalhador deve dispor do seu salário para comprar uma cesta básica, um botijão de gás ou encher o tanque do carro.

Para essa estimativa foram fixados alguns valores para o gasto médio de cada produto pesquisado, no caso da cesta básica foram consideradas 2 cestas básicas por mês, mesmo que, de acordo com a DIEESE, uma cesta básica, para uma família de 4 pessoas, deve durar 10 dias. No caso do gás de cozinha, segundo algumas revendedoras, o gás de um botijão de 13kg pode durar de 30 a 50 dias considerando um fogão de 4 bocas sendo usado 3h/dia, sabendo disso, foi considerado para os cálculos 1 botijão de gás por mês. Já no caso da gasolina, foi considerado um abastecimento semanal de 35 litros de gasolina, que equivale à capacidade média do tanque de um carro popular, totalizando 140 litros por mês.

Tendo em vista esses valores foi criada uma outra planilha reunindo os valores do salário mínimo junto com os valores da gasolina, do gás e da cesta básica lado a lado e junto a eles, a porcentagem que cada um representa em relação ao salário mínimo de cada cidade ao longo dos cinco anos.

Figura 11 – TABELA DE PORCENTAGENS

Mês-Ano												
	A	B	C	D	E	F	G	H	I	J	K	L
1	Mês-Ano	Salário mínimo	Cesta básica	% CESTA	Gás	% GÁS	Gasolina	% GASOLINA				
2	jan/17	R\$ 1.190,20	R\$ 397,69	33,41	R\$ 54,47	2,29	R\$ 3,397	39,96				
3	fev/17	R\$ 1.190,20	R\$ 387,27	32,54	R\$ 54,47	2,29	R\$ 3,397	39,96				
4	mar/17	R\$ 1.190,20	R\$ 389,52	32,73	R\$ 54,47	2,29	R\$ 3,397	39,96				
5	abr/17	R\$ 1.269,40	R\$ 404,55	31,87	R\$ 54,47	2,15	R\$ 3,397	37,47				
6	mai/17	R\$ 1.269,40	R\$ 403,51	31,79	R\$ 54,47	2,15	R\$ 3,397	37,47				
7	jun/17	R\$ 1.269,40	R\$ 394,49	31,08	R\$ 54,47	2,15	R\$ 3,397	37,47				
8	jul/17	R\$ 1.269,40	R\$ 399,00	31,43	R\$ 56,22	2,21	R\$ 3,724	41,07				
9	ago/17	R\$ 1.269,40	R\$ 385,11	30,34	R\$ 56,22	2,21	R\$ 3,724	41,07				
10	set/17	R\$ 1.269,40	R\$ 376,46	29,66	R\$ 56,22	2,21	R\$ 3,724	41,07				
11	out/17	R\$ 1.269,40	R\$ 388,06	30,57	R\$ 56,22	2,21	R\$ 3,724	41,07				
12	nov/17	R\$ 1.269,40	R\$ 381,26	30,03	R\$ 56,22	2,21	R\$ 3,724	41,07				
13	dez/17	R\$ 1.269,40	R\$ 374,94	29,54	R\$ 56,22	2,21	R\$ 3,724	41,07				
14	jan/18	R\$ 1.269,40	R\$ 399,72	31,49	R\$ 64,68	2,55	R\$ 4,070	44,89				
15	fev/18	R\$ 1.269,40	R\$ 392,91	30,95	R\$ 64,68	2,55	R\$ 4,070	44,89				
16	mar/18	R\$ 1.293,60	R\$ 401,65	31,05	R\$ 64,68	2,50	R\$ 4,070	44,05				
17	abr/18	R\$ 1.293,60	R\$ 394,07	30,46	R\$ 64,68	2,50	R\$ 4,070	44,05				
18	mai/18	R\$ 1.293,60	R\$ 397,17	30,70	R\$ 64,68	2,50	R\$ 4,070	44,05				
19	jun/18	R\$ 1.293,60	R\$ 412,44	31,88	R\$ 64,68	2,50	R\$ 4,070	44,05				
20	jul/18	R\$ 1.293,60	R\$ 391,32	30,25	R\$ 65,46	2,53	R\$ 4,226	45,74				
21	ago/18	R\$ 1.293,60	R\$ 389,15	30,08	R\$ 65,46	2,53	R\$ 4,226	45,74				
22	set/18	R\$ 1.293,60	R\$ 387,39	29,95	R\$ 65,46	2,53	R\$ 4,226	45,74				
23	out/18	R\$ 1.293,60	R\$ 406,42	31,42	R\$ 65,46	2,53	R\$ 4,226	45,74				
Belém Curitiba Goiânia Salvador São Paulo (+)												

FONTE: O Autor (2022)

Foi criada também outra tabela, acrescentando os valores fixados para simular o gasto mensal de um cidadão comum. Essa tabela contém a porcentagem que esses valores representam do salário mínimo individualmente e também somadas.

FIGURA 12 – TABELA SIMULAÇÃO

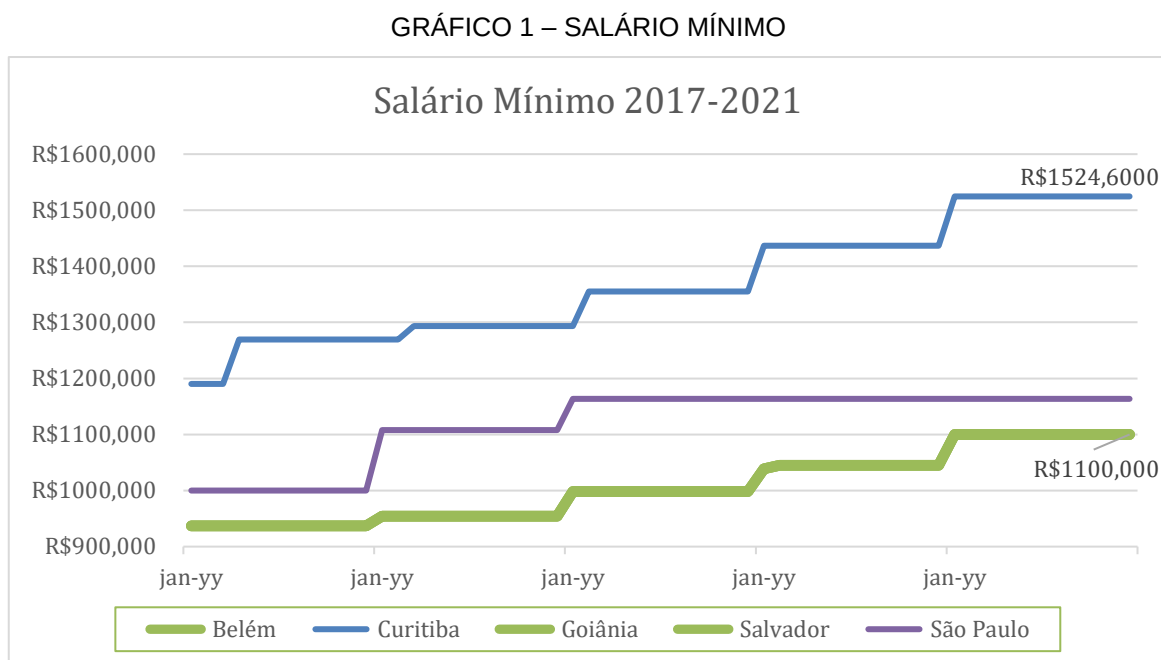
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Mês-ANO	Salário mínimo	Cesta básica	2 Cestas Básicas	% 2 Cestas Básicas	Gás	% Gás	Gasolina	Gasolina (140l)	% GASOLINA (140l)	cas + Gasolina (140l)	cas + Gasolina (140l) + Gás (13kg)		
2	jan/17	R\$ 937,00	R\$ 386,04	R\$ 771,68	82,36%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.357,66	144,89%		
3	fev/17	R\$ 937,00	R\$ 380,19	R\$ 760,36	81,15%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.346,34	143,69%		
4	mar/17	R\$ 937,00	R\$ 388,31	R\$ 776,62	82,88%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.362,60	145,42%		
5	abr/17	R\$ 937,00	R\$ 388,83	R\$ 777,66	82,99%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.363,64	145,53%		
6	mai/17	R\$ 937,00	R\$ 382,72	R\$ 765,44	83,82%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.371,42	146,36%		
7	jun/17	R\$ 937,00	R\$ 380,19	R\$ 780,20	83,27%	R\$ 57,47	6,13%	R\$ 3,775	R\$ 528,512	56,40%	R\$ 1.366,18	145,80%		
8	jul/17	R\$ 937,00	R\$ 386,67	R\$ 773,34	82,53%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.414,53	150,96%		
9	ago/17	R\$ 937,00	R\$ 386,81	R\$ 773,62	82,56%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.414,81	150,99%		
10	set/17	R\$ 937,00	R\$ 372,89	R\$ 745,98	79,61%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.387,17	148,04%		
11	out/17	R\$ 937,00	R\$ 382,58	R\$ 725,16	77,39%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.366,35	145,82%		
12	nov/17	R\$ 937,00	R\$ 381,96	R\$ 723,92	77,26%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.365,11	145,69%		
13	dez/17	R\$ 937,00	R\$ 380,70	R\$ 721,40	76,99%	R\$ 63,74	6,80%	R\$ 4,125	R\$ 577,450	61,63%	R\$ 1.362,59	145,42%		
14	jan/18	R\$ 954,00	R\$ 382,23	R\$ 724,46	75,94%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.431,76	150,08%		
15	fev/18	R\$ 954,00	R\$ 384,01	R\$ 728,02	76,31%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.435,32	150,45%		
16	mar/18	R\$ 954,00	R\$ 380,43	R\$ 720,86	75,56%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.428,16	149,70%		
17	abr/18	R\$ 954,00	R\$ 385,81	R\$ 731,62	76,69%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.438,92	150,83%		
18	mai/18	R\$ 954,00	R\$ 385,65	R\$ 733,30	76,87%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.440,60	151,01%		
19	jun/18	R\$ 954,00	R\$ 385,81	R\$ 731,62	76,69%	R\$ 75,58	7,92%	R\$ 4,512	R\$ 631,721	66,22%	R\$ 1.438,92	150,83%		
20	jul/18	R\$ 954,00	R\$ 386,38	R\$ 732,76	76,81%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.476,16	154,73%		
21	ago/18	R\$ 954,00	R\$ 382,47	R\$ 724,94	75,99%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.468,34	153,91%		
22	set/18	R\$ 954,00	R\$ 384,11	R\$ 708,22	74,24%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.451,62	152,16%		
23	out/18	R\$ 954,00	R\$ 387,72	R\$ 715,44	74,99%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.458,84	152,92%		
24	nov/18	R\$ 954,00	R\$ 388,06	R\$ 736,12	77,16%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.479,52	155,09%		
25	dez/18	R\$ 954,00	R\$ 388,86	R\$ 777,72	81,52%	R\$ 75,49	7,91%	R\$ 4,771	R\$ 667,909	70,01%	R\$ 1.521,12	159,45%		

FONTE - O Autor (2022)

4 RESULTADOS

Os dados foram coletados e organizados de uma maneira lógica, de acordo com os objetivos da pesquisa, contudo, o formato de tabelas não mostra com clareza o comportamento dos dados, para melhor análise e entendimento sobre eles, aplicando-se os conceitos de visualização da informação, os dados foram convertidos em gráficos de linhas, utilizando o próprio aplicativo Excel e apresentados a seguir.

Com relação ao salário mínimo, através da tabela montada, foi possível produzir um gráfico para comparar a evolução do salário mínimo entre as cidades, sabendo que apenas Curitiba e São Paulo possuem um valor Regional diferente, enquanto que Belém, Goiânia e Salvador seguem o piso salarial federal e por isso aparecem sobrepostas no gráfico.



FONTE: O Autor (2022)

Observando o gráfico 1, podemos notar que Curitiba possui o maior piso salarial, com reajustes em todos os últimos cinco anos, atingindo o valor de R\$1524,20 no último ano, São Paulo por sua vez, apesar de manter um piso salarial regional maior que o federal, não sofre reajustes nos últimos dois anos. No ano

passado, o piso salarial de São Paulo quase se equiparou ao valor federal, com uma diferença de apenas R\$63,00.

O próximo gráfico (GRÁFICO 2) mostra a evolução do preço da cesta básica, afim de comparar as diferenças ou semelhanças entre as cidades pesquisadas.



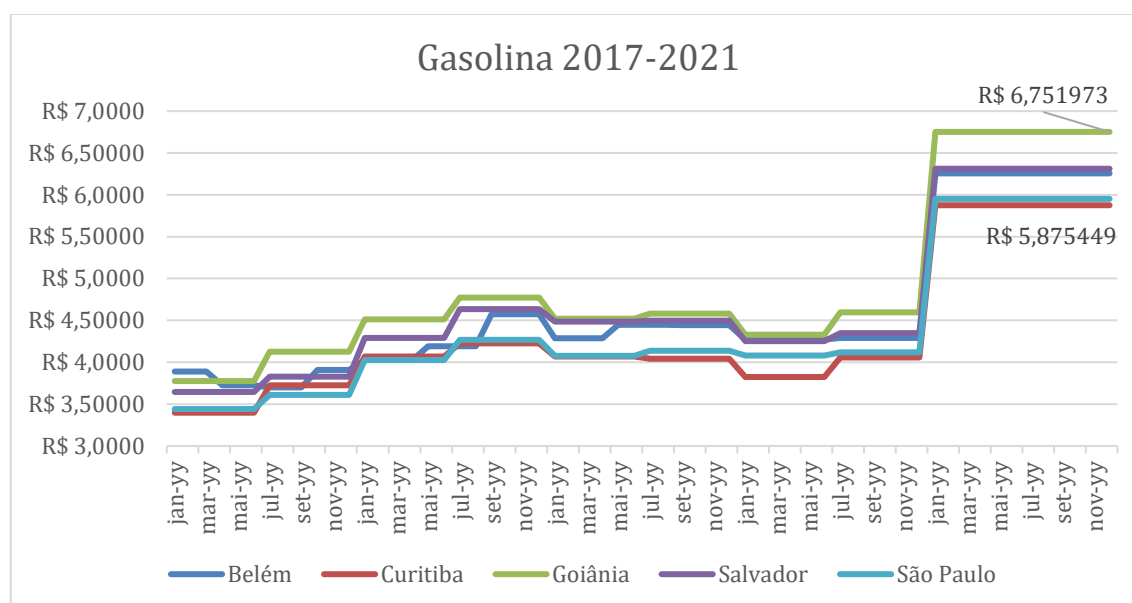
FONTE: O Autor (2021)

O gráfico 2 de linhas mostra que a variação dos preços acontece de maneira bastante proporcional por praticamente todo o período observado, tendo as cidades quase sempre nas mesmas posições. São Paulo apresenta os maiores valores durante os cinco anos, com uma cesta básica custando quase R\$700 reais no final de 2021, seguida de Curitiba, custando R\$628, e com o custo mais baixo, fica Salvador, custando cerca de R\$520 reais uma cesta básica no mesmo período.

Podemos já notar uma certa discordância de valores se levarmos em consideração também o gráfico do salário mínimo, o gráfico anterior (Gráfico1), mostra São Paulo com um salário mínimo pouco corrigido nos últimos anos, e mesmo assim com os maiores preços da cesta básica. Por outro lado, as cidades que utilizam o piso federal vigente em seus estados também possuem preços menores para a cesta básica.

Seguindo as comparações por cidade, a tabela com os preços da gasolina resultou no seguinte gráfico:

GRÁFICO 3 - GASOLINA



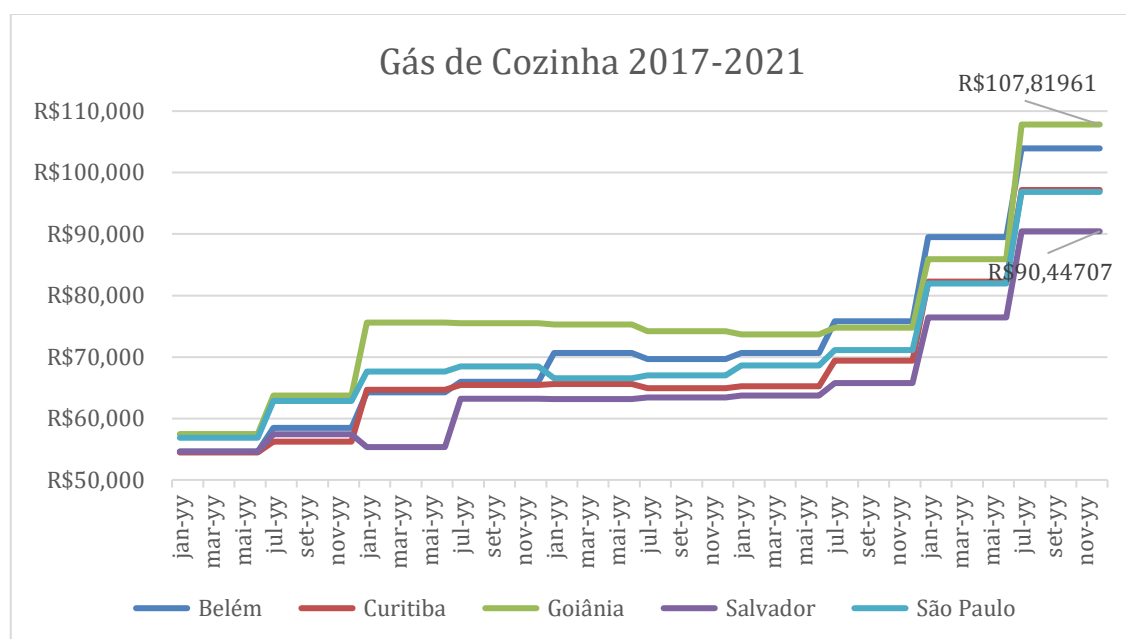
FONTE: O Autor (2022)

Neste gráfico podemos observar que os preços variam em todas as cidades na mesma proporção, sendo que Goiânia e Salvador apresentam os maiores preços e permanecem nas posições acima da tabela por quase todo período observado enquanto que Curitiba e São Paulo permanecem mais abaixo, com preços menores. Em 2021 foi constatada uma enorme alta nos preços da gasolina, Goiânia apresentou o maior preço, custando em média R\$6,752 o litro gasolina, enquanto que Curitiba apresentou a menor média, custando R\$5,875 o litro da gasolina no mesmo período, em 2021.

Se levarmos em consideração o salário mínimo das cidades, podemos afirmar que Curitiba tem as melhores condições para um trabalhador que utiliza o carro diariamente, pois além da cidade apresentar o maior piso salarial, tem também as menores médias de preço do combustível, junto com São Paulo, enquanto que as cidades com médias mais altas no preço da gasolina seguem o piso salarial mínimo federal.

O gráfico a seguir traz a comparação entre cidades, agora em relação ao preço do gás de cozinha.

GRÁFICO 4 - GÁS



FONTE: O Autor (2022)

Com relação ao gás de cozinha podemos observar novamente que as variações são proporcionais por todas as cidades por quase todo período pesquisado, com exceção de Salvador, que no início de 2018, foi a única com uma baixa nos preços, enquanto as outras cidades tiveram altas, em especial Goiânia, com uma alta bem superior às outras. É observada uma grande alta nos preços para todas as cidades no final de 2020 e em 2021. Goiânia permanece com os maiores preços durante quase todos os cinco anos, fechando 2021 com o maior preço médio, R\$107 por botijão de gás, Salvador permanece com os menores preços, custando em média R\$90 no mesmo período e São Paulo e Curitiba encerram ficando praticamente empatados com a segunda menor média de preços, cerca de R\$97. Mais uma vez podemos notar uma discrepância entre os custos que o trabalhador tem e o salário mínimo vigente no seu estado, Goiânia apresentou pela segunda vez as maiores altas nos preços e segue com o piso salarial mais baixo, em contrapartida, Curitiba, que possui um piso salarial maior que as outras cidades, ficou entre as menores médias no preço do gás, ou seja, podemos dizer que em Curitiba existe um custo de vida relativamente menor.

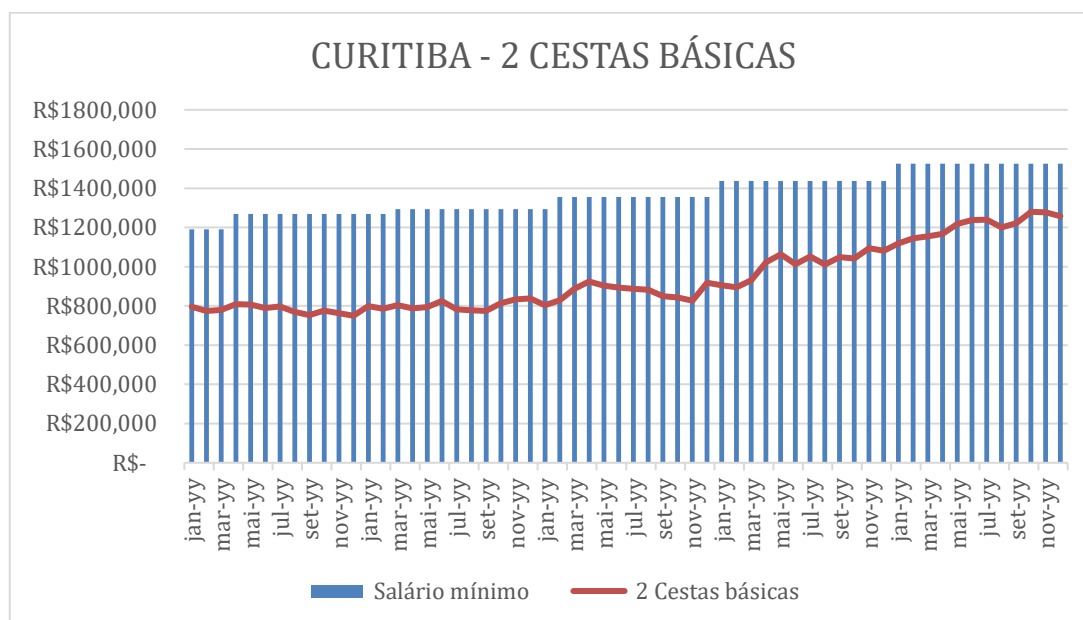
Com os gráficos produzidos, podemos notar que Curitiba e Goiânia representam os dois extremos em relação aos valores analisados, Curitiba com o

maior piso salarial e menores preços dos produtos, e Goiânia com menor piso salarial e maiores preços dos produtos. Para avaliar melhor essa diferença, foi utilizado a tabela com as porcentagens de cada valor observado, ou seja, a porcentagem que os produtos representam no salário mínimo destas duas cidades.

Para essa comparação foi utilizado como base os valores antes fixados para cada produto na etapa de tratamento dos dados para estimar o gasto mensal médio do trabalhador (2 cestas básicas, 1 botijão de cozinha (13 kg) e 140 litros de gasolina).

Na tabela abaixo temos o piso salarial de Curitiba representado pelas colunas e o valor de 2 cestas básicas, representadas pela linha laranja. No seu ponto mais baixo, em setembro de 2017, o preço de duas cestas básicas representa 59% do piso salarial, e no final de 2021 chega a 84%.

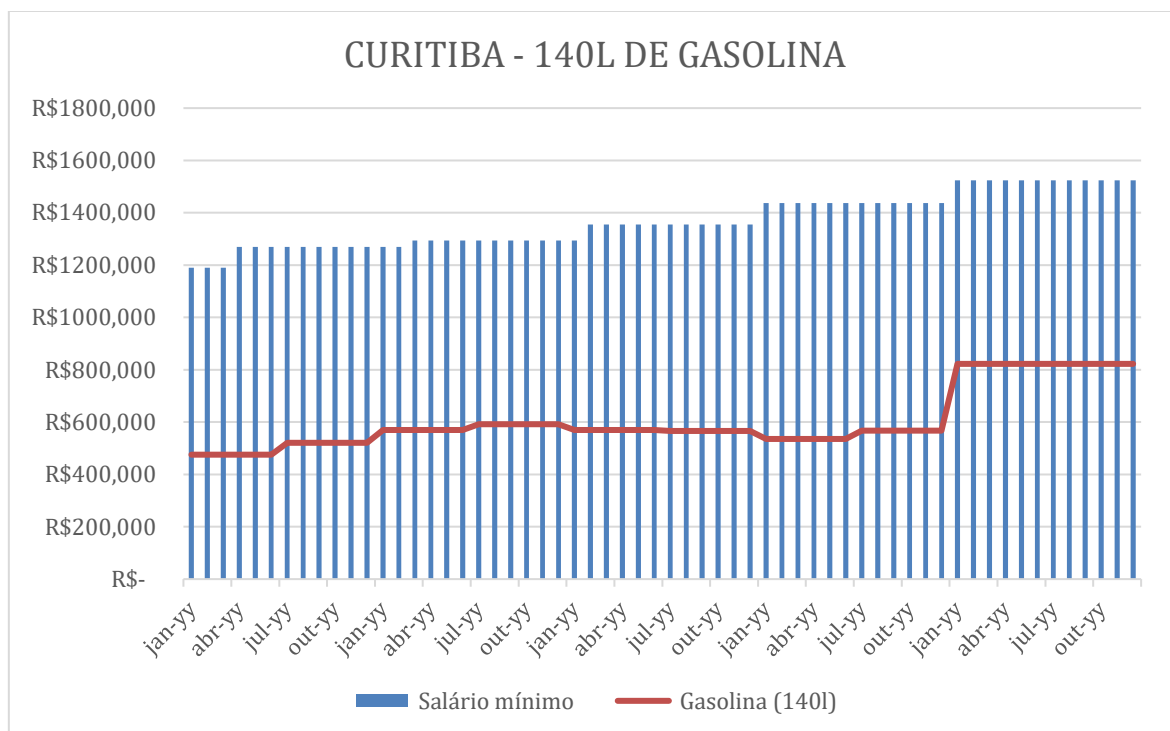
GRÁFICO 5 – CURITIBA, DUAS CESTAS BÁSICAS



FONTE: O Autor (2022)

No gráfico a seguir podemos observar o valor do piso salarial em Curitiba junto com a evolução do preço de 140 litros de gasolina. Em 2017 essa quantia representava 39% do salário mínimo e no final de 2021 chega a 53%.

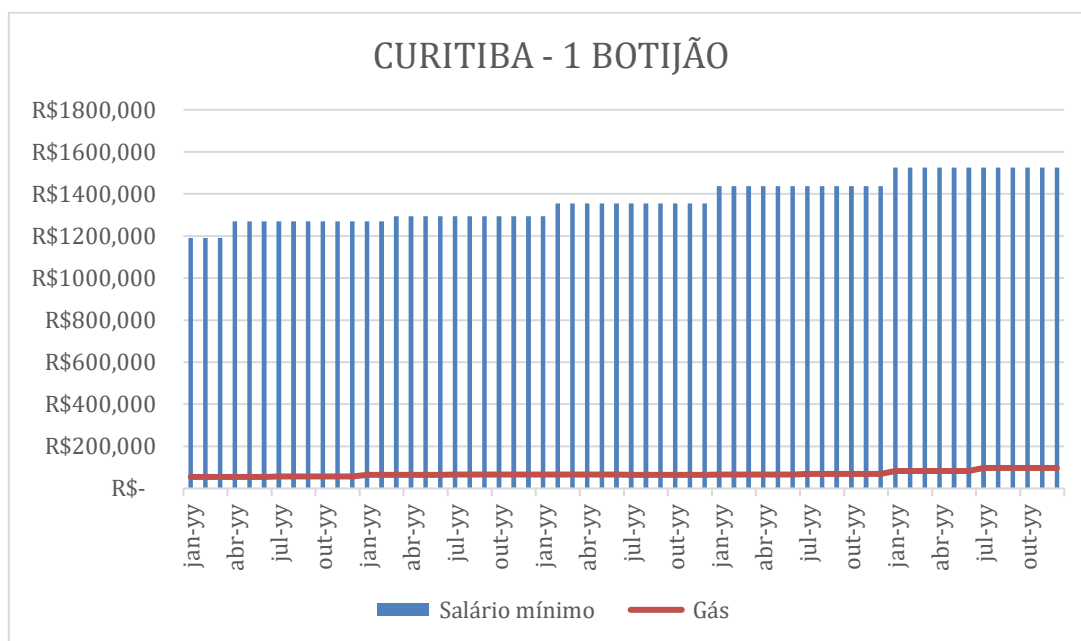
GRÁFICO 6 – CURITIBA, 140L DE GASOLINA



FONTE: O Autor (2022)

A seguir o gráfico, mais uma vez mostra a evolução do salário mínimo de Curitiba, agora em comparação com o valor do Botijão de gás de cozinha, que em 2017 representava 4,5%, e em 2021 passou a 6,4%

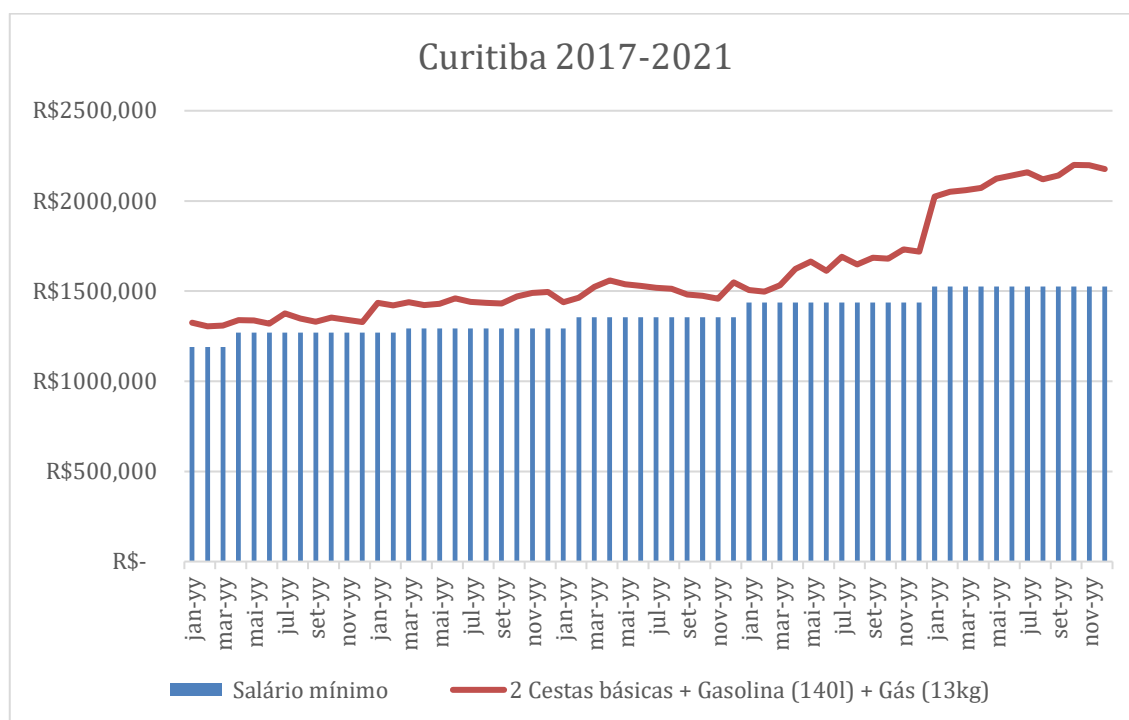
GRÁFICO 7 – CURITIBA, 1 BOTIJÃO



FONTE: O Autor (2022)

O gráfico a seguir une os valores dos três produtos e compara a evolução desta soma a evolução do Salário mínimo em Curitiba.

GRÁFICO 8 – COMPARATIVO GERAL CURITIBA



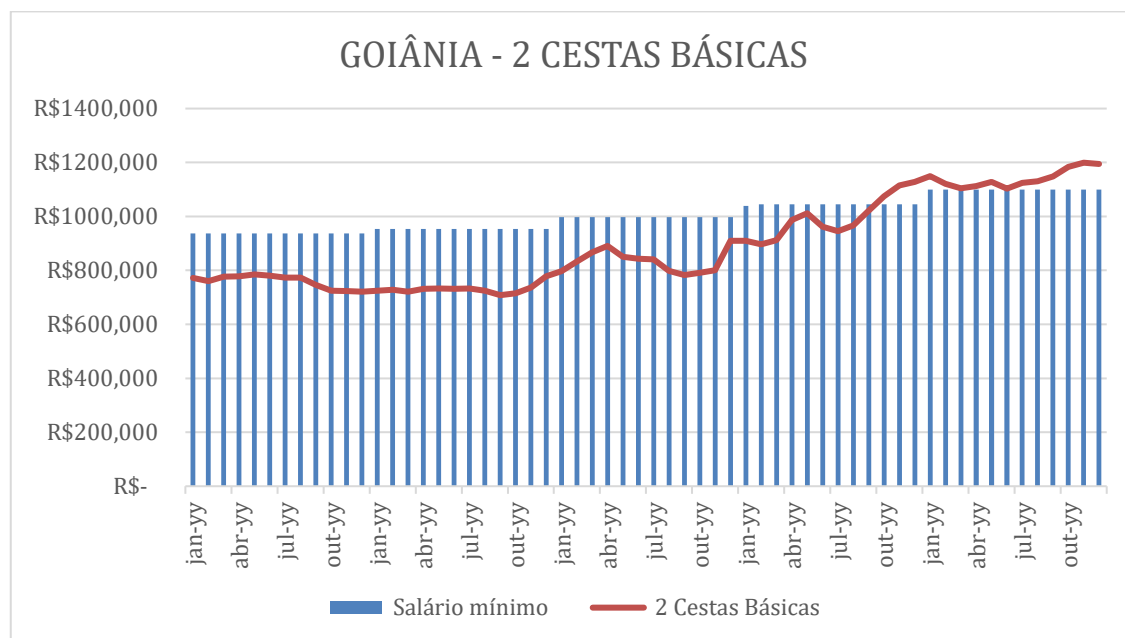
FONTE: O Autor (2022)

No gráfico 8 podemos observar a variação do piso salarial em Curitiba (azul) junto com a variação do preço dos produtos pesquisados somados (laranja), ilustrando qual seria a diferença entre o salário mínimo e valor necessário para comprar este conjunto de produtos. É possível notar que desde 2017 o piso salarial de Curitiba não supre o valor mínimo para adquirir todos os produtos, o melhor cenário fica no início de 2018, onde a linha de custo quase encosta na coluna do salário e, segundo a tabela de origem dos valores, o custo representa 108% do salário mínimo, enquanto que no final de 2021 essa diferença sobe para 144%. Isso mostra talvez por que um trabalhador com renda mínima, mesmo possuindo um veículo, é obrigado a adotar outros meios de transporte para o dia a dia, pois comida e gás certamente se tornam uma prioridade maior.

A seguir os gráficos sobre Goiânia, primeiramente com o valor de 2 cestas básicas em comparação ao piso salarial. A linha que representa o valor das duas

cestas básicas, no seu ponto mais baixo, em 2018, representa 78% do salário mínimo, e essa diferença sobe para 109% em 2021.

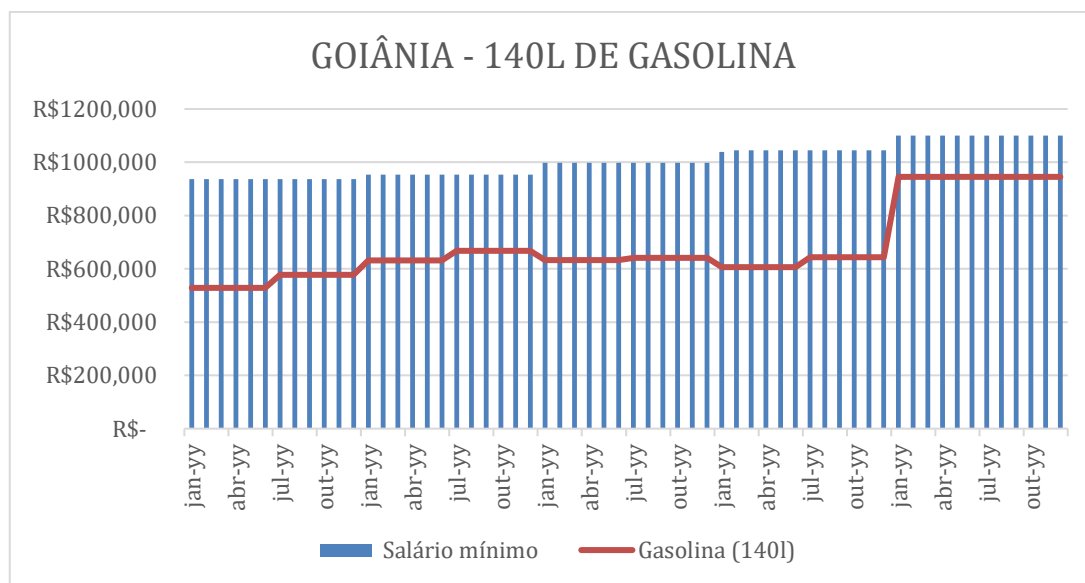
GRÁFICO 9 – GOIÂNIA 2 CESTAS BÁSICAS



FONTE: O Autor (2022)

A seguir, o gráfico compara o piso salarial de Goiânia com o valor de 140 litros de gasolina, que em 2021 representa 86% do salário mínimo.

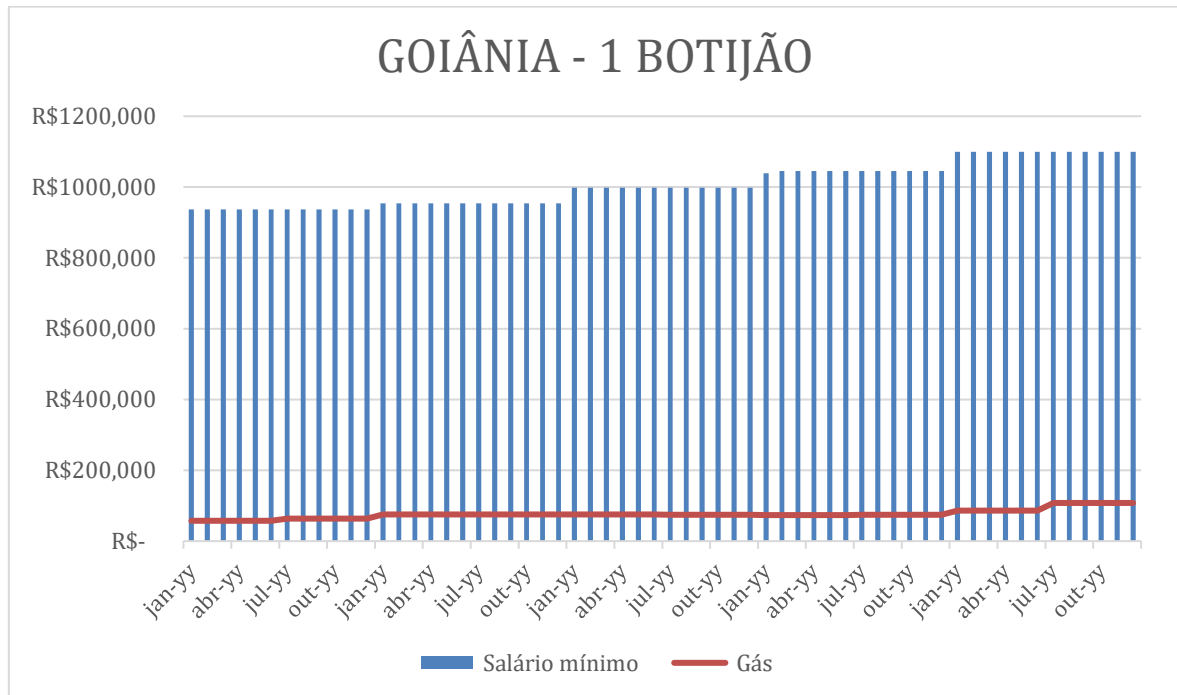
GRÁFICO 10 – GOIÂNIA 140L DE GASOLINA



FONTE: O Autor (2022)

O gráfico a seguir compara o piso salarial de Goiânia com o botijão de gás, que representa 6% do salário em 2017 e chega a 10% em 2021.

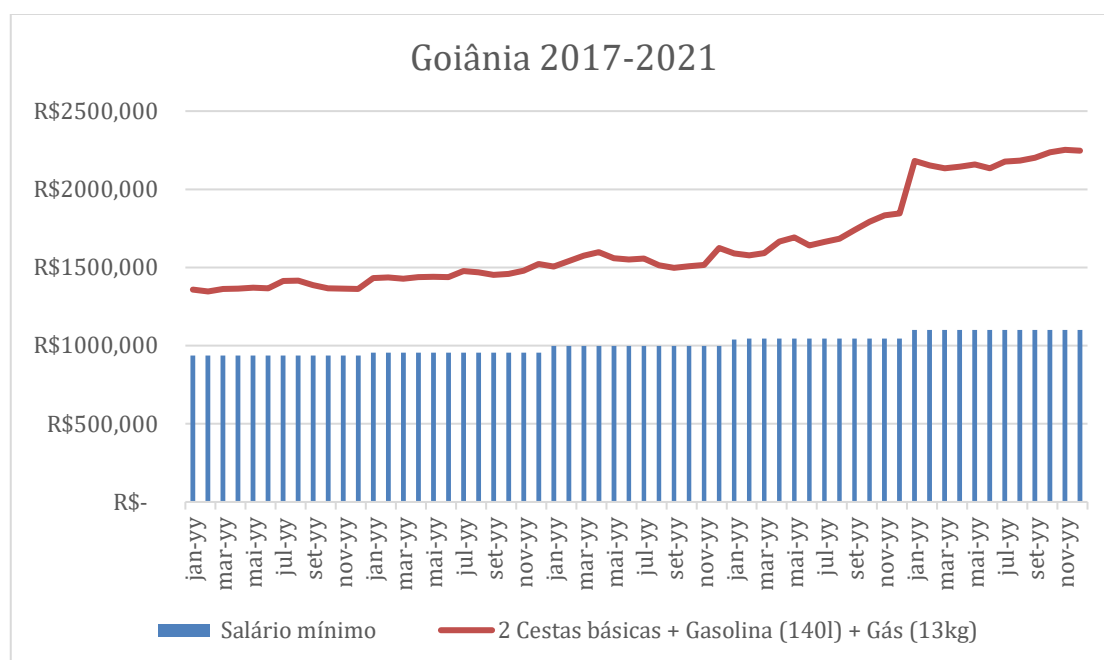
GRÁFICO 10 – GOIÂNIA 1 BOTIJÃO



FONTE: O Autor (2022)

O gráfico a seguir mostra os valores somados dos três produtos em comparação a evolução do salário mínimo em Goiânia.

GRÁFICO 11 – COMPARATIVO GERAL GOIÂNIA



FONTE: O Autor (2022)

No gráfico 11, podemos notar que a distância entre a linha de custo e as colunas do piso salarial possuem uma distância bem maior comparado a Curitiba, mostrando uma divergência ainda maior entre esses dois valores. O custo do conjunto de produtos, na sua menor diferença com o piso salarial, representa 143% do seu valor, e essa diferença chega a 204% no final de 2021, isso mostra que o custo de vida em Goiânia, considerando os produtos pesquisados, é bem maior do que as outras cidades pesquisadas, em especial Curitiba, onde a maior diferença entre os custos o salário é equivalente a menor diferença observada em Goiânia.

5 CONSIDERAÇÕES FINAIS

A utilização de técnicas de descoberta de conhecimento e de visualização da informação, mesmo que utilizadas apenas para fins acadêmicos, contribuem para compreensão de diversos fatores que envolvem essa tarefa e seus objetivos. No caso deste trabalho, a pesquisa por fontes de dados exploratórios revelou uma enorme quantidade de bases de dados abertos, públicos, disponíveis para análise e manipulação. A notoriedade das agências, como universidades públicas e agências federais, que depositam estas bases nos repositórios tornam as bases valiosas para estudos das mais diversas áreas do conhecimento, em particular para a ciência da informação.

As técnicas de visualização da informação e descoberta de conhecimento, aplicadas em bases de dados abertas, permitiram coletar e tratar de forma objetiva os dados referentes aos preços de produtos essenciais para o trabalhador brasileiro como a cesta básica e o combustível, mensurar a evolução desses preços, comparar entre diferentes cidades do Brasil e comparar com os reajustes do piso salarial. A pesquisa realizada apontou uma diferença significativa entre as cidades pesquisadas em relação ao custo da gasolina, do gás de cozinha e da gasolina, principalmente considerando a diferença entre os pisos salariais regionais. Com o auxílio dos gráficos, um recurso visual que valoriza e torna mais acessíveis os dados, podemos observar as discrepâncias dos preços praticados no país e o salário mínimo vigente. Quando os dados foram cruzados e observados desta maneira, foi possível notar de forma explícita as diferenças entre as cidades e regiões do país, Curitiba se mostrou uma cidade com preços mais acessíveis e salários maiores enquanto Goiânia apresentou custos muito maiores e um piso salarial inferior. Insights como esse são frutos valiosos em estudos de ciências sociais, políticas e econômicas e são possíveis graças ao tratamento visual e a aplicação de técnicas de descoberta de conhecimento, que neste trabalho foram aplicados em uma pequena amostra de cinco cidades, mas que podem ainda ser aplicadas em proporções muito maiores, como todo o Brasil e com isso, obter resultados ainda mais significativos para o país.

Vale registrar que, até o momento da conclusão deste trabalho (Maio/2022), os preços dos produtos pesquisados neste trabalho sofreram altas consideráveis, como a gasolina, que até o final de 2021, em Curitiba teve sua média em torno de

R\$5,50, e em 2022 chegou a quase R\$8,00. Diversos produtos que compõe a cesta básica também sofreram altas e isso certamente representa uma divergência ainda maior entre o salário do trabalhador e seu custo de vida.

REFERÊNCIAS

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMITH, P. From datamining to knowledge discovery in databases. *AI Magazine*, v. 17, n. 3, p. 37–54, Fall 1996. Disponível em: <<https://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf>> Acesso em: 18 Nov. 2021.

HALMENSCHLAGER, Carine. Um algoritmo para indução de árvores e regras de decisão. 112 f. Tse (Mestrado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, 2002. Disponível em: <<https://www.lume.ufrgs.br/handle/10183/2755>> Acesso em: 18 Nov. 2021.

FONSECA, S. O; NAMEN, A. A. Um modelo de classificação para a descoberta de conhecimento relacionado a professores de matemática do ensino básico. In: ENCONTRO DE MODELAGEM COMPUTACIONAL, 16. 2013, Ilhéus, Bahia, Ilhéus. Rio de Janeiro: Rede Sirius, UERJ, 2013, v.1. Disponível em: <<https://www.bdtu.uerj.br:8443/handle/1/13790#preview-link0>> Acesso em: 20 Nov. 2021.

GÓES, Anderson; STEINER, Roges. Proposta de metodologia para a criação de etiqueta de classificação – estudo de caso: desempenho escolar. Artigo, [s. l.], 2016. DOI <https://doi.org/10.1590/0104-530X810-13>. Disponível em: <<https://www.scielo.br/j/gp/a/YB3VKkyhbnPhkmNYW9MxP3P/?lang=pt>> Acesso em: 15 Nov. 2021.

MIRANDA, R. C. da R. "O uso da informação na formulação de ações estratégicas pelas empresas". *Ciência da Informação*, Brasília, v.28, n.3, p.286-292, set./dez. 1999. Disponível em: <<https://www.scielo.br/j/ci/a/r7L9msHr6FfrYpJ5PKk8fsS/?format=pdf&lang=pt>> Acesso em: 15 Nov. 2021.

DAVENPORT, T. H.; PRUSAK, P. Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual. Rio de Janeiro: Elsevier, 2003.

SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. Sistema de banco de dados. Rio de Janeiro: Elsevier, 2006. Disponível em: <<https://www.docsity.com/pt/banco-de-dados-silberschatz/4784644/>> Acesso em: 21 Nov. 2021.

CARD, S.K, MACKINLAY, J. D. e SHNEIDERMAN, B. Readings in Information Visualization: Using Vision to Think. Morgan Kaufman Publishers, 1999.

FREITAS, C. M. dal S. et al. Introdução à visualização de informações. RITA, Porto Alegre, v. 8, n. 2, 2001. Disponível em: <<https://lume.ufrgs.br/bitstream/handle/10183/19398/000300210.pdf?sequence=1&isAllowed=y>> . Acesso em: 25 Nov. 2021.

BARIONI, M. C. N. Visualização de operações de junção em sistemas de bases de dados para mineração de dados. 2002. Dissertação (Mestrado)-Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2002. Disponível em <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-04122002-114251/publico/dissertacao_MCNB.pdf> Acesso em: 22 Nov. 2021.

WONG, P.C. Visual data mining. IEEE Computer Graphics and Applications, Los Alamitos, v.19, no.5, p. 20-21, Sep./Oct. 1999. Disponível em: <<http://ivizlab.sfu.ca/arya/Papers/IEEE/C%20G%20&%20A/1999/October/Visual%20Data%20Mining.pdf>> Acesso em: 25 Nov. 2021.

SHIMABUKURU, H. M. Visualização temporal em uma plataforma software extensível e adaptável. 2004. Tese (Doutorado)-Instituto de Ciência Matemática e de Computação. Universidade de São Paulo, São Carlos, 2004. Disponível em <<https://teses.usp.br/teses/disponiveis/55/55134/tde-03042005-132736/pt-br.php>> Acesso em: 25 Nov. 2021.

RODRIGUES, H. L. S.; FERREIRA, J. R. B. Visualização da informação como ferramenta de apoio ao tratamento de dados empresariais. Colloquium Exactarum, Pres. Prudente, v.9, n.2, p.114-130, 2017. Disponível em: <https://revistas.unoeste.br/index.php/ce/article/view/1656/1985>. Acesso em: 12 mai. 2022.