

## **1. A QUESTÃO DO NEGÓCIO**

### **1.1 Qual a minha meta?**

A empresa que me contratou, tem como objetivo expandir uma de suas lojas, em uma rede de 45.

Minha meta, com os dados fornecidos, é realizar uma avaliação e investigar qual loja tem o melhor potencial para que valha a pena ser expandida, esclarecendo as razões pelas quais ela é a escolhida.

Para isso, desenvolvi a análise inicial para procurar a loja ideal, a qual se trata de buscar e detectar qual loja tem tido a maior demanda ao longo do período analisado e as características que são favoráveis, para que respalde a sua expansão.

### **1.2 Como posso chegar?**

Utilizando-me dos dados fornecidos, o caminho a ser encontrado é aquele onde haja a oportunidade de expandir a loja onde as vendas e a demanda estejam maiores, com uma região de economia saudável e forte, para justificar a expansão no longo prazo. A lógica abordada foi, primeiramente, realizar uma filtragem que verifique os acúmulos de vendas totais, ao longo de todo período, analisando também o desempenho individual anual, para fins de comparação do crescimento.

Com este primeiro filtro estabelecido, após selecionar as que tiveram os melhores desempenhos, verificarei a saúde da economia da região por meio dos dados fornecidos, por exemplo, inflação e preço do combustível, para justificar a expansão em uma área que seja sustentável.

*(Ref. Colab: Grupamento 1)*

## **2. O ENTENDIMENTO DO NEGÓCIO**

### **2.1 Quais dados eu tenho disponível?**

Os dados disponibilizados para a realização dessas análises são os dados oficiais do Walmart, com informações referentes a economia e desempenho de suas 45 lojas, como inflação da região, preço dos combustíveis, números de vendas, dentre outros.

Verifiquei seus formatos bem como observei as necessidades iniciais de tratamento dos dados.

*(Ref. Colab: Grupamento 1)*

### **2.2 Eles são relevantes para o problema?**

Alguns dados se tornam importantes de acordo com a estratégia escolhida, outros, porém, se tornam irrelevantes por não constituírem importância para o alvo do estudo.

Os mais importantes, na primeira análise, foram aqueles que demonstraram melhor a evolução das vendas de cada loja ao longo do período. Por meio deles, foi possível detectar as melhores lojas em desempenho de vendas.

*(Ref. Colab: Grupamento 1)*

### **2.3 Eles me trazem uma solução direta?**

Não, é necessário um tratamento e filtragem dos dados escolhidos para que tragam as informações e levem ao conhecimento necessário para a tomada de decisão mais assertiva.

Neste caso, realizei a correção das datas e filtrei as quantidades de vendas, para posteriormente comparar seus desempenhos e conseguir selecionar as melhores concorrentes para expansão.

*(Ref. Colab: Grupamento 1)*

## **2.4 Qual o meu setor ou range de negócio?**

O setor de varejo, faz parte do setor terciário, o qual é o setor econômico responsável pelas vendas dos produtos voltados ao consumidor final, já tendo passado pela extração e produção.

## **3. A COLETA DE DADOS**

### **3. 1 Os dados que eu tenho fazem sentido?**

Após uma avaliação criteriosa e separação dos dados mais relevantes, notei que algumas informações possuem sentidos conectados que podem trazer insights.

Sendo assim, alguns itens demonstram claramente o desenvolvimento da loja, elevando a necessidade de expansão.

*(Ref. Colab: Grupamento 2 a 4)*

### **3. 2 Estão no formato que eu gostaria?**

A maioria sim, sendo o suficiente para as análises desejadas, com exceção da data, que estava em formato de Object, a qual foi convertida para datetime para facilitar operações.

*(Ref. Colab: Grupamento 1)*

### **3. 3 O que mais eu consigo obter de informações destes dados?**

Com os dados, pude observar que algumas lojas, embora possuíssem maior acúmulo, não necessariamente eram as que mais se desenvolveram ao longo do período.

*(Ref. Colab: Grupamento 1 a 6)*

## **4. A LIMPEZA DE DADOS**

### **4. 1 Remoção de dados outliers e valores ausentes.**

Foi utilizado um filtro ao longo do projeto, que serviu para selecionar as lojas com o melhor desempenho. Todos mais detalhados no próprio colab.

No 'Grupamento 1' (do colab) verifiquei se haviam valores nulos, e não haviam. Observei também a quantidade de dados por loja, que trouxe 143 dados, ou seja, 3 anos de dados semanais guardados individualmente.

No primeiro momento, tratei o formato das datas, para que facilitasse cálculos que funcionam com datetime.

Como optei por seguir uma linha de indicação de lojas baseada no desenvolvimento das vendas ano a ano individualmente, as lojas que possuíam um fraco desempenho foram eliminadas do estudo, para focar a expansão nas melhores lojas, mantendo apenas a loja mais fraca, isto é, a loja 33, para fins de comparação.

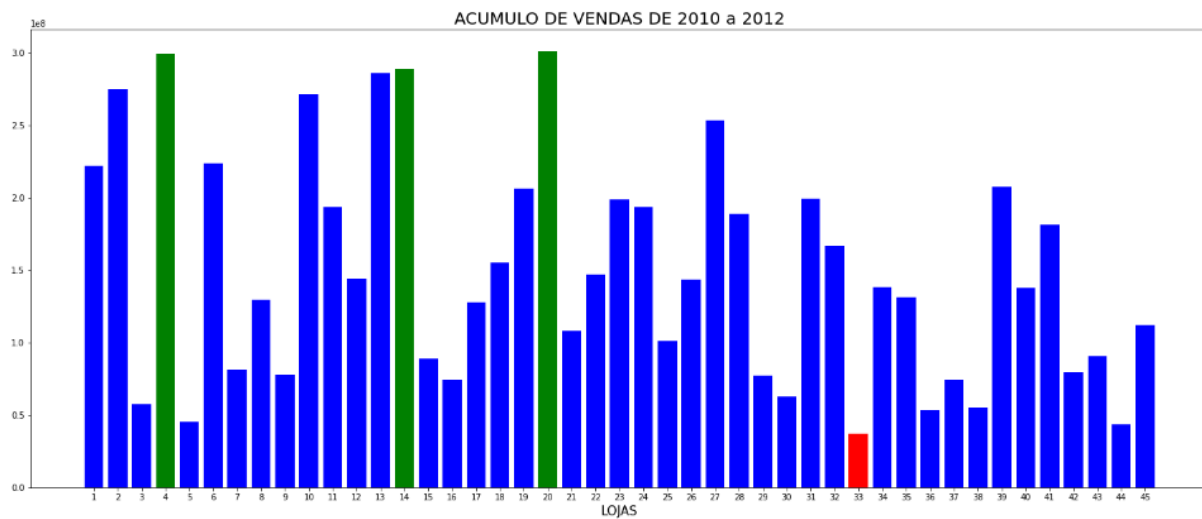
*(Ref. Colab: Grupamento 1 a 4)*

## **5. A EXPLORAÇÃO DE DADOS**

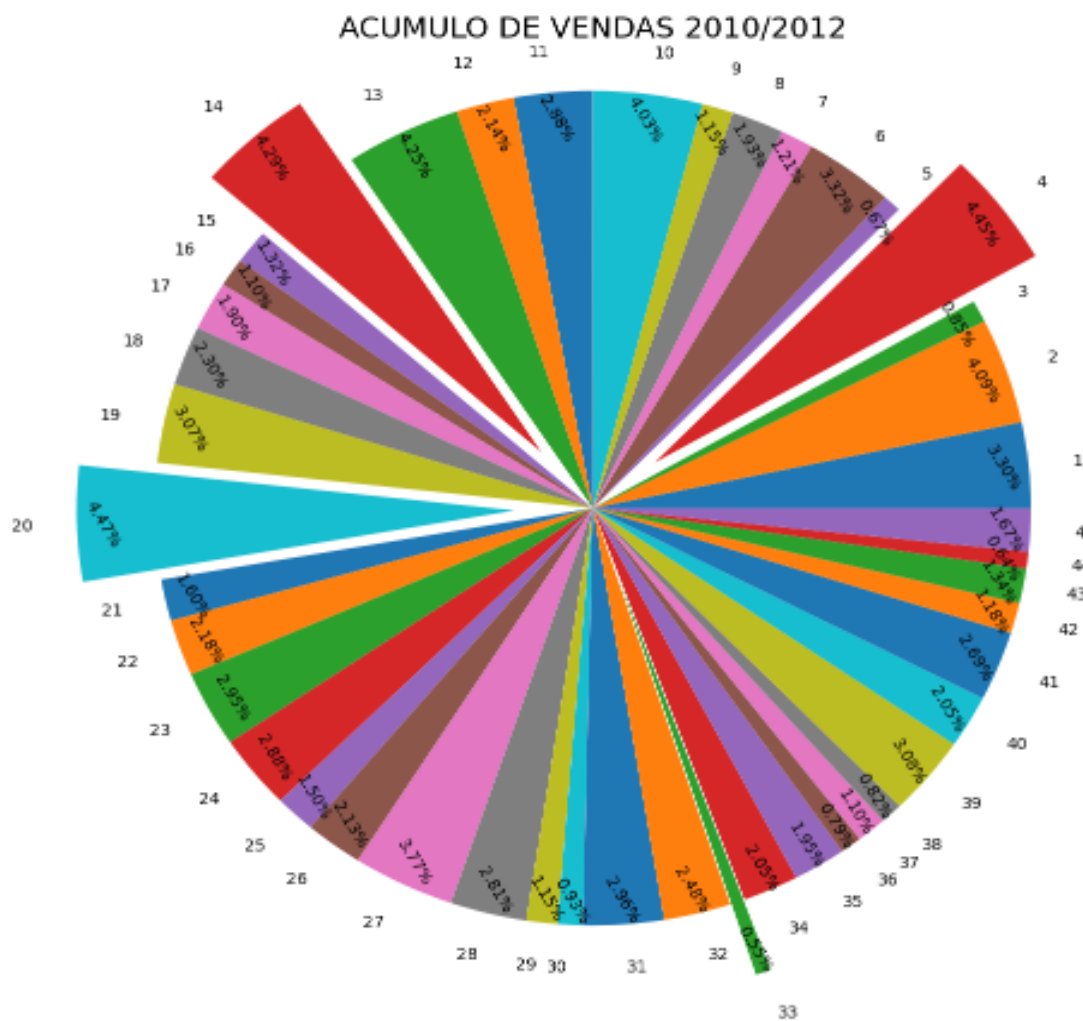
### **5. 1 Visualização dos dados.**

Para apoiar as conclusões e visualizar a percepção coletada dos dados, realizei alguns gráficos, os quais estão no colab também.

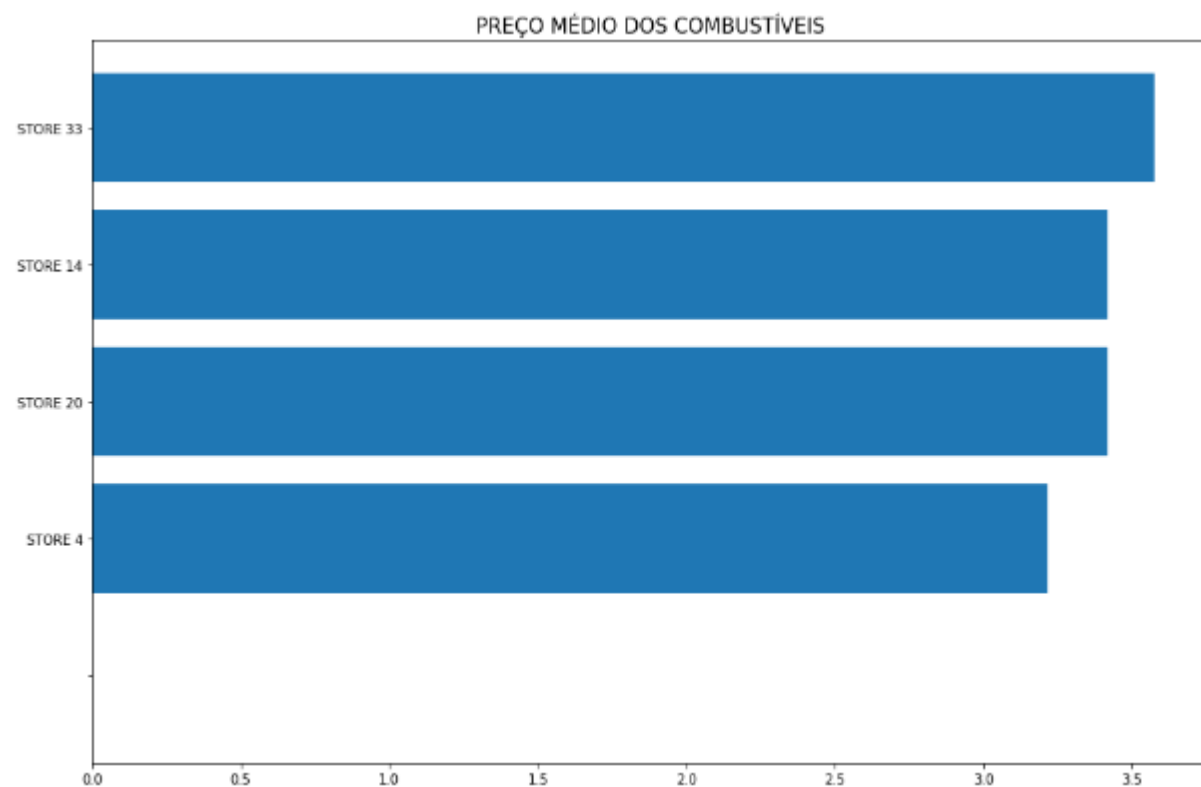
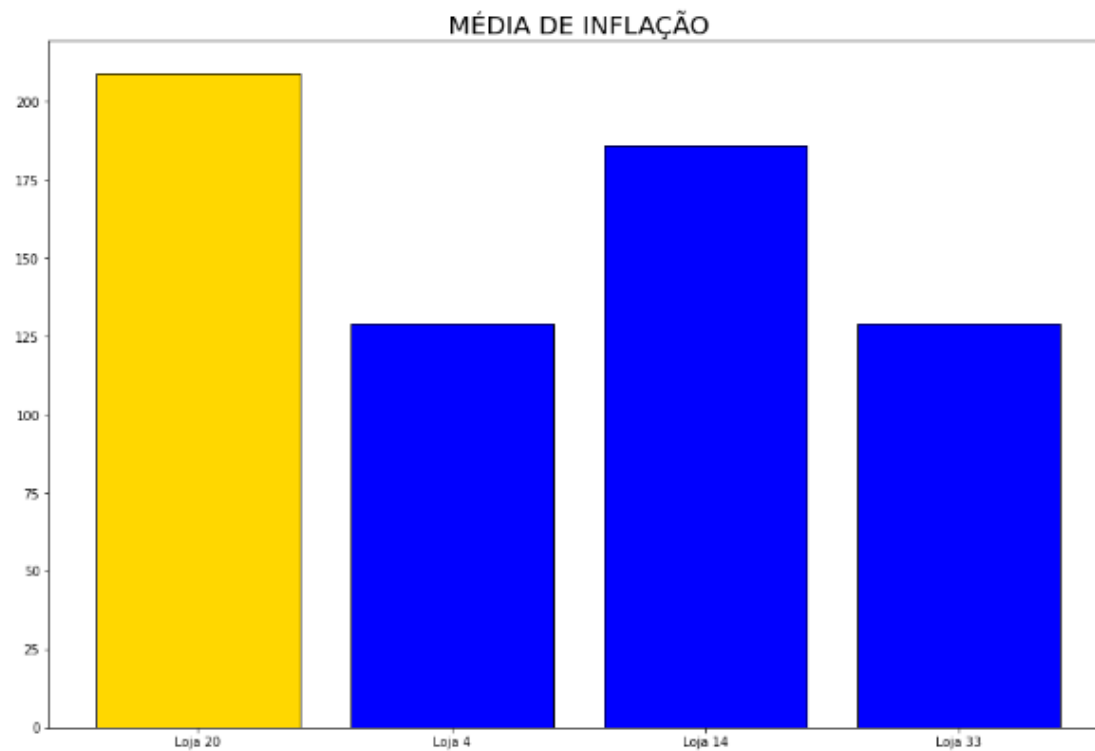
O primeiro e o segundo exibem o acúmulo total das vendas de cada loja, o qual notamos a Loja 20 em destaque claramente acima das demais, entretanto, isso não reflete o ideal de expansão, pois o desempenho pode ser afetado por um dado fora do padrão, assim como a média é afetada por outliers.

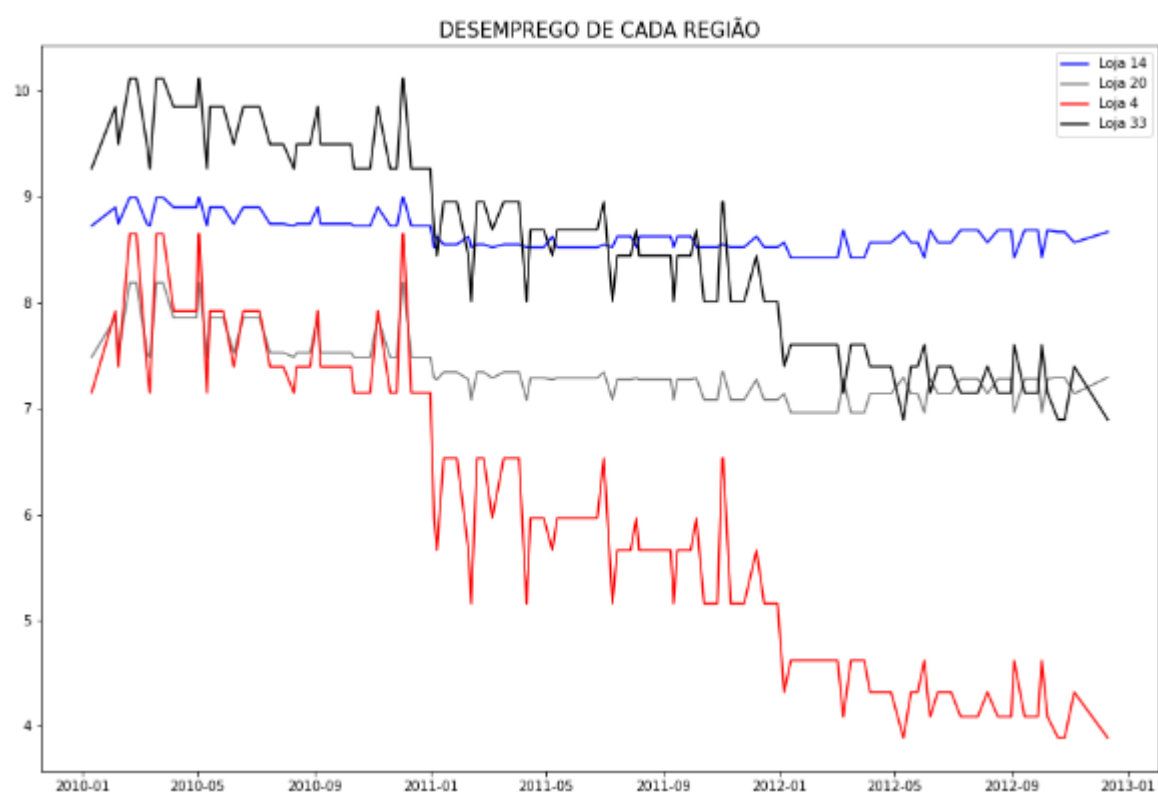
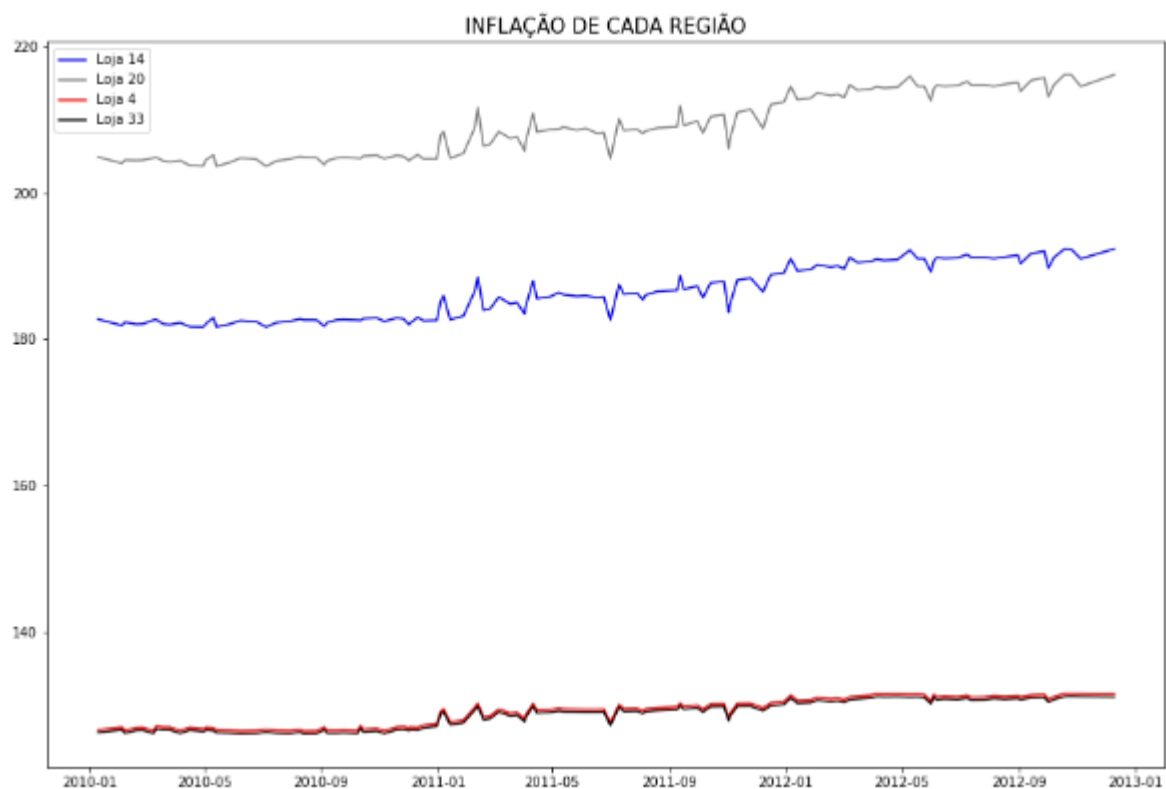


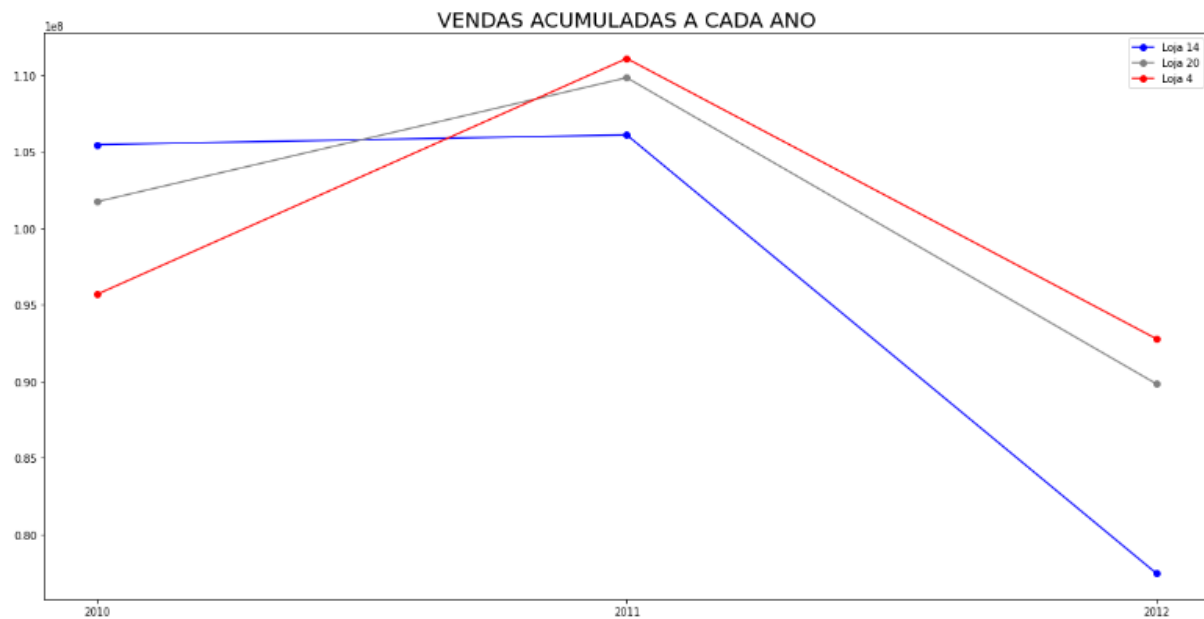
Os totais das vendas acumulados, destacando as 3 melhores: Loja 4, 14 e 20. E a pior: 33.



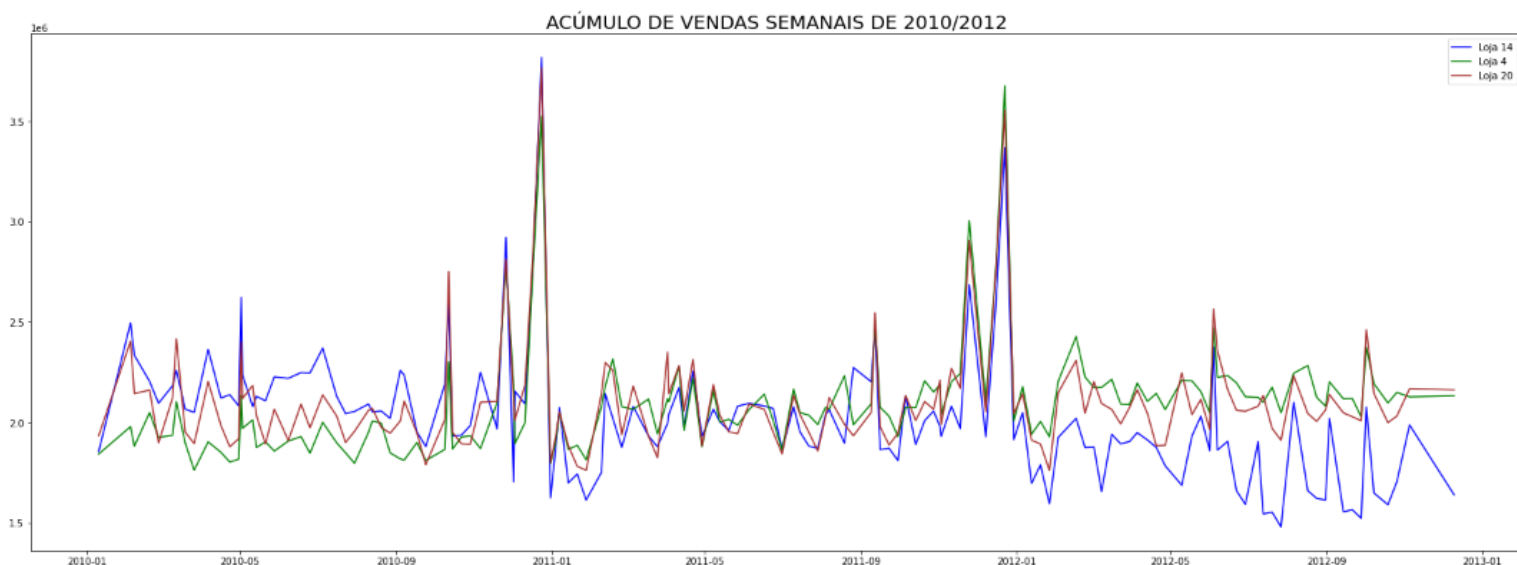
Porém, estas informações não são suficientes, fazendo a comparação, notei que a Loja 4 é a que mais possui crescimento anual, superando até mesmo a Loja 20, que é a maior em acúmulo de vendas no período somado. Bem como, possui a inflação mais baixa, preço do combustível mais baixo e também o menor índice de desemprego, como destacado nos gráficos abaixo:







Inclusive, além de detectar que a loja 4 teve um desempenho superior ao da loja 20 anualmente, detectei também que o pico mais alto de 2010, foi referente ao feriado de natal na loja 14, o qual não se repetiu mais em nenhum dos anos seguintes



	Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unemployment	Temp_Celsius	Mean_Sales
1905	14	2010-12-24	3818886.45	0	30.59	3.141	182.54459	8.724	-0.783333	2020878.0

O maior pico de vendas na véspera do Natal de 2010, ajudou a elevar a Loja 14 para terceira colocada em vendas no total acumulado, porém, como vemos no gráfico acima, o desempenho de vendas dela foi decaindo com o passar do tempo.

## **Análise Final + Sugestão de Store para expansão.**

Avaliando os dados, após filtrar todas essas informações com base no escopo estratégico mencionado, em um primeiro momento, a que parecia ser melhor para expandir, era a Loja 20, pois analisando superficialmente, era a loja com o melhor desempenho e com maior acúmulo de vendas.

Porém, ao compararmos os dados regionais e os desempenhos anuais, tudo muda. Observe:

1 - Ficou destacado que, embora no acumulado do período a Loja 20 seja a melhor, a Loja 4 em 2011 alcança e em 2012 ultrapassa as vendas da Loja 20 em 3,28%. Uma alta que vem ganhando força desde sempre. (Ref: Colab - Grupamento 6)

2 - A inflação da região da Loja 20 é mais alta do que da Loja 4, o que talvez segure o desempenho da Loja 20, enquanto que o custo da expansão levando-se em conta a inflação da região, na Loja 4 custaria mais barato.

3 - A média dos preços de combustível da região da Loja 4 é a mais baixa de todas as melhores comparadas.

4 - O número de desempregados na região da Loja 4 é o mais baixo de todos, demonstrando uma região com desenvolvimento e atividade econômica mais forte, o que possivelmente resulte no aumento da demanda pelos produtos da Walmart na região, que justificaria muito bem nossa expansão.

Com todos estes motivos analisados, concluo e indico a expansão da STORE 4, pois está em um momento de crescimento da região, que nossa expansão pode vir a aproveitar com mais eficiência o que já tem sido sentido nos últimos resultados.

(Ref: Grupamento 6)

Segue link do colab:

<https://drive.google.com/file/d/1Fy6gOog3WLANSxqK4qbNYDS6xyHW3vnt/view?usp=sharing>