**DELIVERABLE #2 – CASO WALLMART**

**Link do Colab**

https://colab.research.google.com/drive/1tZrn07e0NQCZPrzlHvXeybyi\_MbduvCM?usp=sharing

**Contextualização**

Você foi contratado pelo Wallmart para fazer um levantamento do faturamento das lojas nos USA e apontar qual loja seria melhor para expandir seu tamanho.

É preciso analisar as vendas semanais de cada loja, calcular algumas informações importantes que serão perguntadas e ao final de tudo, indicar em qual loja deva ser investida.

Walmart, Inc., é uma multinacional estadunidense de lojas de departamento.

A companhia foi fundada por Sam Walton em 1962, incorporada em 31 de outubro de 1969 e feita capital aberto na New York Stock Exchange, em 1972. No ano de 2021, obteve um um lucro de $13.51 Bilhões. Sendo uma das principais lojas de varejo do mundo, os dados contemplam as vendas semanais de 45 lojas espalhadas pelos Estados Unidos. O Walmart realiza vários eventos promocionais de descontos ao longo do ano. Essas remarcações precedem feriados importantes, os quatro maiores de todos, que são o Super Bowl, o Dia do Trabalho, o Dia de Ação de Graças e o Natal. As semanas que incluem esses feriados têm um peso maior.

**Sobre os Dados**

Dataset: <https://www.kaggle.com/yasserh/walmart-dataset>

Os dados apresentam as vendas semanais de 05/02/10 até 01/11/12 em 45 lojas varejistas da rede Wallmart e também algumas métricas econômicas e meteriológicas por semana.

**A Questão de Negócio**

- Qual a minha meta?

- Encontrar a loja para receber investimento de expansão. São 45 lojas ao todo.

- Como posso chegar?

- Entender o histórico de faturamento das 45 lojas

- Entender a correlação do faturamento com as outras informações do dataset.

- Criar novos indicativos que possam auxiliar da decisão.

**O Entendimento do Negócio**

- Os dados que eu tenho são relevantes para o problema?

- Date: Consigo ter noção da evolução de todos os indicativos ao longo do tempo.

- Holiday\_Flag: Posso filtrar as vendas com e sem feriado. Calcular a porcentagem de aumento de faturamento nos feriados. Ver a correlação das vendas com feriado.

- Temperature: Fazer correlação com aumento do faturamento.

- Fuel\_Price: Fazer correlação com aumento do faturamento.

- CPI: Fazer correlação com aumento do faturamento.

- Unemployment: Fazer correlação com aumento do faturamento.

- Eles me trazem uma solução direta?

- Não, precisaremos calcular e organizar as informaçoes.

- Qual o meu setor ou range de negócio?

- Mercado de varejo.

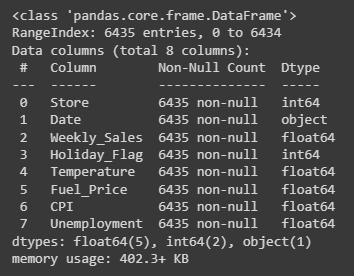
**A Coleta de Dados**

- Os dados que eu tenho fazem sentido?

- Todos os dados do dataset fazem sentido.

- Estão no formato que eu gostaria?

-Ao executar a função .info() percebi que as datas não estão no formato desejado.



Fiz a conversão da data para to\_datetime.

**A Limpeza de Dados**

- Remoção de dados outliers e valores ausentes

- Não havia nenhum valor nulo e decidi não remover os outliers pois eram poucos.

**A Exploração de Dados**

**- Visualização dos dados**

Primeiramente criei uma tabela para entender a correlação de todos os dados referentes às vendas semanais totais por loja. Já foi possível perceber que muitas lojas possuem uma correlação bem baixa com os dados disponíveis, o que acarreta num menor controle de análise com as informações do dataset.

Por isso, usando a função .describe() e auterando os percentiles mínimo para 0.05 e máximo para 0.95 consegui verificar a tendência de correlação para cada dado com a venda.

Filtrei os dados da seguinte maneira:

Holiday\_Flag - valores acima de 0.24

Mesmo sendo um dado de marcação, é possível saber se as vendas vão aumentar por conta do feriado. Como não me interessa correlação negativa, filtrei por 0.24 que representa o percentile 95%. Dessa forma, selecionei apenas as lojas que realmente respondem positivamente a um feriado. Considero que é um fator fundamental para analisar se a população é compradora ou não.

Temperature – valores acima 0.17e abaixo de -0.378

Fuel\_Price – valores acima de 0.32 e abaixo de -0.416

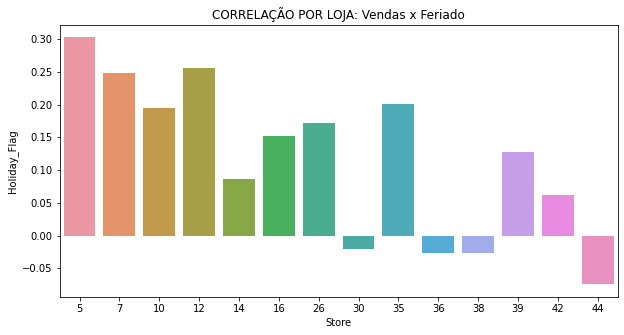
CPI'] – valores acima de 0.42 e abaixo de -0.396

Unemployment – valores abaixo de -0.38: Considerei que os valores de Taxa de desemprego acima disso não seriam boas opções para investimento pois se há um crise inesperada a taxa de desemprego pode subir e isso influencia nas vendas.

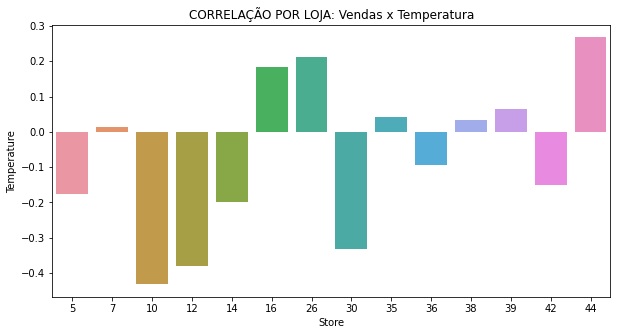
Dessa forma, consegui filtrar muitas lojas chegando a 14 na lista.

Lojas: 5, 7, 10, 12, 14, 16, 26, 30, 35, 36, 38, 39, 42, 44.

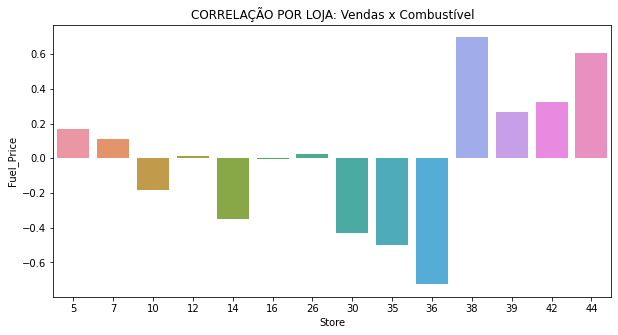
Ainda analisando correlação, agora apenas das lojas escolhidas, fiz um gráfico para cada dado.



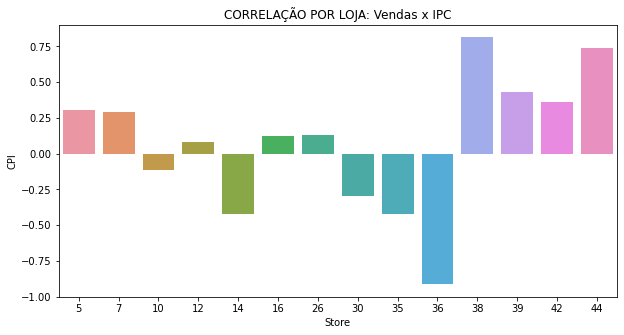
Nesse gráfico percebo que as lojas 30, 36, 38 e 44 possuem uma correlação negativa com o feriado. Decido retirá-las da lista na próxima etapa.



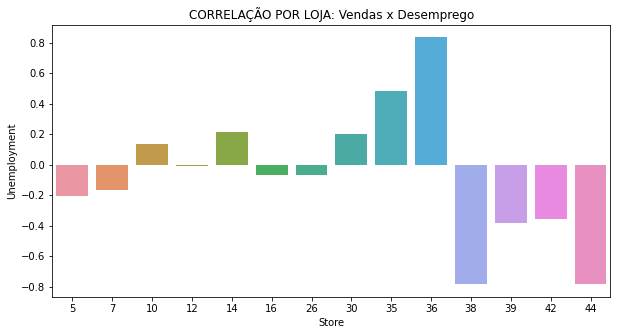
As lojas que mais apresentam correlação com a Temperatura apresenta uma correlação negativa. Nessa análise não dá pra saber se faz sentido ou não, tendo em vista que não sabemos a localização das lojas.



A correlação vendas x combustível se mostra, também, algo um pouco subjetivo. As pessoas não necessariamente vão fazer compras de carro. Elas podem ir de transporte público.



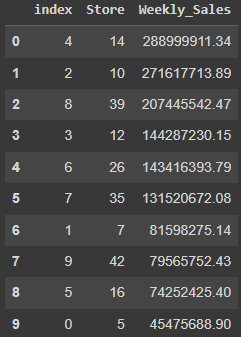
A correlação entre as vendas e o IPC positiva pode significar uma população que está consumindo. Sendo assim, prefiro não levar em conta neste momento para tomada de decisão pois os números está divididos entre positivos e negativos.

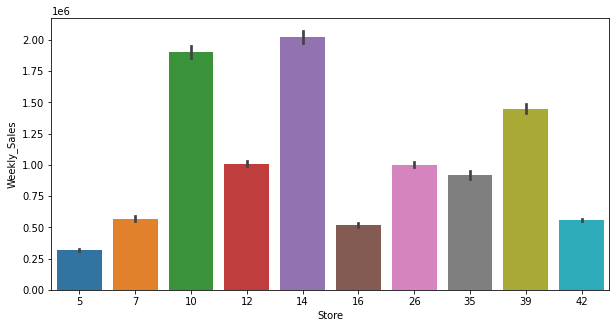


A correlação negativa entre vendas x desemprego é bem relevante. Neste gráfico me atento a possíveis candidatas como as lojas 49 e 42. Lembrando que as lojas 38 e 44 (que apresentam bons resultados para essa correlação) foram descartadas por conta da correlação negativa com o feriado. Entendo que quanto menor a taxa de desemprego, mais pessoas poderão consumir.

Após estas análises, decido eliminar apenas as lojas que respondem negativamente ao feriado. Ficando com as seguintes lojas: 5, 7, 10, 12, 14, 16, 26, 35, 39 e 42.

Neste momento, agrupo as lojas com a soma total das vendas e ordeno de forma decrescente para encontrar as lojas que possuem os maiores faturamentos.



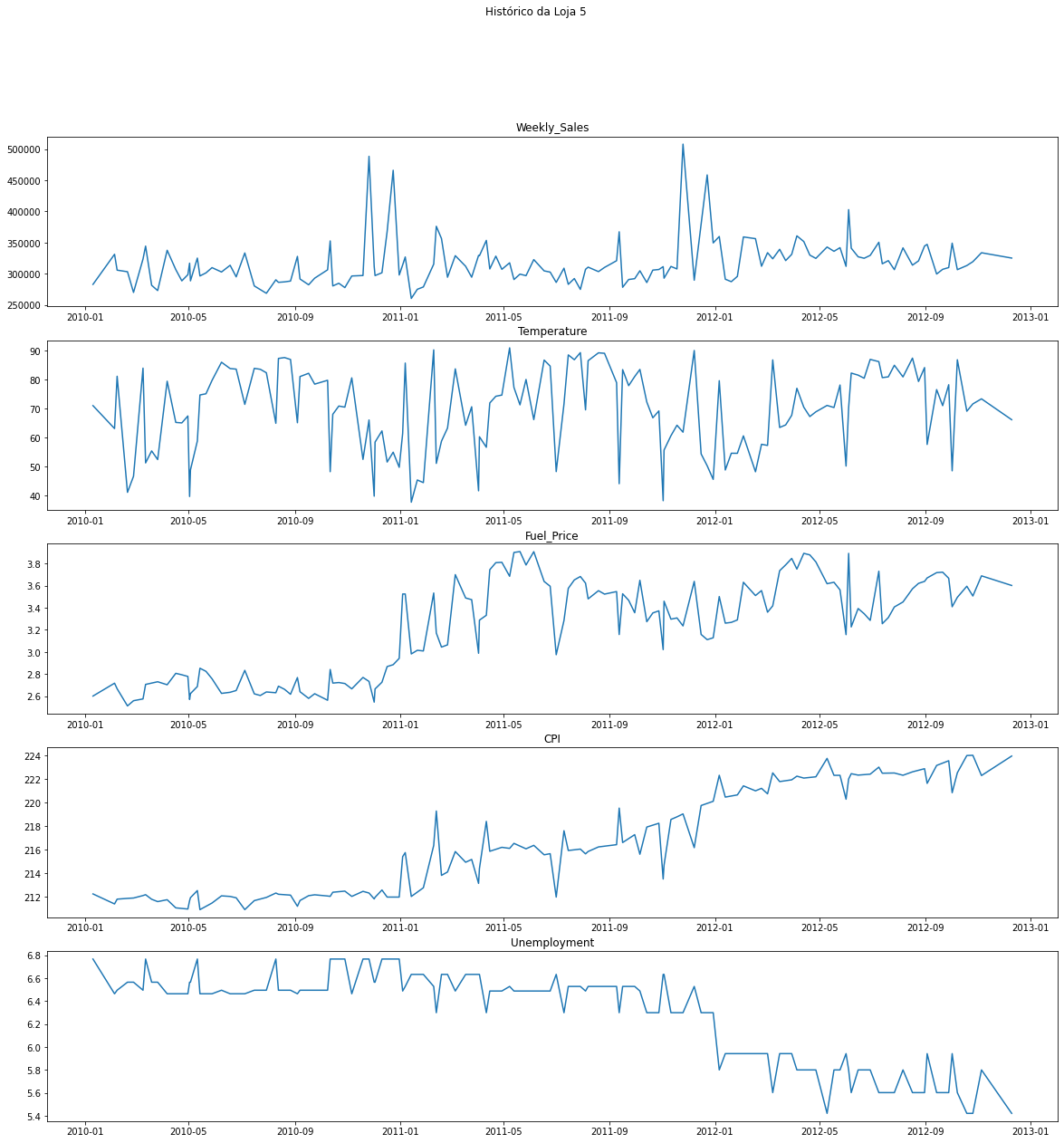


Lojas top3 em faturamento: 14, 10 e 39.

Decido selecionar apenas 5 lojas para a decisão final: as top3 em faturamento 14, 10 e 39 e as duas que melhor reagem ao feriado: 5 e 12.

Realizo uma análise da evolução de cada fator ao longo do tempo. Um dos fatores que vou mais levar em conta é o faturamento da loja hoje diante do seu histórico. Para cada loja ainda determino a média do faturamento.

**LOJA 5: Média do faturamento: US$ 310,338.17**



Faturamento: Atuando na média para cima nos últimos meses.

Temperatura: Oscila muito, inconclusiva.

Preço do combustível: Em alta

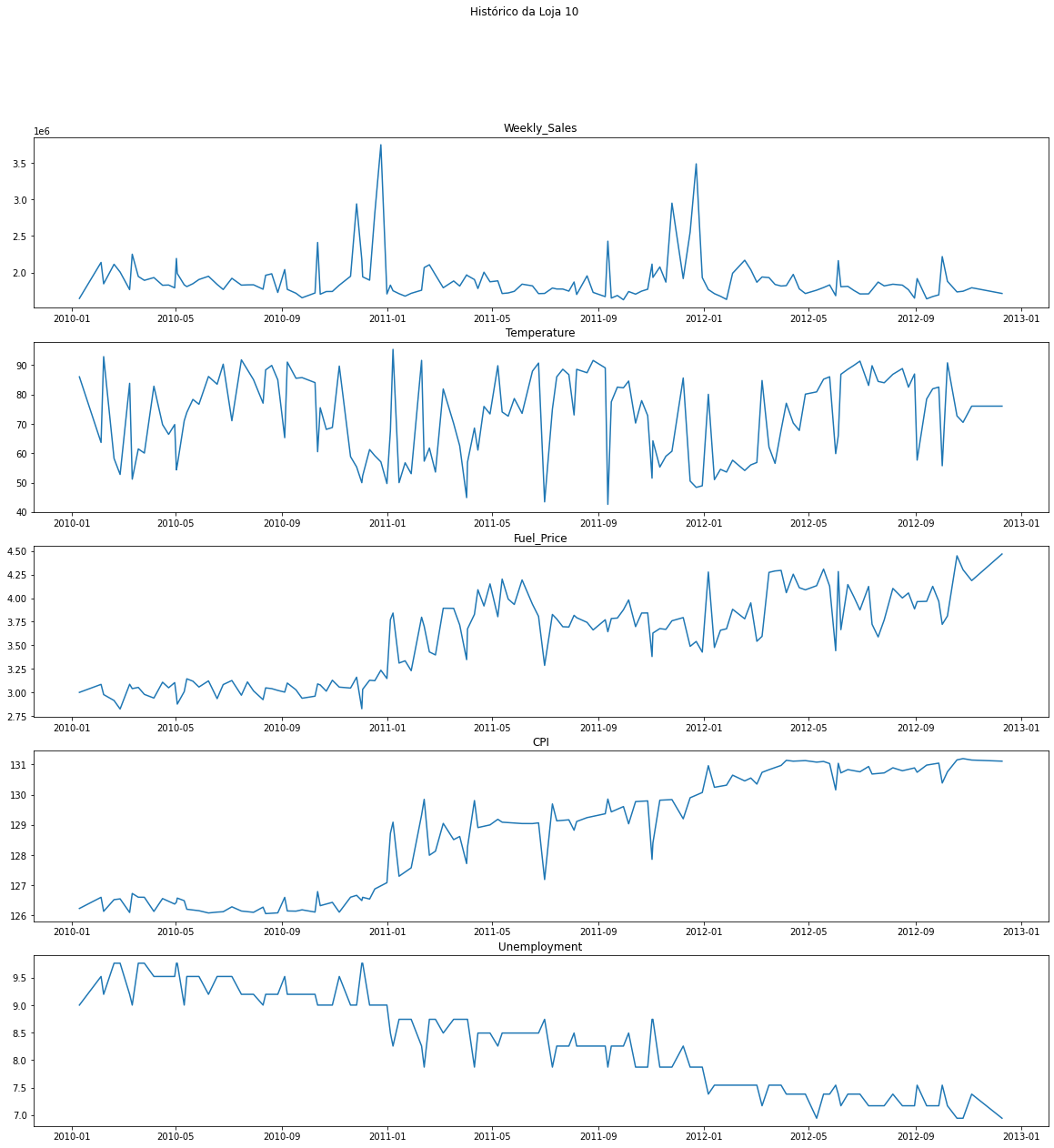
IPC: Em alta

Taxa de Desemprego: Em baixa

Olhando para análises anteriores:

A Loja 5 tem uma correlação positiva com o preço do combustível, mas, como já apontado, pode não ser uma correlação verdadeira. Tem uma boa correlação positiva com IPC e negativa com taxa de Desemprego.

**LOJA 10: Média do faturamento: US$ 1,827,521.71**



Faturamento: Atuando na média para baixo nos últimos meses.

Temperatura: Oscila muito, inconclusiva.

Preço do combustível: Em alta

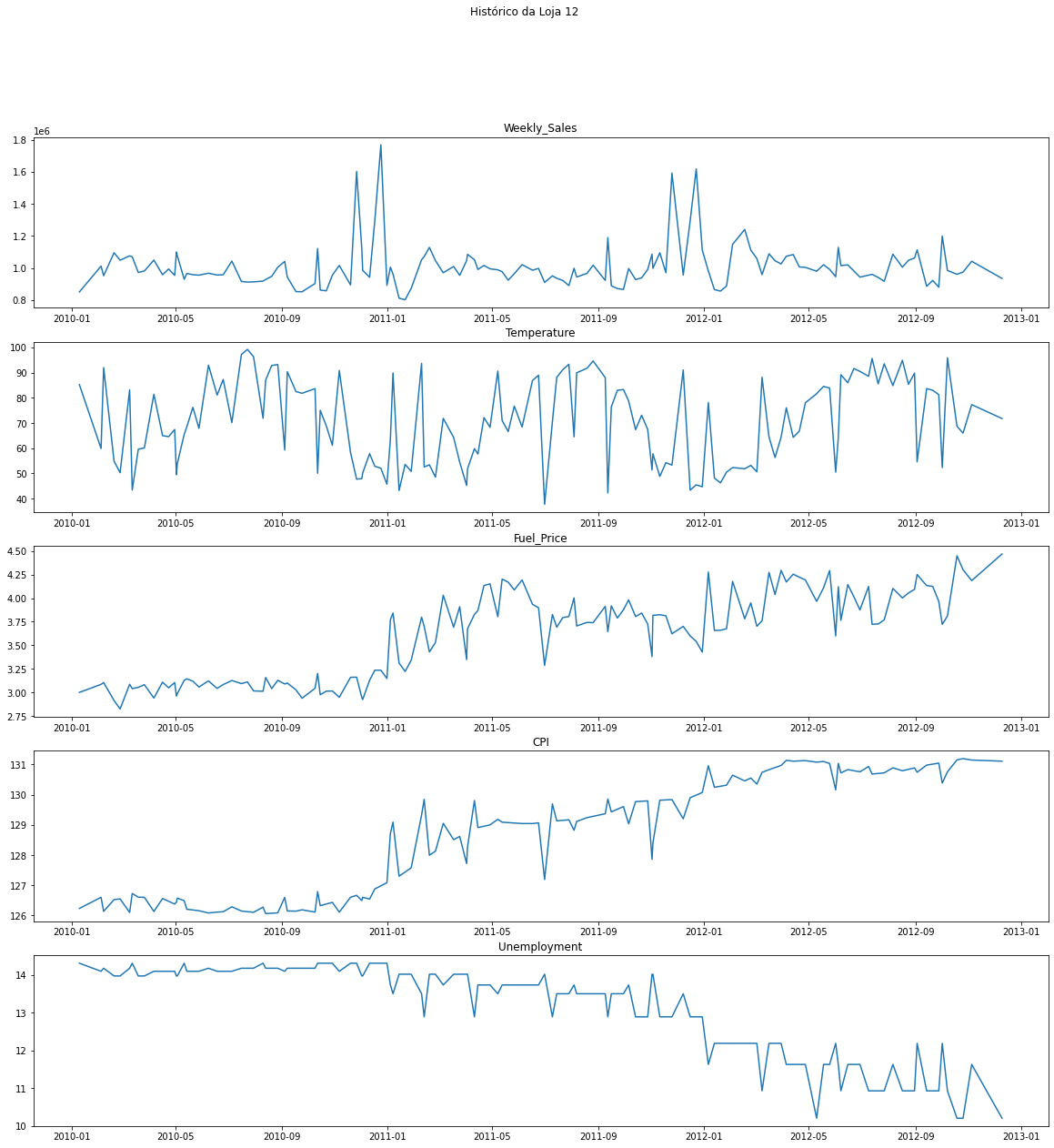
IPC: Em alta

Taxa de Desemprego: Em baixa

Olhando para análises anteriores:

A Loja 10 tem uma correlação negativa com o preço do combustível e, mesmo que possa não ser uma correlação totalmente verdadeira, é um fator negativo pois o preço do combustível está em alta. Mesma coisa vale para o IPC. A Loja 10 não é uma boa opção

**LOJA 12: Média do faturamento: US$ 981,615.81**



Faturamento: Atuando na média nos últimos meses.

Temperatura: Oscila muito, inconclusiva.

Preço do combustível: Em alta

IPC: Em alta

Taxa de Desemprego: Em baixa

Olhando para análises anteriores:

A Loja 12 tem uma correlação irrisória, comparada com as outras 4, com os outros. A análise decisiva não mostrou nada que possa melhoras este cenário. Desta forma a Loja 12 não é uma boa opção.

**LOJA 14: Média do faturamento: US$ 2,004,330.30**



Faturamento: Atuando abaixo da média nos últimos meses.

Temperatura: Oscila muito, inconclusiva.

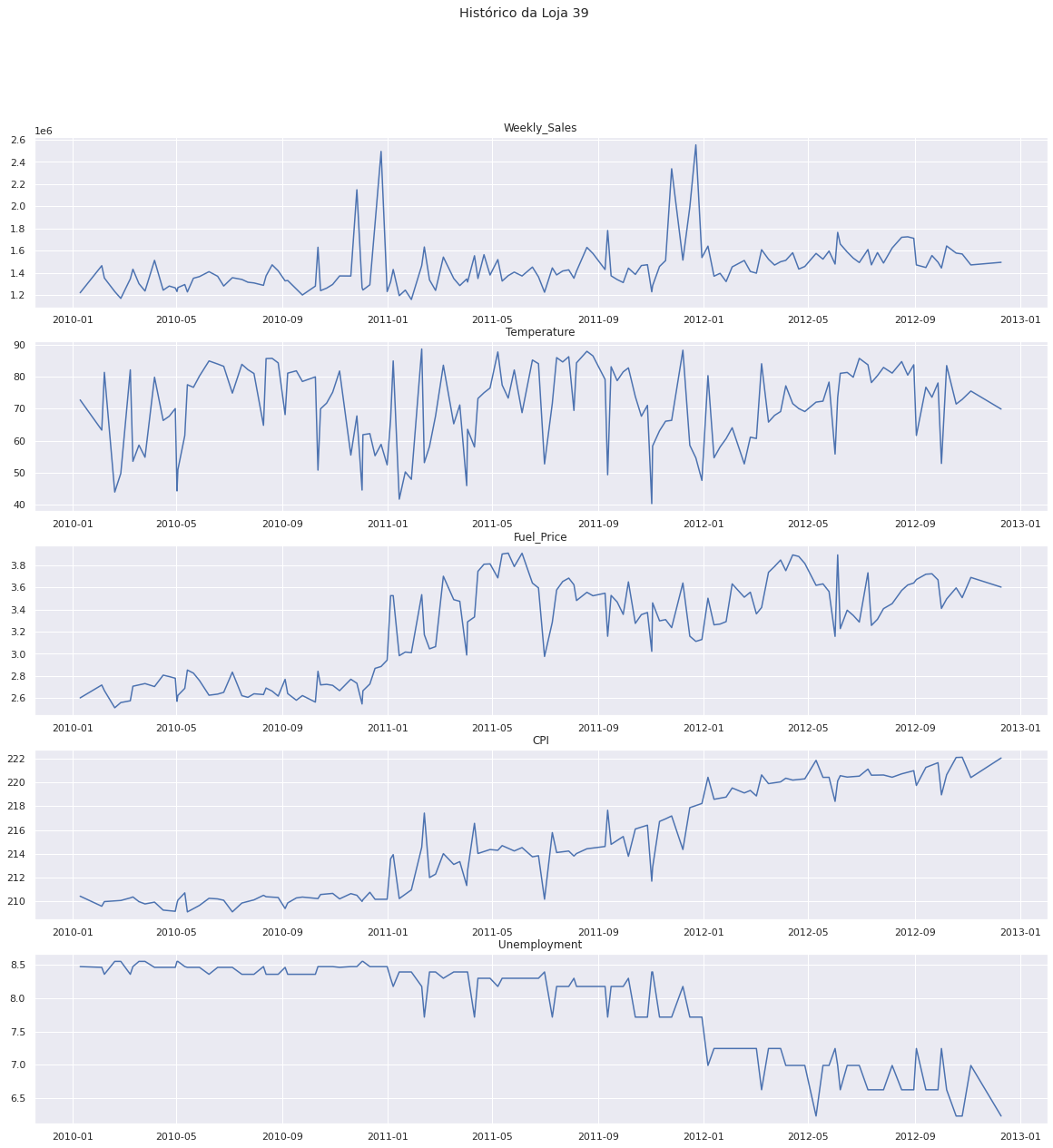
Preço do combustível: Em alta

IPC: Em alta

Taxa de Desemprego: Em baixa porém com fortes oscilações.

Essas informações são suficientes para decidir que a loja 14 não é uma boa opção.

**LOJA 39: Média de faturamento: US$ 1,416,005.59**



Faturamento: Atuando acima da média nos últimos meses.

Temperatura: Oscila muito, inconclusiva.

Preço do combustível: Em alta mas com oscilações

IPC: Em alta

Taxa de Desemprego: Em baixa

Olhando para análises anteriores:

A loja 39 responde bem aos feriados, tem uma boa correlação positiva com o IPC e uma boa correlação negativa com a taxa de desemprego.

**- Baseado nos dados, qual imóvel ou tipo de imóveis eu deveria comprar.**

Decidi que o melhor investimento é na **LOJA 39**. Ela está atuando acima do faturamento. A taxa de desemprego caindo, ela se mostra com uma boa correlação negativa, podendo faturar mais ainda com essa baixa do desemprego. O IPC na região está em um nível aceitável e as vendas respondem bem a este dado. Além disso, é uma loja que responde bem aos feriados. É uma boa opção para investimento.