**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**

**NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**

**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**Leandro da Silva Câmara**

**PREVISÃO DE VENDAS COM PROPHET**

Belo Horizonte

2020

**Leandro da Silva Câmara**

**PREVISÃO DE VENDAS COM PROPHET**

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Ciência de Dados e Big Data como requisito parcial à obtenção do título de especialista.

Belo Horizonte

2020

**LISTA DE ILUSTRAÇÕES**

[**Figura 1** - Carregamento dos *datasets*. 7](#_Toc40624871)

[**Figura 2** - Registros com valores negativos. 9](#_Toc40624872)

[**Figura 3** - Invertendo o "sinal" dos valores negativos. 9](#_Toc40624873)

[**Figura 4** - Convertendo a coluna "*date*" para o formato "*datetime*". 9](#_Toc40624874)

[**Figura 5** - Criação das colunas "*year*", "*month*" e "*total\_value*". 10](#_Toc40624875)

[**Figura 6** - Adicionando a coluna "*item\_category\_id*" ao *dataset* "*sales*". 10](#_Toc40624876)

[**Figura 7** - Dataset "*sales*" após a inclusão das novas colunas. 10](#_Toc40624877)

[**Figura 8** - Agrupamento mensal com as somatórias das vendas e da receita. 11](#_Toc40624878)

[**Figura 9** - Separação dos agrupamentos mensais do total de vendas e receita. 12](#_Toc40624879)

[**Figura 10** - Obtendo o total mensal de receita e de vendas (2013-2015). 12](#_Toc40624880)

[**Figura 11** - Função para retorno da diferença média entre os anos. 14](#_Toc40624881)

[**Figura 12** - Função para retorno de uma linha melhor ajuste (regressão linear). 15](#_Toc40624882)

[**Figura 13** - Obtendo o total de receita e de vendas de cada loja. 15](#_Toc40624883)

[**Figura 14** - Obtendo o total de receita e de vendas de cada categoria de produtos. 17](#_Toc40624884)

[**Figura 15** - Obtendo o total mensal de receita e de vendas. 18](#_Toc40624885)

[**Figura 16** - Obtendo as tendências e as sazonalidades das vendas e da receita. 20](#_Toc40624886)

[**Figura 17** - Adequando o nome das colunas para o uso do “*Prophet*”. 23](#_Toc40624887)

[**Figura 18** - Instanciando os modelos "*Prophet*". 23](#_Toc40624888)

[**Figura 19** - Realizando o treinamento dos modelos. 24](#_Toc40624889)

[**Figura 20** - Construindo o "*DataFrame*" com o histórico e os próximos 5 meses. 24](#_Toc40624890)

[**Figura 21** - Realizando a previsão dos 5 meses incluídos no histórico. 24](#_Toc40624891)

[**Figura 22** - Valor previsto para as vendas dos próximos 5 meses. 25](#_Toc40624892)

[**Figura 23** - Valor previsto para a receita dos próximos 5 meses. 26](#_Toc40624893)

**LISTA DE GRÁFICOS**

[**Gráfico 1** - Vendas por Mês (2013-2015). 13](#_Toc40624894)

[**Gráfico 2** - Receita por Mês (2013-2015). 14](#_Toc40624895)

[**Gráfico 3** - Vendas/Receita por Loja (2013-2015). 16](#_Toc40624896)

[**Gráfico 4** - Vendas/Receita por Categoria de Produto (2013-2015). 17](#_Toc40624897)

[**Gráfico 5** - Total de Vendas (Jan/2013 - Out/2015). 19](#_Toc40624898)

[**Gráfico 6** - Total de Receita (Jan/2013 - Out/2015). 20](#_Toc40624899)

[**Gráfico 7** - Total de Vendas (Tendência/Sazonalidade). 21](#_Toc40624900)

[**Gráfico 8** - Total de Receita (Tendência/Sazonalidade). 22](#_Toc40624901)

[**Gráfico 9** - Previsão de Vendas. 27](#_Toc40624902)

[**Gráfico 10** - Eficácia - Vendas (*Prophet*). 28](#_Toc40624903)

[**Gráfico 11** - Previsão de Receita. 29](#_Toc40624904)

[**Gráfico 12** - Eficácia - Receita (*Prophet*). 30](#_Toc40624905)

[**Gráfico 13** - Previsão de Vendas/Receita (Tendência/Sazonalidade). 31](#_Toc40624906)

**SUMÁRIO**

[1. Introdução 6](#_Toc40607761)

[1.1. Contextualização 6](#_Toc40607762)

[1.2. O problema proposto 6](#_Toc40607763)

[2. Coleta de Dados 7](#_Toc40607764)

[3. Processamento/Tratamento de Dados 8](#_Toc40607765)

[3.1. Pré-Processamento - Análise e Exploração de Dados 8](#_Toc40607766)

[3.2. Pré-Processamento - *Machine Learning* 11](#_Toc40607767)

[4. Análise e Exploração dos Dados 12](#_Toc40607768)

[5. Modelo de *Machine Learning* 22](#_Toc40607769)

[6. Apresentação dos Resultados 26](#_Toc40607770)

[7. Links 32](#_Toc40607771)

[REFERÊNCIAS 33](#_Toc40607772)

# 1. Introdução

## 1.1. Contextualização

Diversas informações são utilizadas pelas empresas para analisar seu desempenho, a quantidade de vendas e a receita gerada, por exemplo, podem ser consideradas na medição de sua performance.

Com a ausência de uma prévia análise dos dados, existe o risco de que as informações disponíveis sejam insuficientes para a compreensão acerca da situação da empresa, dificultando na tomada de decisões.

## 1.2. O problema proposto

Uma das maiores empresas russas de software – *1C Company*, disponibilizou seus dados por meio da plataforma *Kaggle*[[1]](#footnote-1), pois necessitava de informações que auxiliassem no entendimento da situação atual da empresa, acerca das vendas e da receita. Além disso, definir quais as projeções de futuro para contribuir em sua tomada de decisões.

Diante das dificuldades que empresas têm em analisar seus dados e ter informações relevantes à disposição de seus gestores, este trabalho tem como objetivo demonstrar como as análises de dados e a previsão de séries temporais podem auxiliar no entendimento do desempenho das vendas e da receita da empresa.

Ter à disposição projeções de futuro, por exemplo, das vendas e da receita, pode contribuir para melhores decisões, avaliação de potenciais ameaças, identificação de oportunidades, e proporcionar à empresa tempo hábil para planejamento e execução de ações, que possam trazer benefícios aos interesses da mesma.

O período analisado refere-se aos meses entre Janeiro de 2013 a Outubro de 2015, totalizando 34 meses. Por meio da exploração dos dados, propõe-se identificar informações que possam ser relevantes para a empresa, como por exemplo, a diferença do desempenho das vendas e da receita entre os anos de 2013 a 2015. Além disso, realizar análises para verificação da existência de tendências e sazonalidades.

Com essas informações, pretende-se elaborar um modelo, utilizando a biblioteca para construção de séries temporais do *Facebook* – denominada *Prophet*, que seja capaz de prever os próximos 5 meses de vendas e receita da empresa.

# 2. Coleta de Dados

Os dados referentes às vendas diárias da empresa *1C Company* foram obtidos no dia 14 de Janeiro de 2020, por meio de arquivos em formato CSV, disponibilizados na plataforma *Kaggle*. O conjunto de dados (*dataset*) abrange as vendas realizadas entre Janeiro de 2013 e Outubro 2015. O *dataset* é composto por **2.935.849 registros** de vendas e 3 arquivos, sendo estes, “*items.csv*”, “*sales.csv*” e “*item\_categories.csv*”. O carregamento dos *datasets* se deu através do uso da biblioteca “*Pandas*”, conforme apresentado na Figura 1.

**Figura 1** - Carregamento dos datasets.



Fonte: Autor da pesquisa.

* ***sales.csv*** - Dados históricos (vendas diárias) de Janeiro de 2013 a Outubro de 2015;

|  |  |  |
| --- | --- | --- |
| **Nome da coluna** | **Descrição** | **Tipo** |
| *date* | Data no formato “*dd.mm.YYYY*”. | *String* |
| *date\_block\_num* | Número sequencial dos meses entre Janeiro de 2013 (0) e Outubro de 2015 (33). | *Integer* |
| *shop\_id* | Identificador da loja. | *Integer* |
| *item\_id* | Identificador do produto. | *Integer* |
| *item\_price* | Preço do produto. | *Float* |
| *item\_cnt\_day* | Quantidade de produtos vendidos. | *Integer* |

* ***items.csv*** - Informações complementares sobre os itens/produtos;

|  |  |  |
| --- | --- | --- |
| **Nome da coluna** | **Descrição** | **Tipo** |
| *item\_id* | Identificador do produto. | *Integer* |
| *item\_name* | Nome do produto. | *String* |
| *item\_category\_id* | Identificador da categoria do produto. | *Integer* |

* ***item\_categories.csv*** - Informações complementares sobre as categorias de itens.

|  |  |  |
| --- | --- | --- |
| **Nome da coluna** | **Descrição** | **Tipo** |
| *item\_category\_id* | Identificador da categoria do produto. | *Integer* |
| *item\_category\_name* | Nome da categoria do produto. | *String* |

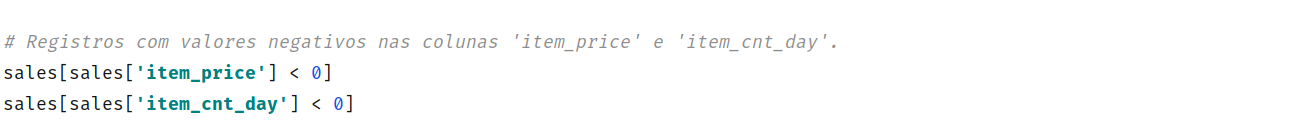
# 3. Processamento/Tratamento de Dados

Durante o processo de análise/exploração dos dados e construção do modelo de previsão de vendas, foi fundamental a realização de alguns processamentos/tratamentos de dados.

## 3.1. Pré-Processamento - Análise e Exploração de Dados

Um dado importante, mas ausente no *dataset*, é o valor total da venda (diária). Portanto, uma nova coluna foi criada - denominada “*total\_value*”, por meio da multiplicação das colunas “*item\_price*” e “*item\_cnt\_day*”. Entretanto, foram identificados **7357 registros** (Figura 2) com valores inválidos (negativos) em tais colunas.

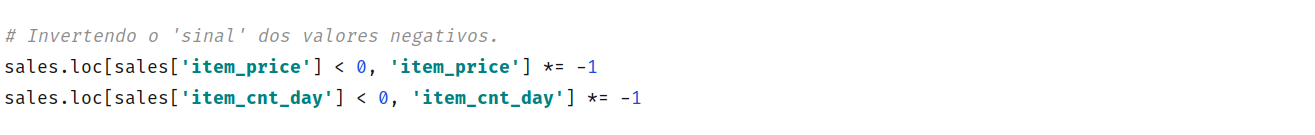
**Figura 2** - Registros com valores negativos.



Fonte: Autor da pesquisa.

Levando em consideração a premissa de que todas as vendas estão relacionadas com pelo menos um produto, todas as vendas com quantidades e preços negativos de produtos foram transformadas para valores positivos, como mostra a Figura 3.

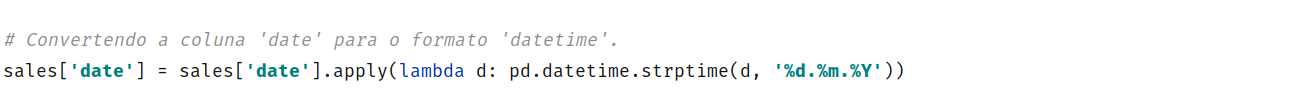
**Figura 3** - Invertendo o "sinal" dos valores negativos.



Fonte: Autor da pesquisa.

Devido ao fato da coluna “*date*” ser do tipo “*string*”, fez-se necessário a conversão (Figura 4) dos dados provenientes de tal coluna para o formato “*datetime*”, pois neste novo formato, torna-se possível análises com base em meses e anos.

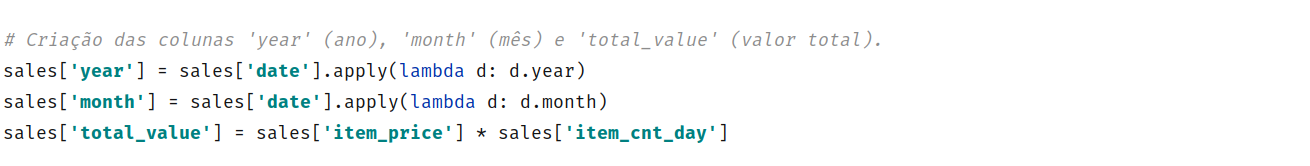
**Figura 4** - Convertendo a coluna "date" para o formato "datetime".



Fonte: Autor da pesquisa.

Após a conversão realizada na coluna “*date*”, conforme a Figura 5, foi possível adicionar à cada registro as colunas “*year*” (ano) e “*month*” (mês). Além disso, foi incluído a coluna “*total\_value*” (valor total).

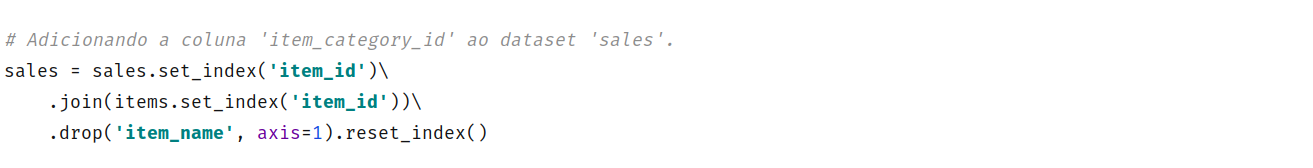
**Figura 5** - Criação das colunas "year", "month" e "total\_value".



Fonte: Autor da pesquisa.

Com o objetivo de realizar análises através do relacionamento das vendas com as categorias de produtos, foi incluído ao *dataset* “*sales*” (Figura 6) a coluna “*item\_category\_id*”, mediante à junção dos *datasets* “*sales*” e “*items*”, por meio da coluna “*item\_id*”.

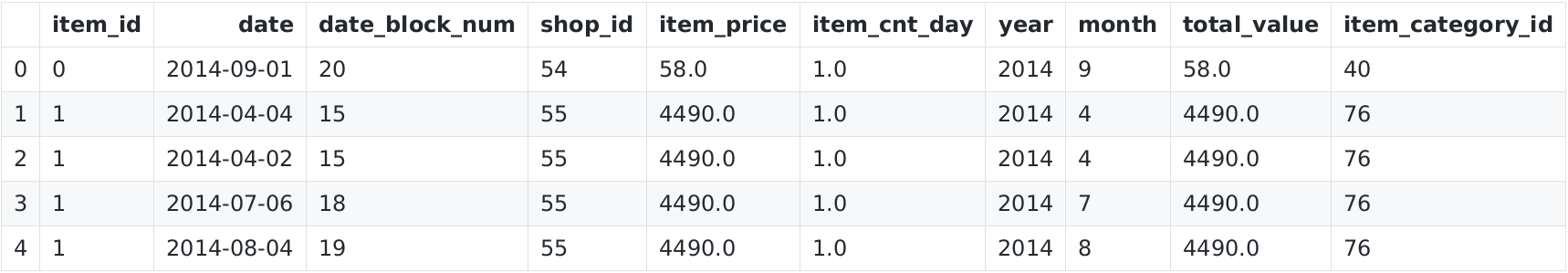
**Figura 6** - Adicionando a coluna "item\_category\_id" ao dataset "sales".



Fonte: Autor da pesquisa.

A Figura 7 apresenta o *dataset* após os tratamentos de dados, necessários para os procedimentos relacionados à exploração dos dados e construção de gráficos, possibilitando assim uma melhor compreensão das vendas e da receita.

**Figura 7** - Dataset "sales" após a inclusão das novas colunas.



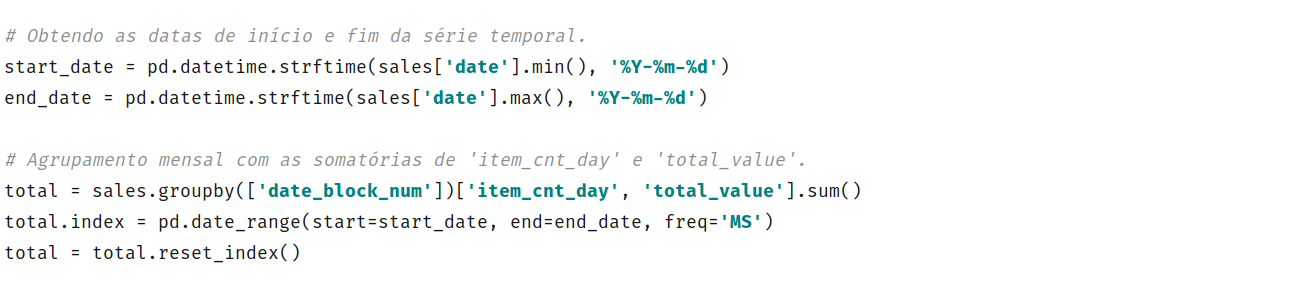
Fonte: Autor da pesquisa.

## 3.2. Pré-Processamento - *Machine Learning*



Para criação do modelo de *Machine Learning*, foi necessário um pré-processamento no *dataset* “*sales*”, para que as vendas fossem agrupadas mensalmente (Figura 8), de modo a obter o somatório das vendas e da receita.

**Figura 8** - Agrupamento mensal com as somatórias das vendas e da receita.



Fonte: Autor da pesquisa.

Após o agrupamento mensal, foi necessário dividir resultado, de modo que o total das vendas ficasse separada da receita, como mostra a Figura 9. Esta partição é fundamental, pois será criado um modelo (*Machine Learning*) para cada *dataset* gerado.

**Figura 9** - Separação dos agrupamentos mensais do total de vendas e receita.

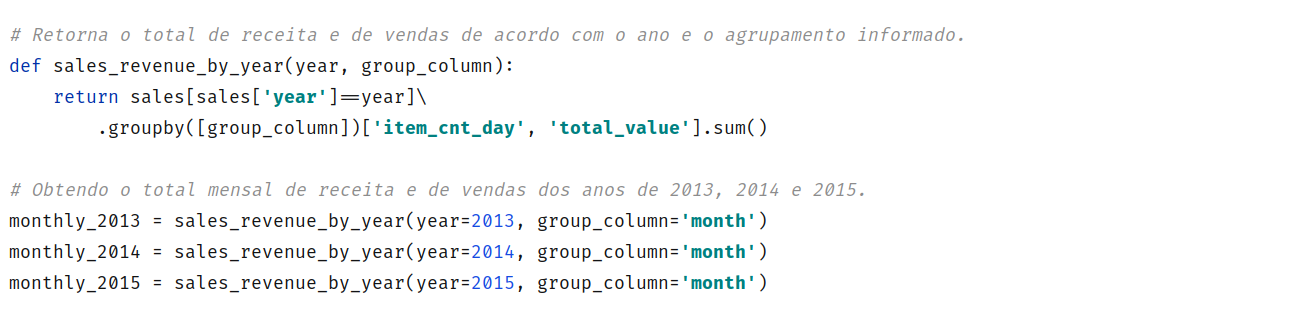


Fonte: Autor da pesquisa.

# 4. Análise e Exploração dos Dados

Após a realização do Processamento/Tratamento dos Dados, análises foram elaboradas, com objetivo de obter *insights* e identificar padrões no *dataset* “*sales*”.

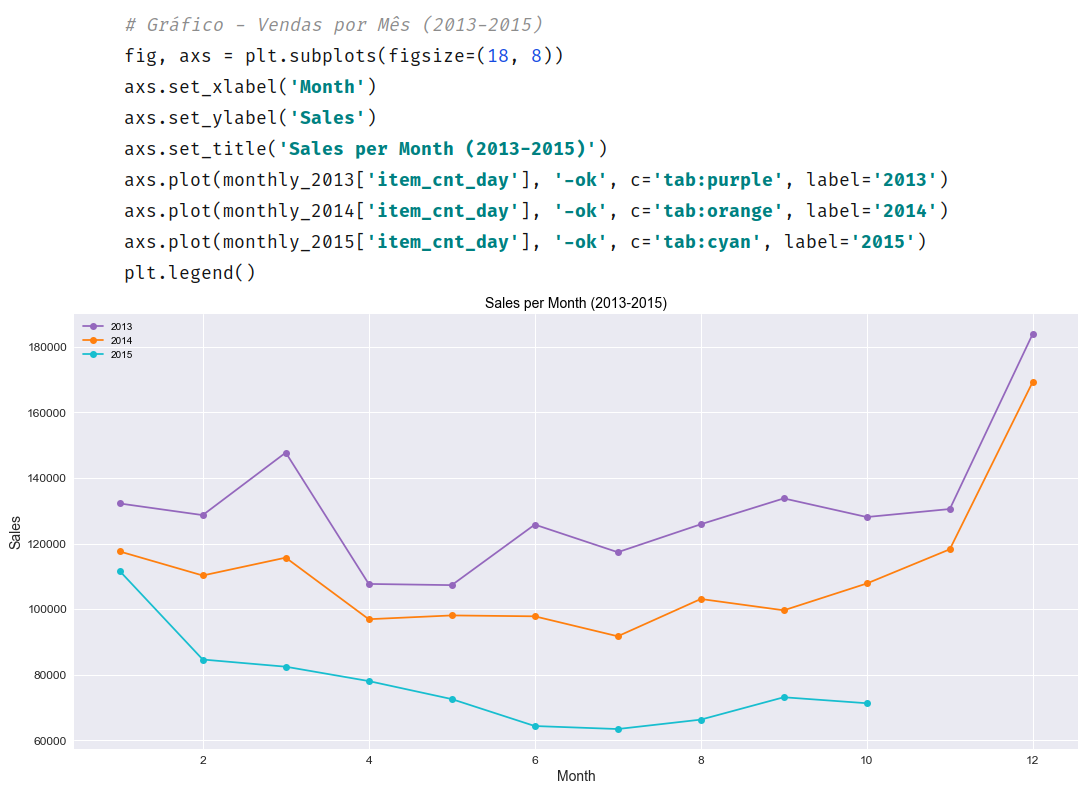
**Figura 10** - Obtendo o total mensal de receita e de vendas (2013-2015).



Fonte: Autor da pesquisa.

Para obter uma visão do desempenho das **vendas** e da **receita** nos anos de 2013, 2014 e 2015, foram construídos gráficos de linhas, a fim de permitir a comparação entre os anos. Portanto, fez-se necessário, o somatório mensal das vendas e da receita, de cada ano, conforme mostra a Figura 10).

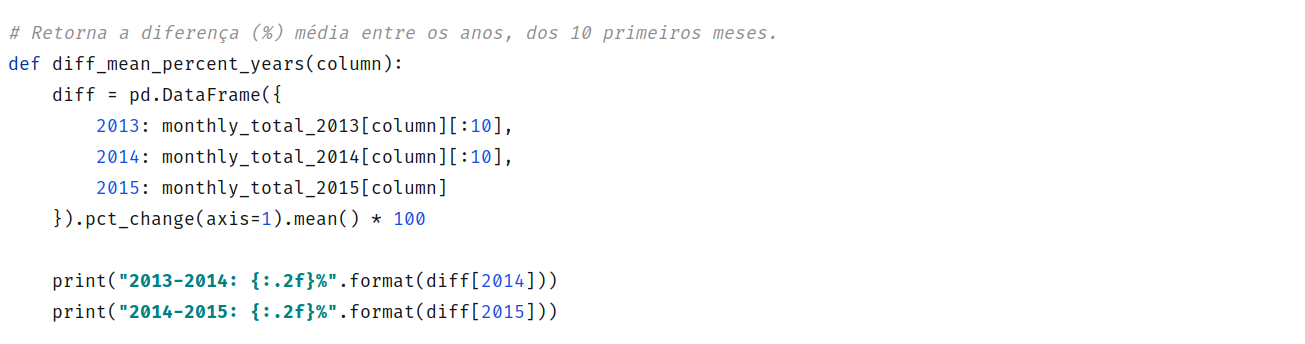
**Gráfico 1** - Vendas por Mês (2013-2015).



Fonte: Autor da pesquisa.

Conforme apresentado no Gráfico 1, com o decorrer dos anos, as vendas diminuíram significativamente. Para obter o percentual referente à variação das vendas e da receita, foi calculado a diferença média entre os anos, dos 10 primeiros meses (Janeiro a Outubro). No cálculo realizado, não foi levado em consideração os meses de Novembro e Dezembro, pois no ano de 2015, os dados referente a tais meses não estavam presentes no *dataset*.

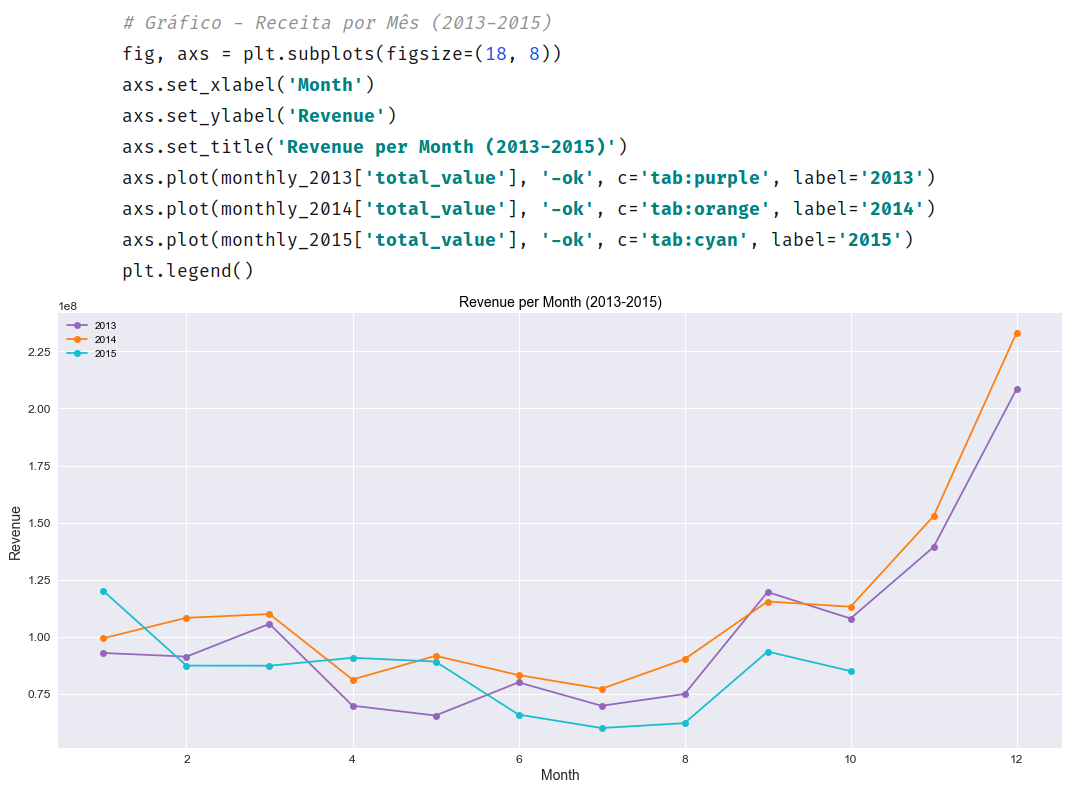
**Figura 11** - Função para retorno da diferença média entre os anos.



Fonte: Autor da pesquisa.

Após a execução do cálculo, por meio da função apresentada na Figura 11, identificou-se que houve uma queda média de 16,92% nas vendas entre os anos de 2013 e 2014. Entre os anos 2014 e 2015, a queda nas vendas foi maior que o período anterior, alcançando em média 26,37%.

**Gráfico 2** - Receita por Mês (2013-2015).



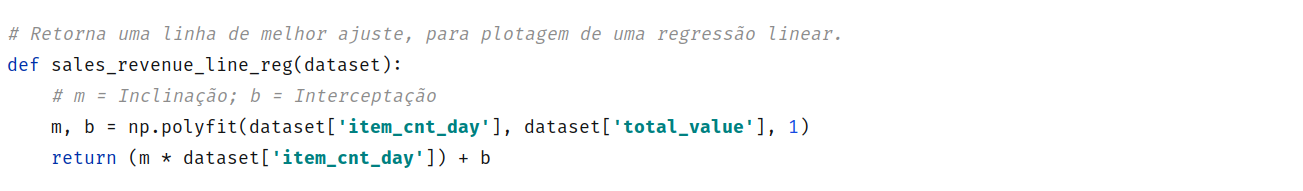
Fonte: Autor da pesquisa.

Com o Gráfico 2, percebe-se que a receita permaneceu quase estável durante os meses de Janeiro e Outubro. Após realizar o cálculo da diferença média, verificou-se que, houve um aumento médio de 12,24% (Figura 11) entre os anos de 2013 e 2014. Entretanto, entre os anos de 2014 e 2015, um decréscimo médio de 12,82%.

Após a análise das vendas e da receita da empresa como um todo, foram elaborados gráficos de dispersão para verificar a distribuição anual das vendas e da receita, relacionada às lojas e às categorias de produtos.

Além disso, foi incluído aos gráficos, para cada ano, uma linha para representação da distribuição dos dados, por meio de Regressão Linear, a fim de facilitar a interpretação dos resultados.

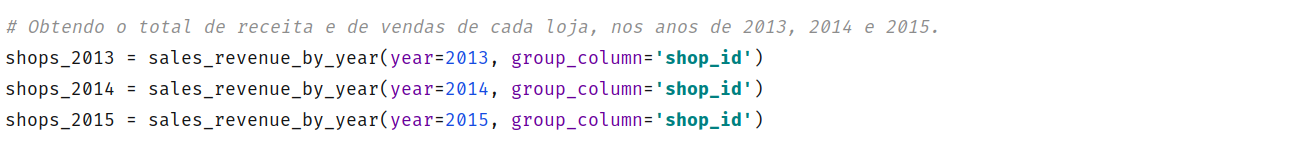
**Figura 12** - Função para retorno de uma linha melhor ajuste (regressão linear).



Fonte: Autor da pesquisa.

Após a definição da função responsável por retornar a Regressão Linear (Figura 12), obteve-se o total de vendas e receita, relacionada à cada loja, de cada ano, como mostra a Figura 13.

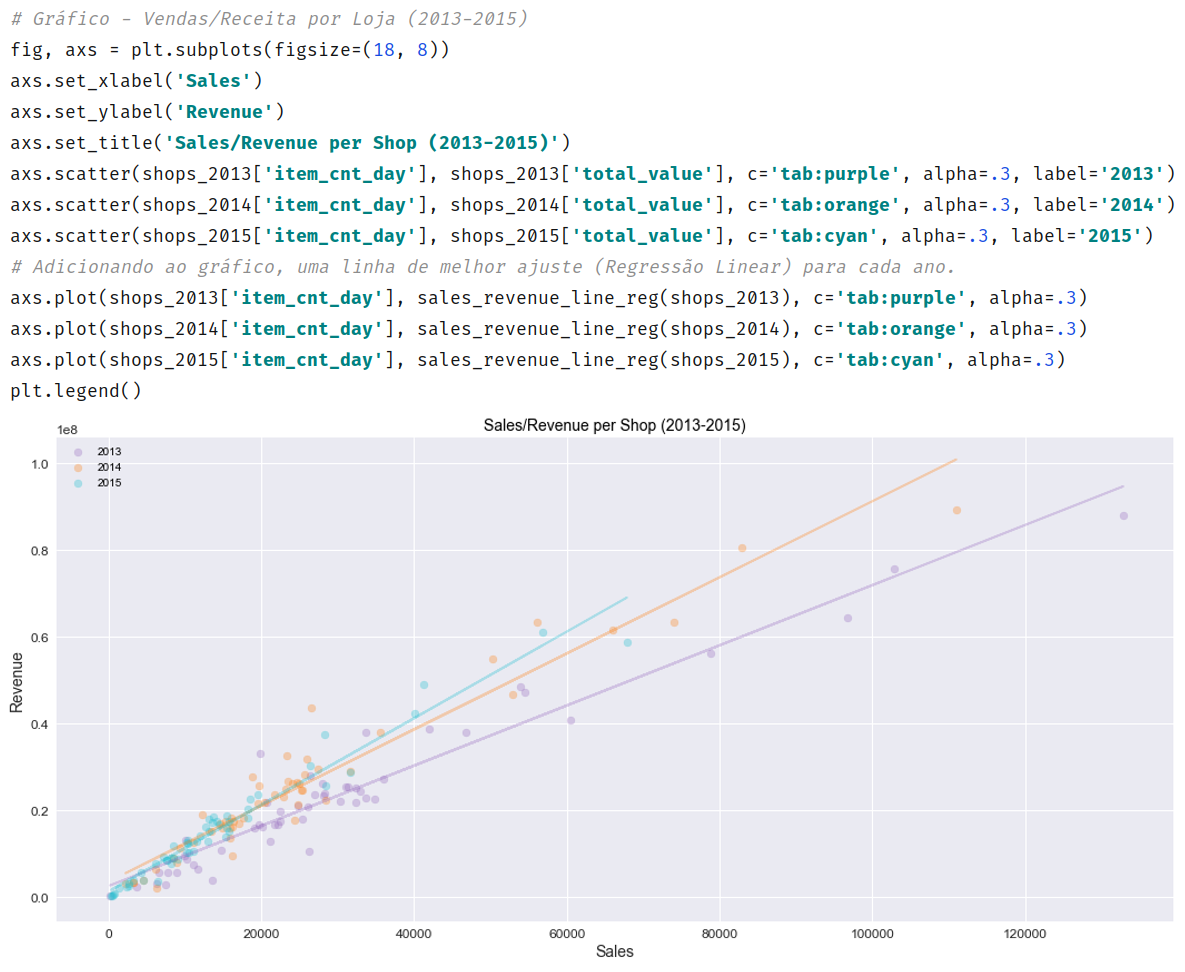
**Figura 13** - Obtendo o total de receita e de vendas de cada loja.



Fonte: Autor da pesquisa.

Por meio do Gráfico 3, é possível observar que, com o passar dos anos, existe uma tendência de queda nas vendas das lojas. Entretanto, há a tendência contrária (de aumento) com relação à receita gerada pelas lojas.

**Gráfico 3** - Vendas/Receita por Loja (2013-2015).

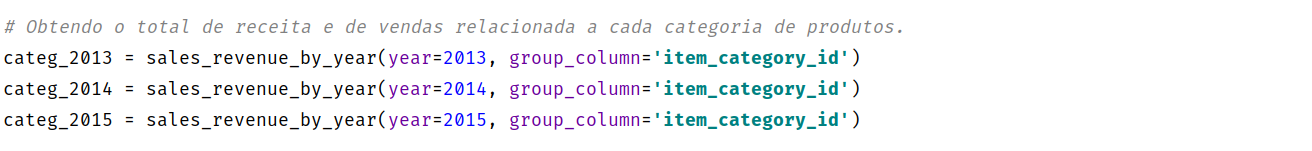


Fonte: Autor da pesquisa.

Além disso, observa-se que durante os anos, existem dois grupos, que possuem características similares e representam a maioria das lojas da empresa. O primeiro grupo, realizou até 20.000 vendas, com geração de receita de até 20.000.000. Enquanto o segundo grupo, realizou entre 20.000 e 40.000 vendas e com geração de receita entre 20.000.000 e 40.000.000.

Após as análises relacionadas ao desempenho anual das lojas, obteve-se o total anual de vendas e receita de cada categoria de produtos (Figura 14), assim como, a representação da distribuição dos dados, por meio de Regressão Linear.

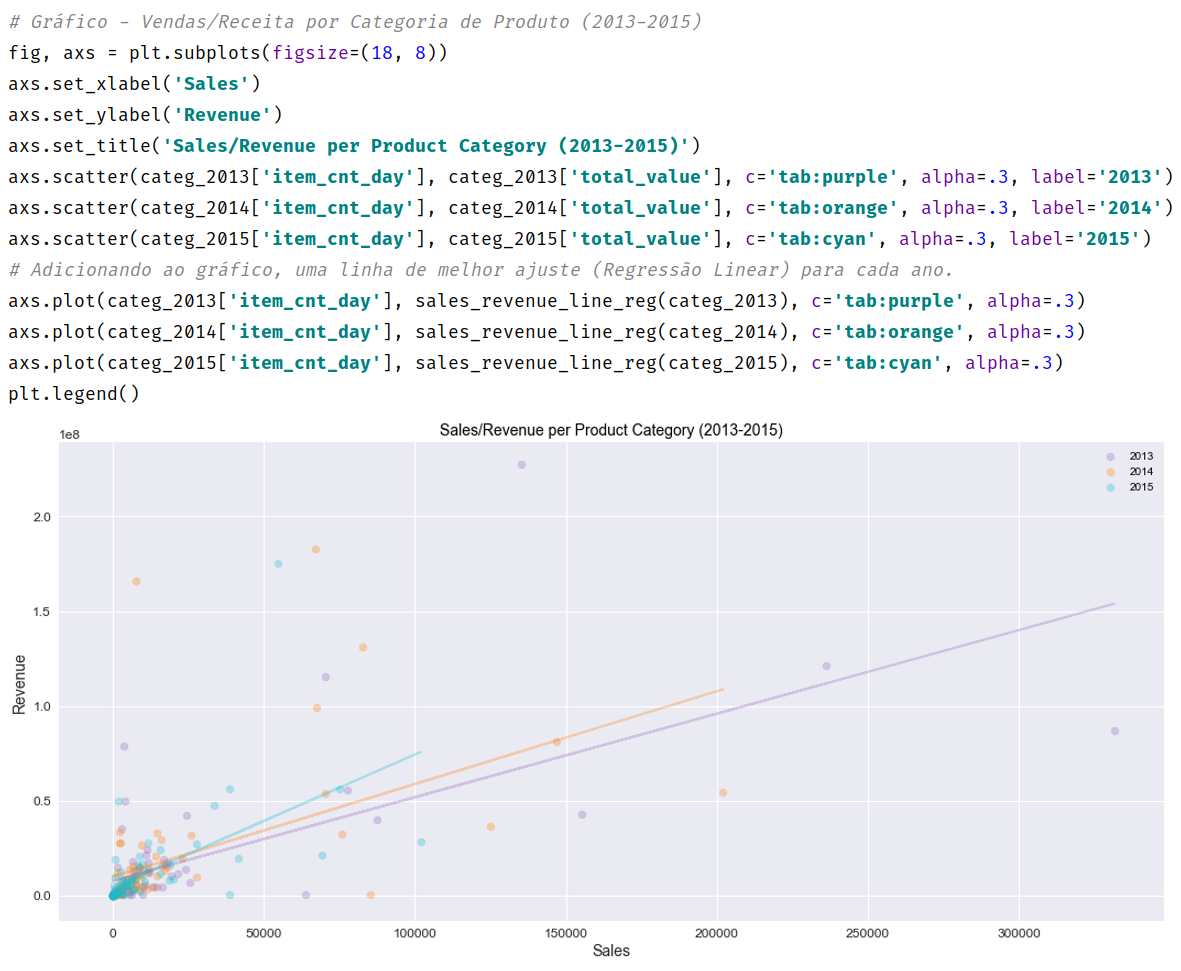
**Figura 14** - Obtendo o total de receita e de vendas de cada categoria de produtos.



Fonte: Autor da pesquisa.

Por meio do Gráfico 4, é possível observar que, com o passar dos anos, existe uma tendência de crescimento na receita, quando associada às categorias de produtos. Entretanto, há a tendência contrária (de queda) com relação às vendas, quando associadas às categorias de produtos.

**Gráfico 4** - Vendas/Receita por Categoria de Produto (2013-2015).



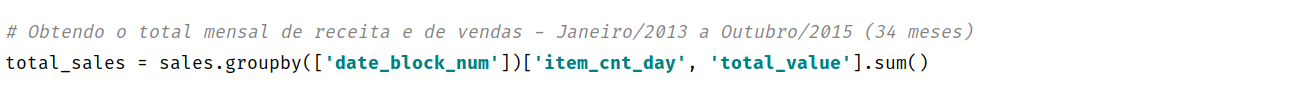
Fonte: Autor da pesquisa.

Além disso, observa-se que, nos anos de 2013 e 2014, houveram categorias de produtos que representaram aumentos significativos nas vendas e na geração de receita. Por exemplo, em 2013, algumas dessas categorias estiveram relacionadas a mais de 300.000 das vendas realizadas, e outras, relacionadas à geração de receita de mais de 200.000.000.

Devido ao fato de que, os registros de vendas de 2015 estavam limitados aos meses de Janeiro a Outubro, é provável que as tendências identificadas nos gráficos de dispersão foram influenciadas pela ausência de registros de vendas referentes aos meses de Novembro e Dezembro, principalmente quando se comparado aos anos anteriores.

Após a análise dos dados e dos gráficos relacionados às vendas e à receita de cada ano, foi examinado o desempenho geral da vendas e da receita, de modo a observar as **tendências** e as **sazonalidades**. Para tal, foi necessário obter o somatório total mensal das vendas e da receita, conforme apresentado na Figura 15.

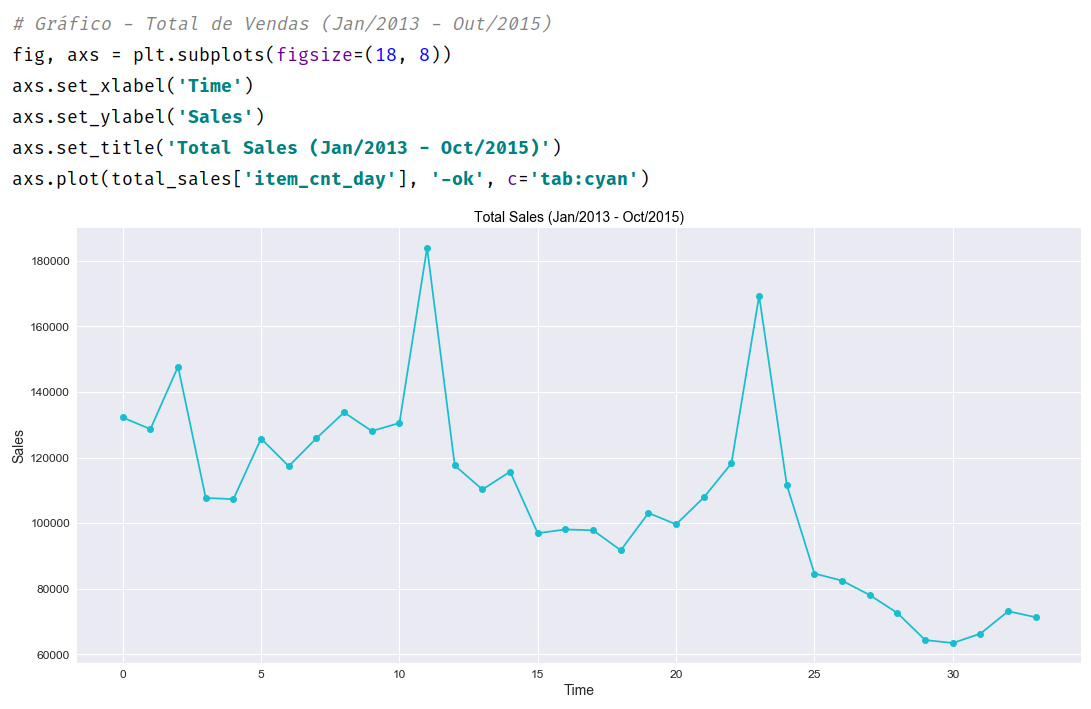
**Figura 15** - Obtendo o total mensal de receita e de vendas.



Fonte: Autor da pesquisa.

Com o resultado do agrupamento, tornou-se possível a construção de dois gráficos de linhas. O primeiro, relacionado às vendas, enquanto o segundo, relacionado à receita gerada. Ambos os gráficos representam o desempenho da empresa no decorrer dos três anos, entre Janeiro de 2013 a Outubro de 2015, totalizando 34 meses.

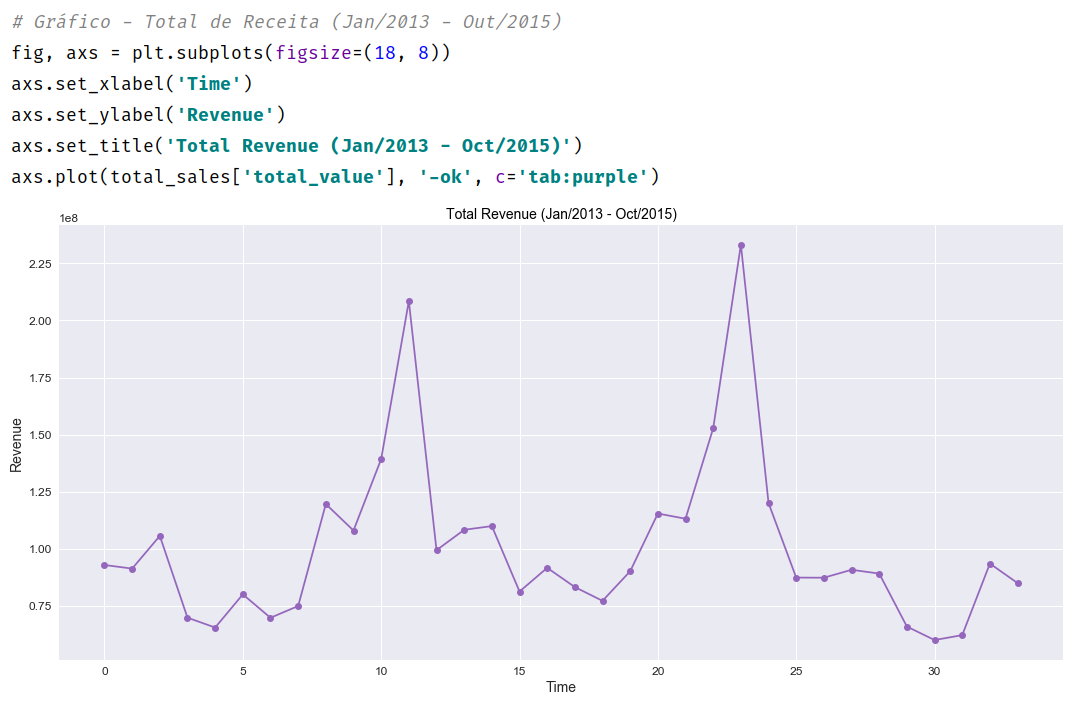
**Gráfico 5** - Total de Vendas (Jan/2013 - Out/2015).



Fonte: Autor da pesquisa.

Conforme apresentado no Gráfico 5, é possível identificar a queda nas vendas com o passar do tempo, assim como indicado nos gráficos anteriores. Vale ressaltar que, existem picos elevados nas vendas, ocorridos nos meses de Dezembro de 2013 e 2014.

**Gráfico 6** - Total de Receita (Jan/2013 - Out/2015).

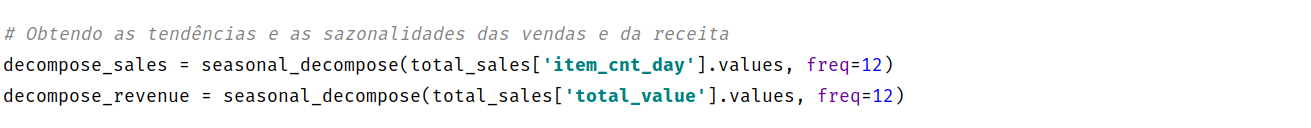


Fonte: Autor da pesquisa.

No Gráfico 6, a receita demonstra estar mais estável, se comparado com as vendas, assim como mencionado anteriormente, visto que as diferenças médias entre os anos de 2013, 2014 e 2015, praticamente se anulam.

Em seguida, a partir da função “*seasonal\_decompose*”, disponível através da biblioteca “*statsmodel*”, obteve-se informações sobre a tendência, a sazonalidade e os resíduos, referentes aos dados do total de vendas e receitas, como mostra a Figura 16.

**Figura 16** - Obtendo as tendências e as sazonalidades das vendas e da receita.



Fonte: Autor da pesquisa.

No Gráfico 7, é possível observar que, tornou-se mais evidente, a existência de uma tendência de queda nas vendas da empresa. Além disso, percebe-se que há uma sazonalidade anual, ou seja, que os fenômenos (vendas) se repetem, de forma idêntica, a cada 12 meses.

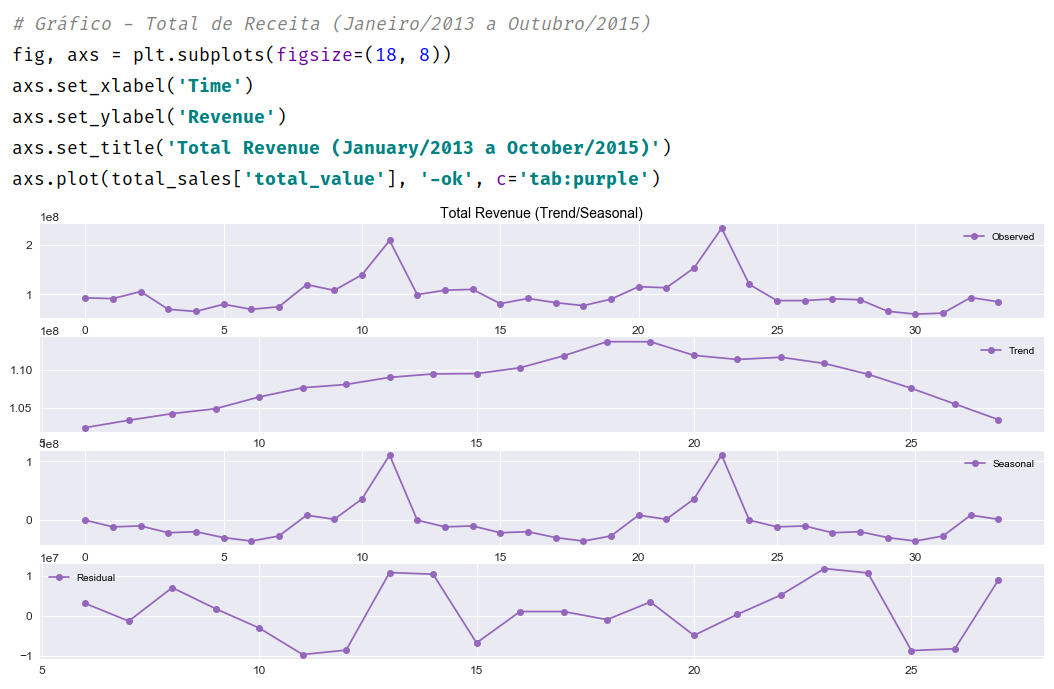
**Gráfico 7** - Total de Vendas (Tendência/Sazonalidade).



Fonte: Autor da pesquisa.

Apesar de que, no Gráfico 8 é perceptível a existência de uma sazonalidade (anual), há uma maior dificuldade quanto à interpretação da tendência, referente à geração de receita da empresa.

**Gráfico 8** - Total de Receita (Tendência/Sazonalidade).



Fonte: Autor da pesquisa.

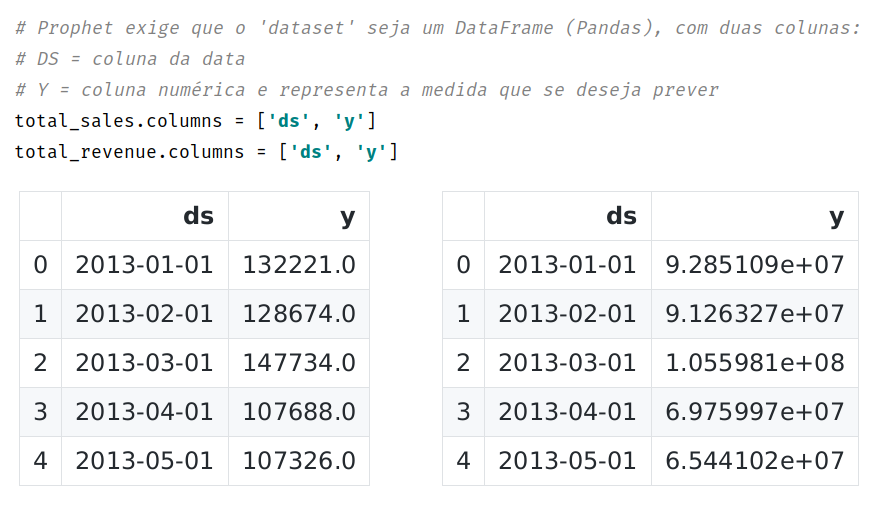
# 5. Modelo de *Machine Learning*

Para a criação do modelo de *Machine Learning* e realização da previsão de vendas, foi utilizado a biblioteca *Prophet*, criada pela equipe *Core Data Science* (CDS), do *Facebook* (FACEBOOK, 2020).

*Prophet* é um procedimento para previsão de séries temporais com base em um modelo aditivo, no qual tendências não lineares se ajustam à sazonalidade anual, semanal e diária, além de efeitos de férias. Além disso, *Prophet* é uma biblioteca de código aberto, implementado nas linguagens *R* e *Python* (PROPHET, 2020).

Após o pré-processamento dos dados necessários para criação do modelo (com *Prophet*), apresentado no capítulo [3.2](#_3xt2g1wb77xd), era preciso que o nome das colunas dos datasets “*total\_sales*” e “*total\_revenue*” fossem renomeadas, pois o *Prophet* exige que o *DataFrame* tenha obrigatoriamente duas colunas, “*DS*”, que representa a coluna da data, e “*Y*”, representando a coluna numérica, a qual deseja-se prever.

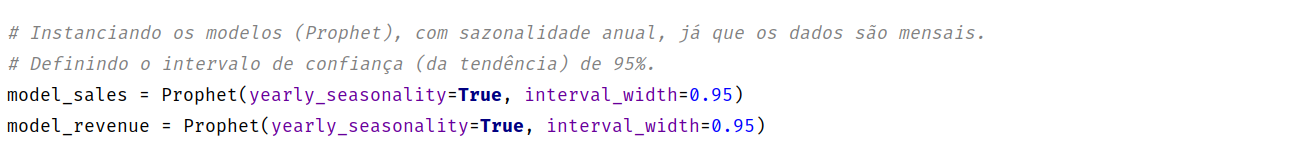
**Figura 17** - Adequando o nome das colunas para o uso do “Prophet”.



Fonte: Autor da pesquisa.

Para criação dos modelos de previsão com o *Prophet*, considerou-se que a sazonalidade é anual, já que os dados são mensais, e para a tendência, foi definido um intervalo de confiança de 95%, pois o intervalo padrão do *Prophet* é 80%.

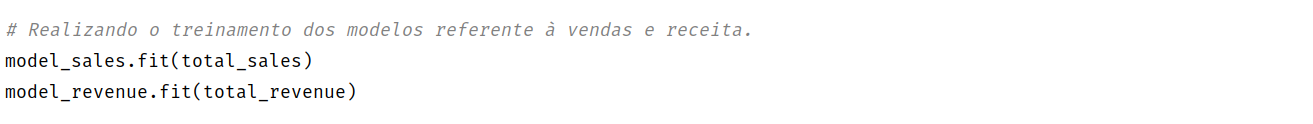
**Figura 18** - Instanciando os modelos "Prophet".



Fonte: Autor da pesquisa.

Uma vez que os *datasets* atende ao formato exigido e os modelos já foram instanciados, é realizado a etapa de treinamento dos modelos com os *datasets* informados.

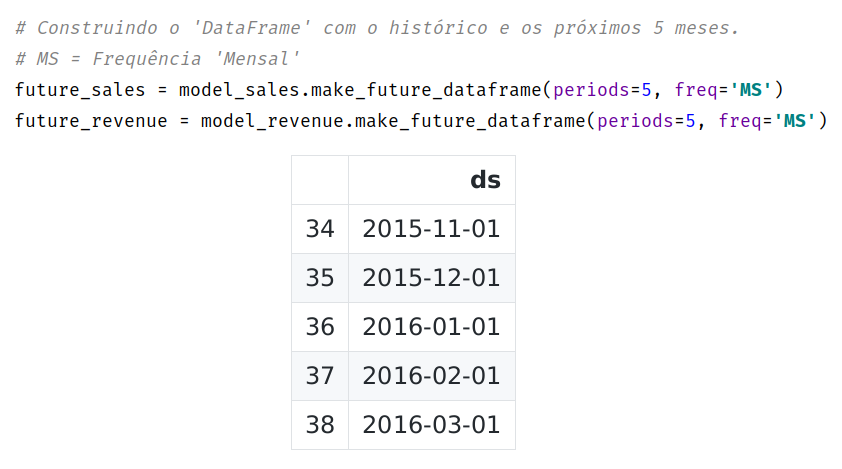
**Figura 19** - Realizando o treinamento dos modelos.



Fonte: Autor da pesquisa.

Após o treinamento, o *Prophet* permite construir um novo *DataFrame*, em que inclui o histórico e as datas dos períodos futuros. Neste caso, a frequência informada é “Mensal”, podendo assim, incluir os próximos 5 meses, a quais se referem à Novembro de 2015 a Março de 2016.

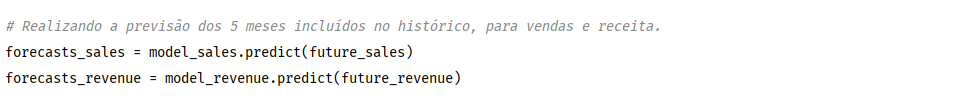
**Figura 20** - Construindo o "DataFrame" com o histórico e os próximos 5 meses.



Fonte: Autor da pesquisa.

Por meio da função “*predict*”, será incluída para cada registro do *dataset* (*future*) informado, uma nova coluna - denominada “*yhat*”, com o valor previsto.

**Figura 21** - Realizando a previsão dos 5 meses incluídos no histórico.



Fonte: Autor da pesquisa.

Juntamente com o valor previsto (*yhat*), o *Prophet* inclui os intervalos de confiança, “*yhat\_lower*” e “*yhat\_upper*”, que foram calculados com base no intervalo (95%) definido previamente, no momento de instanciação dos modelos. Na figura abaixo, é apresentado as previsões referentes às vendas da empresa.

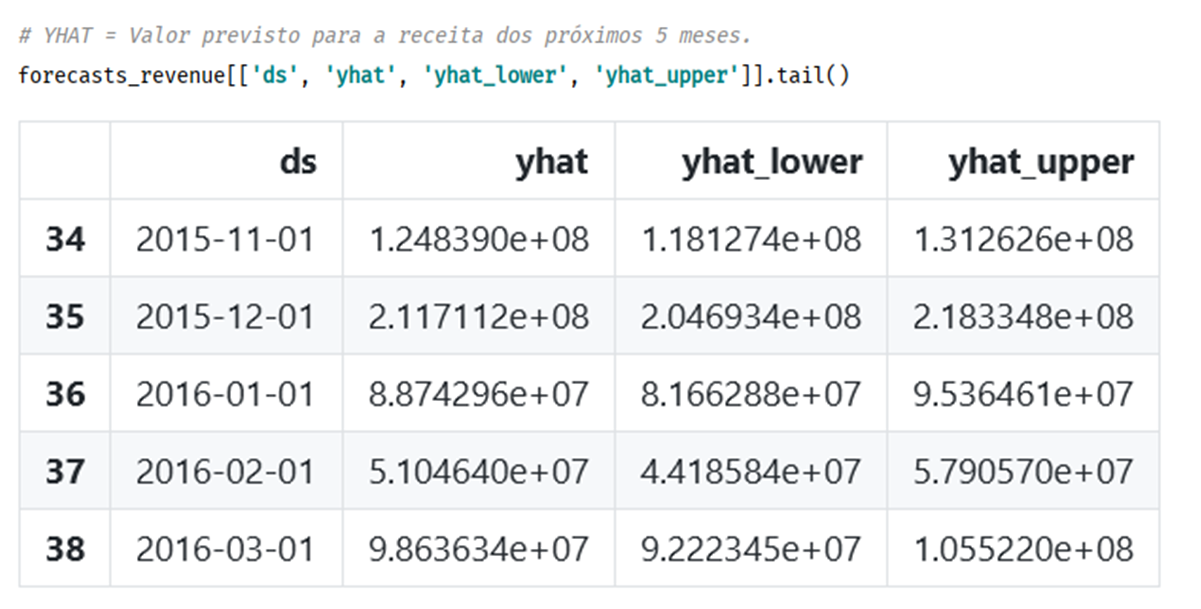
**Figura 22** - Valor previsto para as vendas dos próximos 5 meses.



Fonte: Autor da pesquisa.

A função “*tail*” retorna os últimos registros do *dataset*, por padrão, os 5 últimos, que coincide com a quantidade de meses previstos. Entretanto, é possível definir a quantidade a ser retornada, por meio dos parâmetros da função.

**Figura 23** - Valor previsto para a receita dos próximos 5 meses.



Fonte: Autor da pesquisa.

# 6. Apresentação dos Resultados

Após obter as previsões das vendas para os próximos 5 meses, foram plotados os dados históricos (já existentes), juntamente com os valores previstos e intervalos de confiança, calculados nos modelos criados com o *Prophet*, conforme apresentado no Gráfico 9.

**Gráfico 9** - Previsão de Vendas.



Fonte: Autor da pesquisa.

Além disso, para uma melhor compreensão da eficácia do modelo sobre os dados históricos, o Gráfico 10 mostra os valores previstos para todos os meses da série. Percebe-se que as previsões do modelo, estão próximas com os dados históricos de vendas, aumentando a confiabilidade sobre os valores previstos para os próximos meses.

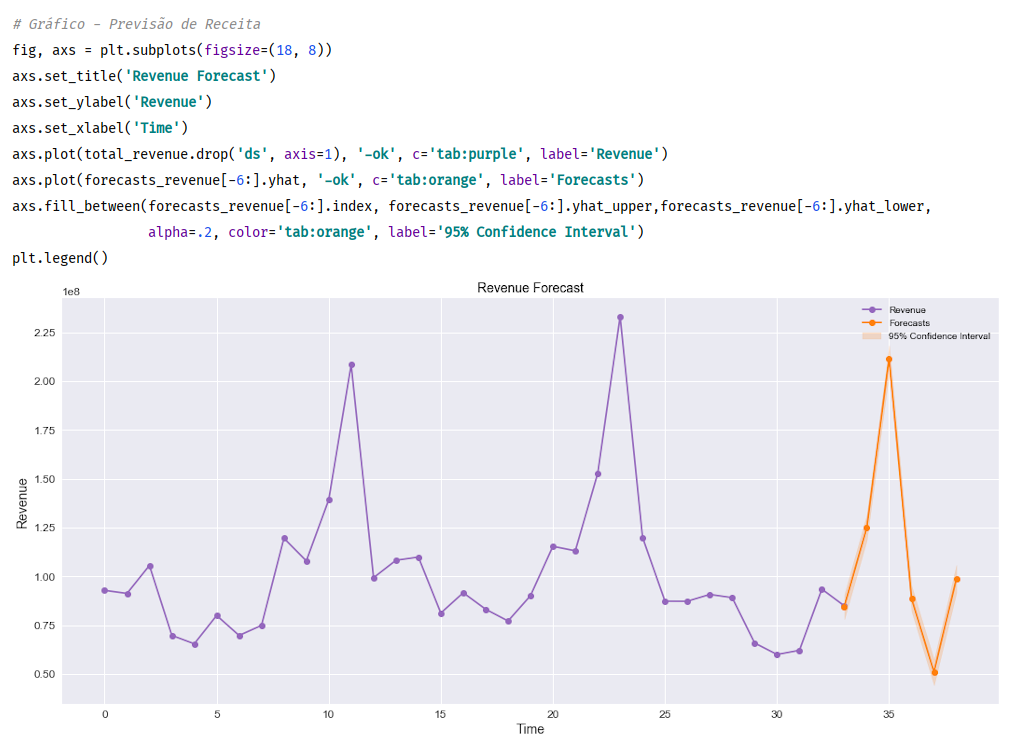
**Gráfico 10** - Eficácia - Vendas (Prophet).



Fonte: Autor da pesquisa.

Do mesmo modo, o Gráfico 11 foi criado, para representar as previsões obtidas pelo modelo (*Prophet*) dos meses futuros, em conjunto com as receitas anteriores da empresa.

**Gráfico 11** - Previsão de Receita.



Fonte: Autor da pesquisa.

No que diz respeito à eficácia do modelo referente à receita, observa-se, por meio do Gráfico 12, que há uma similaridade entre dados históricos da receita e os valores previstos pelo modelo.

**Gráfico 12** - Eficácia - Receita (Prophet).



Fonte: Autor da pesquisa.

Com os valores previstos, para vendas e receita, referente aos meses de Novembro de 2015 a Março de 2016, foi possível observar, através do Gráfico 13, as tendências e sazonalidades, atualizadas com as previsões obtidas com os modelos do *Prophet*, permitindo assim uma análise em relação à um provável futuro para empresa.

**Gráfico 13** - Previsão de Vendas/Receita (Tendência/Sazonalidade).



Fonte: Autor da pesquisa.

Tendo em vista os aspectos observados, conclui-se que, mesmo que sejam considerados as previsões dos meses seguintes, é incontestável a existência de tendência de queda nas vendas e na receita. Embora haja uma previsibilidade acerca do desempenho futuro da empresa, as análises e previsões demonstradas neste trabalho, pode auxiliar em tomadas de decisões e ações preventivas, por parte dos gestores, de modo que sejam minimizados os impactos negativos das vendas e da receita sobre a empresa.

# 7. Links

1. Segue abaixo o *link* para acesso ao **repositório** do projeto, contendo os arquivos utilizados no desenvolvimento do projeto, e o *link* para acesso ao **vídeo** de apresentação:

* Repositório dos Dados: <https://www.kaggle.com/c/competitive-data-science-predict-future-sales/data>
* Repositório GitHub: <https://github.com/leandrocamara/future-sales-forecast>
* Vídeo de Apresentação:

# REFERÊNCIAS

FACEBOOK. ***Core Data Science (CDS)***. Disponível em: <https://research.fb.com/teams/core-data-science/>. Acessado em: jan. 2020.

KAGGLE. ***Kaggle Inc***. Disponível em: <https://www.kaggle.com/>. Acessado em: jan. 2020.

PROPHET. ***Forecasting at scale***. Disponível em: <https://facebook.github.io/prophet/>. Acessado em: jan. 2020.

1. *Kaggle* refere-se à uma subsidiária da *Google LLC*, é uma comunidade *on-line* de cientistas de dados e profissionais de aprendizado de máquina (KAGGLE, 2020). [↑](#footnote-ref-1)