



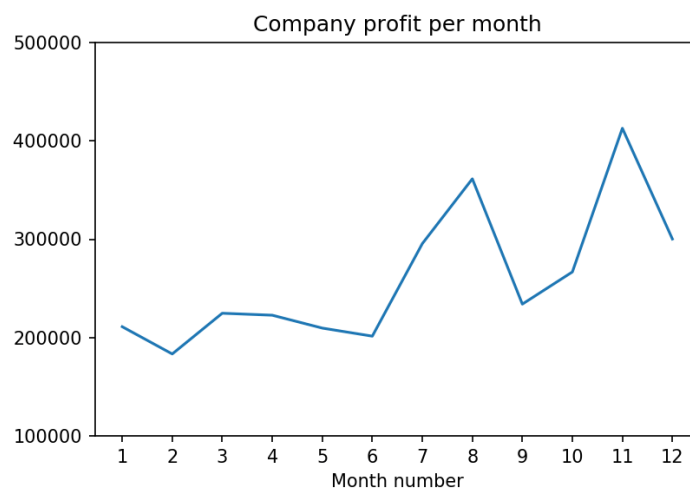
## LISTA DE EXERCÍCIO 07 – Matplotlib e Seaborn

**Exercício 01** - Leia o lucro total de todos os meses e mostre-o usando um gráfico de linhas.

**Arquivo:** *company\_sales\_data.csv*

Dados de lucro total fornecidos para cada mês. O gráfico de linha gerado deve incluir as seguintes propriedades:

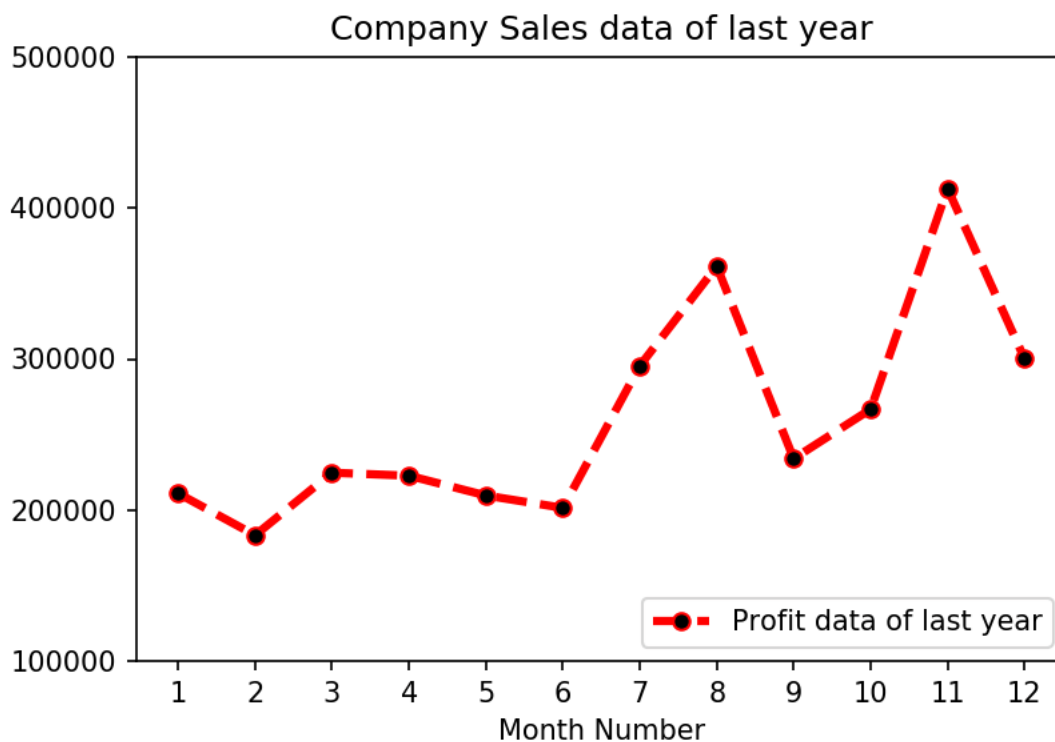
- **Nome do rótulo X** = Número do mês
- **Nome do rótulo Y** = lucro total



**Exercício 02** - Obtenha o lucro total de todos os meses e mostre o gráfico de linhas com as seguintes propriedades de estilo.

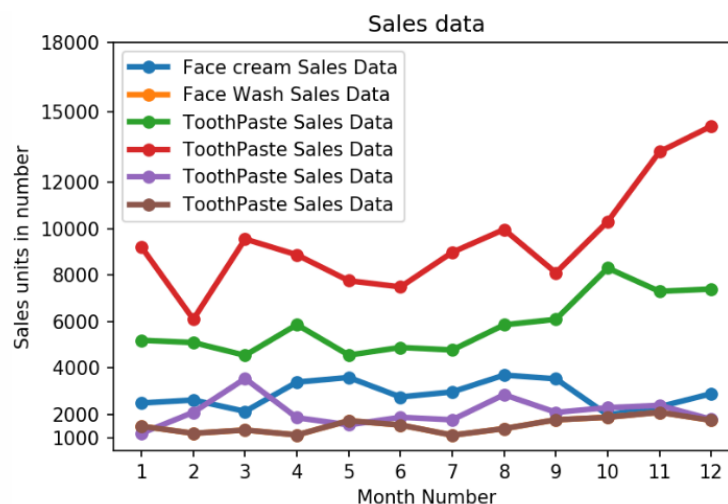
O gráfico de linha gerado deve incluir as seguintes propriedades de estilo:

- O estilo da linha pontilhada e a cor da linha devem ser vermelhas
- Mostrar legenda no local inferior direito.
- Nome do rótulo X = Número do mês
- Nome do rótulo Y = número de unidades vendidas
- Adicione um marcador de círculo.
- Cor do marcador de linha conforme lida
- A largura da linha deve ser 3



**Exercício 03** – Leia todos os dados de vendas de produtos e mostre-os usando um gráfico multilinha.

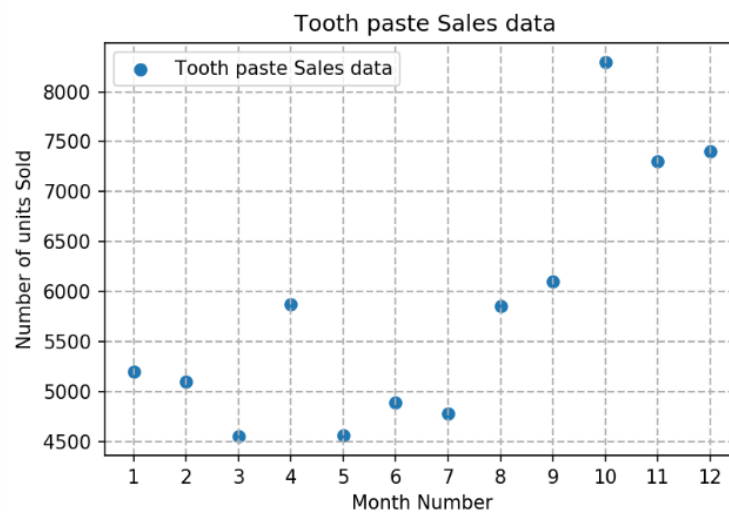
Exiba o número de unidades vendidas por mês para cada produto usando gráficos multilinhas. (ou seja, faça a plotagem das linhas separadas para cada produto).



**Exercício 04** - Leia os dados de vendas de pasta de dente de cada mês e mostre-os usando um gráfico de dispersão.

Além disso, adicione uma grade no gráfico. o estilo da linha de grade deve “-”.

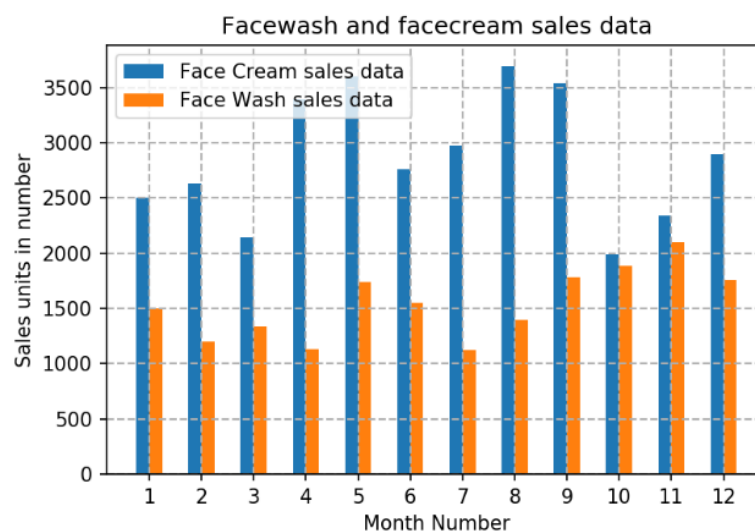
O gráfico de dispersão deve ficar assim.



**Exercício 05:** Leia os dados de vendas de cremes faciais e produtos de lavagem facial e mostre-os usando o gráfico de barras.

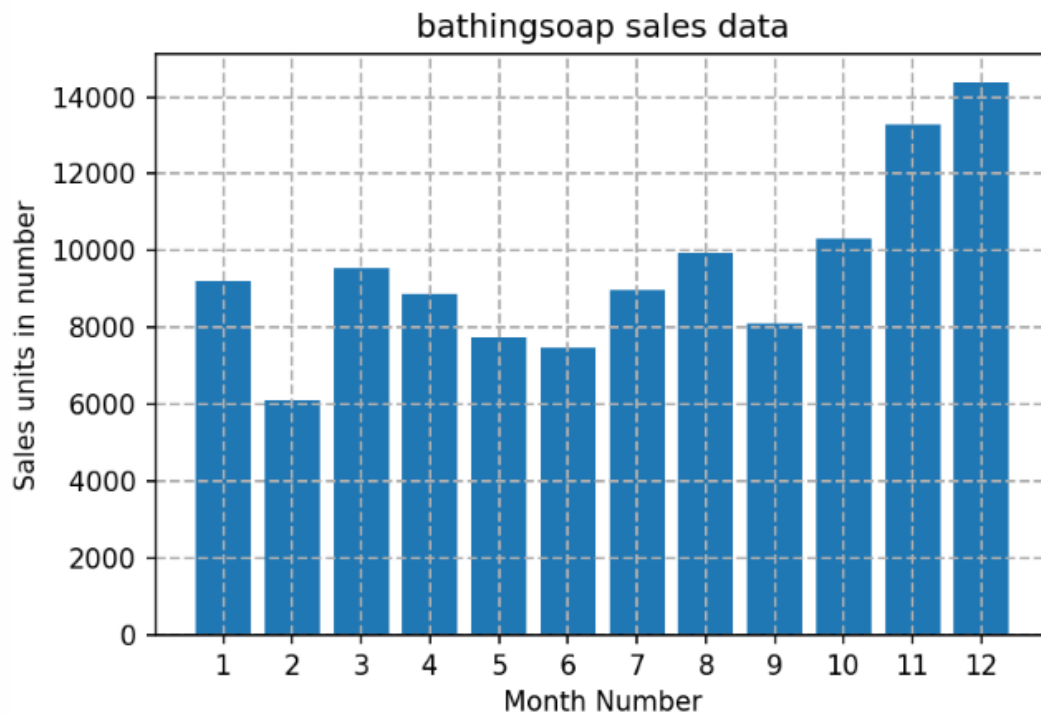
O gráfico de barras deve exibir o número de unidades vendidas por mês para cada produto. Adicione uma barra separada para cada produto no mesmo gráfico.

O gráfico de barras deve ficar assim.

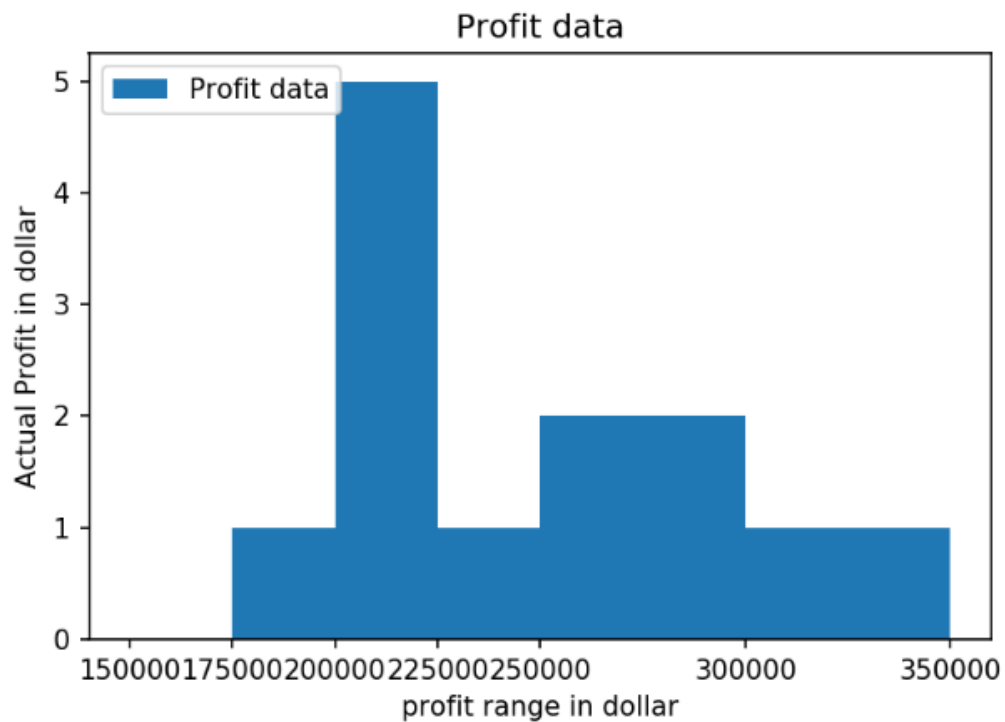




**Exercício 06** - Leia os dados de vendas de sabonetes de todos os meses e mostre-os em um gráfico de barras. Salve este gráfico em um diretório.



**Exercício 07:** Leia o lucro total de cada mês e mostre-o usando o histograma para ver as faixas de lucro mais comuns.

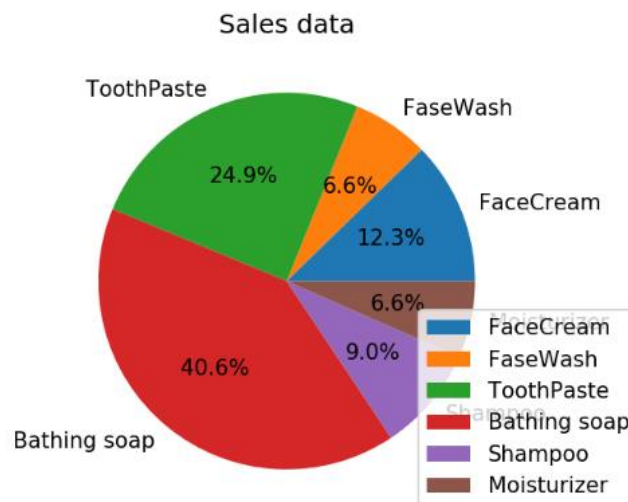




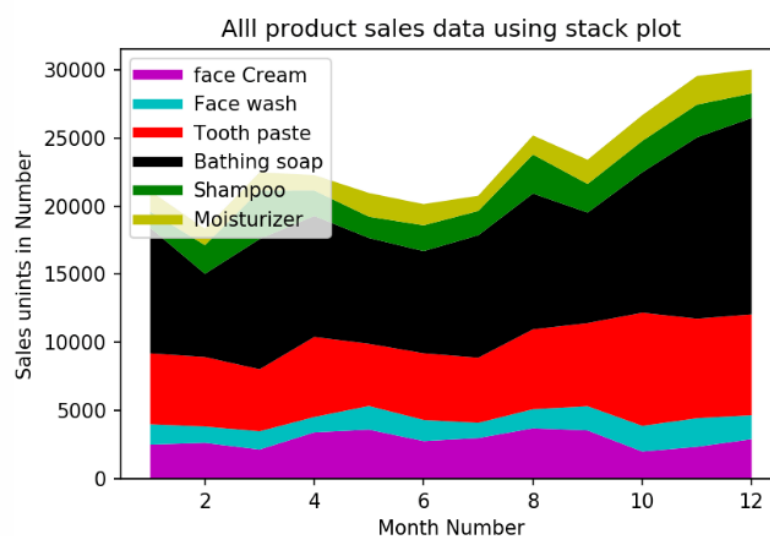
**Exercício 08:** Calcule os dados totais de vendas do ano passado para cada produto e mostre-os usando um gráfico de pizza.

**Nota:** No gráfico de pizza, exibe o número de unidades vendidas por ano para cada produto em porcentagem.

O gráfico de pizza deve ficar assim.



**Exercício 09:** Leia todos os dados de vendas de produtos e mostre-os usando o gráfico de pilha.





**Exercício 10:** Você é Analista de Dados na empresa XYZ Corporation International, uma revendedora de automóveis de luxo com sede em São Paulo. A empresa começou sua operação no Brasil em 2016 e atua nos quatro estados da região sudeste, mais os estados do Paraná e Bahia. Seu gerente vai apresentar os resultados da equipe comercial para o novo CEO da empresa e precisa da sua ajuda para construir uma análise que represente os dados de vendas no período de 2016 a 2019.

Sua fonte de dados é um arquivo Excel com dados coletados do sistema de vendas e CRM da empresa, com as seguintes colunas:

Coluna	Descrição
DataNotaFiscal	Data de emissão da nota fiscal
Fabricante	Fabricante do veículo
Estado	Estado onde foi realizada a venda
PrecoVenda	Preço de venda do veículo
PrecoCusto	Preço de custo do veículo para a empresa
TotalDesconto	Total de Desconto fornecido sobre o preço de venda
CustoEntrega	Custo de entrega do veículo ao proprietário
CustoMaoDeobra	Custo de Mão de Obra (secretária, mecânico, etc...)
NomeCliente	Nome do cliente que comprou o veículo
Modelo	Modelo do veículo
Cor	Cor do veículo
Ano	Ano de fabricação do veículo

Seu gerente precisa das seguintes informações:

- 1- Total de Vendas por Ano
- 2- Custo de Entrega do Veículo Por Fabricante
- 3- Custo de Mão de Obra Por Estado
- 4- Relatórios em Excel para anl

Além disso, pode ser interessante, se o CEO puder visualizar o total de vendas por estado e se as vendas estão acima ou abaixo da média. Seu gerente já sabe que um assunto será abordado pelo CEO durante a apresentação. O CEO está avaliando se continua ou não com a venda de automóveis da marca Jaguar e ele gostaria de saber como evoluíram as vendas de automóveis deste fabricante por ano e por estado.

#### Etapas:

- 1) Importar as bibliotecas Pandas, Matplotlib, Seaborn e OS.
- 2) Carregar os dados do arquivo DadosVendaCarros.xlsx;
- 3) Visualizar o cabeçalho do dataframe;
- 4) Renomear as colunas conforme a tabela abaixo:



DE	PARA
Coluna 1	DataNotaFiscal
Coluna 2	Fabricante
Coluna 3	Estado
Coluna 4	ValorVenda
Coluna 5	ValorCusto
Coluna 6	TotalDesconto
Coluna 7	CustoEntrega
Coluna 8	CustoMaoDeObra
Coluna 9	NomeCliente
Coluna 10	Modelo
Coluna 11	Cor
Coluna 12	Ano
Coluna 13	Mês

- 5) Verificar se há dados nulos ou faltantes.
- 6) Criar um gráfico de colunas para representar o total vendido por ano;
- 7) Criar um gráfico de linhas para analisar as vendas por mês separando por ano;
- 8) Criar um gráfico de barras para analisar o custo de entrega do veículo por fabricante;
- 9) Criar um gráfico de barras para representar o custo de mão de obra por estado;
- 10) Criar um código para gerar relatórios, filtrando a base de forma automática para exibir os dados por Fabricante, Estado e Modelo.

**Relatório Fabricante:**

Exibir as vendas por fabricante ao longo dos anos. O relatório deve exibir as seguintes colunas: Ano, Fabricante, PreçoVenda, Preço de Custo, Total de Desconto, Custo de Entrega e Custo da Mão de Obra.

**Relatório Estado:**

Deve exibir as vendas totais realizadas por cada filial, demonstrando as vendas durante o ano de 2018. O relatório deve exibir as seguintes colunas: Ano, Estado, PreçoVenda, Preço de Custo, Total de Desconto, Custo de Entrega e Custo da Mão de Obra.

**Relatório Modelo:**

Deve exibir as vendas por modelo de veículo. O relatório deve exibir as seguintes colunas: Fabricante, Modelo e PreçoVenda.

**Exercício 11:** Você é Analista de Dados na empresa PontoMaximo, uma rede de varejo que vende produtos eletrônicos e eletrodomésticos com lojas espalhadas por diversas cidades do Brasil.

A empresa começou sua operação no Brasil em 2012 e atua nos quatro estados da região sudeste mais os estados do Paraná e Bahia. A empresa está montando a estratégia de vendas para o próximo ano e precisa saber qual dos fabricantes dos produtos vendidos, apresenta melhor desempenho nas vendas. O objetivo é descartar os fabricantes cujos produtos possuem poucas vendas e tentar negociar melhores condições com os principais fabricantes.

Em paralelo a isso, a empresa gostaria de ter diferentes visões das vendas realizadas nos últimos 4 anos (período de 2012 a 2015). Deve ser possível segmentar as análises de vendas por diferentes informações e por diferentes ângulos. Estas informações irão suportar as estratégias da empresa para o próximo ano. Sua fonte de dados é um arquivo Excel com dados coletados do sistema de vendas, CRM e ERP da empresa. O conjunto de dados foi entregue pelo departamento de TI com as seguintes colunas.

Coluna	Descrição
ID-Produto	Identificador único de cada produto
Produto	Nome do produto
Categoria	Categoria do produto
Segmento	Segmento do produto
Fabricante	Fabricante do produto
Loja	Loja onde foi efetuada a venda
Cidade	Cidade da loja onde foi efetuada a venda
Estado	Estado da loja onde foi efetuada a venda
Vendedor	Nome do vendedor
ID-Vendedor	ID-Vendedor
DataVenda	Data da venda
ValorVenda	Valor da venda



**ETAPAS:**

- 1) Importar as bibliotecas Pandas, Matplotlib, Seaborn e OS.
- 2) Carregar os dados do arquivo DadosVendaCarros.xlsx;
- 3) Visualizar o cabeçalho do dataframe;
- 4) Renomear as colunas conforme a tabela abaixo:

DE	PARA
Coluna 1	ID-Produto
Coluna 2	Produto
Coluna 3	Categoria
Coluna 4	Segmento
Coluna 5	Fabricante
Coluna 6	Loja
Coluna 7	Cidade
Coluna 8	Estado
Coluna 9	Vendedor
Coluna 10	ID-Vendedor
Coluna 11	Data Venda
Coluna 12	ValorVenda
Coluna 13	ANO
Coluna 14	Mês

- 5) Verificar se há dados nulos ou faltantes.
- 6) Criar um gráfico de colunas para representar o total vendido por ano;
- 7) Criar um gráfico de linhas para analisar as vendas por mês separando por ano;
- 8) Criar um gráfico de pizza para representar o total vendido por segmento;
- 9) Crie um gráfico de barras para representar o valor de venda de obra por estado;
- 10) Crie um gráfico de barras para representar o a média de venda de obra por estado;
- 11) Crie um gráfico para cada ano, para representar o total de vendas por loja.



## Exercício 12: Análise de Satisfação de Bancos

Imagine que você foi contratado por uma instituição financeira de grande porte para analisar a satisfação dos clientes com os serviços prestados pela empresa. Como analistas de dados, sua missão é coletar, processar e analisar um grande volume de dados para obter insights valiosos sobre o que os clientes pensam sobre a empresa, suas necessidades e expectativas.

Para começar, vamos coletar dados de diversas fontes, como pesquisas de satisfação, registros de atendimento ao cliente, feedbacks nas redes sociais e reclamações registradas em plataformas de avaliação de empresas. Com esses dados em mãos, podemos criar um banco de dados robusto e padronizado para análise.

### Faça as seguintes análises

- 1) Analisar NPS de classificação de cada instituição
- 2) Calcular a % de elogios do app para cada instituição
- 3) Calcular a % de reclamações do app para cada instituição
- 4) Calcular a % de elogios a instituição para cada instituição
- 5) Calcular a % de reclamação a instituição para cada instituição
- 6) Plote um gráfico de linhas para entender a tendência das classificações por instituição

### Etapas:

- 1) Importar as bibliotecas Pandas, Seaborn e Matplotlib;
- 2) Carregue o arquivo dados-satisfação.xlsx;
- 3) Realizar o tratamento dos dados categóricos, converta-os para dados numéricos. As colunas que terão os dados convertidos serão: Elogio quanto ao app, Reclamação quanto ao app, Elogio a Instituição, Reclamação a Instituição e Não Classificável.

Nestas colunas, você deve substituir NÃO para 0 e SIM para 1. Converta os dados para o tipo INT.



### Análise de NPS por instituição

- 4) Faça as seguintes substituições na coluna classificação:
  - $2 > -1$
  - $1 > -1$
  - $5 > 1$
  - $4 > 1$
  - $3 > 0$
- 5) Crie um dataframe chamado `df_classificacao`, para agrupar os dados por instituição, calculando a média da da coluna Classificação.
- 6) Multiplique os valores da coluna Classificação por 100.
- 7) Carregue a paleta de cores Accent, com o comando `sns.set_palette`;
- 8) Carregue o estilo darkgrid, utilizando o comando `sns.set_style`;
- 9) Gere um gráfico de colunas para apresentar a classificação média por instituição, e escreva um texto com sua análise do NPS.

### Análise do % de elogios e relcamações do app para cada instituição

- 10) Repita os passos 5, 6 e 9, considerando a coluna Elogio quanto ao app e Reclamação quanto ao App;

### Análise do % de elogios e reclamações por instituição

- 11) Repita os passos 5, 6 e 9, considerando a coluna Elogio quanto a instituição e Reclamação quanto a instituição;

### Análise temporal da classificação por instituição

- 12) Crie um gráfico de linhas utilizando a biblioteca Seaborn, para demonstrar a evolução da classificação por instituição.

**Exercício 13:** Uma grande empresa pretende adquirir uniformes para todos os seus 400 funcionários. Como as pessoas possuem diferentes estaturas, precisam de uniformes de tamanhos diferentes. Ao invés de medir a altura de todos os funcionários, o que levaria muito tempo e daria muito trabalho, a empresa decidiu utilizar um método estatístico, neste caso, utilize o arquivo *alturas.xlsx* e gere um histograma para analisar a distribuição das alturas.

**Exercício 14: Histogramas para dados de produção**

A empresa TODOSUNIDOS recebeu um lote inicial de eixos para montagem de motores destinados a servomecanismos de aeronaves. Foi necessária uma verificação para identificar a capacidade de cada um de três fornecedores, visto que pretendia-se firmar um futuro contrato de parceria. Optou-se por utilizar histograma da distribuição dos diâmetros externos de uma amostra aleatória de 240 eixos, proveniente dos três fornecedores. O limite de especificação dos diâmetros é  $5 \pm 2$  (cm).

Arquivo: eixos.xls

Planilha: eixos

- a) Que conclusões você obtém por meio de uma análise preliminar do histograma com todos os eixos?
- b) Que conclusões você obtém por meio de uma análise preliminar dos histogramas por fornecedor?
- c) Que outro gráfico auxiliaria você na análise solicitada no item anterior?

Se você fosse o gestor desse processo, baseando-se somente nessa ferramenta, qual(is) o(s) fornecedor(es) que você selecionaria para o contrato de parceria?

**Exercício 15:** Uma empresa tem um conjunto de dados sobre a dimensão de um determinado produto feito com matéria-prima de três fornecedores. Você foi designado a desenvolver uma análise para conseguir analisar quais desses fornecedores tem a matéria prima que não garante o dimensional final deste produto, a média máxima é de 120,15 mm e a dimensão mínima é de 119,9 mm. Crie um histograma e um boxplot para analisar as medidas de cada fornecedor e um boxplot para comparar os três fornecedores. Utilize o *arquivo fornecedores.xlsx*



## Exercício 16 – Gráfico de dispersão

Uma indústria alimentícia, produtora de peças de presunto e de outros embutidos, iniciou o giro do ciclo PDCA de melhorias com a meta de reduzir para 5% a porcentagem de quebra durante o fatiamento do produto, até o final do ano. Esta decisão foi tomada após a confirmação de que a atual porcentagem de 12% não atendia às expectativas do cliente.

Na etapa observação do problema, levantou-se a hipótese de que a quebra no fatiamento após o resfriamento do presunto poderia ser dependente da temperatura interna do produto (decorrente do processo de fabricação). Com base nesta suposição, a equipe responsável pela solução do problema providenciou um levantamento de dados a fim de conhecer melhor a relação entre a temperatura interna do produto, a quebra no fatiamento e o tipo de produto fabricado.

Fonte de dados: *alimentos.xlsx*

- a) Analisar a relação entre temperatura interna do produto e % de quebra durante fatiamento do produto considerando todas as observações.
- b) Analisar a relação entre temperatura interna do produto e % de quebra durante fatiamento do produto estratificando por tipo de produto.