Em primeiro lugar, foi utilizado a importação das bibliotecas necessárias:

* Para a Manipulação de Dados:
  + Pandas
  + Numpy
* Para a visualização de Dados:
  + Seaborn
  + Matplotlib.pyplot

Uma imagem com texto

Descrição gerada automaticamente

Após importar as bibliotecas necessárias, foi carregado o Dataset utilizando, onde foi utilizado um DataSet com todos os dados, e outro dataSet sem os anos 2020 e 2021, devido à situação de COVID-19.

Uma imagem com texto

Descrição gerada automaticamente

Após a importação das bibliotecas, e dos *DataSets* com os dados que iremos utilizar, foi começado a criação das variáveis por cada mês, onde *janeiro\_comCOVID* é a variável que utilizamos para colocar apenas o mês de janeiro de todos os anos, onde o resto das variáveis fazem o mesmo para o resto dos meses, e *janeiro\_semCOVID*, que é a mesma variável, mas sem os anos 2020 e 2021.

Uma imagem com texto

Descrição gerada automaticamente

**Número de Chamadas por ano**

Como já exposta no capítulo de Identificação do Problema, foi criado um gráfico para verificar o número de chamadas por ano, de 2012 a 2022. A partir deste gráfico, é possível observar uma tendência de crescimento no número de chamadas ao longo dos anos, com um aumento expressivo no ano 2022, quando a pandemia de COVID-19 acabou.

Uma imagem com gráfico

Descrição gerada automaticamente

Explicação do código: Na primeira linha, é utilizado o método *groupby* para agrupar as chamadas por ano e somar a quantidade de chamadas de cada ano. O resultado é armazenado na variável *calls\_by\_year*. Em seguida, é utilizado a biblioteca *Matplotlib* para criar o gráfico de barras. A função *plot* é utilizada para criar o gráfico, especificando que o tipo de gráfico é 'bar', a cor é azul e a transparência é de 0,7. O eixo x do gráfico mostra o ano e o eixo y mostra o número de chamadas. Após adicionar o título, rótulos para os eixos x e y, e uma legenda, é retirado a notação científica dos valores do eixo y e rodado num angulo de 45 graus os valores do eixo x, evitando assim sobreposições dos mesmos. Foi então depois criado um *loop for*, para adicionar o valor de cada barra no topo da barra correspondente, com a função *texto*, tornando o gráfico mais informativo. Por fim, o gráfico é guardado num ficheiro de imagem com o nome ‘Aumento de Chamadas Recebidas.png’.

Este gráfico mostra uma coisa importante, sendo que existe um aumento de chamadas recebidas por ano, esse valores de chamadas diminuiu nos anos 202 e 2021, sendo que foram os anos da pandemia, e aumentou expressivamente no ano 2022, sendo esse ano o final da pandemia

De seguida, utilizando a variável ax, foi criado o gráfico de barrar para visualizar o número de chamadas recebidas por ano, onde existe um claro aumento de chamadas até ao ano 2020, onde no ano 2022 voltou a subir drasticamente.

Uma imagem com texto

Descrição gerada automaticamente

Para podermos ver a tendência do aumento do número de chamadas recebidas por ano, foi feita uma linha de Regressão, mostrada na imagem abaixo. Foi feita esta linha de Regressão com todos os anos, e retirando os anos 2020 e 2021, onde é notável que a linha de regressão do gráfico sem os dados de 2020 e 2021 tem uma maior inclinação, sendo que o número de chamadas recebidas durante o COVID foi menor.

Uma imagem com gráfico

Descrição gerada automaticamente

Uma imagem com gráfico

Descrição gerada automaticamente

Explicação do código: Após criar a variável *df\_grouped*, que agrupa a soma do número de chamadas por ano, é então criado o gráfico e a linha de regressão com a biblioteca seaborn

Uma imagem com texto, monitor, cobertura, captura de ecrã

Descrição gerada automaticamenteDe seguida, foi feita uma linha de regressão Linear do número total de chamadas por ano mas apenas do mês de janeiro, com e sem os anos 2020 e 2021. Para fazer este gráfico, foi primeiro feito duas funções, *numeroChaPorAno\_Mes\_semCOVID* e *numeroChaPorAno\_Mes\_comCOVID*, onde calcula o número de chamadas por ano apenas no mês escolhido. A função *numeroChaPorAno\_Mes\_semCOVID* é utilizado as variáveis feitas anteriormente sem os anos 2020 e 2021 e a função *numeroChaPorAno\_Mes\_comCOVID* é utilizado as variáveis feitas anteriormente com todos os anos.

Apos fazermos estas funções, criamos o gráfico igualmente como foi criado o gráfico anterior, calculando o total por ano, mas apenas do mês de Janeiro, com e sem os anos 2020 e 2021.

IMG

**Número total de chamadas por mês**

Para determinar em que meses existe uma maior media de chamadas, foi feito este próximo gráfico, que demonstra que Janeiro e Dezembro são os meses com um maior número de chamadas recebidas, e o mês de Abril o mês com menor número de chamadas recebidas.

Uma imagem com gráfico

Descrição gerada automaticamente

Explicação do Código: Igualmente como feito o gráfico do número total de chamadas por ano, a variável *calls\_by\_mouth* é um dataSet agrupado por Mês, onde é feito a medio do numero de chamadas recebidas por cada mês entre 2012 e 2022.

Uma imagem com texto

Descrição gerada automaticamente

**Flutuação do Número de Chamadas Recebidas por mês entre 2016 e 2022**

Para este gráfico, foi criado novamente uma função para determinar o numero de chamadas por mês de um ano especifico.

Uma imagem com texto

Descrição gerada automaticamente

É então chamada esta função para vermos o número de chamadas por mês dos anos 2016, 2017, 2018, 2019 e 2022, sendo que mostra uma flutuação bastante comum entre estes vários anos. É comum nestes anos haver uma descida do número de chamadas para fevereiro, abril, junho, setembro e novembro. Pelo contrário, é comum também haver um aumento do número de chamadas recebidas para os meses de março, maio, julho, outubro e dezembro.

Uma imagem com gráfico

Descrição gerada automaticamente

Explicação do Código: Após definir o tamanho do gráfico, são criadas várias variáveis com dataSets do numero de chamadas recebidas por mês, do ano escolhido. É depois então criado para cada uma dessas variáveis uma linha no gráfico.

Uma imagem com texto

Descrição gerada automaticamente

**Linhas de Regressão Polinomial:**

Para ser possível prever o número de chamadas nos próximos meses/anos, decidimos criar uma linha de regressão pelo menos para os três primeiros meses, onde fazemos um gráfico com o numero total de cada ano, entre 2012 e 2022 do mês de janeiro, fevereiro e março, criamos uma linha de regressão polinomial e prevemos o numero total de 2023 dos respetivos meses, sendo que o resultado foi o seguinte

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Mês** | **Chamadas previstas para 2023 (incluindo 2020 e 2021)** | **Chamadas previstas para 2023 (excluindo 2020 e 2021)** | **Número de chamadas Real** | **Erro estimado (incluindo 2020 e 2021)** | **Erro estimado (excluindo 2020 e 2021)** |
| Janeiro | 125 508 | 127 354 | 129 474 | 3,06% | 1,64% |
| Fevereiro |  |  |  |  |  |
| Março |  |  |  |  |  |