## Introdução

Em nossos estudos, fizemos o cálculo da média, mediana, variância e desvio padrão da amostra de voos feitos entre 2014 e 2023, saindo pelos aeroportos de Guarulhos ou Congonhas. Iremos então aplicar uma análise vertical para entender a evolução dos parâmetros conforme os anos passam.

## Automação de processo

O processo do cálculo das medidas de dispersão foi automatizado usando linguagem Python, com o seguinte script:

```
tabelaMediaPorMes = contagemSemFiltro.groupby(['Ano', 'Mes'])['quantidade'].sum().reset_index()
print(tabelaMediaPorMes)

estatisticas_ano = tabelaMediaPorMes.groupby('Ano')['quantidade'].agg(
    mean='mean',
    median='median',
    var='var',
    std=lambda x: x.std(ddof=0) # ou ddof=0 para população
).reset_index()
```

Nele, temos um dataframe da biblioteca Pandas, com os dados do nº de voos por ano, chamado contagemSemFiltro. Ele é então filtrado por cada ano e mês, e o cálculo é feito usando funções nativas do Pandas.

O processo da análise Vertical foi feito a partir dessa tabela, que foi importada para o Excel, onde foi usado uma fórmula simples para comparar anos subsequentes.

| Tabela | WC III | MIMES | uc | DODICE | <br>WID. | JULUU |
|--------|--------|-------|----|--------|----------|-------|

| Ano  | MÉDIA       | MEDIANA | VAR         | DESV PADR   |
|------|-------------|---------|-------------|-------------|
| 2014 | 16277,18182 | 17968   | 29650434,96 | 5191,814796 |
| 2015 | 18085,58333 | 18089,5 | 478991,7197 | 662,6279069 |
| 2016 | 17184,41667 | 17233,5 | 426473,9015 | 625,2474788 |
| 2017 | 17099,16667 | 17287   | 595356,8788 | 738,7447499 |
| 2018 | 18407,16667 | 18673,5 | 2519405,606 | 1519,689159 |
| 2019 | 18669,66667 | 18795,5 | 1385440,97  | 1126,937246 |
| 2020 | 7994,583333 | 5780,5  | 44621529,36 | 6395,550686 |
| 2021 | 10281,33333 | 11037,5 | 8448675,697 | 2782,915627 |
| 2022 | 15293,41667 | 15448,5 | 2745780,811 | 1586,494798 |
| 2023 | 18102,83333 | 18255   | 1133721,788 | 1019,433653 |

Tabela de Análise Vertical

| Ano  | Evolução<br>Média | Evolução<br>Mediana | Evolução<br>Variância | Evolução Desvio<br>Padrão |
|------|-------------------|---------------------|-----------------------|---------------------------|
| 2014 | 0,00%             | 0,00%               | 0,00%                 | 0,00%                     |
| 2015 | 111,11%           | 100,68%             | 1,62%                 | 12,76%                    |
| 2016 | 95,02%            | 95,27%              | 89,04%                | 94,36%                    |
| 2017 | 99,50%            | 100,31%             | 139,60%               | 118,15%                   |
| 2018 | 107,65%           | 108,02%             | 423,18%               | 205,71%                   |
| 2019 | 101,43%           | 100,65%             | 54,99%                | 74,16%                    |
| 2020 | 42,82%            | 30,75%              | 3220,75%              | 567,52%                   |
| 2021 | 128,60%           | 190,94%             | 18,93%                | 43,51%                    |
| 2022 | 148,75%           | 139,96%             | 32,50%                | 57,01%                    |
| 2023 | 118,37%           | 118,17%             | 41,29%                | 64,26%                    |