

## Relatório de Amostragem (A3)

1. Este é um relatório referente a Avaliação A3 sobre amostragem. Neste relatório, gerei uma base de dados aleatória utilizando a linguagem de programação Python sobre idades com 500 linhas de dados.
2. A população escolhida, seria provavelmente, pessoas próximas a mim em redes sociais, círculos de amigos, funcionários da empresa em que trabalho, colegas da faculdade, vizinhos etc.
3. Para a análise, escolhi os métodos de amostragem aleatória Simples e Estratificada.
4. Utilizei as ferramentas Python e Excel, para exemplificar e tratar os dados da coleta e separação dos dados da amostragem
5. Para a coleta de dados, criei um formulário online para exemplificar o formato que seria feita a estimativa de idades.

Link: [https://docs.google.com/forms/d/e/1FAIpQLScuOJ3PDjUGVH-VYgW7q57EciB6Da2UWiaWU-29RTt8F7PpoQ/viewform?usp=sf\\_link](https://docs.google.com/forms/d/e/1FAIpQLScuOJ3PDjUGVH-VYgW7q57EciB6Da2UWiaWU-29RTt8F7PpoQ/viewform?usp=sf_link)

6. A amostragem simples, traz dados mais genéricos e sem muito tratamento, é possível estimar a média de idades utilizando-a, verificar qual a idade mínima e máxima etc.


Já na amostra estratificada, temos dados mais separados e tratados, podendo separar por faixas etárias, por exemplo.

Neste caso, se precisa de uma análise mais simples e genérica, a amostragem aleatória simples pode ser interessante, mas caso queira uma amostragem mais detalhada e dividida, a amostragem aleatória estratificada seria mais indicada.

Obs.: A base de dados tem um total de 500 dados, retirei 20% deles, o que gerou 100 linhas e fiz a análise destas 100 linhas, representando a amostragem.

Calculei a média de idades para cada amostragem também.

Terminal gerado pela amostragem aleatória simples:

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS SQL CONSOLE  Code

Amostra Aleatória Simples (20% da base):
●   ID  Idade
2   3    14
7   8     1
12  13    9
17  18    8
22  23    5
..  ...   ...
477 478   86
482 483   67
487 488   87
492 493   73
497 498   76

[100 rows x 2 columns]
Média de idade da amostra aleatória simples:
32.0
Média de idade da base de dados completa:
32.0
Amostra Aleatória Simples (20% da base) salva em amostra_aleatoria_simples.csv
```

Terminal gerado pela amostragem aleatória estratificada:

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS SQL CONSOLE

Médias de idade por faixa etária:
Faixa etária: 0-14 anos, Média de idade: 7
Faixa etária: 15-24 anos, Média de idade: 18
Faixa etária: 25-34 anos, Média de idade: 29
Faixa etária: 35-44 anos, Média de idade: 39
Faixa etária: 45-54 anos, Média de idade: 48
Faixa etária: 55-64 anos, Média de idade: 57
Faixa etária: 65+ anos, Média de idade: 80

Porcentagem de cada faixa etária:
25-34 anos    22.45
0-14 anos     18.37
15-24 anos    18.37
35-44 anos    16.33
45-54 anos    10.20
55-64 anos     8.16
65+ anos      6.12
Name: Faixa_Etaria, dtype: float64
```

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS SQL CONSOLE Code +

Quantidade de cada faixa etária:
25-34 anos    22
0-14 anos     18
15-24 anos    18
35-44 anos    16
45-54 anos    10

Amostra Estratificada por Faixa Etária (20% da base):
  ID  Idade  Faixa_Etaria
7    8     1    0-14 anos
12   13     9    0-14 anos
18   19    11    0-14 anos
23   24     2    0-14 anos
28   29     3    0-14 anos
..   ...   ...      ...
477  478    86    65+ anos
482  483    67    65+ anos
487  488    87    65+ anos
492  493    73    65+ anos
497  498    76    65+ anos

[98 rows x 3 columns]

Amostra estratificada por faixa etária salva em amostra_aleatoria_estratificada.csv
```

Arquivos do código, para fins de teste:



Análise A3.zip