

Obtendo os betas

Considere o seguinte conjunto de dados:

UF	Y	X
Rondônia	670	1,59
Acre	220	0,56
Amazonas	1202	2,68
Roraima	188	0,47
Pará	1869	5,2
Amapá	248	0,58
Tocantins	477	1,32
Maranhão	1294	3,88
Piauí	816	2,11
Ceará	2671	5,53
Rio Grande do Norte	1403	2,6
Paraíba	1586	2,94
Pernambuco	3468	6,62
Alagoas	973	1,91
Sergipe	701	1,48
Bahia	5310	10,64
Minas Gerais	10950	22,39
Espírito Santo	2008	4,2
Rio de Janeiro	9574	21,9
São Paulo	28863	59,66
Paraná	6466	14,22
Santa Catarina	4274	9,57
Rio Grande do Sul	6432	14,67
Mato Grosso do Sul	1326	3,28
Mato Grosso	1423	3,49
Goiás	3211	6,94
Distrito Federal	2140	6,25

Onde temos o número de domicílios que possuem computador (Y) e a renda total recebida por pessoas de 10 anos ou mais (X), em cada um dos estados brasileiros.

Acredita-se que a posse de computadores tenha uma relação linear com a renda das famílias e para comprovar isso, podemos estimar uma regressão linear com os dados acima e avaliar os resultados obtidos.

Neste problema, precisamos, inicialmente, obter as estimativas dos parâmetros de regressão.

Informações extras:

O valor da renda está dividido por 1.000.000, isto é, se um estado tem uma renda total de R\$ 4,2, significa que ele tem uma renda total de R\$ 4.200.000.

Utilize o ferramental do **statsmodels** para obter os parâmetros.

Segue o *dataset* em formato de dicionário Python para facilitar a solução do problema:

```
dataset = {  
    'Y': [670, 220, 1202, 188, 1869, 248, 477, 1294, 816, 2671, 1403, 1586, 3468, 973, 701,  
    'X': [1.59, 0.56, 2.68, 0.47, 5.2, 0.58, 1.32, 3.88, 2.11, 5.53, 2.6, 2.94, 6.62, 1.91,  
}]
```

Assinale a alternativa abaixo que apresenta as estimativas dos parâmetros de regressão.

Selecione uma alternativa

A $\beta_1 = 0,374282$
 $\beta_2 = 0,002071$

B $\beta_1 = 0$
 $\beta_2 = 474,691893$

C $\beta_1 = -168,321156$
 $\beta_2 = 481,390397$

D $\beta_1 = 481,390397$
 $\beta_2 = -168,321156$