Questão 2

Pietro Piva Vieira

1° de Agosto de 2022

# Introdução

Para a Questão 2, a associação entre idade do imóvel e taxas cobradas deve ser avaliada através do diagrama de dispersão e coeficiente de correlação. Para tanto, um banco de dados é fornecido:

# Realizamos a leitura do banco de dados  
zCityData = read.table("TRABALHO2.txt", header = TRUE)  
  
# Com a biblioteca knitr, estruturamos uma tabela  
knitr::kable(zCityData, caption = "Dados de Imóveis Residenciais comercializados na Cidade Z em 2019")

Dados de Imóveis Residenciais comercializados na Cidade Z em 2019

| casa | local | tam | prv | util | id | ter | tax | prc |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | bairroD | media | 142.0 | 260.1 | 60 | 840 | 3167 | 114.1 |
| 2 | bairroC | media | 175.0 | 167.2 | 12 | 1290 | 4033 | 163.8 |
| 3 | bairroB | pequena | 129.0 | 120.8 | 41 | 990 | 1471 | 115.8 |
| 4 | bairroC | pequena | 138.0 | 157.9 | 22 | 1380 | 3204 | 176.2 |
| 5 | bairroD | media | 232.0 | 232.3 | 5 | 6150 | 3613 | 272.4 |
| 6 | bairroD | media | 135.0 | 167.2 | 9 | 1710 | 3028 | 110.5 |
| 7 | bairroA | media | 150.0 | 185.8 | 18 | 12000 | 3131 | 148.9 |
| 8 | bairroB | media | 207.0 | 204.4 | 16 | 6660 | 5158 | 227.6 |
| 9 | bairroD | grande | 271.0 | 278.7 | 30 | 1590 | 5702 | 276.6 |
| 10 | bairroD | pequena | 89.0 | 92.9 | 43 | 900 | 2054 | 94.8 |
| 11 | bairroA | pequena | 153.0 | 204.4 | 18 | 1140 | 4127 | 129.2 |
| 12 | bairroB | pequena | 86.5 | 148.6 | 50 | 1950 | 1445 | 85.0 |
| 13 | bairroD | media | 234.0 | 232.3 | 2 | 4830 | 2087 | 198.6 |
| 14 | bairroC | media | 105.5 | 185.8 | 13 | 660 | 2818 | 87.0 |
| 15 | bairroA | media | 175.0 | 204.4 | 15 | 6180 | 3917 | 194.1 |
| 16 | bairroC | media | 165.0 | 157.9 | 33 | 1380 | 2220 | 159.9 |
| 17 | bairroD | media | 166.0 | 213.7 | 37 | 810 | 3498 | 170.9 |
| 18 | bairroC | pequena | 136.0 | 176.5 | 22 | 1890 | 3607 | 141.0 |
| 19 | bairroB | pequena | 148.0 | 157.9 | 13 | 1080 | 3648 | 107.0 |
| 20 | bairroC | media | 151.0 | 176.5 | 24 | 1020 | 3561 | 171.8 |
| 21 | bairroB | media | 180.0 | 223.0 | 10 | 4650 | 4681 | 223.2 |
| 22 | bairroB | media | 293.0 | 241.5 | 6 | 1380 | 7088 | 263.0 |
| 23 | bairroA | media | 167.0 | 185.8 | 46 | 1380 | 3482 | 159.2 |
| 24 | bairroC | grande | 190.0 | 204.4 | 37 | 1440 | 3920 | 164.8 |
| 25 | bairroA | grande | 184.0 | 195.1 | 27 | 3900 | 4162 | 199.5 |
| 26 | bairroD | media | 157.0 | 185.8 | 7 | 900 | 3785 | 127.3 |
| 27 | bairroC | media | 110.0 | 148.6 | 26 | 870 | 3103 | 84.2 |
| 28 | bairroB | media | 135.0 | 167.2 | 35 | 1290 | 3363 | 120.7 |
| 29 | bairroB | media | 567.0 | 594.6 | 4 | 2550 | 12192 | 627.6 |
| 30 | bairroD | media | 180.0 | 185.8 | 11 | 3000 | 3831 | 211.6 |
| 31 | bairroA | pequena | 183.0 | 157.9 | 16 | 9000 | 3564 | 171.3 |
| 32 | bairroB | media | 185.0 | 185.8 | 56 | 19470 | 3765 | 182.8 |
| 33 | bairroD | media | 152.0 | 157.9 | 33 | 2100 | 3361 | 140.1 |
| 34 | bairroA | pequena | 148.0 | 120.8 | 22 | 1170 | 3950 | 143.5 |
| 35 | bairroA | pequena | 152.0 | 139.4 | 25 | 1770 | 3055 | 169.4 |
| 36 | bairroC | pequena | 146.0 | 148.6 | 31 | 1080 | 2950 | 137.6 |
| 37 | bairroC | media | 170.0 | 223.0 | 33 | 1710 | 3346 | 193.9 |
| 38 | bairroB | media | 127.0 | 185.8 | 65 | 1200 | 3334 | 154.3 |
| 39 | bairroA | grande | 265.0 | 334.4 | 33 | 3600 | 5853 | 233.5 |
| 40 | bairroD | media | 157.0 | 167.2 | 12 | 3390 | 3982 | 164.5 |
| 41 | bairroD | media | 128.0 | 157.9 | 25 | 1560 | 3374 | 159.3 |
| 42 | bairroD | media | 110.0 | 139.4 | 11 | 1770 | 3119 | 114.4 |
| 43 | bairroB | media | 123.0 | 167.2 | 43 | 1170 | 3268 | 117.0 |
| 44 | bairroC | grande | 212.0 | 362.3 | 202 | 12870 | 3648 | 212.6 |
| 45 | bairroC | media | 145.0 | 167.2 | 44 | 660 | 2783 | 136.2 |
| 46 | bairroD | pequena | 129.0 | 92.9 | 15 | 3000 | 2438 | 162.0 |
| 47 | bairroA | media | 143.0 | 195.1 | 10 | 3600 | 3529 | 155.3 |
| 48 | bairroB | media | 247.0 | 269.4 | 4 | 3750 | 4626 | 225.5 |
| 49 | bairroB | pequena | 111.0 | 139.4 | 97 | 3330 | 3205 | 116.3 |
| 50 | bairroA | pequena | 133.0 | 241.5 | 42 | 1080 | 3059 | 119.8 |
| 51 | bairroA | grande | 160.0 | 260.1 | 60 | 840 | 3169 | 162.7 |
| 52 | bairroA | media | 180.0 | 167.2 | 12 | 1290 | 4035 | 151.2 |
| 53 | bairroD | pequena | 132.0 | 120.8 | 41 | 990 | 1473 | 127.7 |
| 54 | bairroB | pequena | 140.0 | 157.9 | 22 | 1380 | 3206 | 139.4 |
| 55 | bairroC | media | 240.0 | 232.3 | 5 | 6150 | 3615 | 278.3 |
| 56 | bairroC | media | 140.0 | 167.2 | 9 | 1710 | 3030 | 153.3 |
| 57 | bairroC | media | 160.0 | 185.8 | 18 | 12000 | 3133 | 153.0 |
| 58 | bairroC | media | 225.0 | 204.4 | 16 | 6660 | 5160 | 210.4 |
| 59 | bairroD | grande | 285.0 | 278.7 | 30 | 1590 | 5704 | 195.5 |
| 60 | bairroA | pequena | 90.0 | 92.9 | 43 | 900 | 2056 | 81.0 |
| 61 | bairroB | pequena | 157.0 | 204.4 | 18 | 1140 | 4129 | 161.3 |
| 62 | bairroC | pequena | 90.0 | 148.6 | 50 | 1950 | 1447 | 85.2 |
| 63 | bairroD | media | 238.0 | 232.3 | 2 | 4830 | 2089 | 265.8 |
| 64 | bairroD | media | 116.0 | 185.8 | 13 | 660 | 2820 | 117.0 |
| 65 | bairroB | media | 180.0 | 204.4 | 15 | 6180 | 3919 | 195.6 |
| 66 | bairroC | media | 170.0 | 157.9 | 33 | 1380 | 2222 | 140.9 |
| 67 | bairroC | media | 170.0 | 213.7 | 37 | 810 | 3500 | 180.6 |
| 68 | bairroB | pequena | 140.0 | 176.5 | 22 | 1890 | 3609 | 87.8 |
| 69 | bairroB | pequena | 160.0 | 157.9 | 13 | 1080 | 3650 | 198.0 |
| 70 | bairroA | media | 153.0 | 176.5 | 24 | 1020 | 3563 | 139.1 |
| 71 | bairroA | media | 190.0 | 223.0 | 10 | 4650 | 4683 | 167.4 |
| 72 | bairroA | media | 305.0 | 241.5 | 6 | 1380 | 7090 | 239.3 |
| 73 | bairroC | media | 170.0 | 185.8 | 46 | 1380 | 3484 | 171.2 |
| 74 | bairroC | grande | 193.0 | 204.4 | 37 | 1440 | 3922 | 180.7 |
| 75 | bairroC | grande | 190.0 | 195.1 | 27 | 3900 | 4164 | 170.7 |
| 76 | bairroB | media | 165.0 | 185.8 | 7 | 900 | 3787 | 164.9 |
| 77 | bairroA | media | 115.0 | 148.6 | 26 | 870 | 3105 | 81.4 |
| 78 | bairroA | media | 145.0 | 167.2 | 35 | 1290 | 3365 | 139.9 |
| 79 | bairroC | media | 625.0 | 594.6 | 4 | 2550 | 12194 | 627.3 |
| 80 | bairroB | media | 185.0 | 185.8 | 11 | 3000 | 3833 | 187.9 |
| 81 | bairroD | pequena | 188.0 | 157.9 | 16 | 9000 | 3566 | 201.3 |
| 82 | bairroD | media | 193.0 | 185.8 | 56 | 19470 | 3767 | 175.2 |
| 83 | bairroD | media | 155.0 | 157.9 | 33 | 2100 | 3363 | 177.7 |
| 84 | bairroA | pequena | 153.0 | 120.8 | 22 | 1170 | 3952 | 158.5 |
| 85 | bairroC | pequena | 159.0 | 139.4 | 25 | 1770 | 3057 | 129.5 |
| 86 | bairroA | pequena | 150.0 | 148.6 | 31 | 1080 | 2952 | 132.6 |
| 87 | bairroB | media | 190.0 | 223.0 | 33 | 1710 | 3348 | 179.9 |
| 88 | bairroC | media | 130.0 | 185.8 | 65 | 1200 | 3336 | 140.2 |
| 89 | bairroD | grande | 270.0 | 334.4 | 33 | 3600 | 5855 | 252.5 |
| 90 | bairroB | media | 163.0 | 167.2 | 12 | 3390 | 3984 | 141.1 |
| 91 | bairroA | media | 135.0 | 157.9 | 25 | 1560 | 3376 | 149.5 |
| 92 | bairroC | media | 120.0 | 139.4 | 11 | 1770 | 3121 | 146.9 |
| 93 | bairroB | media | 130.0 | 167.2 | 43 | 1170 | 3270 | 119.8 |
| 94 | bairroC | grande | 230.0 | 362.3 | 202 | 12870 | 3650 | 214.3 |
| 95 | bairroA | media | 145.0 | 167.2 | 44 | 660 | 2785 | 117.6 |
| 96 | bairroA | pequena | 135.0 | 92.9 | 15 | 3000 | 2440 | 86.4 |
| 97 | bairroD | media | 145.0 | 195.1 | 10 | 3600 | 3531 | 154.6 |
| 98 | bairroB | media | 252.0 | 269.4 | 4 | 3750 | 4628 | 276.8 |
| 99 | bairroD | pequena | 120.0 | 139.4 | 97 | 3330 | 3207 | 109.9 |
| 100 | bairroC | pequena | 145.0 | 241.5 | 42 | 1080 | 3061 | 104.1 |

# Coeficiente de Correlação de Pearson

Para calcular o Coeficiente de Correlação de Pearson, armazenamos e ordenamos os valores das variáveis em estudo:

# Questão 2:  
# Separamos as variáveis de idade e taxas pagas  
propertiesAges = zCityData$id  
propertiesTaxes = zCityData$tax  
  
# Realizamos ordenação dos dados  
sort(propertiesAges)

## [1] 2 2 4 4 4 4 5 5 6 6 7 7 9 9 10 10 10 10  
## [19] 11 11 11 11 12 12 12 12 13 13 13 13 15 15 15 15 16 16  
## [37] 16 16 18 18 18 18 22 22 22 22 22 22 24 24 25 25 25 25  
## [55] 26 26 27 27 30 30 31 31 33 33 33 33 33 33 33 33 35 35  
## [73] 37 37 37 37 41 41 42 42 43 43 43 43 44 44 46 46 50 50  
## [91] 56 56 60 60 65 65 97 97 202 202

sort(propertiesTaxes)

## [1] 1445 1447 1471 1473 2054 2056 2087 2089 2220 2222 2438 2440  
## [13] 2783 2785 2818 2820 2950 2952 3028 3030 3055 3057 3059 3061  
## [25] 3103 3105 3119 3121 3131 3133 3167 3169 3204 3205 3206 3207  
## [37] 3268 3270 3334 3336 3346 3348 3361 3363 3363 3365 3374 3376  
## [49] 3482 3484 3498 3500 3529 3531 3561 3563 3564 3566 3607 3609  
## [61] 3613 3615 3648 3648 3650 3650 3765 3767 3785 3787 3831 3833  
## [73] 3917 3919 3920 3922 3950 3952 3982 3984 4033 4035 4127 4129  
## [85] 4162 4164 4626 4628 4681 4683 5158 5160 5702 5704 5853 5855  
## [97] 7088 7090 12192 12194

E então calculamos o Coeficiente:

# Calculamos o coeficiente de correlação linear de Pearson  
correlationCoefficient = cor(propertiesAges, propertiesTaxes, method = "pearson")  
correlationCoefficient

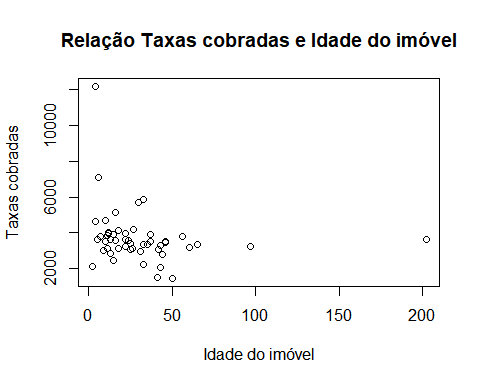
## [1] -0.1892211

Por ser um valor negativo muito próximo de zero, podemos inferir que as duas variáveis não tem nenhuma relação entre si!

# Diagrama de Dispersão

Com as variáveis armazenadas e ordenadas, e o coeficiente de correlação calculado, podemos gerar o *Diagrama de Dispersão*:

# Geramos um gráfico de dispersão!  
plot(propertiesAges, propertiesTaxes, ylab = "Taxas cobradas", xlab = "Idade do imóvel", main = "Relação Taxas cobradas e Idade do imóvel")



O Diagrama nos confirma o que foi “dito” pelo Coeficiente de Correlação, mostrando, enfim, que não há correlação entre as duas variáveis.

# Conclusão

Conclui-se que não há nenhuma relação entre as duas variáveis.