

Faça como eu fiz: Organizando dados na planilha

Para a análise de dados, organizamos os dados do questionário para facilitar a análise.

Começando com os títulos das colunas. A análise usando um título com separação por underline ou traço, sem espaços e sem acentos, ajuda para referência em linguagens como Python ou R ou softwares de análise estatística.

Para dados com mais de uma opção de seleção, é interessante que sejam separados em colunas diferentes, como o objetivo.

idade	objetivo_1	objetivo_2	objetivo_3
Menos de 18 anos: Sim	Passatempo	Quero aprender teorias musicais	Quero aprender teorias musicais
Menos de 18 anos: Sim	Passatempo	Desejo ter um grupo ou banda	Conhecer um novo estilo musical
Menos de 18 anos: Sim	Passatempo	Quero aprender teorias musicais	Conhecer um novo estilo musical
Menos de 18 anos: Sim	Passatempo	Quero melhorar minhas habilidades musicais	
Menos de 18 anos: Sim	Tocar/cantar na igreja	Quero melhorar minhas habilidades musicais	

Após o ajuste nos títulos das colunas, nós organizamos os dados em números para facilitar a análise por diferentes ferramentas ou métodos. Dados como 'Sim' e 'Não' foram colocados como valores de 1 para Sim e 0 para Não. Dados categóricos foram escolhidos a partir do valor 2. Os dados relacionados às notas não foram mexidos, pois já estão no formato adequado.

Menos de 18 anos -> 40	Passatempo -> 2
18 - 24 anos -> 41	Tocar/cantar na igreja -> 3
25 - 34 anos -> 42	Desejo ter um grupo ou banda -> 4
35 - 44 anos -> 43	Quero melhorar minhas habilidades musicais -> 5
45 - 54 anos -> 44	Fazer novas amizades -> 6
55 - 64 anos -> 45	Desejo dar aula um dia -> 7
65 anos ou mais -> 46	Quero aprender teorias musicais -> 8
	Conhecer um novo estilo musical -> 9

idade	objetivo_1	objetivo_2	objetivo_3	frequencia_aulas	tempo_trein
40	1	2	8	11	54
40	1	2	4	9	12
40	1	2	8	9	11
40	1	2	5	-1	12
40	1	3	5	-1	11

Após a aplicação do algoritmo K-Means nos dados organizados, podemos adicionar uma coluna com o resultado dos grupos de dados.

cluster	Qual a sua idade?	Você assina o plano anual ou semestral?
0	Menos de 18 anos	Anual
4	Menos de 18 anos	Anual
0	Menos de 18 anos	Semestral
4	Menos de 18 anos	Anual
4	Menos de 18 anos	Anual
4	Menos de 18 anos	Semestral