

Marcos Geraldo Braga Emiliano

19.1.4012

Preços de diamantes

O que é o conjunto de dados de preços de diamantes?

Este documento explora um conjunto de dados contendo preços e atributos para aproximadamente 54.000 diamantes de lapidação redonda.

Existem 53.940 diamantes no conjunto de dados com 10 recursos (quilate, corte, cor, clareza, profundidade, tabela, preço, x, y e z). A maioria das variáveis são de natureza numérica, mas as variáveis corte, cor e clareza são variáveis de fator ordenadas com os seguintes níveis.

Sendo que o objetivo deste trabalho seria desenvolver um algoritmo que teria, com certa precisão, a capacidade de prever o preço de um diamante tendo em vista suas características físicas, com a motivação de poder ser usada para detectar transações incomuns no mercado de forma simples, analisando possíveis fraudes.

Onde as variáveis são: quilate (float64), corte (object), cor (object), clareza (object), profundidade (float64), tabela (float64), preço (int64), x (float64), y (float64) e z (float64).