

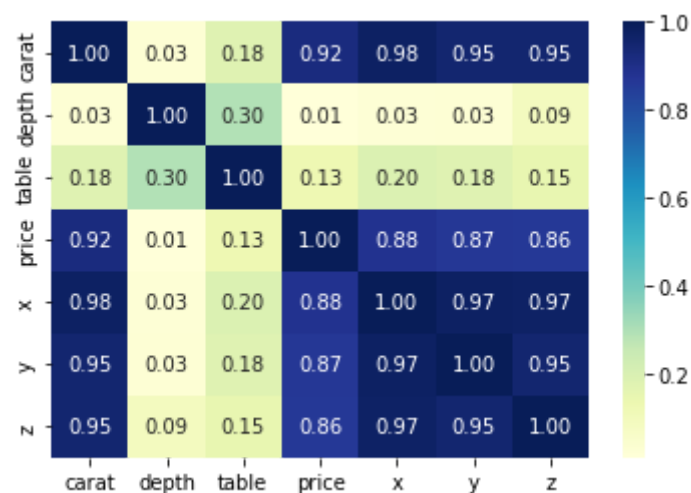
Actividad 1

Análisis de correlación utilizando plataformas de analítica

Elena Margarita Nivón Hernández - A0117466

Dentro del análisis realizado para esta actividad se realizó el tratado de algunos valores nulos, donde cada variable fue sustituida según su origen o contexto, por lo que cada una fue considerada de diferentes maneras. Posteriormente, se realizó el trato de outliers, donde se utilizó el método de varianza para la identificación de estos mismos.

A grandes rasgos, las correlaciones de las variables se ven reflejadas en el siguiente gráfico de color:



La siguiente tabla muestra la comparación sobre los coeficientes de correlación y determinación, de esta manera conocer la mayor correlación entre las variables analizadas.

Modelo	Coeficiente correlación		Coeficiente determinación	
price - carat	0.9234	92%	0.8527	85%
price - depth	0.0130	1%	0.00017	0%
price - table	0.1272	12%	0.0161	1%
price - carat - depth	0.9224	92%	0.851	85%

price - carat - table	0.9224	92%	0.851	85%
price - depth - table	0.1303	13%	0.017	1%
price - tamaño	0.8848	88%	0.783	78%
price - tamaño - carat - depth	0.9252	92%	0.856	85%
price - tamaño - price - depth - table	0.9268	92%	0.859	85%

Después de realizar el análisis correspondiente, se puede determinar que al relacionar las variables price, tamaño, carat y depth, así como price, tamaño, carat, depth, table , la correlación se mantiene de manera positiva y con un valor fuerte, pues el porcentaje obtenido muestra cercanía al 100%.

Finalmente, se puede decir que existen variables que logran estar correlacionadas con otras, consiguiendo un análisis extenso de lo que ocurre dentro del precio y movimientos en una industria tan volátil, como lo es la de los diamantes ; Sin embargo la correlación no significa causalidad.