

Dentro de este entregable, se realizó un análisis de correlación entre variables pertenecientes al data frame asignado por el profesor, el cual, resultó ser: “diamonds”, mismo que contenía información sobre atributos destacables de distintos tipos de diamantes, tales como: precio, tamaño, quilate, profundidad, corte, claridad, etectera; cabe destacar que dicho análisis arrojó información relevante sobre el impacto que ciertas variables tienen entorno al precio, la cual, se presenta a continuación:

VARIABLES	COEFICIENTE DE CORRELACIÓN	COEFICIENTE DE DETERMINACIÓN (r^2)	NIVEL DE CORRELACIÓN
Carat.	0.723736317781156	0.522729	ALTO
Depth.	5.5834952963351725e-06	3.11364e-13	BAJO
Table.	0.01640691694328733	0.00026896	BAJO
Carat y Depth.	0.8508818954473059	0.724	ALTO
Carat y Table.	0.85146931829632	0.725	ALTO
Depth y Table.	0.1341640786499874	0.018	BAJO
Tamaño (x,y,z).	0.8449852069711044	0.714	ALTO
Tamaño, Carat y Depth.	0.8573214099741123	0.735	ALTO
Tamaño, Carat, Depth y Table.	0.8579044235810886	0.736	ALTO

Teniendo en cuenta lo expuesto dentro de la tabla anexa, hemos de decir que, únicamente 3 variables presentan una correlación baja con respecto al precio, las cuales son: depth, table y las anteriores en conjunto. Por otra parte, podemos destacar que, si bien, el mejor modelo para este caso de estudio, de acuerdo a los valores de correlación y determinación obtenidos, es el último, el cual considera las variables anteriormente mencionadas, mismas que no generan un impacto significativo por sí solas, podemos afirmar que, cuando se juntan cada una con la variable Carat, se genera un modelo de regresión lineal igualmente favorecedor, destacando Carat y table por pocos puntos decimales en comparación con Carat y Depth, puesto que explica el 84% de las variaciones de Y a partir de X, donde la primera incógnita es el precio y la segunda, Carat y Table, es por ello que, se recomienda utilizar este último modelo para la predicción del precio, teniendo en cuenta el principio de parsimonia, el cual establece que, en igualdad de condiciones, la solución más sencilla suele ser la mejor.