

Análisis de regresión múltiple: Credicel

El siguiente código realiza un estudio de regresión lineal, en el que nos ayuda a generar un sistema de predicción para una variable dependiente en base a varias independientes.

Utilizaremos la base de datos de Credicel, con las variables numéricas relacionadas con el valor monetario.

Las variables por utilizar tienen correlaciones específicas que nos hacen hacer un estudio anticipado de su comportamiento dentro del estudio, estas son:

- Enganche: Con una correlación alta con varias variables del modelo, enganche es una de las opciones por analizar como variable dependiente dentro de la regresión múltiple.
- Precio: Con una correlación alta con varias variables del modelo, precio es una de las opciones por analizar como variable dependiente dentro de la regresión múltiple.
- Costo total: Con una correlación alta con varias variables del modelo, Costo total es una de las opciones por analizar como variable dependiente dentro de la regresión múltiple.
- Pagos realizados: Para el caso de pagos realizados, se nota una baja relación con casi todas las variables excepto una, en este caso es recomendable un estudio de regresión simple.
- Riesgo: Para el caso de riesgo, se nota una baja relación con casi todas las variables excepto una, en este caso es recomendable un estudio de regresión simple.
- Plazo: Para el caso de plazo, se nota una baja relación con casi todas las variables excepto una, en este caso es recomendable un estudio de regresión simple.
- Monto financiado: Con una correlación alta con varias variables del modelo, Monto financiado es una de las opciones por analizar como variable dependiente dentro de la regresión múltiple.
- Porcentaje de enganche: Para el caso de porcentaje de enganche, se nota una baja relación con casi todas las variables excepto una, en este caso es recomendable un estudio de regresión simple.

Iniciando este análisis se prueban las variables que son opción como dependientes y se generan conclusiones acerca de estas, una de las más sobresalientes es donde tomamos precio como variable dependiente, ya que tiene 0.99 en el valor del coeficiente de determinación, siendo el modelo más efectivo para la predicción de los valores futuros, y este está formado por todas las variables descritas como independientes excepto por precio, ya que funciona como variable dependiente.