Comentarios iniciales

1. Reescribir la tabla de forma tal que sea útil para el algoritmo apriori, teniendo en cuenta los renglones, se puede usar la columna venta para hacerlo.

2. Prod\_ID, cuidado con el formato de la columna.

3. ¿hay ruido en el product ID?

4. Propuestas:

1. Hacer las reglas sobre el ID del producto. (No lo veo tan conveniente)
2. Concatenar DescGen, DescAdic y Marca para hacer las reglas, creo que al no saber del negocio es lo más conveniente para determinar las reglas que son “interesantes”.

5. No veo negativos en el precio (Ya se por qué)

6. Los de precio 0 son solamente de las marcas KUNNAN, NATWAY, PENN y WATERDOG

7. Siguiendo la pregunta de Nicolás sobre el ID de las ventas: Falta del 6061 al 7300, del 8112 al 10160

8. Tampoco veo negativos para la columna renglón (Ya se por qué)

9. Cuando UM2 es negativo, UM1 también, ¿qué significa esto? Podría ser parte del ruido o como dice Nicolás podría ser parte de ventas canceladas.

10. Hay que categorizar los precios.