**Algoritmo Csv**

**Primer paso:**

Importar las librerías que utilizaremos para limpiar los datos ("import panda as pd").

**Segundo paso:**

Cargar los datos o archivos.csv al dataframe “df = pd.read\_csv” (‘dbo.DIM\_Customer’)

(dbo.DIM\_Customer,dbo.DIM\_Date,dbo.DIM\_Orders,dbo.DIM\_Product, dbo.DIM\_SalesPerson, dbo.DIM\_Supplier , dbo.DIM\_Sales.)

**Tercer paso:**

Explorar los datos y analizar los datos para identificar los problemas :

“df.head”

“df.describe”

“data.info”

**Cuarto paso:**

Identificar valores ausentes y así mismo eliminar columnas y filas con datos null

“data.dropna(inplace=True)

“data.info()”

**Quinto paso:**

Empezar la limpieza de los datos según los problemas identificados.

**sexto paso:**

Eliminar los datos duplicados o columnas duplicadas

**Séptimo paso:**

Guardar los nuevos csv que se han limpiado