Para missing values los representa entre ? 0 N/A o NaN

Dropna()

Axis 0, filas, Axis= 1 para columnas

Inplace = true hace las modificaciones en el mismo DF

Para reemplazarlos

.replace(Missing value, new value)

Para dar Formato a la información

Df[“columna”] = operación con mismo df DF[“columna”]

Df.rename(columns={“tituloant”}:{titulonvo}, inplace = true)

Para cambio de tipo

Df[“columna”] = df[“columna”].astype(“int”)

Normalizacion

Para que todos los datos tengan la misma importancia

Para scalaar

Df[col] = df[col] / df[col].max()

Minmax

Df[col] = (df-df-min()) / (df.max-df.min())

Zscore

Df = (df-df.mean())/df.std()

Bining (agrupar)

Bins = np.linspace(min(df),max(df),4) para 3 segmentos

Groupnames = [“1”,”2”,”3”]

Df[“pricebinned”] = pd.cut(df[price).Bins\_labels=Gruponames, include\_lowets=true

Categorical a cuantitativas

Convertir tipos en una columna a columnas separadas con 1unos y ceros

Pd.get\_dummies(df[col])