Data Preprocessing Checkpoint

1. import pandas as pd

file = pd.read\_csv(r'C:\Users\belve\Downloads\STEG\_BILLING\_HISTORY.csv')

client\_0\_bills = file.iloc[0:10]

client\_0\_bills

1. type(client\_0\_bills)
2. file.describe()
3. num\_rows = file.shape[0]

print(f'Nombre de ligne : {num\_rows}')

num\_columns = file.shape[1]

print(f'Nombre de colonnes : {num\_columns}')

1. num\_categorical\_features = len(file.select\_dtypes(include = 'object').columns)

print('Nombre de caractéristiques catégorielle :', num\_categorical\_features)

1. file.info()
2. num\_missing\_values = file.isnull().sum()

print(f'Nombre de valeure manquante par colonne : \n{num\_missing\_values}')

1. file.select\_dtypes(include = 'number').describe()
2. train\_Client\_0 = file.loc[file['client\_id'] == 'train\_Client\_0']

print(train\_Client\_0)

1. from sklearn.preprocessing import LabelEncoder

encoder = LabelEncoder()

file['counter\_type'] = encoder.fit\_transform(file['counter\_type'])

file

1. file.drop('counter\_statue', axis=1)