

```
len(df)
```

```
22092
```

Step 3: Data Preprocessing

```
df.columns
```

```
Index(['Code', 'Name', 'Substance', 'Unit',
       'Supply Chain Emission Factors without Margins',
       'Margins of Supply Chain Emission Factors',
       'Supply Chain Emission Factors with Margins', 'Unnamed: 7',
       'DQ ReliabilityScore of Factors without Margins',
       'DQ TemporalCorrelation of Factors without Margins',
       'DQ GeographicalCorrelation of Factors without Margins',
       'DQ TechnologicalCorrelation of Factors without Margins',
       'DQ DataCollection of Factors without Margins', 'Source', 'Year'],
      dtype='object')
```

```
df.isnull().sum()
```

	0
Code	0
Name	0
Substance	0
Unit	0
Supply Chain Emission Factors without Margins	0
Margins of Supply Chain Emission Factors	0
Supply Chain Emission Factors with Margins	0
Unnamed: 7	22092
DQ ReliabilityScore of Factors without Margins	0
DQ TemporalCorrelation of Factors without Margins	0
DQ GeographicalCorrelation of Factors without Margins	0
DQ TechnologicalCorrelation of Factors without Margins	0
DQ DataCollection of Factors without Margins	0
Source	0
Year	0

```
dtype: int64
```