

import pandas as pd import seaborn as sns

import matplotlib.pyplot as plt

data=pd.read\_csv(“greenhouse\_gas\_inventory\_data\_data.csv“) df=pd.DataFrame(data)

df

country\_or\_area year

**vaLue**

**category**

0 Australia 2014 393126.946994

1 Australia 2013 396913.936530

2 Australia 2012 406462.847704

3 Australia 2011 403705.528314

4 Australia 2010 406200.993184

carbon dioxide co2 emissions without land use ... carbon dioxide co2 emissions without land use ... carbon dioxide co2 emissions without land use ... carbon\_dioxide\_co2\_emissions\_without\_land\_use\_... carbon dioxide co2 emissions without land use ...

84‹1 United States of

|  |  |  |
| --- | --- | --- |
| 1994 | 593.222570 | unspecified mix of hydrofluorocarbons hfcs and... |
| 1993 | 586.939752 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |
| 1992 | 574.414624 | unspecified mix of hydrofluorocarbons hfcs and... |
| 1991 | 566.900968 | unsoecified mix of hvdrofluorocarbons hfcs and... |

America

84‹2 United States of

America

United States of

84‹3

America yyyy United States of

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| df.isnull() | country\_or\_area | yean | value | category |
| 0 | False | False | False | False |
| 1 | False | False | False | False |
| 2 | False | False | False | False |
| 3 | False | False | False | False |
|  | False | False | False | False |
| 8401 | False | False | False | False |
| 8402 | False | False | False | False |
| 8403 | False | False | False | False |
| 8404 | False | False | False | False |
| 8405 | False | False | False | False |

8406 rows 4 columns

df.describe()

|  |  |  |
| --- | --- | --- |
|  | **year** | **va1ue** |
| count | 8406.000000 | 8.406000e+03 |
| mean | 2002.188437 | 2.055472e+05 |
| std | 7.151605 | 8.041504e+05 |
| min | 1990.000000 | 1.103750e-06 |
| 25°/» | 1996.000000 | 2.976980e+02 |
| 50% | 2002.000000 | 7.840318e+03 |
| 75°/» | 2008.000000 | 6.589885e+04 |
| max | 2014.000000 | 7.422208e+06 |

df['country\_or\_area']=df['country\_or\_area'].map({'Australia' :0, 'United States of America' :1}).astype(float) df

country\_or\_area year

value

category

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 | 0.0 | 2014 | 393126.946994 | carbon dioxide co2 emissions without land use ... |
| 1 | 0.0 | 2013 | 396913.936530 | carbon\_dioxide\_co2\_emissions\_without\_land\_use\_... |
|  | 0.0 | 2012 | 406462.847704 | carbon dioxide co2 emissions without land use ... |
| 3 | 0.0 | 2011 | 403705.528314 | carbon dioxide co2 emissions without land use ... |
| 4 | 0.0 | 2010 | 406200.993184 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| **8401** | 1.0 | 1994 | 593.222570 | unspecified mix of hydrofluorocarbons hfcs and... |
| **8402** | 1.0 | 1993 | 586.939752 | unspecified mix of hydrofluorocarbons hfcs and... |
| 8403 | 1.0 | 1992 | 574.414624 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |
| **8404** | 1.0 | 1991 | 566.900968 | unspecified mix of hydrofluorocarbons hfcs and... |
| 8405  R4f)fi rnws x 4 columns | 1.0 | 1990 | 293.205663 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |

df=df[df['value']›1000] df

country\_or\_area year

value

category

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 | 0.0 | 2014 | 393126.946994 | carbon dioxide co2 emissions without land use ... |
| 1 | 0.0 | 2013 | 396913.936530 | carbon dioxide co2 emissions without land use ... |
|  | 0.0 | 2012 | 406462.847704 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 3 | 0.0 | 2011 | 403705.528314 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 4 | 0.0 | 2010 | 406200.993184 | carbon dioxide co2 emissions without land use ... |
| 8396 | 1.0 | 1999 | 4597.097543 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |
| **8397** | 1.0 | 1998 | 4470.729386 | unspecified mix of hydrofluorocarbons hfcs and... |
| 8398 | 1.0 | 1997 | 3904.524176 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |
| **8399** | 1.0 | 1996 | 3236.982773 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |
| **8400**  '5fi9fi rnws x 4 rnI‹imn\* | 1.0 | 1995 | 1764.731336 | unspecified\_mix\_of\_hydrofIuorocarbons\_hfcs\_and... |

non=df[-df.duplicated('category')] non

0

**1074**

**2023**

**3097**

**4072**

**5311**

**5394**

6477

**7341**

8336

df.head()

country\_or\_area year

0.0 2014

0.0 2014

0.0 2014

0.0 2014

0.0 2014

NaN 2013

0.0 2014

0.0 2005

NaN 1997

NaN 2009

value 393126.946994

522397.090711

522397.090711

10787.350138

98076.109401

1360.957366

20084.542834

1791.701120

1086.396974

1110.088250

category carbon dioxide co2 emissions without land use ... greenhouse\_gas\_ghgs\_emissions\_incIuding\_indire... greenhouse gas ghgs emissions without land use... hydrofluorocarbons\_hfcs\_emissions\_in\_kilotonne... methane\_ch4\_emissions\_without\_Iand\_use\_Iand\_us... nitrogen trifluoride nf3 emissions in kilotonn... nitrous oxide n2o emissions without land use 1... perfIuorocarbons\_pfcs\_emissions\_in\_kiIotonne\_c... sulphur hexafluoride sf6 emissions in kilotonn... unspecified mix of hydrofluorocarbons hfcs and...

|  |  |  |  |
| --- | --- | --- | --- |
| country\_or\_area | year | **value** | **category** |
| 0.0 | 2014 | 393126.946994 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 1 0.0 | 2013 | 396913.936530 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 2 0.0 | 2012 | 406462.847704 | carbon dioxide co2 emissions without land use ... |
| 3 0.0 | 2011 | 403705.528314 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 0.0 | 2010 | 406200.993184 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |

df['value']=pd.factorize(df['value'])[0] df

<ipython-input-26-eeb2e9bcd642>:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[now\_indexen,col\_indexen] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user guide/indexing.html#returning-a-view-versus-a-cc df['value']=pd.factorize(df['value'])[0]

country\_or\_area year value category

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 | 0.0 | 2014 | 0 | carbon dioxide co2 emissions without land use ... |
| 1 | 0.0 | 2013 | 1 | carbon dioxide co2 emissions without land use ... |
| 2 | 0.0 | 2012 | 2 | carbon dioxide co2 emissions without land use ... |
| 3 | 0.0 | 2011 | 3 | carbon\_dioxide\_co2\_emissions\_without\_Iand\_use\_... |
| 4 | 0.0 | 2010 | 4 | carbon dioxide co2 emissions without land use ... |
| 8396 | 1.0 | 1999 | 5025 | unspecified\_mix\_of\_hydrofluorocarbons\_hfcs\_and... |
| **8397** | 1.0 | 1998 | 5026 | unspecified mix of hydrofluorocarbons hfcs and... |
| 8398 | 1.0 | 1997 | 5027 | unspecified mix of hydrofluorocarbons hfcs and... |
| 8399 | 1.0 | 1996 | 5028 | unspecified\_mix\_of\_hydrofluorocarbons\_hfcs\_and... |
| **8400**  5696 rows • 4 columns | 1.0 | 1995 | 5029 | unspecified\_mix\_of\_hydrofluorocarbons\_hfcs\_and... |
|  |  |  |  |  |

df=df.drop(['category'],axis=1) df

country\_or\_area year value

|  |  |  |  |
| --- | --- | --- | --- |
| 0 | 0.0 | 2014 | 0 |
| 1 | 0.0 | 2013 | 1 |
| 2 | 0.0 | 2012 | 2 |
| 3 | 0.0 | 2011 | 3 |
|  | 0.0 | 2010 | 4 |
| 8396 | 1.0 | 1999 | 5025 |
| **8397** | 1.0 | 1998 | 5026 |
| **8398** | 1.0 | 1997 | 5027 |
| **8399** | 1.0 | 1996 | 5028 |
| 8400  5696 rows • 3 columns | 1.0 | 1995 | 5029 |

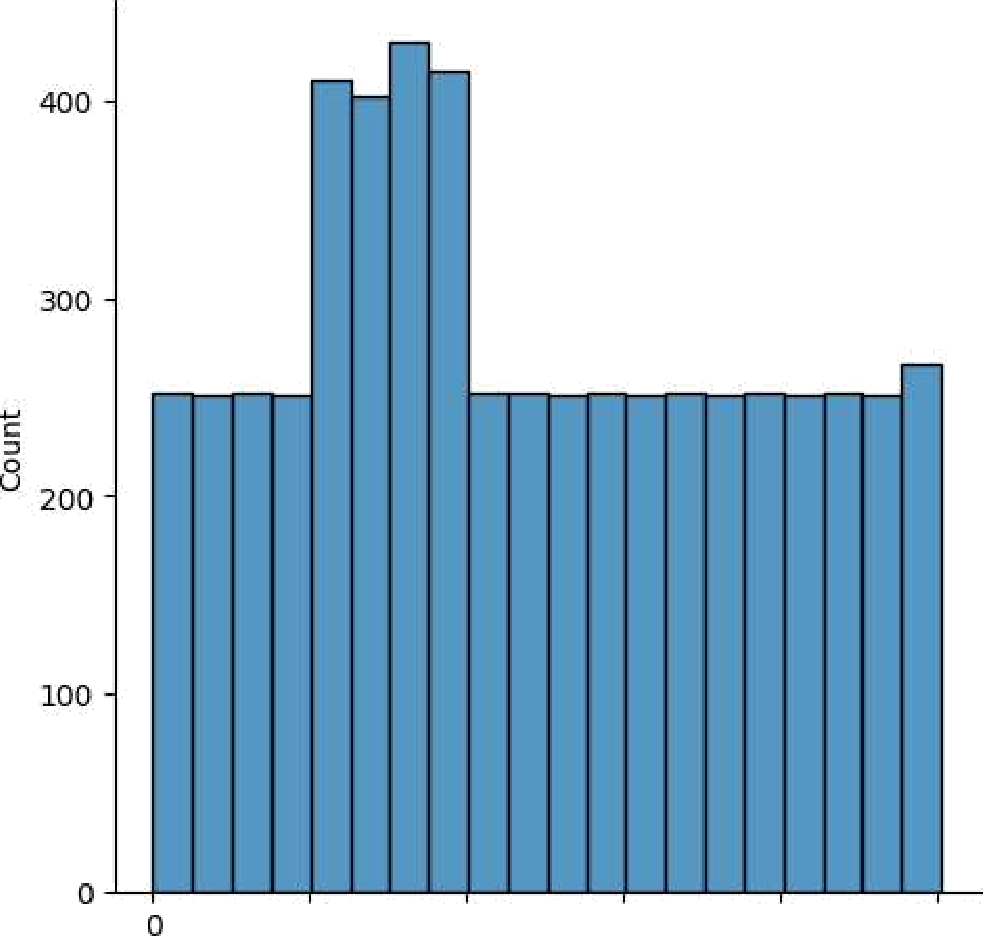
sns.displot(df['value']) plt.show

matplotlib.pyplot.show def show(’args, \*\*kwargs)

Display all open figures.

Parameters

block : bool, optional

Whether to wait for all figures to be closed before returning.

IDD0 2 DDO 3000 4D0 0 5000

value