

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [2]: df = pd.read_csv('Temperature_change_Data.csv')
df.head()
```

	Country Code	Country Name	year	tem_change
0	AFG	Afghanistan	1961	-0.080
1	ALB	Albania	1961	0.631
2	DZA	Algeria	1961	0.186
3	ASM	American Samoa	1961	-0.014
4	AND	Andorra	1961	0.749

```
In [3]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16756 entries, 0 to 16755
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   Country Code    14042 non-null object  
1   Country Name    16756 non-null object  
2   year            16756 non-null int64   
3   tem_change      14975 non-null float64  
dtypes: float64(1), int64(1), object(2)
memory usage: 523.8+ KB
```

```
In [4]: #it makes a quick report about data
import pandas_profiling as pdp
report=pdp.ProfileReport(df)
report

Summarize dataset:   0%|          | 0/5 [00:00<?, ?it/s]

Generate report structure:   0%|          | 0/1 [00:00<?, ?it/s]

Render HTML:   0%|          | 0/1 [00:00<?, ?it/s]
```

```
In [5]: report.to_file("temp _report")

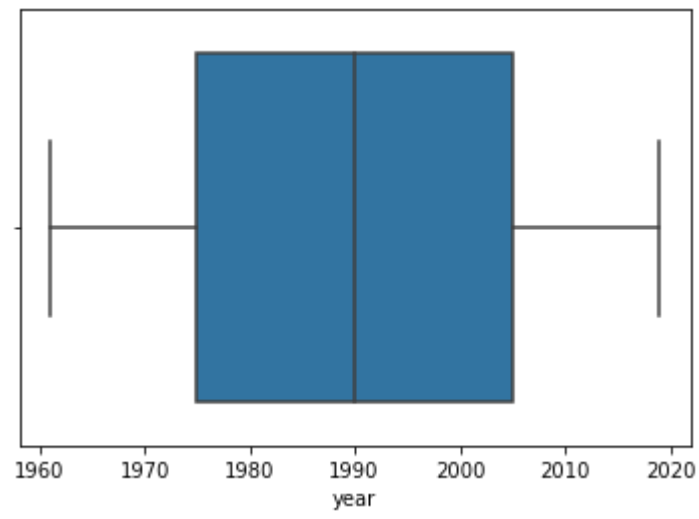
C:\Users\Mcs\anaconda3\lib\site-packages\pandas_profiling\profile_report.py:262: UserWarning: Extension  not supported. For now we assume
.html was intended. To remove this warning, please use .html or .json.
  warnings.warn(

Export report to file:   0%|          | 0/1 [00:00<?, ?it/s]
```

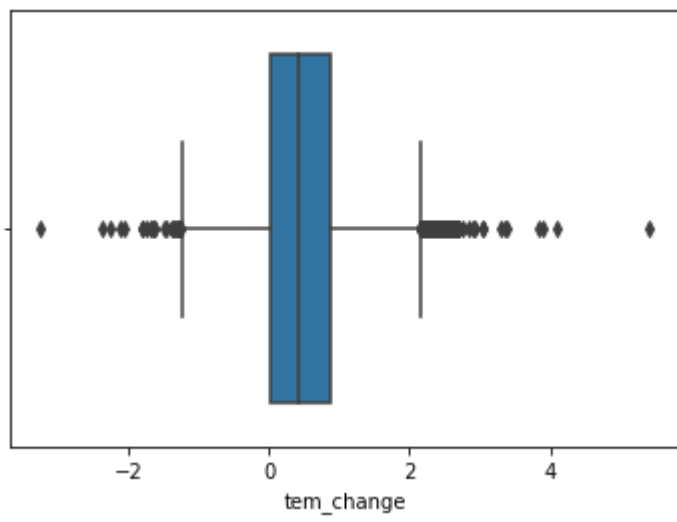


```
In [6]: for i , coltype in df.dtypes.iteritems():
        if coltype != object:
            print(sns.boxplot(x=df[i]))
            plt.show()
```

AxesSubplot(0.125,0.125;0.775x0.755)



AxesSubplot(0.125,0.125;0.775x0.755)



```
In [9]: for i , coltype in df.dtypes.iteritems():
        print(df[i].value_counts())
        print("_ "*10)
```

```
AFG      59
PRY      59
NER      59
NGA      59
NIU      59
..
GIB      59
GRC      59
GRL      59
GRD      59
X21      59
Name: Country Code, Length: 238, dtype: int64
```

```
-- -- -- -- --
Afghanistan      59
Saint Pierre and Miquelon  59
Senegal          59
Saudi Arabia     59
Sao Tome and Principe  59
..
Honduras         59
Hungary          59
Iceland          59
India            59
OECD             59
Name: Country Name, Length: 284, dtype: int64
```

```
-- -- -- -- --
1961      284
2005      284
1993      284
1994      284
1995      284
1996      284
1997      284
1998      284
1999      284
2000      284
2001      284
2002      284
2003      284
2004      284
2006      284
1991      284
2007      284
2008      284
2009      284
2010      284
2011      284
2012      284
2013      284
2014      284
2015      284
2016      284
2017      284
2018      284
1992      284
1990      284
1962      284
1975      284
1963      284
1964      284
1965      284
1966      284
1967      284
1968      284
1969      284
1970      284
1971      284
1972      284
1973      284
1974      284
1976      284
1989      284
1977      284
1978      284
1979      284
1980      284
1981      284
1982      284
1983      284
1984      284
1985      284
1986      284
1987      284
1988      284
2019      284
Name: year, dtype: int64
```

```
0.307    20
0.221    19
0.130    19
0.215    18
0.305    18
..
-0.974    1
0.876     1
1.375     1
-1.079     1
1.297     1
Name: tem_change, Length: 2855, dtype: int64
```

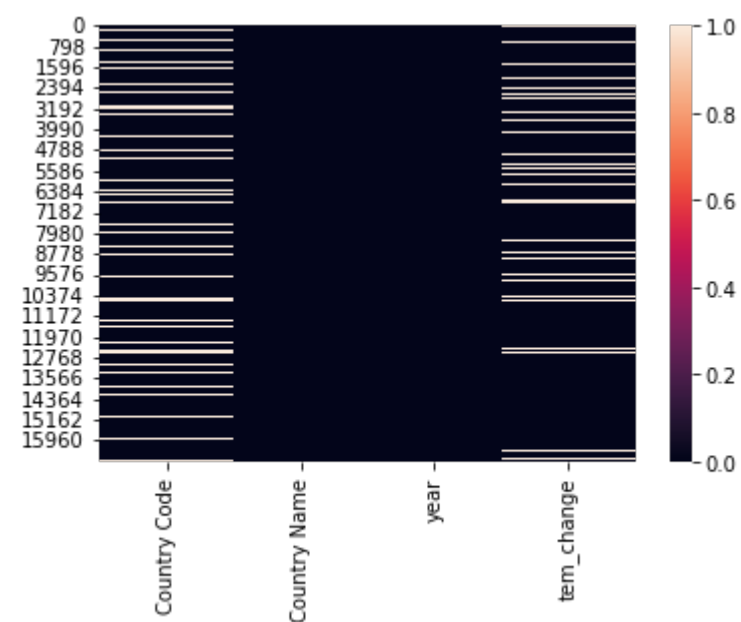
```
In [11]: df.isnull().sum()
```

```
Country Code    2714
Country Name      0
year             0
tem_change      1781
dtype: int64
```

```
In [ ]:
```

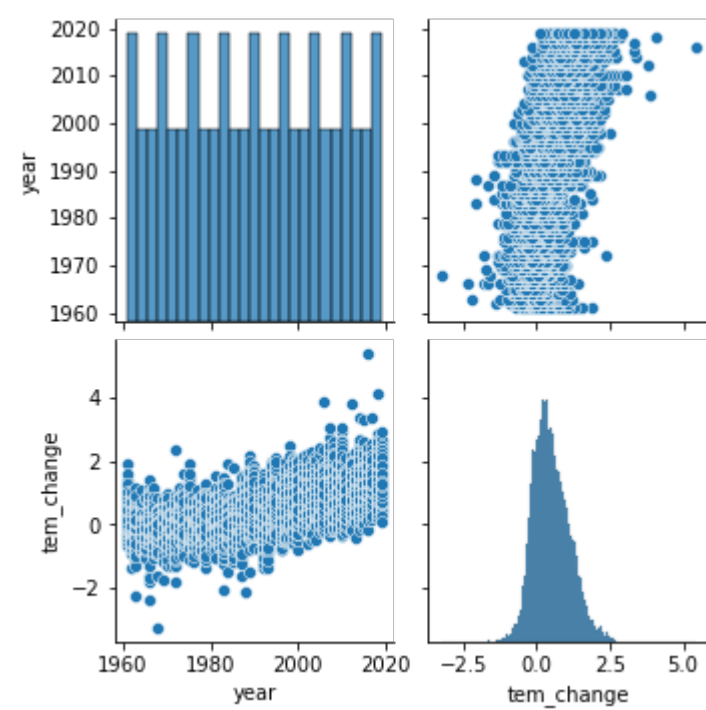
```
In [12]: sns.heatmap(df.isnull())
```

<AxesSubplot:>



```
In [15]: sns.pairplot(df)
```

<seaborn.axisgrid.PairGrid at 0x22ebf262f40>



In [18]:

<AxesSubplot:xlabel='year', ylabel='Count'>

C:\Users\Mcs\anaconda3\lib\site-packages\IPython\core\pylabtools.py:151: UserWarning: Creating legend with loc="best" can be slow with large amounts of data.

fig.canvas.print_figure(bytes_io, **kw)

In []:

In []:

In []:

In []:

In []:

In []:

In []: