

KITAVI DUNCAN GITAU

SCT211-0031/2021

```
import pandas as pd
data = pd.read_csv('climate.csv')
data

{"type": "dataframe", "variable_name": "data"}

data.head()

{"type": "dataframe", "variable_name": "data"}

data.tail(3)

{"type": "dataframe"}

data.sample(10)

{"type": "dataframe"}

data.columns
Index(['ObjectId', 'Country', 'IS02', 'IS03', 'Indicator', 'Unit',
      'Source',
      'CTS Code', 'CTS Name', 'CTS Full Descriptor', '1961', '1962',
      '1963',
      '1964', '1965', '1966', '1967', '1968', '1969', '1970', '1971',
      '1972',
      '1973', '1974', '1975', '1976', '1977', '1978', '1979', '1980',
      '1981',
      '1982', '1983', '1984', '1985', '1986', '1987', '1988', '1989',
      '1990',
      '1991', '1992', '1993', '1994', '1995', '1996', '1997', '1998',
      '1999',
      '2000', '2001', '2002', '2003', '2004', '2005', '2006', '2007',
      '2008',
      '2009', '2010', '2011', '2012', '2013', '2014', '2015', '2016',
      '2017',
      '2018', '2019', '2020', '2021', '2022'],
      dtype='object')

data.columns = ['ID', 'Country', 'IS02', 'IS03', 'Indicator', 'Unit',
               'Source',
               'CTS_Code', 'CTS_Name', 'CTS_Full_Descriptor', '1961', '1962',
               '1963', '1964', '1965', '1966', '1967', '1968', '1969', '1970',
               '1971', '1972', '1973', '1974', '1975', '1976', '1977', '1978',
               '1979', '1980', '1981', '1982', '1983', '1984', '1985', '1986',
               '1987', '1988', '1989', '1990', '1991', '1992', '1993', '1994',
               '1995', '1996', '1997', '1998', '1999', '2000', '2001', '2002',
```

```
'2003', '2004', '2005', '2006', '2007', '2008', '2009', '2010',
'2011', '2012', '2013', '2014', '2015', '2016', '2017', '2018',
'2019', '2020', '2021', '2022'
]
```

```
columns_selected = data.iloc[:, [1, 11, 16, 21, 26, 31, 36, 41, 46,
51, 56, 61, 66, 71]]
```

```
columns_selected
```

```
{"summary": "{\n  \"name\": \"columns_selected\",\n  \"rows\": 225,\n  \"fields\": [\n    {\n      \"column\": \"Country\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 225,\n        \"samples\": [\n          \"Armenia, Rep. of\",\n          \"Spain\",\n          \"Maldives\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1962\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.3418123874389091,\n        \"min\": -0.908,\n        \"max\": 0.998,\n        \"num_unique_values\": 169,\n        \"samples\": [\n          0.093,\n          -0.118,\n          -0.776\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1967\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.3394841957312319,\n        \"min\": -1.048,\n        \"max\": 1.134,\n        \"num_unique_values\": 160,\n        \"samples\": [\n          0.159,\n          0.032,\n          -1.028\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1972\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.3844737091788439,\n        \"min\": -1.796,\n        \"max\": 0.933,\n        \"num_unique_values\": 167,\n        \"samples\": [\n          -0.703,\n          -0.268,\n          0.304\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1977\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.25342808077817564,\n        \"min\": -0.599,\n        \"max\": 1.079,\n        \"num_unique_values\": 167,\n        \"samples\": [\n          0.039,\n          -0.022,\n          -0.279\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1982\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.31829855255580203,\n        \"min\": -0.682,\n        \"max\": 1.135,\n        \"num_unique_values\": 174,\n        \"samples\": [\n          0.709,\n          0.623,\n          0.238\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1987\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.47508396314884604,\n        \"min\": -1.652,\n        \"max\": 1.562,\n        \"num_unique_values\": 172,\n        \"samples\": [\n          0.368,\n
```

```

n          0.505,\n          0.964\n          ],\n\n\"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n\n          }, \n          {\n          \"column\": \"1992\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.5652350970020732, \n          \"min\": -1.344, \n          \"max\": 1.601, \n          \"num_unique_values\": 191, \n          \"samples\": [\n          0.875, \n          0.533, \n          0.807\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"1997\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.4830883203093068, \n          \"min\": -0.429, \n          \"max\": 1.933, \n          \"num_unique_values\": 196, \n          \"samples\": [\n          0.079, \n          1.321, \n          0.359\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"2002\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.37852904375599056, \n          \"min\": 0.009, \n          \"max\": 2.255, \n          \"num_unique_values\": 191, \n          \"samples\": [\n          1.135, \n          1.13, \n          1.376\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"2007\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.5463201550082332, \n          \"min\": -0.219, \n          \"max\": 2.729, \n          \"num_unique_values\": 200, \n          \"samples\": [\n          0.547, \n          1.08, \n          1.242\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"2012\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.44211269170531653, \n          \"min\": -0.128, \n          \"max\": 2.144, \n          \"num_unique_values\": 191, \n          \"samples\": [\n          0.533, \n          0.795, \n          0.508\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"2017\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.39399890441060664, \n          \"min\": 0.017, \n          \"max\": 2.493, \n          \"num_unique_values\": 194, \n          \"samples\": [\n          0.79, \n          1.124, \n          1.184\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          }, \n          {\n          \"column\": \"2022\", \n          \"properties\": {\n\n          \"dtype\": \"number\", \n          \"std\": 0.6692789604244717, \n          \"min\": -1.305, \n          \"max\": 3.243, \n          \"num_unique_values\": 206, \n          \"samples\": [\n          1.216, \n          1.707, \n          1.074\n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n          } \n          } \n          ] \n          } \n          ], \n          \"type\": \"dataframe\", \"variable_name\": \"columns_selected\"}

```

```
columns_selected.describe()
```

```

{\"summary\": \"{\n  \"name\": \"columns_selected\", \n  \"rows\": 8, \n  \"fields\": [\n    {\n      \"column\": \"1962\", \n

```

```

{"properties": {"dtype": "number", "std": 66.80791808693246, "min": -0.908, "max": 189.0, "num_unique_values": 8, "samples": [-0.01347619047619048, -0.056, 189.0]}, "semantic_type": "", "description": ""}, {"column": "1967", "properties": {"dtype": "number", "std": 67.53527830500084, "min": -1.048, "max": 191.0, "num_unique_values": 8, "samples": [-0.11083246073298428, -0.146, 191.0]}, "semantic_type": "", "description": ""}, {"column": "1972", "properties": {"dtype": "number", "std": 67.92137580160988, "min": -1.796, "max": 192.0, "num_unique_values": 8, "samples": [-0.08490625, -0.045, 192.0]}, "semantic_type": "", "description": ""}, {"column": "1977", "properties": {"dtype": "number", "std": 65.33820846769545, "min": -0.599, "max": 185.0, "num_unique_values": 8, "samples": [0.1658162162162162, 0.182, 185.0]}, "semantic_type": "", "description": ""}, {"column": "1982", "properties": {"dtype": "number", "std": 67.80739834083896, "min": -0.682, "max": 192.0, "num_unique_values": 8, "samples": [0.17649479166666668, 0.181, 192.0]}, "semantic_type": "", "description": ""}, {"column": "1987", "properties": {"dtype": "number", "std": 67.07211253384104, "min": -1.652, "max": 190.0, "num_unique_values": 8, "samples": [0.405021052631579, 0.491, 190.0]}, "semantic_type": "", "description": ""}, {"column": "1992", "properties": {"dtype": "number", "std": 73.4486350743146, "min": -1.344, "max": 208.0, "num_unique_values": 8, "samples": [0.2364903846153846, 0.2985, 208.0]}, "semantic_type": "", "description": ""}, {"column": "1997", "properties": {"dtype": "number", "std": 72.97889741647639, "min": -0.429, "max": 207.0, "num_unique_values": 8, "samples": [0.5439951690821256, 0.547, 207.0]}, "semantic_type": "", "description": ""}, {"column": "2002", "properties": {"dtype": "number", "std": 74.64205429898776, "min": -0.429, "max": 207.0, "num_unique_values": 8, "samples": [0.5439951690821256, 0.547, 207.0]}, "semantic_type": "", "description": ""}

```

```

{"min": 0.009, "max": 212.0, "num_unique_values": 8, "samples": [0.9249999999999998, 0.84, 212.0], "semantic_type": "", "description": "", "column": "2007", "properties": {"dtype": "number", "std": 76.3771649102675, "min": -0.219, "max": 217.0, "num_unique_values": 8, "samples": [1.0225483870967744, 0.921, 217.0], "semantic_type": "", "description": "", "column": "2012", "properties": {"dtype": "number", "std": 75.71665953709281, "min": -0.128, "max": 215.0, "num_unique_values": 8, "samples": [0.9022232558139536, 0.808, 215.0], "semantic_type": "", "description": "", "column": "2017", "properties": {"dtype": "number", "std": 75.2585459050908, "min": 0.017, "max": 214.0, "num_unique_values": 8, "samples": [1.2807850467289719, 1.282, 214.0], "semantic_type": "", "description": "", "column": "2022", "properties": {"dtype": "number", "std": 74.90859590103369, "min": -1.305, "max": 213.0, "num_unique_values": 8, "samples": [1.382112676056338, 1.315, 213.0], "semantic_type": "", "description": ""}}, "type": "dataframe"}

```

```

from pandas.api.types import is_numeric_dtype
for col in columns_selected:
    if is_numeric_dtype(columns_selected[col]):
        print('%s:'%(col))
        print('\t Mean = %.2f'% columns_selected[col].mean())
        print('\t Standard Deviation=%.2f'%columns_selected[col].std())
        print('\t Minimum Value =%.2f'%columns_selected[col].min())
        print('\t Maximum Value =%.2f'%columns_selected[col].max())

```

```

1962:
    Mean = -0.01
    Standard Deviation=0.34
    Minimum Value =-0.91
    Maximum Value =1.00
1967:
    Mean = -0.11
    Standard Deviation=0.34
    Minimum Value =-1.05
    Maximum Value =1.13
1972:

```

|       |                         |
|-------|-------------------------|
|       | Mean = -0.08            |
|       | Standard Deviation=0.38 |
|       | Minimum Value =-1.80    |
|       | Maximum Value =0.93     |
| 1977: |                         |
|       | Mean = 0.17             |
|       | Standard Deviation=0.25 |
|       | Minimum Value =-0.60    |
|       | Maximum Value =1.08     |
| 1982: |                         |
|       | Mean = 0.18             |
|       | Standard Deviation=0.32 |
|       | Minimum Value =-0.68    |
|       | Maximum Value =1.14     |
| 1987: |                         |
|       | Mean = 0.41             |
|       | Standard Deviation=0.48 |
|       | Minimum Value =-1.65    |
|       | Maximum Value =1.56     |
| 1992: |                         |
|       | Mean = 0.24             |
|       | Standard Deviation=0.57 |
|       | Minimum Value =-1.34    |
|       | Maximum Value =1.60     |
| 1997: |                         |
|       | Mean = 0.54             |
|       | Standard Deviation=0.48 |
|       | Minimum Value =-0.43    |
|       | Maximum Value =1.93     |
| 2002: |                         |
|       | Mean = 0.92             |
|       | Standard Deviation=0.38 |
|       | Minimum Value =0.01     |
|       | Maximum Value =2.25     |
| 2007: |                         |
|       | Mean = 1.02             |
|       | Standard Deviation=0.55 |
|       | Minimum Value =-0.22    |
|       | Maximum Value =2.73     |
| 2012: |                         |
|       | Mean = 0.90             |
|       | Standard Deviation=0.44 |
|       | Minimum Value =-0.13    |
|       | Maximum Value =2.14     |
| 2017: |                         |
|       | Mean = 1.28             |
|       | Standard Deviation=0.39 |
|       | Minimum Value =0.02     |
|       | Maximum Value =2.49     |

2022:

Mean = 1.38  
Standard Deviation=0.67  
Minimum Value =-1.30  
Maximum Value =3.24

data['1963'].value\_counts()

|        |   |
|--------|---|
| 0.169  | 3 |
| -0.211 | 3 |
| 0.330  | 3 |
| 0.288  | 3 |
| 0.168  | 2 |
| ..     |   |
| 0.308  | 1 |
| 0.309  | 1 |
| 0.157  | 1 |
| -0.188 | 1 |
| -0.390 | 1 |

Name: 1963, Length: 167, dtype: int64

data['Country'].value\_counts()

|                              |   |
|------------------------------|---|
| Afghanistan, Islamic Rep. of | 1 |
| Libya                        | 1 |
| New Zealand                  | 1 |
| Nicaragua                    | 1 |
| Niger                        | 1 |
| ..                           |   |
| Grenada                      | 1 |
| Guadeloupe                   | 1 |
| Guatemala                    | 1 |
| Guinea                       | 1 |
| Zimbabwe                     | 1 |

Name: Country, Length: 225, dtype: int64

columns\_selected.describe(include='all')

```
{"summary":{"name": "columns_selected", "rows": 11, "fields": [{"column": "Country", "properties": {"dtype": "category", "num_unique_values": 3, "samples": ["225", "Afghanistan, Islamic Rep. of", "1"]}], "semantic_type": "", "description": ""}, {"column": "1962", "properties": {"dtype": "number", "std": 66.80791808693246, "min": -0.908, "max": 189.0, "num_unique_values": 8, "samples": ["-0.01347619047619048", "-0.056", "189.0"]}], "semantic_type": "", "description": ""}, {"column": "1967", "properties":
```

```
{
  "dtype": "number",
  "std": 67.53527830500084,
  "min": -1.048,
  "max": 191.0,
  "num_unique_values": 8,
  "samples": [
    -0.11083246073298428,
    -0.146,
    191.0
  ],
  "semantic_type": "",
  "description": "",
  "column": "1972",
  "properties": {
    "dtype": "number",
    "std": 67.92137580160988,
    "min": -1.796,
    "max": 192.0,
    "num_unique_values": 8,
    "samples": [
      -0.08490625,
      -0.045,
      192.0
    ],
    "semantic_type": "",
    "description": "",
    "column": "1977",
    "properties": {
      "dtype": "number",
      "std": 65.33820846769545,
      "min": -0.599,
      "max": 185.0,
      "num_unique_values": 8,
      "samples": [
        0.1658162162162162,
        0.182,
        185.0
      ],
      "semantic_type": "",
      "description": "",
      "column": "1982",
      "properties": {
        "dtype": "number",
        "std": 67.80739834083896,
        "min": -0.682,
        "max": 192.0,
        "num_unique_values": 8,
        "samples": [
          0.17649479166666668,
          0.181,
          192.0
        ],
        "semantic_type": "",
        "description": "",
        "column": "1987",
        "properties": {
          "dtype": "number",
          "std": 67.07211253384104,
          "min": -1.652,
          "max": 190.0,
          "num_unique_values": 8,
          "samples": [
            0.405021052631579,
            0.491,
            190.0
          ],
          "semantic_type": "",
          "description": "",
          "column": "1992",
          "properties": {
            "dtype": "number",
            "std": 73.4486350743146,
            "min": -1.344,
            "max": 208.0,
            "num_unique_values": 8,
            "samples": [
              0.2364903846153846,
              0.2985,
              208.0
            ],
            "semantic_type": "",
            "description": "",
            "column": "1997",
            "properties": {
              "dtype": "number",
              "std": 72.97889741647639,
              "min": -0.429,
              "max": 207.0,
              "num_unique_values": 8,
              "samples": [
                0.5439951690821256,
                0.547,
                207.0
              ],
              "semantic_type": "",
              "description": "",
              "column": "2002",
              "properties": {
                "dtype": "number",
                "std": 74.64205429898776,
                "min": 0.009,
                "max": 212.0,
                "num_unique_values": 8,
                "samples": [
                  0.9249999999999998,
                  0.84,
                  212.0
                ],
                "semantic_type": "",
                "description": "",
                "column": "2007",
                "properties": {
                  "dtype": "number",
                  "std": 76.3771649102675,

```



```

{"min\\": -0.219,\\n      \\\"max\\\": 217.0,\\n
\\\"num_unique_values\\\": 8,\\n      \\\"samples\\\": [\\n
1.0225483870967744,\\n      0.921,\\n      217.0\\n      ],\\n
\\\"semantic_type\\\": \\\"\\\",\\n      \\\"description\\\": \\\"\\\"\\n      }\\n
n      },\\n      {\\n      \\\"column\\\": \\\"2012\\\",\\n      \\\"properties\\\": {\\n
\\\"dtype\\\": \\\"number\\\",\\n      \\\"std\\\": 75.71665953709281,\\n
\\\"min\\\": -0.128,\\n      \\\"max\\\": 215.0,\\n
\\\"num_unique_values\\\": 8,\\n      \\\"samples\\\": [\\n
0.9022232558139536,\\n      0.808,\\n      215.0\\n      ],\\n
\\\"semantic_type\\\": \\\"\\\",\\n      \\\"description\\\": \\\"\\\"\\n      }\\n
n      },\\n      {\\n      \\\"column\\\": \\\"2017\\\",\\n      \\\"properties\\\": {\\n
\\\"dtype\\\": \\\"number\\\",\\n      \\\"std\\\": 75.25854590509083,\\n
\\\"min\\\": 0.017,\\n      \\\"max\\\": 214.0,\\n
\\\"num_unique_values\\\": 8,\\n      \\\"samples\\\": [\\n
1.2807850467289719,\\n      1.282,\\n      214.0\\n      ],\\n
\\\"semantic_type\\\": \\\"\\\",\\n      \\\"description\\\": \\\"\\\"\\n      }\\n
n      },\\n      {\\n      \\\"column\\\": \\\"2022\\\",\\n      \\\"properties\\\": {\\n
\\\"dtype\\\": \\\"number\\\",\\n      \\\"std\\\": 74.90859590103369,\\n
\\\"min\\\": -1.305,\\n      \\\"max\\\": 213.0,\\n
\\\"num_unique_values\\\": 8,\\n      \\\"samples\\\": [\\n
1.382112676056338,\\n      1.315,\\n      213.0\\n      ],\\n
\\\"semantic_type\\\": \\\"\\\",\\n      \\\"description\\\": \\\"\\\"\\n      }\\n
n      }\\n      ]\\n      }\", \"type\": \"dataframe\"}

```

```

print('covariance')
columns_selected.cov()

```

covariance

<ipython-input-22-897f9c76ffdf>:2: FutureWarning: The default value of numeric\_only in DataFrame.cov is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric\_only to silence this warning.

```
columns_selected.cov()
```

```

{"summary": "{\\n  \\\"name\\\": \\\"columns_selected\\\",\\n  \\\"rows\\\": 13,\\n
\\\"fields\\\": [\\n    {\\n      \\\"column\\\": \\\"1962\\\",\\n
\\\"properties\\\": {\\n        \\\"dtype\\\": \\\"number\\\",\\n        \\\"std\\\":
0.05398062634638313,\\n        \\\"min\\\": -0.1026298231850117,\\n
\\\"max\\\": 0.11683570820668696,\\n        \\\"num_unique_values\\\": 13,\\n
\\\"samples\\\": [\\n          -0.01557000046554935,\\n          -
0.04402819897179786,\\n          0.11683570820668696\\n          ],\\n
\\\"semantic_type\\\": \\\"\\\",\\n        \\\"description\\\": \\\"\\\"\\n        }\\n
n      },\\n      {\\n      \\\"column\\\": \\\"1967\\\",\\n      \\\"properties\\\": {\\n
\\\"dtype\\\": \\\"number\\\",\\n      \\\"std\\\": 0.05426708734470303,\\n
\\\"min\\\": -0.06292953596757853,\\n      \\\"max\\\": 0.11524951915128133,\\n
n      \\\"num_unique_values\\\": 13,\\n      \\\"samples\\\": [\\n
0.03687327994657277,\\n      0.06983007230456262,\\n      -
0.06292953596757853\\n      ],\\n      \\\"semantic_type\\\": \\\"\\\",\\n
\\\"description\\\": \\\"\\\"\\n      }\\n      },\\n      {\\n      \\\"column\\\":

```

```
\n1972\", \n      \"properties\": { \n        \"dtype\": \"number\", \n        \"std\": 0.06020954071368436, \n        \"min\": -0.07287558865248227, \n        \"max\": 0.14782003304973818, \n        \"num_unique_values\": 13, \n        \"samples\": [ \n          -0.010008131982811525, \n          0.0286546531251797, \n          -0.07287558865248227 \n        ], \n        \"semantic_type\": \"\", \n        \"description\": \"\", \n        \"column\": \"1977\", \n        \"properties\": { \n          \"dtype\": \"number\", \n          \"std\": 0.021654488016240137, \n          \"min\": -0.017544460853858777, \n          \"max\": 0.06422579212690954, \n          \"num_unique_values\": 13, \n          \"samples\": [ \n            0.0027514761813045704, \n            0.010465799441340785, \n            0.006324622548722003 \n          ], \n          \"semantic_type\": \"\", \n          \"description\": \"\", \n          \"column\": \"1982\", \n          \"properties\": { \n            \"dtype\": \"number\", \n            \"std\": 0.04274222724876838, \n            \"min\": -0.03815610968255774, \n            \"max\": 0.10424059177409459, \n            \"num_unique_values\": 13, \n            \"samples\": [ \n              0.03815578440896119, \n              0.02747148318095567, \n              -0.03815610968255774 \n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\", \n            \"column\": \"1987\", \n            \"properties\": { \n              \"dtype\": \"number\", \n              \"std\": 0.08575082933616673, \n              \"min\": -0.114341266387337, \n              \"max\": 0.22570477204121409, \n              \"num_unique_values\": 13, \n              \"samples\": [ \n                -0.0030425963474524334, \n                0.08949663051438536, \n                0.014925239664194123 \n              ], \n              \"semantic_type\": \"\", \n              \"description\": \"\", \n              \"column\": \"1992\", \n              \"properties\": { \n                \"dtype\": \"number\", \n                \"std\": 0.09680380235408993, \n                \"min\": -0.1026298231850117, \n                \"max\": 0.31949071488294306, \n                \"num_unique_values\": 13, \n                \"samples\": [ \n                  0.050641779346733654, \n                  0.1417888160075776, \n                  -0.1026298231850117 \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\", \n                \"column\": \"1997\", \n                \"properties\": { \n                  \"dtype\": \"number\", \n                  \"std\": 0.06813318797379933, \n                  \"min\": -0.03664672576652066, \n                  \"max\": 0.23337432521926746, \n                  \"num_unique_values\": 13, \n                  \"samples\": [ \n                    0.10366619787202597, \n                    0.05456464624581541, \n                    -0.03664672576652066 \n                  ], \n                  \"semantic_type\": \"\", \n                  \"description\": \"\", \n                  \"column\": \"2002\", \n                  \"properties\": { \n                    \"dtype\": \"number\", \n                    \"std\": 0.064308862055709, \n                    \"min\": -0.028815683832264207, \n                    \"max\": 0.16606322704238705, \n                    \"num_unique_values\": 13, \n                    \"samples\": [ \n                      0.09511327458843391, \n                      0.15098181828029794, \n                      -0.016600601810926943 \n                    ], \n                    \"semantic_type\": \"\", \n                    \"description\": \"\", \n                    \"column\": \"2007\", \n                    \"properties\": { \n                      \"dtype\": \"number\", \n                      \"std\": 0.10981954347368977, \n                      \"min\": -0.08949663051438536, \n                      \"max\": 0.29846571176821973, \n
```

```

n      \ "num_unique_values\ ": 13,\n      \ "samples\ ": [\n
0.10568593006093435,\n      0.29846571176821973,\n      -
0.04402819897179786\n      ],\n      \ "semantic_type\ ": \ "\",\n
\ "description\ ": \ "\",\n      },\n      {\n      \ "column\ ":
\ "2012\ ",\n      \ "properties\ ": {\n      \ "dtype\ ": \ "number\ ",\n
\ "std\ ": 0.0760875119496455,\n      \ "min\ ": -0.0554315465970493,\n
\ "max\ ": 0.19546363216692023,\n      \ "num_unique_values\ ": 13,\n
\ "samples\ ": [\n      0.06215838639383471,\n
0.1414238134668693,\n      0.018802302394106803\n      ],\n
\ "semantic_type\ ": \ "\",\n      \ "description\ ": \ "\",\n
n      },\n      {\n      \ "column\ ": \ "2017\ ",\n      \ "properties\ ": {\n
\ "dtype\ ": \ "number\ ",\n      \ "std\ ": 0.06488451501246204,\n
\ "min\ ": -0.01557000046554935,\n      \ "max\ ": 0.19168502575337576,\n
n      \ "num_unique_values\ ": 13,\n      \ "samples\ ": [\n
0.15523513667675834,\n      0.10568593006093435,\n      -
0.01557000046554935\n      ],\n      \ "semantic_type\ ": \ "\",\n
\ "description\ ": \ "\",\n      },\n      {\n      \ "column\ ":
\ "2022\ ",\n      \ "properties\ ": {\n      \ "dtype\ ": \ "number\ ",\n
\ "std\ ": 0.14796246902213392,\n      \ "min\ ": -0.114341266387337,\n
\ "max\ ": 0.4479343268668614,\n      \ "num_unique_values\ ": 13,\n
\ "samples\ ": [\n      0.19168502575337576,\n
0.25177002230576445,\n      -0.006743258081727445\n      ],\n
\ "semantic_type\ ": \ "\",\n      \ "description\ ": \ "\",\n
n      },\n      ],\n      },\n      \ "type\ ": "dataframe"}

```

```

print('correlation')
columns_selected.corr()

```

correlation

<ipython-input-23-001bd0c92dbe>:2: FutureWarning: The default value of numeric\_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric\_only to silence this warning.

```
columns_selected.corr()
```

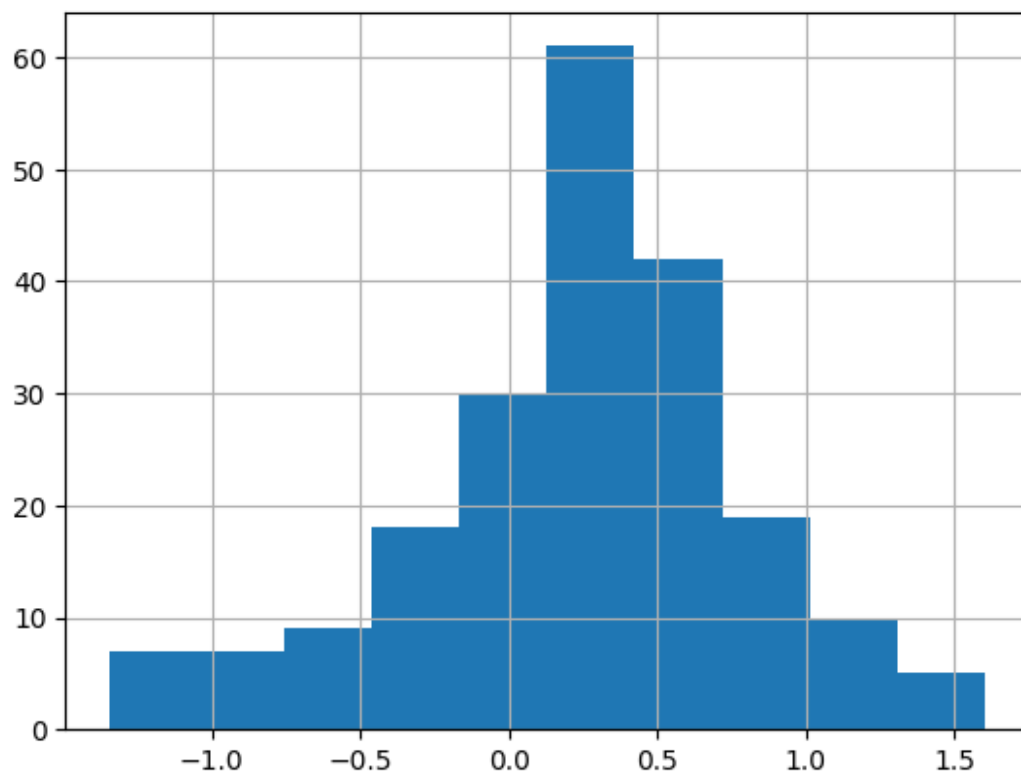
```

{"summary": "{\n  \ "name\ ": \ "columns_selected\ ",\n  \ "rows\ ": 13,\n
\ "fields\ ": [\n    {\n      \ "column\ ": \ "1962\ ",\n
\ "properties\ ": {\n      \ "dtype\ ": \ "number\ ",\n      \ "std\ ":
0.414178160665609,\n      \ "min\ ": -0.5570005826546497,\n
\ "max\ ": 1.0,\n      \ "num_unique_values\ ": 13,\n
\ "samples\ ": [\n      -0.11167489232220965,\n      -
0.2827423243155094,\n      1.0\n      ],\n
\ "semantic_type\ ": \ "\",\n      \ "description\ ": \ "\",\n
n      },\n      {\n      \ "column\ ": \ "1967\ ",\n      \ "properties\ ": {\n
\ "dtype\ ": \ "number\ ",\n      \ "std\ ": 0.3911285699988312,\n
\ "min\ ": -0.5395866556261767,\n      \ "max\ ": 1.0,\n
\ "num_unique_values\ ": 13,\n      \ "samples\ ": [\n
0.2656641724117877,\n      0.4493517582290633,\n      -
0.5395866556261767\n      ],\n      \ "semantic_type\ ": \ "\",\n

```

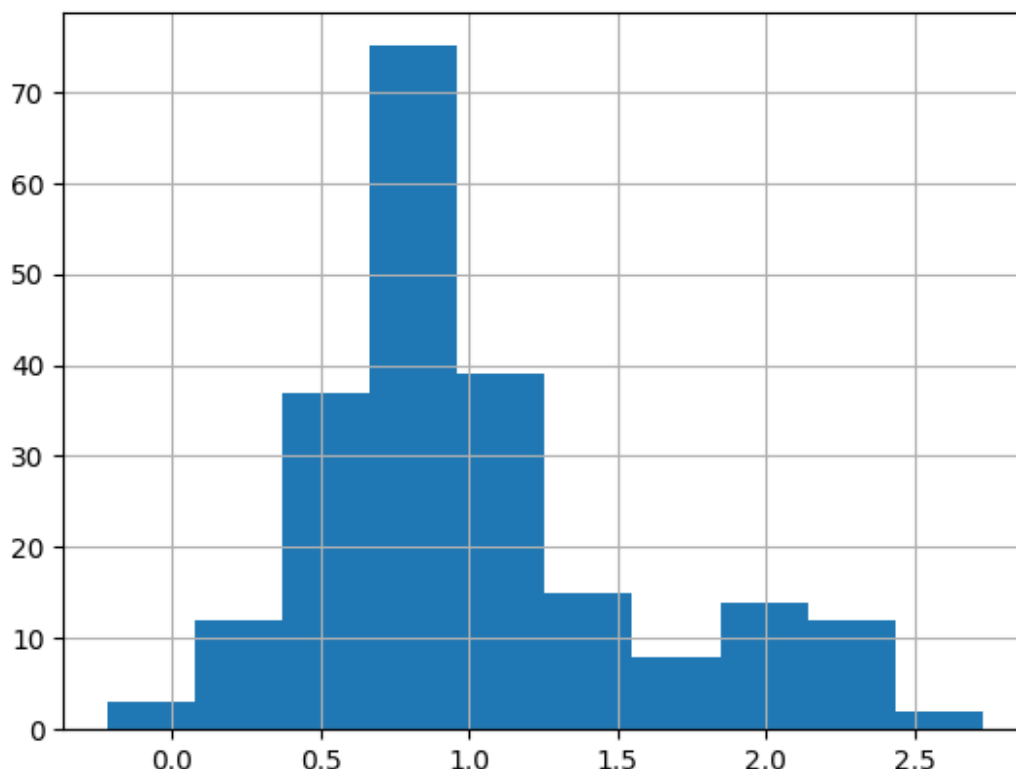
```
\\"description\\": \\'\\'\\n    }\\n        \\\"column\\\":  
\\\"1972\\\",\\n        \\\"properties\\\": {\\n            \\\"dtype\\\": \\\"number\\\",\\n  
\\\"std\\\": 0.3873177285870638,\\n            \\\"min\\\": -0.5546517277963406,\\n  
\\\"max\\\": 1.0,\\n            \\\"num_unique_values\\\": 13,\\n  
\\\"samples\\\": [\\n                -0.06362939309893603,\\n  
0.1646174080865412,\\n                -0.5546517277963406\\n            ],\\n  
\\\"semantic_type\\\": \\\"\\\",\\n            \\\"description\\\": \\\"\\\"\\n        }\\n  
n    },\\n    {\\n        \\\"column\\\": \\\"1977\\\",\\n        \\\"properties\\\": {\\n  
\\\"dtype\\\": \\\"number\\\",\\n            \\\"std\\\": 0.2902131222472633,\\n  
\\\"min\\\": -0.11095843511250704,\\n            \\\"max\\\": 1.0,\\n  
\\\"num_unique_values\\\": 13,\\n            \\\"samples\\\": [\\n  
0.027545855618948525,\\n                0.09041632055110377,\\n  
0.07250482919605752\\n            ],\\n            \\\"semantic_type\\\": \\\"\\\",\\n  
\\\"description\\\": \\\"\\\"\\n        }\\n    },\\n    {\\n        \\\"column\\\":  
\\\"1982\\\",\\n        \\\"properties\\\": {\\n            \\\"dtype\\\": \\\"number\\\",\\n  
\\\"std\\\": 0.343639499331109,\\n            \\\"min\\\": -0.3508971755858687,\\n  
\\\"max\\\": 1.0,\\n            \\\"num_unique_values\\\": 13,\\n  
\\\"samples\\\": [\\n                0.2943882471540574,\\n  
0.1897725116715811,\\n                -0.3508971755858687\\n            ],\\n  
\\\"semantic_type\\\": \\\"\\\",\\n            \\\"description\\\": \\\"\\\"\\n        }\\n  
n    },\\n    {\\n        \\\"column\\\": \\\"1987\\\",\\n        \\\"properties\\\": {\\n  
\\\"dtype\\\": \\\"number\\\",\\n            \\\"std\\\": 0.39436101848632815,\\n  
\\\"min\\\": -0.4131341185607546,\\n            \\\"max\\\": 1.0,\\n  
\\\"num_unique_values\\\": 13,\\n            \\\"samples\\\": [\\n                -  
0.015664304538638825,\\n                -0.4131341185607546,\\n  
0.09114736305841736\\n            ],\\n            \\\"semantic_type\\\": \\\"\\\",\\n  
\\\"description\\\": \\\"\\\"\\n        }\\n    },\\n    {\\n        \\\"column\\\":  
\\\"1992\\\",\\n        \\\"properties\\\": {\\n            \\\"dtype\\\": \\\"number\\\",\\n  
\\\"std\\\": 0.37114069979990205,\\n            \\\"min\\\": -0.5570005826546497,\\n  
\\\"max\\\": 1.0,\\n            \\\"num_unique_values\\\": 13,\\n  
\\\"samples\\\": [\\n                0.21986583061122933,\\n  
0.47049353257935705,\\n                -0.5570005826546497\\n            ],\\n  
\\\"semantic_type\\\": \\\"\\\",\\n            \\\"description\\\": \\\"\\\"\\n        }\\n  
n    },\\n    {\\n        \\\"column\\\": \\\"1997\\\",\\n        \\\"properties\\\": {\\n  
\\\"dtype\\\": \\\"number\\\",\\n            \\\"std\\\": 0.31837803581656665,\\n  
\\\"min\\\": -0.2277604062070909,\\n            \\\"max\\\": 1.0,\\n  
\\\"num_unique_values\\\": 13,\\n            \\\"samples\\\": [\\n  
0.5332819216501354,\\n                0.20562587051888986,\\n  
0.2277604062070909\\n            ],\\n            \\\"semantic_type\\\": \\\"\\\",\\n  
\\\"description\\\": \\\"\\\"\\n        }\\n    },\\n    {\\n        \\\"column\\\":  
\\\"2002\\\",\\n        \\\"properties\\\": {\\n            \\\"dtype\\\": \\\"number\\\",\\n  
\\\"std\\\": 0.3472091931235733,\\n            \\\"min\\\": -0.19028252564309875,\\n  
\\\"max\\\": 1.0,\\n            \\\"num_unique_values\\\": 13,\\n  
\\\"samples\\\": [\\n                0.6247718461160864,\\n  
0.7264864142077491,\\n                -0.1504196236880361\\n            ],\\n  
\\\"semantic_type\\\": \\\"\\\",\\n            \\\"description\\\": \\\"\\\"\\n        }\\n  
n    },\\n    {\\n        \\\"column\\\": \\\"2007\\\",\\n        \\\"properties\\\": {\\n  
\\\"dtype\\\": \\\"number\\\",\\n            \\\"std\\\": 0.3979578787733069,\\n  
\\\"min\\\": -0.4131341185607546,\\n            \\\"max\\\": 1.0,\\n
```





```
data['2007'].hist(bins=10)
```

```
<Axes: >
```



```
columns_selected.head()
```

```
{
  "summary": {
    "name": "columns_selected",
    "rows": 225,
    "fields": [
      {
        "column": "Country",
        "properties": {
          "dtype": "string",
          "num_unique_values": 225,
          "samples": [
            "Armenia, Rep. of",
            "Spain",
            "Maldives"
          ],
          "semantic_type": "",
          "description": ""
        }
      },
      {
        "column": "1962",
        "properties": {
          "dtype": "number",
          "std": 0.3418123874389091,
          "min": -0.908,
          "max": 0.998,
          "num_unique_values": 169,
          "samples": [
            0.093,
            -0.118,
            -0.776
          ],
          "semantic_type": "",
          "description": ""
        }
      },
      {
        "column": "1967",
        "properties": {
          "dtype": "number",
          "std": 0.3394841957312319,
          "min": -1.048,
          "max": 1.134,
          "num_unique_values": 160,
          "samples": [
            0.159,
            0.032,
            -1.028
          ],
          "semantic_type": "",
          "description": ""
        }
      },
      {
        "column": "1972",
        "properties": {
          "dtype": "number",
          "std": 0.3844737091788439,
          "min": -1.796,
          "max": 0.933,
          "num_unique_values": 167,
          "samples": [
            -0.703,
            -0.268,
            0.304
          ]
        }
      }
    ]
  }
}
```

```

\"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"1977\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.25342808077817564, \n      \"min\": -0.599, \n      \"max\": 1.079, \n      \"num_unique_values\": 167, \n      \"samples\": [ \n        0.039, \n        -0.022, \n        -0.279 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"1982\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.31829855255580203, \n      \"min\": -0.682, \n      \"max\": 1.135, \n      \"num_unique_values\": 174, \n      \"samples\": [ \n        0.238, \n        0.709, \n        0.623 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"1987\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.47508396314884604, \n      \"min\": -1.652, \n      \"max\": 1.562, \n      \"num_unique_values\": 172, \n      \"samples\": [ \n        0.368, \n        0.505, \n        0.964 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"1992\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.5652350970020732, \n      \"min\": -1.344, \n      \"max\": 1.601, \n      \"num_unique_values\": 191, \n      \"samples\": [ \n        0.875, \n        0.533, \n        0.807 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"1997\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.4830883203093068, \n      \"min\": -0.429, \n      \"max\": 1.933, \n      \"num_unique_values\": 196, \n      \"samples\": [ \n        0.079, \n        1.321, \n        0.359 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"2002\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.37852904375599056, \n      \"min\": 0.009, \n      \"max\": 2.255, \n      \"num_unique_values\": 191, \n      \"samples\": [ \n        1.135, \n        1.13, \n        1.376 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"2007\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.5463201550082332, \n      \"min\": -0.219, \n      \"max\": 2.729, \n      \"num_unique_values\": 200, \n      \"samples\": [ \n        0.547, \n        1.08, \n        1.242 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  }, \n  { \n    \"column\": \"2012\", \n    \"properties\": { \n      \"dtype\": \"number\", \n      \"std\": 0.44211269170531653, \n      \"min\": -0.128, \n      \"max\": 2.144, \n      \"num_unique_values\": 191, \n      \"samples\": [ \n        0.533, \n        0.795, \n        0.508 \n      ], \n      \"semantic_type\": \"\", \n      \"description\": \"\" \n    } \n  } \n}

```



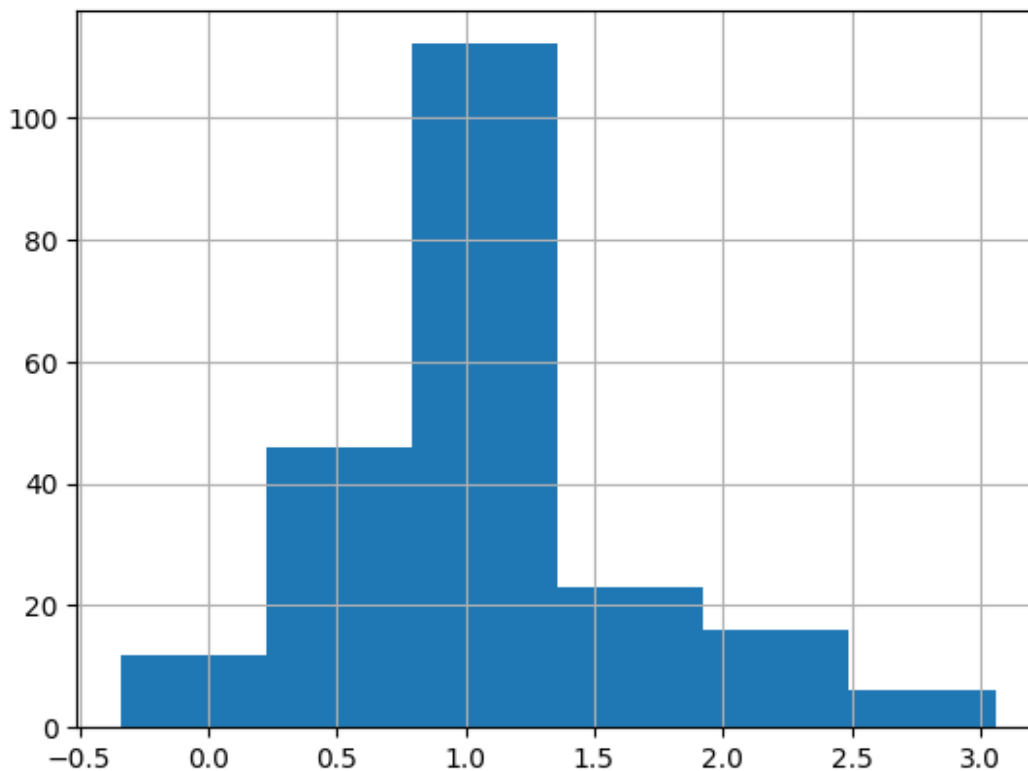
```

n    },\n    {\n        \"column\": \"2017\", \n        \"properties\": {\n            \"dtype\": \"number\", \n            \"std\": 0.39399890441060664, \n            \"min\": 0.017, \n            \"max\": 2.493, \n            \"num_unique_values\": 194, \n            \"samples\": [\n                0.79, \n                1.124, \n                1.184\n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\" \n        }, \n        {\n            \"column\": \"2022\", \n            \"properties\": {\n                \"dtype\": \n                \"number\", \n                \"std\": 0.6692789604244717, \n                \"min\": -\n                1.305, \n                \"max\": 3.243, \n                \"num_unique_values\": 206, \n                \"samples\": [\n                    1.216, \n                    1.707, \n                    1.074\n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\" \n            }, \n            {\n            }\n        }\n    ], \n    \"type\": \"dataframe\", \"variable_name\": \"columns_selected\"}

```

```
data['2010'].hist(bins=6)
```

```
<Axes: >
```

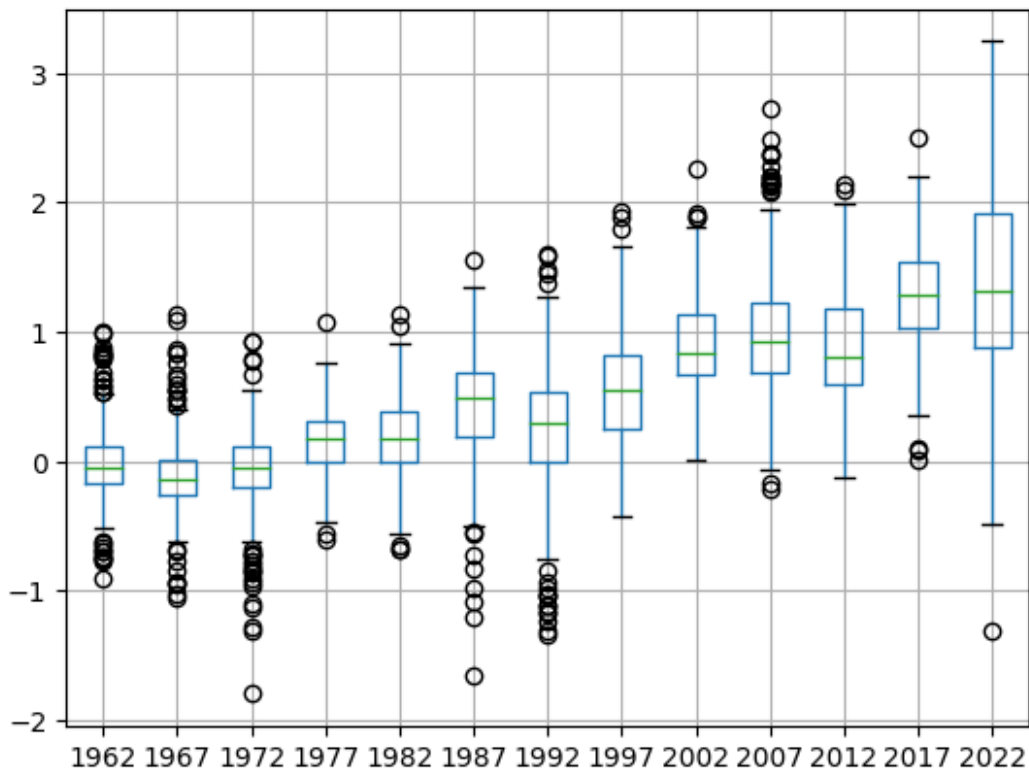


```

x = columns_selected['1977']
type(x)
pd.core.series.Series
y = pd.DataFrame(x)
type(y)
pd.core.frame.DataFrame
columns_selected.boxplot()

```

<Axes: >



```
type(data)
pd.core.frame.DataFrame
columns_selected['1977'].boxplot()
```

```
-----
-----
AttributeError                                Traceback (most recent call
last)
```

```
<ipython-input-38-0c8d66558987> in <cell line: 3>()
```

```
1 type(data)
```

```
2 pd.core.frame.DataFrame
```

```
----> 3 columns_selected['1977'].boxplot()
```

```
/usr/local/lib/python3.10/dist-packages/pandas/core/generic.py in
```

```
__getattr__(self, name)
```

```
5900         ):
```

```
5901             return self[name]
```

```
-> 5902         return object.__getattribute__(self, name)
```

```
5903
```

```
5904     def __setattr__(self, name: str, value) -> None:
```

```
AttributeError: 'Series' object has no attribute 'boxplot'
```

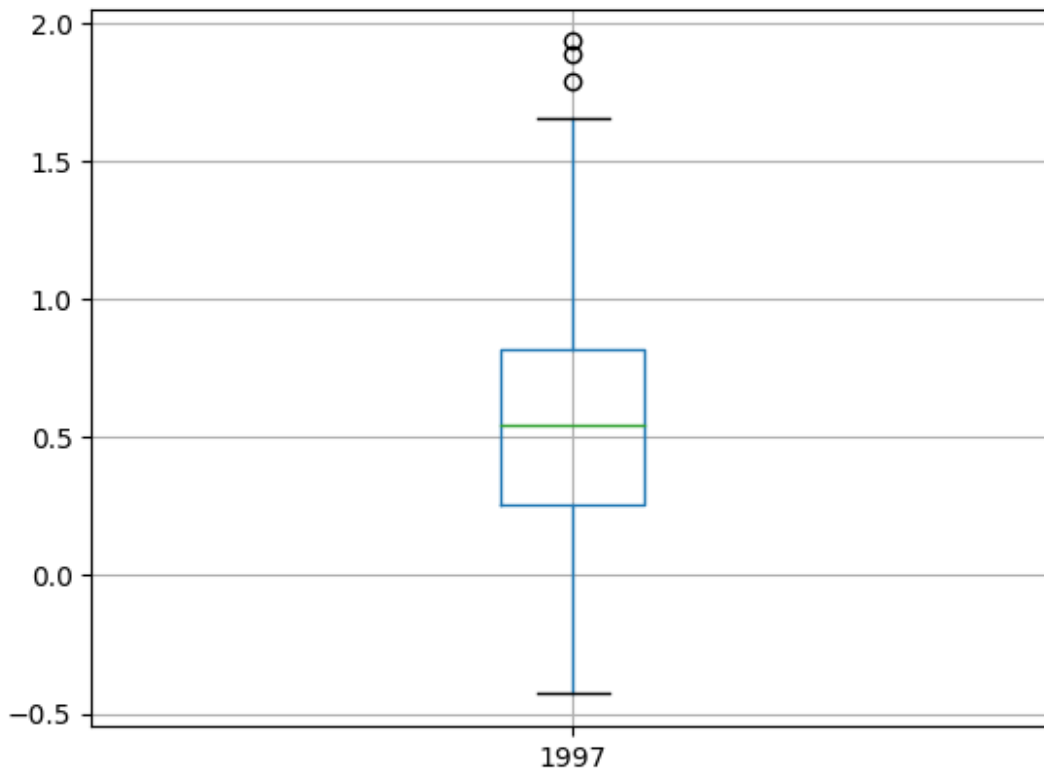
```

x = pd.DataFrame(data['1997'])
type(x)
pd.core.frame.DataFrame
x = (columns_selected['1997'])
x = pd.DataFrame(x)
type(x)

pd.core.frame.DataFrame
type(x)
pd.core.frame.DataFrame

x.boxplot()
<Axes: >

```



```

import matplotlib.pyplot as plt
columns_selected.head(2)

{"summary": "{\n  \"name\": \"columns_selected\",\n  \"rows\": 225,\n  \"fields\": [\n    {\n      \"column\": \"Country\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 225,\n        \"samples\": [\n          \"Armenia, Rep. of\",\n          \"Spain\",\n          \"Maldives\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1962\",\n      \"properties\":

```

```
{\n  \"dtype\": \"number\", \"std\": \n0.3418123874389091, \"min\": -0.908, \"max\": \n0.998, \"num_unique_values\": 169, \"samples\": [\n0.093, \n-0.118, \n-0.776\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1967\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.3394841957312319, \n\"min\": -1.048, \"max\": 1.134, \n\"num_unique_values\": 160, \"samples\": [\n0.159, \n0.032, \n-1.028\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1972\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.3844737091788439, \n\"min\": -1.796, \"max\": 0.933, \n\"num_unique_values\": 167, \"samples\": [\n\n0.304, \n-0.703, \n-0.268\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1977\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.25342808077817564, \n\"min\": -0.599, \"max\": 1.079, \n\"num_unique_values\": 167, \"samples\": [\n\n-0.039, \n-0.022, \n-0.279\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1982\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.31829855255580203, \n\"min\": -0.682, \"max\": 1.135, \n\"num_unique_values\": 174, \"samples\": [\n\n0.238, \n0.709, \n0.623\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1987\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.47508396314884604, \n\"min\": -1.652, \"max\": 1.562, \n\"num_unique_values\": 172, \"samples\": [\n\n0.368, \n0.505, \n0.964\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1992\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.5652350970020732, \n\"min\": -1.344, \"max\": 1.601, \n\"num_unique_values\": 191, \"samples\": [\n\n0.875, \n0.533, \n0.807\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"1997\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.4830883203093068, \n\"min\": -0.429, \"max\": 1.933, \n\"num_unique_values\": 196, \"samples\": [\n\n0.079, \n1.321, \n0.359\n], \n\"semantic_type\": \"\", \"description\": \"\"\n}, {\n  \"column\": \"2002\", \"properties\": {\n\"dtype\": \"number\", \"std\": 0.37852904375599056, \n
```

```

{"min": 0.009, "max": 2.255, "num_unique_values": 191, "samples": [1.135, 1.13, 1.376], "semantic_type": "", "description": "", "column": "2007", "properties": {"dtype": "number", "std": 0.5463201550082332, "min": -0.219, "max": 2.729, "num_unique_values": 200, "samples": [0.547, 1.08, 1.242], "semantic_type": "", "description": "", "column": "2012", "properties": {"dtype": "number", "std": 0.44211269170531653, "min": -0.128, "max": 2.144, "num_unique_values": 191, "samples": [0.533, 0.795, 0.508], "semantic_type": "", "description": "", "column": "2017", "properties": {"dtype": "number", "std": 0.39399890441060664, "min": 0.017, "max": 2.493, "num_unique_values": 194, "samples": [0.79, 1.124, 1.184], "semantic_type": "", "description": "", "column": "2022", "properties": {"dtype": "number", "std": 0.6692789604244717, "min": -1.305, "max": 3.243, "num_unique_values": 206, "samples": [1.216, 1.707, 1.074], "semantic_type": "", "description": ""}}], "type": "dataframe", "variable_name": "columns_selected"}

```

```

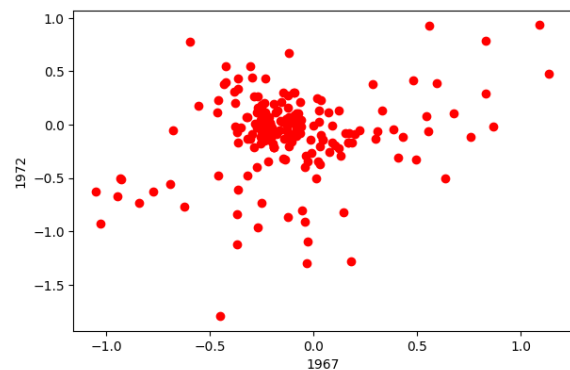
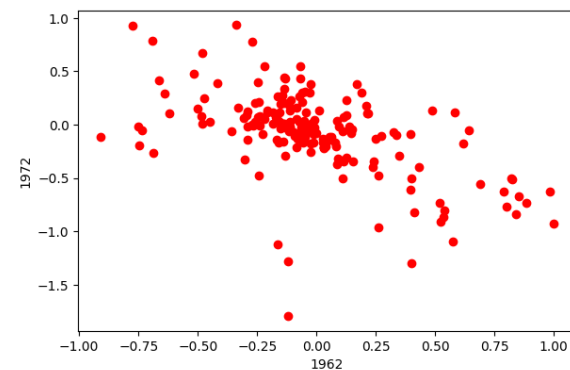
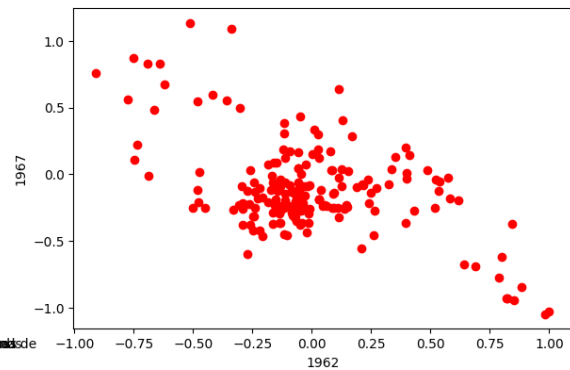
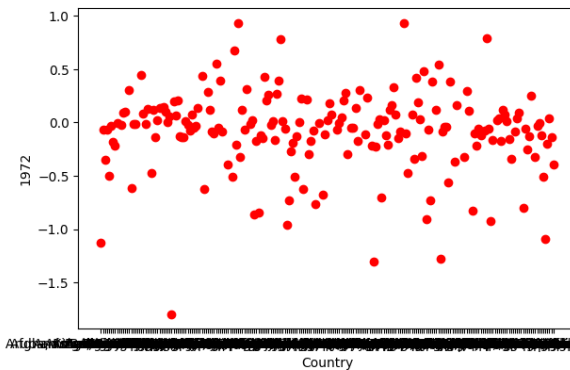
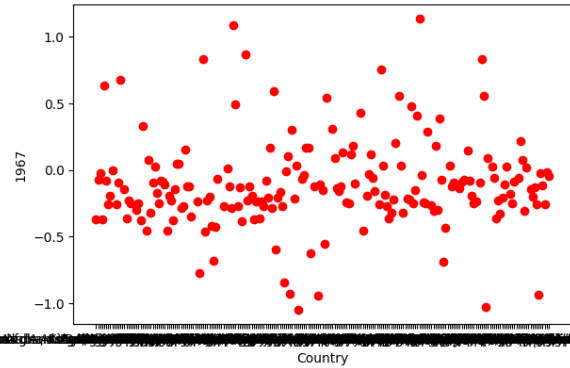
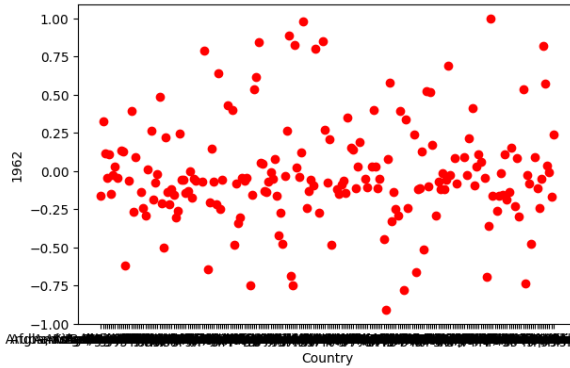
fig, axes=plt.subplots(3,2, figsize=(15,15))
index=0

```

```

for i in range(3):
    for j in range(i+1,4):
        ax1=int(index/2)
        ax2=int(index % 2)
        axes[ax1]
        [ax2].scatter(columns_selected[columns_selected.columns[i]],columns_selected[columns_selected.columns[j]], color='red')
        axes[ax1][ax2].set_xlabel(columns_selected.columns[i])
        axes[ax1][ax2].set_ylabel(columns_selected.columns[j])
        index = index +1

```



```
columns_selected.describe()
```

```
{
  "summary": {
    "name": "columns_selected",
    "rows": 8,
    "fields": [
      {
        "column": "1962",
        "properties": {
          "dtype": "number",
          "std": 66.80791808693246,
          "min": -0.908,
          "max": 189.0,
          "num_unique_values": 8,
          "samples": [
            -0.01347619047619048,
            -0.056,
            189.0
          ],
          "semantic_type": "\"\"",
          "description": "\"\""
        }
      },
      {
        "column": "1967",
        "properties": {
          "dtype": "number",
          "std": 67.53527830500084,
          "min": -1.048,
          "max": 191.0,
          "num_unique_values": 8,
          "samples": [
            -0.11083246073298428,
            -0.146,
            191.0
          ],
          "semantic_type": "\"\"",
          "description": "\"\""
        }
      }
    ]
  }
}
```

```
{\n      },\n      {\n        \"column\": \"1972\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 67.92137580160988,\n          \"min\": -1.796,\n          \"max\": 192.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            -0.08490625,\n            -0.045,\n            192.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"1977\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 65.33820846769545,\n          \"min\": -0.599,\n          \"max\": 185.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.1658162162162162,\n            0.182,\n            185.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"1982\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 67.80739834083896,\n          \"min\": -0.682,\n          \"max\": 192.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.17649479166666668,\n            0.181,\n            192.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"1987\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 67.07211253384104,\n          \"min\": -1.652,\n          \"max\": 190.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.405021052631579,\n            0.491,\n            190.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"1992\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 73.4486350743146,\n          \"min\": -1.344,\n          \"max\": 208.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.2364903846153846,\n            0.2985,\n            208.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"1997\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 72.97889741647639,\n          \"min\": -0.429,\n          \"max\": 207.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.5439951690821256,\n            0.547,\n            207.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"2002\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 74.64205429898776,\n          \"min\": 0.009,\n          \"max\": 212.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            0.9249999999999998,\n            0.84,\n            212.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"2007\",\n        \"properties\": {\n          \"dtype\": \"number\", \n          \"std\": 76.3771649102675,\n          \"min\": -0.219,\n          \"max\": 217.0,\n          \"num_unique_values\": 8,\n          \"samples\": [\n            1.0225483870967744,\n            0.921,\n            217.0\n          ],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"2012\",
```

```

\"dtype\": \"number\",\\n          \"std\": 75.71665953709281,\\n
\"min\": -0.128,\\n          \"max\": 215.0,\\n
\"num_unique_values\": 8,\\n          \"samples\": [\\n
0.9022232558139536,\\n          0.808,\\n          215.0\\n          ],\\n
\"semantic_type\": \"\",\\n          \"description\": \"\"\\n          }\\n
n      },\\n      {\\n          \"column\": \"2017\",\\n          \"properties\": {\\n
\"dtype\": \"number\",\\n          \"std\": 75.2585459050908,\\n
\"min\": 0.017,\\n          \"max\": 214.0,\\n
\"num_unique_values\": 8,\\n          \"samples\": [\\n
1.2807850467289719,\\n          1.282,\\n          214.0\\n          ],\\n
\"semantic_type\": \"\",\\n          \"description\": \"\"\\n          }\\n
n      },\\n      {\\n          \"column\": \"2022\",\\n          \"properties\": {\\n
\"dtype\": \"number\",\\n          \"std\": 74.90859590103369,\\n
\"min\": -1.305,\\n          \"max\": 213.0,\\n
\"num_unique_values\": 8,\\n          \"samples\": [\\n
1.382112676056338,\\n          1.315,\\n          213.0\\n          ],\\n
\"semantic_type\": \"\",\\n          \"description\": \"\"\\n          }\\n
n      }\\n  ]\\n}\" , \"type\": \"dataframe\"}

```

```
import pandas as pd
```

```
data=pd.read_csv('climate.csv',header=None)
```

```
data
```

```
{\"type\": \"dataframe\", \"variable_name\": \"data\"}
```

```
data.head(20)
```

```
{\"type\": \"dataframe\", \"variable_name\": \"data\"}
```

```
data.tail(10)
```

```
{\"type\": \"dataframe\"}
```

```
data.sample(20)
```

```
{\"type\": \"dataframe\"}
```

```
from pandas.api.types import is_numeric_dtype
```

```
for col in data.columns:
```

```
    if is_numeric_dtype(data[col]):
```

```
        print('%s: '%(col))
```

```
        print('\\t Mean = %.2f'%data[col].mean())
```

```
        print('\\t Standard Deviation= %.2f'%data[col].std())
```

```
        print('\\t Minimum = %.2f'%data[col].min())
```

```
        print('\\t Maximum = %.2f'%data[col].max())
```

```
10:
```

```
    Mean = 10.54
```

```
    Standard Deviation= 142.63
```

```
    Minimum = -0.69
```



Maximum = 1961.00  
11: Mean = 10.31  
Standard Deviation= 142.34  
Minimum = -0.91  
Maximum = 1962.00  
12: Mean = 10.38  
Standard Deviation= 142.79  
Minimum = -1.27  
Maximum = 1963.00  
13: Mean = 10.32  
Standard Deviation= 142.87  
Minimum = -0.88  
Maximum = 1964.00  
14: Mean = 10.15  
Standard Deviation= 142.95  
Minimum = -1.06  
Maximum = 1965.00  
15: Mean = 10.29  
Standard Deviation= 141.51  
Minimum = -1.80  
Maximum = 1966.00  
16: Mean = 10.13  
Standard Deviation= 141.96  
Minimum = -1.05  
Maximum = 1967.00  
17: Mean = 10.05  
Standard Deviation= 142.04  
Minimum = -1.63  
Maximum = 1968.00  
18: Mean = 10.47  
Standard Deviation= 142.46  
Minimum = -0.90  
Maximum = 1969.00  
19: Mean = 10.46  
Standard Deviation= 142.91  
Minimum = -1.29  
Maximum = 1970.00  
20: Mean = 10.07  
Standard Deviation= 142.26

Minimum = -0.87  
Maximum = 1971.00

21: Mean = 10.13  
Standard Deviation= 141.95  
Minimum = -1.80  
Maximum = 1972.00

22: Mean = 10.40  
Standard Deviation= 141.64  
Minimum = -0.99  
Maximum = 1973.00

23: Mean = 10.07  
Standard Deviation= 142.10  
Minimum = -0.98  
Maximum = 1974.00

24: Mean = 10.43  
Standard Deviation= 143.66  
Minimum = -1.09  
Maximum = 1975.00

25: Mean = 10.16  
Standard Deviation= 143.37  
Minimum = -0.96  
Maximum = 1976.00

26: Mean = 10.79  
Standard Deviation= 144.95  
Minimum = -0.60  
Maximum = 1977.00

27: Mean = 10.48  
Standard Deviation= 143.49  
Minimum = -0.87  
Maximum = 1978.00

28: Mean = 10.65  
Standard Deviation= 143.56  
Minimum = -1.24  
Maximum = 1979.00

29: Mean = 10.56  
Standard Deviation= 142.88  
Minimum = -0.76  
Maximum = 1980.00

30: Mean = 10.50

```
Standard Deviation= 142.95
Minimum = -0.91
Maximum = 1981.00
31:
Mean = 10.45
Standard Deviation= 142.66
Minimum = -0.68
Maximum = 1982.00
32:
Mean = 10.72
Standard Deviation= 143.46
Minimum = -2.06
Maximum = 1983.00
33:
Mean = 10.58
Standard Deviation= 144.31
Minimum = -1.46
Maximum = 1984.00
34:
Mean = 10.57
Standard Deviation= 144.38
Minimum = -1.19
Maximum = 1985.00
35:
Mean = 10.55
Standard Deviation= 143.69
Minimum = -0.77
Maximum = 1986.00
36:
Mean = 10.81
Standard Deviation= 143.75
Minimum = -1.65
Maximum = 1987.00
37:
Mean = 10.90
Standard Deviation= 143.81
Minimum = -0.50
Maximum = 1988.00
38:
Mean = 10.67
Standard Deviation= 143.90
Minimum = -1.54
Maximum = 1989.00
39:
Mean = 11.03
Standard Deviation= 144.33
Minimum = -0.74
Maximum = 1990.00
40:
```

```
Mean = 10.90
Standard Deviation= 144.80
Minimum = -0.70
Maximum = 1991.00
41:
Mean = 9.77
Standard Deviation= 137.77
Minimum = -1.34
Maximum = 1992.00
42:
Mean = 9.71
Standard Deviation= 137.52
Minimum = -1.35
Maximum = 1993.00
43:
Mean = 10.15
Standard Deviation= 137.89
Minimum = -0.42
Maximum = 1994.00
44:
Mean = 10.08
Standard Deviation= 137.30
Minimum = -0.33
Maximum = 1995.00
45:
Mean = 9.74
Standard Deviation= 137.39
Minimum = -0.79
Maximum = 1996.00
46:
Mean = 10.14
Standard Deviation= 138.43
Minimum = -0.43
Maximum = 1997.00
47:
Mean = 10.44
Standard Deviation= 137.48
Minimum = -0.61
Maximum = 1998.00
48:
Mean = 10.26
Standard Deviation= 137.89
Minimum = -0.27
Maximum = 1999.00
49:
Mean = 10.19
Standard Deviation= 137.97
Minimum = -0.72
Maximum = 2000.00
```

50:  
Mean = 10.42  
Standard Deviation= 138.35  
Minimum = -0.19  
Maximum = 2001.00

51:  
Mean = 10.32  
Standard Deviation= 137.11  
Minimum = 0.01  
Maximum = 2002.00

52:  
Mean = 10.16  
Standard Deviation= 136.55  
Minimum = -0.25  
Maximum = 2003.00

53:  
Mean = 10.14  
Standard Deviation= 136.94  
Minimum = -0.62  
Maximum = 2004.00

54:  
Mean = 10.26  
Standard Deviation= 137.32  
Minimum = -0.39  
Maximum = 2005.00

55:  
Mean = 10.16  
Standard Deviation= 136.43  
Minimum = -0.51  
Maximum = 2006.00

56:  
Mean = 10.22  
Standard Deviation= 135.86  
Minimum = -0.22  
Maximum = 2007.00

57:  
Mean = 10.23  
Standard Deviation= 137.53  
Minimum = -0.14  
Maximum = 2008.00

58:  
Mean = 10.34  
Standard Deviation= 137.59  
Minimum = -0.32  
Maximum = 2009.00

59:  
Mean = 10.40  
Standard Deviation= 136.69  
Minimum = -0.34

Maximum = 2010.00  
60: Mean = 10.04  
Standard Deviation= 136.15  
Minimum = -0.48  
Maximum = 2011.00  
61: Mean = 10.21  
Standard Deviation= 136.84  
Minimum = -0.13  
Maximum = 2012.00  
62: Mean = 10.20  
Standard Deviation= 136.59  
Minimum = 0.12  
Maximum = 2013.00  
63: Mean = 10.39  
Standard Deviation= 136.64  
Minimum = -0.09  
Maximum = 2014.00  
64: Mean = 10.55  
Standard Deviation= 136.70  
Minimum = -0.43  
Maximum = 2015.00  
65: Mean = 10.85  
Standard Deviation= 137.71  
Minimum = 0.25  
Maximum = 2016.00  
66: Mean = 10.66  
Standard Deviation= 137.47  
Minimum = 0.02  
Maximum = 2017.00  
67: Mean = 10.73  
Standard Deviation= 137.86  
Minimum = 0.24  
Maximum = 2018.00  
68: Mean = 10.87  
Standard Deviation= 137.92  
Minimum = 0.05  
Maximum = 2019.00  
69: Mean = 11.03  
Standard Deviation= 138.30

```

    Minimum = 0.23
    Maximum = 2020.00
70:
    Mean = 10.78
    Standard Deviation= 138.06
    Minimum = -0.42
    Maximum = 2021.00
71:
    Mean = 10.82
    Standard Deviation= 138.13
    Minimum = -1.30
    Maximum = 2022.00

```

```
data['Country'].value_counts()
```

```

Afghanistan, Islamic Rep. of    1
Libya                          1
New Zealand                    1
Nicaragua                     1
Niger                         1
..
Grenada                       1
Guadeloupe                   1
Guatemala                    1
Guinea                       1
Zimbabwe                     1
Name: Country, Length: 225, dtype: int64

```

```
columns_selected.describe(include='all')
```

```

{"summary":{"\n  \"name\": \"columns_selected\",\n  \"rows\": 11,\n  \"fields\": [\n    {\n      \"column\": \"Country\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 3,\n        \"samples\": [\n          \"225\",\n          \"Afghanistan, Islamic Rep. of\",\n          \"1\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1962\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 66.80791808693246,\n        \"min\": -0.908,\n        \"max\": 189.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n          -0.01347619047619048,\n          -0.056,\n          189.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1967\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 67.53527830500084,\n        \"min\": -1.048,\n        \"max\": 191.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n          -0.11083246073298428,\n          -0.146,\n          191.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"1972\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\":

```

```
67.92137580160988,\n        \"min\": -1.796,\n        \"max\": 192.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        -0.08490625,\n        -0.045,\n        192.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"1977\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 65.33820846769545,\n        \"min\": -0.599,\n        \"max\": 185.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.1658162162162162,\n        0.182,\n        185.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"1982\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 67.80739834083896,\n        \"min\": -0.682,\n        \"max\": 192.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.17649479166666668,\n        0.181,\n        192.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"1987\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 67.07211253384104,\n        \"min\": -1.652,\n        \"max\": 190.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.405021052631579,\n        0.491,\n        190.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"1992\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 73.4486350743146,\n        \"min\": -1.344,\n        \"max\": 208.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.2364903846153846,\n        0.2985,\n        208.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"1997\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 72.97889741647639,\n        \"min\": -0.429,\n        \"max\": 207.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.5439951690821256,\n        0.547,\n        207.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"2002\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 74.64205429898776,\n        \"min\": 0.009,\n        \"max\": 212.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        0.9249999999999998,\n        0.84,\n        212.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"2007\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 76.3771649102675,\n        \"min\": -0.219,\n        \"max\": 217.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n        1.0225483870967744,\n        0.921,\n        217.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n    },\n    {\n        \"column\": \"2012\",\n        \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 75.71665953709281,\n        \"min\": -0.128,\n        \"max\": 215.0,\n
```



```

\"num_unique_values\": 8,\n      \"samples\": [\n0.9022232558139536,\n      0.808,\n      215.0\n      ],\n\"semantic_type\": \"\",\n      \"description\": \"\"\n      },\n      {\n      \"column\": \"2017\",\n      \"properties\": {\n\"dtype\": \"number\",\n      \"std\": 75.25854590509083,\n\"min\": 0.017,\n      \"max\": 214.0,\n\"num_unique_values\": 8,\n      \"samples\": [\n1.2807850467289719,\n      1.282,\n      214.0\n      ],\n\"semantic_type\": \"\",\n      \"description\": \"\"\n      },\n      {\n      \"column\": \"2022\",\n      \"properties\": {\n\"dtype\": \"number\",\n      \"std\": 74.90859590103369,\n\"min\": -1.305,\n      \"max\": 213.0,\n\"num_unique_values\": 8,\n      \"samples\": [\n1.382112676056338,\n      1.315,\n      213.0\n      ],\n\"semantic_type\": \"\",\n      \"description\": \"\"\n      }\n      ]\n    },\n    \"type\": \"dataframe\"}

```