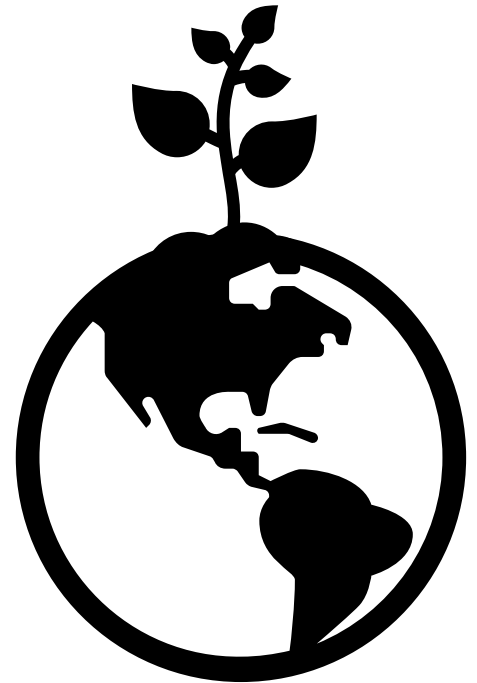


Degrees of Climate Change Project Summary

5 June 2018

Todd Schultz, Rahul Birmiwal, Abhishek Anand



Outline

- Background
- User persona/Research Questions
- Data Sources
- Demo
- Design
- Project Structure
- Lessons learned and future work

Background

- Climate change, fundamentally, is defined by
 - Historical changes in statistical properties of a climate system(s)
 - Often measured by sample mean/sample variance (spread) of historical global temperatures
 - Does climate change equal global warming and what caused it?
- Ramifications reach to the basic ability of Earth to support life
- Corrections cost enormous sums of money
- Corrections threaten entire industries and nations

Background

- Potential bias in the data
 - Independent databases, such as the World Bank, are “centrist”?
 - NOAA or Berkeley Earth, given their sources of funding, are biased towards data evincing global warming?
 - China, Russia, the Middle East, whose economies wholly depend on the prolonged sustenance of oil, tend towards conservative climate estimates?
- So what is the truth?
- And who watches the watchers?

Sally the Scientist, Activist

- Concerned about climate change
- Professional scientist, engineer, or mathematician
- Self exploration
- Novice with computer languages
- Novice with web service technology




Research Questions


- What is the average global temperature from a certain agency?
- How does the estimates from various agencies compare?
- How do I add another agency's estimate?
- What other quantities correlate to the average global temperature?

Data Sources

- NOAA (<https://www.ncdc.noaa.gov/cag/global/time-series>)
- Berkeley Earth (<http://berkeleyearth.org>)
- The World Bank (<http://data.worldbank.org>)
- Scripps Institute of Oceanography (<https://scripps.ucsd.edu/>)

National Oceanic and Atmospheric Administration (NOAA)

**NOAA** NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION



Home Climate Information Data Access Customer Support Contact About

Search

Home > Climate Monitoring > Climate at a Glance

May US Release: Wed, 6 Jun 2018, 11:00 AM EDT

Climate at a Glance

Climate Monitoring

State of the Climate

Temp, Precip, and Drought

Climate at a Glance

Extremes

Societal Impacts

Snow and Ice

Teleconnections

GHCN Monthly

Monitoring References

Global

National

Regional

Statewide

Divisional

City

Mapping

Time Series

Data Information

Background

Global Time Series

Choose from the options below and click "Plot" to create a time series of surface temperature anomalies.

Please note, global and hemispheric anomalies are with respect to the 20th century average. Continental anomalies are with respect to the 1910 to 2000 average. Coordinate anomalies are with respect to the 1981 to 2010 average.

Timescale: 1-Month

Month: April

Start Year: 1880

End Year: 2018

Region/
Continent: Global

Latitude: 0.0 Longitude: 0.0
-90.0 to 90.0 -180.0 to 180.0

Surface: Land and Ocean

Plot

Options

☐ Display Trend

☒ per Decade ☐ per Century

Start: 1880 End: 2018

National Oceanic and Atmospheric Administration (NOAA)

- Scientific agency within the United States Department of Commerce
- Historic record goes back to 1850
- Provided calculated global average temperature anomalies
 - Earth divided into blocks <http://data.worldbank.org>
 - Computed average temperature for each block
 - Normalized to the 20th century average
 - Calculated estimates provided monthly
- Zhang, H.-M., B. Huang, J. Lawrimore, M. Menne, Thomas M. Smith, NOAA Global Surface Temperature Dataset (NOAAGlobalTemp), Version 4.0 NOAA Global Surface Temperature Data. NOAA National Centers for Environmental Information. doi:10.7289/V5FN144H.

Berkeley Earth

2017 - 3rd Hottest Year on Record. Read our summary here.

BERKELEY EARTH™

A Measured Approach:
CLIMATE SCIENCE + STRATEGIC ANALYSIS

[About Us](#) · [Air Pollution](#) · [Data](#) · [Findings](#) · [Press & Opinions](#)

Berkeley Earth An independent non-profit

Berkeley Earth was created to address potential biases in the land surface temperature record. We are now expanding scientific investigations, educating and communicating about climate

Land + Ocean Data Set

Summary of Findings

Air Pollution

Result

Independent

Berkeley Earth is independently operated

Transparent

We continue to lower the barriers to entry into climate

Complete

Berkeley Earth took a data-driven approach to

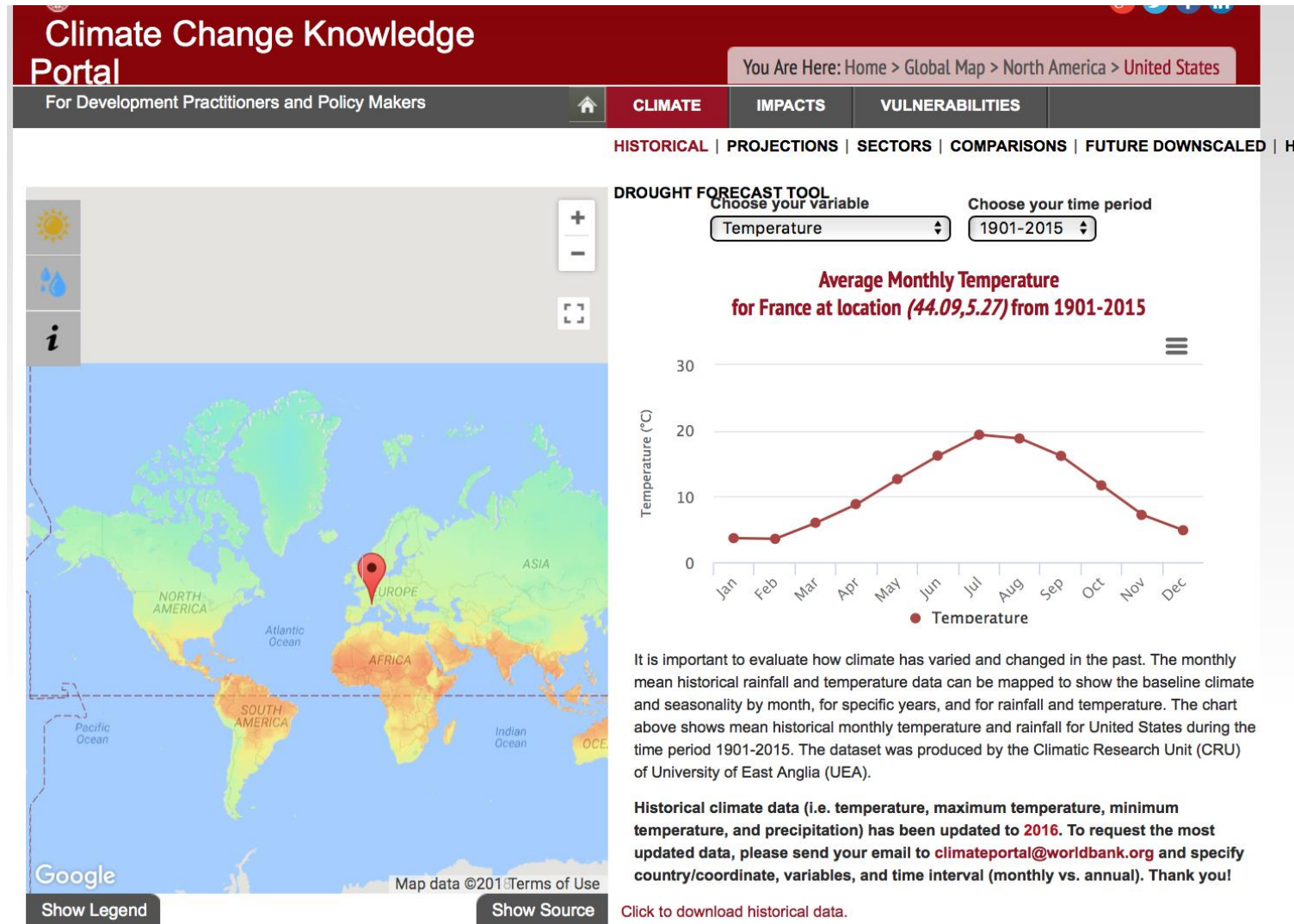
Blog

Horrible Air Pollution in Europe Reaches 7

Berkeley Earth

- Berkeley Earth is independently operated, funded primarily by unrestricted educational grants
- Mission: To pursue objectivity without concern for policies of government, industry or philanthropic ventures
- Provides data repositories for various climate and environmental metrics

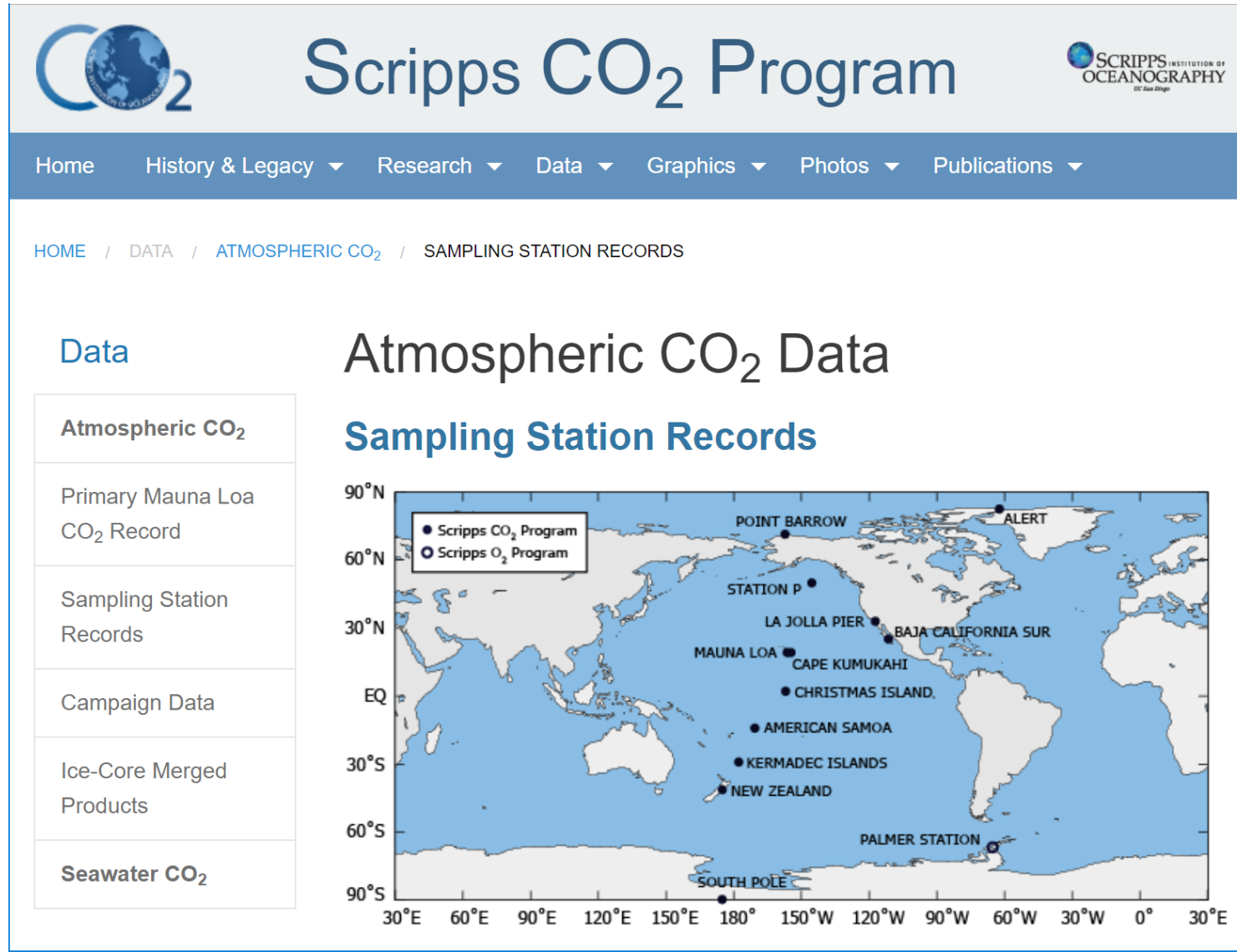
The World Bank Group



The World Bank Group

- Multinational partnership amongst 170 countries worldwide
- Mission: Seeking to improve global welfare
- Provides data repositories for various things:
 - Population
 - Health
 - Social demographics
 - Economics
 - Education
 - Climate data!!

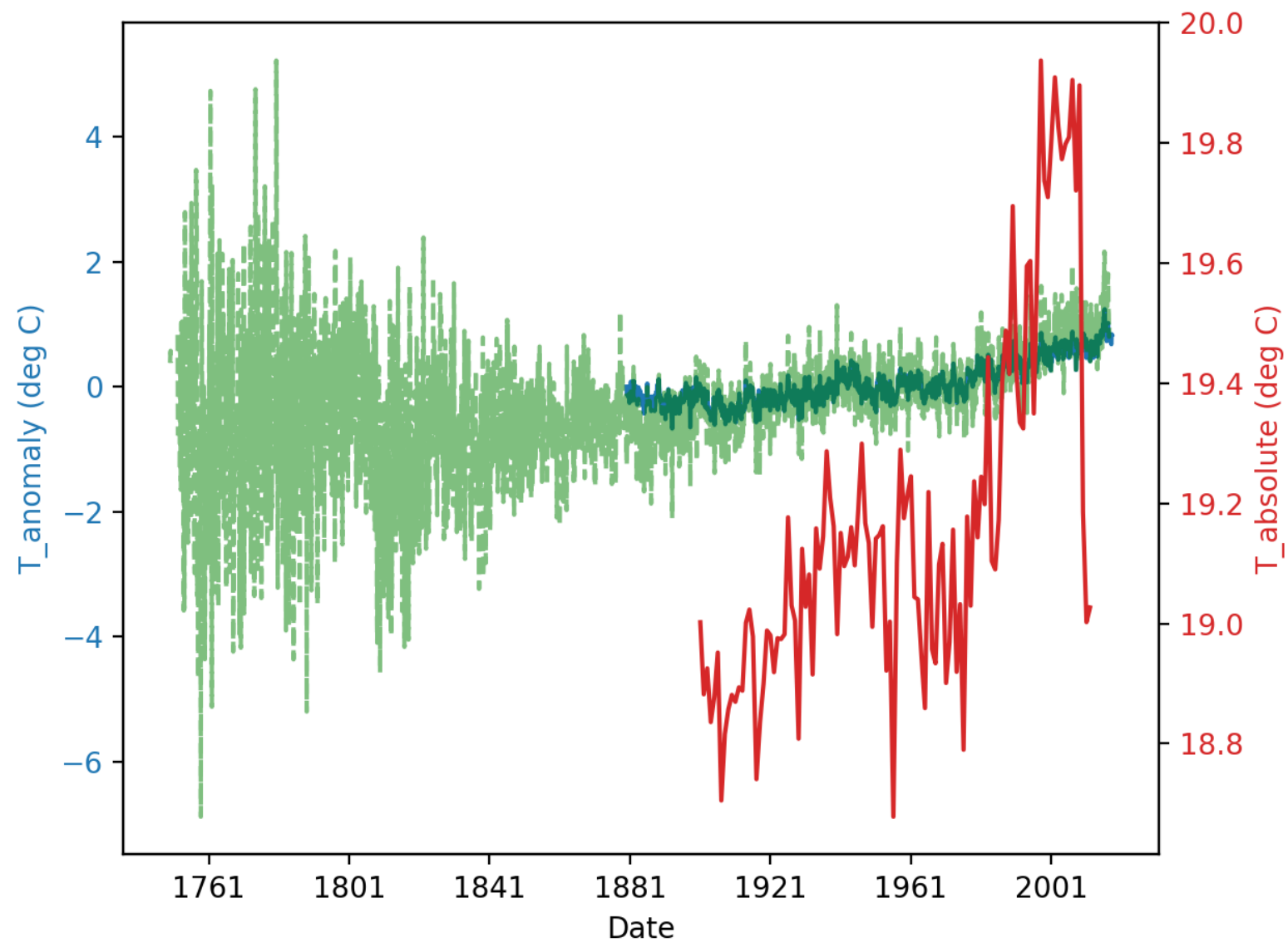
Scripps Institute of Oceanography



Scripps Institute of Oceanography

- The Scripps Institute for Oceanography is a heralded forum for carbon dioxide research
- Founder, Charles Keeling, considered first to alert about potential risks of greenhouse gas emissions
- Operates has several sampling stations in the western hemisphere to gather CO₂ data in parts per million

Demo



Design

- Created functions to communicate with each of data sources with minimal functional signature
 - Use Pandas dataframe to ease data formatting
- Created script to combine the average global temperature estimates into single graph
- Encapsulated all functionality into Jupyter notebook for easy of use and expansion

grab_noaa

```
1  # -*- coding: utf-8 -*-
2  """grab_noaa retrieves global average temperatures from NOAA.
3
4  This python module contains a single function, grab_NOAA, that
5  retrieves the global average temperature anomaly estimates from NOAA. This
6  is the climate data that is used to evaluate global climate change.
7
8  Syntax
9  import grab_noaa
10 df_noaa = grab_noaa.grab_noaa()
11
12 An overview of the data access available from NOAA is at:
13 https://www.ncdc.noaa.gov/cag/global/time-series
14
15 The data is provided as a delimited, plain text webpage that can be easily
16 read by various text interpreters. This module uses the Pandas package and
17 the Pandas.read_csv function to retrieve the data into a Pandas DataFrame.
18
19 Written by Todd Schultz
20 2018
21 """
22
23 import datetime
24 import pandas as pd
25
26
27 def grab_noaa():
28     """Retrieves global average temperatures from NOAA.
29     Inputs
30     None required
31     Outputs
```

grab_worldbank

```
def grab_worldbank(start_date = 1901, end_date = 2012):  
    """Returns a dataframe of (Year, GlobalAverageTemperature) tuples with data  
    from the WorldBank database. https://data.worldbank.org/topic/climate-change  
    Args:  
        start_date (int): Starting year for data retrieval; minimum 1901. Defaults to 1901  
        end_date (int): End year for data retrieval; maximum 2012. Defaults to 2012  
    Returns:  
        pandas dataframe: Dataframe pointing to the results from the worldbank  
                           Columns are of type Date (yyyy-mm-dd string); Tabsolute_C (float)  
                           NOTE: January 1st chosen as a dummy month-date for each year
```

Examples:

```
>>> df = grab_worldbank()  
>>> print(df.head())  
   Date      Tabsolute_C  
0  1901-01-01    19.002034  
1  1902-01-01    18.882094  
2  1903-01-01    18.925365  
3  1904-01-01    18.835930  
4  1905-01-01    18.877793  
  
>>> df = grab_worldbank(2011,2012)  
>>> print(df.head())  
   Date      Tabsolute_C  
0  2011-01-01    19.002201  
1  2012-01-01    19.026535
```

"""

grab_scripps_co2_data

```
def grab_scripps_co2_data():
    """Returns a dataframe of (Year, Mean CO2 Level (ppm)) tuples with data
    from the Scripps Institute sampling stations.
    Note: Scripps uses the mnemonic '-99.99' to represent missing data!!

    Args:
        None

    Returns:
        pandas dataframe: Dataframe pointing to the CO2 measurement per annum

    Examples:
        >>> df = grab_co2_scripps()
        >>> print(df.head())

           Date      CO2
0    1957-01-01  313.625000
1    1958-01-01  314.706667
2    1959-01-01  315.635000
3    1960-01-01  316.616667
4    1961-01-01  317.387500
5    1962-01-01  317.364444
    """
    # url links to data
    links = ['http://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/flask_co2/monthly/
    monthly_flask_co2_alt.csv',
             'http://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/flask_co2/monthly/
    monthly_flask_co2_ljo.csv',
             'http://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/flask_isotopic/monthly/
    monthly_flask_co2_isotopic.csv']
```

grab_berkeley

```
27 def grab_berkeley():
28     """
29     Returns a dataframe of (Date, Tanomaly_C) tuples with data
30     from the Berkeley Earth.
31     Note: Temperatures are in Celsius and reported as anomalies
32           relative to the Jan 1951-Dec 1980 average.
33           Estimated Jan 1951-Dec 1980 absolute temperature (C): 8.64
34
35     Args:
36         None
37     Returns:
38         pandas dataframe: Dataframe pointing to the temperature measurement on monthly basis
39     Examples:
40         >>> df_Berekely = grab_berkeley()
41         >>> print(df_Berekely.head())
42             Date  Tanomaly_C
43         0  1750-1-01      0.382
44         1  1750-2-01      0.539
45         2  1750-3-01      0.574
46         3  1750-4-01      0.382
47         4  1750-5-01       NaN
48
49     """
50     url = "http://berkeleyearth.1bl.gov/auto/Global/Complete_TAVG_complete.txt"
51     df_berkeley = pd.read_csv(url, delim_whitespace=True, index_col=None, skiprows=34, header=None, lineterminator='\n')
```

Project Structure

- <https://github.com/toddschultz/DegreesOfClimateChange>

The screenshot shows the GitHub repository page for 'DegreesOfClimateChange' by user 'toddschultz'. The repository has 220 commits, 1 branch, 0 releases, 3 contributors, and is licensed under MIT. The file list shows various directories and files, including .github, DegreesOfClimateChange, doc, examples, .coveragerc, .gitignore, .travis.yml, LICENSE, README.md, requirements.txt, requirements2.txt, and setup.py. The commit history shows recent updates to the project structure and documentation.

```
.
├── DegreesOfClimateChange
│   ├── __init__.py
│   ├── __pycache__
│   │   ├── grab_berkeley.cpython-36.pyc
│   │   ├── grab_co2_scripps.cpython-36.pyc
│   │   ├── grab_noaa.cpython-36.pyc
│   │   └── grab_worldbank.cpython-36.pyc
│   ├── country_iso_codes.csv
│   ├── doe.json
│   ├── grab_berkeley.py
│   ├── grab_co2_scripps.py
│   ├── grab_noaa.py
│   ├── grab_worldbank.py
│   ├── tests
│   │   ├── __init__.py
│   │   ├── test_grab_berkeley.py
│   │   ├── test_grab_co2_scripps.py
│   │   ├── test_grab_noaa.py
│   │   └── test_grab_worldbank.py
│   └── version.py
├── LICENSE
├── README.md
├── doc
│   ├── Components.md
│   ├── DataSources.txt
│   ├── DegreesOfClimateChangeSummary.pptx
│   ├── GithubStarCount.png
│   ├── Ideas.md
│   ├── TechnologyReview.md
│   ├── TechnologyReview.pptx
│   ├── UseCases.md
│   ├── UserPersonas.md
│   ├── Web-scraping\ Comparison\ Example.ipynb
│   ├── design
│   └── designworkflow.md
├── examples
│   ├── GlobalTemperatureComparison-Final.ipynb
│   ├── __init__.py
│   ├── berkeleyearth.ipynb
│   ├── co2_data_retrieval.ipynb
│   ├── grab_WorldBank_Module.ipynb
│   └── noaa_co2_data.ipynb
├── requirements.txt
└── setup.py
```

6 directories, 44 files

Lessons Learned

- Data source providers don't provide sufficient documentation on their web service APIs
- Surprised by the difference in temperature metric between the agencies
- Setting up imports to work with modules in different folders can be challenging
- Setting up imports to work with Travis CI can be challenging

Future Work

- Additional data sources
 - Increase geopolitical diversity of data source origins
 - Increase time scale of climate data
 - Geological time scale very large (~millions of years)
 - Add potential causal data sources
 - CO2 emissions
 - Solar activity
- Improve data handling and visualizations
 - Find consistent estimates of the temperature anomaly instead of mixing deviations with absolute temperature values

Thanks