

team was helping the City of Cincinnati in Ohio to improve their Emergency

predictor in a model to predict level of emergency response throughout the city. We'll be getting our weather data from the most reliable source, NOAA (National Oceanic and Atmospheric Administration). Since I now live in beautiful Long Beach, California, let's guide our exploration

by trying to answer the following question: How has average daily temperature in Long Beach changed since 2015? Let's get started!

1. Request an Access Token from NOAA In order to request data from NOAA, we will need to request an access token, which basically affirms that we are a trusted source asking for data. Navigate to

. . .

For more information about CDO Web Services read the documentation for CDO

Request Web Services Token

To gain access to NCDC CDO Web Services, register with your email address. An email will be sent with a unique token which will allow access RESTful services.

Long Beach, CA, USA

Select Dataset

Daily Summaries

Select Date Range

Data Categories

Evaporation

Land

Air Temperature

2015-01-01 to 2019-06-01

the Request a Token website.

Web Services guide.

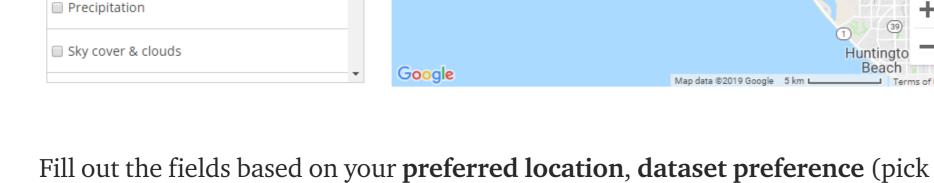
Please enter your email address

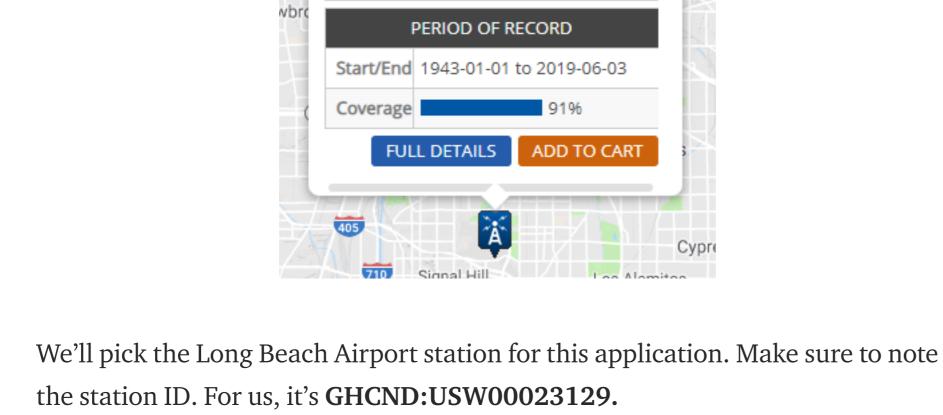
The next thing you'll probably want to do is figure out which weather station to gather data from. The easiest way to find this station would be to navigate to the Find a Station website **Enter Location** Santa Fe South Whittier (72)

Torrance

Lakewood

Long Beach





4 #store the data we get as a dataframe 5 import pandas as pd #convert the response as a strcuctured json 8 import json

```
Token = 'YOUR_ACCESS_TOKEN'
18
   #Long Beach Airport station
    station_id = 'GHCND:USW00023129'
```

15

import numpy as np

3. Grab the Data!

Python's *requests* library for this task.

#mathematical operations on lists

from datetime import datetime

weather_imports.py hosted with ♥ by GitHub

#parse the datetimes we get from NOAA

#add the access token you got from NOAA

First let's import the needed libraries in Python.

```
In order to request data from NOAA we need to use NOAA's API (Application
communicate with NOAA to ask for data.
Let's take a look at the anatomy of a NOAA API request.
```

is GHCND:USW00023129 • startdate and enddate, specifying the date range we wish to get data from. We will be calling the API once for each year because 1000 items is not enough for our full 3.5 year date range from 2015 until now.

In all, our first request for all data in 2015 will look like this:

SW00023129&startdate = 2015-01-01&enddate = 2015-12-31

https://www.ncdc.noaa.gov/cdo-web/api/v2/data?

5 prcp = []

11

12

13

14

7 #for each year from 2015-2019 ...

print('working on year '+year)

#load the api response as a json

8 for year in range(2015, 2020):

#make the api call

#initialize dataframe

2 df_temp = pd.DataFrame()

year = str(year)

• limit, specifying the maximum number of items to include in the response.

• stationid, specifying which station(s) we would like data from. Our station id

The default is 25 and the maximum is 1000, which is what we will set.

1 #initialize lists to store data 2 dates_temp = [] 3 dates_prcp = []

17 18 #get the date field from all average temperature readings 19 dates_temp += [item['date'] for item in avg_temps] 20 temps += [item['value'] for item in avg_temps] 21

And, now the code to store the collected data as a dataframe:

```
Jan 2015
                             Jul 2016
                                               Jul 2017
                                                                 Jul 2018
                                                                                   Jul 2019
                                                                          Jan 2019
Nice! So we successfully requested, processed, and reported on the average daily
temperature in Long Beach, California since January 2015.
In truth, there are a ton more options built into the NOAA API. Check out the <u>full</u>
documentation. I figured it would a good start to walk through one complete
```

codes. Connor Shorten · Jun 10, 2019

Benny Friedman · Jun 10, 2019 **Creating a TensorFlow CNN in C++ (Part 2)** In this post I'll show how to create, train and test a Convolutional Neural Network using the TensorFlow C++ API — Background In the last two

Ashutosh Tripathi · Jun 10, 2019

Gender gap in IT

Ramya Subramanya

Leveraging data: 7 Factors for

Opportunities in the Mining...

Maximizing Digital

defaulter

Gojek

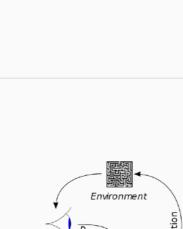
IFC

Get this newsletter

More from Towards Data Science

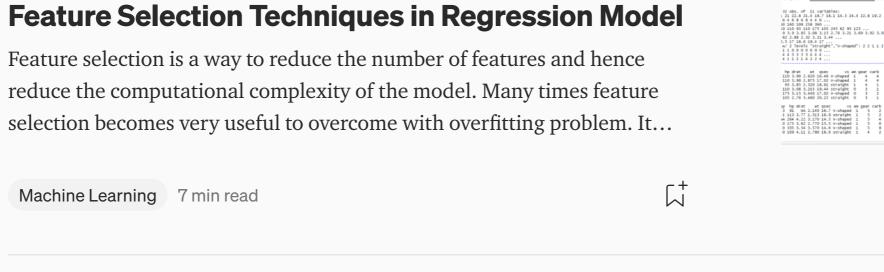
ubiquitous but has diminishing marginal returns. More data does not necessarily translate into new information, in fact, it may sometimes only... Big Data 2 min read

Share your ideas with millions of readers.



Follow

Machine Learning 6 min read



Recommended from Medium

4dresults live **How to predict 4d numbers** accurately in Malaysia erin malone Using Mapping to Scope a

Get started Sign In Q Search

Ritvik Kharkar 602 Followers Follow

Software Engineer — Mathemagician — Home Chef More from Medium

Angelica Lo D... in Towards Data Sci... **Model Evaluation in Scikit-learn** Pankaj Ku... in Python in Plain Engl... Analysis of the arXiv Papers for a **Topic Using the arXiv API**

Angelica Lo D... in Towards Data Sci...

Is a Small Dataset Risky? Jake in Towards Data Science **Recursion for Data Scientists**

Help Status Writers Blog Careers Privacy Terms About

I personally found the need to access weather data during a project where my Medical Response system. My team needed to use daily weather data as a

 \times SUBMIT Request an Access Token from NOAA Fill in an email address and you'll shortly get an email with your access token. Store this somewhere safe and keep it *private*. Otherwise, others could make requests to NOAA on your behalf and you don't necessarily want that. 2. Figure Out Which Weather Station to Access

fewer stations will have the entirety of that data), and data categories (we care just about air temperature for the task at hand). The map on the right of the page will auto-update with available stations. Aligeles Montebello STATION DETAILS Name LONG BEACH DAUGHERTY

AIRPORT, CA US

Lat/Lon 33.8116, -118.1463

ID GHCND:USW00023129

"Daily Summaries" for daily data), time range (the longer range you pick the

#needed to make web requests 2 import requests

Now that we have our access token and we know the ID of the station we wish to

access, all that is left to do is request the data from NOAA. We'll be using

Programming Interface) which is basically just the specific language to The base request looks like: https://www.ncdc.noaa.gov/cdo-web/api/v2/data? After the question mark, '?', we will need to put all the options specifying the exact data we are looking for. These will include: • datasetid, which for us will be GHCND (Global Historical Climatology Network Daily). • datatypeid, which is a list of the variables we are after. For us, that will be **TAVG** (average temperature)

view raw

Let's take a look at the code to make the API calls:

datasetid = GHCND&datatypeid = TAVG&limit = 1000&stationid = GHCND:U

d = json.loads(r.text) 15 #get all items in the response which are average temperature readings 16 avg_temps = [item for item in d['results'] if item['datatype']=='TAVG'] #get the actual average temperature from all average temperature readings get_weather_data.py hosted with ♥ by GitHub view raw

r = requests.get('https://www.ncdc.noaa.gov/cdo-web/api/v2/data?datasetid=GHCND&dat

#populate date and average temperature fields (cast string date to datetime and convert df_temp['date'] = [datetime.strptime(d, "%Y-%m-%dT%H:%M:%S") for d in dates_temp] 6 $df_{emp}['avgTemp'] = [float(v)/10.0*1.8 + 32 for v in temps]$ store_weather_data.py hosted with ♥ by GitHub view raw A chart made from the resulting dataframe looks like this: Average Daily 52 09 59 04

specific use case of the NOAA API. ~ Happy data gathering!:) 290 | 7 290 🔾 7 Sign up for The Variable By Towards Data Science Every Thursday, the Variable delivers the very best of Towards Data Science: from handson tutorials and cutting-edge research to original features you don't want to miss. Take a look.

Your home for data science. A Medium publication sharing concepts, ideas and **Deep Compression** In their current form, Deep Neural Networks require enormous memory to fund their massive over-parameterization. Classic Neural Networks such as AlexNet and VGG-16 require around 240 and 552 MB, respectively. Many... Deep Learning 3 min read

Write on Medium

years, Google's TensorFlow has been gaining popularity. It is by far the m... Machine Learning 13 min read Prajakta KN · Jun 10, 2019 ★ The next frontier in Big Data Analytics Cross-industry analytics and its three pillars — In our world, data is

Ryan Gross · Jun 10, 2019 ★ Reinforcement learning is going mainstream. Here's what to expect. The same tech that's beating world champions in games will soon revolutionize anything that can be simulated (and because everything is...

selection becomes very useful to overcome with overfitting problem. It... Machine Learning 7 min read Read more from Towards Data Science

Predicting Personal Loan System Sabita Langkam Lawrence Kurnia... in Gojek Produ... **Descriptive Statistics for Data Efficient Experimentation at Science Part-2**

> **Misinterpretations and Legends About Sadness**

Additor