

Source the Vancouver Permitting Data for Scraping

Date : 23rd March 2023

Author : Parul Sharma

Overview : Our objective is to retrieve the “records” key using GET API and then automate the process using python and the libraries we are thinking to use are requests(for API interaction), json(read response), pandas and csv (output file).

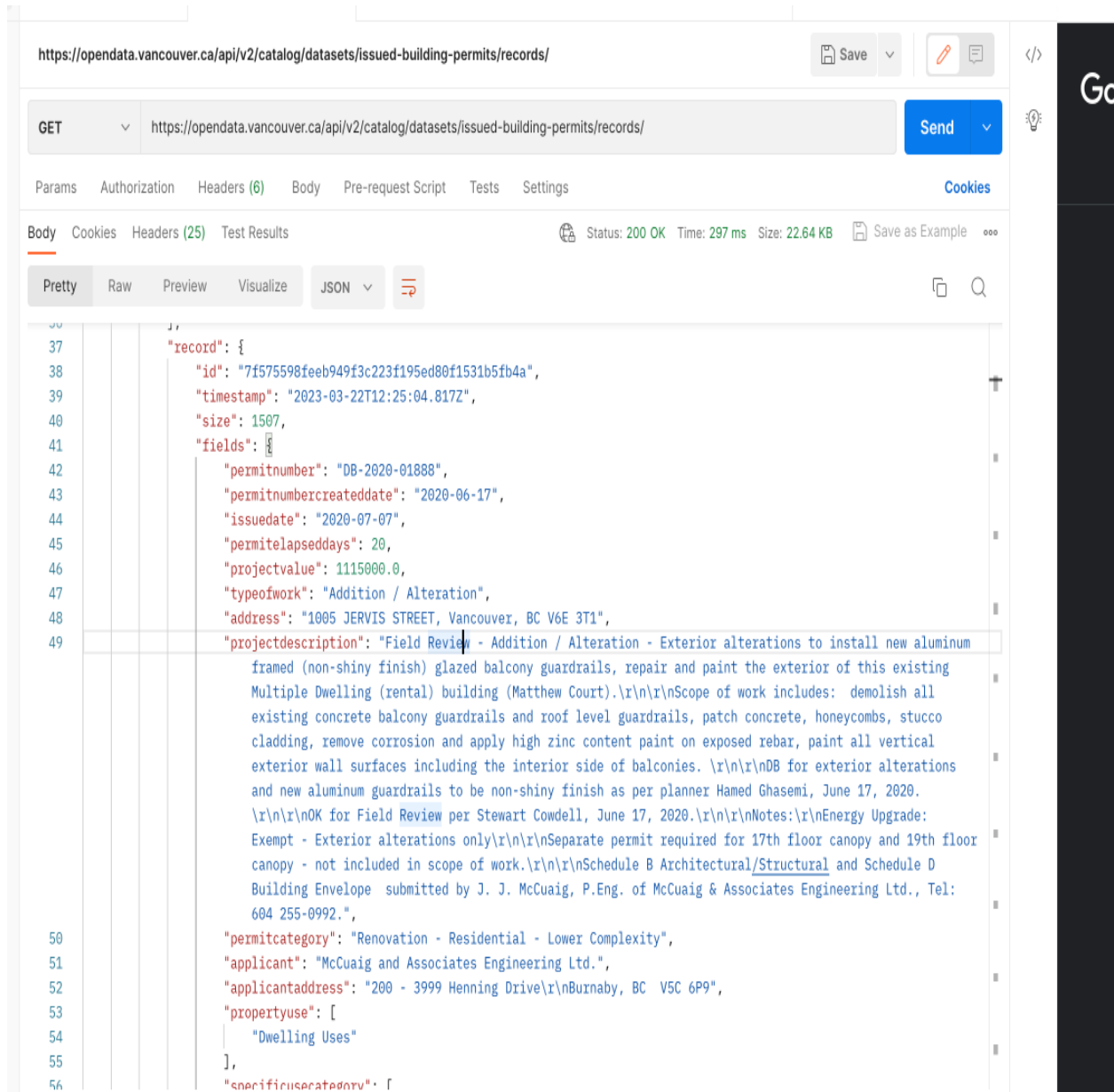
One of the data source that we can use is	DataSource
	Full API console
	API documentation

Step 1) To retrieve the records, we fired “GET” command in postman and the output is shown below :-

The screenshot displays the Postman interface with a GET request to the URL `https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-building-permits/records/`. The response is a JSON object with the following structure:

```
21  "records": [  
22    {  
23      "links": [  
24        {  
25          "rel": "self",  
26          "href": "https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-building-permits/records/  
27                    7f575598feeb949f3c223f195ed80f1531b5fb4a"  
28        },  
29        {  
30          "rel": "datasets",  
31          "href": "https://opendata.vancouver.ca/api/v2/catalog/datasets"  
32        },  
33        {  
34          "rel": "dataset",  
35          "href": "https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-building-permits"  
36        }  
37      ],  
38      "record": {  
39        "id": "7f575598feeb949f3c223f195ed80f1531b5fb4a",  
40        "timestamp": "2023-03-22T12:25:04.817Z",  
41        "size": 1507,  
42        "fields": {  
43          "permitnumber": "DB-2020-01888",  
44          "permitnumbercreateddate": "2020-06-17",  
45          "issuedate": "2020-07-07",  
46          "permitlapseddays": 20,  
47          "projectvalue": 1115000.0,  
48          "typeofwork": "Addition / Alteration",  
49          "address": "1005 JERVIS STREET, Vancouver, BC V6E 3T1",  
          "projectdescription": "Field Review - Addition / Alteration - Exterior alterations to install new aluminum  
framed (non-shiny finish) glazed balcony guardrails, repair and paint the exterior of this existing
```

The “records” key contains a nested json and the required fields like “permitnumber”, “permitdescription”, “applicant” are keys of “fields”, key of “record” of “records” of “issued-building-permits” dataset id of “datasets”. We need to flatten the nested json using python in order to extract the required data fields.

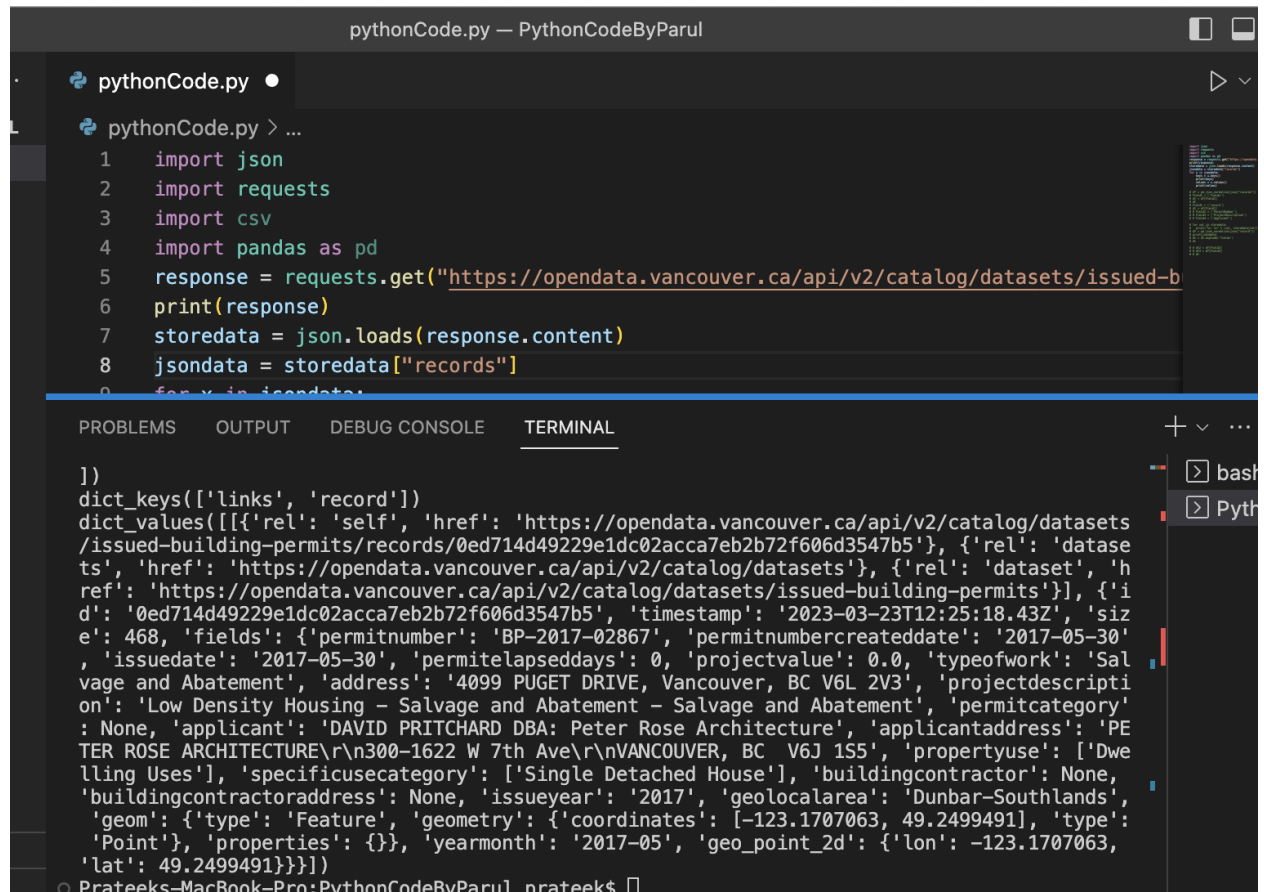


The screenshot shows a REST client interface with a GET request to the URL `https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-building-permits/records/`. The response is a JSON object with the following structure:

```
{
  "record": {
    "id": "7f575598feeb949f3c223f195ed80f1531b5fb4a",
    "timestamp": "2023-03-22T12:25:04.817Z",
    "size": 1507,
    "fields": {
      "permitnumber": "DB-2020-01888",
      "permitnumbercreateddate": "2020-06-17",
      "issuedate": "2020-07-07",
      "permitelapsedsdays": 20,
      "projectvalue": 1115000.0,
      "typeofwork": "Addition / Alteration",
      "address": "1005 JERVIS STREET, Vancouver, BC V6E 3T1",
      "projectdescription": "Field Review - Addition / Alteration - Exterior alterations to install new aluminum framed (non-shiny finish) glazed balcony guardrails, repair and paint the exterior of this existing Multiple Dwelling (rental) building (Matthew Court).\r\n\r\nScope of work includes: demolish all existing concrete balcony guardrails and roof level guardrails, patch concrete, honeycombs, stucco cladding, remove corrosion and apply high zinc content paint on exposed rebar, paint all vertical exterior wall surfaces including the interior side of balconies. \r\n\r\n\r\nDB for exterior alterations and new aluminum guardrails to be non-shiny finish as per planner Hamed Ghasemi, June 17, 2020. \r\n\r\n\r\nOK for Field Review per Stewart Cowdell, June 17, 2020.\r\n\r\n\r\nNotes:\r\n\r\nEnergy Upgrade: Exempt - Exterior alterations only\r\n\r\n\r\nSeparate permit required for 17th floor canopy and 19th floor canopy - not included in scope of work.\r\n\r\n\r\nSchedule B Architectural/Structural and Schedule D Building Envelope submitted by J. J. McCuaig, P.Eng. of McCuaig & Associates Engineering Ltd., Tel: 604 255-0992.",
      "permitcategory": "Renovation - Residential - Lower Complexity",
      "applicant": "McCuaig and Associates Engineering Ltd.",
      "applicantaddress": "200 - 3999 Henning Drive\r\n\r\nBurnaby, BC V5C 6P9",
      "propertyuse": [
        "Dwelling Uses"
      ],
      "specificusecategory": [

```

Step 2) Snapshot (below) of the automation we can do using python. The python library “requests” helps to interact with the API and the result is stored in a variable “response”. Using “response.content” we can store our data, using the json library and json.loads(), in a variable “storedata”. We can also flatten the json data and store it in pandas dataframe and write/export it in a csv file resulting in our required data fields.



The image shows a code editor window titled "pythonCode.py — PythonCodeByParul" with a Python script and a terminal window below it.

pythonCode.py

```
1 import json
2 import requests
3 import csv
4 import pandas as pd
5 response = requests.get("https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-b
6 print(response)
7 storedata = json.loads(response.content)
8 jsondata = storedata["records"]
9 for x in jsondata:
```

TERMINAL

```
dict_keys(['links', 'record'])
dict_values([{'rel': 'self', 'href': 'https://opendata.vancouver.ca/api/v2/catalog/datasets
/issued-building-permits/records/0ed714d49229e1dc02acca7eb2b72f606d3547b5'}, {'rel': 'dataset', 'h
ref': 'https://opendata.vancouver.ca/api/v2/catalog/datasets'}, {'rel': 'dataset', 'h
ref': 'https://opendata.vancouver.ca/api/v2/catalog/datasets/issued-building-permits'}], {'id': '0ed714d49229e1dc02acca7eb2b72f606d3547b5', 'timestamp': '2023-03-23T12:25:18.43Z', 'size': 468, 'fields': {'permitnumber': 'BP-2017-02867', 'permitnumbercreateddate': '2017-05-30', 'issuedate': '2017-05-30', 'permitlapseddays': 0, 'projectvalue': 0.0, 'typeofwork': 'Salvage and Abatement', 'address': '4099 PUGET DRIVE, Vancouver, BC V6L 2V3', 'projectdescription': 'Low Density Housing - Salvage and Abatement - Salvage and Abatement', 'permitcategory': None, 'applicant': 'DAVID PRITCHARD DBA: Peter Rose Architecture', 'applicantaddress': 'PETER ROSE ARCHITECTURE\\r\\n300-1622 W 7th Ave\\r\\nVANCOUVER, BC V6J 1S5', 'propertyuse': ['Dwelling Uses'], 'specificusecategory': ['Single Detached House'], 'buildingcontractor': None, 'buildingcontractoraddress': None, 'issueyear': '2017', 'geolocalarea': 'Dunbar-Southlands', 'geom': {'type': 'Feature', 'geometry': {'coordinates': [-123.1707063, 49.2499491], 'type': 'Point'}, 'properties': {}}, 'yearmonth': '2017-05', 'geo_point_2d': {'lon': -123.1707063, 'lat': 49.2499491}}})
Prateeks-MacBook-Pro:PythonCodeByParul prateek$
```