

Assignment-2

Specification

Make Submission

Check Submission

Collect Submission

Data Service for World Bank Economic Indicators

In this assignment, you are asked to develop a Flask-Restplus data service that allows a client to read and store some publicly available economic indicator data for countries around the world, and allow the consumers to access the data through a REST API.

IMPORTANT : For this assignment , you will be asked to access the given Web content programmatically. Some Web hosts do not allow their web pages to be accessed programmatically, some hosts may block your IP if you access their pages too many times. During the implementation, download a few test pages and access the content locally - try not to perform too many programmatically.

- The source URL: <http://api.worldbank.org/v2/> (<http://api.worldbank.org/v2/>)
- (<http://api.worldbank.org/v2/>) Documentations on API Call Structure:
(<https://datahelpdesk.worldbank.org/knowledgebase/articles/898581-api-basic-call-structure>)
<https://datahelpdesk.worldbank.org/knowledgebase/articles/898581-api-basic-call-structure>
(<https://datahelpdesk.worldbank.org/knowledgebase/articles/898581-api-basic-call-structure>)

(<https://datahelpdesk.worldbank.org/knowledgebase/articles/898581-api-basic-call-structure>)

The World Bank indicators API provides 50 year of data over more than 1500 indicators for countries around the world.

- List of indicators can be found at: <http://api.worldbank.org/v2/indicators> (<http://api.worldbank.org/v2/indicators>)
- List of countries can be found at: (<http://api.worldbank.org/v2/countries>) <http://api.worldbank.org/v2/countries>
(<http://api.worldbank.org/v2/countries>)

As the documentation shows , you can construct URLs specifying different parameters to acquire necessary data. Use a custom URL to get information about a specific indicator. For example, the data on the GDP of all countries from 2012 to 2017 can be acquired via this URL: (<http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017>) [http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?](http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017)
[date=2012:2017&format=json&per_page=1000](http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000) ([http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?](http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000)
[date=2012:2017&format=json&per_page=1000](http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000))

For this assignment we are interested in annual values of a user specified indicator from 2012 to 2017 for one or all countries. The content of the file is quite straightforward. The data service specification will ask you to 'import' this data into a 'data storage'. You should inspect the content of the file carefully and decide on a data model and storage method. You will find that a JSON format is suitable to accessing data and publishing it.

Assignment Specification

The data service should use JSON format to publish its data and implement following operations.

Question-1: Import a collection from the data service (2.5 marks)

This operation can be considered as an on-demand 'import' operation. The service will download the JSON data for all countries (http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000) respective to the year 2012 to 2017 and identified by the indicator id given by the user and process the content into an internal data format and store it in the database; use **sqlite** for storing the data (the name of the database should be **YOURZID.db**) .

For example, the following link can be used to obtain data for the indicator called: **NY.GDP.MKTP.CD** (http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000) [http://api.worldbank.org/v2/countries/all/indicators/ **NY.GDP.MKTP.CD** ?date=2012:2017&format=json&per_page=1000](http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000) (http://api.worldbank.org/v2/countries/all/indicators/NY.GDP.MKTP.CD?date=2012:2017&format=json&per_page=1000)

To import data you require an indicator as a parameter given by the API user. The parameter should be a query parameter named:

- **indicator_id** : an indicator (<http://api.worldbank.org/v2/indicators>) <http://api.worldbank.org/v2/indicators> (<http://api.worldbank.org/v2/indicators>)

After importing the collection, the service should return a response containing at least the following information:

- **uri** : the URL with which the imported collection can be retrieved
- **id** : a unique integer identifier automatically generated
- **creation_time** : the time the collection stored in the database

Example:

```
HTTP operation: POST /collections?indicator_id=NY.GDP.MKTP.CD
```

Returns: 201 Created

```
{
  "uri" : "/collections/1",
  "id" : 1,
  "creation_time": "2019-04-08T12:06:11Z",
  "indicator_id" : "NY.GDP.MKTP.CD"
}
```

- You should return appropriate responses (JSON formatted) in case of already imported collections, invalid indicators or any invalid attempts to use the endpoint (e.g. If the input indicator id doesn't exist in the data source, return error 404)
- **UPDATE: Please remove those entries where the value is NONE**

Question 2- Deleting a collection with the data service (2.5 marks)

This operation deletes an existing collection from the database. The interface should look like as below:

```
HTTP operation: DELETE /collections/{id}
```

Returns: 200 OK

```
{
  "message" : "The collection 1 was removed from the database!",
  "id": 1
}
```

Question 3 - Retrieve the list of available collections (2.5 marks)

This operation retrieves all available collections. The interface should look like as like below:

```
HTTP operation: GET /collections?order_by={+id,+creation_time,+indicator,-id,-creation_time,-ind
```

"order_by" is a comma separated string value to sort the collection based on the given criteria. Each segment of this value indicates how the collection should be sorted, and it consists of two parts (+ or -, and the name of column e.g., id). In each segment, + indicates ascending order, and - indicates descending order. Here are some samples:

+id,+creation_time	order by "id ascending" and then "creation_time ascending" In this case sorting by "id" has priority over "creation_time "
--------------------	---

-creation_time	order by "creation_time descending"
----------------	-------------------------------------

Returns: 200 OK

```
[
  {
    "uri" : "/collections/1",
    "id" : 1,
    "creation_time": "2019-04-08T12:06:11Z",
    "indicator" : "NY.GDP.MKTP.CD"
  },
  {
    "uri" : "/collections/2",
    "id" : 2,
    "creation_time": "2019-05-08T12:16:11Z",
    "indicator" : "2.0.cov.Math.pl_3.all"
  },
  ...
]
```

Question 4 - Retrieve a collection (2.5 marks)

This operation retrieves a collection by its ID . The response of this operation will show the imported content from world bank API for all 6 years. That is, the data model that you have designed is visible here inside a JSON entry's content part.

The interface should look like as like below:

```
HTTP operation: GET /collections/{id}
```

Returns: 200 OK

```
{
  "id" : 1,
  "indicator": "NY.GDP.MKTP.CD",
  "indicator_value": "GDP (current US$)",
  "creation_time" : "2019-04-08T12:06:11Z"
  "entries" : [
    {
      "country": "Arab World", "date": 2016, "value": 2500164034395.78 },
      {"country": "Australia", "date": 2016, "value": 780016444034.00 },
      ...
    ]
  }
}
```

Question 5 - Retrieve economic indicator value for given country and a year (2.5 marks)

The interface should look like as like below:

```
HTTP operation: GET /collections/{id}/{year}/{country}
```

Returns: 200 OK

```
{
  "id": 1,
  "indicator" : "NY.GDP.MKTP.CD",
  "country": "Arab World",
  "year": 2016,
  "value": 780016444034.00
}
```

Question 6 - Retrieve top/bottom economic indicator values for a given year (2.5 marks)

The interface should look like as like below:

```
HTTP operation: GET /collections/{id}/{year}?q=<query>
```

Returns: 200 OK

```
{
  "indicator": "NY.GDP.MKTP.CD",
  "indicator_value": "GDP (current US$)",
  "entries" : [
    {
      "country": "Arab World",
      "value": 2500164034395.78
    },
    ...
  ]
}
```

The <query> is an optional integer parameter which can be either of following:

- **+N (or simply N)** : Returns top N countries sorted by indicator value (highest first)
- **-N** : Returns bottom N countries sorted by indicator value

where N can be an integer value between 1 and 100. For example, a request like " GET /collections/1 /2015?q=+10 " should returns top 10 countries according to the indicator value.

Notice:

- Your code must implemented in **flask-restplus** and automatically generate swagger doc for testing the endpoints.
- Your code must be executable in CSE machines
- Your code must not require installing any software (python libraries are fine)
- Your code must be written in Python 3.5 or newer versions.
- Your operations (API methods) must return appropriate responses in JSON format, and appropriate HTTP response code! e.g., **500 Internal Server Error** is inappropriate response!
- Make sure you are using right datatypes in the database and in you API methods (e.g., not string for years '2017')
- Install flask_restplus on cse machines by pip3 command, and make sure you install **pip3 install werkzeug~=0.16.1**

Submission:

- One and only one Python script file named " YOUR_ZID .py" which contains your code

- **How I can submit my assignment?**

Go to the assignment page click on the "Make Submission" tab; pick your files which must be named "YOUR_ZID.py". Make sure that the files are not empty, and submit the files together.

- **Can I submit my file after deadline?**

Yes, you can. But 25% of your assignment will be deducted as a late penalty per day. In other words, if you be late for more than 3 days, you will not be marked.

Resource created 5 months ago (Thursday 27 February 2020, 12:54:52 PM), last modified 4 months ago (Tuesday 31 March 2020, 10:35:01 PM).

Comments

Q (/COMP9321/20T1/forums/search?forum_choice=resource/43071) (/COMP9321/20T1/forums/resource/43071)

Add a comment



Juan Chen (/users/z5148637) 3 months ago (Wed Apr 22 2020 11:14:42 GMT+1000 (澳大利亚东部标准时间))

When can we get results of assignment2?

Reply



Rachel Hwang (/users/z5162384) 4 months ago (Sun Apr 12 2020 15:11:33 GMT+1000 (澳大利亚东部标准时间))

Out of curiosity, roughly when would we be getting the results for this assignment?

Reply



Karl Nicolas (/users/z5117434) 4 months ago (Tue Apr 14 2020 10:24:06 GMT+1000 (澳大利亚东部标准时间))

Bump

Reply



Tianyang Men (/users/z5192600) 4 months ago (Sat Apr 04 2020 18:07:58 GMT+1100 (澳大利亚东部夏令时间)), last modified 4 months ago (Sat Apr 04 2020 18:22:03 GMT+1100 (澳大利亚东部夏令时间))

Hi,

My zid is z5192600 I just got some code wrong in my assignment,so i corrected it and submit again,however it has past 24 hours,i wonder if i exceeded the deadline,if i need to be punished,i would better to use the previous version submitted at Fri Apr 3 13:14:29 2020.And i think the best situation to be punished is one day late submission because when i fixed the problem the time was 18:15 2020/4/4.Thanks a lot!

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) 4 months ago (Sat Apr 04 2020 20:23:33 GMT+1100 (澳大利亚东部夏令时间))

Hi,

As long as, if you submit late less then 24 hours, you will be penalized by 25%. If it is less than 1 hour late, you may not be penalized.

Reply



Tianyang Men (/users/z5192600) 4 months ago (Sat Apr 04 2020 20:35:55 GMT+1100 (澳大利亚东部夏令时间))

Thank you

Reply



Jingfeng Yan (/users/z5158278) [4 months ago \(Fri Apr 03 2020 18:06:15 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Hi,

I just realised that I have missed two lines of codes, could I add them back right now without late penalty?

regards,

Reply



Yixuan Ma (/users/z5194164) [4 months ago \(Fri Apr 03 2020 14:10:16 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Hi, I would like to know how the assignment will be tested and marked, because my url requests may not be exactly like the example, but it runs well on the web by typing the parameters in. So I hope it can be a manual testing option on the web.

Reply



Syd Pleno (/users/z3063250) [4 months ago \(Fri Apr 03 2020 13:20:01 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Hi,

Q3 - is the order_by parameter optional?

Thanks.

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) [4 months ago \(Fri Apr 03 2020 14:00:40 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Yes

Reply



Jackie Banh (/users/z5061994) [4 months ago \(Fri Apr 03 2020 01:32:47 GMT+1100 \(澳大利亚东部夏令时间\)\)](#), last modified [4 months ago \(Fri Apr 03 2020 01:33:39 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Should indicators being added to the collection table be unique (can't add the same collection if it is currently in the table)?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) [4 months ago \(Fri Apr 03 2020 07:42:49 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

how you manage your database is up to you

Reply



Meishi Chen (/users/z5269035) [4 months ago \(Thu Apr 02 2020 23:34:55 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Hi, should we consider the situation that the server will be restarted during the test?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) [4 months ago \(Fri Apr 03 2020 07:43:13 GMT+1100 \(澳大利亚东部夏令时间\)\)](#)

Yes! otherwise, why you are using database!

Reply



Juan Chen (/users/z5148637) 4 months ago (Fri Apr 03 2020 11:22:09 GMT+1100 (澳大利亚东部夏令时间))

if server restarts, the data already in the database should be dropped or not ?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) 4 months ago (Fri Apr 03 2020 11:29:18 GMT+1100 (澳大利亚东部夏令时间))

No, it should be there

Reply



Sophia Lev (/users/z5178140) 4 months ago (Thu Apr 02 2020 16:17:02 GMT+1100 (澳大利亚东部夏令时间))

Hi Mohammadali,

In Q4, do the "entries" have to be in one line?

Thanks!

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) 4 months ago (Thu Apr 02 2020 16:29:11 GMT+1100 (澳大利亚东部夏令时间))

No, but response should be a valid json (formatted as json, and response datatype is application/json)

Reply



Junyu Ren (/users/z5195715) 4 months ago (Thu Apr 02 2020 15:37:18 GMT+1100 (澳大利亚东部夏令时间))

Can we only focus on first 1000 records of a given indicator?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) 4 months ago (Thu Apr 02 2020 16:15:17 GMT+1100 (澳大利亚东部夏令时间))

Yes, only the first page will be enough

Reply



Lloyd Richardson (/users/z5188286) 4 months ago (Thu Apr 02 2020 16:06:56 GMT+1100 (澳大利亚东部夏令时间))

<https://webcms3.cse.unsw.edu.au/COMP9321/20T1/forums/2757580>

Reply

Load More Comments