Resources  /  Assignment-2

# Assignment-2

| Specification | Make Submission | Check Submission | Collect Submission |

# Data Service for TV Shows

In this assignment, you are asked to develop a Flask-Restx data service that allows a client to read and store some TV Shows, and allow the consumers to access the data through a REST API.

*The assignment is based on The TV Maze API, which provides a detailed list of TV shows. You can explore the TVmaze API using the following links*

- The source URL: (http://api.worldbank.org/v2/) http://api.tvmaze.com/shows (http://api.tvmaze.com/shows)
- (http://api.worldbank.org/v2/) Documentations on API Call Structure: https://www.tvmaze.com/api (https://www.tvmaze.com/api)

***Disclaimer: We are using an extremal API provided by TV Maze (https://www.tvmaze.com/). We want the students to interact with real life web services to add to the learning experience and hence we are not responsible nor liable for the wording or inclusion/exclusion of TV shows and descriptions within the API. This is a "building your REST API" exercise and should be treated as such.

In this assignment, you are going to use the information provided in this API and add a few functionalities as listed below:

# Assignment Specification

## Question-1: import a TV Show (2 marks)

This operation can be considered as an on-demand 'import' operation to get the details of a TV show and store it in your application. The service will download the JSON data for the given TV show (by its name) ; You must use **sqlite** for storing the data (the name of the database should be **YOURZID.db** ) locally after importing the TV show .

You can use the following to query the API: http://api.tvmaze.com/search/shows?q= (http://api.tvmaze.com/search/shows?q=TITLE) ?
For example, you can check the following query: http://api.tvmaze.com/search/shows?q=good%20girls (http://api.tvmaze.com/search/shows?q=good%20girls)

To import a tv show, your API accepts a query parameter called "name".:

- **name** : title for the tv show

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

- **id** : a unique integer identifier automatically generated (this might be different than tvmase_id)

- **tvmaze-id** : the id of the tv show in tvmaze API
- **last-update** : the time the collection stored in the database
- **_links** : the URL with which the imported collection can be retrieved (as shown in below example)

*Important* : *For this assignment , you are 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. Check the documentation of the tvmaze API to get insights about their rate limiting policy.*

*Example:*

```
POST /tv-shows/import?name=good girls
```

An example response [This is not what you store in DB, it is the call's response]

: 201 Created

```
{
    "id" : 123,
    "last-update": "2021-04-08-12:34:40",
    "tvmaze-id" : 23542,
    "_links": {
        "self": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows/123"
        }
    }
}
```

- You must import a TV show if only the given name matches a valid TV show (good girls, Good Girls, Good-Girls all match each other but they should not match a TV show like "Good Boys"). Be noted that the TVMaze API provides a fuzzy search and hence tolerates typos, capital/small, dashes...etc. and at the same time provides more results than the exact TV Show. You should only import the matching one (ignoring cases, and any character except English alphabet and numbers ). In case there are more than one TV show with the same one, it is up to you how to deal with it (e.g., importing both, importing the latest one, etc)
- What and how you store the data in DB is up to you, but take a look at the rest of questions to know what attributes you need to keep for each TV show
- Do not get confused with having two identifiers ("id", and "tvmaze-id"); "id" is a unique identifier in your service and all of your operations relay on this id to work; "tvmaze-id" is just a reference to the original data.
- CLARIFICATION: You should never change an ID of a resource (do not update ids)
- CLARIFICATION: You should replace [HOST_NAME]:[PORT] with correct values
- A brief description about _links here : https://developer.wordpress.org/rest-api/using-the-rest-api/linking-and-embedding (https://developer.wordpress.org/rest-api/using-the-rest-api/linking-and-embedding/) / <u></u>

---

## Question 2 - Retrieve a TV Show (2 marks)

This operation retrieves a collection by its ID (the ID that is generated by your application) . The response of this operation will show the details of TV show. Please see the provided response example below to see what attributes should be included in the response. "_links" gives the links for previous, next, and current resource if they exist . The next and previous resources are based on the sequential ID generated by your application.

The interface should look like as like below:

```
GET /tv-shows/{id}
```

Example Response returns: 200 OK

```
{
    "tvmaze-id" :23542,
    "id": 124,
    "last-update": "2021-04-08-12:34:40",
    "name": "Good Girls",
    "type": "Scripted",
    "language": "English",
    "genres": [
        "Drama",
        "Comedy",
        "Crime"
    ],
    "status": "Running",
    "runtime": 60,
    "premiered": "2018-02-26",
    "officialSite": "https://www.nbc.com/good-girls",
    "schedule": {
        "time": "22:00",
        "days": [
          "Sunday"
        ]
    },
    "rating": {
        "average": 7.4
    },
    "weight": 100,
    "network": {
        "id": 1,
        "name": "NBC",
        "country": {
          "name": "United States",
          "code": "US",
          "timezone": "America/New_York"
        }
    },
    "summary": "<p><b>Good Girls</b> follows three \"good girl\" suburban wives and m
    "_links": {
        "self": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows/124"
        },
        "previous": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows/123"
        },
        "next": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows/125"
        }
    }
}
```

## Question 3- Deleting a TV show (2 marks)

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

```
DELETE /tv-shows/{id}
```

Returns: 200 OK

```
{
    "message" :"The tv show with id 134 was removed from the database!",
    "id": 134
}
```

## Question 4 - Update a TV Show (2 marks)

This operation partially updates the details of a given TV Show.

The interface should look like the example below:

```
PATCH /tv-shows/{id}
{
    "name": "Good Girls",
    "language": "English",
    "genres": [
        "Drama",
        "Comedy",
        "Crime"
    ]
}
```

The above payload is just an example; it can contain any of the TV show attributes. Take a look at the example response in Question 2 to know the existing attributes.

Returns: 200 OK

```
{
    "id" : 123,
    "last-update": "2021-04-08-12:34:50",
    "_links": {
        "self": {
            "href": "http://[HOST_NAME]:[PORT]/tv-shows/123"
        }
    }
}
```

## Question 5 - Retrieve the list of available TV Shows 4 marks)

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

```
GET /tv-shows?order_by=<CSV-FORMATED-VALUE> & page=1 & page_size=100 & filter=<CSV-FORM
```

All four parameters are optional with default values being "order_by=+id", "page=1", and "page_size =100", filter="id,name". "page" and "page_size" are used for pagination; "page_size" shows the number of TV shows per page. "order_by" is a comma separated string value to sort the list based on the given criteria. The string consists of two parts: the first part is a special character '+' or '-' where '+' indicates ordering ascendingly, and '-' indicates ordering descendingly . The second part is an attribute name which is one of {id,name,runtime,premiered,rating-average}. Here are some sample values of "order_by" :

| | |
|---|---|
| +rating-average,+id | order by "rating-average ascending" and then "id ascending" <br> In this case sorti ng by "rating-average" has p riority over "id". This is similar to SQL order by clause : <br> " rating-average ASC, id ASC " |
| -premiered | order by "premiered descending" |

"filter" is also another comma separated values (only consider= tvmaze_id ,id ,last-update ,name ,type ,language ,genres ,status ,runtime ,premiered ,officialSite ,schedule ,rating ,weight ,network ,summary), and show what attribute should be shown for each TV show accordingly. Take a look at the following example to know how response should be like:

```
GET /tv-shows?order_by=+id&page=1&page_size=100&filter=id,name
```

All four parameters are optional with default values being "order_by=+id", "page=1", and "page_size=100", "filer=id,name"

Returns: 200 OK

```
{
    "page": 1,
    "page-size": 100,
    "tv-shows": [
          {
           "id" : 1,
           "name" : "Good Girls"
           },
           {
           "id" : 2,
           "name" : "Brilliant Girls"
           },
           ...
        ],
    "_links": {
        "self": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows?order_by=+id&page=1&page_size=100
        },
        "next": {
          "href": "http://[HOST_NAME]:[PORT]/tv-shows?order_by=+id&page=2&page_size=100
        }
      }
}
```

## Question 6 - get the statistics of the existing TV Show (3 marks)

This operations accepts a parameter called "format" which can be either "json" or "image". Depending on the format your operation should provide the following information: In case when the the format is image, your operation should return an image (can be in any image format) and the image illustrates the requested information in a visualization (apply all your knowledge when creating the visualization such as choosing appropriate visualization type and making sure that it is human readable, clear, and informative).

- TV shows break down by an attribute determined by the "by" parameter; this parameter can be any of the following TV show attributes: "language" (showing the percentage of TV shows per Language), "genres", "status", and "type". In case of "genres", a TV show can have multiple values; you should come up with a solution to visualise it. For instance you can think of h *ow to visualise what percentage of movies belong to both "Comedy" and "Crime" genres, etc.*
- Total Number of TV shows
- Total Number of TV shows updated in the last 24 hours

The interface should look like as like below when the format is JSON:

```
GET /tv-shows/statistics?format=json&by=language
```

Returns: 200 OK

```
{
    "total": 1241,
    "total-updated": 24,
    "values" : { "English": 60.7, "French": 19.2, ... }
}
```

# Notice:

- TVMAZE API is only used in Question 1 for importing a new TV show. The rest of operations relay on the existing TV Shows in your local database
- There is not template code for this assignment, you submission should stick to the assignment specifications.
- Your submission will be marked manually.
- You should adhere to the best design guidelines for REST API mentioned in the lecture (e.g., appropriate responses based on JSON format with proper status codes, full API documentation : the generated Swagger should be fully self-explanatory with operation summary, parameter descriptions, default values).
- You should consider cases such as invalid titles or any invalid attempts to use the endpoint ( e.g. If the input title doesn 't exist in the data source, return error 404)
- You should return appropriate responses (JSON) in case of already imported collections
- Your code must implemented in **flask-restx** 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')
- Some of the response of some operations contain "_links", depending on the response this should include links to "self", "next", and "previous" resources.
-

# Submission:

- The Deadline is **Saturday the 3rd of April 2021 17:59**
- 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 22 days ago (Wednesday 17 March 2021, 01:08:22 PM), last modified 7 days ago (Thursday 01 April 2021, 10:30:28 AM).

---

## Comments

🔖   🔍 (/COMP9321/21T1/forums/search?forum_choice=resource/59281)

💬 (/COMP9321/21T1/forums/resource/59281)

---

💬 Add a comment

---

Yixi Chen (/users/z5337291) a day ago (Tue Apr 06 2021 18:10:01 GMT+1000 (澳大利亚东部标准时间))

Hi professor, I made a submission accidentally after 18:00 but I've submitted it before. Just like to double confirm you will accept the last submission before the deadline, am I right?

Reply

> Mohammadali Yaghoubzadehfard (/users/z5138589) a day ago (Wed Apr 07 2021 06:54:42 GMT+1000 (澳大利亚东部标准时间))
>
> We will only consider the latest submission. However, if it is a bit late, we will mark without any impact on your mark,
>
> Reply

Jiayi Jiang (/users/z5277675) a day ago (Tue Apr 06 2021 17:08:32 GMT+1000 (澳大利亚东部标准时间)), last modified a day ago (Tue Apr 06 2021 17:09:27 GMT+1000 (澳大利亚东部标准时间))

Is this ok? I think it is in 3 days, but it metions '3 days and 1 hour'.

```
Submission datestamp:    Tue Apr  6 17:06:50 2021
Assignment deadline:     Sat Apr  3 17:59:00 2021
This submission is 3 days and 1 hour late
Your submission is ACCEPTED.
```

Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) a day ago (Wed Apr 07 2021 06:54:56 GMT+1000 (澳大利亚东部标准时间))

This is fine

Reply

Di Wu (/users/z5247036) a day ago (Tue Apr 06 2021 17:30:31 GMT+1000 (澳大利亚东部标准时间)), last modified a day ago (Tue Apr 06 2021 17:32:07 GMT+1000 (澳大利亚东部标准时间))

I also have this question, the latest version I submitted is Apr 6 17:24, which is before Apr 6 18:00 in Sydney time, I think the time difference between Beijing time (China time) and Sydney time is the reason why it showed it's 3 days +1 hour late. I think this won't affect your submission

```
Session: 21T1
Assignment: assign2  Submission ID: 5247036


Current day and time:   Tue Apr  6 17:31:18 2021
Assignment deadline:    Sat Apr  3 17:59:00 2021
A submission now would be 3 days and 1 hour late
Event           Day and time            Details
=================================================================
Submission      Mon Apr  5 23:46:56 2021  5247036 all assign2 [2 days +7 hours late]
Submission      Tue Apr  6 00:37:06 2021  5247036 all assign2 [2 days +8 hours late]
Submission      Tue Apr  6 16:03:21 2021  5247036 all assign2 [3 days late]
Submission      Tue Apr  6 16:39:02 2021  5247036 all assign2 [3 days late]
Submission      Tue Apr  6 17:03:05 2021  5247036 all assign2 [3 days +1 hour late]
Submission      Tue Apr  6 17:24:47 2021  5247036 all assign2 [3 days +1 hour late]

Most recent submission:
-rw------- z5247036/z5247036 19196 2021-04-06 17:24 z5247036.py
```

Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) a day ago (Wed Apr 07 2021 06:55:45 GMT+1000 (澳大利亚东部标准时间))

This is because the real deadline was on 3th not today; but we will not penalize for this

Reply

Di Wu (/users/z5247036) a day ago (Tue Apr 06 2021 17:36:39 GMT+1000 (澳大利亚东部标准时间))

am I right?

Reply

Devi Nair (/users/z5291618) a day ago (Tue Apr 06 2021 17:28:26 GMT+1000 (澳大利亚东部标准时间))

I got the same thing

Reply

Lin Chen (/users/z5286873) 2 days ago (Tue Apr 06 2021 16:07:45 GMT+1000 (澳大利亚东部标准时间))

Can I still submit it now?

Reply

**Mohammadali Yaghoubzadehfard (/users/z5138589)** 2 days ago (Tue Apr 06 2021 16:12:40 GMT+1000 (澳大利亚东部标准时间))

Yes, till 18:00 we are still accepting submissions

Reply

**Lin Chen (/users/z5286873)** 2 days ago (Tue Apr 06 2021 16:19:42 GMT+1000 (澳大利亚东部标准时间))

thanks

Reply

**Ahmad El Majzoub (/users/z5292964)** 2 days ago (Tue Apr 06 2021 13:59:20 GMT+1000 (澳大利亚东部标准时间))

We won't be penalized if we don't have enough comments in our python file right? The professor mentioned that we are not graded on the quality of the code so I am assuming code comments are part of that?

Thanks

Reply

**Mohammadali Yaghoubzadehfard (/users/z5138589)** 2 days ago (Tue Apr 06 2021 16:12:59 GMT+1000 (澳大利亚东部标准时间))

You will not be penalised

Reply

**Guanzhou Zhao (/users/z5242942)** 2 days ago (Tue Apr 06 2021 02:36:56 GMT+1000 (澳大利亚东部标准时间))

Hi,

For Q6, I was wondering how should we output the charts? Because there is an error python exit suddenly. How should we fix it?

Thank you!

Reply

**Mohammadali Yaghoubzadehfard (/users/z5138589)** 2 days ago (Tue Apr 06 2021 07:18:58 GMT+1000 (澳大利亚东部标准时间))

What is the error about?

Reply

**Guanzhou Zhao (/users/z5242942)** 2 days ago (Tue Apr 06 2021 09:16:15 GMT+1000 (澳大利亚东部标准时间))
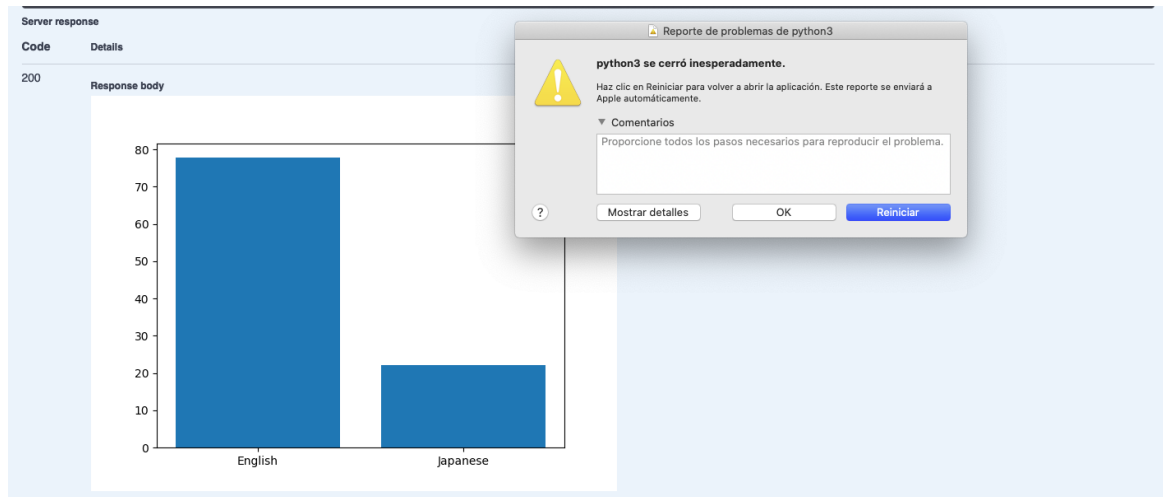
Python will exit unexpectedly, but the image will still be saved locally.

Reply

Matias Alfredo Morales Armijo (/users/z5216410) <u>2 days ago (Tue Apr 06 2021 01:32:58 GMT+1000 (澳大利亚东部标准时间))</u>

When I send the response of the image I got this error and the computer close the server. The image is saved and it is possible to see in the response. Do you know what could be happening? thanks



Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) <u>2 days ago (Tue Apr 06 2021 07:18:26 GMT+1000 (澳大利亚东部标准时间))</u>

Can you translate the message for me?

Reply

Matias Alfredo Morales Armijo (/users/z5216410) <u>2 days ago (Tue Apr 06 2021 09:47:00 GMT+1000 (澳大利亚东部标准时间))</u>

appears this message: Assertion failed: (NSViewIsCurrentlyBuildingLayerTreeForDisplay() != currentlyBuildingLayerTree), function NSViewSetCurrentlyBuildingLayerTreeForDisplay, file /AppleInternal/BuildRoot/Library/Caches/com.apple.xbs/Sources/AppKit/AppKit-1894.60.100/AppKit.subproj/NSView.m,

and the message says "python3 quit unexpectdely"

However I tested in the lab computers and it works. But I dont know why I have that problem in my computer.

Thanks,

Reply

Flynn Zhang (/users/z5256265) <u>2 days ago (Mon Apr 05 2021 23:21:41 GMT+1000 (澳大利亚东部标准时间))</u>, last modified <u>2 days ago (Tue Apr 06 2021 00:42:31 GMT+1000 (澳大利亚东部标准时间))</u>

For some reason on firefox in vlab, my requests are sent multiple times when i click 'execute' in swagger but not in chrome. Does this matter?

Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) 2 days ago (Tue Apr 06 2021 07:17:39 GMT+1000 (澳大利亚东部标准时间))

Do not worries about he client part, make sure your API does what it supposed to do

Reply

Siying Huang (/users/z5253264) 2 days ago (Mon Apr 05 2021 23:18:26 GMT+1000 (澳大利亚东部标准时间))

for Q5, should we consider order by +rating-sum? If we should consider sum part, how about the sum format, should like SUM,sum?

Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) 2 days ago (Tue Apr 06 2021 07:16:46 GMT+1000 (澳大利亚东部标准时间))

"rating-average" is a single attribute; there is no "sating-sum" in the question spec

Reply

Tai-Jung Yeh (/users/z5236552) 3 days ago (Mon Apr 05 2021 16:46:58 GMT+1000 (澳大利亚东部标准时间)), last modified 3 days ago (Mon Apr 05 2021 16:53:41 GMT+1000 (澳大利亚东部标准时间))

For q5, when I try to insert the parameters as a string, sqlite seems to treat it as a string literal.

For example:

_filter = "id,name"

c. execute("select ? from table", (_filter,) )

fetching the rows get me something like this:

( 'id,name', )

It seems to ignore my order by string as well. How should I rectify this?

Edit: Found the problem. I can't pass parameter requests for table names, column names etc, only values. Using string concat works fine for this.

Reply

Daniel Fan (/users/z5114117) 3 days ago (Mon Apr 05 2021 12:28:54 GMT+1000 (澳大利亚东部标准时间))

Hi,

Will tutors start marking after 3 days from the submission deadline, in case students resubmit since there is no penalty?

Reply

Mohammadali Yaghoubzadehfard (/users/z5138589) 3 days ago (Mon Apr 05 2021 22:01:48 GMT+1000 (澳大利亚东部标准时间))

Yes

Reply

Load More Comments