Generated at Sat Apr 27 02:41:09 2024.

Source at https://jupyterhub-dev.cheme.cmu.edu/user/kjamunap@andrew.cmu.edu/lab/tree/s24-06643/sse/assignments/project/project.ipynb.

Project

The project should provide some useful functionality in science or engineering. It could be a command line utility, or a package to use in notebooks. There should be some substance, but it does not need to be extensive. I don't expect it should take more than a few hours to write the code. It is not necessary to write very sophisticated code. Overall the project should demonstrate you have learned something in this class.

1. The project must be pip installable
2. Your project should utilize git, and there should be a version history.
3. Your project should have some tests.
4. Your project should use at least one code quality tool.
5. Your project should have a readme.md and LICENSE file.
6. The code should be well documented.
7. The code should be original work.
8. You should push it to a GitHUB repo.

At the end of the mini you will give a live demonstration of your project and what it does.

Before you turn this problem in, make sure everything runs as expected. First, **restart the kernel** (in the menubar, select Kernel→Restart) and then **run all cells** (in the menubar, select Cell→Run All).

Make sure you fill in any place that says `YOUR CODE HERE` or "YOUR ANSWER HERE", as well as your name and collaborators below:

In [3]:
```
NAME = "Kiran Prasad J P"
COLLABORATORS = ""
```

Title: Concept Trend Analyzer API

The project is to develop an OpenAlex API for Concept Search and Trend Analysis. This API will help researchers and users to search for concepts, retrieve related concepts, and analyze trends in scholarly works count and cited by count over ten years. The API will allow users to search for concepts using a query parameter, retrieve related concepts and concept ID, and analyze trends in works count and cited by count for the top related concepts. Leveraging the OpenAlex database, this project aims to facilitate concept exploration and trend analysis, enabling users to gain valuable insights into research trends and identify emerging topics of interest.

Key Features

- **Concept Search**: Users can search for concepts using a query parameter. The API will parse the search query and return relevant concept IDs and Wikipedia link.
- **Related Concepts Retrieval**: After receiving a concept ID, the API will fetch the top ten related concepts according to their score, along with their display names. This will be achieved by calling the `/concepts/{concept_id}?select=related_concepts`.
- **Trend Analysis**: Using the retrieved concept IDs and same API, user can give command for the API to retrieve the values of works count and cited by count for the last ten years, binned by year. This data will be visualized to show the trend over the years. This will be achieved by calling the `/concepts/{concept_id}?select=counts_by_year`.
- **Visualization**: The API will generate visualizations (e.g., line charts) to display the trend in works count and cited by count over the last ten years for each related concept.

Project demo

In these cells show how your project is installed, and how it is used. If it is a CLI use %%bash cells to illustrate it. Or import the library and show what it does.

Check your present working directory

In [4]:
```
pwd
```
Out[4]:
```
'/home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project'
```

Remove and uninstall any previously installed package folders

In [10]:
```
%%bash
rm -fr ConceptTrendAnalyzer
```

In [11]:
```
!pip uninstall CTA --yes
```
```
WARNING: Skipping CTA as it is not installed.
```

Clone your GIT Repository to the HUB

In [12]:
```
! git clone https://github.com/captainKHSH/ConceptTrendAnalyzer.git
```
```
Cloning into 'ConceptTrendAnalyzer'...
remote: Enumerating objects: 461, done.
remote: Counting objects: 100% (290/290), done.
remote: Compressing objects: 100% (190/190), done.
remote: Total 461 (delta 181), reused 172 (delta 100), pack-reused 171
Receiving objects: 100% (461/461), 138.12 KiB | 5.12 MiB/s, done.
Resolving deltas: 100% (258/258), done.
```

Show the outline of your project

Use the `tree` command to show the structure of your project here.

In [13]:
```
!tree ConceptTrendAnalyzer
```
```
ConceptTrendAnalyzer
├── analyze
│   ├── concept_search.py
│   ├── concept_trend_analyzer.py
│   ├── __init__.py
│   └── relatedconcepts.py
├── __init__.py
├── LICENSE
├── main.py
├── README.md
├── setup.py
├── test
│   └── test_concept_search.py
├── utils.py
└── visualization
    ├── generate_charts.py
    ├── __init__.py
    └── name.py

3 directories, 14 files
```

Install the Concept-Trend-Analyzer package

In [14]:
```
! cd ConceptTrendAnalyzer && pip install .
```
```
Defaulting to user installation because normal site-packages is not writeable
Processing /home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project/ConceptTrendAnalyzer/Concep
tTrendAnalyzer
  Preparing metadata (setup.py) ... done
Requirement already satisfied: requests in /opt/tljh/user/lib/python3.9/site-packages (from CTA==1.0.0) (2.31.
0)
Requirement already satisfied: matplotlib in /opt/tljh/user/lib/python3.9/site-packages (from CTA==1.0.0) (3.5.
3)
Requirement already satisfied: cycler>=0.10 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib->CTA
==1.0.0) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib
->CTA==1.0.0) (4.36.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib
->CTA==1.0.0) (1.4.4)
Requirement already satisfied: numpy>=1.17 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib->CTA=
=1.0.0) (1.23.2)
Requirement already satisfied: packaging>=20.0 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib->
CTA==1.0.0) (21.3)
Requirement already satisfied: pillow>=6.2.0 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib->CT
A==1.0.0) (8.4.0)
Requirement already satisfied: pyparsing>=2.2.1 in /opt/tljh/user/lib/python3.9/site-packages (from matplotlib-
>CTA==1.0.0) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in /opt/tljh/user/lib/python3.9/site-packages (from matplot
lib->CTA==1.0.0) (2.8.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /opt/tljh/user/lib/python3.9/site-packages (from req
uests->CTA==1.0.0) (2.0.0)
Requirement already satisfied: idna<4,>=2.5 in /opt/tljh/user/lib/python3.9/site-packages (from requests->CTA==
1.0.0) (3.1)
Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/tljh/user/lib/python3.9/site-packages (from requests-
>CTA==1.0.0) (1.26.7)
Requirement already satisfied: certifi>=2017.4.17 in /opt/tljh/user/lib/python3.9/site-packages (from requests-
>CTA==1.0.0) (2023.7.22)
Requirement already satisfied: six>=1.5 in /opt/tljh/user/lib/python3.9/site-packages (from python-dateutil>=2.
7->matplotlib->CTA==1.0.0) (1.16.0)
Building wheels for collected packages: CTA
  Building wheel for CTA (setup.py) ... done
  Created wheel for CTA: filename=CTA-1.0.0-py3-none-any.whl size=7320 sha256=18dda140a75f4b179b9bac664ecf79c00
e0411b8190ce9b81835dd80964af8cc
  Stored in directory: /tmp/pip-ephem-wheel-cache-_h2svbwf/wheels/21/35/54/eb223bdfc3caf55acad7ce313681fc436e8b
31d585078867bb
Successfully built CTA
WARNING: Error parsing requirements for jupyter-core: [Errno 2] No such file or directory: '/opt/tljh/user/lib/
python3.9/site-packages/jupyter_core-4.11.1.dist-info/METADATA'
DEPRECATION: nb-black 1.0.7 has a non-standard dependency specifier black>='19.3'; python_version >= "3.6". pip
 24.1 will enforce this behaviour change. A possible replacement is to upgrade to a newer version of nb-black o
r contact the author to suggest that they release a version with a conforming dependency specifiers. Discussion
 can be found at https://github.com/pypa/pip/issues/12063
DEPRECATION: plotlyhtmlexporter 0.0.2 has a non-standard dependency specifier nbformat>=4.2traitlets. pip 24.1
 will enforce this behaviour change. A possible replacement is to upgrade to a newer version of plotlyhtmlexpor
ter or contact the author to suggest that they release a version with a conforming dependency specifiers. Discu
ssion can be found at https://github.com/pypa/pip/issues/12063
Installing collected packages: CTA
Successfully installed CTA-1.0.0
```

In [15]:
```
cd ConceptTrendAnalyzer
```
```
/home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project/ConceptTrendAnalyzer/ConceptTrendAnaly
zer
```

Import main to get CODE usability

In [16]:
```
from main import *
```
```
# Usage code

To Search Concept
CS.search('CONCEPT')
To get related concepts
RC.related('concept_id')
To find trend over years
y,w,c = CAT.year('concept_id',Number of years)
CAT.ytable(y,w,c)
To Visualization
chart.chart('concept_id',Number of years)
```

Importing the utilities for the Concept Trend Analyzer package where we'll import all modules at once and get usable functions

In [17]:
```
from utils import *
```
```
Importing utilities for the Concept Trend Analyzer package...
```

Search query with complete concept

In [18]:
```
query = "Cars"
CS.search(query)
```

ID	Display Name	Wikipedia Link
C125557594	Refrigerator car	https://en.wikipedia.org/wiki/Refrigerator%20car
C2781360296	Car ownership	https://en.wikipedia.org/wiki/Car%20ownership
C64093975	Floating car data	https://en.wikipedia.org/wiki/Floating%20car%20data
C2983061281	Car sharing	https://en.wikipedia.org/wiki/Carsharing
C2987494969	Electric cars	https://en.wikipedia.org/wiki/Electric%20car

Out[18]:
```
True
```

Search query with auto-complete concept

In [19]:
```
query = "Comp"
CS.search(query)
```

ID	Display Name	Wikipedia Link
C41008148	Computer science	NA
C159985019	Composite material	NA
C38652104	Computer security	NA
C31258907	Computer network	NA
C98045186	Process (computing)	NA

Out[19]:
```
True
```

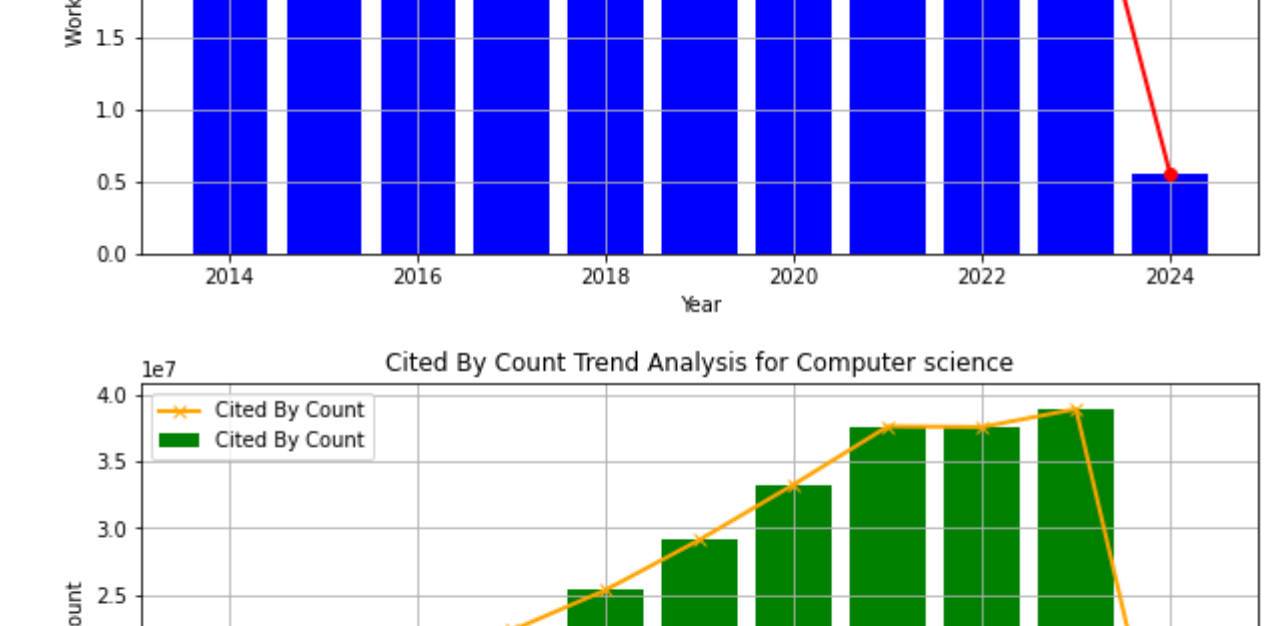
Search for related Concepts with ID obtained

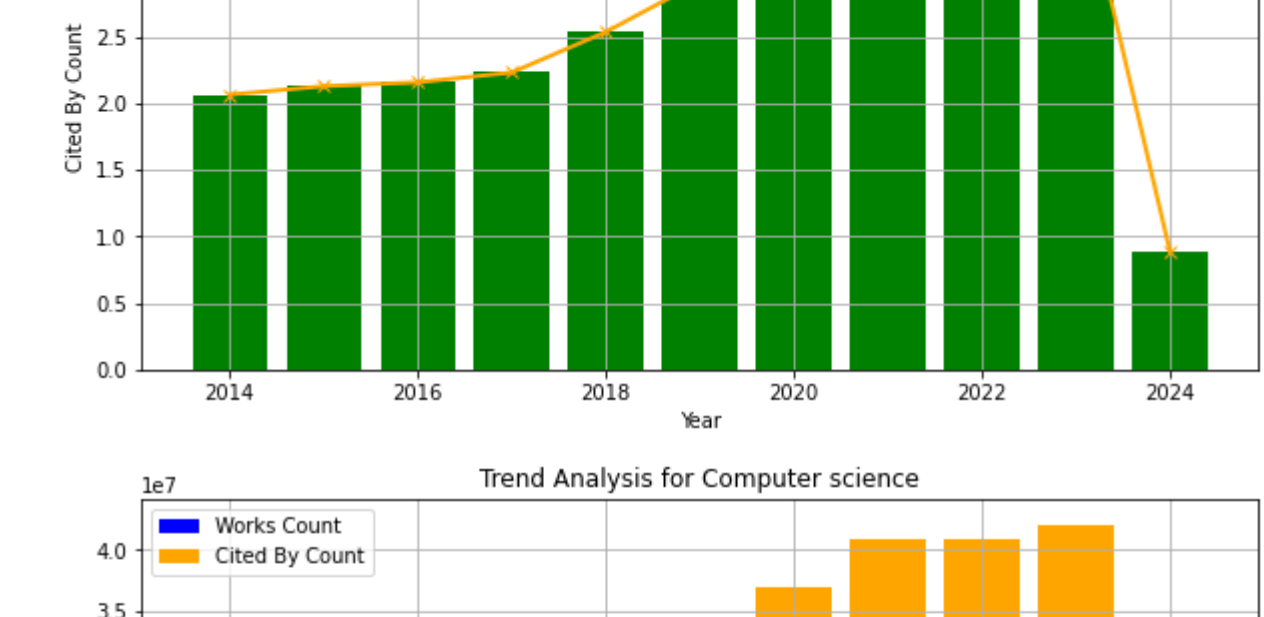
In [20]:
```
concept_id = "C41008148"
RC.related(concept_id)
```
```
ID              Related Concept Names
==============================================
C33923547       Mathematics
C119599485      Electrical engineering
C121332964      Physics
C127413603      Engineering
C78519656       Mechanical engineering
C154945302      Artificial intelligence
C11413529       Algorithm
C86803240       Biology
C147176958      Civil engineering
C185592680      Chemistry
```

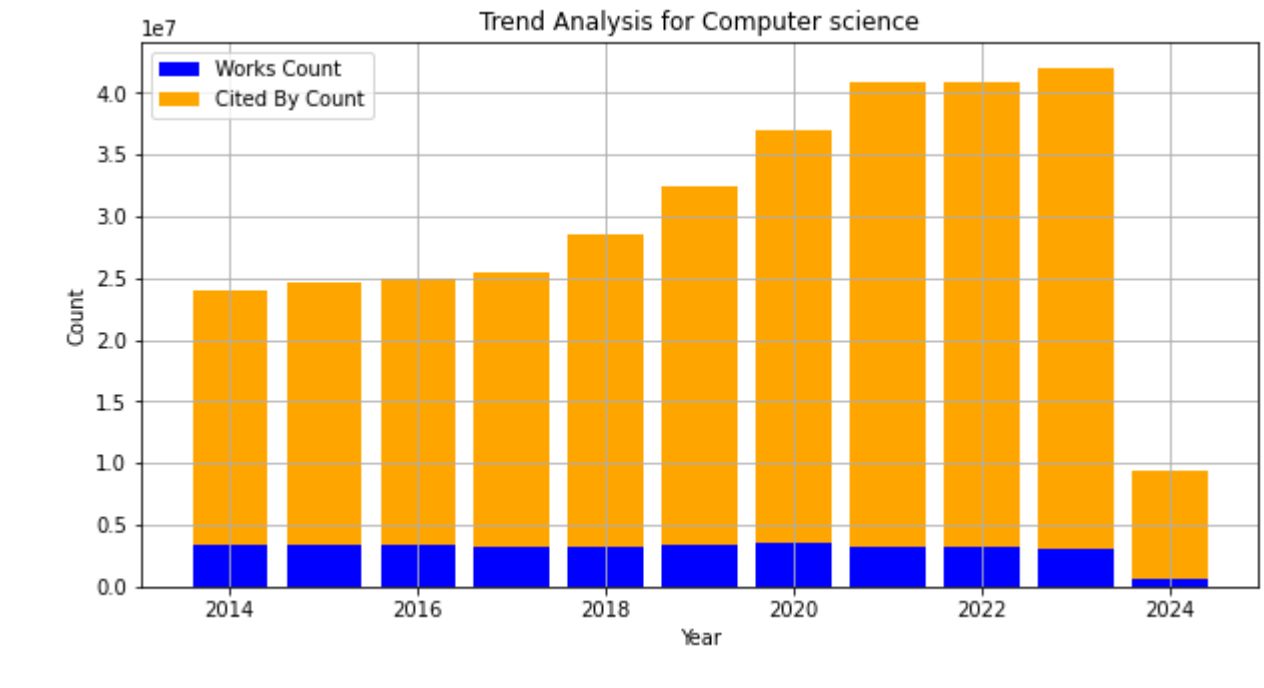
Search for Works and citations over years with ID obtained

In [23]:
```
Num = 7
y,w,c = CAT.year(concept_id,Num)

# Print results
CAT.ytable(y,w,c)
```
```
Year        Works Count      Cited Count
=============================================
2024        557611           8841987
2023        3068414          38944805
2022        3241673          37589263
2021        3195909          37647271
2020        3610755          33285545
2019        3320806          29136110
2018        3200363          25359538
```

In [25]:
```
chart.chart(concept_id,num_years = 11)
```
```
 CONCEPT NAME: Computer science

 DESCRIPTION: study of computation
```







GitHUB

Put a link to the GitHUB repo here.

https://github.com/captainKHSH/Concept-Trend-Analyzer.git

Version control

Describe your version control approach, and show that there is git history using the git log command.

In [26]:
```
!git log --graph --oneline
```
```
* 9fd8af7 (HEAD -> main, origin/main, origin/HEAD) Update concept_search.py
* 5397aa1 Update concept_search.py
* d128c28 Update test_concept_search.py
* 185d77a Update test_concept_search.py
* f93224d Update concept_search.py
* 9aeed98 Update test_concept_search.py
* 6bb6416 Update concept_search.py
* 7c5b73f Update concept_search.py
* 1f30503 Update test_concept_search.py
* 9fc8389 Update concept_search.py
* e73d16d Update concept_search.py
* 1ae654e Update concept_search.py
* c354f71 Update concept_search.py
* 3276be2 Update concept_search.py
* 3810f05 Update concept_search.py
* a5c3232 Update concept_search.py
* 1f42dae Update concept_search.py
* c835ddc Update concept_search.py
* d36d644 Update concept_search.py
* 37911d9 Update concept_search.py
* ff98301 Update concept_search.py
* d25f522 Update concept_search.py
* 9377f47 Update concept_search.py
* bb72b84 Update concept_search.py
* 0e75bcf Update concept_search.py
* 417a05a Update concept_search.py
* d1bf0eb Update concept_search.py
* 2ce545f Update concept_search.py
* 94bb0a6 Update concept_search.py
* e4ecf7c Update concept_search.py
* f7d1f6a Update concept_search.py
* aa57aaf Update concept_search.py
* 23156b3 Update test_concept_search.py
* 167a91c Update concept_search.py
* 933c08a Update concept_search.py
* eaa5446 Update concept_search.py
* df870d1 Update concept_search.py
* 617c776 Update concept_search.py
* 80c3a20 Update test_concept_search.py
* 778b8f3 Update test_concept_search.py
* 3d38bf3 Update test_concept_search.py
* be0e453 Update test_concept_search.py
* d2f9536 Update test_concept_search.py
* 9fc41e3 Update concept_search.py
* ec04c7d Update __init__.py
* bc5b6dc Update setup.py
* c79126e Update setup.py
* 92aa011 Update test_concept_search.py
* 0a4d0c7 Update test_concept_search.py
* 4303a77 Update generate_charts.py
* 780e0e6 Update utils.py
* b082630 Update utils.py
* 893fc86 Update utils.py
* f832ec9 Update __init__.py
* 1964a8c Update utils.py
* 5039109 Update __init__.py
* 6600363 Update __init__.py
* 18ac075 Update __init__.py
* bac5fff Update __init__.py
* 02991b4 Update __init__.py
* f0f7df9 Update __init__.py
* 763e877 Update utils.py
* 40d45d7 Update utils.py
* cb90114 Update utils.py
* 06490b2 Update __init__.py
* fc3d019 Update utils.py
* 16cefd6 Update generate_charts.py
* 2c98560 Update utils.py
* 8f47d7a Update generate_charts.py
* 036727c Update utils.py
* 5f1fb7f Update utils.py
* 10a8a6f Update generate_charts.py
* cc2c07e Update generate_charts.py
* 9cc30d3 Update utils.py
* 0c3864e Update __init__.py
* 13f4e09 Update __init__.py
* 962db0f Update __init__.py
* 755fac0 Update __init__.py
* 660778c Update __init__.py
* 01f3bdf Update generate_charts.py
* ba16231 Update test_concept_search.py
* 3f36902 Update __init__.py
* abb8fdf Update __init__.py
* 41c6a7e Update utils.py
* c0d792e Update __init__.py
* 71f98f5 Update generate_charts.py
* b54c51b Update utils.py
* 2fee8d5 Update test_concept_search.py
* c51fd14 Rename testy.py to test_concept_search.py
* a6e1e73 Update test_search.py
* da91802 Update .pre-commit-config.yaml
* c09451f Create testy.py
* bd42173 Update generate_charts.py
* 8f43855 Update generate_charts.py
* 8920a11 Update main.py
* 91db56c Update setup.py
* 975d49b Update and rename makefile.yaml to .pre-commit-config.yaml
* e09b3cb Create makefile.yaml
* ca3c5d5 Update README.md
* 9963f3d Update README.md
* 97d7eaa Update README.md
* f28a5fd Update README.md
* 6e971d7 Update README.md
* 23e007c Update generate_charts.py
* 359cfae Create name.py
* 3764e5c Update concept_trend_analyzer.py
* d1b21eb Update relatedconcepts.py
* 206c4fa Update concept_search.py
* 5bd25b6 Update __init__.py
* 53f66b2 Update __init__.py
* 9aaa824 Create main.py
* 12483f3 Update utils.py
* ef76121 Update __init__.py
* 97dfa37 Update setup.py
* 10f71b9 Update concept_search.py
* 37656c3 Update concept_search.py
* 7406d0e Update concept_search.py
* 18cd420 Update concept_search.py
* 78d70d2 Update concept_search.py
* e7d3cb4 Update generate_charts.py
* 1e8bb9c Create __init__.py
* 02e77b0 Create concept_trend_analyzer.py
* a802b7a Create relatedconcepts.py
* bf7af87 Update __init__.py
* ea3ca87 Update __init__.py
* 9124eed Create concept_search.py
* 2ea8eca Create __init__.py
* 907bd9e Create __init__.py
* 74043ef Create utils.py
* b02a3de Create setup.py
* d3533d6 Initial commit
```

Code quality tool

Describe the code quality tool you used in your project, and show evidence here of how it is implemented and how it is used.

In [27]:
```
!tree
```
```
.
├── analyze
│   ├── concept_search.py
│   ├── concept_trend_analyzer.py
│   ├── __init__.py
│   ├── __pycache__
│   │   ├── concept_search.cpython-39.pyc
│   │   ├── concept_trend_analyzer.cpython-39.pyc
│   │   ├── __init__.cpython-39.pyc
│   │   └── relatedconcepts.cpython-39.pyc
│   └── relatedconcepts.py
├── build
│   ├── bdist.linux-x86_64
│   └── lib
│       ├── analyze
│       │   ├── concept_search.py
│       │   ├── concept_trend_analyzer.py
│       │   ├── __init__.py
│       │   └── relatedconcepts.py
│       └── visualization
│           ├── generate_charts.py
│           ├── __init__.py
│           └── name.py
├── CTA.egg-info
│   ├── dependency_links.txt
│   ├── PKG-INFO
│   ├── requires.txt
│   ├── SOURCES.txt
│   └── top_level.txt
├── __init__.py
├── LICENSE
├── main.py
├── __pycache__
│   ├── main.cpython-39.pyc
│   └── utils.cpython-39.pyc
├── README.md
├── setup.py
├── test
│   └── test_concept_search.py
├── utils.py
└── visualization
    ├── generate_charts.py
    ├── __init__.py
    ├── name.py
    └── __pycache__
        ├── generate_charts.cpython-39.pyc
        ├── __init__.cpython-39.pyc
        └── name.cpython-39.pyc

12 directories, 35 files
```

In [28]:
```
cd ..
```
```
/home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project/ConceptTrendAnalyzer
```

In [29]:
```
! black ConceptTrendAnalyzer --diff --color
```
```
All done! ✨ 🍰 ✨
12 files would be left unchanged.
```

In [30]:
```
! flake8 --extend-ignore F401 --exclude CTA/build,CTA/.ipynb_checkpoints --docstring-convention numpy ConceptTr
```

Black and Flake8 doesnot show any errors

Tests

Describe the tests you built into your project and how they help ensure the project works, and that changes don't break functionality. In the cells below show how you run the tests, and that they work.

Two tests are conducted to check weather the Concept_Search is fetching results for completed query and auto-complete query

In [31]:
```
!pytest ConceptTrendAnalyzer
```
```
============================= test session starts ==============================
platform linux -- Python 3.9.7, pytest-7.2.2, pluggy-1.3.0
rootdir: /home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project/ConceptTrendAnalyzer/ConceptT
rendAnalyzer
plugins: typeguard-2.13.3, anyio-3.6.1
collected 2 items

ConceptTrendAnalyzer/test/test_concept_search.py ..                      [100%]

============================== 2 passed in 0.59s ===============================
```

In [34]:
```
cd ..
```
```
/home/jupyter-kjamunap@andrew.cm-10011/s24-06643/sse/assignments/project
```

When you are done, download a PDF and turn it in on Canvas. Make sure to save your notebook, then run this cell and click on the download link.

In [35]:
```
%run ~/s24-06643/s24.py
%pdf
```

Open project.pdf

download project.pdf