This scraper is implemented in Python 3.7.3. Please also have the BeautifulSoup (>= 4.9.0) package installed prior to running the scraper. For the extra credit parts, please have Matplotlib and NetworkX installed as well.

```
Mirandas-MBP:fa20-cs242-assignment2 mirandaliu$ python3 main.py -h
usage: main.py [-h] --url  [--nbook] [--nauthor]

CS242 Assignment 2 User Inputs.

required arguments:
  --url       URL to begin.

optional arguments:
  --nbook     Number of books to scrape.
  --nauthor   Number of authors to scrape.
Mirandas-MBP:fa20-cs242-assignment2 mirandaliu$
```

Direct to the working folder, and in your terminal, please run **python3 main.py -h** to look at all options for running the scraper. You should be asked to give an url of a book to begin with. The parameters nbook and nauthor to scrape are optional.

An example to run the scraper with *Clean Code: A Handbook of Agile Software Craftsmanship* would be:

python3 main.py --url https://www.goodreads.com/book/show/3735293-clean-code

After the scraper has finished running, there should be two files generated: book.json & author.json:

{} book.json > ...


```
[{"book_url": "https://www.goodreads.com/book/show/3735293-clean-code", "title": "Clean Code: A Handbook of Agile Software Craftsmanship",
"book_id": "3735293", "isbn": "9780132350884", "author_url": "https://www.goodreads.com/author/show/45372.Robert_C_Martin", "author":
"Robert C. Martin", "rating": 4.4, "rating_count": 14891, "review_count": 876, "image_url": "https://i.gr-assets.com/images/S/compressed.
photo.goodreads.com/books/1436202607l/3735293._SX318_.jpg", "similar_books": ["4099.The_Pragmatic_Programmer", "85009.Design_Patterns",
"44936.Refactoring", "58128.Head_First_Design_Patterns", "4845.Code_Complete", "34927404-effective-java", "44919.
Working_Effectively_with_Legacy_Code", "179133.Domain_Driven_Design", "387190.Test_Driven_Development", "23215733-the-software-craftsman",
"2998152-javascript", "85012.Enterprise_Integration_Patterns", "70156.Patterns_of_Enterprise_Application_Architecture", "13629.
The_Mythical_Man_Month", "22512931-building-microservices", "127932.Java_Concurrency_in_Practice", "67833.Extreme_Programming_Explained",
"6487349-the-art-of-unit-testing"]}, {"book_url": "https://www.goodreads.com/book/show/4099.The_Pragmatic_Programmer", "title": "The
Pragmatic Programmer: From Journeyman to Master", "book_id": "4099", "isbn": "9780201616224", "author_url": "https://www.goodreads.com/
author/show/2815.Andy_Hunt", "author": "Andy Hunt", "rating": 4.32, "rating_count": 15716, "review_count": 907, "image_url": "https://i.
gr-assets.com/images/S/compressed.photo.goodreads.com/books/1401432508l/4099.jpg", "similar_books": ["3735293-clean-code", "85009.
Design_Patterns", "4845.Code_Complete", "44936.Refactoring", "10284614-the-clean-coder", "58128.Head_First_Design_Patterns", "13629.
The_Mythical_Man_Month", "44919.Working_Effectively_with_Legacy_Code", "67833.Extreme_Programming_Explained", "18043011-clean-architecture",
"2998152-javascript", "34927404-effective-java", "387190.Test_Driven_Development", "179133.Domain_Driven_Design",
"6399113-the-passionate-programmer", "43713.Structure_and_Interpretation_of_Computer_Programs", "23232941-soft-skills", "515601.
The_C_Programming_Language"]}, {"book_url": "https://www.goodreads.com/book/show/85009.Design_Patterns", "title": "Design Patterns: Elements
of Reusable Object-Oriented Software", "book_id": "85009", "isbn": "9780201633610", "author_url": "https://www.goodreads.com/author/show/
48622.Erich_Gamma", "author": "Erich Gamma", "rating": 4.18, "rating_count": 9845, "review_count": 351, "image_url": "https://i.gr-assets.
com/images/S/compressed.photo.goodreads.com/books/1348027904l/85009.jpg", "similar_books": ["3735293-clean-code", "44936.Refactoring", "4099.
The_Pragmatic_Programmer", "58128.Head_First_Design_Patterns", "4845.Code_Complete", "387190.Test_Driven_Development", "515601.
The_C_Programming_Language", "70156.Patterns_of_Enterprise_Application_Architecture", "105125.Effective_C_", "108986.
Introduction_to_Algorithms", "10284614-the-clean-coder", "179133.Domain_Driven_Design", "34927404-effective-java", "44919.
Working_Effectively_with_Legacy_Code", "127932.Java_Concurrency_in_Practice", "18043011-clean-architecture", "112251.
The_C_Programming_Language", "2998152-javascript"]}, {"book_url": "https://www.goodreads.com/book/show/44936.Refactoring", "title":
"Refactoring: Improving the Design of Existing Code", "book_id": "44936", "isbn": "9780201485677", "author_url": "https://www.goodreads.com/
author/show/25215.Martin_Fowler", "author": "Martin Fowler", "rating": 4.25, "rating_count": 6912, "review_count": 276, "image_url": "https:/
/i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1386925632l/44936.jpg", "similar_books": ["44919.
Working_Effectively_with_Legacy_Code", "85009.Design_Patterns", "3735293-clean-code", "4099.The_Pragmatic_Programmer", "387190.
Test_Driven_Development", "4845.Code_Complete", "179133.Domain_Driven_Design", "58128.Head_First_Design_Patterns", "67833.
Extreme_Programming_Explained", "18043011-clean-architecture", "127932.Java_Concurrency_in_Practice", "10284614-the-clean-coder",
"34927404-effective-java", "85041.Refactoring_to_Patterns", "13629.The_Mythical_Man_Month",
"4268826-growing-object-oriented-software-guided-by-tests", "1069827.Release_It_", "45280021-clean-agile", "582858.
Managing_Software_Requirements", "968277.ATL_Internals", "20453509-adopting-the-rational-unified-process", "198009.
Agility_and_Discipline_Made_Easy", "1007058.Applying_Use_Case_Driven_Object_Modeling_with_UML", "1617464.The_Art_of_Clearcase_Deployment",
"380743.The_Art_of_Objects", "462050.Aspect_Oriented_Software_Development_with_Use_Cases", "1657073.
Building_J2ee_Applications_with_the_Rational_Unified_Process", "1920160.Building_Web_Applications_with_UML"]}, {"book_url": "https://www.
goodreads.com/book/show/58128.Head_First_Design_Patterns", "title": "Head First Design Patterns", "book_id": "58128", "isbn":
"9780596007126", "author_url": "https://www.goodreads.com/author/show/32731.Eric_Freeman", "author": "Eric Freeman", "rating": 4.28,
```


{} author.json > {} 0 > [ ] author_books > 123 3


```
[{"name": "Robert C. Martin", "author_url": "https://www.goodreads.com/author/show/45372.Robert_C_Martin", "author_id": "45372", "rating": 4
34, "rating_count": 26988, "review_count": 1763, "image_url": "https://images.gr-assets.com/authors/1490470967p5/45372.jpg",
"related_authors": ["2815.Andy_Hunt", "3307.Steve_McConnell", "25201.Michael_C_Feathers", "25211.Kent_Beck", "25215.Martin_Fowler", "32731.
Eric_Freeman", "48622.Erich_Gamma", "48660.Joshua_Kerievsky", "104368.Eric_Evans", "7127583.Sandro_Mancuso"], "author_books": ["Clean Code:
A Handbook of Agile Software Craftsmanship", "The Clean Coder: A Code of Conduct for Professional Programmers", "Clean Architecture", "Agile
Software Development, Principles, Patterns, and Practices", "Agile Principles, Patterns, and Practices in C#", "Clean Agile: Back to
Basics", "The Robert C. Martin Clean Code Collection", "UML for Java Programmers", "Pattern Languages of Program Design 3", "Design
Principles and  Design Patterns"]}, {"name": "Andy Hunt", "author_url": "https://www.goodreads.com/author/show/2815.Andy_Hunt", "author_id":
"2815", "rating": 4.22, "rating_count": 20515, "review_count": 1169, "image_url": "https://images.gr-assets.com/authors/1215370372p5/2815.
jpg", "related_authors": ["302.Chad_Fowler", "3307.Steve_McConnell", "16360.Ethan_Marcotte", "25201.Michael_C_Feathers", "25211.Kent_Beck",
"25215.Martin_Fowler", "27837.Greg_Mortenson", "32731.Eric_Freeman", "38238.Tom_DeMarco", "42537.37_Signals", "45372.Robert_C_Martin",
"48622.Erich_Gamma", "55215.Junot_D_az", "101967.Scott_Carrier", "104368.Eric_Evans", "160418.Judith_S_Beck", "278960.Eric_Ries", "377625.
Julia_Tavalaro", "2938301.Roy_Osherove", "3174788.Frederick_P_Brooks_Jr_", "5037537.Michael_Scotto", "13168387.Megan_McLachlan", "14408819.
Ijeoma_Oluo"], "author_books": ["The Pragmatic Programmer: From Journeyman to Master", "Pragmatic Thinking and Learning: Refactor Your
Wetware", "Pragmatic Unit Testing in C# with NUnit", "Conglommora", "Learn to Program with Minecraft Plugins: Create Flying Creepers and
Flaming Cows in Java", "Conglommora Found", "Learn to Program with Minecraft Plugins: Create Flaming Cows in Java Using CanaryMod",
"\u7a0b\u5e8f\u5458\u7684\u601d\u7ef4\u4fee\u70bc", "Managing Your Project: Achieving Success with Minimal Stress", "Programming Ruby: The
Pragmatic Programmers' Guide"]}, {"name": "Steve McConnell", "author_url": "https://www.goodreads.com/author/show/3307.Steve_McConnell",
"author_id": "3307", "rating": 4.21, "rating_count": 11356, "review_count": 519, "image_url": "https://s.gr-assets.com/assets/nophoto/user/
m_200x266-d279b33f8eec0f27b7272477f09806be.png", "related_authors": ["2815.Andy_Hunt", "16770.Alan_Cooper", "25201.Michael_C_Feathers",
"25211.Kent_Beck", "25215.Martin_Fowler", "29393.Jon_L_Bentley", "32731.Eric_Freeman", "38238.Tom_DeMarco", "45372.Robert_C_Martin", "48622.
Erich_Gamma", "60805.Joshua_Bloch", "60832.Scott_Meyers", "60841.Thomas_H_Cormen", "104368.Eric_Evans", "153350.Brian_W_Kernighan", "229855.
Andy_Oram", "312165.Peter_Seibel", "1290380.Douglas_Crockford", "3174788.Frederick_P_Brooks_Jr_", "4149510.Jez_Humble", "5409448.
Harold_Abelson"], "author_books": ["Code Complete", "Rapid Development: Taming Wild Software Schedules", "Software Estimation: Demystifying
the Black Art", "Software Project Survival Guide", "Professional Software Development: Shorter Schedules, Higher Quality Products, More
Successful Projects, Enhanced Careers", "After the Gold Rush: Creating a True Profession of Software Engineering", "More Effective Agile",
"K\u014ddo Konpur\u012bto:  003.Kanzenna Puroguramingu O Mezashite Maikurosofuto K\u014dshiki", "Sofut\u014dea Mitsumori: Ningetsu No
Anmokuchi O Tokiakasu", "K\u014ddo Konpur\u012bto:  001.Kanzenna Puroguramingu O Mezashite Maikurosofuto K\u014dshiki"]}, {"name": "Michael
C. Feathers", "author_url": "https://www.goodreads.com/author/show/25201.Michael_C_Feathers", "author_id": "25201", "rating": 4.03,
"rating_count": 5484, "review_count": 302, "image_url": "https://images.gr-assets.com/authors/1442450392p5/25201.jpg", "related_authors":
["302.Chad_Fowler", "2815.Andy_Hunt", "3307.Steve_McConnell", "6089.Michael_T_Nygard", "25211.Kent_Beck", "25215.Martin_Fowler", "32731.
Eric_Freeman", "38238.Tom_DeMarco", "45372.Robert_C_Martin", "48622.Erich_Gamma", "48627.Gregor_Hohpe", "48660.Joshua_Kerievsky", "60805.
Joshua_Bloch", "73409.Brian_Goetz", "104368.Eric_Evans", "312165.Peter_Seibel", "2938301.Roy_Osherove", "3174788.Frederick_P_Brooks_Jr_",
"3362774.Sam_Newman", "4149510.Jez_Humble", "7127583.Sandro_Mancuso", "7969625.Martin_Kleppmann", "9824504.Steve_Freeman", "17037914.
Nicole_Forsgren"], "author_books": ["Working Effectively with Legacy Code", "Brutal Refactoring: More Working Effectively with Legacy Code",
"Beautiful Code: Leading Programmers Explain How They Think", "Working Effectively with Unit Tests", "Coder Proprement", "The Art of Unit
Testing: Deutsche Ausgabe"]}, {"name": "Kent Beck", "author_url": "https://www.goodreads.com/author/show/25211.Kent_Beck", "author_id":
"25211", "rating": 4.1, "rating_count": 17192, "review_count": 784, "image_url": "https://images.gr-assets.com/authors/1345546877p5/25211.
jpg", "related_authors": ["2620.Mike_Cohn", "2815.Andy_Hunt", "3307.Steve_McConnell", "6089.Michael_T_Nygard", "18257.Scott_Rosenberg",
```

To run the extra credit part, use the following command after you have done the above scraping:

python3 extra.py

This should generate a figure like the one below: