

05-szymanski-similar_books_analysis

January 21, 2019

1 Important

make `data/processed/similar_books.csv` has to be run before any cell in this notebook

2 Imports

```
In [1]: import pandas as pd
```

3 Helper functions

```
In [2]: def retrieve_similar_books(book_id, similar_books_data, book_data):
        similar_books = similar_books_data.groupby('book_id').get_group(
            book_id)['similar_book_id']
        selected_book_title = book_data.loc[book_id, 'title']
        similar_books_titles = book_data.loc[similar_books, 'title'].values

        print(f'Selected Book: {selected_book_title}:\n')
        print(f'\tSimilar Books:\n')
        for similar_book_title in similar_books_titles:
            print(f'\t\t{similar_book_title}')
```

```
In [3]: def get_series_similar_books(series_ids, similar_books_data, book_data):
        for book_id in series_ids:
            retrieve_similar_books(book_id, similar_books_data, book_data)
        print(90*'- ' + '\n')
```

4 Analysis

```
In [4]: similar_books_data = pd.read_csv('../data/processed/similar_books.csv',
        index_col='book_id')
        book_data = pd.read_csv('../data/processed/book.csv', index_col='book_id')
```

```
In [5]: similar_books_data.head()
```

```
Out[5]:
```

	similar_book_id
book_id	
1237	3136
6050	3136
5573	3136
1237	3003
8090	3003

5 Popular book series

In order to check if the data from the `similar books` tag is appropriate a manual analysis for popular book series of different genres will be made. If the proposed similar books will belong to the same series or be similar in terms of settings/plots/genre then it will be considered as a valid position in the data.

5.1 Harry Potter

```
In [6]: harry_potter_book_ids = [2, 18, 21, 23, 24, 25, 27]
```

```
In [7]: get_series_similar_books(
        harry_potter_book_ids, similar_books_data, book_data)
```

Selected Book: Harry Potter and the Sorcerer's Stone (Harry Potter, #1):

Similar Books:

```

    Calling on Dragons (Enchanted Forest Chronicles, #3)
    Rise of the Evening Star (Fablehaven, #2)
    The Legend of Luke (Redwall, #12)
    The Mysterious Benedict Society and the Prisoner's Dilemma (The
Mysterious Benedict Society, #3)
    The Field Guide (The Spiderwick Chronicles, #1)
    The Lost Colony (Artemis Fowl, #5)
    Molly Moon's Incredible Book of Hypnotism (Molly Moon, #1)
    The Woman Who Rides Like a Man (Song of the Lioness, #3)
    The White Dragon (Pern, #3)
    Krew elfów (Saga o Wiedminie, #3)
    The Mark of Athena (The Heroes of Olympus, #3)
    Rebel Angels (Gemma Doyle, #2)
    Dragons of Autumn Twilight (Dragonlance: Chronicles, #1)
```

Selected Book: Harry Potter and the Prisoner of Azkaban (Harry Potter, #3):

Similar Books:

```

    The King of Attolia (The Queen's Thief, #3)
    Abarat (Abarat, #1)
    Sign of the Unicorn (The Chronicles of Amber #3)
    The Riddle (The Books of Pellinor, #2)
    Son of the Shadows (Sevenwaters, #2)
    Lady Knight (Protector of the Small, #4)
    The Grim Grotto (A Series of Unfortunate Events, #11)
    Gregor and the Code of Claw (Underland Chronicles, #5)
```

Selected Book: Harry Potter and the Order of the Phoenix (Harry Potter, #5):

Similar Books:

```

    Child of the Prophecy (Sevenwaters, #3)
    Gregor and the Marks of Secret (Underland Chronicles, #4)
    Emperor Mage (Immortals, #3)
    The Queen of Attolia (The Queen's Thief, #2)
```

A Ring of Endless Light (Austin Family, #4)
The Indigo Spell (Bloodlines, #3)
The Atlantis Complex (Artemis Fowl, #7)
Syren (Septimus Heap, #5)
Haunted (The Mediator, #5)
The Subtle Knife (His Dark Materials, #2)

Selected Book: Harry Potter and the Chamber of Secrets (Harry Potter, #2):

Similar Books:

Behemoth (Leviathan, #2)
Talking to Dragons (Enchanted Forest Chronicles, #4)
Gregor and the Prophecy of Bane (Underland Chronicles, #2)
Queste (Septimus Heap, #4)
Grip of the Shadow Plague (Fablehaven, #3)
Tales From the Hood (The Sisters Grimm, #6)
The Riddle (The Books of Pellinor, #2)
Twilight (The Mediator, #6)
Mortal Coil (Skulduggery Pleasant, #5)
Daja's Book (Circle of Magic, #3)
The Penultimate Peril (A Series of Unfortunate Events, #12)

Selected Book: Harry Potter and the Goblet of Fire (Harry Potter, #4):

Similar Books:

In the Hand of the Goddess (Song of the Lioness, #2)
The Sweet Far Thing (Gemma Doyle, #3)
Darkest Hour (The Mediator, #4)
The Last Battle (Chronicles of Narnia, #7)
Pippi in the South Seas
Physik (Septimus Heap, #3)
Jacob Have I Loved

Selected Book: Harry Potter and the Deathly Hallows (Harry Potter, #7):

Similar Books:

Emperor Mage (Immortals, #3)
Keys to the Demon Prison (Fablehaven, #5)
Inkspell (Inkworld, #2)
Ptolemy's Gate (Bartimaeus, #3)
The Last Olympian (Percy Jackson and the Olympians, #5)
The High Lord (Black Magician Trilogy, #3)
The Authoritative Calvin and Hobbes: A Calvin and Hobbes Treasury
Queen of Sorcery (The Belgariad, #2)

Selected Book: Harry Potter and the Half-Blood Prince (Harry Potter, #6):

Similar Books:

Squire (Protector of the Small, #3)
Mossflower (Redwall, #2)
The Eternity Code (Artemis Fowl, #3)
The Horse and His Boy (Chronicles of Narnia, #5)
City of Masks (Stravaganza, #1)
Rilla of Ingleside (Anne of Green Gables, #8)
The Amber Spyglass (His Dark Materials, #3)
The Carnivorous Carnival (A Series of Unfortunate Events, #9)
The Lives of Christopher Chant (Chrestomanci, #2)
Lord of the Shadows (Cirque Du Freak, #11)

5.2 Twilight

```
In [8]: twilight_book_ids = [3, 49, 52]
```

```
In [9]: get_series_similar_books(  
        twilight_book_ids, similar_books_data, book_data)
```

Selected Book: Twilight (Twilight, #1):

Similar Books:

Twilight Director's Notebook : The Story of How We Made the Movie
Based on the Novel by Stephenie Meyer
The Twilight Saga Breaking Dawn Part 1: The Official Illustrated Movie
Companion (The Twilight Saga: The Official Illustrated Movie Companion, #4)
Dance with a Vampire (Vampire Kisses, #4)
Hourglass (Evernight, #3)
Lover Eternal (Black Dagger Brotherhood, #2)
Shadow Kiss (Vampire Academy, #3)
Midnight Awakening (Midnight Breed, #3)
Bite Club (The Morganville Vampires, #10)
The Killing Dance (Anita Blake, Vampire Hunter, #6)
Dead to the World (Sookie Stackhouse, #4)
Chosen (House of Night, #3)
Wisdom (My Blood Approves, #4)
The Fury / Dark Reunion (The Vampire Diaries, #3-4)

Selected Book: New Moon (Twilight, #2):

Similar Books:

Stargazer (Evernight, #2)
Destined for an Early Grave (Night Huntress, #4)
The Awakening / The Struggle (The Vampire Diaries, #1-2)
Jessica's Guide to Dating on the Dark Side (Jessica, #1)
The Dead Girls' Dance (The Morganville Vampires, #2)
Tempted (House of Night, #6)
All Together Dead (Sookie Stackhouse, #7)
The Golden Lily (Bloodlines, #2)
Fate (My Blood Approves, #2)
Misguided Angel (Blue Bloods, #5)

Blood Noir (Anita Blake, Vampire Hunter #16)
Lover Eternal (Black Dagger Brotherhood, #2)
Kiss of Midnight (Midnight Breed, #1)

Selected Book: Eclipse (Twilight, #3):

Similar Books:

Awakened (House of Night, #8)
Kiss of Crimson (Midnight Breed, #2)
The Coffin Club (Vampire Kisses, #5)
The Awakening / The Struggle (The Vampire Diaries, #1-2)
The Dead Girls' Dance (The Morganville Vampires, #2)
Frostbite (Vampire Academy, #2)
Dead to the World (Sookie Stackhouse, #4)
Lover Revealed (Black Dagger Brotherhood, #4)
Flutter (My Blood Approves, #3)
City of Glass (The Mortal Instruments, #3)

6 Rober Langdon series - Dan Brown

```
In [10]: robert_langdon_book_ids = [9, 26, 201, 240]
```

```
In [11]: get_series_similar_books(  
         robert_langdon_book_ids, similar_books_data, book_data)
```

Selected Book: Angels & Demons (Robert Langdon, #1):

Similar Books:

Patriot Games (Jack Ryan Universe, #2)
The Venetian Betrayal (Cotton Malone, #3)
Scarecrow (Shane Schofield, #3)
Extreme Measures (Mitch Rapp, #11)
The Bourne Supremacy (Jason Bourne, #2)
Map of Bones (Sigma Force, #2)
Gone Tomorrow (Jack Reacher, #13)
Up Country
Moscow Rules (Gabriel Allon, #8)
Jackdaws

Selected Book: The Da Vinci Code (Robert Langdon, #2):

Similar Books:

The Man From St. Petersburg
The Judas Strain (Sigma Force, #4)
Seven Deadly Wonders (Jack West Jr, #1)
The Venetian Betrayal (Cotton Malone, #3)
Visions in Death (In Death, #19)

The Big Bad Wolf (Alex Cross, #9)
The Lion's Game (John Corey, #2)
Valhalla Rising (Dirk Pitt, #16)
If Tomorrow Comes (Tracy Whitney Series, #1)
The Cabinet of Curiosities (Pendergast, #3)
False Impression
The Chamber

Selected Book: The Lost Symbol (Robert Langdon, #3):

Similar Books:

The Closers (Harry Bosch, #11; Harry Bosch Universe, #13)
Amazonia
Red Rabbit (Jack Ryan Universe, #3)
The Charlemagne Pursuit (Cotton Malone, #4)
Still Life With Crows (Pendergast, #4)
The 5 Greatest Warriors (Jack West Jr, #3)
London Bridges (Alex Cross, #10)

Selected Book: Inferno (Robert Langdon, #4):

Similar Books:

Bloodline (Sigma Force, #8)
Threat Vector (Jack Ryan Universe, #15)
The Hit (Will Robie, #2)
The English Girl (Gabriel Allon, #13)

7 Winnie the Pooh

```
In [12]: winnie_the_pooh_book_ids = [444, 1545, 2648, 7481, 7696]
```

```
In [13]: get_series_similar_books(  
         winnie_the_pooh_book_ids, similar_books_data, book_data)
```

Selected Book: Winnie-the-Pooh (Winnie-the-Pooh, #1):

Similar Books:

The Wolves of Willoughby Chase (The Wolves Chronicles, #1)
A Bear Called Paddington (Paddington, #1)
The Story of Ferdinand
The Faraway Tree Stories (The Faraway Tree #1-3)
Five Children and It (Five Children, #1)

Selected Book: The House at Pooh Corner (Winnie-the-Pooh, #2):

Similar Books:

The Tale of Mrs. Tiggy-Winkle
A Bear Called Paddington (Paddington, #1)
A Child's Garden of Verses
Just So Stories
The Complete Adventures of Curious George
The Story of Babar

Selected Book: The World of Winnie-the-Pooh (Winnie-the-Pooh, #1-2):

Similar Books:

A Child's Garden of Verses
Horton Hatches the Egg

Selected Book: When We Were Very Young (Winnie-the-Pooh, #3):

Similar Books:

The Tale of Mrs. Tiggy-Winkle
A Child's Garden of Verses
Just So Stories

Selected Book: Now We Are Six (Winnie-the-Pooh, #4):

Similar Books:

A Child's Garden of Verses
A Kiss for Little Bear (An I Can Read Book) by Minarik, Else Holmelund
[1984]
The Snowman

8 Summary

The similar books usually does not contain books from the same series. However, the remaining books do seem to be similar in terms of setting or plot e.g. books about vampires in case of similar books to Twilight or children books in case of Winnie the Pooh.