

```
#Import needed libraries
```

```
from bs4 import BeautifulSoup
import requests as requests
import pygal
import time
import pygal
from IPython.display import
```

```
base_html = """
<!DOCTYPE html>
<html>
    <head>
        <script type="text/javascript">
        <script type="text/javascript">
        </head>
        <body>
            <figure>
                {rendered_chart}
            </figure>
        </body>
    </html>
"""
```

```
#define functions for finding book details
def find_book_name(table):
    if table.find('h1') != None:
        name = table.h1.get_text()
    return name.text
```

```
def get_author(table):
    author_name = table.find('p').get_text()
    return author_name
```

```
def get_genre(table):
    if table.find('p') != None:
        genre = table.p.get_text()
    else:
        genre = table.find('p').get_text()
    return genre
```

```
def get_publishing_date(table):
    if table.find('p') != None:
        date = table.p.get_text()
    else:
        date = table.find('p').get_text()
    pattern = re.compile(r'\d{4}')
    year = re.findall(pattern, date)
    return int(year[0])
```

```
def get_pages_count(table):
    pages = table.find('p').get_text()
    return int(pages)
```

```

def parse_wiki_page
    page = rq.get(url)
    return bs(page)

def get_book_info_row(row)
    #To avoid breaking
    try:
        book_soup = bs(row)
        book_table = book_soup.find('table')
    except:
        print(f"Can't parse book info")
        return None

    book_info = {}
    #get info with BeautifulSoup
    values = ['Author', 'Publisher', 'Genre']
    functions = [get_book_info_row, get_book_info_row, get_book_info_row]

    for val, func in zip(values, functions):
        try:
            book_info[val] = func(row)
        except:
            book_info[val] = None

    return book_info

#Get books
url = 'https://en.wikipedia.org/wiki/List_of_fictional_characters'
page = rq.get(url).text
soup = bs(page)
rows = soup.find('table')
books_links = [row['href'] for row in rows.find_all('a')]
base_url = 'https://en.wikipedia.org/wiki/'
books_urls = [base_url + link for link in books_links]

#to store books info
book_info_list = []
#loop first books
for link in books_urls:
    #get book info
    book_info = get_book_info_row(soup.get(link))
    #if everything is ok
    if book_info:
        book_info_list.append(book_info)
    #puase a second book
    time.sleep(1)

#Collect different genres
genres = {}
for book in book_info_list:
    book_gen = book['Genre']

```

```
        if book_gen:
            if 'fiction'
                book_ge
        if book_gen not
            genres[book,
        else:
            genres[book,

print(genres)
#Plot results
bar_chart = pygal.B
[bar_chart.add(k,v)
display(HTML(base_h'
```