

می‌خواهیم تابعی بنویسیم که تعداد لینک‌های یک فایل *HTML* را محاسبه کند.

کریم که نتوانست شطرنج فیزیکی خود را درست کند به اینترنت روی آورده و در این روزهای قرنطینه تصمیم گرفته که با استفاده از اینترنت و فضای آنلاین شطرنج بازی کند. در همان ابتدای استفاده از محیط اینترنت کریم عاشق شمردن تعداد لینک‌های صفحه‌ها شده و می‌خواهد برای صفحه داده شده تعداد لینک‌ها را بشمرد.

هدف ما این است که برنامه‌ای بنویسیم که کار او را راحت کند و با گرفتن یک فایل *HTML* تعداد لینک‌های آن صفحه را محاسبه کند.

جزئیات

- برنامه‌ی شما باید شامل یک تابع به نام `process` باشد.
- تابع `process` به عنوان ورودی نام فایل *HTML* مورد نظر را دریافت می‌کند و تعداد لینک‌های این فایل را برمی‌گرداند (توجه کنید که فایل به صورت لوکال در کنار تست‌ها وجود دارد و شما باید آن را `open` کرده و اطلاعاتش را بخوانید).
- لینک به معنای زنجیر است و برای هدایت کردن کاربران از صفحه جاری به صفحه‌ای دیگر استفاده می‌شود. به این صورت که متن مورد نظر خود را می‌نویسیم و کاری می‌کنیم که وقتی کاربران روی آن کلیک کردند به صفحه مورد نظر ما هدایت شوند.
- برای ایجاد لینک در *HTML* از تگ `a` استفاده می‌کنیم به طور مثال:

```
<a href="https://quera.ir">کلیک کنید</a> | 1
```

در لینک بالا، عبارتی که برای `href` در نظر گرفته می‌شود، همان آدرس صفحه‌ی مقصد است که می‌خواهیم کاربر پس از کلیک بر روی متن "کلیک کنید" به آن هدایت شود.

- در واقع شما باید تعداد تگ‌های `a` فایل *HTML* داده شده را محاسبه کنید.
- در این سوال می‌توانید از کتابخانه‌های مختلف پایتون استفاده کنید و نام آن کتابخانه‌ها را طبق روشی که در قسمت نکات آمده بنویسید تا در هنگام کدنویسی بتوانید آن‌ها را `import` کرده و از آن استفاده کنید.

نکات

- نام فایل پایتون شما باید `solution.py` باشد.
- تست‌های اصلی این سوال، صفحات اصلی سایت‌های مختلف هستند و برنامه شما باید توانایی کارکردن روی همه صفحات را داشته باشد.
- می‌توانید فایل اولیه‌ی خام و کد تست نمونه را با استفاده از این [لینک](#) دانلود کنید.
- برای استفاده از کتابخانه‌های مختلف می‌توانید همراه فایل ارسالی، فایل به نام `python_requirements.txt` بگذارید که در آن نام کتابخانه‌های مورد نیاز و شماره نسخه آن‌ها به فرمت زیر در آن موجود باشد: (اگر شماره نسخه را ننویسید آخرین نسخه آن کتابخانه نصب می‌شود)

```
firstlib==1.2.3
secondlib==4.5.6
...
```

قسمت آموزشی

در این قسمت راهنمایی‌های سوال به ترتیب در روزهای شنبه، دوشنبه و چهارشنبه ساعت ۱۸ اضافه می‌شود. مشکلاتان در راستای حل سوال را می‌توانید از بخش "سوال بپرسید" مطرح کنید.

▼ راهنمایی ۱

سعی کنید از کتابخانه‌های خارجی برای حل این سوال استفاده کنید. می‌توانید با سرچ در گوگل کتابخانه‌های خوبی برای این کار پیدا کنید.

ما استفاده از کتابخانه `beautifulsoup` را پیشنهاد می‌کنیم که یکی از کتابخانه‌های خوب برای حل این سوال است!

▼ راهنمایی ۲

اول از همه کتابخانه `BeautifulSoup` را ایمپورت می‌کنیم.

حالا باید آدرس داده شده در تابع رو باز کنیم و با استفاده از `beautifulsoup` اون رو بخونیم و سعی کنیم از این کتابخونه برای پیدا کردن تعداد لینک‌ها استفاده کنیم و در آخر تعداد لینک‌ها رو برگردونیم!

```
1 from bs4 import BeautifulSoup
2
3 def process(path):
4     with open(path) as html:
5         soup = BeautifulSoup(html.read(), 'html')
6         links = []
7         # Find all links
8         return len(links)
```

▼ راهنمایی ۳

در این مرحله باید تمام `href` ها رو از تگ‌های `a` از طریق کتابخانه `BeautifulSoup` پیدا کنیم که به روش زیر برای آن اقدام می‌کنیم:

```
1 for link in soup.find_all('a'):
2     links.append(link.get('href'))
```

و در نهایت کد ما به صورت زیر میشه:

```
1 from bs4 import BeautifulSoup
2
```

```
3 def process(path):
4     with open(path) as html:
5         soup = BeautifulSoup(html.read(), 'html')
6         links = []
7         for link in soup.find_all('a'):
8             links.append(link.get('href'))
9         return len(links)
```

موفق باشید:دی