Install Guider

1. Python install

<http://ubuntuhandbook.org/index.php/2017/07/install-python-3-6-1-in-ubuntu-16-04-lts/>

Install Python 3.6 via commands:

sudo apt update

sudo apt install python3.6

1. pip requirements

- pip install:

sudo apt-get install python-pip

sudo apt-get install python-dev

- install packages with using commands:

sudo pip install scrapy

sudo pip install request

sudo pip install beautifulsoup4

sudo pip install selenium

sudo pip install googletrans

sudo pip install scrapy\_splash

sudo pip install urllib3

- Chromedriver install

You must install chromedriver for scraping category data from otto.de. The reason is why we use python selenium for scraping.

Please open the terminal , and execute the commands.(You can reference this page <https://tecadmin.net/setup-selenium-chromedriver-on-ubuntu/> for chromedriver install.)

sudo apt-get -y update

sudo apt-get -y install google-chrome-stable

wget [https://chromedriver.storage.googleapis.com/2.41/chromedriver\_lin ux64.zip](https://chromedriver.storage.googleapis.com/2.41/chromedriver_linux64.zip)

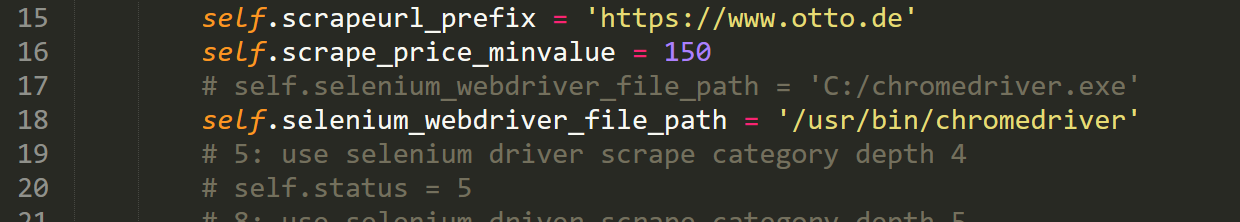
unzip chromedriver\_linux64.zip

sudo mv chromedriver /usr/bin/chromedriver

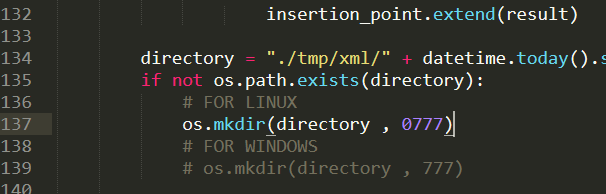
sudo chown root:root /usr/bin/chromedriver

sudo chmod +x /usr/bin/chromedriver

And then you should check line-18 in ./scrapotto/spiders/scrape\_category.py file, it must be self.selenium\_webdriver\_file\_path = '/usr/bin/chromedriver'. so our python scraping works right in linux OS.



And you should check line-137 in ./scrapotto/spiders/generate\_xmdata.py , the mkdir setting must be 0777 so that the error “permission denied” will not be occured in command execute “scrapy crawl generate\_xmldata”.



You should update spider files .

Thanks very much.