DATA SCIENCE

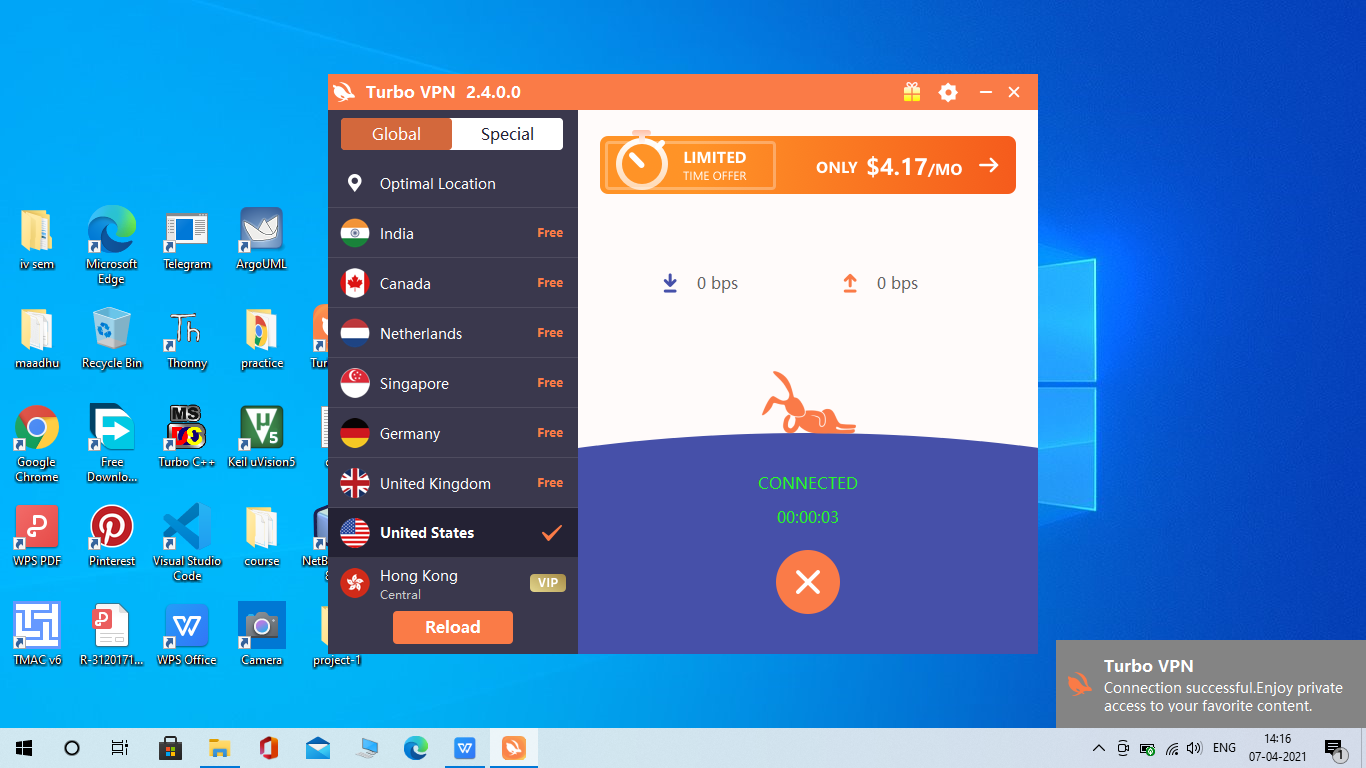
UNSTRUCTURED TO STRUCTED USING WEB SCRAP

2019272037 - SRILAKSHMI P

2019272027 - NOORUL AMEEN

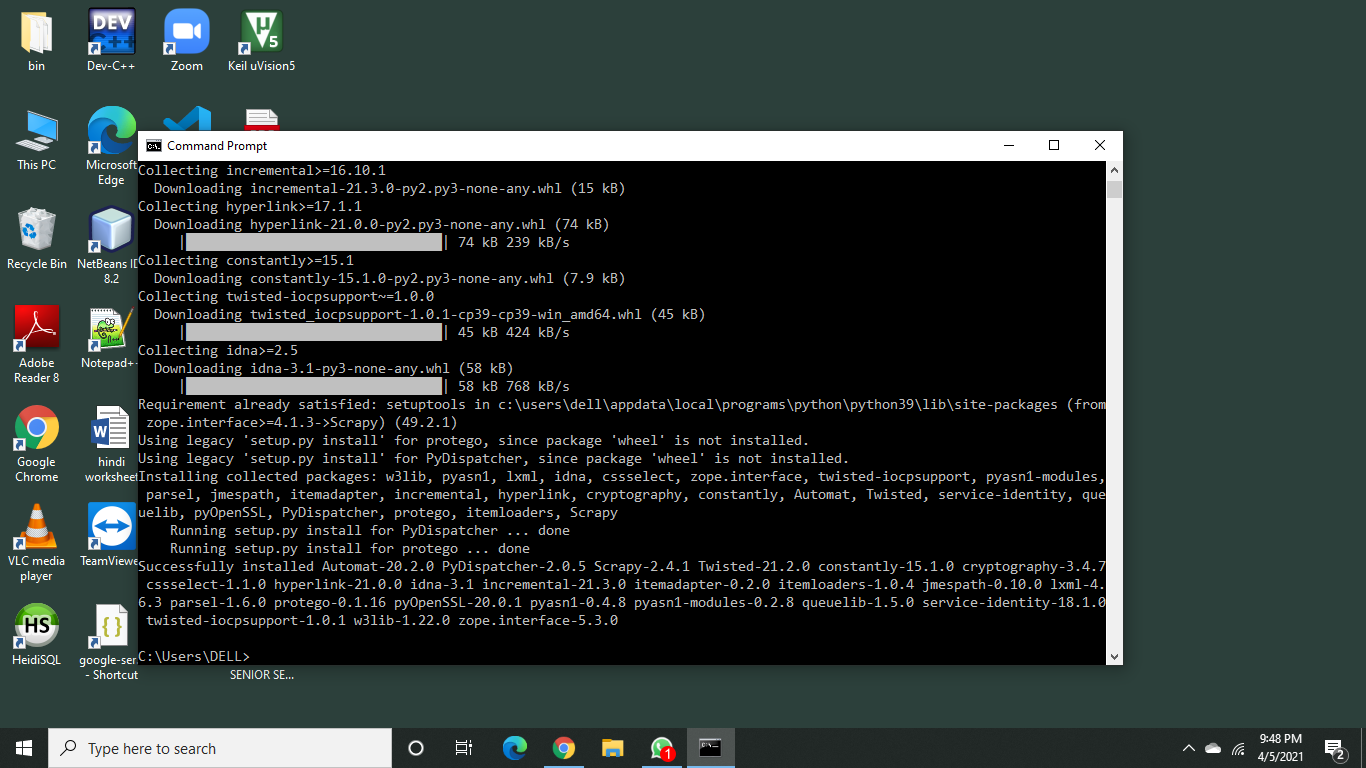
2019272020 - MADHUMITHA S

INSTALLING VPN :



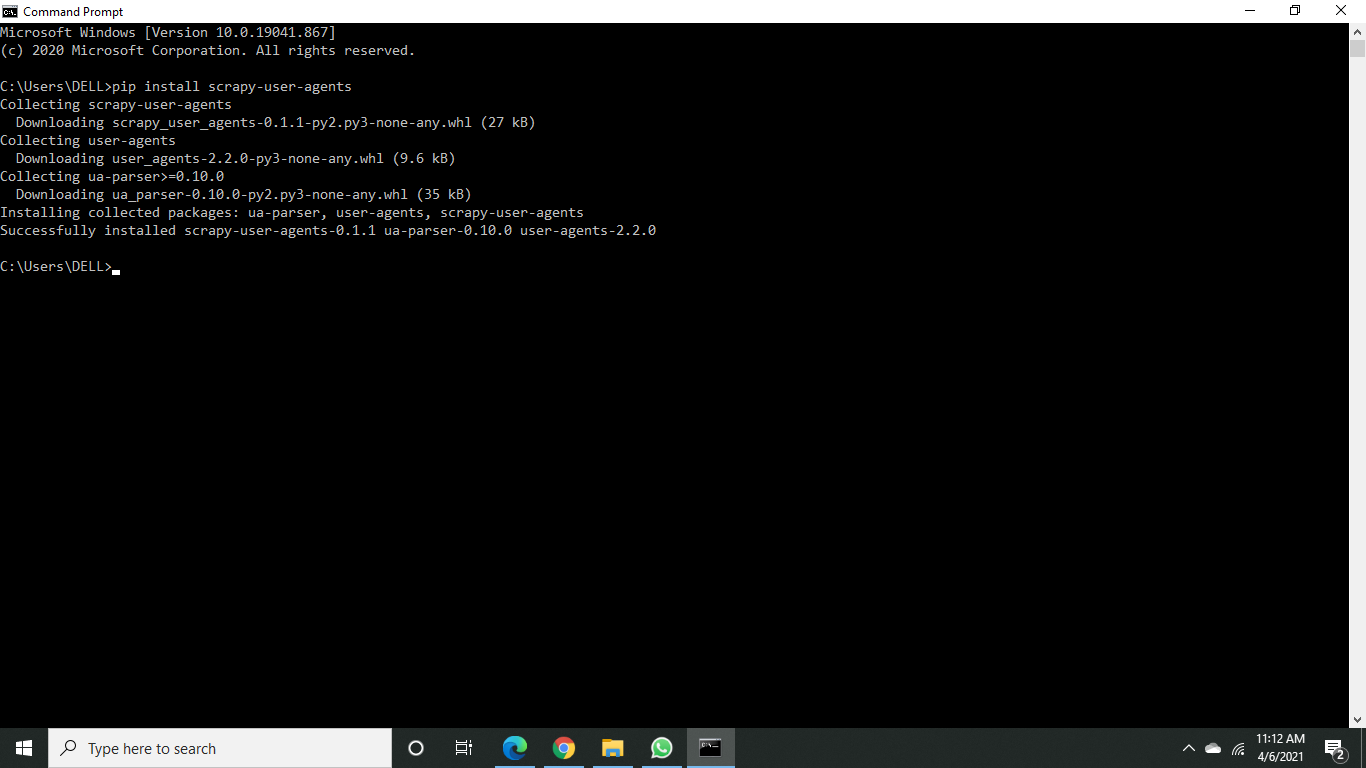
INSTALLING SCRAPY :

Pip install scrapy



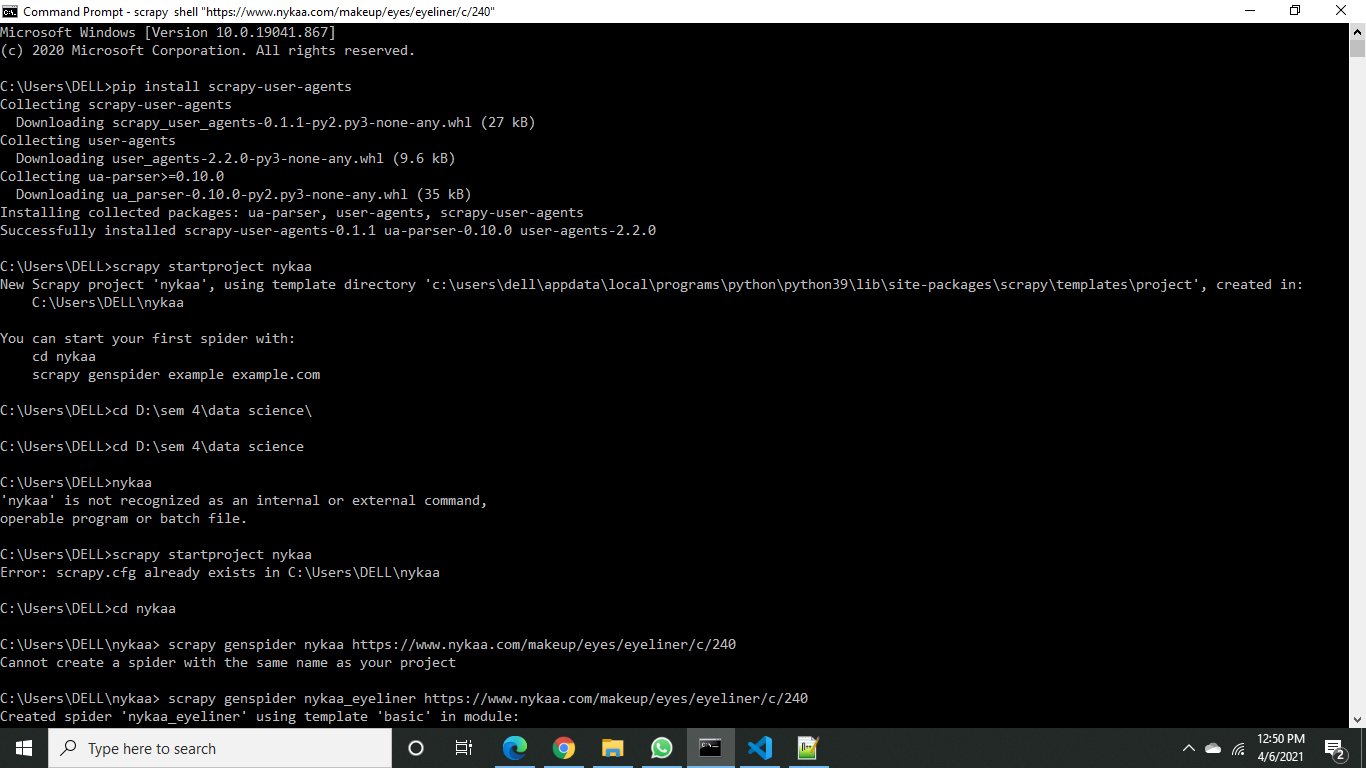
INSTALLING SCRAPY USER AGENT :

Pip install scrapy\_user\_agents

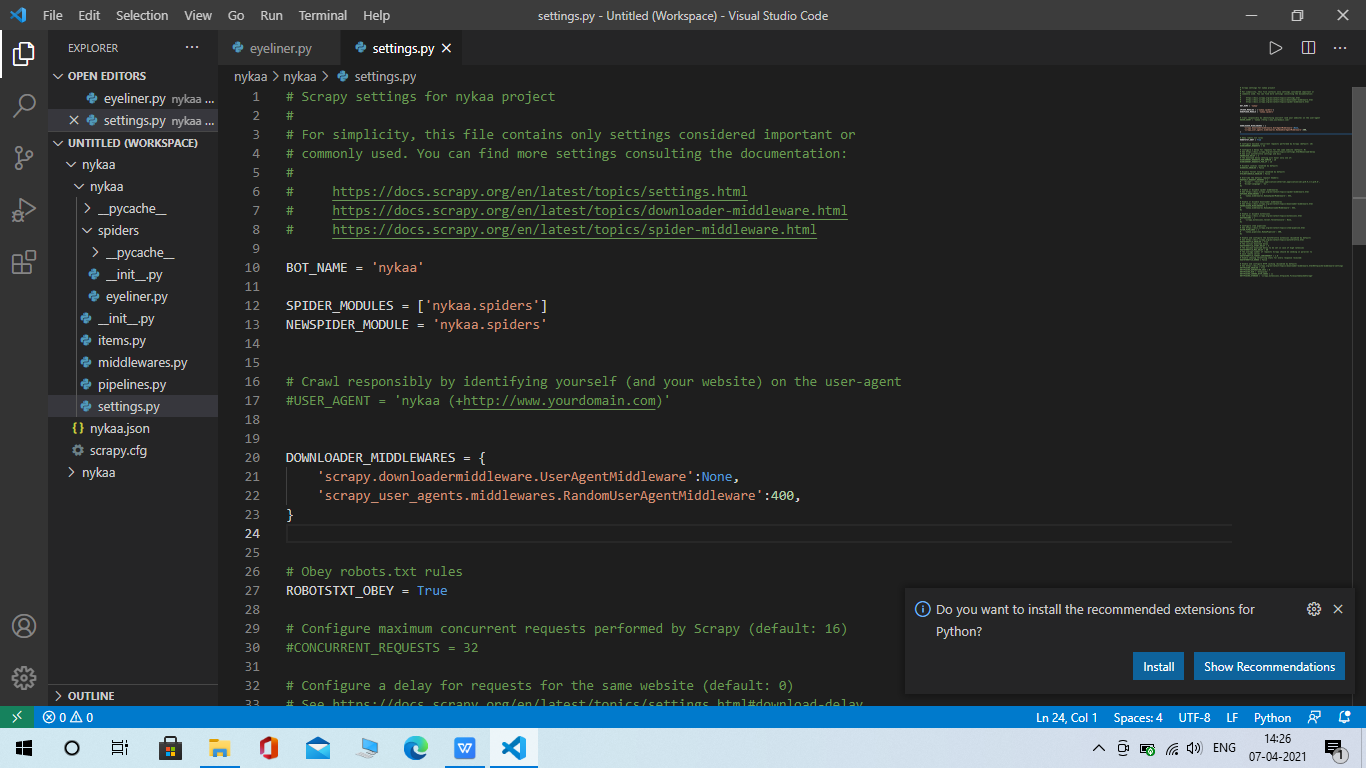


CREATING PROJECT IN SCRAPY :

Scrapy startproject project\_name

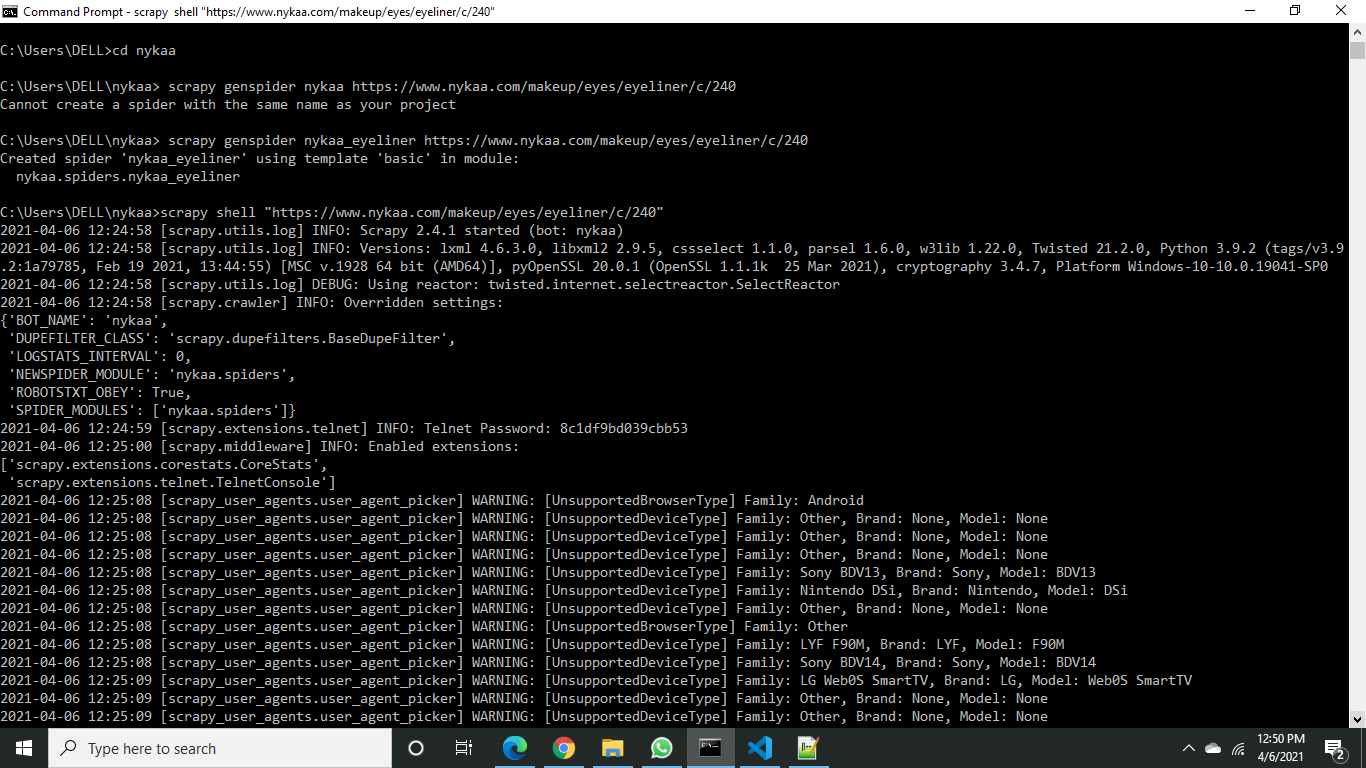


EDIT SETTINGS.PY IN PROJECT:

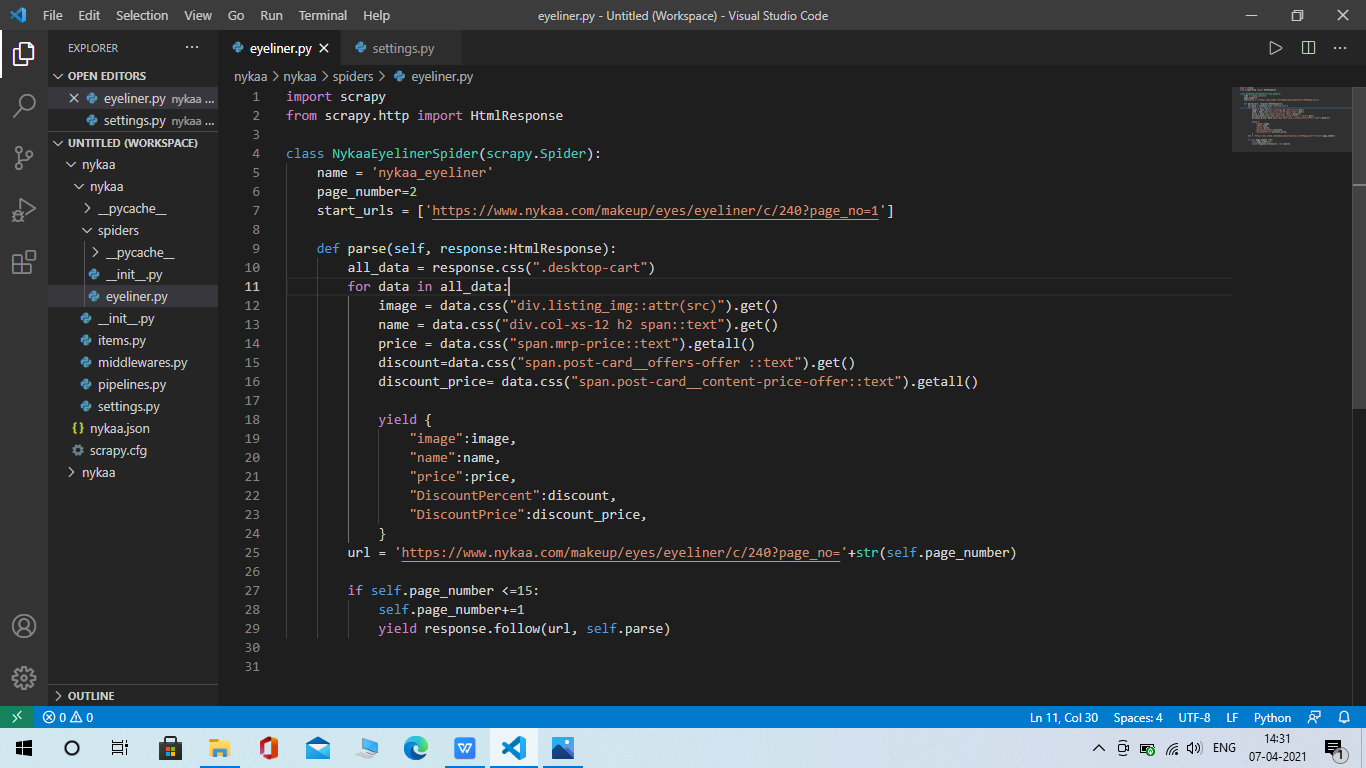


CREATE SCRAPY CRAWL GENSPIDER APP:

Scrapy genspider spider\_name “link of website”

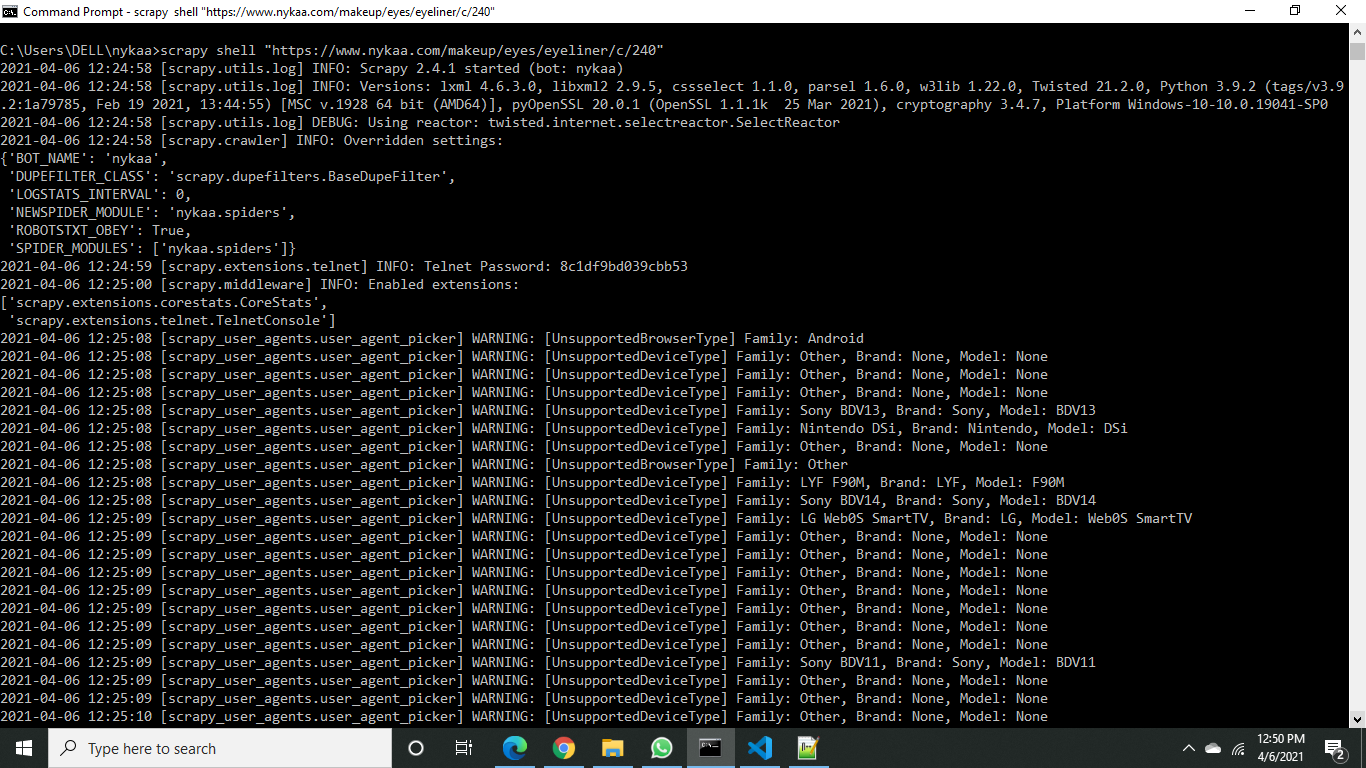


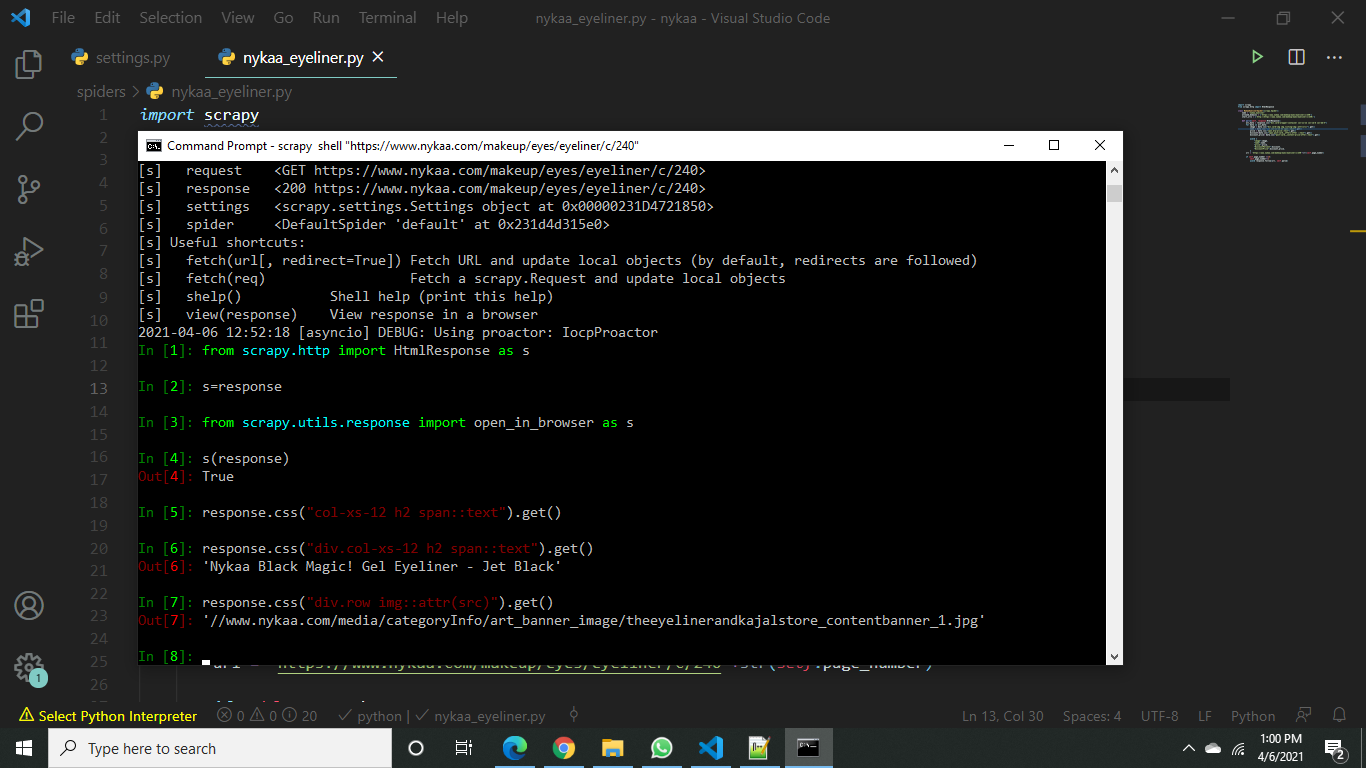
CODE FOR SCRAP DATA FROM WEBSITE:



START SCRAPY SHELL :

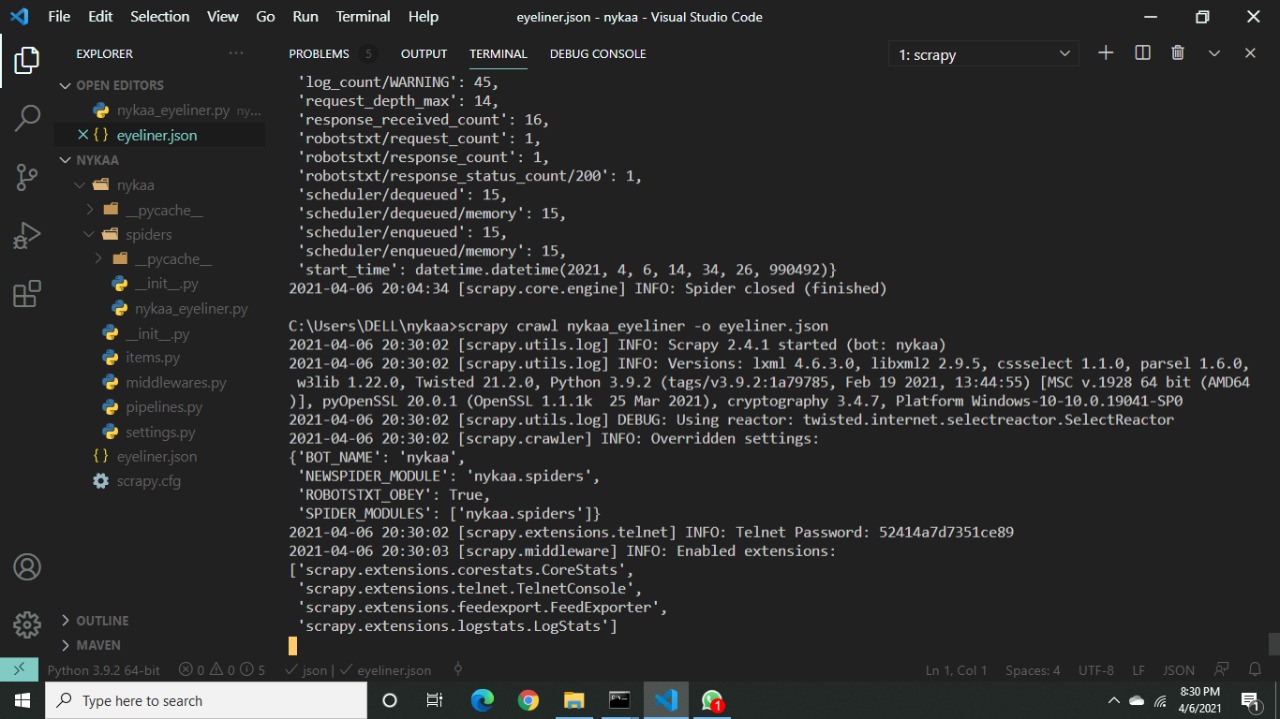
Scrapy shell “link of website”

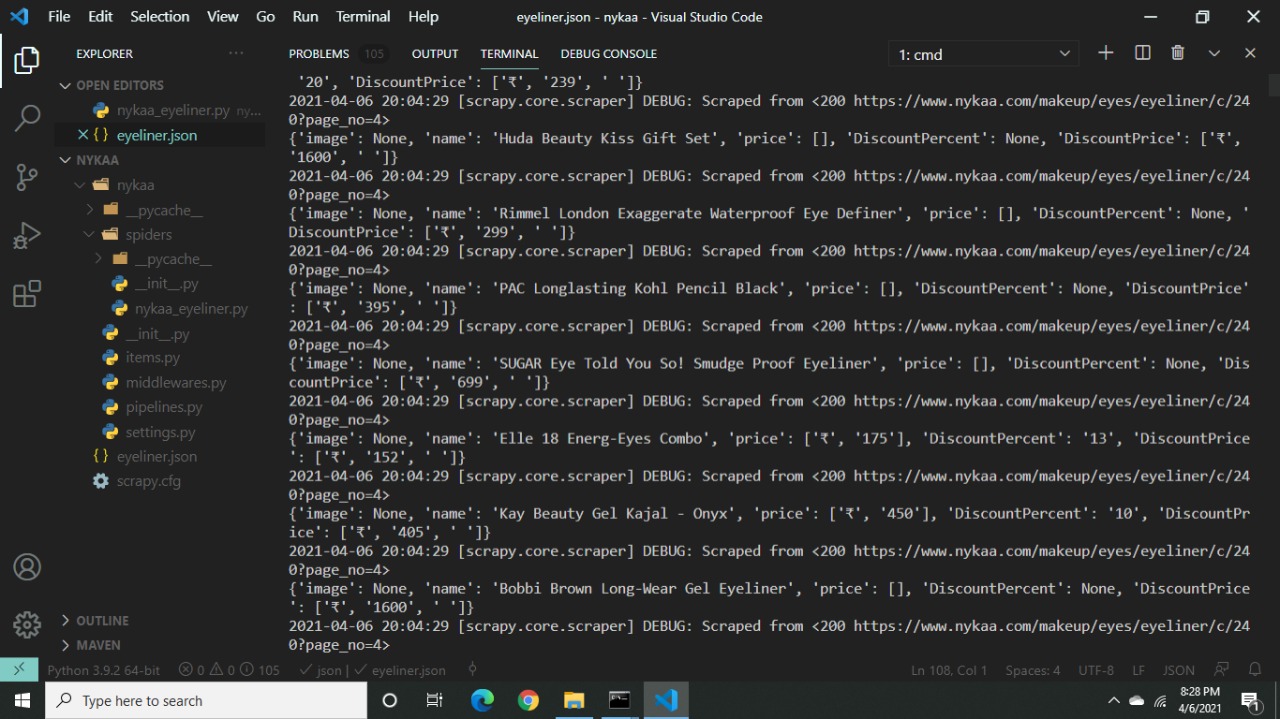




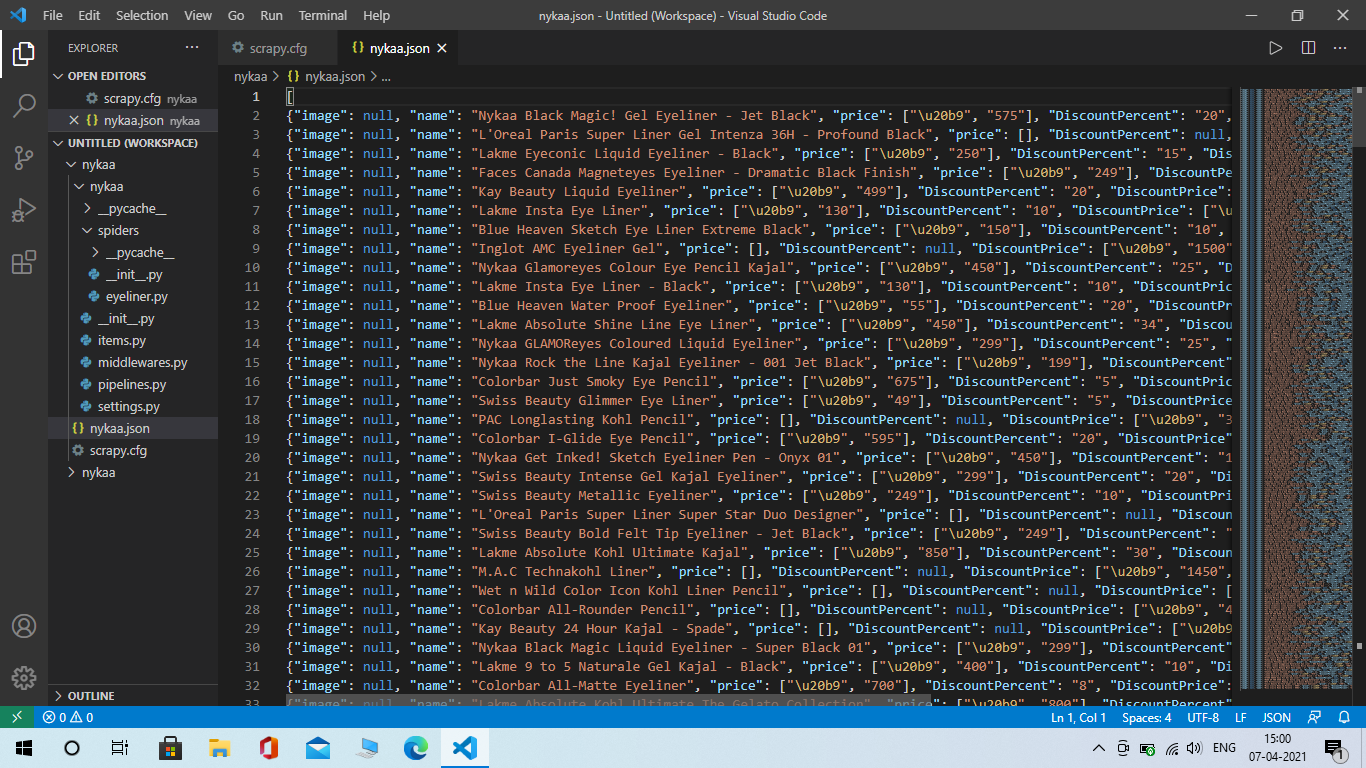
SCRAPY CRAWL AND GET DATA AS JSON :

Scrapy crawl spider\_name -o file\_name.json

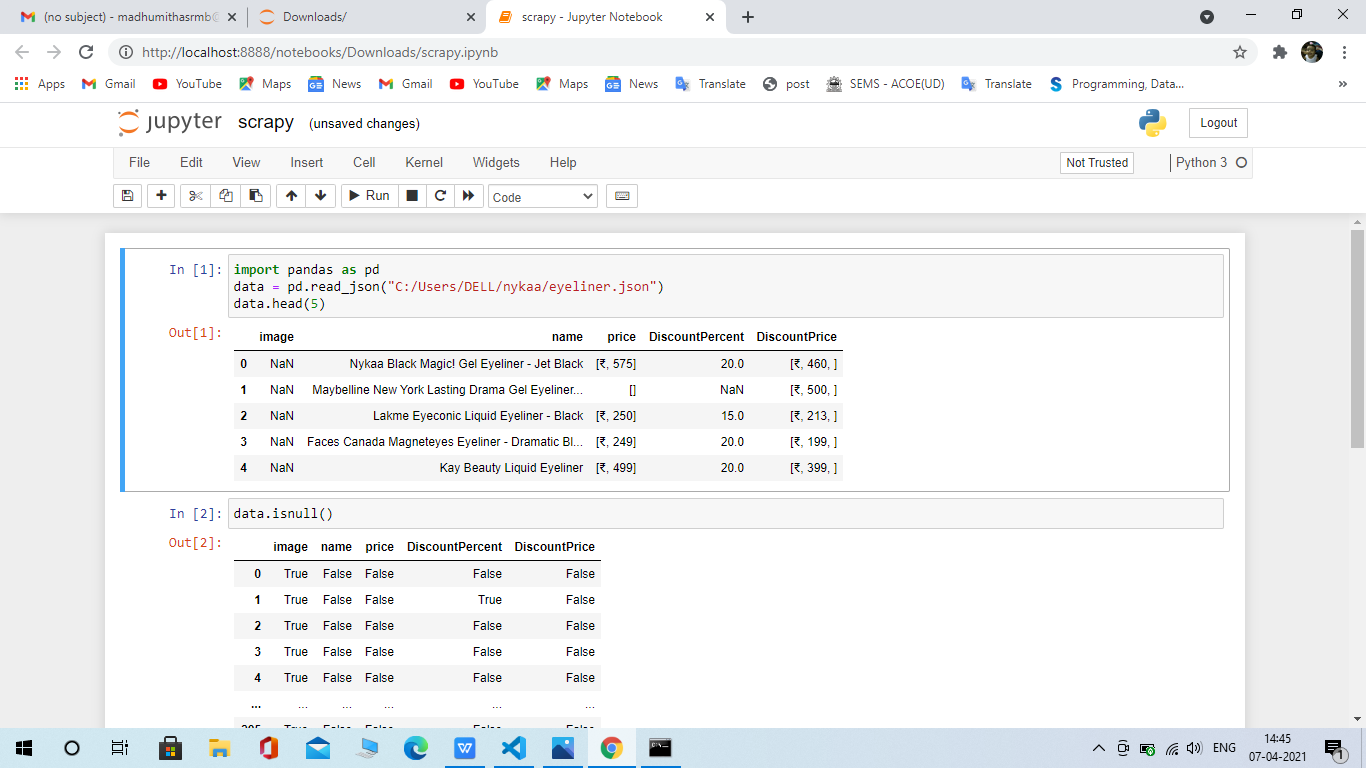


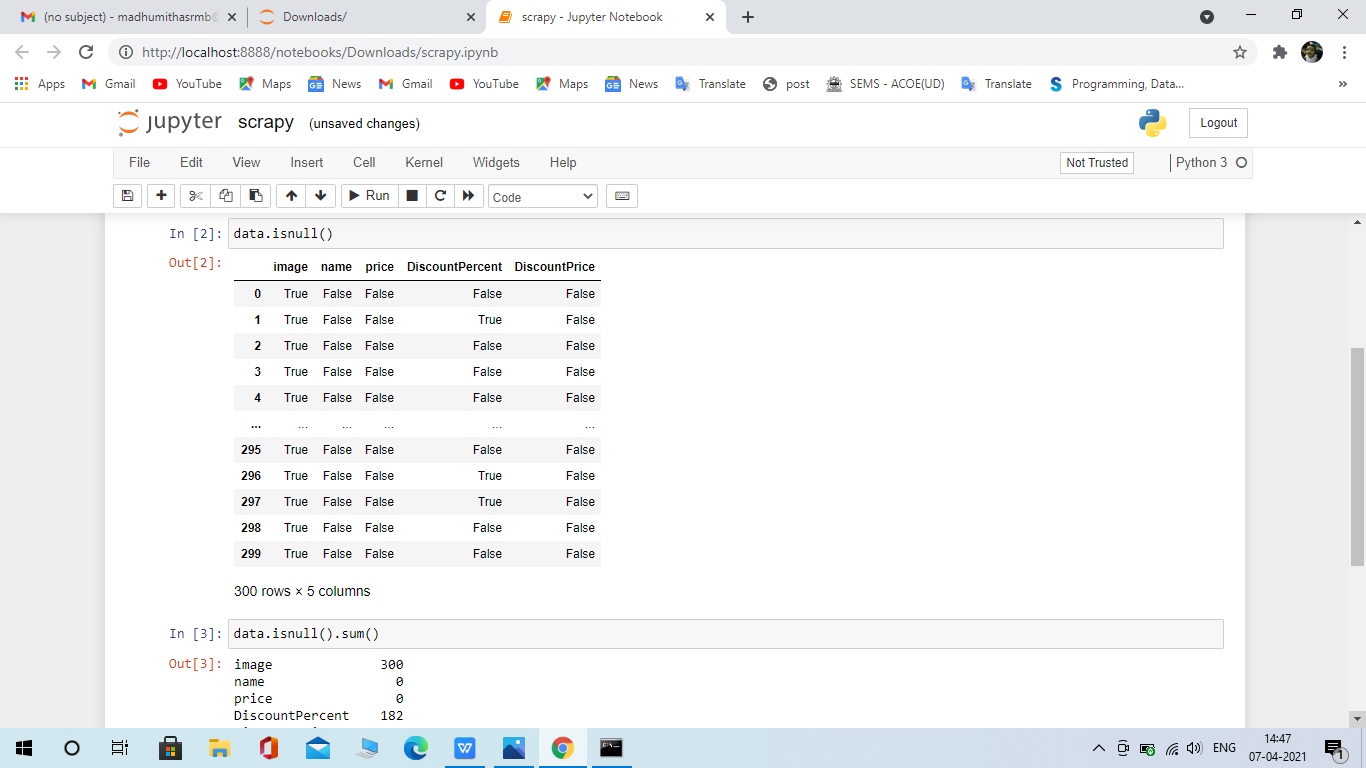


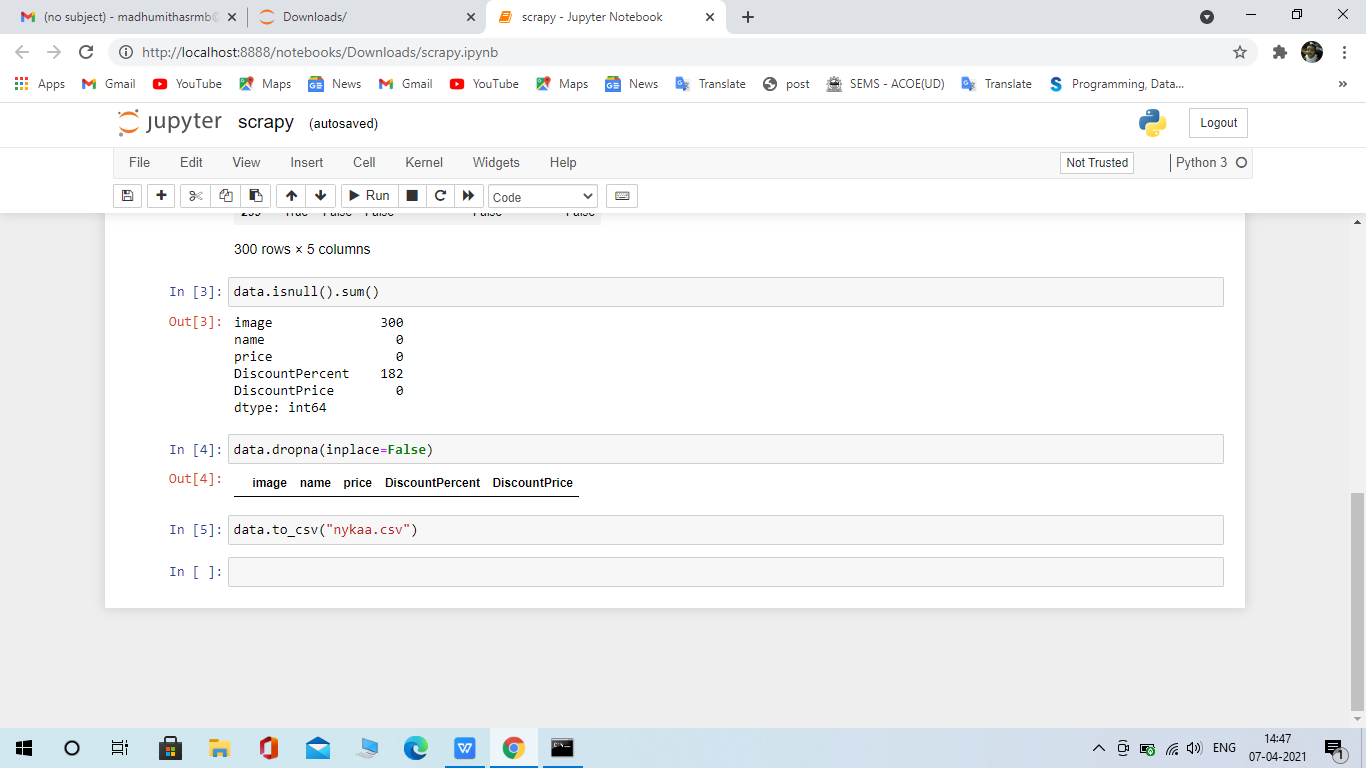
JSON FILE :

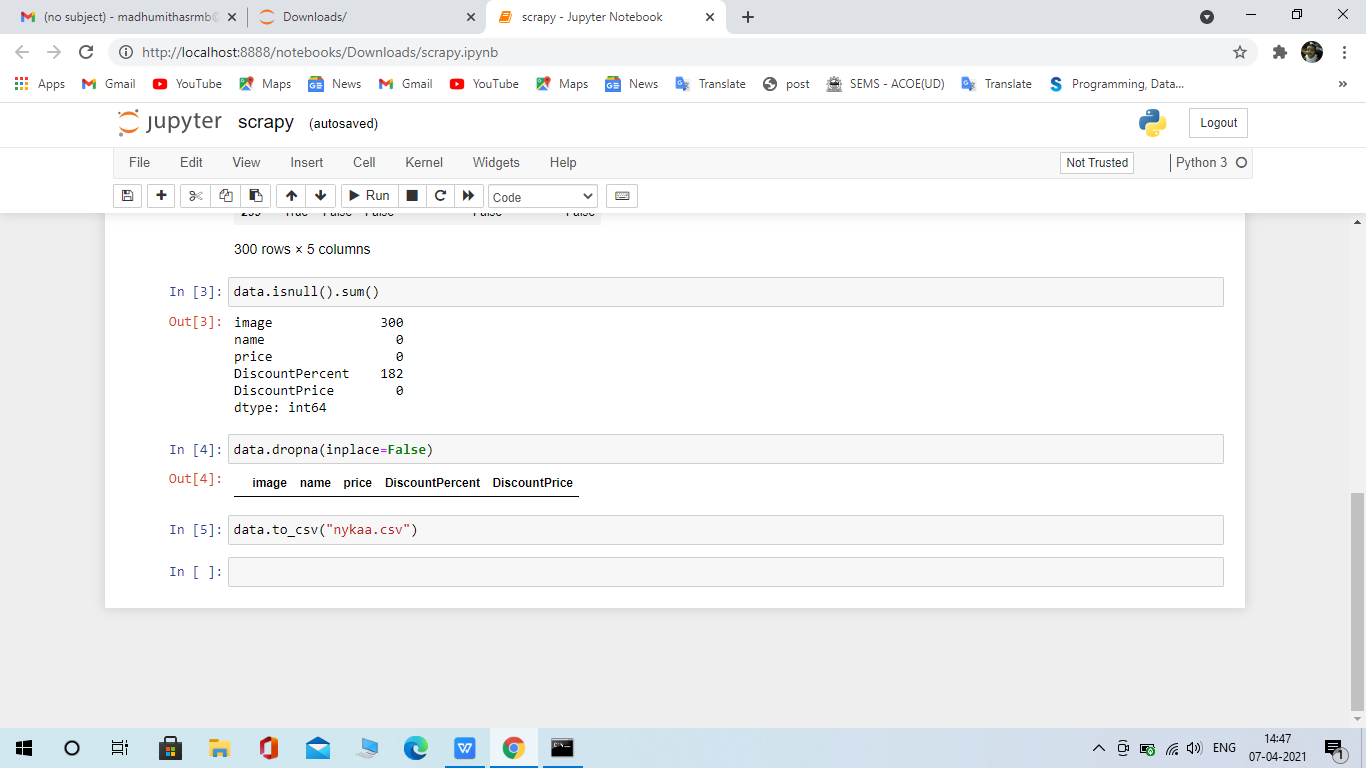


DATA ANALYSIS USING JSON FILE:









CONVERTED JSON TO CSV FILE :

