← Go Back

Project 1(Scrapy and MySQL)

Dear Students, this is the project of Week 4 and Week 5. We have already mentioned the same in first lecture of Week 4.

General Guidelines -

- **The coding needs to be done in Jupyter Notebook (.ipynb) file.**
- **For submitting your assignment you are required to upload a single zip file containing all the files(code file, data files, etc) related to your solution.**
- **Save your assignment with example like- (project1_yourname.zip)**
- **Submit your Project by the end of Week 5 (please check the website for Submission Date)**

We all are very curious and keep ourselves updated regarding what is the current trend? what is the upcoming technology? Which is the most paid job? Which field is booming up? What would be my next dream designation?

Naukri.com can be a good answer to know more about these questions related to jobs and their trends. Naukri.com is a website which regularly post the current requirement based on technology, skills, experience, industry, location etc.

As a part of this Project your task is to-

1. Scrape the Job Listing for Data Analytics position in and around NCR. The scraped data should consists of -
• Job Title
• Experience Required
• Location
• Company Name
• Link to the Detailed Page/Job Description Page
• Keyskills
• Job Description
• Salary
• Posted By (optional)
• Posted on (optional)

2. Store the Data in a new file called
“naukri_dataanalytics.csv” file

3. From the above csv file (in step 2), store the Data in the MySQL Database
a) Create a Database called edulab

b) Create a table called DataAnalyst_ncr containing following columns whose data will be fetched from the csv file
1. Job Title
2. Experience Required
3. Company Name
4. Link to the Detailed Page/Job Description Page
5. Keyskills
6. Job Description
7. Salary
Apart from above columns, DataAnalyst_ncr table will contain below columns -

Job_id will contain integer number as the Primary Key starting from 0
last_updated_on will store the date and time on which row is added/updated in the Database

c) List all the commands in a sql file called naukri_data.sql in order to accomplish the above task in MySQL

4. Create a table called location_jobs table containing the following columns
a) Job_id – Primary Key, Foreign Key (linked with Job_id of DataAnalyst_ncr table)
b) locn_id – Primary Key
c) location – containing location(city) corresponding to a job. If a job has multiple locations in the csv file, these should enter as the separate rows in this table. Job_id and locn_id will be composite Primary Key of this table

Note - List all the commands in a sql file called naukri_data.sql in order to accomplish the above task (given in Steps 4 and 5) in MySQL

5. Write a program in Python which will read the csv file and store the data in the above tables.

Posted on: 17 Sep 2018
Last Submission Date: 28 Sep 2018

Click here to Submit

EduGrad

FEATURED PROGRAM

Python

COMPANY

About us

Blog

SUPPORT

Contact Us

Term Of Use

Legal & Privacy

Copyright