

TABLE OF CONTENTS

I.	EXECUTIVE SUMMARY	2
II.	INTRODUCTION: STREAMLINING JOB SEARCH	2
III.	BUILDING THE WEB SCRAPER	3
IV.	DATABASE ARCHITECTURE	5
V.	UNLOCKING THE BUSINESS IMPACT	7
VI.	FUTURE SCOPE: IMPROVING SEARCH FILTER	9
VII.	CONCLUSION	9
VIII.	APPENDIX I-III	10

EXECUTIVE SUMMARY:

In this project, we have successfully scraped LinkedIn job data using Python, BeautifulSoup, and Selenium, and stored the information of over **1000 jobs** in a MongoDB database. This data contains valuable information on the job location, date posted range, description, company URL, job URL, industry, skills, qualifications, and experience required for different types of jobs in various industries, which can be analyzed and used to inform hiring and training strategies.

The personalized filtering option we developed allows users to filter job listings based on specific criteria, such as keywords, location, job type, skills, or industry. This means that job seekers can easily find and apply for jobs that match their requirements, and companies can identify and target candidates with the most in-demand skills. Companies can also use this data to tailor their job postings. The hiring team can gain insights into the competitor job postings that offer a better package, or how they phrase their job postings to attract the talent pool. By using this data to inform their hiring and training strategies, companies can stay competitive in their respective markets and ensure that they have the right talent in place to drive growth and innovation. Overall, our LinkedIn job scraping project provides a valuable resource for companies, job seekers, and recruiters alike.

INTRODUCTION: STREAMLINING JOB SEARCH

LinkedIn is the world's largest networking platform which is professional. It contains dense information on job openings, industry trends, and talent management. In a world of competitiveness and a tough job market, job seekers need to stay ahead of the curve by identifying the most in-demand skills and qualifications required for the jobs that are of their interest and the industries they wish to step into.

Scraping job data from their job sites can provide students and professionals with valuable insights into the job market, including information on the job type, job post date, the most common keywords and phrases used in job descriptions, the programming languages in demand and technical skills required to

land that dream role. This information can also be used by hiring strategies to target candidates with the most in-demand skills and develop training programs to enhance the skills of the existing employees.

The use of personalized filtering options allows job seekers to quickly and easily find job openings that match their qualifications and interests. By filtering job listings based on specific criteria such as job type, location, or skill requirements, job seekers can streamline their job search and focus on the opportunities that are most relevant to their career goals.

In addition to business and job seekers, recruiters can leverage this data to identify top talent in their industry and develop targeted recruitment plans and strategies to source the best talent pool. Overall, LinkedIn job scraping provides valuable wealth to businesses and job seekers.

BUILDING THE WEB SCRAPER

Automating the Job Search:

The web scraping application prompts the user to input a job of interest and stores it as a variable. It then initializes the Selenium web driver to access the LinkedIn website and navigates to the jobs section using the `click()` function. On clicking a job listing, a new page loads, where the user's input is entered into the search bar, and the enter key is pressed to generate a list of potential jobs.

The “**pyautogui**” module is then used to scroll through the job results page, allowing the application to access more job listings than can fit on a single page. When the end of the page is reached, the `click()` function selects the “**Show More Jobs**” button, loading additional job listings. Using a for loop, the above process is repeated till we load the desired number of jobs. We now have automated the process of searching for job opportunities on LinkedIn making it easier for the user to find relevant roles based on their interests.

Please refer to **Appendix II** to see the results page.

Downloading HTML pages:

After loading all the jobs, the web scraping application downloads the source HTML page as a local file. It then opens this file and creates a BeautifulSoup object to extract the href tag containing the URLs of all the job listings. These URLs are stored in a list called "JobLinks" in the source code. The application then iterates through the JobLinks list and sends a GET request to each URL to access the individual job page. Using BeautifulSoup, it creates a new object to extract information such as the job title, company, location, job description, and any other relevant information.

The individual job page is then downloaded as an HTML file. The web scraping application was used to search for multiple job titles such as **data analyst, data scientist, product manager, and data engineer**. The application was run multiple times using a for loop, allowing it to access and download the HTML pages for more than **1000 job listings**. This approach allowed the application to capture a broad range of job opportunities that matched the user's interests. These HTML pages can be used to parse the required information about the job details.

Parsing the job contents:

After downloading the HTML pages, the web scraping application used a for loop to open each page and create a BeautifulSoup object to parse the contents. Using the BeautifulSoup object, the application accessed various elements of interest on each job page, such as the job title, company name, location, and job description. It found the corresponding **CSS selectors** for each element and extracted the information, storing it in a list.

The parsing step was repeated for all job information, including the name of the job, the company name and URL, the location of the job, the number of applicants, the name and title of the recruiter, the recruiter's LinkedIn URL, job description, type of job, industry, level, function, a job posted date, and the date of job download. Additionally, the application also extracted similar job links that may be of interest

to the user. By parsing the HTML pages in this way, the application was able to extract a wealth of data about each job listing.

DATABASE ARCHITECTURE

Storing scraped job data in MongoDB: Why and How?

From the web scraping application that we developed, we extracted a large amount of job data from LinkedIn, including details such as job titles, company names, locations, and job descriptions. Storing this data in a structured manner is important for organizing and managing the data for intended use in the future. We have chosen MongoDB as the database to store all the job information. It is a famous choice for storing huge amounts of data in a structured manner. It is a NoSQL document database that is known for its flexibility and scalability of data storage.

MongoDB allows for the storage of unstructured data in a document-oriented format. Each document in the database can contain different fields and data types, which makes it ideal for storing the diverse range of job details that were extracted from the LinkedIn job listings.

MongoDB is also known for its scalability and performance, making it well-suited for storing large amounts of data. As our web scraping application downloaded over 1000 HTML pages, we needed a database that could handle this volume of data without slowing down or experiencing performance issues. Overall, MongoDB provides an efficient and scalable solution for storing large amounts of unstructured data, which makes it a suitable choice for storing the job details extracted by our web scraping application.

Designing the LinkedIn database:

To store the job details that were extracted by our web scraping application, we created a database in MongoDB called **LinkedIn_DB**. We used the “**PyMongo**” module to connect our Python environment to the MongoDB instance running on Docker, using port **27017**.

Inside the LinkedIn_DB database, we created **four collections** to store the different categories of job details that we extracted.

1. The **Jobs_Basic collection** contains basic information such as the job ID, title, company name, location, number of applicants, date posted, job URL, and company URL.
2. The **Jobs_Advanced collection** contains more advanced information such as the job level, type, function, and industry, as well as links to similar jobs. This collection helps to provide a more comprehensive picture of the job landscape.
3. The **Jobs_Recruiter collection** contains information on recruiters associated with each job, including their name, title, and profile URL. This information is useful for networking and for understanding the hiring process associated with each job.
4. Finally, the **Jobs_Description collection** contains the job description for each job, which is a key piece of information for understanding the requirements and responsibilities associated with each job.

In addition to the above-mentioned columns, we also have the column signifying the date of download of the job for each collection. By dividing the job details into these different collections, we can organize the data in a structured and meaningful way that allows for efficient querying and further analysis.

Please refer to **Appendix III** for screenshots of the above collections in the database.

Tailoring Job searches to user preferences:

The personalized job filtering feature requires users to fill out a form that prompts them to provide specific information about the type of job they are looking for. Users are asked to specify the job title they are interested in, such as "Data Scientist", "Data Analyst", "Engineer", or "Product Manager". They are also asked to indicate the location they prefer, such as "CA", "San Francisco", "NY", or "United States", as well as the type of job they are seeking, such as "Full-time", "Internship", or "Contract". Additionally, users are prompted to select their preferred industry, which could include "Software", "Technology", "Health Care", or "Music".

After collecting this information, the feature uses the MongoDB LinkedIn job database to search for job listings that match the user's specified criteria. By doing so, the feature helps to personalize the job search experience for each user, allowing them to find jobs that are relevant to their unique interests and qualifications. The filtered jobs can be downloaded in the **filtered_jobs.csv** file.

UNLOCKING THE BUSINESS IMPACT

Leveraging Job data for Efficient Analysis:

The dataset that we illustrated in the previous section can be used to solve multiple business-related questions. Job seekers and businesses can use the dataset of job details that we have extracted and stored in MongoDB to answer various business-relevant questions, such as:

- What are the most common job titles and locations for data-related roles?
- What are the typical requirements and qualifications for data-related job roles?
- What are the top companies hiring for data-related roles, and what are their hiring patterns?
- What are the job trends and patterns in the data-related job market?
- How does the demand for data-related jobs vary across different industries and sectors?

The structured storage of the data in MongoDB allows for efficient querying and analysis, enabling us to quickly and easily extract insights from the data to answer these and other business-relevant questions. Additionally, MongoDB's scalability and flexibility make it easy to add and update data as new job listings become available.

This structured storage allows us to answer important business questions related to industry trends, competitive analysis, candidate sourcing, and diversity and inclusion.

By utilizing MongoDB, organizations can gain valuable insights into emerging job roles, new skill requirements, and changes in compensation across various industries. Competitive analysis can be

conducted to gain insights into their hiring strategies, enabling organizations to refine their own recruitment strategies accordingly. Additionally, candidate sourcing can be improved by leveraging the vast amount of data available through LinkedIn and stored in MongoDB.

Furthermore, MongoDB can help organizations promote diversity and inclusion by identifying patterns of bias and discrimination in job postings and helping to ensure that job listings are inclusive and welcoming to all qualified candidates.

Overall, MongoDB's scalability and flexibility make it a powerful tool for analyzing large volumes of unstructured job data, enabling organizations to gain valuable insights into the job market and improve their recruitment efforts.

Why MongoDB over alternative databases?

Compared to alternative databases such as traditional relational databases, MongoDB's document-oriented approach is well-suited for storing unstructured data such as job listings, which can have varying fields and data types. This allows us to store the diverse range of job details that were extracted from the LinkedIn job listings in a way that is efficient and scalable. Additionally, MongoDB's ability to scale horizontally across multiple servers ensures that the database can handle large volumes of data without performance issues.

The design choices we made for the MongoDB database were driven by the need to organize the job details in a structured manner, making it easy to query and analyze the data. By dividing the job details into four different collections, we were able to store the information in a way that is meaningful and efficient for analysis. For example, the Jobs_Advanced collection allows us to understand the more nuanced aspects of each job listing, such as the job level, function, and industry, which can provide deeper insights into the job market trends.

FUTURE SCOPE: IMPROVING SEARCH FILTER

The current implementation of adding salary range to the Jobs_Description collection using regex has resulted in values that signify yearly as well as hourly pay rates, indicating inaccuracy in the extraction process. As a future scope, a more precise regex function can be developed to accurately extract the job salary range from the job descriptions. This function can handle different formats such as hourly pay rate, base pay, and upper and lower limits for the annual pay rates. The extracted values can then be converted uniformly to hourly or annual pay rates and stored in the database. This improvement can enable the search filter to match user input values with job offerings based on their expected salary.

CONCLUSION

This project involved the ethical web scraping of LinkedIn job information using tools like Selenium and BeautifulSoup. The collected data was stored in a MongoDB database, and we discussed the advantages of this database architecture. The potential uses of this valuable data in the hiring industry and talent businesses were also explored. As a future scope, we proposed an improved approach to sourcing salary information from job descriptions using a precise regex function.

APPENDIX I

Since we are using **ipywidgets** and **IPython.display** packages specifically designed for Jupyter Notebook and other Jupyter environments, we request you to use the Jupyter Notebook for executing the code and not the .py file. The Source code has four user-defined functions in the main function. They are

1. ScrapingToMongoDb()

2. Generate_Form()

3. Filtering()

4. Future_Salary_Filter().

Please run only the functions you are interested in.

The first function **ScrapingToMongoDb()** when run will prompt you to enter a job title you are interested in.

Enter the Job you are interested in:

The **filtering ()** function will prompt you to input the following values.

What Job are you looking for?	Data Scientist/Data Analyst/Engineer/Product Manager
Enter the keyword you are looking for in the Job ...	SQL/Python/Mathematics/Computer Science/AI/neural networks
Which Location?	TX/OH/FL/CA/San Francisco/NY/United States
What type of job are you looking for?	Full-time/Internship/Contract
Which Industry do you prefer?	Software/Technology/Health Care/Music

APPENDIX II

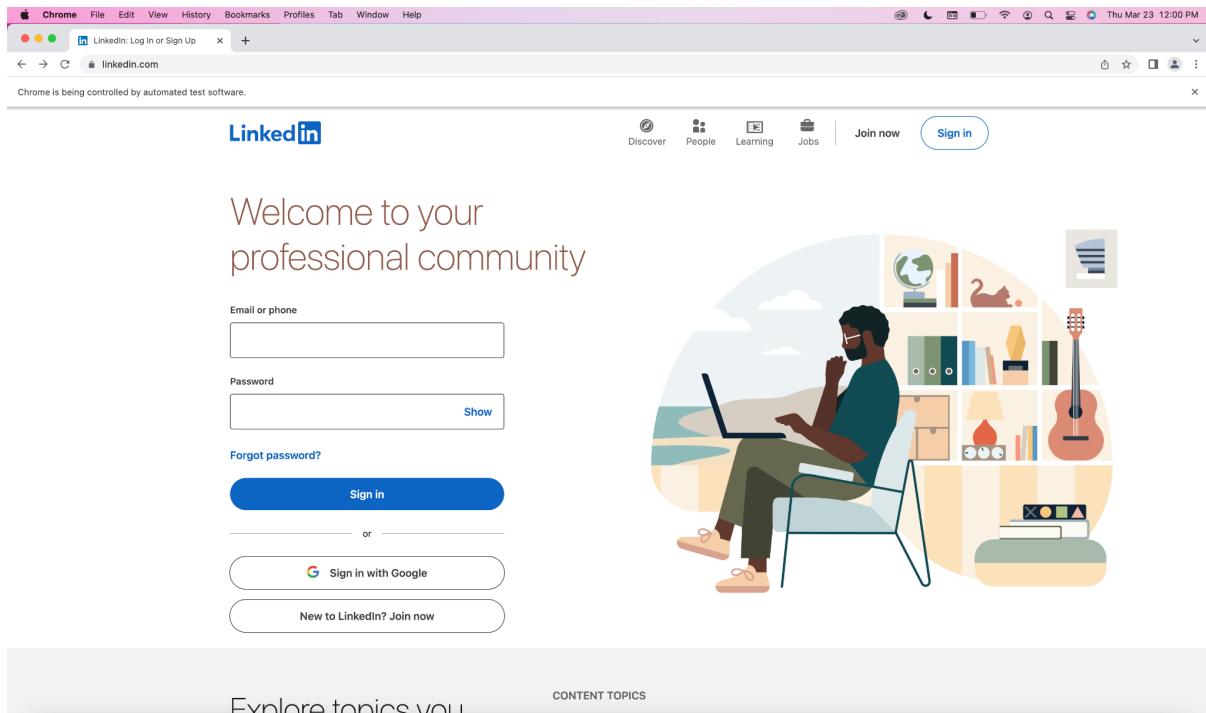


Fig 1.1 LinkedIn Main Page

A screenshot of the LinkedIn job search results page for "Data Analyst" jobs in the United States. The top navigation bar shows the search term "262,000+ Data Analyst jobs in" and the URL "linkedin.com/jobs/search?keywords=data%20analyst&location=United%20States&geolId=103644278&trk=public_jobs_jobs-search-bar_search-submit&position=1&pageNum=0". Below the search bar are notifications about automatic updates and being controlled by automated test software. The main content area displays a list of job postings from various companies, including Red Carrot, Software Technology Inc., Staffing Opportunities Solutions, Inc., Woodland Foods, and Atlas Search. Each listing includes the company name, location, salary range, hiring status, and posting age. To the right of the job list is a detailed view of a job posting for a Data Analyst at Red Carrot in Washington, DC, posted 1 week ago with 41 applicants. The posting details the role's responsibilities and requirements. A sidebar on the right encourages users to sign in or join LinkedIn.

Fig 1.2 LinkedIn Job Results

APPENDIX III

Quickstart × JOBS_Basic ×

MongoDB connection (localhost:27017) > LinkedIn_DB > JOBS_Basic

Run Load query Save query Query history Set default query Copy Paste Visual Query Builder

Projection {} Sort {} Skip {} Limit {}

Result | Query Code | Explain

Documents 1 to 50 Table View

JOBS_Basic > ID

_id	ID	TITLE	COMPANY	LOCATION	#Applicants	Date Posted	Job URL	Company URL	Date_Downloaded
6417513abcfcf86	1	Data Scientist	Sunlighten	Overland Park, KS	null	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	2	Data Scientist	Karge	Waterford, WI	null	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	3	Data Scientist F	Bardess Group L	New York, NY	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	4	Data Scientist I	Hebrew SeniorLi	Boston, MA	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	5	Data Scientist	Sutoer Solutions	Fort Mill, SC	null	2 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	6	Junior Data Scie	HireMatch	Miami, FL	null	2 hours ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	7	Junior Data Scie	Dezign Concept	McLean, VA	null	3 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	8	Machine Learnin	Yakoo	San Francisco, CA	null	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	9	Data Scientist	Spearehead Staff	United States	Over 200 applicants	2 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	10	Data Scientist In	Anika Systems	Leesburg, VA	Over 200 applicants	4 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	11	Data Engineer/D	WalletConnect	New York, United States	null	5 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	12	Data Scientist	The Lab Consult	Greater Houston	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	13	Data Scientist	UCLA Health	Los Angeles, CA	null	6 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	14	Data Scientist	Unlearn.AI	San Francisco Bay Area	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	15	Junior Data Scie	Gryphon Scienti	Chillum, MD	Be among the first 25 applicants	18 hours ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	16	Data Scientist II	Hebrew SeniorLi	Boston, MA	null	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	17	Data Scientist, J	Altamira Techno	Homestead, FL	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	18	Data Scientist	FocusPPI, Inc.	Boston, MA	Be among the first 25 applicants	2 hours ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	19	Machine Learnin	FINESSE	Los Angeles, CA	null	4 hours ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	20	Data Scientist	Company Confic	United States	null	3 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	21	Data Scientist, J	Altamira Techno	Fort Bragg, NC	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	22	Data Scientist In	Indotronics Avani	New Albany, OH	null	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	23	Data Scientist	Ingenio	San Francisco Bay Area	Over 200 applicants	2 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	24	Data Scientist -	Camden Kelly Co	Dallas, TX	null	2 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	25	Data Scientist	TDI Technologie	Greater Philadelphia	Over 200 applicants	3 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	26	Data Scientist	InMoment	United States	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	27	Data Scientist	Alto Neuroscien	Los Altos, CA	Be among the first 25 applicants	4 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	28	Machine Learnin	NovoDynamics	Ann Arbor, MI	Over 200 applicants	5 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	29	Junior Data Scie	Chicago State U	Florida City, FL	Be among the first 25 applicants	5 weeks ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	30	Machine Learnin	Alykmy	United States	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	31	Data Scientist	Janus Health	Chicago, IL	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	32	Junior Data Scie	Lightspeed Syst	Austin, TX	Over 200 applicants	5 days ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19
6417513abcfcf86	33	Machine Learnin	Chemix, Inc.	Sunnyvale, CA	Over 200 applicants	1 week ago	https://www.linkedin.com	https://www.linkedin.com	2023-03-19

1 document selected Count Documents 00:00:00.009

Fig 2.1 JOBS_Basic Collection

Quickstart × JOBS_Basic × JOBS_Advanced ×

MongoDB connection (localhost:27017) > LinkedIn_DB > JOBS_Advanced

Run Load query Save query Query history Set default query Copy Paste Visual Query Builder

Projection {} Sort {} Skip {} Limit {}

Result | Query Code | Explain

Documents 1 to 50 Table View

JOBS_Advanced > ID

_id	ID	LEVEL	TYPE	FUNCTION	INDUSTRY	SIMILAR JOBS	Date_Downloaded
6417513abcfcf86	1	Not Applicable	Full-time	Engineering and Information Technology	Technology, Information and Internet	[37 elements]	2023-03-19
6417513abcfcf86	2	Entry level	Full-time	Engineering and Information Technology	Advertising Services	[37 elements]	2023-03-19
6417513abcfcf86	3	Entry level	Full-time	Engineering and Information Technology	Information Services	[36 elements]	2023-03-19
6417513abcfcf86	4	Entry level	Full-time	Engineering and Information Technology	Hospitals and Health Care	[37 elements]	2023-03-19
6417513abcfcf86	5	Mid-Senior level	Full-time	Engineering and Information Technology	IT Services and IT Consulting	[37 elements]	2023-03-19
6417513abcfcf86	6	Entry level	Full-time	Engineering	Strategic Management Services	[37 elements]	2023-03-19
6417513abcfcf86	7	Mid-Senior level	Full-time	Engineering and Information Technology	Internet Publishing	[37 elements]	2023-03-19
6417513abcfcf86	8	Not Applicable	Full-time	Engineering	Technology, Information and Internet	[36 elements]	2023-03-19
6417513abcfcf86	9	Mid-Senior level	Full-time	Information Technology, Analyst, and Engineering	Staffing and Recruiting	[37 elements]	2023-03-19
6417513abcfcf86	10	Internship	Internship	Engineering and Information Technology	Internet Publishing	[37 elements]	2023-03-19
6417513abcfcf86	11	Not Applicable	Full-time	Engineering	Technology, Information and Internet	[37 elements]	2023-03-19
6417513abcfcf86	12	Mid-Senior level	Full-time	Information Technology	Business Consulting and Services	[37 elements]	2023-03-19
6417513abcfcf86	13	Entry level	Full-time	Engineering and Information Technology	Hospitals and Health Care	[37 elements]	2023-03-19
6417513abcfcf86	14	Mid-Senior level	Full-time	Engineering and Science	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	15	Entry level	Full-time	Engineering and Information Technology	Research	[37 elements]	2023-03-19
6417513abcfcf86	16	Entry level	Full-time	Engineering and Information Technology	Hospitals and Health Care	[37 elements]	2023-03-19
6417513abcfcf86	17	Entry level	Full-time	Engineering and Information Technology	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	18	Entry level	Full-time	Engineering and Information Technology	IT Services and IT Consulting	[37 elements]	2023-03-19
6417513abcfcf86	19	Associate	Full-time	Engineering	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	20	Mid-Senior level	Full-time	Analyst and Information Technology	Musicians	[37 elements]	2023-03-19
6417513abcfcf86	21	Entry level	Full-time	Engineering and Information Technology	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	22	Entry level	Full-time	Engineering and Information Technology	Staffing and Recruiting	[37 elements]	2023-03-19
6417513abcfcf86	23	Associate	Full-time	Analyst and Other	Technology, Information and Internet and Wellness and Fitness	[37 elements]	2023-03-19
6417513abcfcf86	24	Entry level	Full-time	Engineering and Information Technology	Staffing and Recruiting	[37 elements]	2023-03-19
6417513abcfcf86	25	Mid-Senior level	Full-time	Engineering and Information Technology	Defense and Space Manufacturing	[37 elements]	2023-03-19
6417513abcfcf86	26	Entry level	Full-time	Engineering and Information Technology	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	27	Entry level	Full-time	Engineering and Information Technology	Biotechnology	[37 elements]	2023-03-19
6417513abcfcf86	28	Mid-Senior level	Full-time	Information Technology	Software Development	[37 elements]	2023-03-19
6417513abcfcf86	29	Entry level	Full-time	Engineering and Information Technology	Higher Education	[37 elements]	2023-03-19
6417513abcfcf86	30	Mid-Senior level	Full-time	Engineering and Information Technology	Capital Markets, Software Development, and Financial Services	[37 elements]	2023-03-19
6417513abcfcf86	31	Entry level	Full-time	Engineering and Information Technology	Technology, Information and Internet	[37 elements]	2023-03-19
6417513abcfcf86	32	Entry level	Full-time	Engineering and Information Technology	Software Development	[36 elements]	2023-03-19
6417513abcfcf86	33	Not Applicable	Internship	Engineering	Technology, Information and Internet	[37 elements]	2023-03-19

1 document selected Count Documents 00:00:00.032

Fig 2.2 JOBS_Advanced Collection

Quickstart × JOBS_Basic × JOBS_Advanced × JOBS_DESCRIPTION ×

MongoDB connection (localhost:27017) > LinkedIn_DB > JOBS_DESCRIPTION

Run ▾ Load query ▾ Save query ▾ Query history ▾ Set default query ▾ Copy ▾ Paste

Projection ▾ {} Sort ▾ {} Limit ▾ {} Table View

Result | Query Code | Explain

Documents 1 to 50 | Count Documents | 00:00:00.020

JOBS_DESCRIPTION > ID

_id	ID	DESCRIPTIONS	Date_Downloaded	lower_salary	upper_salary			
6417513abc8680aa5e66dc	1	Sunlighten is looking for a Data Scientist to join our growing IT department. They are seeking a dynamic global technology company and its success has been driven by a focus on analytical and modeling tasks. As a junior data scientist at our company, you will collaborate as a team member on analytical and modeling tasks.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	2	Kargo unites the world's leading brands and premium publishers on mobile devices using innovative creative advertising. Bardess is looking for a skilled data scientist.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	3	Bardess is looking for a skilled data scientist.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	4	Biomedical researcher with specialized training in quantitative methodology. Manages, cleans, manipulates, and analyzes data.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	5	Our Client is looking for a dynamic global technology company and its success has been driven by a focus on analytical and modeling tasks.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	6	As a junior data scientist at our company, you will collaborate as a team member on analytical and modeling tasks.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	7	View All Jobs	20221011-1541-01	Salary Range: Up to \$120k	2023-03-19			
6417513abc8680aa5e66dc	8	We're looking for people who want to build for the future of Web3. You'll be building customer-facing products.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	9	Our client is seeking a Data Scientist to join their well-established team due to new contracts. The ideal candidate is experienced in data analysis and has a strong background in machine learning.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	10	Anika Systems is looking for dynamic, well-rounded interns to join our growing team! We are a fast-growing, web-based company.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	11	WalletConnect is the open-source web3 standard to connect blockchain wallets to dapps. Any wallet, any dapp.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	12	We are a mid-sized Management Consulting, Automation, and Data/Process Science firm, established in 1993.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	13	Qualifications:	Master's degree in Computer Science, Mathematics, Statistics, Engineering, or a related field.	□null	□null			
6417513abc8680aa5e66dc	14	At Unlearn, our purpose is to advance artificial intelligence (AI) to eliminate trial and error in medicine. Today, we are seeking a Data Scientist.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	15	This position offers options for in-person (in the District of Columbia (DC) metropolitan area), hybrid or remote work.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	16	Biomedical researcher with specialized training in quantitative methodology. Manages, cleans, manipulates, and analyzes data.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	17	Data Scientist	Altamira delivers a variety of analytic and engineering capabilities to the life sciences industry.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	18	We are searching for a strategic Data Scientist to develop and manage data analysis and reporting.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	19	About The Role:	At FINESSE, we pride ourselves in building the foundations of a consumer data platform.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	20	The Company:	Our client is an independent record label and music publisher with a portfolio of artists and songs.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	21	Data Scientist	Altamira delivers a variety of analytic and engineering capabilities to the life sciences industry.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	22	Position: Data Scientist	Location: New Albany, OH	Duration: 12 Months	2023-03-19	□null	□null	
6417513abc8680aa5e66dc	23	Before we get started:	Here at Ingenuo, we'd love to talk with you regardless of your qualifications.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	24	Data Scientist - Fully Remote Opportunity	We have an opportunity for a fully remote Data Scientist.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	25	TDI Technologies, Inc. is seeking candidates for a Data Analyst position.	TDI Technologies, Inc. is seeking candidates for a Data Analyst position.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	26	Application Deadline:	May 7, 2023	Department:	ML & AI	2023-03-19	□null	□null
6417513abc8680aa5e66dc	27	Company Description:	Alto Neuroscience is a clinical-stage pharmaceutical company.	2023-03-19	□null	□null		
6417513abc8680aa5e66dc	28	Fast-growing AI company is seeking highly-qualified scientists and engineers with experience in ML (Machine Learning).	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	29	LMI is a consultancy dedicated to powering a future-ready, high-performance government, drawing from experts.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	30	At Alkyimi, we're creating products to give everyday business users the power to control the unstructured documents in their organization.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	31	At Janus, we believe in a world where healthcare functions efficiently. Join us on our mission to improve the live.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	32	Lightspeed Systems is looking for a Junior Data Scientist to join our small, multidisciplinary data team.	2023-03-19	□null	□null			
6417513abc8680aa5e66dc	33	Chemix is seeking a highly-motivated machine learning research engineer intern to help develop and expand our product.	2023-03-19	□null	□null			

1 document selected

Fig 2.3 JOBS_Description Collection

Quickstart × JOBS_Advanced × JOBS_Basic × JOBS_DESCRIPTION × JOBS_Recruiter ×

MongoDB connection (localhost:27017) > LinkedIn_DB_Try > JOBS_Recruiter

Run ▾ Load query ▾ Save query ▾ Query history ▾ Set default query ▾ Copy ▾ Paste

Projection ▾ {} Sort ▾ {} Limit ▾ {} Table View

Result | Query Code | Explain

Documents 1 to 50 | Count Documents | 00:00:00.006

JOB Show first page of results

_id	ID	Recruiter Name	Recruiter Title	Recruiter URL	Date_Downloaded
640facd	1	Justin S.	VP Customer Success at DroiSys	https://www.linkedin.com/login?session_redirect=https://www.droisys.com/jobs	2023-03-13
640facd	2	John Sweeney	Senior Technical Recruiter	https://www.linkedin.com/login?session_redirect=https://www.droisys.com/jobs	2023-03-13
640facd	3	■null	■null	■null	2023-03-13
640facd	4	Akhil Vashisth	Talent Acquisition Specialist at MindSource	https://www.linkedin.com/login?session_redirect=https://www.mindsourcelabs.com/jobs	2023-03-13
640facd	5	■null	■null	■null	2023-03-13
640facd	6	Alyssa Bernal	Talent Recruiter	https://www.linkedin.com/login?session_redirect=https://www.mindsourcelabs.com/jobs	2023-03-13
640facd	7	■null	■null	■null	2023-03-13
640facd	8	■null	■null	■null	2023-03-13
640facd	9	■null	■null	■null	2023-03-13
640facd	10	■null	■null	■null	2023-03-13
640facd	11	Stephanie Foster	VP, People Strategy at ClarisHealth	https://www.linkedin.com/login?session_redirect=https://www.clarishealth.com/jobs	2023-03-13
640facd	12	Andrew Condon	Data & Analytics Manager at Franklin Street	https://www.linkedin.com/login?session_redirect=https://franklinstreet.com/jobs	2023-03-13
640facd	13	Maham O'Neill	Lead Recruiter NextGen Application Development Specialist Franklin Street	https://www.linkedin.com/login?session_redirect=https://franklinstreet.com/jobs	2023-03-13
640facd	14	■null	■null	■null	2023-03-13
640facd	15	Jen Hernandez	Director of Talent Acquisition and Development I'm Hiring!	https://www.linkedin.com/login?session_redirect=https://www.imhiring.com/jobs	2023-03-13
640facd	16	■null	■null	■null	2023-03-13
640facd	17	■null	■null	■null	2023-03-13
640facd	18	■null	■null	■null	2023-03-13
640facd	19	■null	■null	■null	2023-03-13
640facd	20	Felicia Gold	AIRS Certified Diversity Recruiter - Putting the "u" and "i" back in diversity	https://www.linkedin.com/login?session_redirect=https://www.airsrecruiting.com/jobs	2023-03-13
640facd	21	Summer Pirkle	Technical Recruiter at Synergis	https://www.linkedin.com/login?session_redirect=https://www.synergis.com/jobs	2023-03-13
640facd	22	■null	■null	■null	2023-03-13
640facd	23	Erica (Watson) Stok	Senior Professional Recruiter at Insight Global	https://www.linkedin.com/login?session_redirect=https://www.insightglobal.com/jobs	2023-03-13
640facd	24	Aaron Tenorio	Executive Recruiter at LHH	https://www.linkedin.com/login?session_redirect=https://www.lhh.com/jobs	2023-03-13
640facd	25	Ifeoluwa Jokotoye	Technical Recruiter at Encore Technologies IT Support	https://www.linkedin.com/login?session_redirect=https://www.encoretech.com/jobs	2023-03-13
640facd	26	■null	■null	■null	2023-03-13
640facd	27	Erica Fuchs	Sr Associate at The Bachrach Group EX D1 Athlete	https://www.linkedin.com/login?session_redirect=https://www.bachrachgroup.com/jobs	2023-03-13
640facd	28	Mario Perez	Expert in Building Cloud Teams in North America	https://www.linkedin.com/login?session_redirect=https://www.bachrachgroup.com/jobs	2023-03-13
640facd	29	■null	■null	■null	2023-03-13
640facd	30	■null	■null	■null	2023-03-13
640facd	31	El hadji diedhiou	Manager at Diamonds International	https://www.linkedin.com/login?session_redirect=https://www.diamondsinternational.com/jobs	2023-03-13
640facd	32	■null	■null	■null	2023-03-13
640facd	33	■null	■null	■null	2023-03-13

1 document selected

Fig 2.4 JOBS_Recruiter Collection