

Vasu Jain

Python-savvy professional exploring AI applications in business contexts. Driven towards a data science career, specializing in finance. Embraces agility in learning to innovate solutions for industry challenges.

Work Experience

S&P Global Market Intelligence Aug 2020 - Present | 3.7 yr.
Gurugram, HR · Web-based platform [Capital IQ] to deliver solutions through global financial intelligence.

Engineer, Application Development 1 yr. 1 mo.

- Developed application to create business efficiencies through use of Python and SQL. Distributed the applications as self-hosted web-apps packaged as executable files.
- Enhanced and developed the internal testing framework providing tools to form connections to databases, cloud frameworks and REST APIs and methods to perform routine tests and analysis. This increased overall efficiencies within the database testing.
- Enhanced Python-based BDD testing developing an internal software suite, automating weekly regression tests for client dataset onboarding. Employed advanced SQL techniques, including procedures, functions, and dynamic queries, for daily large-scale data transformations within a scrum team.
- Conducted rigorous database testing, primarily public company fundamentals and private company data, overseeing transformation pipelines across diverse environments.

Lead Data Analyst 2 yr. 6 mo.

- Played a crucial role in User Acceptance Testing (UAT), vendor onboarding, and creating transformation logics for the ETL pipeline using Databricks (PySpark) for of Bond Reference Data.
- Leveraged Python, SQL, and Advanced Excel to enhance, streamline, and automate security linking workflows, significantly boosting team productivity and accuracy.
- Led lean improvements by applying Python and SQL expertise for process automation and enhancement, contributing significantly to operational efficiency. Collaboratively orchestrated and led a team initiative to swiftly eliminate a 2-year backlog, successfully processing 12,000 pending ingestions within a month.
- Developed new SQL-based Golden Copy reports and enhanced existing ones to maintain data accuracy. Additionally overhauled the process for quality report generation to minimize manual effort by 60-70 minutes per day.
- Managed US and Canada non-portfolio shareholder filings (13D/G, form 4, CEWR) for Capital IQ using Python. Maintained data integrity with a 98% accuracy rate through rigorous internal QA processes. Generated domain-specific QA SQL reports to validate and enhance data quality.
- Improved Gurugram team's core data collection by streamlining processes and automating administrative tasks. Actively involved in site-level events and process enhancement projects.

Certifications

Snowflake Hands-on Essentials 2024
Data Warehousing [↗](#)
Collaboration, Cost Estimation [↗](#)
Certified Lean Facilitator [↗](#) 2022
Lean methodologies
Google Data Analytics [↗](#) 2021
Excel, R, SQL, Tableau

Languages

Python Advanced
- pandas, selenium, sqlalchemy, pyspark, pytest
SQL · MS-SQL Advanced
Regex · PCRE Advanced
Java Intermediate
VBA Intermediate
Gherkin Intermediate
JavaScript Intermediate
HTML/CSS Intermediate
R Functional
Bash/Shell script Functional

Tools

MS Excel Advanced
MS PowerBI Advanced
Tableau Advanced
Databricks Intermediate
Jenkins Intermediate
Adobe Illustrator Intermediate
Git Intermediate
Azure DevOps Server Functional
Docker Functional
SPSS Functional

Projects

Mar 2024	S&P Global – Cloud Comparison Tool Developed an in-house web-app through Python (streamlit, pandas, sqlalchemy, requests) for validating database, DataBricks and Open API connections, executing queries and comparing the query outcomes adding value to database testing processes.
Sep 2023	S&P Global – Bulk Updating Facility (BUF) Validator Developed a web-app through Python (streamlit and pandas) for validating CSV files required for an internal process which allows bulk IUD changes to production databases. This prevented errors and time delays caused due to validation failures.
Dec 2022	S&P Global – BRD N6 Data Loading validation Validating the existing N6 BRD pipeline to ensure that the data is captured, transformed, and represented accurately and completely. This involved replicating the ETL pipeline on Databricks using PySpark and Spark SQL, and subsequently outlining deviations.
Aug 2022	S&P Global – GICS Industry Classification Automation Python project for Industry Classification which undergoes major changes on a bi-annual basis. This project applied machine learning techniques to ascertain the correct company classification based on unstructured business description data.
Jun 2022	S&P Global – Win Automation Python web-scraping Web-scraping project for various US centric public and quasi-government websites requiring outputs in .xls, .xml and .txt formats. This was achieved with pandas, selenium and sqlalchemy. Also deployed the scripts over virtual machine and scheduling runs through windows batch files.
Jan 2021	S&P Global – Timeliness tracker for US Shareholder filings As part of Public Ownership collection process, the SLA deliverables were decided at 18 hours per job. To facilitate administrative supervision over maintenance of SLAs, an excel dashboard was created with dynamic computation of timeliness stats over a continuous basis using advanced Excel, Power Query and DAX.
May 2020	ABS – Project Report on Application of LSTM model for prediction of stock market trends Use of python libraries to apply LSTM model for prediction of long term and short-term trend analysis of major Nifty 50 stocks.
Mar 2019	ABS – Research paper on Impact of Market Volatility on expiry date of Nifty Options Research paper submission for International Conference (AEG) in collaboration with Professor Vipul Mehta (ABS) regarding study of market volatility on expiry date and its impact on computing profitable option strategies.

Activities

Sep 2021	S&P Global – Virtual Volunteering through Essential Impact Topic: Designing math quizzes for primary school students
Mar 2021	S&P Global – Organizing member of Lean Event GGN Topic: Lean Quiz competition with focus on lean wastes and methodologies
Sep 2020	S&P Global – Virtual Volunteering through Essential Impact Topic: Mentoring on digital literacy and email etiquette
Sep 2019	AEG – Coordinator at International Conference Event: VUCA World Issues and Challenges

Education

Post-Graduation	2023-25
MCA (Artificial Intelligence)	Pursuing
Jain University, Bangalore	
Post-Graduation	2018-20
PGDM (Finance, Marketing)	8.1 GPA
Asian Business School, Noida	
Graduation	2014-17
B. Com. (Hons.)	67%
School of Open Learning, Delhi	
Senior Secondary	2014
Class XII (Commerce) CBSE	94.5%
Lovely Public School, Delhi	
Secondary	2012
Class X CBSE	10.0 CGPA
Lovely Public School, Delhi	

Linguistic Proficiency

English	Fluent
Hindi · हिन्दी	Native

Personal Information

Indian · Male · 28 years old
Delhi NCR, India

Contact Details

+91-7503075220 [📞](tel:+917503075220)
vasujain.vj@gmail.com [✉](mailto:vasujain.vj@gmail.com)
linkedin.com/in/vasujain8 [🌐](https://www.linkedin.com/in/vasujain8)
github.com/vasujan [🌐](https://github.com/vasujan)

Work email:
vasu.jain@spglobal.com [✉](mailto:vasu.jain@spglobal.com)