

SAHIL RAAJ KAPUR

Data Scientist | Machine Learning | Predictive Analytics | Python

sahilraajkapur@zohomail.in | +9191717276936 | <https://in.linkedin.com/in/sahil-raaj-kapur-b2428880>

<https://github.com/srkapur>

PROFESSIONAL SUMMARY:

Senior Data Science & Analytics professional with 10+ years of experience designing and delivering predictive models, time-series forecasting solutions, and scalable data workflows across global enterprises. Strong expertise in statistical modeling, machine learning (classification, regression, forecasting), and translating complex datasets into measurable business impact. Hands-on experience in Python, SQL, feature engineering, API-driven data extraction, and model validation. Proven ability to lead analytics teams, drive automation, and align data science initiatives with strategic business objectives.

TECHNICAL SKILLS

Machine Learning & Modeling

Regression, Classification, Random Forest, Gradient Boosting, Time-Series Forecasting, Feature Engineering, Model Evaluation, Cross Validation

Statistical Techniques

Hypothesis Testing, A/B Testing, Forecasting Models, Probability, Scenario Analysis

Programming & Tools

Python (Pandas, NumPy, Scikit-learn), SQL, VBA

Data Engineering & Automation

ETL Workflows, API Integration, Spark, Data Extraction Automation, OCR-based Processing

Business Intelligence & Visualization

Power BI, DAX, Advanced Excel, Data Visualization

PROFESSIONAL EXPERIENCE:

EXL | MANAGER- Data Science [December,2025 - February,2026]

- Led the design and deployment of Python-based automation and data processing solutions, integrating OCR (Tesseract) and structured pipelines to eliminate manual data handling and improve operational efficiency
- Developed and maintained end-to-end Power BI reporting ecosystems integrating multiple structured data sources, translating complex datasets into executive-ready insights.
- Acted as a technical and people manager, bridging business and technology, setting development standards, creating deployment documentation, ensuring scalable, well-governed solutions across projects.

EY GDS | ASSISTANT MANAGER [April,2021 - November,2025]

- Developed API-driven data extraction and transformation pipelines using Python, reducing ESG data retrieval time by 50%
- Designed regression and scenario-based models linking ESG metrics to financial performance indicators, improving benchmarking accuracy and decision support
- Built predictive and scenario-based analytical models to support ESG benchmarking and performance evaluation, enabling data-driven strategic decision-making
- Built automated ETL workflows integrating SQL, Excel, and APIs, reducing processing time by over 60%

PROFESSIONAL EXPERIENCE:

- Optimized high-volume data workflows to improve runtime efficiency and scalability
- Managed and mentored junior analysts while overseeing technical delivery and stakeholder communication.
- Selected for international secondment to Hamburg, Germany (Jan 2023 – Mar 2023) to support on-site analytics delivery and stakeholder collaboration for European engagements.

The Economist | Data Analyst [July,2017 - March,2021]

- Designed and validated macroeconomic time-series forecasting models across 200+ countries using statistical methodologies
- Automated large-scale economic databases using SQL and VBA, improving reporting accuracy and reducing manual intervention
- Conducted quantitative analysis for projects including GDP forecasting and military expenditure modeling
- Supported cross-functional research teams by translating statistical findings into actionable insights

Everest Group | Information Specialist [August,2016 - July,2017]

- Automated quarterly research reporting workflows using Excel and VBA, improving operational efficiency.
- Enhanced and maintained transaction intelligence databases supporting market research initiatives.

PROJECTS

- **Macroeconomic Forecasting Model:** Built time-series forecasting models using statistical techniques to project economic indicators and evaluate predictive performance using RMSE and error comparison frameworks.
- **API-Based Data Extraction Pipeline:** Developed modular Python pipelines to extract, clean, and process structured and unstructured ESG data from financial APIs.
- **Automation & OCR Processing Framework:** Designed Python-based automation tools integrating OCR and structured transformation logic to streamline document data ingestion.

EDUCATION

- BS in Data Science and Applications | IIT Madras (2021–2026)
- MA in Economics | Jamia Millia Islamia (2014–2016)
- BA (Hons) Economics | University of Delhi (2010–2013)

LEADERSHIP & RECOGNITION

- ESG 360° Suite Winner – SAT Olympiad (2023)
- Impact Modelling Winner – SAT Olympiad (2022)
- Led cross-functional analytics teams delivering enterprise-scale data initiatives

SHORT TERM EXPERIENCE

- NHB
- RIS
- GARTNER
- BAHL & Company