Sourav Kumar

souravmail003@gmail.com — +91-8405846421 — Bangalore, India — LinkedIn — GitHub

SUMMARY

- **Data Analyst** with **1+ years of experience** delivering actionable **insights**, process **automation**, and business **impact**.
- Skilled in **Python**, **SQL**, **Power BI**, **ML**, and **NLP** with focus on **scalable model development**.
- Delivered **data-driven solutions** that reduced **manual effort by 25%** and improved **efficiency by 18%**.

EXPERIENCE

- **Rare Metals and Chemicals** — *Junior Data Analyst* — Ranchi, Jharkhand — **June 2024 – June 2025**
  - **Data Analysis & Reporting**: Automated daily and monthly production reports using **Python** and **Excel macros**, streamlining workflows and reducing manual effort by **25%** through **script-based data processing**.
  - **Dashboarding**: Designed interactive dashboards in **Excel** and **Power BI** to monitor **KPIs** like **yield**, **material consumption**, and **batch-wise flow**, enabling real-time insights.
  - **Cross-functional Collaboration**: Collaborated with **process** and **production teams** to identify **inefficiencies** and implement **data-backed improvements** to enhance **output and quality**.
- **LearnVista Pvt Ltd** — *Data Science Intern* — Remote — **Jan 2024 – June 2024**
  - **Data Projects**: Built and deployed **ML models** for **fraud detection**, **healthcare prediction**, and **recommendation systems** using **Python**, **Scikit-learn**, and **Pandas** on real-world datasets for insights and reporting.
  - **NLP & Gen AI**: Applied advanced **NLP techniques** and **generative AI** tools like **ChatGPT** and **Gemini** for **sentiment analysis**, **keyword extraction**, and prompt-based content generation.
  - **Visualization**: Designed **Power BI dashboards** and visualized trends using **Matplotlib**, **Seaborn**, and **Excel** to support **data-driven decision-making**.

EDUCATION

- **I.K. Gujral Punjab Technical University (Main Campus)** — Kapurthala, India
  B.Tech in Computer Science and Engineering — 75% — *Jul 2020 – Jun 2024*

PROJECTS

- **Patient Encounter Cost And Risk Analysis In Healthcare** (**SQL**, **Advanced Excel**, **Power BI**, **DAX**) [GitHub]
  - Conducted deep-dive analysis on encounter cost, payer categories, and procedure risk scores using **SQL queries** across multiple healthcare datasets to identify key spending trends.
  - Built dynamic and interactive **Power BI dashboards** using **DAX**, slicers, and KPI visuals to reveal uncovered costs, supporting better resource allocation and payer strategy.
- **Customer Behaviour and Revenue Analysis in E-commerce** (**MySQL**, **Power BI**, **Pandas**, **DAX**) [GitHub]
  - Engineered a comprehensive data solution by designing and normalizing a **MySQL database schema**, loading raw e-commerce data, and writing complex **analytical SQL queries** to extract key business insights.
  - Developed interactive **Power BI dashboards** to visualize KPIs like revenue trends, AOV, top-selling categories, and cancellation rates, enabling data-driven decisions on seller performance, customer reviews, and seasonal demand.
- **Quick-Commerce Operational Data Analysis (Blinkit)** (**SQL**, **Power BI**, **Joins**, **Views**, **CTEs**) [GitHub]
  - Designed and normalized a **MySQL database schema** for Blinkit's quick-commerce sales data; performed advanced **SQL queries** for **data validation**, identifying anomalies in pricing, sales, and inventory, and enabling accurate reporting.
  - Leveraged SQL to analyze **SKU-level sales trends**, monitor **on-shelf availability**, and assess **city-wise inventory performance** provided actionable insights to enhance **supply chain planning** and support **business decisions**.

TECHNICAL SKILLS

- **Programming & Databases:** Python, SQL, MySQL, MS SQL Server, MongoDB
- **Analytics, ML & Statistics:** EDA, Statistical Analysis, A/B Testing, Hypothesis Testing, Data Modeling, Regression, Classification, Time Series Forecasting, Supervised/Unsupervised Learning, Model Evaluation, Deep Learning, GenAI
- **Libraries & Tools:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly, BeautifulSoup, Streamlit, Git
- **BI Tools & Reporting:** Power BI, Tableau, Advanced Excel, Dashboard Design, KPI Reporting, Data Visualization
- **Cloud & Soft Skills:** AWS, Azure, Data Storytelling, Business Acumen, Stakeholder Communication, Problem Solving

CERTIFICATIONS

- **IBM Advanced Data Science & AI Specialization** — Completed on **Apr 15, 2025**
  Covered **Python**, **ML**, **DL**, **GenAI**, **Data Science Methodology**, and **Data Science Tools**. [Certificate]
- **Tableau & Power BI Certification (Learnbay)** — Completed on **Nov 21, 2024**
  Hands-on with **dashboarding**, **DAX**, **Power Query**, and **data storytelling**. [Certificate]
- **Microsoft Azure Fundamentals** — Completed on **May 24, 2025**
  Covered **cloud concepts**, **Azure services**, **deployment models**, and **cloud security**. [Certificate]

ACHIEVEMENTS

- **Secretary, Entrepreneur Club:** Organized 3+ innovation events and workshops with 100+ student participation.
- **IBM Certified Data Science Professional:** Completed capstone and model deployment module.