MASUD KHAN
Data Analyst

📞 +91-8101365489 ✉ masudist32@gmail.com in linkedin.com/in/masudkhan ⌗ github.com/MasudKhan45

Summary

Results-oriented Data Analyst with a background in Mechanical Engineering and strong proficiency in Python, SQL, and Power BI. Proven ability to translate complex datasets into actionable business insights, optimize operational efficiency, and support stakeholder decision-making. Experienced in statistical analysis and predictive modeling to identify trends and mitigate risks. Seeking to drive data-led growth in a dynamic product or analytics team.

Skills

Languages & Databases: Python, SQL (PostgreSQL, MySQL), PySpark
Libraries & Frameworks: Pandas, NumPy, Scikit-learn, PyTorch, Matplotlib, Seaborn
Visualization & Tools: Power BI, Tableau, MS Excel (Advanced), Streamlit, Git/GitHub
Analytics & ML: Statistical Analysis, Hypothesis Testing, NLP, Predictive Modeling, EDA

Experience

Statistical Analyst / Subject Matter Expert   **Jan 2023 – Jun 2025**
Chegg India Pvt. Ltd. (Remote)   *Tools: Excel, Statistical Methods*
- Delivered high-volume statistical analysis on Regression and Time Series forecasting, maintaining a 95+ accuracy rating while strictly adhering to Service Level Agreements (SLAs).
- Simplified complex quantitative data into actionable documentation for non-technical stakeholders, improving clarity and reducing revision rates.

Data Lead & Technical Mentor (AI/ML)   **Jan 2023 – Jun 2024**
GFG Student Chapter, GGV   *Tools: Excel, Power BI, AI/ML*
- Led a team of 50+ members in executing end-to-end ML projects, overseeing the lifecycle from data collection to dashboard visualization.
- Organized technical workshops on Python and Deep Learning, fostering a data-driven culture and bridging the gap between theory and practical application.

Projects

Restaurant Operations & Profitability Dashboard | *SQL, Power BI*   **Feb 2025 – Mar 2025**
- Designed a holistic operational dashboard in Power BI, querying and transforming raw sales and inventory data using SQL to track business health.
- Calculated high-impact KPIs including per-item profitability and peak service hours, providing stakeholders with clear visibility into revenue trends.
- Delivered actionable insights that led to optimized menu pricing and reduced food waste, directly supporting management in maximizing profit margins.

Fake News Detector (RoBERTa + RAG) | *Python, NLP, PyTorch, FAISS, Streamlit*   **Sep 2025**
- Fine-tuned a RoBERTa transformer model using PyTorch to classify news articles, optimizing the architecture to effectively detect misinformation.
- Engineered a Retrieval-Augmented Generation (RAG) system using FAISS and Sentence Transformers to fetch semantic context and ground predictions in evidence.
- Deployed an interactive Streamlit dashboard for real-time inference, displaying classification probabilities alongside retrieved source snippets for full explainability.

Workforce Attrition Risk Analysis | *Python, Pandas, Hypothesis Testing*   **Oct 2025**
- Analyzed a dataset of 1,470 employee records to identify key drivers of churn, revealing a significant correlation between stock option levels and exit rates.
- Engineered features to quantify the impact of commute distance on job satisfaction, providing data-backed recommendations to improve retention strategies.
- Performed statistical testing to validate risk factors, enabling HR stakeholders to proactively address high-risk employee segments.

Awards & Certifications

- **IBM Data Analyst Professional Certificate** (Coursera, Link)
- **Data Analysis with Tableau** (Coursera, Link)
- **Business Analysis Fundamentals** (Coursera, Link)
- **Deloitte Australia – Data Analytics Job Simulation** (Forage, Link)

Education

B.Tech. Mechanical Engineering   **Graduated: 2026**
Guru Ghasidas Vishwavidyalaya, Bilaspur   *CGPA: 7.0 / 10*