# Shaheer Mahmood

Amherst, MA | (413) 510-7492 | shaheermahmood.5@gmail.com | linkedin.com/in/shaheermahmood | github.com/ShaheerM10

## EXPERIENCE

**International Rescue Committee**  —  New York City, NY, USA
*Data Analyst Intern*  —  *Jan 2025 - May 2025*

- Implemented advanced SQL queries to analyze enrollment patterns, class attendance, and learner progression, uncovering trends that informed program planning and led to a **16%** increase in average class attendance
- Built scalable Tableau dashboards for real-time tracking of weekly crop yields and volunteer participation across three sites, enabling data-driven resource allocation and reducing manual reporting time by **40%**

**Alibaba Group (Daraz)**  —  Karachi, SD, Pakistan
*Regional Product Data Scientist*  —  *Aug 2023 - Aug 2024*

- Communicated with engineering, product managers and cross-functional stakeholders to establish the logic for **24** Seller Platform KPIs, subsequently developed a Tableau dashboard with SQL to monitor the key product metrics
- Executed **7** A/B Testing strategies on the Seller E-Commerce App using statistical modeling and hypothesis testing, resulting in a **25%** reduction in bounce rate and a **6%** improvement in on-time order fulfillment rate
- Spearheaded monthly product presentations for the Seller E-Commerce App, presenting complex analyses as clear, actionable insights to senior leadership and technical audiences to guide engagement and retention strategies
- Built and deployed a LLM model to automatically generate product descriptions from titles and attributes, improving listing quality and seller efficiency while reducing manual listing time by **35%**

**Alibaba Group (Daraz)**  —  Karachi, SD, Pakistan
*Supply Chain Analyst*  —  *Aug 2022 - July 2023*

- Optimized volume allocation across logistics partners by tracking success rates and capacity using Snowflake, rebalanced underperforming regions to boost efficiency and scale delivery by **500,000** packages per month
- Analyzed delivery location data using K-means clustering to identify optimal last-mile delivery hub location, reducing average route distance by **8%** and projecting a **5%** decrease in delivery cost
- Built and optimized ETL pipelines in MaxCompute, validating data workflows and ensuring data quality and seamless integration within large-scale operations management systems
- Applied random forest on large-scale operational data to evaluate the impact of process changes, leveraging distributed data workflows to improve delivery success rates by **2%**

## PROJECTS

**Customer Churn Analysis** | *Python, SQL, Scikit-Learn*

- Developed advanced analytical clustering models to predict churn rates across different customer segmentation on the E-Commerce market, enhancing the effectiveness of sales and marketing promotions strategies

**ETL Pipeline: Reddit Sentiment Analysis** | *Python, Streamlit, Airflow, Apache Spark, Docker*

- Designed a scalable ETL pipeline using Airflow to automate Reddit sentiment analysis, integrating Apache Spark for distributed processing, deploying with Docker, and building an interactive Streamlit dashboard for visualizing results

**Spotify Music Recommendation System** | *Python, SQL, Spotify API, Scikit-Learn*

- Developed a recommendation engine using the Spotify API to analyze listening patterns and audio features, applying clustering and similarity algorithms to generate personalized playlists and highlight trends in user behavior

## EDUCATION

**University of Massachusetts Amherst**  —  Amherst, MA, USA
*Master of Science (M.Sc) in Data Analytics*  —  *Sep 2024 – May 2026*

**Ghulam Ishaq Khan Institute of Engineering Sciences and Technology**  —  Topi, KP, Pakistan
*Bachelor of Science (B.Sc) in Mechanical Engineering*  —  *Aug 2018 – May 2022*

## TECHNICAL SKILLS

**Programming Languages:** Python, SQL, SPL, HTML, C++, R, SAS
**Frameworks:** NumPy, pandas, Scikit-Learn, PyTorch, TensorFlow, SciPy
**Tools:** Splunk, PowerBI, Tableau, Git, Jupyter, Matlab, Jira, Cursor