

# Bushra Jabeen, Data Scientist/Analyst

(289) 625-6886 | b.jabeen023@gmail.com | Mississauga, ON | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

Data Scientist with experience using Python and SQL to clean data, build forecasting models, and create clear reports and dashboards for real projects in government, education, and healthcare. Developed and tested models for prediction and classification, including time series forecasting and image and text data, and checked results carefully to make sure they were reliable and easy to explain. Built end-to-end workflows with data collection, feature preparation, model training, and simple deployment, using tools like Git, Docker, and cloud platforms to keep work organized and repeatable. Worked closely with different teams to understand their questions, turn them into measurable problems, and deliver tables, charts, and written summaries that supported everyday decisions.

## Summary of Qualifications

- Used Python, SQL, Pandas, and Power BI to collect, clean, and visualize data on projects like BeamGroup's open data analytics dashboard
- Developed regression, clustering, and neural network models for forecasting, reaching up to 98% accuracy
- Delivered business results like 50% enrollment growth and 80% better financial recovery through data analysis
- Led cross-functional teams on research and training projects, supporting curriculum design and student performance
- Worked with Agile, Git/GitHub, MLflow, Docker, and dashboards to meet tight deadlines

## Technical Skills

<b>Programming &amp; Data:</b>	<i>Python, SQL, Pandas, NumPy, JSON, Shell/Bash</i>
<b>Machine Learning:</b>	<i>Regression, Classification, Clustering, Time Series, NLP, Scikit-learn, TensorFlow, PyTorch</i>
<b>Deep Learning &amp; Gen AI:</b>	<i>Image Classification, Object Detection, Neural Networks</i>
<b>Data Visualization:</b>	<i>Power BI, Tableau, Matplotlib, Seaborn, Dash</i>
<b>MLOps &amp; DevOps Tools:</b>	<i>ETL Pipelines, Git/GitHub, MLflow, KNIME, Kubernetes, Ansible, Docker, Agile/JIRA</i>
<b>Statistics:</b>	<i>Probability, Descriptive &amp; Inferential Statistics, Time Series Analysis</i>

## Education & Certifications

<b>Data Talent Program – Certificate</b> M2M Tech Connect   Vancouver, BC	<b>Jul 2025 – Oct 2025</b>
A 16-week industry-aligned program designed to build job-ready data professionals. Key highlights include:	
<ul style="list-style-type: none"><li>Built end-to-end data workflows with preprocessing, feature engineering, model training, and deployment using Python, TensorFlow, and Azure ML Studio</li><li>Built and tested ML models with regression and clustering to improve prediction accuracy</li><li>Used MLOps concepts and Git for deployment and documentation in team projects</li></ul>	
<b>IBM DevOps and Software Engineer</b> IBM, Coursera   Mississauga, ON	<b>May 2025</b>
<b>Junior IT Analyst Program</b> NPower Canada   Toronto, ON	
<b>IBM Data Science Professional Certificate</b> IBM, Coursera   Mississauga, ON	
<b>Master's in Public Health</b> Govt. College University   Faisalabad, Punjab	
<b>Sep 2018 - May 2021</b>	

## Work Experience

<b>Data Scientist – Employer-Led Project</b> BeamGroup   M2M Data Talent Program   Toronto, ON	<b>Jul 2025 – Nov 2025</b>
<ul style="list-style-type: none"><li>Worked with BeamGroup to build a province-wide analytics dashboard for Ontario's Service System Managers, pulling together public datasets on social and labor indicators</li><li>Built forecasting models and dashboards to support capacity planning and resource allocation decisions, similar to supply chain and logistics analytics use cases</li><li>Collected, cleaned, and merged open-government data using Python (Pandas, NumPy), SQL, and Excel to make socioeconomic metrics consistent across municipalities</li><li>Created an interactive HTML map of Ontario schools, adding detailed info and website links right into the analytics dashboard</li></ul>	

- Built and tested predictive ML models for unemployment forecasting, optimized regression models to improve forecasting performance
- Uploaded all code, validation scripts, and docs to a shared GitHub repo for easy reuse and teamwork
- Wrote user stories, tested dashboard features, and suggested UI fixes based on feedback during Agile sprints

### **Data Analyst/Scientist Fellow Internship**

Fellowship.AI | United Kingdom

**Dec 2023 – Mar 2024**

- Cleaned and analyzed large datasets, fixed missing or bad data, and created visualizations to find patterns and support model building
- Tested multiple ML algorithms, hitting 95% accuracy with K-Means clustering and improving to 98% using stable diffusion and ControlNet models
- Wrote a [blog post](#) explaining the project steps, methods, and results for a wider audience

### **Data Scientist & Data Transformation Projects**

Freelance/Open-Source Projects | Mississauga, ON

**Dec 2023 – Present**

- Built ETL pipelines with SQL and Python to pull, clean, and reshape data from CSV, APIs, and MySQL for modeling in projects like housing price prediction and stock forecasting
- Added validation, outlier detection (Excel, REGEX, Pandas, NumPy), and schema mapping to ensure clean data; tracked experiments with MLflow and deployed via Docker/Flask
- Trained models with Ridge/Logistic Regression, Random Forest, and auto\_arima, achieving 86% accuracy on Fashion MNIST image classification and 83% on SpaceX Falcon 9 landing prediction
- Created NLP pipeline with NLTK for IMDB sentiment analysis (83.75% Bernoulli Naive Bayes, 86% Logistic Regression) and end-to-end SQLite/SQL/Seaborn pipeline containerized in Docker
- Built visualizations with Matplotlib, Seaborn, Tableau, and Power BI for time series like Australian consumption trends; all projects on GitHub
- Packaged models and dashboards for deployment on cloud-based environments using Docker, GitHub, and CI/CD-style workflows

### **Team Lead & Program Coordinator**

Federal Institute of Health Sciences | Lahore, Punjab

**Mar 2014 – Jul 2021**

- Led a multidisciplinary team to execute projects successfully, boosting team morale and productivity through effective communication and guidance
- Designed and delivered curriculum and training sessions, achieving a 90% success rate in key subject areas through effective teaching methodologies
- Analyzed admission data and implemented a targeted campaign strategy, increasing enrollment by 50%, contributing to the department's growth
- Spearheaded financial recovery initiatives, improving department recoveries by 80% through data-driven decision-making
- Supervised eight research projects, ensuring compliance with deadlines and quality standards, successfully achieving key outcomes
- Directed departmental inspections, achieving a 90% compliance with regulatory and operational benchmarks, ensuring operational excellence
- Streamlined the hiring process for staff, ensuring the selection of skilled professionals aligned with organizational goals, contributing to departmental success

### **Additional Volunteer Activities**

- Open-source contributor, GitHub (small PRs to data/ML repos; code linked in profile)
- Deliver monthly curriculum sessions to over 10 youth participants, preparing materials and leading interactive discussions to build skills and engagement
- Organize extracurricular activities and competitions for the youth group, coordinating logistics, schedules, and events attended by 10+ participants
- Plan and facilitate monthly meetups for over 10 women, guiding group discussions on shared educational content to foster collaboration and knowledge sharing