

# Mehreen Habib

## Computer Science Major

Phone: (405) 394-4937 | Email: [mehreenhabibb1@gmail.com](mailto:mehreenhabibb1@gmail.com) | [Portfolio](#)

### ACADEMIC QUALIFICATION

#### University of Oklahoma

M.S. Computer Science (CGPA: 4.0/4.0)

Norman, Oklahoma

May 2024

#### University of Engineering and Technology

B.S Computer Science (CGPA: 3.25/4.00)

Lahore, Pakistan

August 2020

### CAREER SUMMARY

- Over 4 years of professional experience, specializing in Business/Data Analysis, Data Modeling, and ETL Development.
- Proficient in data visualization tools such as Power BI, Tableau, Python Matplotlib, and AWS QuickSight creating powerful and actionable interactive reports and dashboards.
- Expertise in advanced MS Excel, PowerPoint, and spreadsheets (Pivot Tables & Pivot Charts, Macros, VLOOKUP).
- Demonstrated extensive understanding and experience with the entire Software Development Life Cycle (SDLC) including Agile and Waterfall methodologies, and project management skills.
- Collaborated with diverse business and technical stakeholders to develop and implement enhanced designs and process changes that align with the organization's strategic goals and priorities.
- Played a crucial role in designing and structuring User Acceptance Testing (UAT) and production environments, focusing on formalizing, streamlining, and enhancing newly implemented processes and applications.
- Specialized in crafting business process models and workflows, proficient in interpreting Business Requirements Documents (BRD) and functional specifications for stakeholders in leading financial institutions.
- Hands-on experience with SQL database systems, adept in designing, querying, and managing data across multiple platforms including MS SQL Server, Oracle and NoSQL databases.
- Skilled in data analytics and preprocessing, proficient in using statistical tools and data mining algorithms to interpret, cleanse, transform, and normalize raw data, extracting actionable insights from complex datasets for strategic decision-making.

### TECHNICAL SKILLS

- Programming Languages:** Python, Java, R, C
- Data Visualization Tools:** Tableau, Power BI, Excel
- Databases:** MySQL, Microsoft Azure, MongoDB, Oracle, SQLite, Amazon RDS
- Machine Learning:** NumPy, Pandas, Seaborn, Matplotlib, Scikit-Learn, k-NN, SVM, Random Forest, Decision Trees, K-Means Clustering, Linear Regression, Naïve-Bayes
- Scripting Languages:** Shell Scripting (Unix/Linux), PowerShell
- Cloud Platforms/Services:** Google Cloud Platform (GCP), Microsoft Azure, Amazon Web Services (AWS), AWS S3, IAM, EC2, Amazon QuickSight, Athena, AWS Glue, Amazon Redshift
- Other skills and Tools:** Eclipse, Visual Studio, Microsoft Office, Jupyter notebook, PyCharm, Anaconda, MATLAB, Git, SQL Profiler, Audit Reporting, JIRA, Statistical Analysis ANOVA, SPSS, Access, SAS, Snowflake, PySpark, Databricks, Data Structures

### PROFESSIONAL EXPERIENCE

#### Data Institute for Societal Challenges

##### Graduate Research Assistant (GRA)

Norman Oklahoma

August 2022 – Present

- Collaborated with data engineers and the operations team to implement ETL processes, writing and optimizing SQL queries for data extraction to meet analytical requirements.
- Actively engaged in AI-driven initiatives focusing on agriculture optimization and COVID-19 forecasting.
- Responsible for extensive data collection and preprocessing, ensuring robust analysis foundations.
- Involved in diverse research activities including manuscript writing and data presentation.
- Enhanced analytical skills and developed strong communication through collaborative, multidisciplinary efforts. Proficient in AWS services like EC2, S3, Lambda, Redshift, and QuickSight, contributing to a 20% improvement in data processing efficiency, cost reduction, and enhanced data-driven decision-making.
- Proficient in BI tools, data warehousing, reporting, querying, data mining, and spreadsheets. Created detailed reports using SQL queries for data analysis.
- Engaged in feature engineering tasks like generating feature intersections, normalizing features, and label encoding using scikit-learn preprocessing tools.
- Responsible for developing ETL pipelines, including the successful design, development, and integration of components within the Snowflake ETL platform.
- Utilized Python, including libraries such as NumPy, pandas, scikit-learn, and seaborn, to develop diverse models and algorithms for analytical purposes.
- Designed comprehensive data visualizations using Tableau and Matplotlib to present data in an accessible, human-readable format.

## **NETSOL Technologies, Pakistan**

June 2021 – August 2022

### **Data Analyst/Quality Assurance Engineer**

- Experienced in working at Pakistan's only CMMI level 5 organization, with robust skills in managing data repositories and client implementations.
- Analyzed data visualization needs from business users and created customized visual representations to provide actionable insights for decision-making.
- Formulated SQL queries to extract pertinent data from Sales data marts, ensuring accuracy and relevance.
- Demonstrated expertise in creating dynamic visualizations with Python libraries like Pandas, NumPy, Matplotlib, and Scikit-learn.
- Created intuitive Power BI visualizations, including Scatter Plots, Geographic Maps, Pie Charts, Bar Charts, and Density Charts.
- Acquired extensive experience in data analysis, warehousing, migration, cleaning, and statistical modeling.
- Created scripts for data reconciliation using SQL and Python to ensure data integrity.
- Specialized in data cleaning and feature scaling with Pandas and NumPy in Python.
- Leveraged Business Intelligence tools, encompassing data warehousing, reporting, querying, data mining, and advanced spreadsheets.
- Expert in automating processes, troubleshooting development issues, and providing long-term solutions to enhance manual operations.
- Proficient in the waterfall model, creating detailed test documentation, and handling client queries related to automation.
- Transferred essential project knowledge and testing techniques to new vendors and continually upgraded skills through in-house training.
- Skillfully enhanced database performance, leveraging a deep understanding of joins and data distribution strategies.

## **Company: Ozoned Digital, Pakistan**

Oct 2020 – June 2021

### **Title: Software Quality Assurance Engineer**

- Performed System testing, Functional, Regression, Smoke testing, Integration, GDPR, Black Box, Cross Browser, UI/UX and Sanity testing.
- Bug logging and follow up with the issues to be resolved.
- Analysis of the Stories of a particular release and deciding on if it qualifies to automate for the regression suits.
- Utilized Snowflake to automate ETL workflows and streamline the ingestion and transformation of billions of rows of data, significantly minimizing manual effort and improving operational efficiency by 70%.
- Used Tableau to create interactive and compelling visual representations of data which tracks 20+ KPIs.
- Conducted in-depth root cause analysis for production issues and successfully addressed over 300 defects. Utilized automated SQL queries to analyze logs, resulting in a notable enhancement of the team's issue resolution efficiency by approximately 20%.

## **ACADEMIC PROJECTS**

### **Smart Agricultural Decision Maker using Data Mining and Python/HTML, CSS**

- By checking statistical significance of models developed from scratch and 95% confidence interval suggested the best crop to grow based on various environmental factors.
- Developed a web application to enhance farmer decision-making through data analysis.

### **JSP web application using Java and SQL**

- Developed a dynamic JSP web application integrating Azure SQL databases, Eclipse.
- Java application for managing employees, product, customer, account, complaint, and accident records.
- User-friendly interfaces for inserting data, retrieving information, and handling transactions.
- Organized database support, efficient operations, and decision-making.

### **Smart City Data Analysis**

- This project aims to download and clean pdf documents, create, and explore clustering models, perform topic modeling to derive meanings.
- Analyzed and cleaned Smart City Challenge dataset, extracting insights on themes and concepts.
- Leveraged data collection, preprocessing and visualizations for analysis.

### **Road Accidents Analysis using Power BI and Excel**

- Developed a Road Accident Analysis Dashboard in Excel.
- Involving data cleaning, pivot tables, and formula driven KPIs to visualize casualty trends, driving actionable insights for client safety strategies,
- Enhancing strategic insights for client safety measures.

## **CERTIFICATION**

- JPMORGAN CHASE AND CO: Quantitative Research Virtual Experience Program

August 2023