

# M. SIDDIQ KHAN

[LinkedIn](#) | [m.siddiq2000@gmail.com](mailto:m.siddiq2000@gmail.com) | +18127785538 | Irvine, CA, United States

## Education

### Indiana University, Bloomington, IN, United States

May 2024

· Master of Science in Data Science

GPA 3.71/4.00

· Relevant Coursework: Data Mining, Statistics, Applied Machine Learning, Advanced Database administration

### University of Mumbai, India

May 2022

· Bachelor of Engineering in Mechanical Engineering

GPA 8.7/10.0

## Certifications and Licenses

### Power BI Data Analyst Associate PL-300

Microsoft Certified

### Tableau Desktop Certified Associate

Tableau

### Data Science & Machine Learning Career Track

Coding Ninjas India

### Neural Networks and Deep Learning

DeepLearning.AI

### TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning

DeepLearning.AI

## Skills

### Programming Languages

Python, R, SQL

### Technologies

Tableau, Power BI, Azure, AWS, MS-Office Suite

### Familiar OS

Windows, iOS

### Courses

Exploratory Data Analysis, Computer Vision

Database Management, Data Visualization, Tableau

## Professional Experience

### Data Engineer - Sirah Project

May 2023 - August 2023

- Collaborated on a digital humanities research project led by professor at IUB Professor, that compiled and reconstructed 1.2 million words of early Islamic literature, significantly advancing digital humanities research, and analyzed over 147 historical text copies, showcasing robust data management skills.
- Collaborated with global experts in digital humanities and IT, and networked with leading institutions like Aga Khan University and the London School of Economics.
- Engineered an algorithm to boost the precision of text reconstruction processes and designed a user-friendly interface with right-to-left text support for Arabic, meeting a critical web development need.
- Created a prototype foundational for future grant proposals, targeting funds from prominent organizations like the NEH, underscoring the project's scalability and potential impact.

### Data Science Intern(TA) - Coding Ninjas India

May 2021 - August 2021

- Delivered comprehensive tutorials on SQL, Tableau Desktop, Python, Numpy, Pandas, and Seaborn, enhancing students' practical skills and theoretical knowledge.
- Conducted weekly office hours, offering personalized support and guidance to over 25 students, improving comprehension and project completion rates.
- Actively assisted in debugging projects, contributing to a significant reduction in project submission delays.
- Graded assignments and projects, maintaining rigorous academic standards and providing constructive feedback to ensure continuous improvement.

## Projects

### Football Transfer Market Analysis

- Employed advanced statistical methods, machine learning algorithms, and predictive modeling techniques to analyze the football transfer market, providing valuable insights for strategic decision-making.
- Managed the end-to-end project life cycle, including data wrangling, analysis, and dynamic visualization using Plotly and Matplotlib, enhancing accessibility and user experience.

### Water Quality Portal

- Aggregated and cleaned a comprehensive dataset of 173,140 records, merging Water Quality Portal data with IRS income data, and designed a user-friendly interface for detailed analysis of 50 water quality characteristics.
- Conducted statistical analysis and developed an interactive PowerBI dashboard featuring visualizations to analyze key metrics, across counties from 2018 to 2023.

### Distracted Driver Detection

- Developed a classifier to evaluate driver behavior using Computer Vision and CNNs, achieving 75% accuracy.