# Mohammad Saqib Ansari

# ansarimdsaqib@gmail.com | Mumbai, Maharashtra

### Portfolio | HackerRank | GitHub | Linkedin | Blog

#### EDUCATION

Indian Institute of Technology, Kanpur

Kanpur-UP, India July 2022 - July 2024 Statistics Masters

CGPA: 8.1

**Ewing Christian College** Allahabad-India Statistics and Mathematics Aug-2018 - July-2021

Percentage: 72.22%

Dawn The Higher (CBSE) Allahabad-India PCM HSC Jul 2016 - Jul 2017

Percentage: 77%

EXPERIENCE

#### ICICI GIC | Associate Data Scientist

Mumbai, Maharashtra | Jul-2024 - Present

Conducted EDA on PhonePe Insurance Data, deriving actionable insights using Python and Tableau. Utilized LLMs and Retrieval-Augmented Generation (RAG) to extract information from PDFs. Built a Cross-Sell Data Dashboard in Power BI for customer analytics and visualization. Leveraging Google Analytics 2 wheeler data for lead scoring and building a chatbot with RAG and LangChain.

#### TutorBro | Subject Matter Expert

Prayagraj , Remote | Jun 2024 - Jul 2024

• Assisted University Students in Solving Statistics Problems using Excel, R, and Statistical Software

#### Meritshot | Data Science intern

Kanpur, UP | Sep-2023 - Nov-2023

Established a Medium blog focusing on 'Data Science in IoT,' featuring a business case study on Lenskart, a comprehensive Data Preprocessing Cheatsheet, and an exploration of Statistical Functions in R Programming.

# Chegg | Subject Matter Expert

Kanpur, UP | Oct 2022 - Jun 2023

Showcased specialized expertise as a Subject Matter Expert of Statistics over a period of 9 months, illustrating proficiency and mastery in the subject matter.

#### SKILLS

Programming Languages: Python, R, MATLAB, LaTeX

Libraries/Frameworks: Pandas, matplotlib, numpy, Streamlit, langchain, RAG Tools / Platforms: Git, Kaggle, Tableau, Microsoft Power Bi, Excel

Databases: SQL

PROJECTS / OPEN-SOURCE

#### Data Visualization of Pharmaceutical Company | Link

R Studio, rvest, Shinny

- Leveraged web scraping tools such as rvest and R programming to extract comprehensive data from pharmaceutical companies.
- Conducted thorough data preparation and cleansing to ensure accuracy and reliability.
- Developed an interactive Shiny app that provides insightful visualizations and analysis of the pharmaceutical sector and its growth trends.

## Machine Learning Projects | Link

Jupyter

- Completed a self-paced machine learning course on Internshala, leveraging advanced data analysis and machine learning techniques to solve real-world problems:
- Performed data cleaning, feature scaling, and feature engineering to implement Linear Regression for house price prediction
- Applied data preprocessing, model evaluation, and Logistic Regression to predict bank loan churn rates

• Executed unsupervised learning using Clustering algorithms, with dimensionality reduction techniques for population classification

## CERTIFICATIONS

- $\bullet$  Machine Learning  $\bf Internshala$
- Introduction to Artificial Intelligence Simplilearn
- $\bullet$  Introduction to Prompt Engineering  $\mathbf{Simplilearn}$
- ChatGPT Advanced Course Simplilearn
- CCC National Instiite of Electronics and Information Technology
- Introduction to Psychology Yale (Coursera)
- Maulvi UP Board of Madrasa Education Lucknow