

# MOHAMMAD SAQIB ANSARI

ansarimdsaqib@gmail.com | Mumbai, Maharashtra

Portfolio | HackerRank | GitHub | LinkedIn | Blog

## EDUCATION

---

### Indian Institute of Technology, Kanpur

Statistics Masters

CGPA: 8.1

Kanpur-UP,India

July 2022 - July 2024

### Ewing Christian College

Statistics and Mathematics

Percentage: 72.22%

Allahabad-India

Aug-2018 - July-2021

### Dawn The Higher (CBSE)

PCM HSC

Percentage: 77%

Allahabad-India

Jul 2016 - Jul 2017

## EXPERIENCE

---

### ICICI General Insurance Company | Associate Data Scientist

Mumbai, Maharashtra | Jul-2024 - Present

- Conducted EDA on PhonePe Insurance data, deriving actionable insights to inform business strategies.
- Built interactive dashboards in Tableau and Power BI, including a Cross-Sell Data Dashboard for customer analytics.
- Developed deployment-ready code for a parent-child retrieval chatbot using RAG and LangChain, improving user engagement and information access.
- Built a predictive lead scoring model using Google Analytics two-wheeler data, optimizing call center conversions.

### Self Employed | Data Science Freelancer

Remote - India | Oct 2022 - Present

- Collaborated with platforms like **TutorBro**, **Way2Class**, and **Chegg** as a freelancer, delivering data science and statistics support.
- Assisted students globally by solving data science and statistical problems and offering coding guidance.
- Taught courses and prepared **video solutions** to simplify complex concepts for learners.
- Worked on diverse **data science projects**, leveraging analytical and programming skills to provide tailored solutions.

### Meritshot | Data Science intern

Kanpur, UP | Sep-2023 - Nov-2023

- Established a Medium blog titled '**Data Science in IoT**' to share insights and knowledge.
- Published a **business case study on Lenskart**.
- Created a **Data Preprocessing Cheatsheet**, serving as a quick reference for data wrangling techniques.
- Authored an article exploring **Statistical Functions in R Programming**, providing practical use cases and applications.

## 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, Shiny*

- 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

---

- Machine Learning - **Internshala**
- Introduction to Artificial Intelligence - **Simplilearn**
- Introduction to Prompt Engineering - **Simplilearn**
- ChatGPT Advanced Course - **Simplilearn**
- CCC - **National Institute of Electronics and Information Technology**
- Introduction to Psychology - **Yale (Coursera)**