

Soha Tariq

st.sohatariq@gmail.com | (240) 899-8167 | New York, NY

LinkedIn: <https://www.linkedin.com/in/soha-tariq-5226143b/> | GitHub: <https://github.com/SohaT7>

Portfolio: <https://sohat7.github.io/Portfolio/>

Data analyst who recently earned a certificate in Data Analytics from Columbia University, with a background in Mathematics and Economics. Past work includes data analytics, project management, and building stakeholder relationships. Experienced in handling large datasets consisting of structured, semi-structured, and raw data, and collaborating across diverse groups to support creative, detailed and efficient analysis. Strengths include analytical problem-solving, empathic teamwork, and effective communication.

Technical Skills

Languages: Python, R, Structured Query Language (SQL), HTML5, CSS, JavaScript ES6+

(Libraries: Pandas, NumPy, SciPy, Matplotlib, Scikit-Learn, TensorFlow, Plotly.js, D3.js)

Applications, Systems, and Tools: MongoDB, PostgreSQL, SQLAlchemy (ORM), Amazon Web Services (AWS), Google Cloud Platform (GCP), Alteryx, Tableau, Flask, Hadoop, MS Office Suite (Excel, Word, PowerPoint), Visual Basic Application (VBA), GitHub, APIs, ETL, Agile methodologies

Projects

Amazon_Vine_Analysis | GitHub Link: https://github.com/SohaT7/Amazon_Vine_Analysis

Analysis to determine if there is a positivity bias in product reviews written by Amazon Vine members.

- Tools: PySpark, Python (Pandas library), SQL, Amazon Web Services (AWS), pgAdmin

Credit_Risk_Analysis | GitHub Link: https://github.com/SohaT7/Credit_Risk_Analysis

Supervised machine learning models built and evaluated to predict credit loan risk. Resampling and ensemble techniques are applied to the logistic regression classifier models.

- Tools: Python (Scikit-learn, Imbalanced-learn, Pandas, and NumPy libraries)

Bikesharing | GitHub Link: <https://github.com/SohaT7/Bikesharing>

Creates a dashboard to analyze the NYC Citi Bike sharing data, in order to convince investors to start a similar bike sharing program in another city.

- Tools: Tableau and Python (Pandas library)

NYC_Restaurant_Analysis | GitHub Link: https://github.com/SohaT7/NYC_Restaurant_Analysis

Data visualization of the NYC restaurant data, and data analysis to gauge if a restaurant located in a high-income area receives a higher health inspection grade.

- Role: Machine Learning
- Tools: Python (Pandas, Scikit-learn, Imbalanced-learn libraries), PostgreSQL, SQLAlchemy, Tableau, JavaScript (Plotly.js library), HTML, CSS, and Bootstrap

Relevant Experience

Success Academy

Bronx, NY

Sprint Associate

June 2021 - November 2021

- Managed the provision of academic, behavioral, psychological, and socio-emotional support to the students requiring special educational services in a school of 625 students.

- Improved process design by adding data analytics (using Excel) which made early detection of at-risk students much more efficient.
- Trained teachers to conduct the Intervention Cycles more effectively - improved graduation rates from Intervention by about 10%.
- Conducted effective meetings and built stakeholder relationships (with NYC Department of Education (DOE) personnel, teachers, psychologists, and parents) to ensure the provision of special services to the students in need.

Lahore University of Management Sciences (LUMS)

Lahore, Pakistan

Research Assistant

June 2017 - March 2018

- Built relationships over recurrent meetings with local leaders in the community to assess issues faced when engaging in the politics of infrastructure at the local level.
- Conducted data collection, followed by data analysis in Excel and STATA.

Lahore University of Management Sciences (LUMS)

Lahore, Pakistan

Research Assistant

September 2013 - December 2013

- Managed a team of 9 in data collection (surveys, interviews, and focus groups) from 8,000 households in 15 low-income neighborhoods.
- Conducted mapping of the neighborhoods - made data collection faster by about 50%.
- Conducted data analysis and data visualization in Excel and STATA.

Innovative Development Strategies (IDS)

Lahore, Pakistan

Summer Intern

June 2014 - August 2014

- Conducted data collection (questionnaires, interviews, and focus groups) in 7 localities, followed by data analysis and data visualization of the aggregate data.
- Delivered a presentation to government stakeholders on the issues faced by the Internally Displaced People (IDPs), with recommendations to alleviate those issues.

Additional Experience

University of Virginia

Charlottesville, VA

Teaching Assistant

August 2019 - December 2020

United We Reach

Lahore, Pakistan

Curriculum Writer

November 2018 - June 2019

AIESEC in Lahore; AIESEC in Kenya

Lahore, Pakistan

Project Manager

September 2011 - August 2014

Education

Certificate, Data Analytics

Columbia University, New York

December 2021 - June 2022

Masters of Arts, Anthropology

University of Virginia, Charlottesville

August 2019 - December 2020

Bachelor's of Science (Honors), Politics & Economics

Lahore University of Management Sciences, Lahore

Minor in Mathematics, and Anthropology-Sociology

September 2011 - July 2017