

# Soha Tariq

[st.sohatariq@gmail.com](mailto:st.sohatariq@gmail.com) | (240) 899-8167 | Washington DC

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

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

## Summary

---

Past work experience includes data analysis and visualization, databases and data management, data collection and research, project management, building relationships with multiple stakeholders, event management, budgeting, and administrative work. Willing to relocate.

**Technical Skills:** Microsoft Office (PowerPoint, Word, Access, Excel), VBA, R, Tableau, Power BI, SQL, Python, STATA, JavaScript/HTML/CSS, SPSS, SAS, Qualtrics, N-Vivo, MySQL, PostgreSQL, NoSQL (MongoDB), Github, Linux, Google Suite (Google Sheets), Amazon Web Services (AWS), Google Cloud Platform (GCP). Data analysis, data visualization, database management, machine learning.

## Education

---

<b>Data Analytics &amp; Visualization</b>	Columbia University (bootcamp)	June 2022
<b>Masters of Arts, Anthropology</b>	University of Virginia	August 2019 - December 2020
<b>Bachelors of Science, Economics &amp; Politics   Minor in Mathematics</b>	LUMS	Sept 2011 - July 2017

## Relevant Experience

---

**Columbia University** New York, NY

*As part of the Data Analytics & Visualization Bootcamp learning experience:* Dec 2021 - June 2022

- Ran statistical analysis on large datasets to model, predict, forecast, and identify trends with Microsoft Excel, VBA, Python, SQL, PySpark, and R.
- Created data processing pipelines that take data from multiple sources, and cleaned, transformed, summarized, and loaded data into databases, using SQL and Python.
- Constructed web applications and interactive maps, generated data visualizations and dashboards, and conducted web scraping using Python, R, Tableau, and JavaScript/HTML/CSS.
- Developed Machine Learning and Deep Learning predictive models, moved data across multiple platforms, and carried out ETL and data analysis on the data, using Python and AWS.

**Success Academy** New York, NY

Associate, Data Analyst (Special Education) June 2021 - November 2021

- Owned the end-to-end analysis including data collection, data analysis in Excel and SQL, deliverables and reports in Tableau, and made recommendations to the Committee on Special Education.
- Developed advanced analytical solutions which led to improved processes and greater efficiency in early detection of at-risk students in the school by about 25% - improved database querying, stored procedures, automated reporting via dashboards and visualizations to identify trends.
- Collaborated with cross-functional partners and stakeholders, including NYC Department of Education (DOE) government officials, and helped finalize academic plans for students in need.

**Lahore University of Management Sciences (LUMS)** Lahore, Pakistan

Data Research Assistant September 2013 - December 2013; June 2017 - March 2018

- Built outreach to local communities to understand their perspective on infrastructure and intergenerational occupational (economic) mobility.
- Led a team of 9 in data gathering from 8,000 households in 15 low-income neighborhoods.

- Conducted literature reviews and the mapping of neighborhoods unavailable on Google Maps, which made the entire data gathering process faster by about 50%.
- Developed and maintained a database, conducted data mining, data and statistical analysis, and data visualization, in R, STATA, Microsoft Excel, and N-Vivo.

### **Innovative Development Strategies (IDS)**

Lahore, Pakistan

Data Analyst Intern

June 2014 - August 2014

- Conducted data gathering (survey research, interviews, focus groups) in 7 localities, followed by data analysis and graphical representation of the aggregate data in Microsoft Excel.
- Prepared reports and presentations for the government stakeholders on the issues faced by 5000+ Internally Displaced People (IDPs), with recommendations to alleviate those issues.

### **AIESEC**

Lahore, Pakistan; Kenya (remote)

Project Manager; Marketing Advisor

September 2011 - August 2014

- Conducted market research and media analysis, monitored marketing KPIs, measured ROI for the online advertising campaign, and made recommendations for more effective marketing strategies; improved user engagement by about 30%; brought in 25%+ more customers.
- Successfully matched incoming talent from various parts of the world with 15+ local businesses; built strong stakeholder relationships with the client companies.
- Managed three projects on education, women empowerment, and sustainability.

### **Volunteer Work**

**Lahore Conservation Society**

Administrative Associate

January 2012 - May 2017

Orange Train Project in Lahore and the Lahore Canal Project:

- Conducted stakeholder interviews and focus-discussion groups in communities; evaluated project's adverse impacts on the centuries old historic sites in Lahore.
- Prepared presentations providing solutions to alleviate the problem of traffic congestion.
- Community engagement to present communities' viewpoints to the local government.

### **Additional Experience**

**University of Virginia**

Graduate Teaching Assistant

August 2019 - December 2020

**United We Reach**

Curriculum Developer

November 2018 - June 2019

### **Projects**

**Kickstarter Campaigns Analysis | GitHub:** <https://github.com/SohaT7/Kickstarter-Analysis>

Analyzes a dataset consisting of 4,000 crowdfunding projects to discover hidden trends (campaign performance based on launch dates and funding goals) through formulas, pivot tables, pivot charts.

- Tools: Microsoft Excel

**School District Analysis | GitHub:** [https://github.com/SohaT7/School\\_District\\_Analysis](https://github.com/SohaT7/School_District_Analysis)

District-level and school-level analysis: calculates the math, reading, and overall passing percentages in schools, and handles faulty data.

- Tools: Python (Pandas)

**Purchase\_Analysis | GitHub:** [https://github.com/SohaT7/Purchase\\_Analysis](https://github.com/SohaT7/Purchase_Analysis)

Predictive analytics on 900,000 records in the Google Analytics ecommerce dataset; builds a Machine Learning Model to predict whether a visit to the website will result in a sales transaction or not.

- Tools: Google Cloud Platform (GCP), BigQuery, Vertex AI API, MySQL, Python, BigQuery Client

**NYC Restaurant Analysis | GitHub:** [https://github.com/SohaT7/NYC\\_Restaurant\\_Analysis](https://github.com/SohaT7/NYC_Restaurant_Analysis)

Dashboard to visualize the NYC restaurant data, and a Machine Learning Model to predict whether a restaurant located in a high-income area will receive a higher health inspection grade or not.

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