Saba **Suhail**

DATA SCIENTIST

San Francisco Bay Area, CA

☐ (415) 766-2331 | ✉ sabasuhail.1112@gmail.com | ⌂ sabaiitj.github.io | ⌾ github.com/sabaiitj | 🔗 sabasuhail

About

Data Scientist with engineering background. Experienced in drawing insights from data using statistical principles and sound research design methodology to make critical business decisions and build data-driven products. I'm seeking a data science position focused on hard problems with the opportunity to grow into a leadership role.

Work Experience

Research Assistant
Bangalore, India

INDIAN INSTITUTE OF MANAGEMENT
Aug 2020 - Oct 2021

- Worked under various professors of Strategy and was responsible for gathering data, building and analyzing models, running A/B testing and writing papers.
- Analyzed census data of India for female literacy data for year 1909 and microfinance institutions performance data for year 2019 in various districts in India to find causal relationship in various Indian districts for year 2019.
- Collected sustainability data for Fortune 500 companies from 1955 to 2019 and GDP of various countries from 1960-2019, analyzed and made a report. This was added into a EdX Sustainability MOOC course.
- Skills: Data Cleaning, Causal Modeling, Visualization, Operation Research, Strategy Management, Econometrics, Innovation Management, Research Design, Stata, Python.

Scientist C (Data Scientist)
New Delhi, India

DEFENSE RESEARCH DEVELOPMENT ORGANIZATION
Jul 2018 - Jul 2020

- Built statistical and mathematical models for multiple business needs.
- Worked on log data extraction and analysis efforts to reduce the signal loss errors between device communication and hand-offs.
- Worked with cross-functional team of scientists, technical officers and external stakeholders to optimize the inefficient strategic business processes.
- Skills: SQL, Python, Data Analysis, Data Wrangling, Linear Regression, Statistical Analysis, Data Modeling, Report/Dashboards, Frequency Analysis.

Scientist B (Data Analyst)
New Delhi, India

DEFENSE RESEARCH DEVELOPMENT ORGANIZATION
Aug 2014 - Jun 2018

- DRDO is Indian Govt's Defense research organization and I joined as Scientist responsible for cryptographic systems research.
- Responsible for building, testing and deploying cryptographic communication systems, which included extensive data analysis, validation and managing dashboard/reports.
- Built a new practical hybrid random generator using a low frequency oscillator and ISAAC pseudo RNG, validated the statistical properties of the output in MATLAB.
- Skills: SQL, MATLAB, Data Analysis, Data Modeling, Report/Dashboards, Project Management.

Research Intern
Madison, WI

UNIVERSITY OF WISCONSIN
May 2013 - Jul 2013

- Worked on developing a tool for power electronics research by linking together a set of simple electrical components which makes power electronics research safer and more accessible.

Education

Data Science Bootcamp
Remote/San Francisco, CA

GENERAL ASSEMBLY
Apr 2022 - Jul 2022

- Used Natural Language Processing techniques, various classification algorithms and Ensemble Methods to predict the popularity of reddit posts scraped using Reddit API.
- Collaborated with a team of fellow data scientists to use machine learning principles to make predictions about the climate awareness of a typical American citizen.
- Worked on a capstone project to apply Natural Language Processing techniques, classical Machine Learning and Deep Learning methods to identify sarcastic posts.

PhD Strategic Management
Bangalore, India

INDIAN INSTITUTE OF MANAGEMENT
Aug 2020 - Oct 2021

- Relevant Courses: Statistics for Management Research, Econometrics, Technology and Innovation Management, Research Design Methodology, Qualitative Research Methods, Linear Algebra, STATA

B.E. Electrical Engineering
Jodhpur, India

INDIAN INSTITUTE OF TECHNOLOGY
Jul 2010 - Jun 2014

- Relevant Courses: Linear Algebra, Probability and Statistics, Multivariate Calculus, Programming and Data Structures, Game Theory

Relevant Skills

Programming Tools	SQL, Python, R, Numpy, Pandas, Sklearn, Scipy, Keras, Git, AWS
Data Visualization	Matplotlib, Tableau, Seaborn, Plotly, Excel, Google Sheets
Machine Learning	EDA, Linear Regression, Logistic Regression, KNN, Random Forests, Decision Tree, NLP, Neural Nets, Clustering, PCA
Statistical Methods	Predictive Modeling, Causal Modeling, Hypothesis (A/B) Testing, Time Series Analysis