

Ayman Khan

5793 Tyndall Ave, Bronx, NY 10471 | 332-215-2663 | aymankhaninfo@gmail.com | www.linkedin.com/in/aymanawsafkhan | <https://aymankhan2003.github.io/aymansportfolio.github.io/>

Education

Middlebury College

Candidate for Bachelor of Arts, Major in Computer Science

Middlebury, VT

September 2021 - May 2025

Relevant Coursework: Data Structures, Machine Learning, Computer Graphics, Regression, Delving into Data, Statistical Learning, Algorithms and Complexity, Software Development, Natural Language Processing

Languages: Python, R, C++, C, Java, JavaScript, HTML, CSS, Stata, SQL, Excel, Microsoft Suite

Frameworks and Libraries: React, NextJS, Flask, Material-UI, Sci-kit learn, pandas, NumPy, Tesseract, p5.js, HuggingFace, PyTorch, Transformers, TensorFlow, Jupyter, Matplotlib, NodeJS, Docker, Typescript, Git, Jira

Experience

Middlebury Consulting Group

Student Data Analyst - Data Team

Middlebury, VT

September 2024 - May 2025

- Collaborated with a team of 5 to analyze datasets containing over 100,000 entries, using R and Python, improving model accuracy by 15% and delivering client presentations.
- Engaged with our clients, World Data Lab, to research and enhance predictive models to business challenges, implementing innovative approaches that improved data-driven decision-making.
- Developed validation frameworks to assess and compare model performance, ensuring the reliability of new data estimates and communicated findings through pitch decks.

Green Mountain Care Board

Data Analyst Intern

Montpelier, VT

June 2024 - August 2024

- Analyzed datasets using R and SQL, covering over 200,000 healthcare providers and 2 million patients to assess network adequacy, provider availability, and access across various insurance plans.
- Created detailed visualizations and reports, highlighting trends and disparities in provider networks, which informed policy recommendations to improve Vermonters' healthcare access.
- Conducted surveys to gather stakeholder insights, analyzed data trends, and presented findings to the Vermont Health Board, collaborating with policymakers to refine network adequacy metrics.

Fairbanks Museum & Planetarium

Data Analyst Intern

St. Johnsbury, VT

June 2023 - August 2023

- Converted over 1,000 handwritten flashcards into structured Excel data using the Tesseract machine learning library in Python, improving data accessibility and organization.
- Analyzed trends in flower bloom and fruiting dates using R, leveraging dplyr, tidyverse, and ggplot to create visualizations that enhanced ecological research insights.
- Joined research calls and standups to align on goals, managed an online data dashboard, and collaborated on interactive visualizations to enhance data accessibility and public engagement.

Projects

Investigating the Attributes that Influence Player Ratings in FC24

- Utilized Support Vector Machine (SVM) to predict player ratings in FC24, achieving an R^2 of 0.908, indicating high model accuracy.
- Applied hierarchical clustering and principal component analysis to assess attribute impacts on player ratings.
- Conducted data analysis and visualizations using R, identifying key player attributes that influence ratings with minimal error (MSE = 4.44).
- <https://aymankhan2003.github.io/aymansportfolio.github.io/assets/FinalStats.html>

Market Value Predictor

- Built machine learning models (Random Forest Regression and LASSO) to predict soccer player market values, achieving 95% accuracy with RFR and 75% with LASSO, leveraging feature engineering in Python with scikit-learn, pandas, and numpy.
- Conducted extensive data analysis and manipulation in R, including web scraping, to preprocess and refine datasets for model training.
- Developed an interactive UI using Flask (backend) and HTML (frontend) to enable real-time market value predictions.
- <https://aymankhan2003.github.io/posts/project/ProjectBlogPost.html>