MADHAN JEGANATHAN

Los Angeles, CA • madhanjeg.work@gmail.com • linkedin.com/in/madhan-jeganathan • madhan-jeg.github.io

EDUCATION

University of California, Santa Barbara (UCSB), Santa Barbara, CA *Graduated: June 2023*
Bachelor of Science (BS), Statistics and Data Science

- **Relevant Coursework:** Statistical Machine Learning, Big Data Analytics, Regression Analysis, Bayes Data Analysis, Design of Experiments, SAS Base Programming, Stochastic Processes, Time Series, Probability and Statistics

WORK EXPERIENCE

Data Science Intern, Shooting Stars Foundation, Fremont, CA *July 2024 - Present*

- Standardized and cleaned registration data from thousands of student entries, improving data consistency and accuracy across many events.
- Built a database to normalize open-ended college names and accurately track student registration, enabling a more effective college outreach plan.
- Automated data validation and transformation processes to streamline event reporting and ensure high-quality data for analysis.
- ***On-Site Volunteer***, 2024 Amgen STEM Entrepreneurship Program, Thousand Oaks, CA
  - Actively participated in daily staff meetings discussing program planning and organization.
  - Helped teams of students refine their entrepreneurship ideas and presentations by providing feedback and resources.
  - Developed student feedback surveys to collect data crucial for program improvement. Analyzed collected data to extract and present insights daily, resulting in a 15% increase in overall rating throughout the event.
  - Produced a comprehensive event report, using libraries such as Matplotlib and Seaborn to generate clear and informative visualizations.

Flex Associate, Amazon Fresh, Northridge, CA *November 2020 - February 2022*

- Utilized order-processing software to complete and organize online orders.

PROJECT EXPERIENCE

Unsupervised Clustering and Dimensionality Reduction in Python, Los Angeles, CA *March - June 2024*

- Created a unique dataset from a large, diverse, personal Spotify playlist by utilizing Spotify's Web API to extract relevant audio features from songs in a playlist.
- Implemented a clustering algorithm to group similar songs based on their numerical characteristics and tested different k values to find the optimal number of clusters.
- Conducted a principal component analysis to reduce dimensionality and visualize differences between each cluster.
- Ensured a more organized listening experience by curating smaller, more cohesive playlists from a larger playlist.

Statistical Analysis in Python and Interactive Tableau Dashboard, Los Angeles, CA *October 2023 - January 2024*

- Analyzed a World Happiness Report dataset containing data from 165 countries from 2005 to 2022 to identify trends in national happiness levels over time.
- Employed statistical analysis techniques such as correlation matrices and principal component analysis to determine key drivers of happiness across nations.
- Drew insights into how GDP, social support, corruption, and other factors relate to national happiness.
- Created an interactive dashboard with Tableau in which filters like region, country, and column simplify comprehension of the multi-dimensional dataset.

Predictive Modeling with Machine Learning in GCP, UCSB, Santa Barbara, CA *January - March 2023*

- Analyzed US voter participation trends across over 40 million rows of data from national voter files using PySpark and Google Cloud Platform.
- Trained machine learning models with logistic regression and decision trees to predict voter turnout based on education level, income, and other variables.
- Compared feature importances between historically Democratic and Republican states to draw insights about how state political affiliation and other factors such as education level affect voter turnout.

SKILLS

- **Programming Languages:** Python, R, SQL, SAS, Java, C++, JavaScript
- **Other Tools:** Pandas, Scikit-learn, Tableau, MySQL, GCP, PySpark, Excel, Jupyter Notebook, HTML, CSS, LaTeX