Pratham Aggarwal

619-953-7147 | praggarwal@ucsd.edu | in/pratham-agg | github/pratham-aggr

Education

Bachelor of Science, Data Science GPA 4.0

Halıcıoğlu Data Science Institute, University of California San Diego Expected Jun 2027

DSC 10: Principles of Data Science, UC San Diego Fall 2024

- Data cleaning, Exploratory Data Analysis, TF-IDF text analysis, Hypothesis Testing, Data Storytelling, Web Scraping, Multi-Dimensional Linear Regression, A/B testing

DSC 20: Programming and Data Structures for Data Science, UC San Diego Win 2025

- Object-oriented programming (OOP), Time & Space Complexity, Higher-order functions

Projects

Predictive Modeling of Heating and Cooling Loads *Learn more*

- Predicted energy loads with 91% accuracy by utilizing multidimensional linear regression and R-squared as a performance metric, optimizing energy consumption models and improving building energy efficiency.
- Identified a significant shift in energy load behavior by applying k-means clustering, leading to deeper insights into energy consumption patterns and informing more accurate forecasting models.
- Conducted a thorough Exploratory Data Analysis (EDA) on a building features dataset using Python, Pandas, and Matplotlib, uncovering key trends, outliers, and correlations, which enabled more targeted energy optimization strategies.

Rebooting Sitcom Friends with Data-Driven Strategies for Enhanced Engagement *Learn more*

- Made data-driven decisions on episode ratings and viewership trends by applying statistical testing (bootstrapping, hypothesis testing), improving decision-making for a potential reboot and refining content strategies for higher audience engagement.
- Improved storyline development by using sentiment analysis and Bayesian modeling to determine which Friends characters are best suited for specific roles, resulting in more engaging content and better-aligned character development.
- Determined the optimal 80-episode count by developing a linear regression model to maintain audience interest and engagement throughout the series.

TSwift Tunes: Data-Driven Insights and Recommender Systems *Learn more*

- Uncovered trends and visualized feature relationships by conducting exploratory data analysis (EDA) on Taylor Swift's discography, which led to more accurate song categorization and a deeper understanding of her musical patterns.
- Built a recommender system by utilizing audio features, which will optimize the user experience and deliver personalized track recommendations, increasing user satisfaction and engagement.
- Built a TF-IDF-based lyric search tool, which resulted in faster and more accurate identification of key themes in large-scale text data.

Simulating Black Hole Evolution: Comparative Analysis of Light and Heavy Seeds *Learn more*

- Simulated supermassive black hole growth using Eddington and super-Eddington models, processing astrophysical datasets to analyze different formation pathways.
- Wrote a research paper contrasting black hole growth based on heavy and light seed masses, using visualizations to compare their growth trajectories, resulting in clearer insights into the impact of seed mass on black hole development.
- Conducted 12 weeks of research under the guidance of PhD student Matthew Scoggins from Columbia University, contributing to advancements in computational astrophysics.

Experience

Program Manager & Consultant Mar 2025 – Present

Data Science Student Society (DS3) Consulting *Solana Center*

- Enhance data quality by performing comprehensive data cleaning and exploratory data analysis (EDA), extracting useful information from data
- Lead effective communication and coordination with Solana Center team, organizing meetings to ensure alignment between project goals and business objectives.

Honors & Awards

Provost Honors, Revelle College, UC San Diego Fall 2024

AP Scholar with Distinction, College Board Jul 2024

Skills

Programming Languages: MATLAB, SQL, Java, Python

Libraries: Pandas, NumPy, SciPy, Scikit-Learn, BeautifulSoup

Data Visualization: Tableau, Matplotlib, Seaborn

Version Control: Git & Github

Other: Experiment Design, Research Ethics, Documentation, Terminal