

PRATHAM AGGARWAL

San Diego, CA | 619-953-7147 | praggarwal@ucsd.edu | [Portfolio](#) | [in/pratham-aggr](https://in.pratham-aggr) | [github/pratham-aggr](https://github.com/pratham-aggr)

SUMMARY

Aspiring Data Scientist eager to leverage business intelligence and advanced data-driven techniques to develop impactful solutions across various domains, focused on delivering actionable and deployment-ready outcomes.

EDUCATION

Bachelor of Science, Data Science

GPA 4.0

Halicioğlu Data Science Institute, University of California San Diego

Expected Jun 2027

Relevant Course: Data Structures and Algorithms for Data Science (DSC 30), UC San Diego

PROJECTS

Predictive Modeling of Heating and Cooling Loads

[Learn more](#)

- Predicted energy loads with 91% accuracy by utilizing machine learning and 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](#)

- Enhanced song categorization and deepened understanding of musical patterns by conducting exploratory data analysis (EDA) on 5 datasets, uncovering key trends and visualizing relationships between features.
- Built a personalized recommender system using over 10 audio features, optimizing the user experience and driving increased satisfaction and engagement through tailored track recommendations.
- 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, data 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.

EXPERIENCE

Program Manager & Consultant

Mar 2025 – Present

Data Science Student Society (DS3) Consulting

Solana Center

- Collaborated with Solana Center to solve composting-related problems by utilizing advanced analytics, machine learning.
- Communicated technical insights to non-technical stakeholders, articulating complex data into actionable insights to improve customer experience and develop effective business cases.
- Mentored team members to meet weekly targets, managing flexible schedules and ensuring alignment with business goals through calendar coordination and time management.

HONORS & AWARDS

Two-Time Provost Honors, Revelle College, UC San Diego

Win 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: Excel, Experiment Design, Research Ethics, Documentation, Terminal