

Carlos Vega

+1 (917) 943-2594 | carlosvegalemus@gmail.com | [linkedin.com/in/carlos-vega-lemus](https://www.linkedin.com/in/carlos-vega-lemus) | github.com/iCarlosVega

OBJECTIVE

Computer Science and Mathematics double major seeking entry-level Software Engineering roles, leveraging experience in full-stack development, data analysis, and machine learning.

PROJECTS

- Class Schedule Database** | *T-SQL, Docker, Azure Data Studio* Fall 2024
- Developed automated data loading procedures handling 3,200+ building locations and 3,889 room assignments.
 - Processed 4,500 course records by developing T-SQL data validation that identified and corrected hundreds of inconsistencies from the original dataset, achieving 100% data integrity post-migration.
 - Optimized database indexing and query strategies, resulting in an average 40% faster retrieval.
- Apartment Price Prediction Model** | *Python, Scikit-learn, Pandas, NumPy, Matplotlib* Spring 2024
- Developed a Random Forest Regressor (100 estimators) to predict housing sale prices using a 2016-2017 dataset (2,230 records, 55 initial features), achieving an R^2 of 0.903 and an RMSE of \$52,036 on the test set.
 - Implemented an imputation strategy using iterative Random Forest Regression to fill missing values across 13 key features, averaging 75.7% missing values across them.
 - Conducted feature importance analysis, identifying listing price (53.9%), num of full bathrooms (14.9%), and walk score (5.4%) as the top three predictors of sale price.
- AI Admissions Assistant** | *Python, React, Flask, Azure AI Foundry* Fall 2023
- Spearheaded the content acquisition for RAG by recursively web-scraping and processing over 500 unique pages, yielding a comprehensive dataset of 2023 information.
 - Engineered a data pipeline to vectorize over 2GBs of text data using word2vec, creating a specialized knowledge base for the RAG system.
 - Reduced outdated responses from the base LLM by approximately 65% for inquiries requiring 2023 information

EDUCATION

- CUNY Queens College** Queens, NY
- Bachelor of Arts in Computer Science and Mathematics*
- Relevant Coursework:** Data Structures, Object-Oriented Program in Java, Database Systems, SQL Programming and Modern ERD, Probability Theory, Statistical Theory and Computation Stats for Data Science.

EXPERIENCE

- Let's Get Ready** Sept. 2024 - Present
- Core Coach* Remote
- Provide college assistance and support to a cohort of 250+ students through the Mainstay messaging platform.
 - Respond to student inquiries, offering guidance on college-related matters to facilitate their academic journey.
 - Perform data entry tasks, including updating student information and processing opt-out requests to maintain accurate records.
- QC Learning Commons** Sept. 2024 - May 2025
- Tutor* Queens, NY
- Provided academic support to undergraduate students in Machine Learning, Intro to Programming, Object Oriented Programming in Java.
 - Facilitated understanding of complex concepts and problem-solving techniques through one-on-one and small group sessions.

ACTIVITIES

- 3DPrints - Custom 3D Printing Business** January 2025 - Present
- Designed and delivered 200+ promotional products for Big Brothers Big Sisters of NYC.
- Google Software Engineering Program** Oct. 2024 - January 2025
- Gained insights into Google's best coding practices, and software engineering principles through direct mentorship.
- Physics Club, Queens College** Fall 2023 - May 2025
- Lead of Technology - Developed and maintained the club's website.

SKILLS

Languages: Python (Proficient), JavaScript/TypeScript, SQL

Frameworks & Libraries: React, Next.js

Tools: Git/GitHub, Docker, Linux, Azure Data Studio, AWS, AWS S3, PostgreSQL, Node.js, Flask

Languages: English (Native), Spanish (Native)

Soft Skills/Interest: Problem-Solving, Effective Communication, Adaptability, 3D Printing