

# Ryan Voda

818-370-1304 | [ryanvoda@gmail.com](mailto:ryanvoda@gmail.com) | [linkedin.com/in/ryan-voda](https://www.linkedin.com/in/ryan-voda) | [ryanvoda.github.io](https://ryanvoda.github.io)

## EDUCATION

---

### Georgia Institute of Technology

Jan. 2024 – Jan. 2026

*Master of Science in Computer Science*

### University of California, Los Angeles

Sep. 2020 – Jun. 2023

*Bachelor of Science in Mathematics; GPA: 3.5*

- Coursework: Python with Applications, Linear Algebra, Optimization, Algorithms, Introduction to Probability, Introduction to Computer Science, Real Analysis, Applied Numerical Methods, Multivariable Calculus

## EXPERIENCE

---

### Data Management Trainee

Aug. 2023 – Present

*Sony Pictures Entertainment*

- Developed a Python script to clean and manipulate data inputs, streamlining the ingestion process.
- Performed MLOps duties with SQL to build a reliable pipeline for deployment and maintenance of models.
- Identified data discrepancies and failures in Apache Airflow for major YouTube channels.

### Data Management Intern

Jun. 2023 – Aug. 2023

*Sony Pictures Entertainment*

- Automatically ingested social media data and generated financial reports using BigQuery and Excel.
- Created ETL pipeline to integrate data from new vendor into Snowflake and automate workflow.
- Profiled 4 massive datasets and provided visualizations and insights to team members.
- Collaborated with data science team to make new features and metadata for machine learning models.
- Completed ad-hoc requests to maintain databases and assist with model creation in Python.

### Teaching Assistant

Jan. 2023 – Mar. 2023

*(UCLA) Calculus III T.A.*

- Led review sessions in weekly discussion sections for 35 students.
- Held 60-minute office hours to simplify Calculus 3 concepts.
- Coordinated with the TA and professor to implement lesson plans and establish a positive learning environment.

### Research Assistant

Mar. 2022 – Jun. 2022

*(UCLA) California Health Survey Analysis with Prof. Harlin Lee*

- Cleaned and interpreted CHIS-2020 data in Python using pandas and numpy.
- Found correlations between distress levels and 13 factors related to COVID-19 with linear regression.
- Verified 92 percent correlation by applying scipy statistical tests such as spearmanr and chisquare.
- Produced data analysis summary using plotly and matplotlib.

## PROJECTS

---

### Milk Grade Classification | Python

Aug. 2022

- Imported kaggle dataset to predict milk quality using categorical and quantitative characteristics of milk.
- Cross-validated feature selection tests to find the strongest predictive features of milk grade.
- Conducted hyperparameter tuning to apply results on several appropriate Machine Learning models.

### Penguins Species Predictor | Python

Feb. 2022

- Worked with team to build optimal Machine Learning model for predicting penguins species from the Palmer Penguins dataset.
- Performed data analysis and visualization to narrow down our selection of features.
- Designed a function to automatically clean and prepare data for train/test split.
- Evaluated performance of several models, such as Random Forest, Logistic Regression, and Clustering.

## TECHNICAL SKILLS

---

**Languages:** Python, SQL, R, JavaScript, MATLAB, C++

**Technologies:** Excel, Tableau, Periscope, SAS

**Other:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Mosaic, Git