

A.J. Goldman

[linkedin.com/in/aj-goldman](https://www.linkedin.com/in/aj-goldman) | ajgoldman@gmail.com | 818-800-2809 | ajgoldman.github.io/My-Portfolio/

Education & Certifications

University of California, Los Angeles, Los Angeles, CA

- B.S. Chemical & Biomolecular Engineering, Minor in Theatre June 2019
 - **Relevant Coursework:** Computer-Aided Design and Analysis with Pro/II, Computer Programming with C++, Computer Programming with MATLAB, Introduction to Geographic Information Systems

Board for Professional Engineers, Land Surveyors, and Geologists

July 2019

- Engineer-in-Training: Certificate #168486

Thinkful, Online

- Data Science Flex Program, Certificate of Completion January 2020
 - **Relevant Coursework:** Programming in Python, Data Visualization & Experimentation, Supervised/Unsupervised Machine Learning, SQL Fundamentals, Statistics & Probability

Work Experience

Deloitte Risk & Financial Advisory, *Consultant - Data Scientist*

June 2021 – Present

Federal Law Enforcement Agency (Contract)

- Created and improved upon Python applications, Flask APIs, NiFi flows, and Kibana dashboards, focused on using Natural Language Processing to assist the client in quickly and accurately identifying criminal activities, especially those related to money laundering.
- Developed a Python module with a Command Line Interface to allow forensic accountants to process large CSV files of financial transaction data. Whereas the accountants had previously relied on a basic Excel template with functions, the new tool allowed them to process orders of magnitude larger datasets and used regex patterns to greatly improve on parsing text data into entity name, address, and country of origin.
- Improved upon existing architecture that used HuggingFace transformer models to create a dashboard user interface into which the client could feed text data and relevant questions, which increased efficiency exponentially as the amount of text being processed grew linearly.

Spatial Front Inc., *Data Scientist*

February 2020 – June 2021

U.S. Census Bureau, Economy-Wide Statistics Division (Contract)

- Developed a Python application to effectively address over 70% of the million non-useful entries reported with each Economic Census cycle. With the planned integration of this tool into the 2022 Economic Census, we will be able to quantify up to 20% more revenue data than in previous cycles, accounting for billions of dollars in additional revenue.
- Developed a Tableau-esque Python application that will allow Census analysts to pull relevant data and produce general visualizations for the purpose of pinpointing areas of focus. The tool runs as a Flask app and produces visualizations via Plotly Dash.
- Reviewed old Python code written by members of the team that have since left the Census. Translated the Python scripts to more easily understandable pseudocode and generated reports summarizing the programs' functionality, thus enabling current and future team members to better understand and use sections of the code in related projects.
- Worked on several projects in which Supervised Machine Learning was used to develop classification models and interpolation pipelines for matching publicly available data and imputing missing records in Census and IRS databases.

NovellusDX, *Data Science Intern*

Summer 2018

- Designed and implemented Python functions utilizing NumPy and Pandas to compare feature data collected by the image recognition software searching plates of cells for specific, introduced mutations.
- Utilized logistic regression, principal component analysis, and t-SNE analysis, as well as other statistical models, to analyze the relationship between the dozens of features from the cells being analyzed and the algorithm's probability of successfully predicting whether a specified mutation was present.
- Built a Python program to automate the data analysis process by determining the most significant experimental features and their appropriate ranges, improving the algorithm's accuracy while reducing the amount of employee time required to tune it.

Skills

Programming Languages: Python, SQL, Javascript, HTML/CSS, MATLAB, C++

Libraries: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, SpaCy, Scrapy, NLTK, Bokeh, Flask, Transformers, Haystack

Software Tools: QGIS, Linux/ Red Hat, Jenkins, Vim, Elasticsearch, Kibana, NiFi, Git