

# Jonathan Papir

[papir805@gmail.com](mailto:papir805@gmail.com) | [Portfolio](#) | [GitHub](#) | [LinkedIn](#) | San Jose, CA

---

## SUMMARY

Data analyst and former math teacher with 2-4 years of experience using SQL, Excel, Tableau, R, and Python and eight years of communicating complex ideas to technical and non-technical audiences. Skilled in developing stakeholder relationships, organizing and collaborating with others to solve problems with keen attention to detail, and storytelling to present actionable insights. Enrolled in a rigorous master's program in analytics to sharpen existing and learn new analytical and technical skills. Eager to leverage these talents and a unique teaching background in a data scientist role.

---

## EXPERIENCE

### Desert Oasis Healthcare

**Palm Springs, CA**

*Financial Data Analyst*

*February 2025 – Present*

- Gather stakeholder requirements, translate business rules into SQL logic, and collaborate with developers to create stored procedures and build and edit financial and healthcare quality reports in SQL Server Reporting Services (SSRS).
- Conduct ad hoc financial reporting on medical claims data for the CFO and other senior executives by analyzing procedure codes, places of service, vendors, and more to identify high-cost drivers using SQL, Excel, and Python.
- Manage resource utilization, provider performance, and claims reconciliation projects and maintain 12 associated dashboards for the VP and Program Manager of Arizona Priority Care.
- Wrote object-oriented Python script using Pandas and NumPy to automate transforming three regularly distributed Centers for Medicare and Medicaid Services datasets for ingestion into 10 SQL tables.

### Fusion Academy

**Los Gatos, CA**

*Mathematics Teacher*

*September 2016 – August 2024*

- Authored and delivered thousands of presentations tailored to diverse audiences and mentored five colleagues on presentation best practices for achieving maximum impact.
- Designed and implemented initiatives to digitize the math department after COVID forced instruction online. Led monthly department meetings to set goals and weekly check-ins with colleagues to ensure deadlines were met.

---

## PROJECTS

- **Predicting World Happiness** | [[full report](#)]: Compiled data from the International Labour Organization, International Monetary Fund, and other sources to predict Gallup World Poll happiness scores. Compared linear regression with principal component analysis, ridge regression, random forest, and XGBoost models to address multicollinearity, and the best model achieved 94% testing accuracy after tuning hyperparameters via cross-validation.
- **Generative AI Whitepaper Search Tool** | [[Medium article](#)]: Decreased the time required to answer technical product questions after constructing and deploying an automated tool using Streamlit, Python, Langchain, and ChatGPT that searches, identifies, and preprocesses the ten most relevant whitepapers and returns answers along with source materials.
- **Craigslist Web Scraper** | [[full report](#)]: Boosted my math tutoring business' revenue by 25% by building a Python web scraper with Requests, BeautifulSoup, and Psycopg2 to extract ~5k daily Craigslist posts into PostgreSQL, using NumPy and SKlearn to deduplicate posts across sub-regions by calculating TF-IDF scores and cosine similarity, and analyzing competing market rates in Pandas.

---

## EDUCATION

### Georgia Institute of Technology

*Master of Science, Analytics – Computational Data Analytics – GPA: 4.0*

*Expected June 2027*

### San Jose State University

*Bachelor of Science, Business Administration – International Business*

---

## SKILLS

**Languages:** SQL (4 years): PostgreSQL, BigQuery; Python (4 years): Pandas, NumPy, Matplotlib, Scikit-Learn; R (2 years)

**Tools:** Microsoft Excel (4 years); Tableau (2 years); Salesforce; Jupyter Notebook; Git; Google Cloud Platform

**Machine Learning:** Linear and Logistic Regression; K-Nearest Neighbors; Support Vector Machines; Classification and Regression Trees; Random Forests; K-Means Clustering; Large Language Models; Natural Language Processing, XGBoost

**Other:** Exploratory Data Analysis; Data Cleaning; Data Visualization; Reporting; Writing Documentation; Written and Oral Communication; Presentation Skills; Teaching; Mentorship; Statistical Analysis; Hypothesis Testing