ANDREW JOHNSON

Mobile: (214)-457-3338 | Email: andrewjohnson9393@gmail.com | GitHub | LinkedIn | Portfolio

SUMMARY

Data Analyst skilled in SQL, Python, data visualization, and dashboard development, with experience supporting data engineers, analyzing complex datasets, and developing data products that drive business decisions. Strong ability to interpret data, identify trends, create KPIs, and communicate insights to stakeholders. Experienced in building predictive models, optimizing data quality, and maintaining thorough analytical documentation. Passionate about innovation, problem-solving, and delivering high-impact data solutions.

EDUCATION

B.S. in Information Technology – Software Engineering
Middle Georgia State University, Macon, GA
Relevant Courses: System Analysis, & Design, Software Engineering, Operating Systems, Database Principles

SKILLS

- **Programming Languages:** Python, R, Django, SQL, Java (Spring Boot), JavaScript, TypeScript
- **Python Libraries:** NumPy, Pandas, Matplotlib, openpyxl, psycopg2, scikit-learn
- **Database Systems:** PostgreSQL, MongoDB
- **Data Visualization:** Tableau, Power BI
- **Data Analysis Tools:** Advanced Excel
- **Version Control:** Git
- **Cloud & DevOps:** Docker, AWS (S3, Amplify, EC2, Elastic Beanstalk)
- **Other:** API & Webhook Development, Automation, Infrastructure Architecture, SaaS Platforms, Agile (Scrum, Sprint Planning, Code Reviews), Security Best Practices, Jira, Figma
- **AI & Machine Learning:** Applied machine learning concepts, predictive modeling, feature engineering, model evaluation, and data-driven decision support using Python and analytical workflows
- **AI Tools & Automation:** Experience leveraging AI-assisted analytics tools, building data-driven AI workflows, integrating AI capabilities into analytics solutions, and developing intelligent automation for reporting and insights

EXPERIENCE

Full Stack Developer & Software Tester Internship

Next Play Nation | Apr 2025 – Jun 2025 | Web & Mobile Internship App |Internship recommendation letter

• Developed and maintained React micro-frontends integrated with Node.js and Spring Boot APIs in a microservice-style architecture.

• Implemented RESTful and GraphQL APIs supporting data integrations and internal tooling.
• Applied Test Driven Development (TDD) with Jest, React Testing Library, and Supertest to validate features.
• Optimized system performance and resolved technical issues, applying debugging and system architecture troubleshooting skills.
• Participated in code reviews and Agile ceremonies, contributing to a collaborative engineering culture.

PROJECTS

Large-Scale Audit & Error Detection System for Government Operations

- Built an end-to-end data analytics pipeline analyzing 500,000+ simulated government financial records (procurement, payroll, benefits, inspections) using Python, SQL, and R to identify errors, inefficiencies, and financial risk.
- Designed and implemented data generation, cleaning, and validation workflows in Python using pandas and faker, performing duplicate removal, missing value handling, and error flagging to improve data quality.
- Applied machine learning–based anomaly detection (Isolation Forest) to identify suspicious transactions, uncovering 4,900+ high-risk records representing $1.4B+ in potential financial exposure.
- Conducted financial impact analysis to quantify risk severity, calculate total and average suspicious transaction values, and produce executive-ready audit summaries.
- Created two Tableau Public visualizations to communicate insights to non-technical stakeholders:
  - Overall Government Spend (agency, transaction type, approval status)
  - Risk & Anomalies Analysis (high-risk transactions and financial exposure)

Data Warehouse Analytics: Sales Data Warehouse & Dashboards

- Built a mini data warehouse in PostgreSQL with a star schema, including fact and dimension tables, to organize sales data.
- Developed an ETL pipeline in Python to clean, transform, and load sales data into the warehouse.
- Created interactive dashboards in Tableau and Power BI highlighting sales trends, top products, and customer insights.
- Performed feature engineering and a baseline machine learning model using Python and R.
- Utilized libraries and tools: pandas, psycopg2, scikit-learn, R, Tableau, Power BI, PostgreSQL, pgAdmin.

Healthcare Data Analytics: Billing, Visits & Outcomes

- **Cleaned and processed 10,000+ rows of healthcare visit data using Python (Pandas, NumPy, openpyxl)** to remove missing values, standardize fields, and prepare the dataset for analysis.
- **Performed exploratory data analysis (EDA)** in Jupyter Notebook to calculate billing statistics, department visit breakdowns, and patient outcome patterns.
- **Designed a relational SQL table and imported the dataset into PostgreSQL (pgAdmin)** using CREATE TABLE and COPY to run data validation and analytical queries.
- **Developed three interactive Tableau Public visualizations** (billing trends, department visit distribution, and patient outcomes) and published them in a combined dashboard.
- **Version-controlled the entire project with Git/GitHub,** including Python scripts, SQL files, notebooks, and a public Tableau dashboard link.