ANDREW JOHNSON

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

SUMMARY

Data Analyst with strong experience in Python, SQL, R, and data visualization, specializing in building end-to-end analytics solutions for finance, fintech, and government use cases. Proven ability to generate, clean, validate, and analyze large datasets to support fraud detection, risk analysis, audit compliance, and operational decision-making. Skilled in developing KPI-driven dashboards using Tableau, Power BI, and Excel (Power Query, pivot tables), and translating complex data into clear, executive-ready insights. Adept at data modeling, ETL workflows, anomaly detection, and ensuring data quality across analytical pipelines.

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

CERTIFICATIONS

Data Analytics Certification – Introduction to Data Analytics

Simplilearn | 2025

Skills covered: Data Analysis, SQL, Python, Data Visualization, Business Analytics Fundamentals

SKILLS

1. **Programming Languages:** Python, R, SQL, Java (Spring Boot), JavaScript, TypeScript, Django
2. **Python Libraries:** Pandas, NumPy, Matplotlib, openpyxl, psycopg2, scikit-learn
3. **Database Systems & Warehousing:** PostgreSQL, MongoDB, **Snowflake (data warehousing concepts & analytics workloads)**
4. **Data Modeling & ETL:** Star schema design, fact & dimension tables, ETL pipelines, data transformation, data validation
5. **Data Visualization & BI:** Tableau, Power BI
6. **Data Analysis Tools:** Advanced Excel (Power Query, Pivot Tables)
7. **Version Control:** Git
8. **Cloud & DevOps:** AWS (S3, EC2, Amplify, Elastic Beanstalk), Docker
9. **Other:** API & Webhook Development, Automation, SaaS Platforms, Agile (Scrum, Sprint Planning, Jira, Slack, Code Reviews), Security & Data Privacy Best Practices
10. **AI & Machine Learning:** Applied machine learning concepts, anomaly detection, feature engineering, predictive modeling, model evaluation, and data-driven decision support
11. **AI Tools & Automation:** AI-assisted analytics, intelligent reporting workflows, and automated data insights

EXPERIENCE

Full Stack Developer & Software Tester Internship
Next Play Nation | Apr 2025 – Jun 2025 | Web & Mobile Internship App |Internship recommendation letter

1. Developed and maintained React micro-frontends integrated with Node.js and Spring Boot APIs in a
2. microservice-style architecture.
3. Implemented RESTful and GraphQL APIs supporting data integrations and internal tooling.
4. Applied Test Driven Development (TDD) with Jest, React Testing Library, and Supertest to validate
5. features.

6. Optimized system performance and resolved technical issues, applying debugging and system
7. architecture troubleshooting skills.
8. Participated in code reviews and Agile ceremonies, contributing to a collaborative engineering culture.

PROJECTS

Large-Scale Audit & Error Detection System for Government Operations

1. 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.
2. 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.
3. 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.
4. Conducted financial impact analysis to quantify risk severity, calculate total and average suspicious transaction values, and produce executive-ready audit summaries.
5. Created two Tableau Public visualizations to communicate insights to non-technical stakeholders:
6. Overall Government Spend (agency, transaction type, approval status)
7. Risk & Anomalies Analysis (high-risk transactions and financial exposure)

Finance & Fintech Data Analytics

1. Built an end-to-end finance analytics project by generating, cleaning, and analyzing customer, account, transaction, and loan data using Python (pandas, NumPy), SQL schemas, and R (tidyverse) to simulate real financial firm operations.
2. Designed automated data pipelines to export analysis-ready datasets to CSV and Excel (.xlsx) formats, enabling seamless use in Excel, Power BI, and Tableau for executive-level reporting and visualization.
3. Analyzed key business metrics including fraud rate detection, customer balance aggregation, and loan risk exposure, directly supporting day-to-day operations and strategic financial decision-making.
4. Prepared recruiter-ready dashboards and visuals by structuring data for pivot tables, Power Query transformations, and Tableau dashboards, demonstrating strong data cleaning, validation, and visualization best practices across tools.

Healthcare Data Analytics: Billing, Visits & Outcomes

1. 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.
2. Performed exploratory data analysis (EDA) in Jupyter Notebook to calculate billing statistics, department visit breakdowns, and patient outcome patterns.
3. Designed a relational SQL table and imported the dataset into PostgreSQL (pgAdmin) using CREATE TABLE and COPY to run data validation and analytical queries.
4. Developed three interactive Tableau Public visualizations (billing trends, department visit distribution, and patient outcomes) and published them in a combined dashboard.
5. Version-controlled the entire project with Git/GitHub, including Python scripts, SQL files, notebooks, and a public Tableau dashboard link.

Data Warehouse Analytics: Sales Data Warehouse & Dashboards

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