CHARLES GEORGE – DATA SCIENTIST | DATA ENGINEER

703-302-0185 | cgdesigned@gmail.com | in/charlesgcv | github.com/87designer | mydatavault.streamlit.app

SUMMARY

Innovative data professional with 10+ years of experience specializing in Python, SQL, and ETL development for data engineering and analytics. Expert in building scalable data pipelines, automating workflows, and transforming complex datasets into actionable insights. Strong background in data integration, process automation, and analytics, with growing experience in cloud platforms. Demonstrated ability to problem-solve and learn new technologies quickly. Excited to continue expanding expertise in cloud-based data solutions while driving impactful data-driven strategies.

TECHNICAL SKILLS

Statistics: Hypothesis testing, linear regression, logistic regression, Bayesian statistics
Data Cleaning and Processing: Imputation, normalization, transformation, PDF Scraping, Web Scraping
Machine Learning: Feature selection, feature engineering, regression, classification, supervised learning, Natural Language Processing, time-series analysis, algorithms
Programming and Tools: SQL, PostgreSQL, Oracle, NoSQL, Python, Pandas, NumPy, Scikit-Learn, Statsmodels, SciPy, LangChain, XLSXWriter, Selenium, HTML, CSS, Jupyter, Posit, Azure Data Studio
Cloud and DevOps: AWS, Azure, GCP, DataBricks, CI/CD, FastAPI, Power Automate
Visualization and BI: Plotly, Dash, Streamlit, Looker, Power BI, Qlik Sense, Visio, Excel
Project Management: Agile, Scrum, Jira, Azure DevOps, Confluence, SharePoint, GitHub

PROFESSIONAL EXPERIENCE

National Heart, Lung, and Blood Institute (Office of Planning Analytics and Evaluation) — Remote
Senior Management Analyst | Data Science & Automation — 05/2024 – 03/2025

- Engineered machine learning models (Linear Regression, Random Forest) to automate workforce analytics and budget forecasting, enhancing predictive accuracy and decision-making.
- Developed & deployed interactive data applications using Dash and Streamlit, enabling near real-time workforce planning and financial insights.
- Automated ETL pipelines, eliminating 50% of manual processing and enhancing data reliability.
- Directed GitHub repository management and Python 3.12 migration, ensuring zero downtime and improved code versioning.
- Designed & implemented a geospatial analysis tool, cutting processing time from 30 hours to under 15 minutes for workforce distribution optimization.
- Enhanced workforce analysis accuracy by improving position categorization logic, increasing data reliability by 15% for strategic workforce planning.
- Developed and presented key analytical tools and workforce insights at OPAE monthly meetings, informing leadership decision-making and driving strategic planning improvements.
- Authored Standard Operating Procedures (SOPs) for FTE Usage Analysis, streamlining workforce planning and reducing cross-team communication delays by 20%.

ECS Federal (Analytics, Automation & Artificial Intelligence) — *Remote*
Data Scientist – Manager — 10/2023 - 05/2024
Data Scientist – Senior Consultant — 03/2022 - 10/2023
Data Scientist – Consultant II — 07/2021 - 03/2022
Data Scientist – Consultant — 03/2020 - 07/2021

- Led and mentored a team of three data science professionals, overseeing workforce analytics, financial planning, and system development for $1.7M+ in contract revenue.
- Developed and implemented strategic tools for capturing, migrating, and cleansing large datasets, improving efficiency and productivity in workforce analytics and financial planning.
- Implemented Git version control and code formatting best practices to streamline automation workflows.
- Led a Clinical Trial NLP Prediction Project, automating grant reviews and cutting manual workloads by 90%.
- Researched and optimized the Program Officer Workload analysis by reconciling conflicting data sources, reducing processing time from 6–8 weeks to under 1 week using Python and data science algorithms.

- Built a Python-powered Dash application to monitor fiscal trends & prevent budget overruns, supporting high-profile research funding.
- Designed custom data visualizations, including multi-trace plots, threshold overlays, and Sankey diagrams, improving leadership insights.
- Collaborated with IT staff to assess database requirements, successfully provisioning new databases.
- Researched and evaluated Azure ML Studio for cloud-based AI workflows, effectively communicating technical needs and findings, ensuring alignment between data infrastructure and machine learning capabilities.
- Drove technical contributions for a $6.5M contract re-compete, contributing to capture planning, technical writing, and business development for analytics and clinical informatics growth opportunities.
- Managed senior-level client relationships, driving $413K in contract growth and expanding project scope to enhance contractor team support.
- Oversaw program staffing and subcontractor relationships, supporting eight subcontractors and 30+ FTEs, ensuring effective project execution and IT system development.
- Strengthened Data Science Community of Practice, mentoring junior staff and fostering a culture of knowledge-sharing through brown bags, training sessions, and career coaching.
- Championed workforce engagement initiatives, developing onboarding materials, process improvements, and mentorship programs, while leading bi-weekly 1:1 meetings to provide direct reports with career coaching and structured feedback.
- Served as a thought leader in data science, advising client leadership on AI, software application best practices, and emerging technologies, driving innovation in government analytics programs.

American Media Source   Hybrid
Data Management & Analysis Specialist   10/2019 - 03/2020
- Automated data migration to Google Cloud Platform, reducing transfer time by 91%.
- Developed Python scripts and interactive dashboards, boosting performance monitoring efficiency by 66%.
- Managed relational databases and automated ETL workflows, streamlining legacy system migration.
- Led team training in Jupyter Notebooks, enhancing data analysis productivity.

American Target Advertising   Manassas, VA
Data Analyst   08/2018 - 10/2019
- Implemented Python-based data automation, reducing manual efforts by 40%.
- Designed & developed a workflow automation system, reducing project request response times by 75%.
- Built API-driven solutions, enhancing customer data segmentation & analytics.
- Led SQL-based segmentation modeling, optimizing campaign performance tracking.

FFC Properties LLC   Woodbridge, VA
Property Management Analyst   09/2012 - 08/2018
- Designed & developed a property database, streamlining renovation status tracking.
- Created a data collection form, improving data accuracy by 25%.

PROJECTS

Resume Reactor (AI-Powered Resume Optimization) - http://bit.ly/3FycBla
Engineered an LLM-integrated web application with CI/CD pipelines, assisting federal employees in transitioning to the private sector using prompt engineering and API integrations.
My Data Vault (AI-Driven Interactive Tech Portfolio) - https://bit.ly/3FxQz8s
Developing a full-stack portfolio application showcasing expertise in data science, engineering, and visualization, integrating LangChain-powered AI querying and Plotly dashboards.
DS Leadership Kit (Ongoing Repo) - https://bit.ly/4hFYQVc
A comprehensive toolkit for effective leadership in the realm of Data Science. Empower your data science team with essential resources, best practices, and collaborative tools.
School Days Calculator (Simple Python Web App) - https://bit.ly/4268j3r
A simple yet powerful homeschool calendar calculator that helps parents and educators efficiently plan and manage their school year.
Python AI Dente - A Programmer's Guide to Perfect Code (Programming Book):
Currently authoring a hands-on Python book that simplifies complex coding concepts through real-world, interactive examples.

EDUCATION

Thinkful   Online
Certificate, Data Science.   06/2019 - 01/2020
- Gained practical mastery of Python, Pandas, PostgreSQL, and industry best practices under the guidance of a senior data scientist mentor.
- Developed and presented projects involving exploratory data analysis, data cleaning, machine learning, statistical modeling, study design, and statistics/probability.

Pensacola Christian College   Pensacola, FL
B.S., Communications.   09/2006 - 06/2010

Additional Courses
Microsoft Azure AI Fundamentals – AI-900
Microsoft Power BI Data Analyst – Course PL-300T00-A
Qlik Development 102

CERTIFICATIONS

IBM Data Engineering Professional Certificate (In progress)
Databricks Fundamentals (Expires 2026)
AWS Certified Cloud Practitioner (Previously Certified, Expired 2024)
Certified ScrumMaster (CSM) (Previously Certified, Expired 2024)
Python Pro Intensive – codecademy (Issued 2019)
Web Development Pro Intensive – codecademy (Issued 2018)