JUSTIN GROVES, PhD

jagroves95@gmail.com | (864)634-1029 | JustinGroves.com | LinkedIn | GitHub

Senior Data Scientist | Machine Learning & Applied AI | Data Engineering | End-to-End Analytics

PROFESSIONAL EXPERIENCE

Senior Data Scientist   ReefPoint Group (Remote)
`PyTorch, GitHub, Power Platform/Power BI, Foundry`   Aug 2023 - Present

- **Delivered a cloud-integrated database system and built Power BI reports in 6 months** tracking all VA form templates**;** collaborated with government stakeholders to define KPIs and enable downstream analytics.
- **Engineered a pilot data pipeline, securing a $900K client contract (potential extensions totaling $2M)**, developing a VA file inventory system using Selenium, BeautifulSoup, and NLP**.**
- **Deployed end-to-end solution to identify and correct 200+ business model inconsistencies** using Palantir Foundry, eliminating a manual review process and freeing human resources.
- **Reduced manual review time cost by 83% through automation,** deploying a classification algorithm leveraging Optical Character Recognition (OCR), NLP, and regular expressions (RegEx) analyzing 1,400+ documents.
- **Fostered data science team growth** by interviewing candidates, authoring technical articles and presentations, and mentoring team members through personalized Python package development.

Retirement Analyst Intern   WTW (Remote)
`Excel, Radar (Proprietary)`   May 2022 – Aug 2022

- **Calculated actuarial liabilities** for employee retirement plans, leveraging advanced Excel functions and proprietary software, ensuring accurate financial forecasting for client portfolios.
- **Developed and delivered a strategic actuarial plan,** presenting a comprehensive fiscal year report to senior leadership, facilitating data-driven decision-making for future budgeting and resource allocation.

Graduate Teaching Associate   University of Tennessee (Knoxville, TN)
`Python, R, Excel, SageMath`   Aug 2017 – Aug 2023

- **Developed Python-based data analysis modules** using pandas and NumPy to clean, validate, and analyze large datasets implementing algorithmic approaches to complex problems.
- **Presented complex technical concepts** to diverse audiences through multiple talks, translating advanced "mathematical and data-driven ideas into clear, actionable insights.
- **Led instruction and performance tracking for 500+ individuals**, using Excel-based reporting and analysis to monitor outcomes, predict performance, and implement improvements.

PROJECTS

Scalable ETL and AI-Powered Speech Transcription Pipeline
`AWS, PostgreSQL, Docker, FastAPI`

- **Co-designed and implemented a Python-based ETL pipeline** to scrape web content, process media files, and populate a normalized PostgreSQL (Supabase) database via SQLAlchemy for structured reporting.
- **Integrated AI-powered transcription (OpenAI Whisper) and orchestrated AWS EC2 compute,** enabling large-scale NLP workflows and processing 100 hours of audio content per hour.

EDUCATION

PhD in Mathematics (Minor: Data Science & Statistics)   University of Tennessee, Knoxville

CLEARANCES AND CERTIFICATIONS

VA Public Trust Security Clearance   2023 – Present

Data Engineer, Associate – Palantir Foundry   2024