

# Cole Campbell

+14259028188 | colejcampbell1738@gmail.com  
GitHub | LinkedIn

## SUMMARY

New-grad data scientist and ML engineer with a deep applied portfolio spanning generative AI, deep learning, and statistical modeling. Builds end-to-end systems from data pipelines through model integration in production backends. Pursuing an M.S. in Data Science atop a 2026 B.S. from Arizona State University, seeking an entry-level role in tech, engineering, and data/AI.

## EDUCATION

### Arizona State University

B.S. in Data Science (Concentration in Computer Science), 2026  
GPA: 3.52

### Arizona State University

M.S. in Data Science, Analytics, and Engineering (Computational Mathematics & Data), 2027

## TECHNICAL SKILLS

<b>Languages:</b>	Python, R, Dart, C/C++, JavaScript (Node.js), SQL, Bash, MATLAB
<b>ML / AI:</b>	Deep learning (CNNs, RNNs/LSTMs, Transformers), generative models (Diffusion, GANs, StyleGAN), LLMs & prompt engineering, LoRA/QLoRA/PEFT, transfer learning, neural style transfer, audio/music generation, model evaluation (FID, CLIP, LPIPS, BLEU, perplexity)
<b>Frameworks:</b>	PyTorch (torchvision/torchaudio), TensorFlow, Hugging Face (Transformers, Diffusers, Accelerate), scikit-learn, NumPy, Pandas, SciPy, StatsModels, FastAPI, Flutter/Riverpod
<b>Data Science:</b>	Regression & statistical modeling, classification, time series, Bayesian, clustering, longitudinal/panel analysis, feature engineering, A/B testing, causal inference, hypothesis testing, visualization (Matplotlib, Seaborn, Plotly)
<b>Data Engineering:</b>	ETL & data pipelines, REST API integration, SQLAlchemy + Alembic, PostgreSQL, Firebase, Docker, CUDA, cloud (AWS, GCP, Azure ML)
<b>Tools:</b>	Git, Linux, Jupyter, VS Code, OpenCV, librosa, ParaView/VTK, pytest, Weights & Biases, Kaggle/NBA APIs, REST/OpenAPI
<b>Math:</b>	Linear algebra, probability, optimization, numerical methods, stochastic processes, statistical theory

## PROJECTS

### Multi-Architecture Generative Content Studio — Deep Learning Capstone

*PyTorch • Transformers • Diffusion • GANs • CNNs • RNNs*

- Built a unified generative system producing story text, scene illustrations, character portraits, and background music from a single prompt.
- Integrated GPT-2, Stable Diffusion, StyleGAN2/3, VGG16/19, and LSTM into one cohesive pipeline.
- Fine-tuned GPT-2 for narrative generation, entity extraction, and structured scene descriptions; implemented Stable Diffusion with LoRA + ControlNet for composition-controlled scenes.
- Trained StyleGAN2 for character portraits, applying GAN inversion and latent editing for cross-scene consistency; used CNN neural style transfer to unify visual aesthetics.
- Built an LSTM music generator trained on 10k+ MIDI files for mood-aligned audio.
- Evaluated outputs with FID, CLIP score, perplexity, and lexical-diversity metrics (0.97 unique-word ratio on generated narratives); benchmarked DCGAN against Stable Diffusion and measured ~390x higher image sharpness (1306 vs 3.3) for Diffusion, driving the pipeline's generator choice.

### Tempo — AI Scheduling Assistant

*FastAPI • Flutter • PostgreSQL • LLM / NLP • Constraint Optimization*

- Built a cross-platform scheduling app (Flutter front end, FastAPI back end) that generates alternative day/week plans, compares them, and applies the one the user picks.
- Designed a constraint-based scheduling engine that orders tasks, builds the daily timeline, resolves conflicts, and merges blocks into a coherent schedule.
- Engineered a scenario pipeline supporting generation, diff-based comparison, trade-off analysis, and undo/redo state management.
- Added an "Ask Tempo" conversational layer for natural-language task parsing and schedule interaction.
- Backed by SQLAlchemy + Alembic schema management and Firebase auth; hardened with 360+ backend tests at ~84% coverage.

### PSID Panel Dataset Analysis — Longitudinal Economics Project

*Python • Statistical Modeling • Longitudinal Analysis*

- Analyzed a 17,000-observation PSID panel (1999–2023) to study income and wealth dynamics.
- Built the full pipeline: cleaning, transformation, feature engineering, regression modeling, and visualization.
- Found wealth roughly twice as unequal as income (wealth Gini ~0.85 vs income Gini ~0.45), with the median wealth-to-income ratio rising from 1.25 to 2.05 across the panel.

### Scientific Visualization Pipeline (ParaView) — Self-Directed

*ParaView • paraview.simple • VTK • NumPy*

- Built a NumPy/VTK dataset (3D Gaussian concentration field + analytic ABC-flow velocity) and a fully Python-scripted ParaView pipeline (paraview.simple).
- Generated volume renderings, Contour isosurfaces, slices, and Stream Tracer streamline tubes, with automated figure and orbit-animation export.

**Portfolio & demos:** [github.com/cjcamp16](https://github.com/cjcamp16)

## EXPERIENCE

### Handshake

*AI Trainer — May 2026 – Present*

- Trained AI models based on their ability to reason and evaluate tasks across mathematics, codebases, and visualizations.