JULIAN MERCER

PYTHON DATA SCIENTIST

San Francisco, CA | (415) 555-0148 | julian.mercer@email.com | linkedin.com/in/julianmercer

PROFESSIONAL SUMMARY

Python Data Scientist with 7+ years building forecasting, ranking, and risk models in pandas, scikit-learn, and PyTorch. Ships tested, packaged Python code that moves from notebook to production service and holds up under monitoring.

CORE SKILLS

Python | pandas | NumPy | Polars | scikit-learn | XGBoost | LightGBM | PyTorch | statsmodels | SciPy | pytest | FastAPI | Jupyter | Airflow | Docker | SQL

PROFESSIONAL EXPERIENCE

Senior Data Scientist Aug 2023 - Present
Instacart, San Francisco, CA

- Built a delivery-time model in Python with LightGBM and scikit-learn that cut mean absolute error by 11% across 5,000 metro zones.
- Packaged feature code as a versioned internal library with pytest coverage above 85%, so four teams reuse one tested set of transforms.
- Moved batch scoring jobs from ad hoc notebooks to Airflow tasks with typed Python modules, shrinking the weekly run from six hours to under two.
- Mentor three junior scientists on code review, profiling, and vectorized pandas, which removed several slow row-by-row loops from shared jobs.
- Reviewed 60 pull requests from peers in a year, adding linting, type hints, and a short style guide to the shared data science repository.

Data Scientist Jul 2021 - Jul 2023

- Wrote a Python search-ranking experiment toolkit using statsmodels and SciPy, shortening readouts for product partners from five days to two.
- Trained a substitution recommender in PyTorch that raised accepted replacements by 4% in a controlled test of two million orders.
- Profiled memory-heavy pandas pipelines and rewrote them with Polars and chunked reads, which cut cloud compute spend on the job by about a third.

Data Scientist Jun 2019 - Jun 2021
Capital One, McLean, VA

- Developed a credit-line increase model in Python with XGBoost, reviewed by model risk, and lifted approved-account profitability in the pilot segment.
- Automated monthly drift reports with Jupyter, papermill, and matplotlib, replacing a manual process that took an analyst two days each cycle.
- Containerized three scoring services with Docker and FastAPI, giving engineers stable endpoints with documented request and response schemas.
- Replaced a copy-pasted feature script with a pip-installable package and release notes, so model retraining no longer broke when a column changed.
- Presented model findings to business leads and wrote plain-language notes on assumptions, limits, and the data behind every decision.

EDUCATION

M.S. in Computational Data Science, Carnegie Mellon University, Pittsburgh, PA

CERTIFICATIONS

Google Cloud Professional Machine Learning Engineer | Python Institute PCAP: Certified Associate in Python Programming