

# Yashar Mansouri

## Senior Data Engineer

---

- Portfolio: [yasharmansouri.github.io](https://yasharmansouri.github.io)
- Phone: (202) 602-8464
- Email: [yasharmansouri99@gmail.com](mailto:yasharmansouri99@gmail.com)
- LinkedIn: <https://www.linkedin.com/in/yasharmansouri>
- Certifications: <https://www.linkedin.com/in/yasharmansouri/details/certifications/>
- Book a meeting with me: [calendly.com/yasharmansouri/30min](https://calendly.com/yasharmansouri/30min)

## PROFESSIONAL SUMMARY

---

Senior Data Engineer with 6+ years of experience building high-performance scalable data products that collect, process, and govern high-volume data for analytics and AI purposes. Background in AI/ML workflows and expert in data warehousing and administration, dimensional data modeling, performance optimization using distributed processing and predictive analytic frameworks. Deep proficiency in Python and SQL with a strong software engineering mindset, and direct experience with batch and streaming patterns including CDC, incremental loads, and event-driven architectures. Quick to pick up new technologies and able to thrive in fast-paced environments. Proven ability to drive complex cross-functional projects autonomously, mentor junior engineers, and partner effectively with non-technical stakeholders to deliver reliable, actionable data products across healthcare, technology, finance, and public sectors.

## PROFESSIONAL EXPERIENCE

---

### Data Engineer (Contract) | DC Public Charter School Board

Washington, DC | February 2022 – September 2025

- Served as sole data architect designing 70+ production data models for school performance evaluation, compliance reporting, and financial analytics using dimensional modeling and data governance best practices; collaborated with data analysts and stakeholders to gather requirements and iterate on designs
- Architected and maintained 75+ active Apache Airflow DAGs processing 10M+ daily records across student attendance, outcomes, assessments, enrollment projections, performance metric calculations and compliance reporting with 99%+ uptime serving 130+ charter schools
- Administered PostgreSQL databases with role-based access control and automated permission management; optimized queries and resolved performance bottlenecks improving database throughput by 50-80%
- Resolved production and technical debt issues through code refactoring, error handling, and performance tuning, reducing incident response time by 60%
- Developed automated file processing infrastructure handling 50+ data types including student records and financial reporting through multi-layered validation frameworks and data quality constraints
- Managed enterprise BI external user facing platforms (Looker, QuickBase, Tableau Cloud) serving 200+ users; migrated 12+ financial dashboards from Looker to Tableau reducing licensing costs by 70%
- Mentored 4+ junior engineers through code reviews and technical guidance; authored architecture documentation and standard operating procedures for organization handbook

### AI & Data Engineering Consultant | Deloitte

McLean, VA | December 2021 – April 2023

- Mayo Clinic: Designed ML pipeline POCs on GCP using Dataflow, BigQuery, Vertex AI for data science teams; created technical documentation and training materials on deployment best practices
- Meta: Built privacy incident monitoring dashboards and automated biweekly scorecard reports using Python and Presto SQL; onboarded junior data engineers through shadowing and knowledge transfer documentation

- Apple: Developed automated financial reporting pipelines for Wallet, Payments & Commerce team using Teradata and Trino SQL; resolved 10+ critical data quality issues during Teradata-to-Trino migration
- Deloitte: Engineered Python-based synthetic data generation extensions for NVIDIA Omniverse Isaac Sim, enabling randomized 3D robotics scenarios for ML model training

## Data Engineer | DC Public Charter School Board

Washington, DC | May 2021 – December 2021

- Developed automated Airflow DAGs for batch synchronization, ELT workflows, and infrastructure maintenance; managed AWS infrastructure (RDS, EC2, S3) with 99.9% uptime
- Designed transactional and analytical PostgreSQL data models for different departments; developed QuickBase and Looker dashboards enabling self-service analytics for schools and finance teams
- Led audit initiatives consolidating 15+ Looker repos into main codebase, deprecating legacy modules and reducing maintenance overhead
- Established unified data integrity and logging frameworks across several pipelines automating quality checks, error handling, and user notifications

## Data Engineer (Intern) | DC Public Charter School Board

Washington, DC | March 2021 – May 2021

- Contributed to PostgreSQL schema design and automated validation pipelines for 20+ DAGs; improved data model generation using SQLAlchemy and Jinja
- Supported database provisioning, troubleshooting pipeline failures, and legacy system migrations

## Data Scientist (Intern) | Allwyn Corporation

Herndon, VA | January 2020 – October 2020

- Developed AI/ML models predicting hospital readmission rates with 82% accuracy using; engineered ETL scripts processing 50M+ medical records and generating predictive features
- Optimized data pipelines with PySpark reducing processing time by 60%; authored 3 technical white papers on healthcare analytics and mentored 10+ graduate capstone students with AI workflows

## TECHNICAL SKILLS

---

- **Programming / Scripting:** SQL, Python, Bash, YAML
- **Data Engineering & Orchestration:** Airflow, dbt, Spark, Kafka, ETL/ELT, Dimensional Data Modeling, Batch / Stream Processing, Data Quality Frameworks, Data Governance
- **Databases & Data Warehouses:** PostgreSQL, DuckDB, Snowflake, Databricks
- **BI & Low-Code:** Tableau, Looker, Microsoft Fabric, QuickBase, Alteryx
- **Infrastructure & DevOps:** Docker, Kubernetes, Terraform, Infrastructure as Code (IaC), Git
- **AWS:** S3, EC2, RDS, Lambda, Glue, Athena, Step Functions, EMR, Kinesis
- **GCP:** Cloud Storage, BigQuery, Compute Engine, Cloud SQL, Dataflow, Cloud Composer, Vertex AI
- **Languages:** English, Persian, Turkish

## EDUCATION

---

- **Master of Science in Data Science** | American University, Washington, DC | January 2019 - December 2021
- **Master of Science in Biotechnology** | American University, Washington, DC | August 2015 - December 2017
- **Data Science Bootcamp Fellow** | Flatiron School, Washington, DC | August 2019 - December 2019
- **Bachelor of Science in Molecular and Cellular Biology, Genetics** | Azad University, Tehran, Iran | February 2012