ATHARVA YEOLA

📞 +1 8582203508 | ✉ ayeola@ucsd.edu | 🔗 in/atharvayeola | ⬡ /atharvayeola 🌐 /atharvayeola

SUMMARY

Product-focused Data Scientist with experience building recommender systems, predictive models, and causal frameworks that guide product strategy, user retention, and catalog optimization. Proven track record in experimentation, scalable dashboards, and real-time analytics for content & behavioral products.

EDUCATION

University of California San Diego (UCSD)	**Sep 2023 – Jun 2025**
Master of Science (M.S) – Machine Learning & Data Science	*San Diego, USA*
Pune Institute of Computer Technology (PICT)	**Aug 2019 – Jun 2023**
Bachelor of Engineering (B.E) – Electronics & Telecommunication	*Pune, India*

WORK EXPERIENCE

Data Scientist, Analytics — **Jul 2024 – Present**
Qualcomm Institute - UC San Diego Division of Calit2 — *San Diego, USA*
- Headed the causal analysis of impact of the U.S. sports betting on gambling addiction via **synthetic control** and **time series forecasting**; quantifying a **23% nationwide increase** in addiction.
- Created HIVTrends.org, **a live dashboard** that tracks weekly uptake of at-home HIV test kits; adopted by internal product and public health partners, **cutting ad-hoc data pulls by** ∼**50%** and steering campaign-budget decisions.

Data Scientist, Product — **Aug 2022 – Apr 2023**
RhythmFlows Solutions Pvt. Ltd — *Pune, India*
- Devised a **personalized context-aware music recommender** for a mental-health app, curated 75 seed tracks collaborating with psychiatrists, mitigating cold-start; **boosted NDCG@10 by 18%** and **hit-rate@5 by 14%**.
- Led a two-week **A/B test** across 20K patient sessions; context-aware model **increased CTR by 33%** and **average session by 20%** over the collaborative filtering approach, directly informing playlist design and user segmentation strategy.
- **Productionized** the model on **AWS Sagemaker** with continuous training and rollout, ensuring seamless scalability.

Data Scientist — **Jan 2022 – Jun 2022**
Atomic Loops — *Pune, India*
- Spearheaded a **cross-functional team** to launch a real-time computer vision pipeline (YOLOv5 + Raspberry Pi) that surfaced conveyor KPIs, hit **98%** detection accuracy, and **lifted customer footfall by 30%**.
- Partnered with Hardware, Ops, and Product teams to sync belt speed with model latency and ship live dashboards, **boosting revenue 12% MoM** and **reducing food waste by 5x**.

PROJECTS

StorePulse | 📎 *Python, Pandas, PySpark, SQLite, Dash*
- Developed a medallion-architected ETL pipeline (Bronze→Silver→Gold) on 10k+ Superstore rows, automating daily parquet and SQLite table refreshes.
- Modeled a star schema in SQLite (`dim_date`, `dim_category`, `dim_region`, `fact_sales`), powering an interactive Plotly Dash dashboard of sales trends, margins, and top customers.

CommerceFlow | 📎 *PySpark, SQL, Plotly Dash, Cohort Analysis, RFM Segmentation*
- Analyzed 2M+ user events to map the end-to-end funnel (view → cart → purchase); uncovered a 96% checkout drop-off and presented actionable insights to inform conversion optimization strategy.
- Delivered a self-serve dashboard with cohort retention, pricing-tier adoption, and 6-month sales forecasting to support product planning and campaign performance reviews.

RetailLens | 📎 *YOLOv8, SuperGlue, OpenCV, Docker*
- Developed a real-time computer vision system for retail shelves using YOLOv8, achieving 90%+ precision in detecting and localizing on-shelf products.
- Enhanced object matching accuracy with SuperGlue-based feature correspondence, enabling fine-grained product identification across varied layouts.
- Packaged the system with Docker for seamless cross-platform deployment and reproducibility in edge environments.

TECHNICAL SKILLS

Programming – **SQL**, **Python**, R, Pandas, NumPy, PySpark, Scikit-learn, dplyr
Data Management – MySQL/PostGreSQL, **Excel**, ETL, BigQuery, Redshift, Adobe/ Google Analytics
Visualization Tools – **Tableau**, Power BI, DBT, Plotly, Matplotlib, Seaborn, Metabase
Cloud & Deployment – **AWS**, Azure, **Docker**, CI/CD, GitHub Actions
Other – **A/B Testing & Experimental design**, Probability, **Statistics**, Cohort Analysis, Forecasting