Divyam Kataria

602-920-3945 divyam@asu.edu linkedin/divyamkataria github/katsdivi divyamkataria.me

Education

Arizona State University **Expected May 2027**

B.S. Computer Science, Minor in Mathematics ***(GPA: 3.88)*** *Tempe, AZ*

- **Coursework:** Data Structures, Algorithms Analysis, Database Management, Systems Programming, Software Methodology, Artificial Intelligence, Internet Technology, Computer Architecture

Experience

Founder & Quantitative Researcher **May 2026 – Present**

AlternativeIndi *Faridabad, India*

- Built a point-in-time, append-only, entity-resolved panel for US-exposed Indian pharmaceutical equities, stamping every row with the date it became publicly knowable so any historical state reconstructs free of lookahead bias; resolved 89 filer-name variants to tickers and mapped 25 FEI-keyed facilities, 26 regulatory events, and 103 US consignees.
- Designed and ran three pre-registered statistical probes of the shipment-volume-to-revenue thesis using block-permutation significance testing and committed stopping rules; all three returned null, and closed the thesis rather than tuning toward a result.
- Surfaced a data-integrity failure in a widely used market data source reporting zero promoter pledge on nine of ten tested names where NSE filings show real pledge, and rebuilt the ingestion layer onto direct XBRL filings under SEBI LODR and Prohibition of Insider Trading regulations.

Quantitative Research, Contract **Jul 2026 – Present**

Scriptonomy, Fintech Market-Data Research *Remote*

- Built a multi-agent research pipeline in LangChain (coordinator, strategy engineer, research critic) iterating momentum strategies across nine sector ETFs, with LangSmith tracing, subprocess isolation, and separate development, validation, and holdout splits.
- Pre-committed the model selection rule before any parameter sweep ran; the gate rejected the strongest challenger on a 0.0109 validation Sharpe shortfall despite 39% higher development Sharpe and 42% lower turnover, and a 12-backtest ablation showed the candidate volume filter cost 0.099 Sharpe at 4.4x the turnover.

Summer Intern **Jun 2025 – Aug 2025**

Ernst & Young (EY) *Delhi, India*

- Engineered a GenAI pipeline to extract and analyze 1,000+ pages of policy documentation, cutting manual review prep time by an estimated 30%, and systematized a multi-bank sourcing analysis into a repeatable playbook.

Projects

StaleLine | *Cross-venue prediction market research* | *Gator Quant Hacks 2026, team of 4* | *Python* **Oct 2026**

- Recorded the data for the team's pre-registered test of which of Kalshi, polymarket.com, and Polymarket US reprices football markets first (book-midpoint cross-correlation, unrelated-game placebos, sealed holdout): across 88 games the median lag was 0.0 s ($p = 0.709$), and an apparent Kalshi lead traced to a 30 s cached quote feed.
- Isolated three ways a backtest fakes an edge (thin trade prints, cached quotes, and leg-conditional arbitrage that flips from +1.65 to -1.57 cents per contract); both holdout strategies lost (combined paper book Sharpe -7.78), and a passive market-making rule is frozen for an out-of-sample test.

EZParlays | *Prediction market microstructure research* | *Kalshi NFL and college football* **Mar 2025 – Present**

- Tested 141 candidate features across 5,164 NFL games and found none survived conditioning on the market price; decomposed maker execution into +0.507pp spread capture, -0.374pp fees, and -0.839pp adverse selection for a net -0.706pp per round trip.

NeuraMesh Protocol | *Peer-to-peer distributed inference and memory protocol* | *Swift, llama.cpp* **Jul 2026**

- Wrote a custom UDP transport with a Noise IK handshake, AES-256-GCM, and XOR forward error correction, benchmarking 2.6x lower latency than TCP+TLS on encrypted inference traffic, and sharded Llama-2-7B layers across Mac, iPhone, and iPad via a llama.cpp shim to run a model exceeding any single device's memory. Validated by a 474-test suite.

Technical Skills

Languages: Python, TypeScript/JavaScript, Swift, Java, C++, SQL (PostgreSQL, MySQL, SQLite), R, Bash
Quantitative Research: Market Microstructure, Execution Cost Analysis, Adverse Selection, Alternative Data, Backtesting, Point-in-Time Data, Entity Resolution, Pre-Registration, Purged & Embargoed CV, Clustered Standard Errors, Look-Ahead & Survivorship Bias, Sharpe & Sortino
AI & Machine Learning: LangChain, LangSmith, Multi-Agent Systems, LLM APIs (OpenAI, Gemini, Mistral, Groq, Ollama), RAG, Embeddings & Vector Search, LightGBM, XGBoost, scikit-learn, pandas, NumPy
Backend, Data & Infrastructure: FastAPI, WebSockets, REST APIs, PostgreSQL, Parquet, XBRL, ETL/DAG Pipelines, Docker, Git, React, React Native, SwiftUI