Shane Thakkar

Frisco, Texas Shane.Thakkar@gmail.com 469-989-3585 in/ShaneThakkar ShaneThakkar.com

SUMMARY

Analytics & AI professional with 2.5 years of hands-on experience in Python, SQL, Tableau, and machine learning across sports, finance, and telecom. Built automations saving 10+ hours per week, delivered ML models achieving 90% accuracy, and developed venue clustering analyses guiding multimillion-dollar LA28 investments. Skilled at transforming complex data into actionable insights that optimize pricing, inventory, operations, and strategy.

EDUCATION

B.S. Business Analytics and Artificial Intelligence – Data Science Track

University of Texas at Dallas • Richardson, TX • May 2026

SKILLS

Programming & Analytics: Python (Pandas, NumPy, SciPy) SQL, R, Git, Bash/UNIX
Machine Learning: Scikit-learn, TensorFlow, Keras, Optuna
Statistical Methods: Time-Series Forecasting, Clustering, Regression Analysis, Economic Modeling
Visualization & BI: Tableau, Power BI, Matplotlib, Plotly
Big Data & Cloud: Databricks, Hadoop, Spark, Hive, Azure, SQL Server, Snowflake
Tools: Excel, Stata, Alteryx, Power Automate

EXPERIENCE

Intern - Business Intelligence

Legends Global July 2025 - December 2025, Frisco, TX

- Analyzed 1M+ POS records across 8+ professional sports venues using Python and Pandas, cleaning and standardizing data to perform category-level performance analysis that identified high and low-performing products and drove inventory, pricing, and seasonal planning decisions.
- Applied Market Basket Analysis and Pareto principles in Python to evaluate FC Dallas concessions profitability, discovering high-frequency product pairings for bundling strategies and identifying 30% of SKUs as tail products generating <15% of revenue.
- Consolidated 2 years of POS data from 3 systems for Del Mar Racetrack, building reports and SWOT analyses that identified key revenue drivers and performance trends, presenting findings to stakeholders to inform operational strategy.
- Modeled F&B performance for ~15 LA28 venues in Python by normalizing historical per-cap and attendance data, then clustering venues into priority tiers that informed executive investment and risk decisions for LA28 operators.
- Automated weekly SQL Server reconciliation between game-day flash reports and POS transaction data across venues, flagging discrepancies and data gaps in audit logs and eliminating ~10 hours per week of manual validation work.
- Pioneered transparent reporting practices by adding methodology sections to all analyses—documenting data sources, assumptions, and formulas to enable stakeholder validation and error detection—later adopted as the team standard.

Intern - Project Management and Data Analytics

Supreme Lending June 2024 - November 2024, Dallas, TX

- Built Excel-based project management system coordinating 100+ test cases across departments during loan origination software upgrade, tracking completion KPIs through automated dashboards that enabled defect-free production rollout.
- Built Power Automate workflow to automate Azure DevOps user story creation, reducing manual request intake time and enabling automated stakeholder tracking and notifications.
- Led team of 14 interns in 30-minute executive pitch proposing university mortgage lending course, serving as lead speaker and presenting survey analysis of 300+ students to validate market demand and curriculum design.

Intern - Business Analyst

TruGen May 2023 - May 2024, Frisco, TX

- Cleaned and structured wireless network performance data across 10+ regions using SQL, preparing datasets for machine learning models that achieved 90% classification accuracy in signal quality prediction.

RELEVANT COURSEWORK

Advanced Applied AI & Machine Learning - Neural Networks, NLP, Deep Learning, LLMs Applications

Applied Econometrics & Time-Series Analysis - Business Forecasting, Econometric Modeling

Advanced Big Data Analytics - Hadoop, Spark, Hive, UNIX, Distributed Machine Learning

Data Visualization - Decision Support Technologies, Visualization Principles

RELEVANT PROJECTS

Quantitative Trading Platform

Contributed to the launch of an intraday quantitative trading platform by leading backtesting analytics, optimization research, and cross-functional validation to improve robustness and deployment readiness.

- Developed end-to-end forecasting pipeline in Python to predict market movements using technical indicators and regime detection.
- Engineered 50+ features from price/volume data and optimized 1000+ hyperparameter combinations using Optuna's Bayesian optimization across 3 years of historical data.
- Implemented walk-forward validation and probabilistic methods for uncertainty quantification, achieving 22% reduction in forecast error through continuous recalibration.

LLM Applications & Prompt Engineering

- Implemented prompt engineering techniques using Microsoft Phi-3, LangChain, and transformers to optimize LLM outputs for text summarization and classification tasks.
- Built RAG-style retrieval workflows using sentence-transformers and embeddings, achieving 35% improvement in response relevance across 250 test queries.

Classification Model Optimization

- Built and optimized multi-class image classification models using scikit-learn on 70,000-image dataset, implementing SGDClassifier and ensemble methods.
- Performed precision-recall optimization through confusion matrix analysis and cross-validation, achieving 94% classification accuracy with 92% precision.

CERTIFICATIONS

IBM AI Engineering Professional Certificate ● Snowflake Hands-On Essentials: Data Warehousing Workshop ● Professional Scrum Master™ I