TOUHID IMAM

6745 W Coolidge St, Phoenix, AZ 85033 — (712) 577 8084
Touhid.Imam@coyotes.usd.edu
touhid.us | github.com/touhid-imam/data-analyst | linkedin.com/in/touhidimam

Summary

Data Analyst with 5+ years of experience in data integration, predictive modeling, and business intelligence across multiple industries, including healthcare. Expert in SQL, Python, and data visualization via Tableau, with a specialized ability to build automated data pipelines and AI-driven analytical tools. I have a proven track record of turning complex datasets into strategic growth, having improved sales forecasting to 90% accuracy and reduced operational fraud by 25% through advanced statistical modeling.

Skills

- **Data Analysis & Processing:** SQL, Python (Pandas, NumPy, Scikit-learn), Excel, Data Wrangling, Feature Selection
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Excel Dashboards
- **Statistical Analysis:** Hypothesis Testing, Regression Modeling, Predictive Analytics
- **Tools & Platforms:** Jupyter Notebook, Git, VS Code, RStudio, Azure (basic), Office 365
- **Scripting & Automation:** Python, Bash, Process Automation
- **Collaboration & Communication:** Stakeholder Collaboration, Team Coordination, Analytical Reporting

Education

- **Masters in Computer Science**, University of South Dakota *May 2025*
  - ACM Club Member
  - USD Computing Club
  - BSA Treasurer

Experience

- **Intern Data Analyst**
  UpSkill (Remote)
  September 2025 – Present
  - Analyzed healthcare data using Python and SQL, transforming raw datasets into actionable insights to improve clinical and operational efficiency.
  - Designed interactive Power BI and Tableau dashboards to provide real-time strategic insights for healthcare decision-makers.
  - Engineered automated data pipelines and applied NLP to streamline data integration and ensure high data integrity across multiple sources.

- **Web Analytics Developer**
  RockIT Fuel Design And Tech, Ontario, Canada
  November 2020 – August 2023

  - Analyzed healthcare data using Python and SQL, transforming raw datasets into actionable insights to improve clinical and operational efficiency.
  - Designed interactive Power BI and Tableau dashboards to provide real-time strategic insights for healthcare decision-makers.
  - Engineered automated data pipelines and applied NLP to streamline data integration and ensure high data integrity across multiple sources.

- **Web Developer**
  Magic Technologies Group, California, USA
  April 2018 – October 2020

  - Developed responsive websites and web applications using HTML, CSS, JavaScript, PHP, and Flask, ensuring high-quality performance and user-friendly interfaces.
  - Managed the full project lifecycle from initial concept to deployment, collaborating with cross-functional teams to align web solutions with client and business goals.
  - Transitioned from a successful freelance project to a full-time role, focusing on improving website functionality and enhancing the overall user experience through optimized code.

Projects

- **Sales and Profit Performance Analytics Dashboard** *[Live Project]*
  Technologies: Tableau, SQL, Data Visualization, Trend Analysis, KPI Reporting, Data Modeling

  - Designed an interactive Tableau dashboard to visualize multi-year sales and profit performance across various product categories and subcategories.
  - Developed key performance indicators (KPIs) to track year-over-year growth, identifying a 29.47% increase in sales and a 32.74% boost in total profit.
  - Conducted deep-dive trend analysis to pinpoint top-performing products and identify specific subcategories experiencing losses for operational improvement.
  - Integrated dynamic filters and real-time trend graphs to streamline the identification of growth opportunities and enhance strategic decision-making.

- **Junipar - Interactive AI Portfolio Agent** *[Live Project]*
  Technologies: Python, Streamlit, OpenAI API (GPT-4o), Custom CSS, Data Integration

  - Developed a professional AI-driven portfolio application using Streamlit and GPT-4o to provide recruiters with an interactive, conversational "digital twin" experience.
  - Integrated a comprehensive knowledge base including professional history, research publications, and project deep-dives to ensure context-aware responses and data-driven insights.
  - Implemented custom CSS for a branded UI/UX and integrated secure password protection to manage API consumption and safeguard professional assets.

- **Airbnb Seattle Market Analysis Dashboard** *[Live Project]*
  Technologies: Tableau, Excel, Data Preprocessing, Geospatial Analysis, Trend Analysis

  - Developed a comprehensive Tableau dashboard to analyze Seattle Airbnb market data, focusing on pricing structures and revenue distribution.

- Evaluated the relationship between property features and pricing, identifying how bedroom counts and specific zip codes influence average daily rates.
- Tracked annual revenue trends to identify seasonal growth patterns and peak demand periods for strategic market positioning.

- **Bike Sales Performance Dashboard** *[Live Project]*
  Technologies: Microsoft Excel, Pivot Tables, Data Cleaning, Slicers, Trend Analysis
  - Developed an interactive Excel dashboard to analyze customer demographics and income levels, utilizing pivot tables and slicers for dynamic data exploration.
  - Evaluated the impact of commuting patterns on sales, identifying a significant correlation between shorter commute distances and increased purchase frequency.
  - Segmented sales data by age brackets and marital status to pinpoint high-performing consumer profiles and provide actionable insights for targeted marketing.

- **Lung Cancer Prognostication with Machine Learning**
  Technologies: Python, XGBoost, LightGBM, CatBoost, Feature Engineering, ADASYN, Optuna
  - Built a stacking ensemble model combining XGBoost, LightGBM, and CatBoost for lung cancer prediction.
  - Improved prediction accuracy by handling class imbalance using Adaptive Synthetic Sampling (ADASYN).
  - Applied feature selection techniques like RFECV and LASSO to enhance model efficiency.
  - Used Bayesian optimization with Optuna to fine-tune hyperparameters for better performance.
  - Validated the model with Stratified K-Fold Cross-Validation, ensuring robustness for clinical applications.

Research

- **A Multimodal Analytical Approach to Alzheimer's Disease Diagnosis Using Machine Learning and Convolutional Neural Networks on MRI Datasets**
  Published by **IEEE**
  - Investigated early detection of Alzheimer's Disease using machine learning and deep learning techniques on the OASIS MRI dataset.
  - Employed models such as Random Forest, Logistic Regression, Extra Trees, and Convolutional Neural Networks (CNN), achieving the highest accuracy and AUC with CNN.
  - Demonstrated the importance of early diagnosis to enable timely interventions and improve disease management strategies.

- **An Ontological Framework for Lung Carcinoma Prognostication via Sophisticated Stacking and Synthetic Minority Oversampling Techniques**
  Published by **IEEE**
  - Developed a stacking ensemble model combining XGBoost, LightGBM, and CatBoost, achieving improved predictive accuracy for lung cancer detection.
  - Addressed class imbalance with Adaptive Synthetic Sampling (ADASYN) and applied feature selection techniques like RFECV and LASSO Regression to enhance model performance.
  - Optimized hyperparameters using Bayesian techniques with Optuna and validated robustness through Stratified K-Fold Cross-Validation, demonstrating clinical applicability.

- **Recondite Thyroid Pathology Prediction: Hermeneutic Integration of Neural and Machine Learning Architectures**
  Published by **IEEE**
  - Developed an ensemble model integrating CNN, TabNet, and machine learning algorithms like XGBoost, achieving 98.7
  - Addressed class imbalance using advanced preprocessing techniques and optimized performance through feature selection, hyperparameter tuning, and cross-validation.
  - Demonstrated the effectiveness of combining deep learning and ensemble approaches for early and accurate diagnosis of thyroid disorders, improving clinical decision-making.

Certification

- **SQL Intermediate Certificate**
  HackerRank
  Issued - May, 2025
- **Fundamentals of Data Governance**
  Edureka
  Issued - October, 2025
- **Healthcare Data Security, Privacy, and Compliance**
  Johns Hopkins University
  Issued - October, 2025
- **Process Data from Dirty to Clean**
  Google
  Issued - November, 2025
- **Supervised Machine Learning: Regression and Classification**
  Standford CPD
  Issued - August, 2024

References

- **KC Santosh**
  University of South Dakota
  Department Chair
  Email: kc.santosh@usd.edu
  Phone: (929) 264-1429
- **Andrew Oldroyd**
  Rock IT Fuel Design and Tech
  Founder
  Email: andrew@atmwebdesign.ca
- **Mike Robles**
  Magic Web Studios
  CEO
  Email: mrobles@mtgi.net
  Phone: (605) 677-3184

- **Md Mahfujul Islam**
  Fora Financial LLC
  Senior Software Engineer
  Email: Mahfujul.islam@forafinancial.com
  Phone: (929) 264-1429