Nathan Diffey

diffeyn.com | Houston, TX | (281) 685-8505 | diffeyn@gmail.com

SUMMARY

Data Analyst with hands-on experience in data cleaning, statistical modeling, and visualization using R, Python, SQL, and Tableau. Skilled at collecting, processing, and analyzing complex datasets to uncover actionable insights that drive operational improvements. Recognized for strong problem-solving skills, clear communication, and collaborative work across teams. Adaptable, detail-oriented, and proactive in fast-paced environments, with a focus on using data to inform strategic decisions.

PROJECT EXPERIENCE

Major League Soccer Database and match prediction – Personal Project | https://github.com/diffeyn/MLS   **Feb 2025 - Present**
- Built a relational database (MySQL) to manage structured time-series data for matches, players, stats, and commentary, incorporating composite keys and hashed IDs for consistency across datasets.
- Automated data ingestion by developing Python scrapers (Selenium, BeautifulSoup, requests) to extract schedules, rosters, and detailed player/match statistics; maintained pipelines with GitHub and PythonAnywhere schedulers
- Standardized and cleaned multi-season data (regex, feature engineering, R tidyverse) to resolve inconsistencies and reshape raw stats (e.g., converting possession bars into uniform 5-minute intervals).
- Performed exploratory analysis and initial modeling (kNN, decision trees, regression) to assess team and player performance.
- Currently expanding project scope to develop predictive models for match outcomes using advanced machine learning techniques, with planned integration of seasonal and per-match player stats into the pipeline

Predicting Austin Animal Shelter Outcomes- Team Course Project   **Sep 2024 - Dec 2024**
University of Houston- Downtown   *Houston, TX*
- Utilized R for data cleaning, preprocessing, and feature engineering on 170,000+ animal intake and outcome records to prepare data for machine learning.
- Applied multiple supervised and unsupervised models (Logistic Regression, Decision Tree, Random Forest, ANN, Ensemble) to predict animal outcomes and identify key predictive factors.
- Generated actionable insights on how intake type and age influence adoption and relocation trends, improving understanding of shelter operations and outcome patterns.

Learning English in the United States - Course Project   **Sep 2024 - Dec 2024**
University of Houston- Downtown   *Houston, TX*
- Utilized R for data cleaning and preprocessing of 5+ datasets with some containing 10,000+ records, and Tableau for data visualization to create an interactive dashboard to identify learning outcome trends, encompassed by a presentation at the end.
- Derived actionable insights from datasets, improving understanding of how English as a Second Language programs have been performing over a period of time.

EDUCATION

The University of Houston- Downtown, *Master of Data Analytics*   **Jan 2024 - May 2026**
- **GPA:** 3.25
- **Coursework:** Statistical Foundations for Data Analytics (A), Programming Foundations for Data Analytics (A), Data Mining (B), Database Management Systems (B), Big Data Analytics (C), Multivariate-Nonparametric Statistics (A)

The University of Houston- Downtown, *Bachelor's of Science, Psychology*   **Jan 2018 - Dec 2022**
- **GPA:** 3.25

EXPERIENCE

Certified Pharmacy Technician   **Mar 2024 - Present**
- Analyzed prescription and inventory data to identify trends, optimize inventory levels, and improve workflow processes, demonstrating a strong application of data analytics for operational improvements.
- Maintained accurate records using robust data management techniques and detailed logs, ensuring compliance with regulatory guidelines.
- Conducted quality assurance checks by leveraging data analysis methods to resolve discrepancies and implement corrective measures, thereby enhancing process reliability.
- Collaborated with cross-functional teams to integrate data-driven solutions and support decision-making processes.

The Registered Behavior Technician   **Aug 2022 - Jun 2023**
- Collected, monitored, and analyzed behavioral data to identify key trends and support tailored intervention plans, showcasing practical application of data analysis.
- Employed statistical methods to track progress metrics and generate insightful reports for a multidisciplinary team, aligning with data-driven deci-sion-making.
- Designed evidence-based strategies derived from rigorous data analysis, contributing to enhanced client skill development.
- Ensured high-quality data consistency and reliability by implementing robust data collection and monitoring processes.

TECHNICAL SKILLS

- **Programming Languages**: R, Python, SQL
- **Software**: MS Suite, Tableou, PowerBI, R-Studio, VS-Code, Git, GitHub, Linux Operating System
- **Data & Analytical Skills**: Problem Solving, Machine Learning, Data Analytics, Data Analysis, Data Management, Data Modeling, Analytical Skills, Data Visualization