Gainesville, Florida

John McDonald

Transportation Data Analyst

john.mcdonald@ufl.edu
linkedin.com/in/JohnMcDon/

Education

University of Florida, College of Liberal Arts - Gainesville Florida **August 2025**
Bachelor of Science - Mathematics
Bachelor of Science - Statistics

Relevant Coursework: Statistical Learning, Linear Regression, Time Series and Forecasting, Probability Theory, Real Analysis

Experience

GFT Engineering **March 2026 - Present**
Transportation Data Analyst *Jacksonville, Florida*

- Developed a cross-validated LASSO model for FDOT District 7's regional travel demand model using household travel survey data. Increased explained variance from 40% to 78% of the theoretical noise ceiling for home-based work trips, significantly improving predictive performance across trip purpose.
- Built an automated ArcGIS Python pipeline generating 30+ figures incorporated into the Miami-Dade TPO's adopted Transportation Improvement Program (TIP), replacing a manual GIS process and standardizing output across the project.
- Modeled the effect of perimeter barrier scenarios on through-traffic at a federal facility, translating an infrastructure-access change into a network-level traffic analysis.

TECO Energy **November 2023 - May 2024**
Data Engineering Project (via UF Warrington College of Business) *Tampa, Florida*

- Developed an algorithm to identify meters incorrectly assigned to circuits by analyzing anomalous AMI voltage readings and GIS data; 92% of the meters that my algorithm flagged were validated as being correctly identified as on the wrong circuit.
- Refactored historic meter data for migration to modern SQL database infrastructure.

Moffitt Cancer Research Center **May 2019 - August 2019**
Research Intern, Mathematical Oncology *Tampa, Florida*

- Developed a spatial mathematical model for breast cancer dynamics under the guidance of Dr. Jill Gallaher at Moffitt Cancer Research Center.
- Simulated angiogenesis under Paclitaxel treatment to model the effectiveness of various adaptive therapy treatment schedules.

Projects

NVIDIA AI Hackathon competition **September 2023**

- Ranked top 5 among 495 participants.
- Created a Chrome Browser extension utilizing OpenAi API and AWS computer vision tools to create a one-click guide for educational videos.
- Developed the application's backend, including a complete data pipeline that collects and processes raw audiovisual data into a formatted guide.

MCM: Mathematical Contest in Modeling **February, 2023**

- Processed 32 million real player data sets obtained through the Twitter API.
- Developed a player simulator that predicts future player score distributions with 95.4% accuracy.
- Delivered project results and produced a 25-page report detailing the algorithm and significant data insights in 96 hours.

r/AITA Forum Post Classification **May, 2024**

- Developed a classification model to predict "guilty" verdicts from Reddit's r/AITA using natural language processing and machine learning.
- Engineered 768-dimensional semantic embeddings from post text using all-mpnet-base-v2 and applied feature transformations for improved prediction accuracy.
- Trained multiple models—including Random Forest, Neural Networks, and SVM—achieving 86.2% out-of-sample classification accuracy with Random Forest while mitigating class imbalance using SMOTE.

Skills

GIS & Modeling: ArcGIS, Cube Voyager, PTV Visum
Programming Scripting: Python (NumPy, SciPy), R, SQL, Git, LaTeX, scikit-learn
Methods: Traffic Demand Modeling, ARIMA, Test-Driven QAQC, Maximum Likelihood Estimation, Simulation-Based Assignment