WILLIAM GUY
San Luis Obispo, CA 93405
(805) 295-9407 · wguy73@gmail.com
LinkedIn · GitHub · Website

PROFESSIONAL SUMMARY

Analytical professional with an M.S. in Data Science from UC San Diego and a B.A. in Statistics from Cornell University, with applied experience in healthcare, robotics, and sports analytics. Skilled in SQL-driven data analysis, dashboarding, and translating complex datasets into actionable insights. Known for leadership, communication, and teamwork as a four-year varsity athlete.

EDUCATION

Cornell University – Ithaca, NY B.A. in Statistics	**University of California, San Diego** – San Diego, CA M.S. in Data Science

RELEVANT COURSEWORK

- **Foundational Probability/Statistics in Data Science -** Built a predictive model of NFL Quarterback performance using 25 years of data to evaluate career efficiency & longevity.
- **Python for Data Science -** Strengthened A/B testing and power analysis skills, final project modeled the effects of hurricane exposure on population outcomes.
- **Statistical Computing -** Defined success metrics for hospital performance and applied statistical frameworks across industries using real-world clinical datasets.
- **Big Data Analytics Using Spark -** Forecasted NYC taxi demand and rider duration using large-scale, time-based features and distributed data pipelines.
- **Risk Analysis & Decision Modeling**- Assessed financial and operational risk under uncertainty using Monte Carlo simulation, sensitivity analysis, and decision trees.

TECHNICAL SKILLS

- **Core Languages & Tools:** SQL, Excel, Tableau, Power BI, Looker, Python, R
- **Data Analysis & Reporting:** Exploratory Data Analysis, A/B Testing, Correlation Analysis, Power Analysis, Time Series Forecasting
- **Modeling & Statistics:** Predictive Modeling, Feature Engineering, Principal Component Analysis (PCA), Monte Carlo Simulation
- **Business & Productivity Tools:** Slack, Microsoft Word, Visual Studio, GitHub, Google Workspace, VBA
- **Libraries & Frameworks:** Jupyter Notebook, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, ggplot2, TensorFlow

WORK EXPERIENCE

Brave Career – Remote
Data Science Intern
June 2024 – September 2024

- Analyzed multi-variable health datasets from Ontario hospital, identified risk patterns in asthma and COPD patients.
- Conducted correlation and PCA on clinical data, revealing BMI and weight as top predictors of CRD-related mortality.
- Increased hospital's mortality risk identification of respiratory diseases by 45%.

AMBOT – San Luis Obispo, CA
Data Intern (Robotics)
May 2020 – August 2022

- Analyzed autonomous robot testing protocols to refine hardware and development roadmaps.
- Developed a warehouse inventory system, utilizing SQL to track part usage, optimize restocking cycles, and improve logistics efficiency by 75%.
- Analyzed Interact Centaur "visual" data, deployed to the International Space Station through the European Space Agency.

PERSONAL PROJECT - NFL Wide Receiver Career Success Prediction

- Built a predictive model using TensorFlow to predict NFL wide receiver success, training on 250+ players using rookie stats, combine results, and draft position.
- Achieved 78% classification accuracy in identifying top-quartile performers based on 3-year fantasy point averages.
- Analyzed 8 key combine metrics and demonstrated that 2 (40-yard dash, bench press) had <10% contribution to predictive accuracy, challenging conventional scouting assumptions.