

Diego Osborn

 github.com/d2osborn  linkedin.com/diego-osborn  diegoisaacosborn@gmail.com  +1(928) 723-9215

EXPERIENCE

- Baseball Analytics Intern** | *UC San Diego Baseball* Feb 2023 – Present
- Automated 60+ opponent scouting reports to enhance game preparation, contributing to UC San Diego's 2023 NCAA Division-1 Baseball Big West Conference Championship
 - Developed a QDA-based model to automatically classify pitch types, improving manual tagging accuracy and correcting tracking inconsistencies in away games
 - Generated post-game pitching report graphics for our team's pitchers, summarizing each pitch type's metrics (e.g., velocity, movement profiles, whiff%) and visualizing pitch location and break patterns to support coaches in player development
- Data Analytics Intern** | *UC San Diego Career Center* Sep 2024 – Present
- Performed EDA and cleaning on career outcomes data for 12,000+ graduates to support the development of an interactive Power BI dashboard—highlighting insights on salary trends, industries, and top employers for student career planning
 - Built a PostgreSQL database to store and query 100,000+ Handshake accounts; categorized alumni by graduation year and degree type, and facilitated the archival of alumni accounts for improved data organization and historical tracking
- Baseball Analytics Intern** | *Palm Springs Power Baseball* Jun 2023 – Jul 2023
- Developed an interactive [Looker dashboard](#) with visualizations to assist players and coaches in evaluating pitch effectiveness and batted ball metrics by analyzing live in-game pitch data from FlightScope

PROJECTS

- Home Price Predictor** | *Python, Pandas, scikit-learn, Matplotlib, Seaborn* Feb 2025–Mar 2025
- Developed and optimized a machine learning pipeline to predict expected prices across 2.2M+ U.S. real estate listings
 - Created a valuation framework comparing listed vs. predicted prices, identifying 10.5% of listed properties in the US as undervalued
- Spotify Popularity Analysis** | *Python, Statsmodels, Matplotlib, Seaborn* Feb 2025–Mar 2025
- Performed EDA on 100,000+ Spotify tracks to identify key audio features linked to song popularity
- Power Outages** | *Python, Pandas, scikit-learn, Plotly* Jun 2024
- Performed statistical analysis and predictive modeling on U.S. power outage data (2000–2016) to identify key factors affecting major outages and forecast regional energy consumption to help utility companies improve outage response

SKILLS

Languages: Spanish (fluent)
Programming Languages: Python, R, Java, HTML5, \LaTeX
Tools/Databases: SQL, PostgreSQL, SQLite, Tableau, Power BI, Looker
Libraries: Pandas, NumPy, Matplotlib, scikit-learn, Seaborn, Plotly, SciPy, Psycopg2, Statsmodels

EDUCATION

University of California, San Diego Sep 2022 - Jun 2026
B.S. Data Science; Minors: Economics, Mathematics

Relevant Coursework: Probabilistic Modeling and ML, Relational Databases, Statistical Methods, Probability Theory, Exploratory Data Analysis & Statistical Inference, Representation Learning

Honors: UCSD Eleanor Roosevelt College Honors Program, Provost Honors (2x)