

Gustavo Alburez

galburez.pro@gmail.com | Portfolio: galburez.com |

EDUCATION

University of California, San Diego

La Jolla, CA

B.S. Cognitive Science; B.S. Business Psychology

June 2025

- **GPA 3.74**, Provost Honors
 - Relevant Coursework: Data Science in Practice, Business Analytics, Usability & Information Architecture, Interaction Design, Linear Algebra, Research Methods in Psychology, Statistics, Brain Computer Interfaces
-

EXPERIENCE

Language Production Lab, UC San Diego

La Jolla, CA

Research Assistant

September 2024 - June 2025

- Developed and implemented Python scripts to automate data pre-processing, efficiently managing and validating over 30,000 data points, significantly reducing manual processing time while improving accuracy.
 - Conducted 30+ research sessions using research data collection systems, ensuring adherence to established experimental protocol and reliable capture of participant data.
 - Operated, calibrated, and troubleshooted software and hardware systems used for data collection; identified and corrected data discrepancies and performed quality checks to prepare datasets for analysis.
 - Maintained clear, timely communication with the principal investigator and lab team; actively contributed insights during weekly lab meetings.
-

PROJECTS

Multilingual Audio Alignment Pipeline (MAAP) - Python, Praat, Whisper, TorchAudio

- Developed a Python-based data processing pipeline to automate audio ingestion, transcript generation, and forced alignment for English and Spanish recordings, reducing manual processing and increasing scalability.
- Created reusable helper functions and documentation, including workflow diagrams, Jupyter notebooks, and standardized output procedures, enabling researchers to generate alignment data without requiring knowledge of the underlying models.
- Standardized and validated input data by converting audio formats, resampling files, trimming recordings, and normalizing transcripts, ensuring consistent processing and improving output reliability across datasets.
- Presented project results in lab meetings, demonstrating improved alignment accuracy while strengthening science communication skills.

Team Metrics & NBA Finals Success - Python, NumPy, Pandas, Scikit-Learn, Seaborn

- Built a 2500+ point NBA dataset by integrating multiple data sources, cleaning incomplete records, and validating data quality, producing a clean, analysis-ready dataset.
 - Developed and optimized a Random Forest classifier to evaluate relationships between team performance metrics, using statistical and machine learning techniques to identify meaningful trends.
 - Generated reports and visualizations to communicate findings, including feature importance analyses and decade-level performance comparisons that highlighted changes in key metrics over time.
 - Conducted exploratory data analysis using pandas, seaborn and matplotlib to generate descriptive statistics, bar plots, and time-series visualizations of team performance trends.
-

SKILLS

- Research Tools & Methods: PsychoPy, Praat, EyeLink; Usability Testing, Data Collection & Management
- Programming: Python (NumPy, pandas, scikit-learn, TorchAudio, Seaborn), JavaScript, HTML, CSS, MATLAB
- Design and Prototyping: Figma, Framer
- Languages: English (professional), Spanish (professional)