

# William Bai

New York City, NY | (415) 533-6788 | wbai0614@gmail.com

## SUMMARY

Highly Motivated data analyst/engineer looking for mid-level/junior technical role. Skilled in ETL processes, SQL, Python, and cloud platforms (AWS and Azure) and effective communication to bridge technical and business teams. Committed to delivering secure, reliable, and efficient data solutions for business success.

## SKILLS

- **Programming Languages:** Python (Pandas, NumPy, SciPy, Matplotlib), SQL, Spark, Hadoop, R, Java, MongoDB
- **Databases:** SQL (SQL Server, MySQL, PostgreSQL), Snowflake, Redshift, BigQuery
- **BI & Visualization Tools:** Tableau, Microsoft Power BI, Excel (VLookup, Conditional Formatting, Pivot Tables)
- **Cloud Technologies:** AWS (S3, EC2, CloudWatch, RDS/Redshift, Lambda, DynamoDB, Glue), Databricks, Azure

## WORK EXPERIENCE

### PINGPOD Inc. – Sales Growth Specialist (Data)

09/2024 - present

- Built scalable data pipelines in AWS to process millions of rows of session, customer, and PodPlay usage data, reducing processing time by 30% and enabling real-time sales insights (S3, Glue, Redshift)
- Built dashboards to visualize booking trends and customer ratings for franchisees, enhancing decision-making as PingPod scaled to 20+ locations (Tableau, SQL)
- Leveraged queries to identify high cancellation trends in solo-player markets, proposing app-based promotions that reduced cancellations and increased morning slot bookings (SQL, Redshift)

### Edusynth.org – Sales Data Analyst Intern (Paid)

06/2024 - 12/2024

- Developed an automated pipeline, saving 60% of manual analysis time by incorporating automated scheduling, monitoring, and error handling to ensure data accuracy and reliability (S3, EMR, Spark-SQL, Airflow)
- Generated data-driven recommendations based on clustering on age, gender, location, demographics and external factors to enhance ad targeting and strategy (Python, SQL, Tableau)
- Analyzed existing customer data using regression techniques, incorporating variables like temperature, fuel prices, and unemployment to identify actionable insights (PySpark, Scikit-Learn, EMR)
- Conducted A/B testing on ad spending strategies, optimizing budget allocation by analyzing conversion rates, customer engagement, and ROI across multiple channels (Python, SQL, Bayesian Statistics)

### Wrights HR Consulting – Operations Analyst Intern

03/2024 – 06/2024

- Automated data processing for HR workflows to handle text and CSV data, ensuring efficient data extraction and analysis, reducing processing time by 42% (Spark, AWS)
- Led data analysis to optimize hiring, onboarding, and exit processes by interviewing HR personnel and analyzing document creation times (RDS, MySQL)

## EDUCATION

### University of California, Irvine

Irvine, CA

- Bachelor of Arts in Economics

12/2024

- Minor in Computer Science

- Computer Networking

- Software Engineering

- Operating Systems

- Programming Compilers

## CERTIFICATIONS

- AWS Certified Data Analytics – Specialty

## INTERESTS

- **Competitive sport:** US National Team Member in Table Tennis 2018-2021 (Top 8 in US men's, Top 4 in US Juniors)