Alejandro Batres

Email: alejandrobatres001@gmail.com
LinkedIn: linkedin.com/in/alejandro-batres
Github: github.com/alejandrobatres
Website: alejandrobatres.github.io

Skills Summary

- Languages:
  Python, SQL, C++, Typescript
- Frameworks:
  Pandas, Numpy, Airflow, Psycopg2, SQLAlchemy, Boto3 API, PyTorch, React
- Cloud Technologies:
  Airflow, AWS Services (Redshift, Lambda, EC2, S3, Glue), PostgeSQL, Gitlab CI/CD

Experience

- Data Engineer — Remote
  TransUnion — *Dec 2022 - Present*
  - Data Validation Framework:
    Developed and implemented a data validation framework utilizing Redshift and PostgreSQL to perform comprehensive data quality checks throughout an automated Pentaho workflow.
    The framework ensured the accuracy and integrity of data across multiple stages, reducing periodic manual data effort from an average of over 1 week down to 2 days and enabling timely identification and resolution of data issues.
  - UK-Compliant Data Retrieval:
    Created a data retrieval process for UK GDPR (United Kingdom General Data Protection Regulation) compliance by utilizing Python packages psycopg2 and boto3 to identify and retrieve account data based on subject requests, generate a PDF report detailing the data used by the company, and delete or suppress the account as per the subject's decision.
    This process ensured that the company adhered to GDPR regulations, improved data privacy measures, and reduced the risk of non-compliance penalties by automating the handling of subject access requests.
  - Customer Dashboard Development:
    Added backend capabilities for enrolling clients across multiple regions. Leveraged AWS EC2 and Airflow DAGs to streamline the enrollment process, ensuring efficient and scalable processing. Conducted enrollment for various regions, tailoring processes to meet specific regional requirements.
  - Data Lake Hydration:
    Automated the data flow from Amazon Redshift to AWS Glue in Spectrum databases utilizing a scheduled Airflow workflow to execute the data transfer process, ensuring timely and accurate data integration across the platforms.
    This automation reduced manual intervention, enhanced data processing efficiency, and improved the overall reliability of data management within the Spectrum environment.

- Software Engineering Intern — Remote
  Allevia — *Aug 2021 - Jan 2022*
  - Designed an optimized relational model and data pipeline for the collection and modeling of user input utilizing AWS S3 for secured storage of user data and AWS Lambda for the processing of user metrics.
  - Implemented an interactive user-faced mobile application using the React Native for deployment on both IOS and Android with a concentration on accessibility and ease of use.

- Computer Science Tutor — Montclair, NJ
  The Coding School — *Sep 2022 - Dec 2022*
  - Designed and assisted students with projects in Python and Java to develop students' skill programming ability.
  - Conducted lessons to groups of students in Python and Scratch to provide an introduction to coding languages.

- Mathematics and Science Tutor — Montclair, NJ
  Montclair Tutoring — *Aug 2018 - Dec 2022*
  - Instructed students of all ages (K-12, adults) in elementary mathematics, algebra, calculus, statistics, chemistry and physics. Provided test preparation for SSHAT, ISEE, SAT, ACT, ASVAB, Praxis, GED.

Education

- New York University Tandon School of Engineering — Brooklyn, NY
  Bachelor of Science in Mathematics; Minor in Computer Science — *Aug 2018 - May 2022*
- Relevant Coursework:
  Object Oriented Programming, Data Structures, Analysis of Algorithms, Machine Learning, Databases, Linear Algebra, Applied Probability, Applied Statistics, Numerical Analysis, Math Modeling

Projects

- School Scheduling Application:
  - Created a SAT solver command-line application in Python for scheduling students into classes based on their schedules.
  - Found an optimized solution utilizing Google CP-Model. (2019)