# Alonso Jose Gutierrez Moran – Data Engineer

+1 (954) 513 9734 | Brooklyn NY, 11214 | alonsojgutierrez@outlook.com | alonsojg.github.io
LinkedIn/Alonso

## Work Experience

### Data Engineer | Capgemini America | New York, US. (Remote) | Mar, 2022 – Present
- Designed and developed ETL pipelines for data ranging in source, pre-processing and warehousing paradigms with Python, MySQL, PostgreSQL, Databricks and Scala to create solutions such as blockchain-trace dashboard and others.
- Implemented data pipelines using Hadoop and Spark for distributed processing and  cut down blockchain analysis time in half .
- Collaborated with cross-functional teams to understand data requirements and provide solutions via collaborative environments such as Databricks and Jupyter notebook.
- Improved data processing efficiency by optimizing pipelines and implementing ad-hoc data partitioning schemas.
- Presented findings through clear and concise reports and visualizations Tableau , PowerBI and Superset .

### Data Engineer/Business Analyst (BA) | SEI Investments Company | Contract | Pennsylvania US. (Remote) | Sep, 2022 – Present
- Eliminated 50% of manual large dataset BA analysis via implementing Python-based automation.
- Responsible for user acceptance testing (UAT) and production release testing API services.
- Presented findings through clear and concise reports and visualizations PowerBI and Tableau.
- Collaborated with development, quality assurance and Apigee teams to provide Ad-hoc solutions tailored to individual Private banking clientele.

### Data Analyst | Freelance | Florida (Remote) | Jun, 2018 – Mar, 2022
- Collected, organized, and analyzed large sets of data to inform business decisions.
- Conducted statistical analysis and data modeling to inform business decisions.
- Created dashboards and reports using Tableau to track key performance metrics.

## Project Experience

### LottoPlays LLC | Startup Tech Guru and Partner
Lead Data Engineer [Python, Bash, C++, Scrapy], Back-end Developer [Bash, C++], Lead Database Engineer [NoSQL, MongDB], Machine Learning Engineer [Pytorch], Automation Processing
- Versatile role is composed of redesigning ETL Pipelines, Database Warehousing architecture redesigning, Back-end architect along with various tasks ranging from visualization to AB testing. Our subscription based service provides users with the latest data on the scratch-off lottery tickets with their respective winning odds on a single site.

### Processing Enron Debacle Data

Machine Learning, natural language processing , Python, Pandas, Numpy

- Enhanced search ML classifier algorithms for machine learning, along with the primary component analysis, stratified data shuffling, grid-searching and score-based model-selection.

### Gene Expression Visualization

Python, Pandas, Numpy, d3.js/dimple.js, HTML

- Statistical testing optimization using Python modules to quickly sort through a large set of gene data and find most influential genes used to set apart two very similar types of Leukemia.

## Education

Bachelor of Science Degree in Mathematics | University of Central Florida | 2018 - 2021
Others…

## Skills

1. Data Engineering: Proficient in best practices practices, including designing and implementing data pipelines, data integration, and data warehousing for efficient data processing and analysis.

2. Programming Languages and Frameworks: Skilled in utilizing programming languages such as Python, SQL, C++, Scala, and R to manipulate and analyze data, implement machine learning (Scikit-Learn and SparkMLlib), and develop data-centric applications.

3. Big Data Technologies: Experienced in leveraging big data technologies like Hadoop, Spark and Apache Hive to handle large-scale data processing and storage.

4. Cloud Platforms: Demonstrated expertise in working with leading cloud platforms such as AWS and GCP, utilizing services like Snowflake, Amazon S3, Amazon Redshift and Google BigQuery to build scalable and robust data solutions.

5. Database Management: Proficient in utilizing various database technologies, including MySQL, PostgreSQL, MongoDB and Apache Cassandra, to design and manage data storage systems efficiently.

6. Data Visualization: Skilled in utilizing data visualization tools like Tableau, Power BI, Superset and D3.js to create interactive and insightful visual representations of data for effective communication and decision-making.

7. Agile Methodologies: Experienced in working with Agile methodologies, employing collaborative tools like JIRA and Confluence to facilitate efficient project management and teamwork.