Contato

83987389391 (Mobile)
erijames123@gmail.com

www.linkedin.com/in/james-barros-ds (LinkedIn)
github.com/Erijames (Portfolio)

Principais competências

IBM Db2
Cloudera
Shell Scripting

Certifications

AWS Cloud Technical Essentials
Python for Data Science and AI
Watson Query with IBM Cloud Pak for Data (V4.6) Foundations
General English
Databricks Lakehouse Fundamentals

James Almeida

Data Engineer | Data Governance | Data Quality | Python | Spark | SQL | Biga Data | AWS | Azure | Databricks | Informatica Developer
Campina Grande, Paraíba, Brasil

Resumo

I bring a wealth of experience in the realms of data engineering and data governance, underpinned by a passion for technology and innovation. My journey includes:

Crafting robust Data Catalogs, Data Lineage, and Business Glossary solutions.
Expertise in Data Visualization tools such as Looker, Tableau, Mode and Power BI.
Adept at data analysis and ETL processes.
☁# Building governed data lakes on cloud platforms like GCP, AWS, and Databricks.
Proficiency in ensuring data quality and implementing top-tier Master Data Management (MDM) solutions.
Armed with a Bachelor's in Analysis and systems development and an MBA in Data Engineer .

I thrive on curiosity, adaptability, and a friendly disposition. An enthusiast for challenges, I'm a frequent participant in hackathons and datathons. Let's connect and explore the possibilities!

Experiência

Jump Label Solutions
Data Engineer I
agosto de 2023 - Present (1 ano)
São Paulo, Brazil

- Offer services to measure, monitor and report data quality compliance indexes in the analytical environment, promoting continuous improvement and generation of lasting value to the business through our Center of Excellence in Data Quality;

- Development of data quality processes involving different dimensions: Data Profile, Validation, Standardization, Hygiene, Deduplication, Enrichment, Unification;

- Construction and monitoring of data quality panels with analytical vision, providing support for changes and remediations related to data quality;

- Data quality monitoring for various sectors

Accenture
Data Engineer Analyst
julho de 2022 - novembro de 2023 (1 ano 5 meses)
São Paulo, Brazil

-Development of pipelines using Pentaho-PDI to extract, transform and load data using APIs.
-Data modeling for storage in a Data Mart environment, Fact-Dimension tables, Star Schema model.
- Data manipulation in Oracle/SQL Server environment;
- Data manipulation in the Hadoop ecosystem (Hive, Sqoop, HDFS) using HiveSql, Python and PySpark;
- Structuring databases;
- Application of business rules in a BIG DATA environment.
• Results Achieved: Work focused on optimizing the data according to the customer's needs, to reduce the use of resources and maintaining data quality, migrations between servers to better allocate the data

BlueShift Brasil
2 anos 6 meses

Data Scientist PL.
janeiro de 2022 - agosto de 2023 (1 ano 8 meses)
Campina Grande, Paraíba, Brasil

- Natura & Co - Cosmetic products sector
Working in the data scientist team and in the financial services squad, offering solutions for defining strategies.
- Maintenance and development of ML models.

Data Scientist Jr.
maio de 2021 - março de 2022 (11 meses)
Campina Grande, Paraíba, Brasil

* Reverse engineering SAS code for python/pyspark, creating source transcripts.
* Creation of feature stores, calculation of variables (algorithm), application of recycling and
deploy, map, and evaluate source data.

Data Engineer Trainee
março de 2021 - maio de 2021 (3 meses)
Campina Grande, Paraíba, Brasil

ETL development and data modeling
- Support of data pipelines in Azure cloud(Function)
- Programming in Python and SQL
- Data visualization with Power Bi and Google Data Studio
- Azure data engineering (DataFactory, Databricks, Webscraping) ETL development and data modeling - Azure cloud data pipeline support (Function) - Python and SQL programming - Data visualization with Power Bi and Google Data Studio - Engineering Azure Data (DataFactory, Databricks, Webscraping)

Competencies: ETL (Extract, transform and load)

Ford Motor Company
Artificial Intelligence Researcher
outubro de 2022 - abril de 2023 (7 meses)
United States

Develop pre-processing architecture involving steps in Hive, Python and Alteryx.
- Data science applied to big data.
- Feature engineering.
- Data visualization in Qlik Sense
Competências: Altery • Qlik Sense • GitFlow
Big data • Sistemas operacionais embarcados • Python • Ciência de dados

AmorSaúde
Supervisor clínico
janeiro de 2021 - maio de 2021 (5 meses)
Campina Grande, Paraíba, Brasil

grc treinamentos
Supervisor de call center
agosto de 2019 - dezembro de 2020 (1 ano 5 meses)

Agibank
Sales Supervisor
abril de 2017 - agosto de 2019 (2 anos 5 meses)
Campina Grande e Região, Brasil

Caixa Econômica Federal
Assistente de gerente
janeiro de 2011 - junho de 2013 (2 anos 6 meses)
Campina Grande e Região, Brasil

———

Formação acadêmica

FIAP
MBA ENGENHARIA DE DADOS, Tecnologia da Informação · (novembro de
2023 - março de 2025)

Anhanguera Educacional
Postgraduate Degree, Engenharia de Dados · (julho de 2023 - janeiro de
2024)

UNOPAR - Universidade Norte do Paraná
Análise e Desenvolvimento de Sistemas, Tecnologia em Informática/
Software · (agosto de 2021 - julho de 2023)