

Felipe Vidor Vaz

Porto Alegre, RS, Brazil 91020-010

Mobile: +55 51 991019404 / vaz.felipe@gmail.com / www.linkedin.com/in/vazfelipe/

GitHub: <https://github.com/VazFelipe>

Data Modeler | Data Engineering | Statistical Analysis

WHO AM I?

I am a Data Engineer, designing data flow processes, developing ETL pipelines, collaborating with business analysts and data scientists to overcome business challenges. Also, I am a researcher in cognitive sciences, enthusiast of Machine Learning and Statistics.

I have advanced knowledge working with Microsoft Excel Expert, SQL, Python, Oracle Data Integration, Azure Data Factory, Google Dataflow, dbt, Airflow, and developing dashboards using Power BI and Qlik Sense.

My abilities as a Data Engineer are rooted as an Analytical Engineer: since 2018 I have delivered insights using statistical approach, designing data model structures, developing functions, complex SQL queries, Excel workbooks and transformations with Python.

I have done my career transition to the Data Engineer role in 2020, when I have started performing data ingestion, transformation and loading from on-premise, cloud and external databases.

WHAT AM I DOING RIGHT NOW?

Currently, I am working on a company in car rentals building and managing data flows (packages) using SQL Server Integration Services, maintaining OLTP and OLAP processes (SAP BO/BW, Sybase, SQL Server), working with store procedures (e.g., triggers), and migrations from legacy processes through Google Cloud. I support the data science teams doing data engineering tasks that improves the data collection, data processing, and providing cleansed data coding.

Usually, I am troubleshooting, improving, and deploying SSIS packages, evaluating the quality of data, modeling data conceptually, logically, and physically (Power Designer, dbt framework). Typically, the data solutions are being deployed in the data lake or data warehouse on Google Cloud (e.g., Cloud Storage, BigQuery) with Dataflow jobs and Cloud Composer (Airflow) for orchestration.

Finally, using an agile framework with Azure DevOps for team collaboration, I am focus on improving data pipelines that reduces the overall cost, latency and delivers data that met the business requirements and accelerate decision making.

QUALIFICATIONS

- GCP
 - Engineer Data in Google Cloud ([certificate](#))
 - Build and Optimize Data Warehouses with BigQuery ([certificate](#))
 - Perform Foundational Data, ML, and AI Tasks in Google Cloud ([certificate](#))
 - Serverless Data Processing with Dataflow: Develop Pipelines ([certificate](#))
 - Serverless Data Processing with Dataflow: Operations ([certificate](#))
 - Building Batch Data Pipelines on Google Cloud ([certificate](#))
 - Create and Manage Cloud Resources ([certificate](#))
 - Google Cloud Big Data and Machine Learning Fundamentals ([certificate](#))
 - Modernizing Data Lakes and Data Warehouses with Google Cloud ([certificate](#))
 - Serverless Data Processing with Dataflow: Foundations ([certificate](#))

- Data Engineering, Big Data, and Machine Learning on GCP Specialization ([certificate](#))
- Building Batch Data Pipelines on GCP ([certificate](#))
- Building Resilient Streaming Analytics Systems on Google Cloud ([certificate](#))
- Google Cloud Big Data and Machine Learning Fundamentals ([certificate](#))
- Modernizing Data Lakes and Data Warehouses with Google Cloud ([certificate](#))
- Azure
 - Microsoft Azure Databricks for Data Engineering ([certificate](#))
 - Data Integration with Microsoft Azure Data Factory ([certificate](#))
 - Microsoft Azure for Data Engineering ([certificate](#))
- Concepts in Python - loops, functions and returns ([certificate](#))
- Storing, Retrieving, and Processing JSON data with Python ([certificate](#))
- dbt Fundamentals ([certificate](#))
- Microsoft Certified Trainer 2021-2022 ([certificate](#))
- Microsoft Excel Expert (Microsoft 365 Apps and Office 2019) ([certificate](#))
- Mathematics for Machine Learning: Multivariate Calculus ([certificate](#))
- Mathematics for Machine Learning: Linear Algebra ([certificate](#))
- Machine Learning - Stanford University (Andrew Ng) ([certificate](#))
- Scrum Foundation Professional Certificate (Certiprof)
- Data Driven Business Analysis using Canvas (iMaps)

ACADEMIC FORMATION

- Specialization on Statistics and Applied Mathematics (360 hours). 2020-2022
- Bachelor of Business Administration - Pontifical Catholic University of Rio Grande do Sul (PUCRS). 2012-2017 (Magna Cum Laude Score: 9,21)
 - Marketing research group, statistics tutor

EXPERIENCE

Localiza

- Data Engineer – 02/2022 – present
 - Interviewing stakeholders, identifying business requirements for data solutions
 - Designing, building, testing, and maintaining data flow architectures
 - Building data flows packages and managing SQL Server Integration Services
 - Maintaining OLTP and OLAP processes (SAP BO/BW, Sybase, SQL Server, Power BI Report Service)
 - Developing data visuals using Power BI and Excel
 - Defining data conceptually, logically, and physically using Power Designer and dbt models
 - Developing complex SQL queries, SSIS expressions, dotNet and python scripts
 - Supporting Data Science Teams in data wrangling and providing cleansed data coding

Unimed Porto Alegre

- Data Engineer – 04/2021 – 02/2022
 - Understanding, identifying, and documenting data solutions for the business, as well as data processes and pipelines
 - Prototyping, testing and tuning data flows using SQL query
 - Building, testing, and maintaining solutions for Data Warehouses and Data Lakes
 - Designing, building, testing, and maintaining data flow architectures
 - Managing cloud computing platforms with a focus on analytics services
 - Developing algorithms to transform data into useful and actionable information
 - Defining the frequency of data update as needed by the company, always seeking to obtain information in real time
 - Working in collaboration with scientists and data analysts to build solutions that meet the organization's needs
 - Understanding issues related to training machine learning models, algorithm types and performance to understand the needs of data scientists
 - Designing and developing data visualization on Power BI and Excel
 - Developing complex SQL queries, developing data transformations with Python and R
- Data Modeler Analyst – 08/2020
- Data Analytics Specialist – 09/2019
- Relationship Analyst – 03/2018

Pontifical Catholic University of Rio Grande do Sul (PUCRS)

- Data Analyst II – 09/2016 – 04/2017
 - Developed Excel Data Models for budget supporting
 - Conducted process mapping using BPM methodology
- Data Analyst I – 11/2014
- Scientific Initiation Scholarship – 03/2014

LANGUAGE

- English skills advanced for reading, listening, and speaking
- Vanhack b2+ advanced English skill

RESEARCH

- Scientific publication: Cognitive dissonance and sustainable consumption: a systematic review of the literature. <https://hdl.handle.net/10923/14220> (PUCRS Repository)
- Scientific article – unpublished work: A brief history of sunspots and the application of Fourier Series Analysis. <https://bit.ly/3SBdSiM> (My Google Drive)