

**José R. Sousa**

Charlotte, NC 28217

LinkedIn: <https://www.linkedin.com/in/jrsousa2/>

**Phone:** 904-290-2890

**E-mail:** [jrsousa2@gmail.com](mailto:jrsousa2@gmail.com)

**GitHub:** <https://jrsousa2.github.io/>

---

## SUMMARY OF QUALIFICATIONS

---

- Large experience in **database / data warehouse** and **SQL** (query language).
- Great with **Python** programming in **VS Code**. Created executable from code, made **web API** requests to retrieve data, created **GUI's**, developed **REST API's**, loaded **Excel** and **XML** files into **pandas** dataframes, read and wrote to Excel files with **openpyxl**, created formatted **PowerPoint** slides automatically, created **virtual environments**, automated **batch jobs**, etc. [All can be viewed on my GitHub](#).
- Extracted data from **Hadoop** with Dremio / Hive and created VDS's.
- Fitted **logistics regression** model using **PySpark ML**.
- Created **workflows** in **n8n** for Gen AI.
- **Expert at SAS** (**Enterprise Guide** and **Viya**) -- created **stored processes** and **ETL's**, wrote advanced code to produce tables and reports, and truly mastered **SAS macros**. Created dashboard with SAS/GRAPH and reports with **Proc Report**.
- Strong with **Machine Learning** models and **data analytics** and familiar with **actuarial analyses** and **actuarial reports**.
- Expert at **MS Excel**, created advanced spreadsheets with **VBA macros**, **automated charts and graphs**.
- Most of my experience is in the **insurance** industry (10+ years). Created a wide variety of **loss triangles**, estimated **ultimate claims** from LDF's, calculated ceded losses from **reinsurance** contracts, calculated earned premiums and created reports on **loss ratio** and **in-force policies**, among others. Performed **data reconciliation** with the general ledger.
- Exposure to **accounting principles** (GAAP, STAT), created **YTD summary reports** to support quarterly filings (10-Q and 10-K) using the **general ledger** database and performed account and profit center level reconciliations with it.
- Graduated in **Math / Statistics** from the **best university in Brazil** (**U of São Paulo**).
- Very analytical, self-taught, can **work independently**, created new formulas in math (e.g. <https://arxiv.org/html/1810.07877v7>).
- **U.S. citizen** (had a **green card** through **NatGen** previously).

---

## SOFTWARE

---

- **Databases:** **Oracle**, **Teradata**, **SQL Server / SSMS** and their **IDE's**.
- **Big Data:** **Snowflake**, **Databricks**, **PySpark**, Dremio, Hive.
- **Programming:** **Python**, **VBA**, **VBScript**, Unix/Windows batch scripting.
- **Cloud:** Working knowledge of AWS, Azure and GCP.
- **SAS:**
  - Platforms: **Unix**, **Linux**, Windows Server and **Viya** (Cloud).
  - Modules: **BASE**, **STAT**, **GRAPH**, **ACCESS**, **CONNECT**.
  - Expert knowledge of most used SAS Procs, including **Proc Report**. Advanced **macro language** and **stored processes**.
- **BI / Predictive Modeling:** **SAS Enterprise Guide**, **Power BI**.
- **Code Editors:** **VS Code**, Jupyter Notebook.
- **Other:** **Excel**, **PowerPoint**, SharePoint, **Git**, **RegEx**, **Jira**, **XML**, **JSON**.

---

## PROFESSIONAL EXPERIENCE

---

**Wells Fargo (banking)**

*Senior Data Analyst / Data Scientist (contract / hybrid)*

**10/2022 – 12/2025**

- **Second project: 07/2025 through 12/2025**

- Worked on **AML / Financial Crimes** remediations (CIP / BO / CDD / **KYC**).
- Built AML remediation populations (Trust, NRM4 and transportation assets).
- Created population identification workbooks and waterfall charts and presented them to senior directors for approval.
- Wrote complex **SQL** queries to extract data from **Hadoop** using **Dremio** and **Hive** and from various databases (**Teradata**, **Oracle**, SQL Server, DB2, etc). Used database **metadata** to locate tables and fields.
- Developed **SAS** and **SQL** code to create ad-hoc, time-sensitive data extracts and **reports**.
- Created code test scripts and worked with code testers to validate the code.
- **Recognized with a full-time offer at project completion.**
- **First project: 10/2022 through 01/2025**
  - Worked on **customer remediations**, responsible for the creation of the final artifacts from which refunds and letters are issued to customers as required by **regulators**, following strict **Confluence** procedures.
  - Managed **SAS** code used to create final remediation artifacts, made changes to the code to add **enhancements**, **optimize performance** and minimize the number of **defects** and findings in **peer reviews**. As an example of enhancement, when a new analyst takes over a remediation, all the **artifacts** and **SAS** tables created by the previous analyst are **automatically** saved to a new subfolder as **backup**.
  - Created **ETL**'s to produce new artifacts by mapping fields from the execution data requirements (EDR) to data tables. Contributed to design **artifact layouts**.
  - Used **GitHub** for code version control and sharing and **DMS** as a secure repository for storing and sharing artifacts.

### **Truist, Charlotte, NC (banking)**

*Data Analytics SQL/Python (contract / remote)*

**09/2021 – 03/2022**

- Created rules based **decision tool** for the compliance and data analysts to choose their optimal analytics application as a web form created in **Python**.
- Created and maintained current report on user accesses for the various **proprietary applications** and data repositories (**AD**, **ACL**), for **compliance**.
- Ran reconciliation between **ServiceNow/RITM** and data pulled from the **mainframe** to assess correctness of an automatic ticket creation process for abends (abnormal job termination), and presented findings with **PowerPoint**.
- Ran **SQL** queries in **Snowflake** to extract data from the **data lake/data warehouse**.
- Created reports using the **SAS grid** and produced **documentation** following internal **templates**.

### **Florida Blue, Jacksonville, FL (healthcare)**

*SAS/Python Developer (contract / remote)*

**04/2020 – 08/2021**

- Developed **ETL** code to extract data from **Oracle** and **DB2** and others, using **SAS 9.4** on **AIX** (Unix) and **Enterprise Guide**.
- Extracted data from multiple sources to create agent **fraud detection** analytics for the **SIU** dept with various metrics. Created **macro to automate** the formatting of the report for each metric using **Proc Report**.
- Automated and scheduled various **SAS ETL** jobs in **Unix**, using a **SAS macro**.
- Used **HTML** syntax in **SAS** to compose rich text e-mail alerts.
- Combined Windows **batch**, **WinScp** and **VB** scripts to schedule jobs and automate file transfers and moving data between platforms using SFTP.
- Created a neat **SAS stored process** for various canned reports.
- Created **visualizations** and **reports** using **Power BI** and **Python**.

- Gave **training** to team members on how to design **SAS stored processes**, covering how to create the **GUI** and how to handle the variable **user's inputs**.

**National General (Allstate), New York, NY (insurance)**

*Insurance Data Programmer / Data Manager*

**04/2015 – 12/2019**

- **Reported to the holding company's chief actuary.**
- Mastered the **claims data system** of the company (EPIC) and legacy data stored in **Oracle** and **SQL Server** databases.
- **Planned** and **created** datasets in **SAS** by extracting and combining data from various databases, from which loss triangles and **data analytics** were produced.
- Created **actuarial loss triangles** used in reserving analyses with **SAS macros**, and reconciled the numbers with the company's **general ledger**.
- Created quarterly company-wide **YTD summary report** in **Excel** using **general ledger data** extracted with **Oracle BI Discoverer** to support quarterly filings and the annual report preparation (**10K, 10Q**), for the **SEC** and other **regulators**.
- Participated in the **ATLAS project**, with a goal to capture claim and policy level data into the **GL Oracle database**.
- Extracted data from the **SAS server** to create reports with **SAS macros** and **code**. These reports were used internally or provided to consultants, **auditors** and **regulators**.
- Using **SAS**, ran **statistical data analyses** to track performance of products.
- Created complete documentation about work done on legacy data from acquired companies.

**AIG, New York, NY (insurance)**

*Data Scientist / SAS Programmer (contractor / FTE)*

**12/2011 – 12/2014**

- Planned and extracted data from multiple sources (**mainframe, DB2, SQL Server** and **Oracle**) to produce **enhanced tables** used in all **reports** and **models** and **processes**, using **ETL** and **SAS/Access** interface to **OLE DB** and **ODBC**.
- Created **ETL's** to read **VB/VSAM** files on **IBM mainframe** and load them into **SAS** tables on Windows. These files became the data foundation of the **Global Actuarial Pricing** department.
- Wrote **SAS macro** to create a wide variety of **loss triangles** used to derive **LDF's** and **estimate** the **ultimate claims**.
- Worked in the **deep dive** and **profitability analyses** of the book of business, by extracting data on **losses, premium, policy** and **insured attributes** to generate various reports and metrics and to monitor the **performance** of the various **lines of business and segments**.
- Created smart **algorithm for linking policies** based on insured and policy characteristics in **SAS**. This policy linking was the basis for the creation of a **scoring model** called **AQI (Account Quality Index)**.
- Worked with Deloitte consultants and provided them with **loss triangles** used as inputs into the Commercial Auto **pricing tool**. Also worked with Deloitte on the **Specialty Worker's Comp model**, making sure that deliverables such as **SAS code, data** and results of the **predictive model** met the requirements.
- Created and scheduled weekly **jobs** to extract submission data entered by underwriters into the **AQI model** and update a **Qlikview dashboard**.

---

**EDUCATION**

Masters's in **Statistics** - **University of São Paulo, Brazil**

Bachelor's in **Statistics** - **University of São Paulo, Brazil**

Bachelor's in **Mathematics** - **University of São Paulo, Brazil**