José Carlos Soares Junior

Vitória - ES, Brazil

in [LinkedIn](#) | 🌐 [josecarlosinfo.com](#) | ✉ [josecarlok905@gmail.com](#) | ⌖ [GitHub](#)

Skills

Tech
- R | Python | SQL | PostgreSQL | MySQL | Excel | Git | GitHub | Jupyter Notebook

Acquired knowledge (Statistics / Data Science / Data Analysis)
- Time Series Analysis | Stochastic Processes | Statistical Inference | Nonparametric Statistics | Hypothesis Testing | Exploratory Data Analysis (EDA) | Probability Theory | Operational Research | Linear Algebra | Calculus
- Regression Analysis (Linear, Polynomial, Multiple, Logistic) | Random Forest | SVM | Gradient boosting | Decision tree | K-Means | PCA

Languages
- Portuguese (BR) | English – *All professional proficiency or above*

Experience

Data Engineering Scholar   [Fiotec](#)   *Remote, BR*   **10/2023 - Present.**

- I have been working on health data extraction using **ETL** (Extract, Transform, and Load) with the aim of making this data available in a centralized and integrated manner to the **Elasticsearch** for the data analysis team at the Brazilian Obstetric Observatory. This will enable them to perform queries using **SQL**.

Scientific Initiation Scholar in Statistics   [OOBr](#)   *Hybrid, Vitória - ES, BR*   **08/2021 - 01/2023. 1 yr 6 mos**

- I conducted general **statistical analysis** on public health data and was responsible for the development, documentation, deployment, and maintenance of interactive panels/**dashboards** using Shiny. These tools played an exceptionally important role during the COVID-19 pandemic, making public health data easily accessible and comprehensible to professionals and researchers in the field, significantly impacting social policies and government decisions related to public health in Brazil. As an example, in 2021, there was a 61.6% increase in weekly deaths for the general population compared to 2020, while for the maternal population, the increase was 145.4%. This critical information contributed to prioritizing the vaccination of maternal populations against COVID-19 in various regions of the country. Such vital information, along with many others derived from my work, whether directly or indirectly, played a pivotal role in enhancing the quality of public health.
- I carried out **advanced time series analysis** on public health data concerning maternal populations in Brazil. It was possible to apply techniques, typically employed in the finance and economic fields, to comprehend public health data related to pregnant women and postpartum women during the COVID-19 pandemic.
- Planned and taught a course on data manipulation and analysis using **Python** and **R** with public health data for an audience (synchronous and asynchronous) of approximately 40 people at the Federal University of Espírito Santo (UFES). I used the theme of recent achievements in astronomy and cosmology related to black holes, along with examples of computer games and intuitive methods to pique the audience's interest in the content presented, while the actual data used was from the public health field.
- I wrote articles for the corporate blog, contributing with my knowledge of **R** programming, **Python**, and **data analysis**.
- Programming and preparation of technical reports using the **R** programming language, **RStudio** environment, and **RMarkdown** tool.
- Development and writing of chapters on **R** and **non-parametric statistical tests** for the project's book titled "Applied Data Science in Maternal and Child Health".

Scientific Initiation Scholar in Statistics   [FAPES](#)   *Remote, BR*   **08/2020 - 07/2021. 1 yr**

- I spent one year working with functional data analysis, a **longitudinal data analysis** methodology, as a solution to improve research methodologies used in the obstetrics field in Brazil for problems where we typically have a scalar target variable and functional input variables (measures over time). Technical reports with partial results were periodically presented. At the end of the study, it was demonstrated through **quantitative research** using **R programming**, **computational simulations**, **data analysis**, and **statistics** that functional data analysis methods were superior and a more evidence-based approach to address these types of problems in the field, in contrast to the methods adopted thus far. This contribution led to an enhancement in the quality of research in this area.

Education

Bachelor's degree in Statistics   [UFES](#)   *Vitória, ES, BR*   **02/2017 - Current**
- Emphasis on **time series analysis** and **machine learning**.