

José Carlos Soares Junior

Tel: 27 996908342 | Vitória - ES

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

Habilidades

Tecnologias

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

Conhecimentos Adquiridos (Estatística / Ciência de Dados / Análise de Dados)

- Análise de Séries Temporais | Processos Estocásticos | Inferência Estatística | Estatística Não Paramétrica | Teste de Hipóteses | Análise Exploratória de Dados (EDA) | Teoria da Probabilidade | Pesquisa Operacional | Álgebra Linear | Cálculo.
- Análise de Regressão (Linear, Polinomial, Múltipla, Logística) | Floresta Aleatória | SVM | Gradient Boosting | Árvore de Decisão | K-Médias | PCA.

Idiomas

- Português (BR) | Inglês – *Proficiência profissional ou acima.*

Experiência

Bolsista em Engenharia de Dados

[Fiotec](#)

Remoto, BR

10/2023 - Atualmente. 4 meses

- Tenho trabalhado na extração de dados da saúde, utilizando **ETL** (extract, transform, and load) com o intuito de disponibilizar esses dados de forma centralizada e integrada ao **Elasticsearch** para a equipe de análise de dados do Observatório Obstétrico Brasileiro, de modo que possam realizar consultas por meio de **SQL**.

Bolsista de Iniciação Científica em Estatística

[OOBr](#)

Híbrido, BR

08/2021 - 01/2023. 1 ano 6 meses

- Realizei análises estatísticas em dados de saúde pública e fui responsável pelo desenvolvimento, documentação, implantação e manutenção de painéis/**dashboards** interativos usando **Shiny**. Essas ferramentas desempenharam papel crucial durante a pandemia de COVID-19, tornando os dados de saúde pública facilmente acessíveis e compreensíveis para profissionais e pesquisadores, impactando significativamente políticas sociais e decisões governamentais relacionadas à saúde pública no Brasil. Como exemplo, em 2021, houve um aumento de 61,6% nas mortes semanais para a população em geral em relação a 2020, enquanto para a população materna o aumento foi de 145,4%. Esta informação crítica contribuiu para priorizar a vacinação das populações maternas contra a COVID-19 em diversas regiões do país. Essas informações vitais, juntamente com muitas outras derivadas do meu trabalho, direta ou indiretamente, desempenharam um papel fundamental na melhoria da qualidade da saúde pública.
- Conduzi **análises avançadas de séries temporais** em dados de saúde pública relacionados a populações maternas no Brasil. Foi possível aplicar técnicas, tipicamente empregadas nas áreas financeira e econômica, para compreender dados de saúde pública relacionados a gestantes e puérperas durante a pandemia da COVID-19.
- Planejei e ministrei um curso sobre manipulação e análise de dados usando **Python** e **R** com dados de saúde pública para um público de aproximadamente 40 pessoas na Universidade Federal do Espírito Santo (UFES). Utilizei o tema das conquistas recentes em astronomia e cosmologia relacionadas aos buracos negros, juntamente com exemplos de jogos de computador e métodos intuitivos para despertar o interesse do público pelo conteúdo apresentado, enquanto os dados reais utilizados eram da área de saúde pública.
- Escrevi artigos para o blog corporativo, contribuindo com meu conhecimento em **programação** nas linguagens **R** e **Python**, e também com meu conhecimento em **análise de dados**.
- Elaborei **relatórios técnicos** usando a linguagem de programação **R**, ambiente **RStudio** e ferramenta **RMarkdown**.
- Desenvolvimento e redação de capítulos sobre **R** e **testes estatísticos não paramétricos** para o livro do projeto intitulado "Ciência de Dados Aplicada à Saúde Materno-Infantil".

Bolsista de Iniciação Científica em Estatística

[FAPES](#)

Remoto, BR

08/2020 - 07/2021. 1 ano

- Passei um ano trabalhando com **análise de dados funcionais**, uma metodologia de **análise dados longitudinais**, como solução para melhorar metodologias de pesquisa utilizadas na área de obstetria no Brasil para problemas onde normalmente temos uma variável resposta escalar e covariáveis funcionais (medidas ao longo do tempo). Relatórios técnicos com resultados parciais foram apresentados periodicamente. Ao final do estudo, foi demonstrado por meio de pesquisa quantitativa utilizando **programação** em **R**, **simulações computacionais**, **análise de dados** e **estatística** que os métodos de análise de dados funcionais são superiores e uma abordagem mais baseada em evidências para resolver esse tipo de problema, em contraste com os métodos adotados até então. Esta contribuição levou a uma melhoria na qualidade de pesquisa nesta área.

Educação

Bacharelado em Estatística

[UFES](#)

Vitória, ES, BR 02/2017 - 12/2023

- Ênfase em análise de séries temporais e aprendizado de máquina.