

GUSTAVO R. SANTOS

DATA SCIENTIST | Apex, NC 27502 | (984) 238-6940

gurezendes@gmail.com | <https://gustavorsantos.me>

Data Scientist passionate about solving complex problems with machine learning, analytics, and visualization. Skilled in Python, Databricks, R, and SQL, delivering actionable insights to drive business impact. Author, Udemy instructor, and blogger with 20K+ monthly readers. Experienced in statistical modeling, predictive analytics, and data storytelling, translating data into strategic decisions, and collaborating with leadership. At Food Lion, my hands-on experience with Databricks helped support decisions and improve operations. Committed to fostering a data-driven culture and mentoring aspiring professionals. Enjoys sharing knowledge through books and courses on ML and data engineering.

EXPERIENCE

2020 – 2024

Data Scientist | Food Lion | Salisbury, NC

Led data analytics and transformation initiatives with PySpark and SQL, enhancing product growth, monetization, and strategic decision-making.

- Impact: Designed and implemented an ETL pipeline, enhancing data quality and improving model-driven e-commerce store selection accuracy by 20%.

Applied statistical analysis techniques (e.g., hypothesis testing, regression) and optimization algorithms to solve complex business problems and generate key insights.

- Impact: Developed and deployed a checkout lane optimization algorithm, reducing labor costs by 33% and increasing front-end space efficiency.

With hands-on Databricks experience, I engineered data pipelines and created machine learning models in Python and R, delivering insights through Power BI dashboards that supported strategic business decisions.

- Impact: Developed an HR incident tracking dashboard, enabling monitoring by region and store.

Applied NLP techniques to tokenize, vectorize, and process text to extract insights about employee turnover.

- Impact: Sentiment analysis to learn reasons for high turnover in stores.

2012 – 2019

Data Analyst | IBM | Hortolândia, SP, Brazil

- Conducted financial analysis on IT projects using SQL and Excel, providing data-driven insights for operational and strategic decision-making.
- Designed and automated KPI reports, forecast tracking tools, and interactive dashboards using SQL, Excel, and Visual Basic, enhancing executive decision-making.
- Positively impacted the company by implementing six semi-automated reports and templates, reducing lead time by 70% and cutting input errors by 99%+. Streamlined workflows, leading to a 50% reduction in overtime hours.

2011 – 2012

Billing Lead | IBM | Hortolândia, SP, Brazil

- Led a three-member team managing mainframe data consumption billing processes, ensuring accuracy and timely reporting for customers.
- Optimized reporting workflows with VBA, reducing lead time by 40% and automating key Excel processes, significantly decreasing manual workload.
- Increased billing accuracy to 99% by streamlining data workflows and automating error-prone processes, improving client satisfaction.

EDUCATION

2023 | **University of São Paulo (USP)**, Brazil — MBA Data Science and Analytics.

2020 | **Data Science Academy** — Data Scientist Professional.

2020 | **MIT**, Online — Data Science and Big Data Analytics Extension.

2006 | **Fundação Getúlio Vargas (FGV)**, Brazil — MBA Marketing

SKILLS

- | | | |
|--------------|----------|-------------------|
| • Python | • Docker | • Databricks |
| • R Language | • MLFlow | • Spark |
| • SQL | • NLP | • Power BI, Excel |

PUBLICATIONS

2024 | [BOOK] Your AI Partner: How Generative AI Enhances Efficiency, Creativity, and Learning in Everyday Life | A book about what generative AI is, how it works on a high level, and how you can use it in your everyday life. Can be seen at <https://www.amazon.com/dp/B0DKPJFYQ5>

2024 | [ONLINE COURSE] Hands-On Data Science Project Using CRISP-DM | <https://www.udemy.com/course/hands-on-ds-project-crisp-dm/>

2023 | [BOOK] Data Wrangling with R with Packt Publishing | A book about modern techniques to load, clean, and transform data using the R language. Can be seen at <https://www.packtpub.com/product/data-wrangling-with-r/>

2023 | [ONLINE COURSE] Mastering Data Wrangling with PySpark in Databricks | <https://www.udemy.com/course/master-data-processing-pyspark/>

MORE ABOUT ME

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

Medium blog: <https://gustavorsantos.medium.com>

GitHub: <https://bit.ly/3g3Xbde>

Portfolio: <https://gustavorsantos.me>