Marcelo**Sanches**

Insights-Oriented Data Scientist

Highly motivated Data Scientist with 2 years of experience. Supported KPA's (most recent employer) largest technology product launch in 30 years with new integrated BI analytics platform. Recognized by leadership for wide "breath of ingenuity, patience, and technical acumen." Brings a **Master of Science in Business Analytics** and foundation in R, Python, and SQL, as well as the discipline and collaborative spirit of a professional classical musician who successfully pivoted to data science in 2018.

Contact

Lafayette, CO
(303) 931 2796
sanchesm@colorado.edu
linkedin.com/in/big-bang-data

Projects

bigbangdata.github.io
rpubs.com/BigBangData

Tools

R, Python, SQL, Git
Spark, Kafka, Elasticsearch
REST API, BI Tools (Sisense)
RStudio, Jupyter, SSMS/SSRS

Expertise

Dashboards
Data Analytics
Data Visualization
Machine Learning
Inferential Statistics
Predictive Modeling
Querying Databases
Business Intelligence
Unstructured (Big) Data
Natural Language Processing

Additional Skills

Curiosity
Tenacity
Creativity
Teamwork
Cooperation
Collaboration
Growth-Mindset
Detailed Insights
Clear and Concise
Interpersonal Skills
Responsibility & Ownership
Cross-Functional Teamwork
Analytical & Critical Thinking

Experience

01/20 – 03/21 **KPA, Lafayette, Colorado**   **Data Science Engineer**

Launched new integrated BI suite of dashboards for 10,000s of clients. Became subject matter expert on data models and dashboards. Transitioned into product owner role, helping pave the way for future analytics efforts.

- ➤ Supported hyper-care triage and resolution of Zendesk tickets
- ➤ Designed ETL flows for decades of legacy data, moving C# logic to SQL
- ➤ Created high accuracy data validation pipelines for entire client base
- ➤ Built performant and robust data cubes for large, complex data models
- ➤ Collaborated with Product Data Scientist and UX director on BI platform
- ➤ Worked with DevOps and vendor on scaling and customization
- ➤ Sole owner and maintainer of data models and dashboards post launch
- ➤ Advocated for action-oriented tactical dashboards based on workflows
- ➤ Recommended cross-domain standardization to prioritize safety actions
- ➤ Identified applications for predictive analytics for incident prevention
- ➤ Collaborated with management and leadership on data analytics roadmap
- ➤ Provided ad-hoc support for SSRS reports for internal stakeholders
- ➤ Mapped old accident reporting structures to new, generalized incident summary and detailed regulatory compliance reports

07/19 – 12/20 **KPA, Lafayette, Colorado**   **Data Steward Intern**

Played critical role in broad company pivot to SaaS, wrangling 30 years of legacy data. Engaged in cross-functional capacity with multiple departments (finance, operations, client success, etc.), sanitized 1,000,000s of records belonging to 10,000s of clients for multimillion-dollar platform launch.

- ➤ Fully evaluated data infrastructure to recommend fixes
- ➤ Ran data governance meetings with executive supervision
- ➤ Analyzed multiple databases and CRM and ERP systems using SSMS / R
- ➤ Built data cleaning and matching pipelines in Python to handle big data
- ➤ Collaborated with the integration team to create a single source of truth
- ➤ Diagnosed and resolved matching, linking, hierarchical, duplication, and standardization issues across dozens of disparate sources

09/96 – 05/18 **US, India, Brazil**   **Freelance Classical Musician**

Professional cellist, educator, and collaborator with solo, chamber, and orchestral engagements throughout the world. Principal Cellist in the Fort Collins Symphony and Greeley Philharmonic, Assistant Principal in the Boulder Philharmonic. Performed for audiences of up to 20K. Frequent guest speaker and conductor for youth orchestras and amateur ensembles.

Education

07/18 – 05/19 Master of Science in Business Analytics – Security Track
University of Colorado Boulder | Leeds School of Business
Focus on translating data into strategic business decisions by integrating IS and analytics with computer science and statistics. Coursework covered data analysis and visualization, big data, parametric and nonparametric predictive and prescriptive modeling, machine learning, inferential statistics, multivariate and logistic regression, classification, database design, etc.

07/17 – 05/18 Data Science Specialization | Johns Hopkins University, Coursera
Covers concepts and tools throughout the entire data science pipeline, from asking the right kinds of questions to making inferences and publishing reproducible results with data and code.

01/18 – 05/18 Front Range Community College
Detailed course from data collection to goodness-of-fit tests.

08/17 – 02/18 Statistics with R Specialization | Duke University, Coursera
Focus on experimental design, probability, and data projects with R.

09/02 – 05/08 Doctor of Musical Arts, Master of Music | University of Colorado
09/95 – 05/99 Bachelor of Music | Oberlin College

Links

Sanitized code and slides
available upon request

Heroku App
GitHub

GitHub

Rpubs

Sanitized code and slides
available upon request

Rpubs

GitHub

Portfolio

Hunter Douglas Predictive Analytics Projects
Predicted product defect and value of prospective clients for product and CS teams. Achieved high precision with gradient boosting. Predicted significant drivers of gross margin per dealer segment with unsupervised clustering. Presented to leadership.

SMS Spam Detection App
Created end-to-end ML project focused on explainable AI. Launched Heroku app that predicts spam using natural language processing (NLP) and XGBoost.

Elasticsearch Recommender for Online Retail
Created an MVP recommender system for online retail using a dynamic, streaming service. A finals-week two-day project run via Linux terminal in a Docker container, written in Python using Kafka producers and consumers and Elasticsearch.

Exploratory Data Analysis and Visualization
Conducted automated EDA through interactive time series plots on JHU-CSSE data. Scraped pre-aggregated data, enriching with WHO country population data.

Targeted Marketing Analytics
Designed and executed analytics projects using the BRFSS data for targeted marketing. Employed ordinal logistic regression and K-means clustering to develop a wellness program.

Scraping Indeed Salaries
EDA on salaries from Indeed.com for data-related job salaries in the Denver Metro.

Python Distributed File System
Built a low-level DFS in a server-client architecture using TCP. An MVP run via command line, implements basic (not secure) authentication, reliability, and traffic optimization.