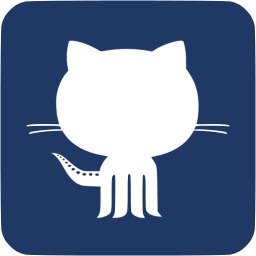
Marcelo Sanches, MS 303.931.2796 Lafayette, CO marcelo.sanches@colorado.edu

linkedin.com/in/big-bang.data bigbangdata.github.io rpubs.com/BigBangData

Insight-Oriented Data Scientist

Passionate about Extracting Insights using Data Intelligence and Driving Business Decision Making

Self-motivated Data Scientist / Data Analyst with 2 years of data science experience. Supported KPA’s (most recent employer) largest technology product launch in 30 years with a new integrated BI analytical platform and recognized by leadership for wide “breadth of ingenuity, patience, and technical acumen.” Brings a Master’s Degree in Business Analytics and a foundation in Python, R, and SQL as well as the discipline and collaborative spirit of a professional classical musician who successfully pivoted to data science in 2018.

* **Languages:** Python (Numpy, Pandas, Scikit-Learn), R (Tidyverse, data.table, knitr), SQL (MySQL, T-SQL), Git-Bash
* **Tools:** Linux, BI Tools (Sisense, Tableau), REST APIs, Spark, Kafka, Elasticsearch, RStudio, Jupyter Notebook, SSMS/SSRS, Markdown/LaTeX, Visio, Visual Studio
* **Expertise:** Quantitative and Qualitative Analysis, Data Analytics, Providing Detailed Insights, Dashboarding & Reporting, Machine Learning Algorithms, Data Modeling, Data Integrity, Data Management, Statistical Analysis

Core Competencies

Business Intelligence Machine Learning Natural Language Processing Data Visualization

Predictive Modeling Testing & Debugging Problem Solving & Troubleshooting Statistical Modeling

Experiment Design (A/B Tests) Prototyping Advanced Spreadsheet Manipulation Documentation

Writing SQL Statements Data Analysis Skills Best Practices & Actionable Insights BI Solutions

Additional Qualifications: Strong Verbal and Written Communication Skills, Clear and Concise, Ability to Communicate at all Levels of an Organization, Interpersonal Skills, Innovative, Attention to Detail, Analytical Skills, Curiosity, Growth-Mindset

Professional History

KPA Jul 2019 – Mar 2021

Data Science Engineer (Jan 2020 – Mar 2021) | Data Science Intern (Jul 2019 – Dec 2019)

EHS software company. Quickly brought on as Data Engineer after serving only five (5) months as intern. Played a critical role in supporting a broad company pivot to SaaS and wrangling / munging 30 years of legacy data in preparation for launch of new, integrated custom applications.

Data Governance: Worked in a cross-functional capacity with multiple departments (finance, operations, etc.) and sanitized 1,000,000s of records belonging to 10,000s of clients. Fully evaluated data processes to recommend new best practices.

* Analyzed data across multiple databases (Salesforce, NetSuite)
* Resolved account hierarchy, nomenclature, and linking issues with sources of truth
* Engaged in cross-functional teamwork on account merges, joining records, de-duplication, and quality assurance

Embedded BI Analytics: Successfully launched new integrated BI application after the full data cleanse.

* Created widgets in dashboards to reference/query data models
* Collaborated with Product Data Scientist on problem solving issues and incidents, and conducted inspections of data models / dashboards
* Standardized naming conventions and initiated quality assurance pipeline to prevent future record duplication, missing records, link errors, and hierarchy errors
* Supported completion of Zendesk tickets (people de-duplication, inactivation, transfers, and client merges

New BI Dashboard: Transitioned into designing robust dashboards for clients. Sole creator of data models. Recommended use of an action-oriented UI with real-time analytics.

* Transformed an off-the-shelf solution with customizations
* Added extra elements and dimension to visual data representations with a custom plugin that employed an ingenious sorting methodology
* Created a new stored procedure to handle previously problematic hierarchical computations

Data Analytics for Incidents: Mapped accident reports for incident summaries and injury/illness data, and set up SSRS reports to track new incidents via robust data model. Researched and identified applications for predictive analytics related to accident prevention.

Musician & Composer 1999 – 2018

Professional cellist and composer with solo and orchestral engagements throughout the world (US, Brazil, India). Principal Cellist in Fort Collins Symphony, Greeley Philharmonic, and Mahlerfest Orchestra; Assistant Principal in Boulder Chamber Orchestra. Performed for audiences of up to XXX. Frequent guest speaker/conductor for youth orchestras across the US.

Education

Master of Science in Business Analytics, Security Track | University of Colorado, Boulder, Leeds School of Business 2019

Focus on translating data into strategic business decisions by integrating IS and analytics with computer science and statistics. Coursework covered data mining and visualization, parametric and nonparametric predictive modeling, machine learning, inferential statistics, testing, multivariate and logistic regression, database design SQL, UML, data analytic methods, cleaning, munging, wrangling and manipulation, R, Python, and use of data and facts-based decision making via regression analysis.

Data Science Specialization Certificates | Johns Hopkins University via Coursera 2018

Statistics with R Specialization | Duke University via Coursera 2017

Courses: R Programming, Getting and Cleaning Data: *Samsung Sensor Activity* tidy data project, [GitHub](https://github.com/BigBangData/SamsungSensorActivity)**;** Exploratory Data Analysis: *Air Quality* project, [GitHub](https://github.com/BigBangData/ExploratoryDataAnalysis)**;** Reproducible Research: *The Impact of Harmful Weather in the U.S.,* [Rpubs](https://rpubs.com/BigBangData/347586)**;** Statistical Inference: a demonstration of the Central Limit Theorem, [Rpubs](https://rpubs.com/BigBangData/370442)**;** Introduction to Probability and Data

**Doctor of Musical Arts & Master of Music | University of Colorado Boulder, College of Music 2008**

**Bachelor of Music | Oberlin College 1999**

Projects

* Predictive analytics projects for Hunter Douglas’ product teams involving a cross-evaluation of models from logistic regression to GBM; produced a full-depth economic impact model to support decision making
* Created an Elasticsearch Recommender for use by online retailers; written in Python using Kafka producers and consumers; [GitHub](https://github.com/BigBangData/ElasticsearchRecommender)
* Designed and executed analytics projects using the BRFSS data for targeted marketing; used ordinal logistic regression and K-means clustering to develop a wellness program
* Developed a Python-distributed file system in a server-client architecture using TCP socket programming; [GitHub](https://github.com/BigBangData/PythonDFS)