JOSEPH JAVIER

DATA SCIENTIST / ANALYST, MACHINE LEARNING ENGINEER, PYTHON PROGRAMMER, DATA MANAGEMENT CONSULTANT, AND REMOTE WORK PRACTITIONER

242 12th Street SE, Apt 1/2
Washington, DC 20003
Cell: (703) 405-5163
Email: joseph.javier520@gmail.com
LinkedIn: /joseph-javier-data-lover
Github: /jjavier1/Springboard-Projects

SUMMARY

I am a certified Data Scientist/Machine Learning Engineer seasoned in data analysis and client relationship management experience in the federal government consulting sector of the DC Metropolitan Area. For almost 5 years, I'm known for working in a collaborating, yet fast-paced environment while being consistent in delivering accurate, clear data visualizations/models that focus on driving sound judgement and solving problems for my key clients. By education, I completed a Data Science and Machine Learning certification program with 600+ hours of hands-on learning, 10 Python/SQL programming mini-projects, and two capstone projects while managed by 1:1 industry expert mentor oversight. This program has driven me to master skills in Python 3 programming, Git/Github version control, data ingestion, data processing, data engineering, basic DevOps, model and system architectures, developing algorithms, ETL, and model development & deployment through Docker along with applications such as artificial intelligence (AI) and natural language processing (NLP). I also hold a Bachelor's in Math where I have taken Computer Science related courses such as Data Structures, Fundamental Java, and Fundamental MATLAB. Professionally, I am also certified in AWS Cloud Computing and Cost Estimating.

PROFESSIONAL EXPERIENCE

SUPPLY CHAIN DATA ANALYST (REMOTE POSITION)

TECOLOTE RESEARCH INC., WASHINGTON, D.C. AND REMOTE, JAN 2020 – PRESENT

- Wrote JSL and Excel VBA code to automate spreadsheet calculations, automate data cleaning based on filtering irrelevant keywords, non-numeric entries, and verifiable outliers, plus reduced time required to work on deliverables from 2-3 days to 45-60 minutes
- Collaborate with a multi-disciplinary team to research approximately 500 suppliers through MorningStar, Bloomberg, Glassdoor, LinkedIn, and company search engines containing company information
- Facilitate team communication, file version control, and time management through Slack, Zoom, MS Outlook, MS SharePoint, and KiteWorks to seamlessly work remotely without sacrificing quality
- Used MS Excel, MS Access, and JMP to analyze gathered supplier data to calculate scores, rankings, and other comparison results determining which suppliers may pose a risk to future client purchasing based on credit risk, compliance issues, delivery/scheduling concerns, and corporate events

MANAGEMENT DATA ANALYST (PARTIAL REMOTE POSITION)

TECOLOTE RESEARCH INC., WASHINGTON, D.C. AND REMOTE, MAY 2019 – JAN 2020

- Wrote Excel VBA code to automate data pulled from source data files, EVM analyses, and data visualization development reducing deliverable work time from 3-4 days to 5-6 hours
- Provide cost analysis and estimate supporting labor to Navy shipbuilding programs worth between $1-5 billion based on new affordability assessments, Government / Contractor Estimates At Completion (EACs), and tradeoff studies/assessments
- Troubleshoot issues/concerns with shipbuilding vendor source data using anomaly detection and pattern detection methods with Navy client engineers on recommendations for sound judgment on further action
- Delivered data visualizations, formal reports in MS Word, and monthly performance briefings MS PowerPoint to advise program/project leadership on cost, schedule, risk, federal contract compliance, and performance trends in addition to reporting and solving problems on any significant shipbuilding issues and deficiencies

FINANCIAL ANALYST

TECOLOTE RESEARCH INC., ARLINGTON, VA, OCT 2018 – APR 2019

- Estimated the cost of converging 20-30 individual Navy IT Systems to 4 individual web applications based on historical data, analogous hardware, client accounting / bookkeeping, and financial operations
- Applied cost and budget data analytics to generate cost and spending growth modeling, respectively
- Delivered MS PowerPoint presentations advising decisive action and next steps based on previous cost estimates, and budget funding worth anywhere between $2-5 billion

JUNIOR COST/BUDGET ANALYST
BOOZ ALLEN HAMILTON, WASHINGTON, DC, OCT 2016 – MAY 2018

- Managed cost and budget portfolio for 7-10 Naval hardware systems programs in charge of development, operation, maintenance, and disposal
- Collaborated with hardware system engineers to estimate and project costs ranging around $200-300 million per new/updated version of hardware systems based on analogous hardware cost, previous Congressional budget funding submissions, and client expense management
- Oversaw contract-specific and sensitive information to ensure fair competition during the bidding process

EDUCATION AND CREDENTIALS

DATA SCIENCE CERTIFICATION, OCTOBER 2020
Springboard, Online (Remote)

CERTIFIED CLOUD PRACTITIONER, AUGUST 2020
Amazon Web Services, Online (Remote)

CERTIFIED COST ESTIMATOR/ANALYST, 2020
International Cost Estimating and Analysis Association, Annandale, Virginia

BACHELOR OF ARTS IN MATHEMATICS, 2015
University of Maryland Baltimore County, Catonsville, Maryland

PROJECTS

RNN PORTFOLIO STRATEGY CAPSTONE PROJECT
Implements an Recurrent Neural Network, also known as RNN, with a simple architecture of Long-Short Term Memory, also known as LSTM, cells to generate trading signals for selected well-known stocks and indices. The RNN was assembled and compiled in Python using Keras with Tensorflow as the backend library. The Pandas library was also used to structure the input and output data. URL: https://bit.ly/36FuVvE

SUPER ZIPS CAPSTONE PROJECT
Models which areas of the U.S. are becoming more or less impoverished and educated over time using descriptive statistics, exploratory data analysis, and K-Means Clustering. The Pandas and Numpy libraries were used to generate statistics to determine data validity and integrity, while the Seaborn and Matplotlib libraries created numerous visualizations to obtain a full understanding of the huge datasets sourced from the IRS and Census Bureau. URL: https://bit.ly/3jxcpJk

BASIC TEXT CLASSIFICATION WITH NAIVE BAYES
Shows the basics of text analysis by performing Naive Bayes Classification on a subset of movie reviews from the Rotten Tomatoes database. The Scikit-Learn and Scipy libraries were used to explain the text classification process and model the performance of the Naïve Bayes classifier. URL: https://bit.ly/2SvzBM2

SPARK MINI-PROJECT: DATABRICKS EDITION
This mini project provides a basic overview on Spark Datasets & Dataframes, showcases some Spark SQL, and builds a couple of binary classification models using the Spark ML library (with some code from the MLlib library, too). This also implements a Databricks Spark cluster to introduce the use of cluster/distributed computing unlike compared to parallel computing. URL: https://bit.ly/3nPMaQX

SQL QUERY MINI-PROJECT
This project shows SQL code required to query datasets from the Springboard database platform for a number of cases. In the beginning, the queries are simple and broad, but toward the end, the queries are more complex as required for more detailed and specific datasets. URL: https://bit.ly/2SBrBt2