

Data Scientist Data Scientist Data Scientist - Capital One Henrico, VA Migrating inhouse projects onto cloud like Snowflake, Redshift, S3 and Databricks. Familiarity and comfort with most aspects of Python3 including object-oriented programming and the data science stack, wrote Python routines to log into the websites and fetch data for selected options. Wrote flexible production-ready code (in R or Python) for different components of the data science process such as data ingestion, data manipulation, feature selection & engineering, model training, model validation, and model deployment. Good experience working with Financial data and creating calculated fields extracted from diverse sources. Deep understanding of modern machine learning techniques and their mathematical underpinnings, such as classification, recommendation systems, and natural language processing. Strong understanding and application of statistical methods and skills like finding Central tendency, a measure of dispersion using various distributions, experimental design, variance analysis, A/B testing, and regression. Exposed solving both classification and regression problems from messy data using Machine Learning algorithms. Experienced building Natural Language Processing models using Linear classifiers like fastText. Delivered business value by translating complex data from BMC remedy tool into meaningful insights using visualization tools (Tableau). Proficient with data visualization and translating complex problems into actionable insights. Implemented python to retrieve and manipulate data, visualize the data for key findings and explained to clients. Worked on IO operations of different file formats like csv, excel, json, txt using python. Good experience with PySpark mainly for getting data from diverse sources for data analysis. Worked with business groups and stakeholders to ensure models can be implemented as part of a delivered solution replicable across many clients with no impact on business. Presented key findings to stakeholders to drive improvements and solutions from concept through to delivery. Displayed strong teamwork and interpersonal skills with the ability to communicate to all levels of management. Implemented SQL Scripts, Stored Procedures and Triggers in MySQL server to handle user requests and work with the data in the database. Experienced documenting all modeling steps in a systematic way including modeling process, insights generated, model validation results and checklists built in the project. Develop software to automate operational

processes along with coding for the shared engineering deliverables. Experience in configuring, deploying the web applications on AWS servers. Performed configuration, deployment and support of cloud services including Amazon Web Services (AWS). Created a database using MySQL, wrote several queries to extract data from the database. Wrote and executed various MYSQL database queries from Python using Python -MySQL connector and MySQL. Involved in the development of Web Services using SOAP for sending and getting data from the external interface. Provide technical guidance, recommendations, and resolutions to new team members and built a new team. Good exposure and business knowledge on AWS services like S3, EC2, Amazon Redshift, Kinesis and ML Azure.

Work Experience Data Scientist Capital One February 2019 to Present

Migrating in-house databases onto AWS cloud, Snowflake and running the jobs on Databricks. Identify, analyze, and interpret patterns in complex data sets using Databricks and Tableau. Developing python models to identify High-Risk Customers and performing comparative analysis over time to highlight High-Risk Customers variations. Creating a Data Pipeline on AWS cloud which allows flexible data movement between different data sources. Moving SQL and Oracle compliance data onto AWS redshift and performing scaling and data validity. Using python database connectors to connect AWS redshift and perform SQL operations from CLI. Creating dashboards, connecting DynamoDB to Tableau and analyzing the changes in realtime. Good experience in Data Blending, Data Joining, Applying Filters and proving granularity in Tableau. Experience using python Logger module to keep track of different warnings, info, debug messages when connected to databases like DynamoDB, SQL and Redshift. This info is basically stored as a log file.

Data Scientist TAMUC - Commerce, TX May 2017 to January 2019

Recruited to work as a Data Scientist to do requirement gathering and hypothesis testing for new implementations in websites. I was also responsible to perform statistical analysis on the datasets we receive and find the relationship between the features and target variables. I also worked on a project Non-Toxic text classification to stop cyberbullying over the social media network. We implemented using fastText linear classifier algorithm and tested its accuracy using other Machine Learning algorithms. I also worked with Dr. Isaac Gang as a Data Engineer to maintain "We Teach CS Collaborative" database.

Key Contributions: Scrapped data from various websites, cleaned and processed unstructured data to structured data and performed exploratory data analysis using python libraries like Pandas, NumPy, Matplotlib, Seaborn. Translated complex unstructured data and analysed the results into compelling visualizations and stories that clearly articulated the strategy and follow-up actions to the stakeholders. Downloaded Twitter tweets using Twitter API, trained manually, segregated positive and negative tweets. Performed hypothesis testing and statistical analysis of various data sets using python libraries and mathematically to find the acceptance and rejection percentage on the null hypothesis and alternate hypothesis. Used Supervised Learning Algorithms and Natural Language Processing to analyze the level of toxicity in the text, particularly in social networks and stop cyberbullying. Worked as Data Engineer for "Texas A&M University-Commerce We Teach_CS Collaborative" funded by UTA. Created dashboards and reports to regularly communicate results and monitor key metrics. Worked closely with various departments to track the survey results and find the insights. Responsible to Monitor Computer Science Department Website changes and ICEL Server. Analyzed the clickstream data and reported the findings to all the departments in form of charts. Conducted text analysis on large datasets of unstructured text and performed statistical analysis using a variety of statistical tools like Minitab and python. Data Analyst/ Python Developer Hewlett Packard Enterprise - Bengaluru, Karnataka August 2014 to December 2016 Key Contributions: Worked independently on complex data analysis, business analytics & data mining tasks on big datasets. Analyzed structured and unstructured data from incidents with complex statistical analysis to find the incident origin and customer behavior towards the incidents. Experienced in collecting data from multiple databases and websites, validated data integrity and accuracy. Leveraged data and performed intensive analysis across all areas of our business to drive growth strategies, including product development and rider engagement strategies. Communicated with clients and store managers in a clear and concise manner for actionable changes focused on improving velocity, predictability, and efficiency of each store. Developed and implemented a database monitoring system web page automation using FLASK and Django which reduced 4 hours of AMOS team manual work. Used legacy mainframes, SQL, Unix

Scripting, Autosys, BMC remedy as part of project enhancement. Responsible for all aspects of the software development life cycle for the applicable projects, including gathering requirements, design, implementation, and deployment. Automated many AMOS tasks reducing 40% manual effort and 30% ticket count increasing efficiency. Experienced using databases such as Postgres and MySQL. Wrote Python routines to log into the websites and fetch data for selected options Built SQL queries to perform various CRUD operations with DDL, DML, DCL and TCL commands.

~~~~~  
~~~~~ HONORS and AWARDS ? Excellence Award- Awarded by Mr. Ben Wishart, AHOLD CIO, for maintaining a good relationship with clients and stakeholders. ? Star Performer- Awarded by Ravi Kumar, AHOLD Project Manager, Hewlett Packard. ? Conduct of Appreciation- Awarded by Dr. Isaac Gang, founder of "WeTeachCS Texas A&M Collaborative" for working as a Data Engineer in this project. ? Texas A&M University Computer Science Scholarship- Awarded for 2 years of master's education for showing good academic and research work on campus. Education M.S. in Computer Science and Information Systems. in Commerce Texas A&M University 2018 Bachelors in Electronics and Communication Engineering in Electronics and Communication Engineering Jawaharlal Nehru Technological University 2014 Skills testing, Excel, Business Intelligence, access, SQL Additional Information Areas of expertise include: \* Python \* PySpark \* SQL \* Tableau \* Minitab \* Statistical Modeling \* Natural Language Processing \* Snowflake \* Pandas \* NumPy \* ScikitLearn \* Keras \* Matplotlib \* NLTK \* PostgreSQL \* AWS (EC2, Redshift, S3 etc.) \* Machine Learning Algorithms \* Decision Trees \* Random Forest \* RNN \* Gradient Descent \* Na ve Bayes \* Logistic Regression \* Collaborative Filtering

Name: Taylor Gomez

Email: loricolon@example.org

Phone: 7204632611