**BRITTANY MARRA**

Data Analyst

**Location: Las Vegas, Nevada | Mobile: (860)-912-6302 | Email: brittanymarra2015@gmail.com**

**Skills**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Coding | **Technologies & Tools** | | **Management** | **Visualization** | **Other** |
| Python | Teradata | MS Access | JIRA | PowerBI | Actuarial: VEE Economics, VEE Statistics Cert |
| R | Snowflake | Salesforce | Agile Methodology | Tableau | Machine Learning |
| SAS | Google Big Query | Excel | GitHub |  | Data Analysis |
| SQL: MySQL, NoSQL | Azure Synapse |  | Trello / Mural |  |  |
|  |  |  | Service Now |  |  |

**Training**

Databricks, Python, Alteryx

**Work Experience**

**Data and Analytics Consultant, *NTT DATA – Las Vegas, NV*** July 2021- Present

*Data Analyst:*

*Healthcare, Manufacturing, and Insurance Industry Clients:*

* Collaborate with stakeholders to define and document business and technical requirements.
* Create and Collect business requirements from cross-functional teams and translate them into data-driven solutions.
* Develop comprehensive strategy, including data mapping, extraction, transformation, and data loading.
* Create data quality and data lineage strategy frameworks, guidelines, and standards to establish a suitable data governance framework.
* Follow data modeling and data handling procedures to maintain compliance with applicable laws and policies across assigned workstreams.
* Experience with managing codebases, pipelines, and underlying data table architectures.
* Responsible for data validation and testing in QA, UAT, and Production environments.
* Validate reporting and data accuracy.
* Create intuitive and interactive dashboards and reports using data visualization tools to facilitate data exploration.
* Accomplishments:
  + Contributed on a data migration project from Teradata to Snowflake for a large-scale enterprise, migrating over 10 TB of data from legacy systems to a cloud-based platform.
  + Utilized web scraping techniques via Python to extract and organize data, resulting in an increase of operational efficiency by 25%
  + Experience with SnowSQL for Teradata to Snowflake conversion, enhanced data quality through query optimization to result in reduction of data anomalies.
  + Develop SnowSQL/Snowflake stored procedures, functions, views, tables, and other database objects as defined by business partner requirements to support overall corporate objectives.
  + Conduct post-migration data audits and service now analysis.

*Data Scientist:*

* Analyze large datasets for exploratory data analysis and leverage Python & SQL for data mining, data profiling, and predictive analytics.
* Modeling:
  + Binary Classification: Logistic Regression, K Nearest Neighbors
  + A/B Testing and Experimental Design
  + Decision Trees: CART, Decision Tree Classification, and Random Forest
  + Gradient Boosting: Catboost, XGBoost, Stochastic Gradient Descent
  + Time Series: AR, MA, ARIMA, ARIMAX
  + Model Evaluation with Shapley values, ROC curves, confusion matrices, and permutation importance.
* Optimize and deploy models
* Leverage and Deploy python code packages: Numpy, Pandas, Matplotlib, SciPy, and Scikit-learn via Github and Jupyter Notebooks: <https://github.com/bmar0821>
* Strong Understanding of Economic principles and Pricing theories
* Accomplishments:
  + Developed artificial intelligence-driven fraud detection model using XGBoost, achieving a 76% accuracy rate in reducing fraudulent transactions.
  + Utilized LGBM model for MLB predictive pitch, attaining a 63% accuracy rate.
  + Engineered a hospital readmissions model employing XGBoost, resulting in a 73% reduction in fraudulent transactions.
  + Designed an auto-prune CART model, generating segmentations for A/B testing to assess the impact of a client's new pricing tool on annual revenue.
  + Trust & Ethics Document, inclusive of ethical frameworks, fairness metrics, and data privacy regulations

**Data Scientist, *IBM* – *Baton Rouge,* LA** June 2019- July 2021

*Retail and Government Industries:*

* Experience with data extraction, manipulation, and blending across different data sources.
* Use of syndicated POS data for attributed waste management model.
* Apply product knowledge for data validation.
* Create and execute MySQL queries via Kubernetes pods.
* Deploy IoT devices and train visualization models.
* Leverage IBM Watson Technology and build multi-node Kubernetes clusters, which include user interfaces on Node-RED, R, and Python analytical scripts.
* Accomplishments:
  + Collaborate with cross functional teams to develop predictive models, resulting in an 18% improvement in demand forecast accuracy
  + Built time series; ARIMA models, that reduced excess inventory by 14% and stock shortages by 6%

**High Level Projects**

Medicare Claims Fraud Prediction, *NTTDATA* April 2024 to Present

International Hospitality Revenue Performance Clustering, *NTTDATA* October 2023 to March 2024

National Health System Data Governance Organizational Design*, NTTDATA* July 2022 to October 2023

National Insurance Company Data Migration,*NTTDATA* August 2021 to June 2022

‘Return to Work’ Kiosk Initiative, *Granted by NIH and the CDC*, IBM March 2020 to July 2021

AI Restaurant Waste Management, *IBM*August 2019 to March 2020

**Education**

**University of Connecticut*, Storrs, CT*** Graduated

B.A Mathematics and Statistics

GPA 3.68