

RUBINA SHRIVASTAVA

Nashville, TN

Phone: 615 693 3107 | Email: rubina.shrivastava@gmail.com

LinkedIn: <https://www.linkedin.com/in/rubina-shrivastava-38332024/>

SUMMARY

Diligent Data Analyst with strong Python. Seeking an opportunity to leverage my diverse skills and experience in data mining and analysis into the open data analyst position at your company. A B.S. in Statistics from a top 50 university coupled with my ample experience in analytics will make me a positive asset to your organization.

TECHNICAL SKILLS

Software/Platforms: Jupyter, Google CoLab, Github, Visual Studio Code, Excel, Google Maps, Graphs, GeoJSON,

Software Languages: Python (Pandas, Numpy, Matplotlib, sci-kit learn), MS SQL, Postgres, Tableau, MS Excel, Java Script (D3, Plotly), HTML, CSS, Machine Learning (Neural Network, Tensorflow, Deep Neural Network, K-means, centroids, elbow curve)

Data Analytics – Large Scale data cleansing, ETL, and analysis

Projects

(Pls refer to the links in project names to access respective GitHub repos)

1. [Mapping Earthquakes](#) - The purpose of this project is to visually show the differences between the magnitudes of earthquakes all over the world over the last seven days, to complete this project, we use an API/Web Service for GeoJSON earthquake data from the USGS website

Responsibilities - Developer

Technologies - GeoJSON, API, D3 javascript

2. [School District Analysis](#) – The purpose of our analysis is to find the overall passing percentages of the students and their performance according to schools, then we have found that some of the test scores from the 9th grade at Thomas High School are missing. we want to replace this with NaN, and after that, we do our analysis with this new change to take all of the scores from the high school so that it doesn't affect our overall numbers.

Responsibilities - Developer

Technologies - Python, Cleaning DataFrames, Groupby(), Pandas, Numpy

3. [Election Analysis](#) – After performing analysis, our script writes the results to a text file, which will be sent to the election commission. By adding any other election result CSV with the same format the as the resource to our script we can use it for other election

Responsibilities - Developer

Technologies - Python, Boolean & Logical operators, Calculations using data types

4. [Kickstarter Analysis](#) - To calculate and visualize the outcomes of successful, canceled, and failed based on launch date and goal

Responsibilities - Developer

Technologies - Excel, Graphs, Analysis, Interpretation, Statistics

5. [Stock Analysis](#) - By using Refactor code we can make the code more efficient by taking fewer steps, using less memory, or improving the logic of the code to make it easier for future users to read

Responsibilities - Developer

Technologies - VBA, Loops, Optimization

6. [Neural Network Charity Analysis](#) - The purpose of this analysis was to explore and implement neural networks using TensorFlow in Python. A neural network is an advanced form of Machine Learning that can recognize patterns and features in the data

Responsibilities - Developer

Technologies - neural network model, Machine Learning classification, TensorFlow, Deep Neural Networks

7. World Weather Analysis - The purpose of this challenge was to improve our app, so customers will be able to use input statements to filter the data for their weather preferences, which will be used to identify potential travel destinations and nearby hotels

Responsibilities - Developer

Technologies - Python libraries, API, JSON, Matplotlib, Google Maps

8. Cryptocurrencies - The purpose of this analysis was to use data from CryptoCompare to provide a report that includes, what cryptocurrencies are on the trading market and how they could be grouped to create a classification system, for this new investment and visualization of currently traded cryptocurrencies that can be grouped to create a new classification system. This report would be used to help Accountability Accounting, offers a new investment portfolio in the exciting world of cryptocurrency to its customers

Responsibilities - Developer

Technologies - Supervised and Unsupervised Learning, Cluster data, K-means algorithm, centroids, elbow curve, PCA

9. Credit Risk Analysis - Credit risk is an inherently unbalanced classification problem, as good loans easily outnumber risky loans. In this project, we want to take a look at how all the factors in our "loanstats.csv", help predict whether someone is low or high-risk status.

Responsibilities - Developer

Technologies - Supervised Machine Learning, sci-kit learn, Random Over Sampler, SMOTE, Cluster Centroids, SMOTEEN, Balanced Random Forest Classifier, Easy Ensemble Classifier

10. Bikesharing - An additional visualization that may help investors determine if Citibikes are a good fit for Des Moines includes correlating weather data between Des Moines and New York City. Both cities have four seasons, and if their temperatures and weather are similar, then this may also provide a stronger case to investors that Citibikes is a great fit for Des Moines, Iowa

Responsibilities - Developer

Technologies - Tableau, dashboards, storytelling

11. Plotly Biodiversity - In this challenge, we are finishing our partially completed dashboard for demographic information and visualizing the bacterial data for each volunteer. Specifically, our volunteers should be able to identify the top 10 bacterial species in their belly buttons

Responsibilities - Developer

Technologies - D3 JSON, Plotly, Parsing, JavaScript Math, event handlers, GitHub Pages

12. UFOs - Our webpage and dynamic table are working as intended, the purpose of this analysis is to provide a more in-depth analysis of UFO sightings by allowing users to filter for multiple criteria at the same time. In addition to the date, you can add table filters for the city, state, country, and shape

Responsibilities - Developer

Technologies - JavaScript, arrow functions, loops, HTML, CSS

13. Surf's Up - We are looking for weather patterns in Hawaii to open a Surf and Ice Cream shop. We choose, the months of June and December in Oahu, to determine if the surf and ice cream shop business is sustainable year-round

Responsibilities - Developer

Technologies - SQL, SQL Alchemy, SQLite, PostgreSQL, Flask, Python

14. Movies ETL - Amazing Prime loves the dataset and wants to keep it updated daily for that we have to create an automated pipeline that takes in new data, performs the appropriate transformations, and loads the data into existing tables.

Responsibilities - Developer

Technologies - ETL, SQL, Python, Pandas, Regex, PostgreSQL, Kaggle, JSON, Parsing

15. Pewlett Packard Analysis - The purpose of this analysis is to prepare Pewlett-Hackard, a large company, for the upcoming "Silver Tsunami", for this we determine the number of retiring employees per title, and identify employees who are eligible to participate in a mentorship program and current employees reach retirement age. The criterion was based on the birth dates ranging from 1952 to 1955 and hired dates from 1985 to 1988.

Responsibilities - Developer

Technologies - ERD, SQL database, pgAdmin, joins, queries

16. Pyper Analysis - Analysed and created visualizations using rideshare data to help improve access to rideshare services and determine affordability for underserved areas. We are creating a multiple-line graph that shows the total weekly fares for each city type. The PyBer Summary DataFrame provides an overview comparison of PyBer's ridesharing services in three types of cities: Rural, Suburban, and Urban cities

Responsibilities - Developer

Technologies - Matplotlib, charts, Numpy, Pandas, SciPy, Python

17. Good Eats - The purpose of this analysis is to predict if a Thai-Indian fusion restaurant would be successful in the San Francisco, CA area

Responsibilities - Developer

Technologies - GeoJSON, API, Python, Cleaning DataFrames, Groupby(), Pandas, Numpy, Supervised and Unsupervised Learning, Microsoft visual studio

EDUCATION

Data Analysis Bootcamp,

Vanderbilt University, Nashville, TN

Nov 2022

An 6 months intensive boot camp focused on gaining skills in data analytics, python, machine learning, Tableau, and Javascript/HTML.

Bachelor of Computer Engineering and Science,

Jabalpur Engineering College (RGPV University)

Jabalpur, India

2000 – 2004