

As a trainee, I embarked on this journey with a fresh perspective, being new to the field of data analysis. However, through the tasks assigned to me, I am steadily gaining familiarity and confidence in this domain.

I have been setting up a Python environment, learning to use Git for version control, and understanding the implementation of CI/CD. I am also getting acquainted with Key Performance Indicators (KPIs) and their role in measuring the success of a project.

In the project planning phase, I have been introduced to the CRISP-DM Framework, a structured approach to data mining projects. Through this, I am not only understanding data but also learning to explore it using exploratory data analysis (EDA) and statistical thinking. The KPIs for this phase have been guiding me to demonstrate proactivity to self-learn, apply EDA techniques to understand data and discover insights, and use suitable statistical distributions and plots to demonstrate my understanding of statistics.

As part of the hands-on experience, I have created a GitHub repository to host all the code for the project. I have also created a new branch called "task-1" for my analysis. I am committed to committing my work at least three times a day with descriptive commit messages.

I have been modifying the starter code to make it suitable for the news data that I am working with. This involves changing class and function names, and moving non-plotting functions into separate Python files.

For the EDA analysis, I am answering several questions about the data, such as identifying the top and bottom 10 websites in terms of news article count and visitor traffic, and identifying the countries with the highest number of news media organizations. I am also analyzing the sentiment of the news articles and comparing the impact of using mean/average and median.

Finally, I am comparing the content metadata across sites, checking the distribution of raw message lengths and number of words in the title across sites, and investigating the impact of frequent news reporting and sentiment on a website's global ranking.

Although I started as a novice in data analysis, the practical tasks and learning opportunities have helped me get familiar with the field. I am excited about the progress I have made and look forward to further enhancing my skills in data analysis.