In my journey through LIS 4761 Intro Data/Text Mining and LIS 4273 Adv Stats & Analytics, two assignments have significantly influenced my preparation for a career in information-related organizations. The sentiment analysis final in LIS 4761 provided an application of data mining techniques within a real-world context. By dissecting a live thread on r/WorldNews discussing the 2023 Israel-Palestine Conflict, I gained skills crucial for data acquisition, preprocessing, and sentiment analysis. In addition to this, the final of e-commerce customer spending behavior in LIS 4273 allowed me to explore the relationship between customer attributes and spending patterns. These assignments were instrumental in equipping me with the skills and perspectives necessary for success in my future career.

## First Assignment

In LIS 4761 Intro Data/Text Mining, one of the assignments that significantly impacted my preparation for a career in information-related organizations was the sentiment analysis final on the r/WorldNews live thread discussing the 2023 Israel-Palestine Conflict. This assignment provided a practical application of data mining techniques in a real-world context. By scraping Reddit comments using an API, my team and I gained valuable experience in acquiring lexicon data from Reddit, a skill crucial for information-related roles. Furthermore, analyzing sentiment from online discussions equipped me with essential text mining skills, including data cleaning, preprocessing, and sentiment classification. These skills are highly relevant for roles involving content analysis, market research, or customer feedback analysis, where extracting insights from unstructured text data is essential.

Through this assignment, I gained a deeper understanding of social dynamics and public opinion, particularly in serious topics like the Israel-Palestine conflict. By examining sentiment variations over time and across different subgroups, I developed insights that are invaluable for

organizations seeking to understand public sentiment or identify emerging trends. Additionally, this project enhanced my programming skills, as I utilized PRAW (Python Reddit API Wrapper) and wrote Python code to interact with the Reddit API. Proficiency in programming, especially in Python, is critical in information-related roles for data manipulation, analysis, and automation tasks.

## Second Assignment

In LIS 4273 Adv Stats & Analytics, I sought to find the factors that influence e-commerce customer spending behavior. The motivation behind this choice stemmed from my prior internship and work experience in e-commerce. The project aimed to understand the relationship between various customer attributes such as age, gender, city, membership type, average rating, discount applied, days since last purchase, and satisfaction level, and the total amount spent by customers. To visually represent the relationships among variables, I opted to create a correlation heatmap matrix. This visualization technique highlighted strong correlations between the rating of the purchase, total amount spent, and number of items purchased. Notably, customer satisfaction, as reflected in the purchase rating, exhibited a positive link to both the total amount spent and the number of items purchased, aligning with the alternative hypothesis.

These assignments were pivotal in shaping my understanding and skill set in the field of data analytics and information-related roles. The sentiment analysis project enabled me to delve into real-world data, applying techniques such as data scraping, preprocessing, and sentiment analysis to extract valuable insights from online discussions. This experience not only honed my technical skills but also provided a practical understanding of social dynamics and public opinion analysis. Similarly, the advanced statistics and analytics final project equipped me with the

ability to conduct multivariate regression analysis and visualize correlations, offering a deeper understanding of customer behavior and spending patterns in the e-commerce sector.