Formula 1 Driver/Team Stats Helper

Theme 5: Free Topics

1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.
    a. **Captian:** Alex Che (Alexmc2)
    b. **NetIDs (UIUC Email):**
        i. Alexmc2 (alexmc2@illinois.edu)
        ii. Savyak2 (savyak2@illinois.edu)
        iii. Xcai17 (xcai17@illinois.edu)
        iv. scw9 (scw9@illinois.edu)
        v. Zeijiajc2 (zejiajc2@illinois.edu)

2. What is your free topic? Please give a detailed description. What is the task? Why is it important or interesting? What is your planned approach? What tools, systems, or datasets are involved? What is the expected outcome? How are you going to evaluate your work?

    a. **Free topic:** F1 Driver statistics
    b. **Topic Description & Motivations:** Formula One is a fast-growing sport that has been around for the past 70+ years. With Netflix's recent show "Drive to Survive", this sport has gained many fans. Formula One fans often enjoy comparing drivers from different teams or on the same team. A website that provides these insights can facilitate these comparisons, letting users assess the performance of past and current drivers or even predict future trends.
    Formula One, a sport with a rich history spanning over 70 years, has recently experienced a surge in popularity, especially with the advent of Netflix's "Drive to Survive". As newcomers find their way into this intricate world, the myriad technicalities, strategies, and histories can be overwhelming. This platform not only serves as an invaluable learning tool to demystify these complexities but also amplifies the fan experience at its core. By offering data, analysis, and insights, it enables enthusiasts to engage in informed debates, make predictions, and revel in a deeper understanding of F1. Fans can also compare drivers, whether from opposing teams or within the same lineup, and this project is poised to facilitate such comparisons, letting users evaluate and project drivers' performances.
    c. **Approach & Expected Outcome:** We plan on making a website that can access and collect live data from several official Formula One websites. This allows the users to have a complete overview of all the statistics of their driver, and they

should be able to compare statistics of their own choice against other drivers on the website.

d. **Tools, Systems, Datasets:** We plan to use Jupiter Notebook and VSCode for all the web crawlings and website creation. **BeautifulSoup** and **request** will be used for Web Scraping. **Flask** will be used for hosting local stat websites.
    i. **Datasets**:
        1. https://www.formula1.com/
        2. https://www.espn.com/f1/
        3. https://ergast.com/mrd/gallery/
        4. https://www.motorsport.com/
        5. https://www.statsf1.com/en/default.aspx

e. **Work Evaluation:** We want to evaluate our work by web traffic metrics, the overall speed of retrieving data, the content quality and accuracy delivered to the users, and the user feedback.

3. Which programming language do you plan to use?
    a. **Python**
4. Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed and the estimated time cost for each task.
    a. Five members == **100 hours of work-load**
        i. **40 hours** – Web Crawling setup & extract website information
        ii. **30 hours** – Website creation and hosting of local website
        iii. **20 hours** – Creating statistics and data visualizations
        iv. **10 hours** – Setup feedback collection system