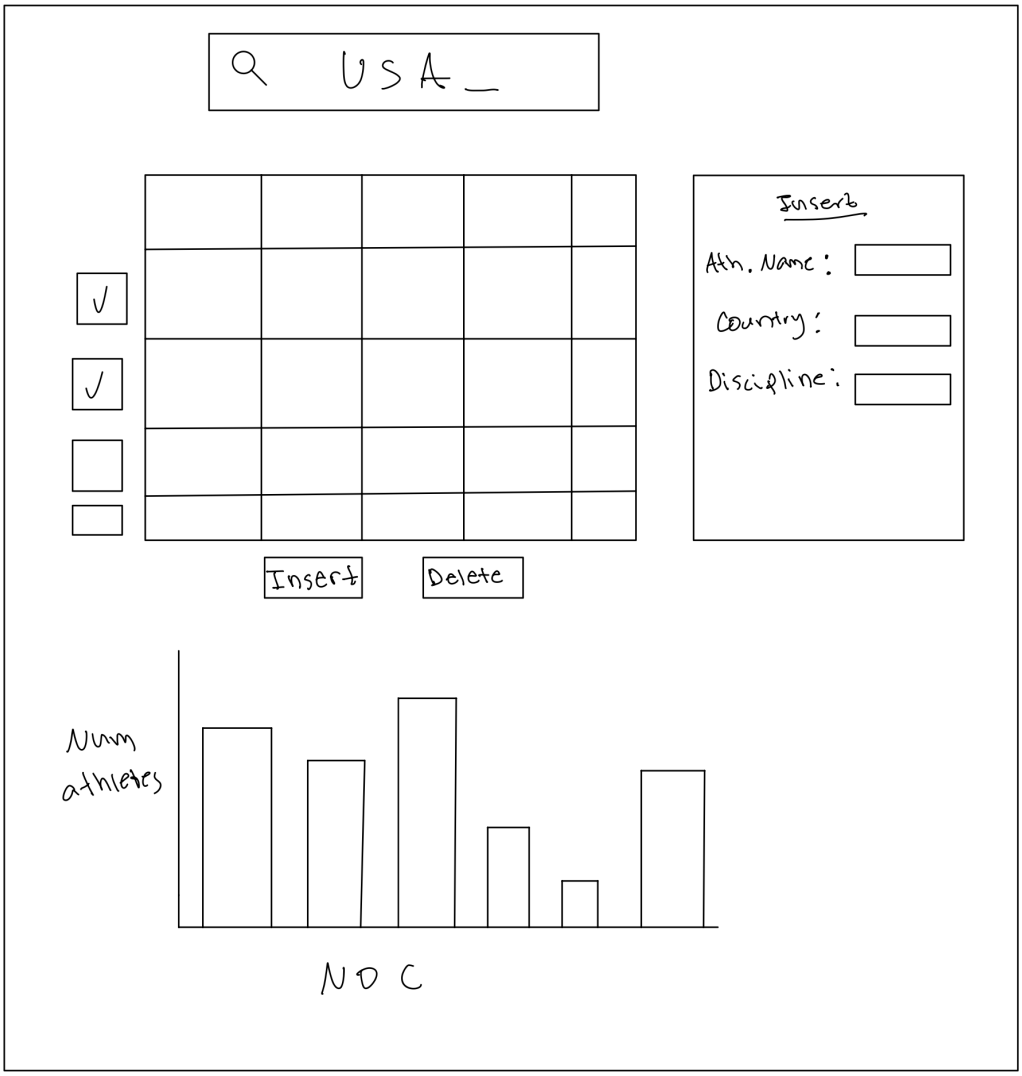
Tokyo Olympics Wiki Project Proposal

1. The data stored in the database contains the details of over 11,000 athletes, with 47 disciplines, along with 743 Teams taking part in the 2021(2020) Tokyo Olympics. This dataset contains the details of the Athletes, Coaches, and Teams participating. It includes their names, countries represented, discipline, and name of the coaches. There are 5 CSVs that will have to be joined and pipelined through the database.
2. Basic Functions:
    a. Search bar: Users can enter a keyword into the search bar and the website will return all relevant data related to the keyword.
    b. Filter features: Users can filter the data according to athletes, country, discipline, name, event, medal, and team.
    c. Visualization: Users can view the data in a bar chart form and ranking tables.
    d. Modification: Users can insert and remove rows of data into the database
    e. Data export: Users can export the data in a variety of formats such as JSON, CSV, and Excel.
3. Creative Component: the web application will be able to create interactive and visual charts based on the selected data using the filters provided. The application will be able to produce statistics for the variable data that is selected.
4. Project Title: Tokyo Olympics Wiki
5. Project Summary:
    a. This project involves creating a web application that allows users to filter and search keywords in data related to the Tokyo 2021 Olympics. This project aims to create a user-friendly platform that allows users to easily and quickly access the data they are looking for. This project will provide users with the ability to filter and search through the data in order to find answers to their questions, while also providing a more accessible way to access the data. This project will enable users to better understand the Olympics, allowing them to make more informed decisions about their Olympic experience.
    b. The web app will contain filtering based on athletes, country, discipline, name, event, medal, and team. The user will be able to quickly search through the data by entering keywords into the search bar. Additionally, this web application will also provide users with the ability to view the data in an interactive and visual way, allowing them to gain a better understanding of the Olympics.
6. This web application will help educate people about the Olympics and form predictions based on organized data for the future. This methodology will be able to be applied to more forms of Olympic data for further analysis when available.
7. This application is useful because it provides a way to organize a bunch of data and then produce more readable information and mathematical analysis of the data. It is important that it is organized with this variable application so that whoever may need statistical analysis on the Olympics will be able to answer their questions more efficiently.
8. The data is provided through Kaggle and sourced directly from the Tokyo Olympics 2020 Website. The data is a set of all the characteristics of the teams competing. It includes 5 different files which need to be joined and cleaned in order to perform analysis.

9. The website will offer a range of features that allow users to interact with the data. Firstly, users will be able to search for keywords in the database. This will allow them to quickly and easily locate the data they are looking for. Secondly, users will be able to filter the data based on athletes, country, discipline, name, event, medal, and team. This will allow them to narrow down the results and view specific data. Thirdly, users will be able to view the data in a bar chart form and ranking tables. This will allow them to easily compare and contrast different variables. Fourthly, users will be able to export the data in a variety of formats such as JSON, CSV, and Excel. Finally, the website will also be able to create interactive and visual charts based on the selected data using the filters provided. This will allow users to gain a better understanding of the data and form predictions for the future.
10. A low-fidelity UI mockup:



- In the low-fidelity UI mockup seen above, there is an example insertion into the database. In this case, an athlete's information is being inserted into the database, however the user will be able to choose what table they would like to insert data into. For

different tables, lets say Coaches for example, a different set of attributes will pop up on the right hand side for the user to insert. In the case above, Athlete Name, Country and

- Users will be able to search up keywords in our database where a table of values containing those keywords, along with the attributes associated with the keyword will pop up. In addition, users will be able to modify the database, by inserting or deleting rows, along with modifying specific entries within each row. Finally, a visualization of a dataset can appear on the bottom of the webpage. For example, the user can see a bar plot of the number of athletes playing for each country at the Olympics.

11. Project work distribution:
    a. Yifeng N and Eduardo Martinez will be responsible for setting up and maintaining the Google Cloud databases. They will also be responsible for creating data processing pipelines to extract data from the databases and store it in a format that is easy to analyze and visualize.
    b. Rohan Vij will be responsible for analyzing the data and creating visualizations to make the results easier to understand. He will also be responsible for creating interactive charts to allow users to gain insights into the data.
    c. Siddhant Sharma will be responsible for developing the user interface, including the navigation menu, search bar, and filter features. He will also be responsible for ensuring that the website is responsive and easy to use.
    d. Finally, everyone will be responsible for testing the web application to ensure that it is working correctly. This will involve testing the functionality of the website across a range of devices and browsers.