CS411 Project Proposal

1. Project Title: *Kosmos*

2. Project Summary:

   Kosmos is a website intended to display predictive information on every country in the world. Using historical data on these countries, we want the public to be able to see where different areas of our globe are heading in terms of their population, military size, and various other fields. Users of the site will have the ability to click on a country (from a spinning globe) and select a year within a range (now to 2075) and see all of the predictions our backend makes through mathematical equations on the different fields. With these fields, users can also see visualizations to detail expected changes across the years.

3. Description:

   Our mission with *Kosmos* is to create a comprehensive source of predictive information on global trends, all in one place. By developing a platform that displays important numerical country data like population, gross domestic product, birth rate, and military size in a sleek dashboard, Kosmos holds value to a vast array of people, whether it is the curious individual surfing the web, an expert data scientist seeking detailed analytical insights, or a government official needing data for policy making.

   Kosmos breaks down large datasets into easily readable and visually engaging formats that anybody with access to the internet can see. Users interact with a dynamic spinning globe and require just a few clicks to explore a variety of critical data for any given country, with the option to view current information or predictions of up to fifty years into the future. Our predictions are powered by time-tested algorithms like the cohort-component model and exponential growth model, in addition to modern machine learning algorithms like K-nearest neighbors (more details below).

4. Creative Components:

   Our team considered numerous features in our application that could incorporate technically challenging features that help our application shine and differentiate from more basic counterparts. One of the standout features was integrating interactive heat maps that dynamically display the population density of countries based on a user-selected year. This feature must utilize several APIs for custom map layering, map tile generation, and heatmap generation. Some libraries and APIs we considered for this feature were Matplotlib, Geopandas, and/or Plotly.

   Another feature we considered was a predictive component that could use our datasets to train machine learning models to predict country-specific metrics, including predicting the country hosting major sporting events (Olympics, for example), political party dominance, and tourism growth potential. Using the Python library sci-kit-learn, we are considering models such as K-Nearest Neighbors, classifiers, and logistic regression to develop a robust prediction engine. Furthermore, this prediction system will process multiple datasets to train the models, proving to truly be a technical challenge.

5. Usefulness:

Kosmos is a web application that provides users with valuable insights into the future trajectories of countries around the world. In today's interconnected society, it's important to have access to clear and user-friendly resources that offer predictive country statistics. Kosmos serves as an educational tool, helping users learn about the future outlook of various countries and the global landscape as a whole. By exploring the site, users can become more informed and develop a deeper awareness of the planet we all share. This includes important concerns like global warming.

The website blends information with interactivity. Users can explore any country on the globe and dive into its detailed statistics, which include dynamic plots and visualizations. These statistics are not just current; they can also be projected into the future, with predictions for variables like population, GDP, and more. Users can filter data by any year, both present and future, with projections extending up to 2075.

While there are other websites that provide country-related information, *Kosmos* stands out in three key ways:

1. It offers predictions for every country worldwide.
2. It provides annual predictions up to 2075.
3. It boasts a more interactive and visually engaging design, making it easier and more enjoyable for users to explore and understand the data.

6. Realness:

Our data comes from various datasets from the world bank. These datasets include carbon dioxide emissions, total population, size of armed forces, GDP, population density, birth rate, death rate, and literacy rate for every country for every year up until 2020 and from at latest, 1990. All of these datasets are in a CSV format; their cardinality is simply the number of locations in each of the datasets which is the countries and controlled territories (265). In regards to the degree, it is the number of years the datasets cover, which is from 30 to 60 years plus the indicator code, indicator name, and country code.

7. Functionality

Kosmos offers an interactive experience for users who want to explore global statistics and predictions for countries around the world. Once the site is opened, users are greeted with an interactive globe / world map, where they can select any country of interest. Once a country is chosen, users are presented with detailed statistics that cover a variety of aspects, such as population, GDP, CO2 emissions, and other key metrics. These statistics can be filtered by year, allowing users to view historical data, current conditions, or future projections based on predictive models.

In addition to this, users have the ability to interact with the website in a multitude of ways. They can select different metrics to compare across countries or years, and the application generates interactive charts and graphs for easy analysis. Additionally, Kosmos allows users to search for specific countries or global trends, helping them quickly navigate to

relevant data. The intuitive design encourages exploration, with options to dive deeper into more complex data or return to high-level global comparisons.

In terms of CRUD capabilities, an additional functionality of our website is the option to create your own predictions and see how they fair versus our model. Furthermore, that newly created data point of the user's prediction can be saved and read back from the database onto site in future circumstances. Lastly, the user can remove or update their prediction on the front end, which will correspondingly "update" the database.

Project Work Distribution:
The main project delegation comes from key components: dataset formatting, algorithm development, frontend development, and backend development.
- Aman: Work on all aspects, leading dataset formatting.
- Arnav: Work on all aspects, leading backend development.
- Dylan: Work on all aspects, leading front-end development.
- Keene: Work on all aspects, leading algorithm development.

Project Mockup: