1. **Project Title & Abstract**
    - **Our project focuses on the Global Growth of Energy Usage and Types.**

      For people all over the world, self-awareness of energy usage is incredibly important. Those living in developed countries should be considerate of the amount of energy they're using in proportion to developing countries. More access to energy in developed countries should mean more attention given to where that energy is being spent. Governments all over the world should constantly revisit their energy consumption – the data that we use may help them do just that.

    - We plan on using time series datasets that feature per-capital energy consumption, country populations, and types of energy used globally. Some questions we would like to answer:
        1. How has energy use grown over the past 50 years?
        2. What countries are using the most energy? The least?
        3. Is there a correlation between energy use and energy type for developed nations? Developing?
    - Our intended audience is quite broad. This project can be used to educate and to help people in power to make decisions. Policy makers, the media, and the general population are our primary audiences because they have the most power. We expect a lot of value to come from this project because many social issues converge on the issue of energy consumption.

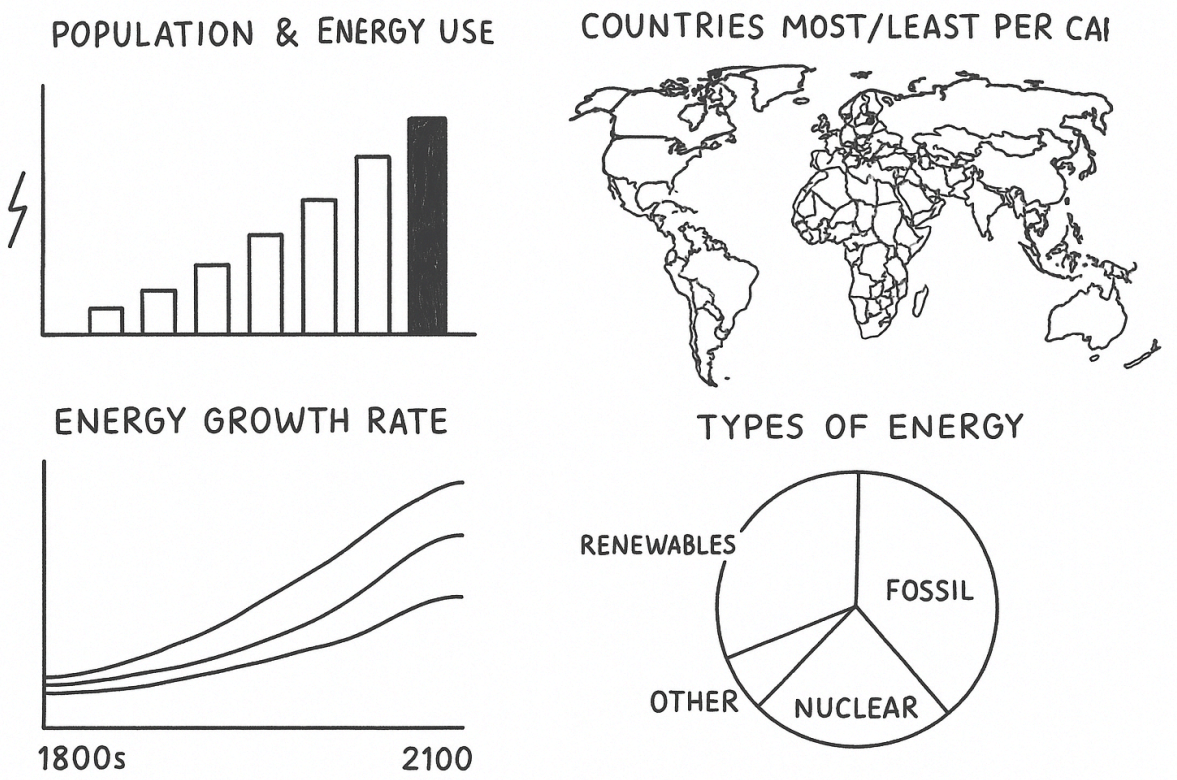
2. **Overall Plan**
    - **Data**: source(s), brief description, size, format, initial quality/limitations.
        i. Primary (Per Capita Energy Use):
            1. This dataset includes energy use (in kilowatt-hours) per capita for each world country as a time series between 1964 to 2024.
            2. There are 4 columns and over 11,000 rows for each year and country.
            3. Not every country has data going all the way back to 1964.

        ii. Secondary (Population):
            1. This dataset includes the population data for each country from 2018 to 2024.
            2. 67 columns, 271 rows. Each row correlates to a country and each year has a column.
            3. Preprocessing may be complicated to align this dataset with the previous one.

        iii. Secondary (Types of Energy):

1. This is a time series of different types of energy and the total usage of that energy by all countries aggregated from 1800 to 2024.
2. 13 columns, 78 rows. Each row represents a year, each column represents a different energy type.
3. There is no specific per-country data

- **Methods**: planned processing/analysis (at a high level).
    i. Processing: Multiple datasets will be combined together, N/A or empty values will not be used in visualizations
    ii. Analysis: Our analysis will have a strong foundation in the time series over multiple decades for various countries.
- **Story angle**: what insight or narrative you expect to communicate.
    i. We would like to tell a story about the energy patterns of different types of countries, including the type of energy and how much is used – this will help differentiate between developing and developed countries.
- **Initial sketches/prototypes**: low-fi sketches or wireframes (attach images or links). If relevant, list *audience* and *KPIs*.
    i. Visualization Ideas:
        1. Population and energy use
        2. Countries most/least per capita
        3. Compare growth rate of energy of each continent since 1800s
        4. Projection for 2100 from the UN data
        5. Types of energy
    ii. Audience: Our audience is very broad, but some examples would include the general public, media, and policy makers.
    iii. KPIs: A KPI includes kilowatt-hours per person.



3. **Collaboration Plan**
   - **Roles** (e.g., data lead, design lead, QA/editor, presenter).
     - Data Lead - John Thornton
     - Design Lead - Medha
     - QA/Editor - Srujana
     - Presenter - Aishwarya
   - **Tools & workflow**: e.g., Google Drive, GitHub, Kaggle Notebooks, Figma, Tableau/Power BI, etc.
     i. Dashboard: Power BI
     ii. Collaboration Hub: Github
   - **Communication**: where/how you'll coordinate (Discord/Slack/GroupMe/Email), and *when* you'll meet (in-class + outside).
     i. Discord/Zoom
   - **Milestone timeline**: weekly tasks & owners through Final Submission.
     - Submit the proposal by November 14th
     - Complete the Google Colab data processing by November 18th
     - Complete the poster design by November 19th

- Complete the Power BI dashboard the November 21st
- Complete the technical report and the design report by November 28th

4. **Group Expectations & Accountability**
   - Norms (responsiveness, meeting etiquette, version control, review before merge).

     Our team of four values collaboration, respect, and accountability. We aim to maintain open and timely communication through **Discord** for daily updates and **Zoom** for weekly check-ins. Each member is expected to respond within 24 hours, participate actively in discussions, and support others to keep progress consistent.

   - Attendance & deliverable expectations.

     We will manage our files and code using **GitHub**, following version control practices such as clear commits and review before merging. Everyone is responsible for completing their assigned work on time and ensuring the overall quality of deliverables.

   - **Consequence ladder** for non-participation. For example: 1) 3 documented warnings → 2) TA/Instructor mediation → 3) Individual letter-grade reduction if unresolved.
   - Dispute resolution plan (how you'll escalate/resolve issues).

     We will first address any conflicts through open discussion within the team. If unresolved, the concern will be brought to the instructor for fair mediation.

5. **Risk & Feasibility Check**
   - Top 2–3 risks (e.g., data availability/cleanliness, scope creep) and mitigation plans.
     - This dataset combines data from different sources. The HYDE is based on archaeological, historical data rather than census records. These estimates can have errors, especially for ancient and medieval periods. So analyzing these for long-term trends may give a misleading report. We decided to use the data from the 1800s to now for accurate representation.
     - Even if we use the data from the 1800s, there is still a break between 1949 to 1950 where after 1950 we rely on the UN world population data. So there might be some error or mixup when they were stitching the datasets together.

   - Accessibility intentions (color-blind safe palettes, alt text, readable type, screen-reader notes if relevant).

        - We will use color palettes that are designed for accessibility, we will avoid red-green colors and also use patterns, line style visualizations so it would be easier to follow the visualizations. We will also use concise conclusions of each analysis rather than long descriptive paragraphs, so we can focus on the main takeaway of the visualization.
    - Ethics & attribution (citing data sources, model limitations, known biases).
        - We will cite our sources on where we got the dataset from, we also cited the sources used in the dataset. We also decided to use MLA format when citing any of these sources in our analysis.
        - Biases might be included in the Projection data from the UN, because they heavily rely on fertility and migration data but they don't include any natural disasters or wars in the projection. So it might not be completely accurate.

6. **Citations**
    - U.S. Energy Information Administration (2025)
    - Energy Institute - Statistical Review of World Energy (2025)
    - Population based on various sources (2024)
    - Energy Institute - Statistical Review of World Energy (2025)
    - Smil (2017)
    - World Population Prospects, United Nations (UN), uri: population.un.org/wpp, publisher: UN
    - Population Division; Statistical databases and publications from national statistical offices, National Statistical Offices, uri: unstats.un.org/home/nso_sites, publisher: National Statistical Offices;
    - Eurostat: Demographic Statistics, Eurostat (ESTAT), uri: ec.europa.eu/eurostat/data/database?node_code=earn_ses_monthly, publisher: Eurostat;
    - Population and Vital Statistics Report (various years), United Nations (UN), uri: unstats.un.org, publisher: UN Statistics Division

References:

- Our World in Data (OWID), Energy Use & Energy Substitution Datasets.
- World Bank Open Data, Population Indicators (2018–2024).
- UN World Population Prospects (2024).

- U.S. Energy Information Administration (2025).
- Energy Institute, Statistical Review of World Energy (2025).
- Smil, Vaclav. Energy and Civilization (2017).