Nayan Kadhre - Milestone 1 – Project Proposal

1. Project Title
Predicting Follower Growth on Twitch: The Impact of Hours Streamed

2. Team or Individual Project

- **Type:** Individual
- **Team Members:** Nayan Kadhre

3. Research Question / Problem Statement

- **Main Question:** Does the number of hours a streamer spends streaming predict the total number of followers gained on Twitch?
- **Hypothesis:** Streamers who spend more hours streaming tend to gain more followers due to increased visibility and audience engagement.

4. Project Description / Focus
This project analyzes the relationship between streaming activity and audience growth on Twitch. Using the Top Streamers dataset, it applies linear regression to determine whether hours streamed significantly predict followers gained. The goal is to identify whether effort directly influences audience expansion.

5. Guiding Framework / Theory / Policy
The project is guided by regression theory, using a simple linear regression model to quantify the relationship between predictor (hours streamed) and outcome (followers gained).

6. Citation Style
This project will use the APA 7th Edition citation style for all references, datasets, and in-text citations.

7. Dataset Information

- **Dataset Source:** Kaggle – Top Streamers on Twitch (Aayush Mishra, 2021)
- **Data Type:** Numerical and categorical
- **Key Variables:** Channel, watch_time, stream_time, peak_viewers, average_viewers, followers, followers_gained, views_gained, partnered, mature.
- **Data Scope:** Esports / Streaming analytics
- **Availability:** Publicly available dataset on Kaggle

8. Data Access / Ethical Considerations

- **IRB Requirement:** Not required — dataset contains no personally identifiable information.
- **Data Security / Access Plan:** Data will be accessed directly from Kaggle and analyzed in a local, secure environment.

9. Feasibility & Timeline

The project is feasible within the semester timeframe. Data cleaning, exploratory data analysis, and regression modeling will be performed using Python (pandas, matplotlib, scikit-learn).

10. References

Mishra, A. (2021). Top Streamers on Twitch [Data set]. Kaggle. https://www.kaggle.com/datasets/aayushmishra1512/twitchdata