2-Page Proposal for the NoSQL Project

1. Team members in alphabetical order of the last name:

   i) Deshpande Abhishek
   ii) Mane Akash
   iii) Purandare Chinmay

2. Briefly and clearly describe the problem you want to work on.

   Answer: We are addressing the challenge of analyzing the 'Top 10000 Songs on Spotify 1960 - Now' dataset using a combination of Tkinter for the development of a robust graphical user interface (GUI) and MongoDB as our NoSQL database. This dataset encompasses a wide range of attributes such as track duration, popularity, danceability, and acousticness, among others. Our primary objective is to tackle four non-trivial data analysis tasks, leveraging these attributes. These tasks involve exploring music trends through data visualization, employing MongoDB's querying capabilities to delve into artist behavior, and utilizing advanced statistical techniques to assess the impact of factors like genre and label on track characteristics. Our project aims to deliver an accessible and insightful solution for users to interact with and gain valuable insights from this rich music dataset.

3. Which NoSQL database you want to use?

   Answer: MongoDB

4. Which public dataset will you use? Provide the URL of the dataset. Does the dataset have at least 20 attributes and 5,000 samples?

   Answer:
   - Dataset Name: Top 10000 Songs on Spotify 1960 -Now
   - No. of Attributes: 35
   - No. of Records: 10000
   - Link: https://www.kaggle.com/datasets/joebeachcapital/top-10000-spotify-songs-1960-now

5. Describe the N+1 non-trivial data analysis queries/tasks you plan to design and implement. Do you take Option 1 or 2 described in the NoSQL project description on Brightspace?

   Answer:
   - We will be selecting Option 1 for our Project.

   - Data Analysis Tasks:
     1. Can we identify any trends in the relationship between the popularity of tracks, their danceability, and energy levels across different albums?
     2. Are tracks released by major record labels generally more danceable and

energetic compared to tracks from independent labels, and does this impact their popularity?

3. Do tracks released on Fridays or weekends tend to have higher tempo and energy levels, and are they more popular than tracks released on other days of the week?

4. Is there a correlation between track duration and track popularity, and has this correlation changed over the years?

6. Will you implement a Web interface or GUI? Remember *that simply plotting graphs or tables, for example, in Jupiter Notebooks doesn't count*, as specified in the NoSQL project description. For more details, refer to the project description.

Answer: We are planning to implement a Tkinter-based graphical user interface (GUI) as specified in the project description. This GUI will provide users with an interactive platform to query our MongoDB NoSQL database and visualize data analysis results. We believe that this user-friendly interface will enhance the overall usability and accessibility of our NoSQL data analysis solution.

7. Briefly describe your milestones to finish the project this semester.

Answer:
September 23 - September 30 (1 week):
- Research and select the technology stack, including MongoDB for the database and Tkinter for the GUI.

October 01 - October 14 (2 weeks):
- Design and implement the Tkinter-based graphical user interface (GUI) for the application.
- Begin writing MongoDB queries for data retrieval and analysis.

October 15 - October 31 (2 weeks):
- Refine the GUI interface for improved usability and aesthetics.
- Continue developing and testing MongoDB queries to ensure accuracy.

November 01 - November 15 (2 weeks):
- Finalize the GUI, incorporating user feedback and making any necessary improvements.
- Test the complete application and begin drafting the initial sections of the Project Report, outlining the project's objectives, methodology, and early progress.

November 15 - November 31 (2 weeks):
- Connect the dataset to MongoDB, ensuring data integrity and consistency.
- Conduct extensive testing and debugging to resolve any issues before the project proposal submission deadline.
- Continue working on the Project Report, documenting the development process, challenges faced, and initial results.