**Milestone2 Agenda**

* **Motivation for the Idea**

In addition to the basic functions of an Olympic Games database, we would like to conduct some interesting queries such as “Is there a limitation on the strength/speed of mankind”, “Is the total performance of a nation in Olympic Games dependent on its population scale” We are curious about whether the sky is the limit, and we regard these questions as a way of interpreting the so-called Olympic Spirit.

* **Basic Features**

Based on the design of the relational database schema, our application mainly provides three functional parts consisting of 10 different query scenarios:

**Part 1. Medal numbers with ranks of all countries:**

1. Over the years, the sum of medals and rank of all the countries
2. Over the years, the sum of medal and rank of all the countries in specific event of specific discipline
3. In specific edition of Olympic Games, the sum of medals and rank of all the countries
4. In specific edition of Olympic Games, the sum of medals and rank of all the countries in specific event of specific discipline

**Part 2. Medal numbers with ranks of all Olympic games**

1. Over the years, the sum of medals and rank of all Olympic games
2. Over the years, the sum of medals and rank of all Olympic games in specific discipline
3. Over the years, the sum of medals and rank of all Olympic games in specific event of specific discipline

**Part 3. Average medal numbers per million nation population**

1. In specific edition of Olympic Games, the sum of medals and rank of all the countries and its average medal numbers per million population
2. In specific edition of Olympic Games, the sum of medals and rank of all the countries in specific discipline and its average medal numbers per million population
3. In specific edition of Olympic Games, the sum of medals and rank of all the countries in specific event of specific discipline and its average medal numbers per million population

For the non-relational part, we have used complementary sources from other websites, which contains the records and personal information of athletes from 17 different disciplines. Their records are quantified, which can be used to plot graphs in the front end.

* **Advanced Features**We visualized the results of the queries and demonstrated them via figures, charts and cartoons. Simple statistical efforts such as fitting and regression was also applied in order to indicate the trends of Olympic records.
* **Technology and Tools Used**The back-end of this database will rely heavily on SQL and relational schemas while Node.js is adopted as the front-end platform of displaying the queries. MongoDB, as an additional, alternative solution, might be used as well if necessary.
* **Complimentary Sources**

In addition to the provided .csv file we will gather information from XML sources such as webpages. We will ingest the data via parsing the source code of the websites and injecting them to the relational database.

Athlete Records: <http://www.theolympicdatabase.nl/olympic/sports>

Population: <https://en.wikipedia.org/wiki/List_of_countries_by_past_population_(United_Nations,_estimates)>

* **Member Responsibility**Jiawei Liu: Query handling, Node.js technical support (22% workload)

Jiapo Tai: Database setup, data parsing and import preprocessing (22% workload)

Hongru Du: Schema design, data parsing, cleaning and insertion (22% workload)

Yuechen Luo: Node.js design and coding (34% workload)

* **Schema Design**

/Users/hongrudu/Downloads/ER_161123 (1).png