# Information Visualization

# CHECKPOINT I: Visualization Proposal

G01-A

**1. Domain**

We have a big amount of data regarding the victories and podium positions of athletes in various sports, throughout the editions of the Olympic Games. We want to give that data some meaning that makes sense to the human eye, to learn things such what countries are best at certain sports and how they compare in general, over the years.

**2. Dataset**

We collected a spread sheet with thousands of entries, containing podium finishes of athletes in each Olympic sport, and the country they represented, from the first Olympic Games in 1896 until 2008. The original chunk of data was obtained from a spread sheet made available in an article by The Guardian. We may also relate the amount of medallists a country has to its population (only since 1960), using another spread sheet on the World Bank website.

Article: <http://www.theguardian.com/sport/datablog/2012/jun/25/olympic-medal-winner-list-data#data>  
Olympics: <https://docs.google.com/spreadsheets/d/1zeeZQzFoHE2j_ZrqDkVJK9eF7OH1yvg75c8S-aBcxaU/edit>  
Population: <http://data.worldbank.org/indicator/SP.POP.TOTL?page=6>

**3. Tasks to be supported**

* + - 1. 1 – Browse – show a list of the countries with the most gold medallists in total in 1896.2 – Identify – show the country with the most medallists in Judo of all time.  
         3 – Locate – show the position of the country in the overall standings.  
         4 – Explore – using the derivative variable nºmedallists/nºpopulation, show the list of countries with the highest coefficient.  
         5 – Compare – show the medals each country won.

**4. Example Questions (matching the supported tasks)**

* + - 1. 1 – What countries had the most gold medallists in the first games, in 1896?  
         2 – What country has the most medallists in Judo?  
         3 – What are the standings of the USSR in 1964?  
         4 – See the countries with the most medallists per capita in 2008.  
         5 – How do the USSR and Russia’s cumulative scores compare?
      2. **5. Data Sample**

(from “all\_medallists.csv”):  
City,Edition,Sport,Discipline,Athlete,NOC,Gender,Event,Event\_gender,Medal  
Athens,1896,Aquatics,Swimming,"HAJOS, Alfred",HUN,Men,100m freestyle,M,Gold

(from “population.csv”):  
China,CHN,"Population, total",SP.POP.TOTL,667070000,660330000,665770000,682335000,698355000,715185000,735400000,754550000,774510000,796025000,818315000,841105000,862030000,881940000,900350000,916395000,930685000,943455000,956165000,969005000,981235000,993885000,1008630000,1023310000,1036825000,1051040000,1066790000,1084035000,1101630000,1118650000,1135185000,1150780000,1164970000,1178440000,1191835000,1204855000,1217550000,1230075000,1241935000,1252735000,1262645000,1271850000,1280400000,1288400000,1296075000,1303720000,1311020000,1317885000,1324655000,1331260000,1337705000,1344130000,1350695000,1357380000,1364270000,