**Capstone Project – Sports Stats**

Chuxuan Quan

10/04/2021

**Milestone 1**

* Preparing for Your Project Proposal
  + Which client did you select and why?
    - I choose the dataset of sports stats client, because I am interested in sports and I spend much of my spare time on sports, like swimming, basketball and baseball. So I want to dig into the dataset to figure out interesting analysis.
    - Also I can find the deeper information of the sports by SQL.
  + Describe the steps you took to import and clean the data.
    - I use jupyter notebook to be my text editor and use read csv method to import data.
    - Then use pandas to sql to store the data.
    - Check the Null value and missing value by info method.
  + Perform initial exploration of data and provide some screenshots or display some stats of the data you are looking at.

图形用户界面, 应用程序

描述已自动生成

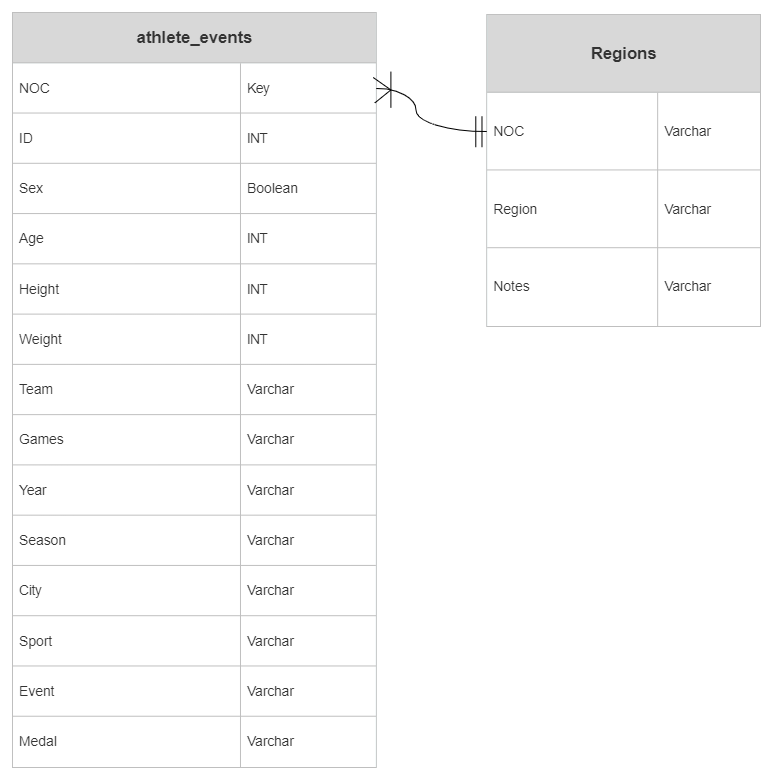
图形用户界面, 应用程序

描述已自动生成

图形用户界面, 文本, 应用程序

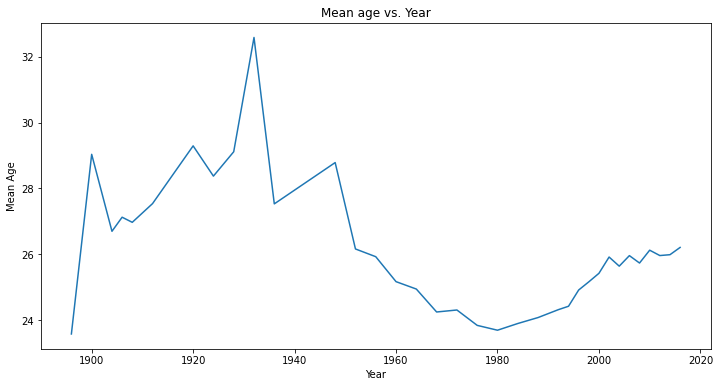
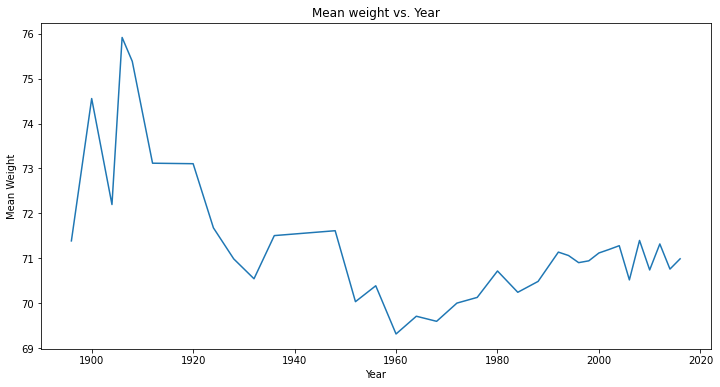
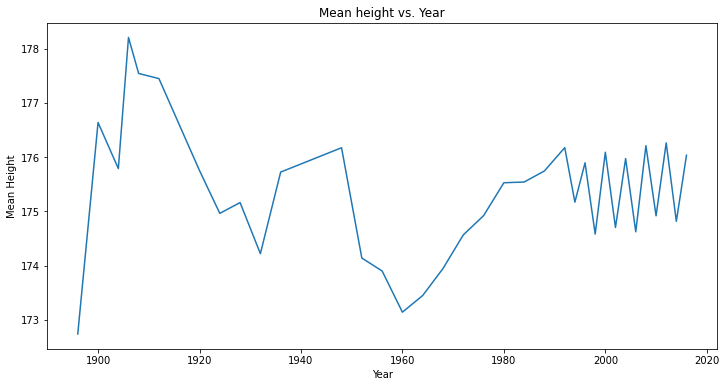
描述已自动生成

* + Create an ERD or proposed ERD to show the relationships of the data you are exploring.

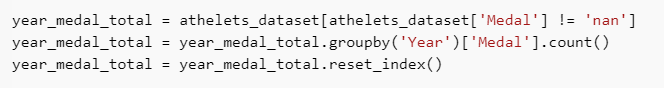


* Develop Project Proposal
  + Description
    - Analyze the trend of medals of top 30 countries with the changing of years.
    - Find the relation between different countries’ performance.
    - The parameters like height, weight and age have effects on the outcome.
  + Questions
    - Is the trend of age getting younger and the trend of weight is getting larger?
    - What countries have a good performance in recent 10 years?
  + Hypothesis
    - The countries basically have no changes during the last 10 years compared it to 20 years ago.
    - The age is getting younger and the height and weight are getting larger.

**Milestone 2**



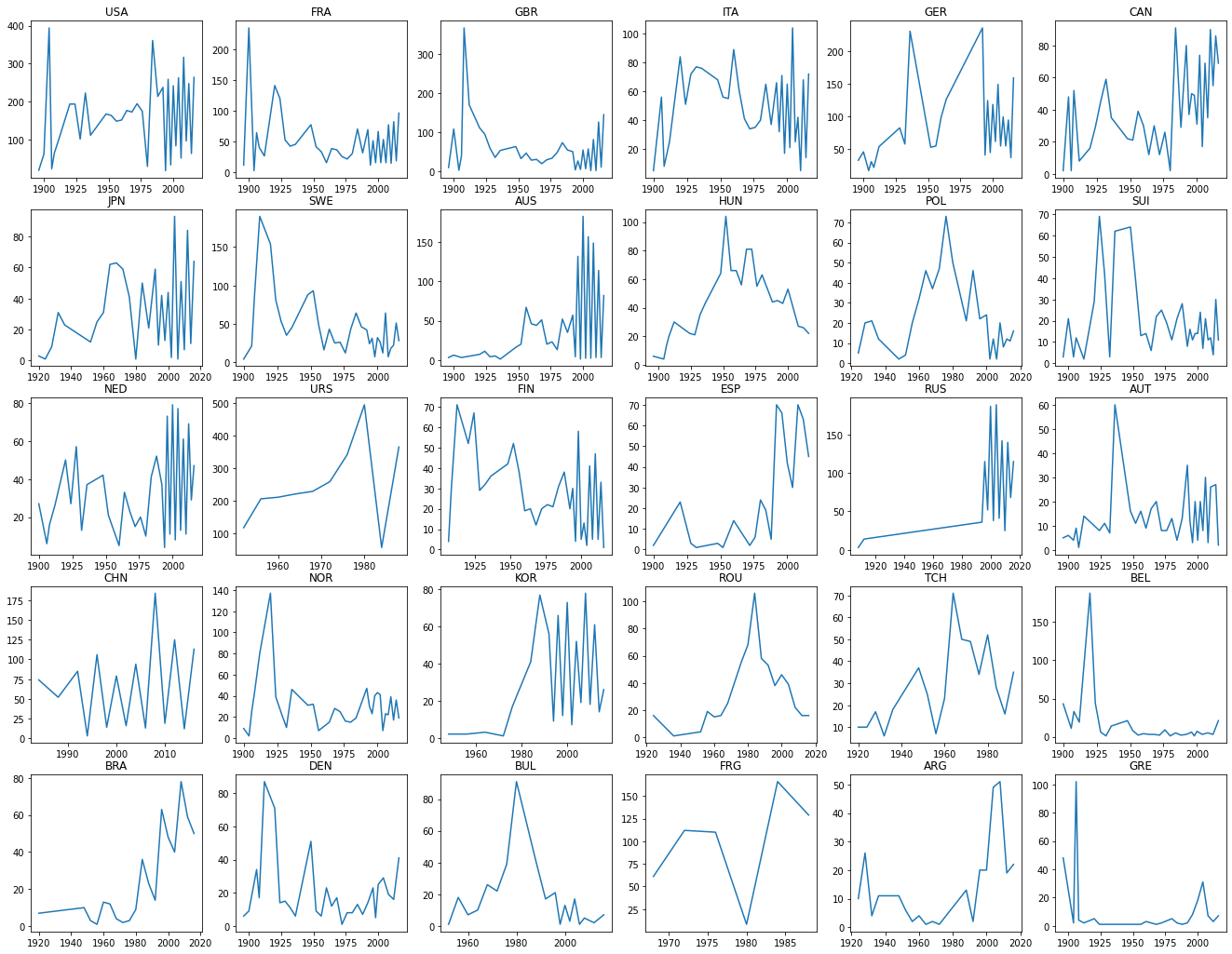
The trend of age, weight and height all experienced a rise before 1920 and a drop then a rise. Average weight and average height are highly correlated, and both reached their lowest point in 1960, and then increased year by year. The decline before 1960 may be related to the world war. In the peaceful era, the height and weight of the contestants increased. These three indicators fluctuated slightly after 2000, but they tended to stabilize as a whole.

文本

低可信度描述已自动生成电子计算器

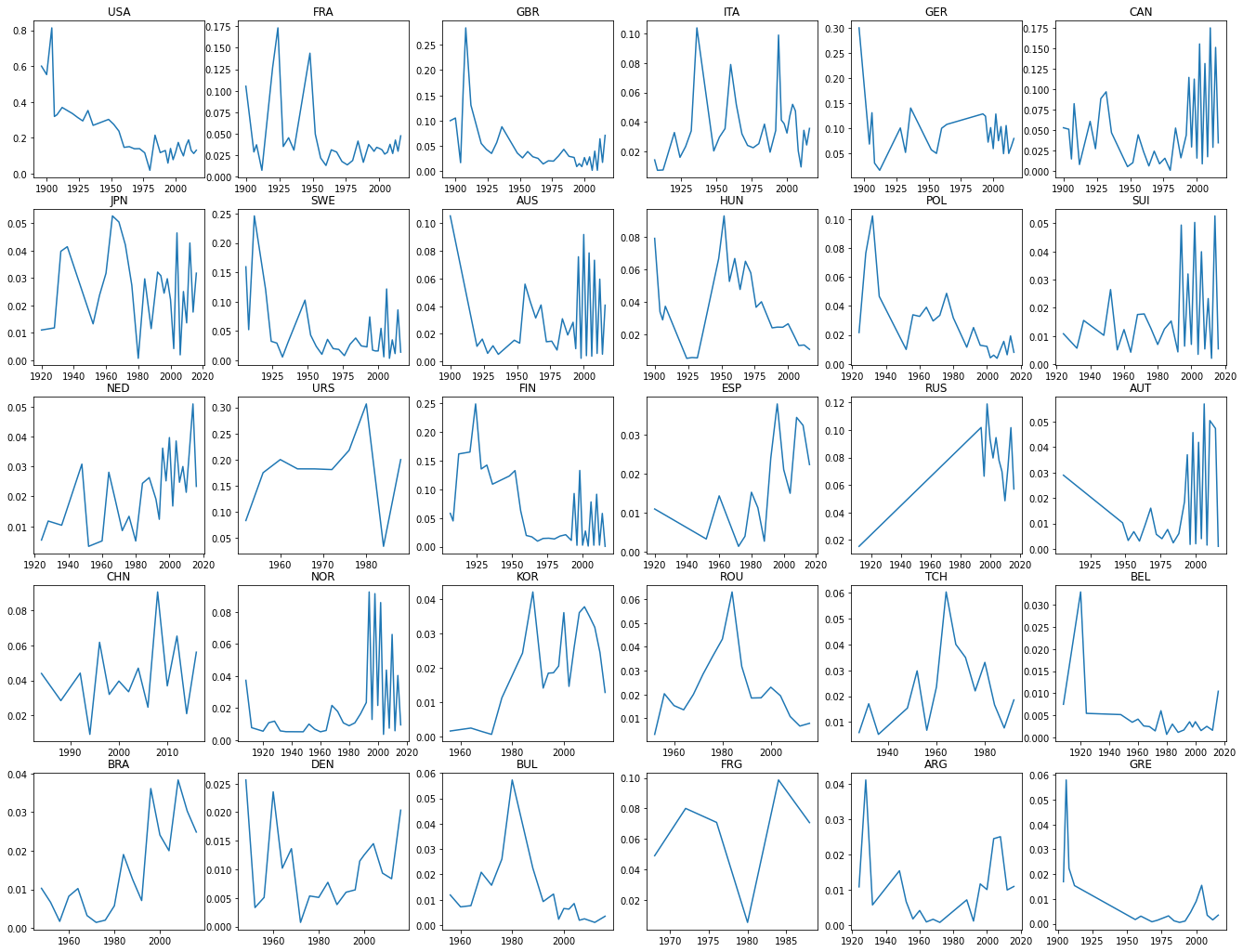
低可信度描述已自动生成





Here is a list of the medals of the top 30 countries in the Olympic Games. The number of medals varies in each country. The performance of most countries is stable or on the rise, and a few countries have seen fewer and fewer medals in recent decades.





This chart shows the ratio of medals of each country to total medals of each year. The first three countries, USA, FRA, GBR, decreases the ratio of medals, while CAN, NED, ESP, RUS, BRA have a growth in the medal ratio.