Olympic Project

Team 1: Xiaoqi Hu, Siqi Zhang, Yanni Lan, Chenhang Niu, Yuyang Shu, Kyle Blackburn (Team Lead)

Raw Data

Data Source (https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results/version/2?)

```
%%bq query
SELECT *
FROM `cohort-b-team-1.olympic_data.athlete_data`
LIMIT 100
```

ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	City	Sport
14840	Andreas Bang Brecke	M	40	0	0.0	Jo	NOR	1920 Summer	1920	Summer	Antwerpen	Sailing
56486	Paal Kaasen	M	36	0	0.0	Jo	NOR	1920 Summer	1920	Summer	Antwerpen	Sailing
103129	Ingolf Richard Rd	M	30	0	0.0	Jo	NOR	1920 Summer	1920	Summer	Antwerpen	Sailing
58576	Godfrey Kenneth Kelly	M	31	175	75.0	Bim	BAH	1960 Summer	1960	Summer	Roma	Sailing
58585	Maurice George Kelly	M	35	179	79.0	Bim	BAH	1960 Summer	1960	Summer	Roma	Sailing
61579	Percival Andrew "Percy" Knowles	M	29	168	70.0	Bim	BAH	1960 Summer	1960	Summer	Roma	Sailing
33921	Philip Vernon le Geyt Falle	M	43	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
100798	Thomas Cooper Riggs	M	24	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
102407	Ernest John Roney	M	27	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
102408	Esmond Richard Roney	M	20	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
102409	Margaret Helena Roney	F	28	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
111668	Thomas G. Skinner	M	0	0	0.0	Feo	GBR	1928 Summer	1928	Summer	Amsterdam	Sailing
22875	Cecil George Cooke	M	41	188	90.0	Gem	BAH	1964 Summer	1964	Summer	Tokyo	Sailing
61572	Durward Randolph Knowles	M	46	176	78.0	Gem	BAH	1964 Summer	1964	Summer	Tokyo	Sailing

44835	Donald Jasper Hains	M	36	0	0.0	Jet	CAN	1952 Summer	1952	Summer	Helsinki	Sailing
50372	Archibald Howie	M	44	0	0.0	Jet	CAN	1952 Summer	1952	Summer	Helsinki	Sailing
101358	John Norman Frank Robertson	M	23	0	0.0	Jet	CAN	1952 Summer	1952	Summer	Helsinki	Sailing
53291	James Peter Jaffe	M	18	0	0.0	Joy	GBR	1932 Summer	1932	Summer	Los Angeles	Sailing
99124	George Colin Ratsey	M	26	180	0.0	Joy	GBR	1932 Summer	1932	Summer	Los Angeles	Sailing
30398	Abraham Everardus "Biem" Dudok van Heel	M	30	172	70.0	Joy	NED	1948 Summer	1948	Summer	London	Sailing
55791	Cornelis Wilhelm "Kees" Jonker	M	39	0	0.0	Joy	NED	1948 Summer	1948	Summer	London	Sailing
124559	Willem Paul "Wim" van Duyl	M	28	189	90.0	Joy	NED	1948 Summer	1948	Summer	London	Sailing
21269	ivind Christensen (Kristensen-)	M	36	170	65.0	KNS	NOR	1936 Summer	1936	Summer	Berlin	Sailing
47831	Sigurd Frithjof Herbern (Johannesen-)	M	35	0	0.0	KNS	NOR	1936 Summer	1936	Summer	Berlin	Sailing
94688	Willy H. A. Pieper	M	25	0	0.0	Kln	SUI	1936 Summer	1936	Summer	Berlin	Sailing

(rows: 100, time: 0.7s, 43MB processed, job: job_QagLsPWLdAdNwrbjyr9SOEKV7xlb)

Problem Statement

The Olympic Games have been a pivotal event in human development, connecting citizens around the world for hundreds of years. Some sports have been around since Ancient Greece (i.e., track and field) while others have been added along the way. Though the Olympic spirit emphasizes competition, inclusion, and the triumph of high-level athletics, statisticians often only look at the quantity of medals as a metric of a country's success. Our team took an interest in this fact and decided to investigate it further. Our research question surrounds this topic: are there patterns in who wins and who doesn't? Are medal wins correlated with other interests such as age, gender, season, host city (country), or sports category (from Athens 1896 to Rio 2016)? In this project we try to identify correlations between these factors and medal wins using free Olympic history data online.

Work Process

1.Gather Olympics data from kaggle.com (above)

2.Download full set in .csv form

3.Create instance on GCP through Cloud shell in order to run Notebook in Google Cloud Datalab

4.Create bucket in Google Cloud Project, to then make a table in Big Query. When adding the data to a table in Big Query, it did not go nicely and would not correctly detect schema. So we individually assigned the correct schema, field by field.

5.Use Big Query to select data relevant to our research questions

6.Create a Dashboard via Google DataStudio

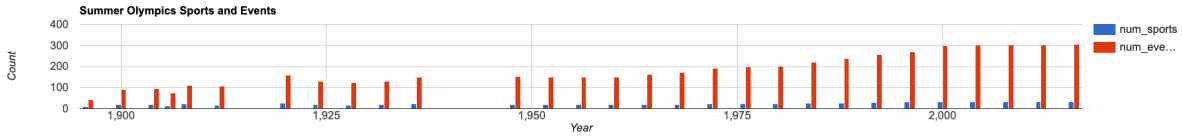
7.Potentially perform further prediction analysis through R/Python for the 2020 Olympics

Summer: Sports and Events

The background behind this query is to provide contextual information of the expansion of the Olympics overtime. A growing number of sports (which also equates to a higher number of events per sport) means a higher number of opportunities to earn medals of any color. Thus, when researching information about medal counts over time, it was important to look at how medal opportunities changed as well. We wanted to include as much information as possible, so we divided up winter and summer Olympics to stratify information groups.

```
%%bq query -n summer_events_time
SELECT Year, COUNT(DISTINCT Sport) as num_sports, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Summer'
GROUP BY Year
ORDER BY Year ASC
```

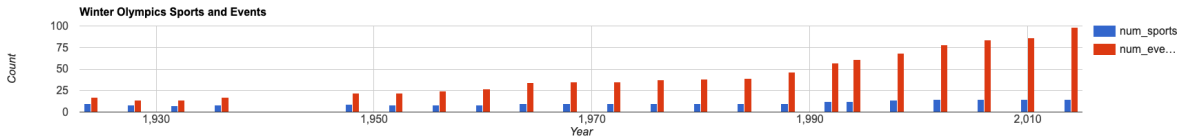
```
%%chart columns --data summer_events_time
title: Summer Olympics Sports and Events
hAxis:
  title: Year
vAxis:
  title: Count
legend: True
```



Winter: Sports and Events

```
%%bq query -n winter_events_time
SELECT Year, COUNT(DISTINCT Sport) as num_sports, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Winter'
GROUP BY Year
ORDER BY Year ASC
```

```
%%chart columns --data winter_events_time
title: Winter Olympics Sports and Events
hAxis:
  title: Year
vAxis:
  title: Count
legend: True
```

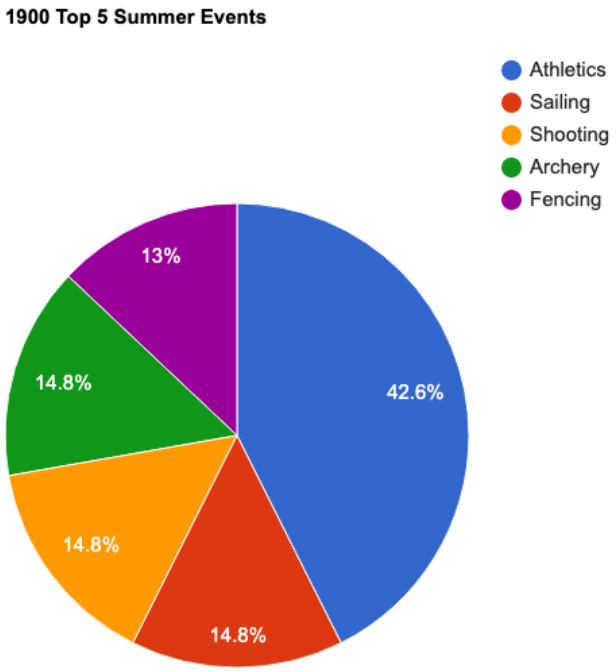


1900 Summer Events Breakdown

```
%%bq query -n 1900_events
SELECT Sport, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Summer'AND Year = 1900
GROUP BY Sport
ORDER BY num_events DESC
```

```
%%bq query -n 1900_breakdown
SELECT
   *,
   ROUND(num_events/(SELECT SUM(num_events) FROM `cohort-b-team-1.olympic_data.1900_events_
breakdown`)*100,2) AS perc
FROM
`cohort-b-team-1.olympic_data.1900_events_breakdown`
Limit 5
```

```
%%chart pie --data 1900_breakdown
title: 1900 Top 5 Summer Events
height: 900
width: 900
legend: True
```

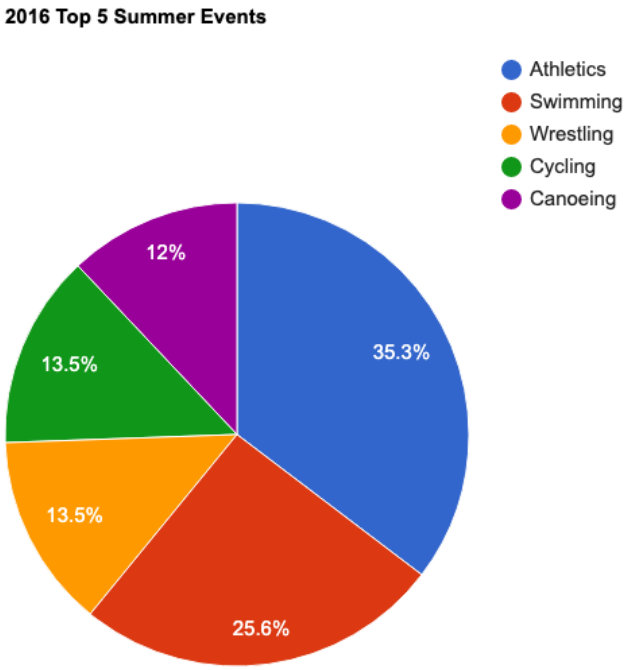


2016 Summer Events Breakdown

```
%%bq query -n 2016_events
SELECT Sport, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Summer'AND Year = 2016
GROUP BY Sport
ORDER BY num_events DESC
```

```
%%bq query -n 2016_breakdown
SELECT
   *,
   ROUND(num_events/(SELECT SUM(num_events) FROM `cohort-b-team-1.olympic_data.2016_events_
breakdown`)*100,2) AS perc
FROM
`cohort-b-team-1.olympic_data.2016_events_breakdown`
LIMIT 5
```

```
%%chart pie --data 2016_breakdown
title: 2016 Top 5 Summer Events
height: 900
width: 900
legend: True
```

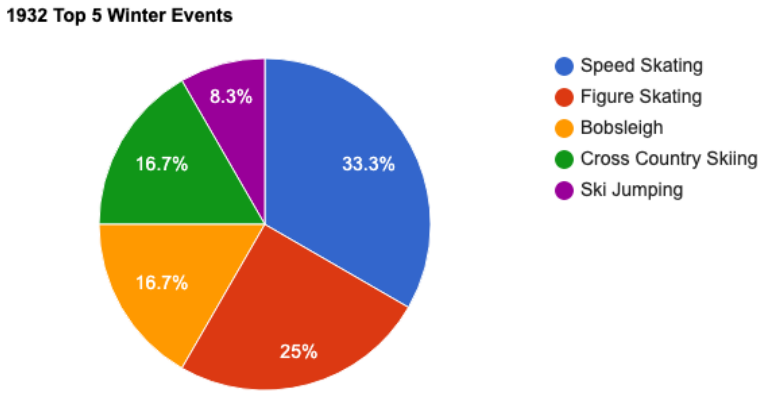


1932 Winter Events Breakdown

```
%%bq query  -n 1932_events
SELECT Sport, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Winter'AND Year = 1932
GROUP BY Sport
ORDER BY num_events DESC
```

```
%%bq query -n 1932_breakdown
SELECT
    *,
    ROUND(num_events/(SELECT SUM(num_events) FROM `cohort-b-team-1.olympic_data.w_1932_even
ts_breakdown`)*100,2) AS perc
FROM
`cohort-b-team-1.olympic_data.w_1932_events_breakdown`
LIMIT 5
```

```
%%chart pie --data 1932_breakdown
title: 1932 Top 5 Winter Events
height: 400
width: 900
legend: True
```

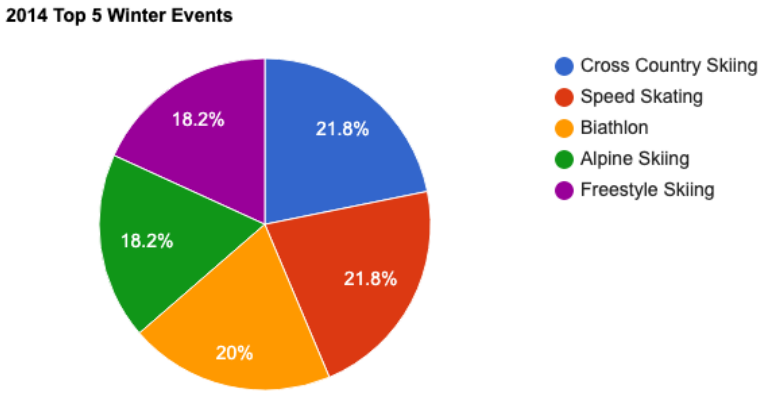


2014 Winter Events Breakdown

```
%%bq query  -n 2014_events
SELECT Sport, COUNT(DISTINCT Event) as num_events
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Season = 'Winter'AND Year = 2014
GROUP BY Sport
ORDER BY num_events DESC
```

```
%%bq query -n 2014_breakdown
SELECT
    *,
    ROUND(num_events/(SELECT SUM(num_events) FROM `cohort-b-team-1.olympic_data.w_2014_even
ts_breakdown`)*100,2) AS perc
FROM
`cohort-b-team-1.olympic_data.w_2014_events_breakdown`
LIMIT 5
```

```
%%chart pie --data 2014_breakdown
title: 2014 Top 5 Winter Events
height: 400
width: 900
legend: True
```



What's the average, minium and maxium BMI for those players who win the medals?

Here, we are trying to determine if there are any patterns between winning medals and physical traits of the athletes. Granted, this is a bit of a generalization as we realize different sports require varying body shapes and sizes, but we were interested in any overarching patterns.

```
%%bq query
SELECT ROUND(min(BMI),1) AS min_BMI, round(max(BMI),1) AS max_BMI, ROUND(avg(BMI),1) A
S avg_BMI,Sex,Medal
FROM `cohort-b-team-1.olympic_data.BMI_VS_Medal`
Group by Sex,Medal
Order by Sex
```

min_BMI	max_BMI	avg_BMI	Sex	Medal
14.4	43.9	21.7	F	Gold
13.5	44.5	21.6	F	Bronze
15.0	51.0	21.6	F	Silver
15.9	50.8	24.0	M	Gold
16.3	56.5	23.9	M	Bronze
16.3	46.8	24.0	M	Silver

(rows: 6, time: 0.8s, 553KB processed, job: job_nb87h8RUIWo5sUJ1Syg2YWmnW_Ue)

```
%%bq query -n BMI_Medal
select Sex,avg_BMI,min_BMI,max_BMI
from(
SELECT ROUND(avg(BMI),1) AS avg_BMI,Sex,ROUND(min(BMI),1) AS min_BMI, round(max(BMI),
1) AS max_BMI
FROM `cohort-b-team-1.olympic_data.BMI_VS_Medal`
Group by Sex
Order by Sex)
```

```
%%chart columns -d BMI_Medal
title: BMI Analysis
hAxis:
  title: Gender
vAxis:
  title: BMI
legend: True
```

Country VS Medal

This part explains the relationship between counrty and medal won by each country

The following query counts three different medals numbers by country and year from the pervious builded sorted database stored in the GCP Bigquery. Order the result by year and country.

```
%%bq query
WITH Gold AS(
SELECT NOC, year, COUNT(Medal) AS gold_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.Sorted_Date_For_Country`
WHERE Medal LIKE 'Gold')
GROUP BY NOC, year),
Silver AS(
SELECT NOC, year, COUNT(Medal) AS silver_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.Sorted_Date_For_Country`
WHERE Medal LIKE 'Silver')
GROUP BY NOC, year),
Bronze AS(
SELECT NOC, year, COUNT(Medal) AS bronze_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.Sorted_Date_For_Country`
WHERE Medal LIKE 'Bronze')
GROUP BY NOC, year)


SELECT Gold.year AS year, Gold.NOC AS country, Gold.gold_num AS gold_num, Silver.silver_num AS si
lver_num, Bronze.bronze_num AS bronze_num
FROM Gold
INNER JOIN Silver
USING (NOC, year)
INNER JOIN Bronze
USING (NOC, year)
Order by year, NOC
```

year	country	gold_num	silver_num	bronze_num
1896	AUT	2	1	2
1896	DEN	1	2	3
1896	FRA	5	4	2
1896	GBR	3	3	3
1896	GER	7	5	2
1896	GRE	10	17	17
1896	HUN	2	1	3
1896	USA	11	6	2
1900	BEL	6	7	5
1900	DEN	2	3	2
1900	FRA	28	40	34
1900	GBR	19	11	12
1900	GER	4	3	1
1900	HUN	1	2	2
1900	NED	1	2	4
1900	SUI	6	2	1
1900	USA	21	16	17
1904	AUT	2	1	1
1904	CAN	4	1	1
1904	GER	4	5	7
1904	HUN	2	1	1
1904	USA	76	79	75
1906	AUT	3	3	3
1906	BEL	2	2	3
1906	DEN	3	2	1

(rows: 757, time: 1.0s, 2MB processed, job: job_kWWEDtFim3yJSENF30mb-MLKvH1b)

The following query counts the total medal number won by each country since 1896 categories the results into Gold, Silver and Bronze. Order the results by number of Gold Medal won.

```
%%bq query
SELECT t1.country as country, SUM(gold_num) as total_gold, t2.total_silver, t3.total_bronze
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`  t1

LEFT JOIN

(SELECT country as country, SUM(silver_num) as total_silver
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`
GROUP BY country) t2

ON t1.country = t2.country

LEFT JOIN

(SELECT country as country, SUM(bronze_num) as total_bronze
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`
GROUP BY country)t3

ON t1.country = t3.country

GROUP BY country,t2.total_silver, t3.total_bronze
ORDER BY total_gold DESC
```

country	total_gold	total_silver	total_bronze
USA	1131	901	791
URS	471	373	353
GER	319	345	340
GBR	286	318	306
FRA	264	283	325
ITA	246	219	229
CHN	239	177	168
RUS	201	179	196
SWE	197	206	234
GDR	192	165	162
HUN	173	151	176
NOR	155	153	140
JPN	151	144	175
FIN	146	137	162
AUS	143	161	178
CAN	124	155	184
NED	122	116	136
KOR	114	95	91
SUI	91	110	101
ROU	88	94	120
AUT	83	110	112
POL	74	87	132
CUB	72	60	70
FRG	67	80	91
EUN	54	44	37

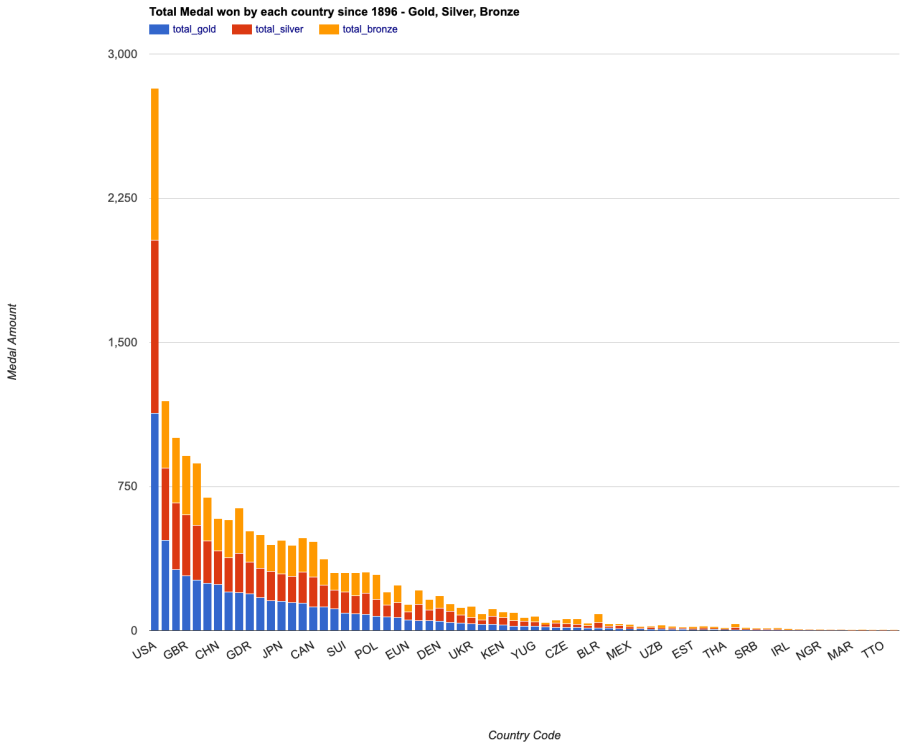
(rows: 71, time: 1.6s, 21KB processed, job: job_qDNsvtyOkK5nLr-ainoPoeJE964R)

The following stores the Result from the query above into "medal_vs_country"

```
%%bq query -n medal_vs_country
SELECT t1.country as country, SUM(gold_num) as total_gold, t2.total_silver, t3.total_bronze
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`  t1

LEFT JOIN

(SELECT country as country, SUM(silver_num) as total_silver
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`
GROUP BY country) t2

ON t1.country = t2.country

LEFT JOIN

(SELECT country as country, SUM(bronze_num) as total_bronze
FROM `cohort-b-team-1.olympic_data.Country_VS_Medal`
GROUP BY country)t3

ON t1.country = t3.country

GROUP BY country,t2.total_silver, t3.total_bronze
ORDER BY total_gold DESC
```

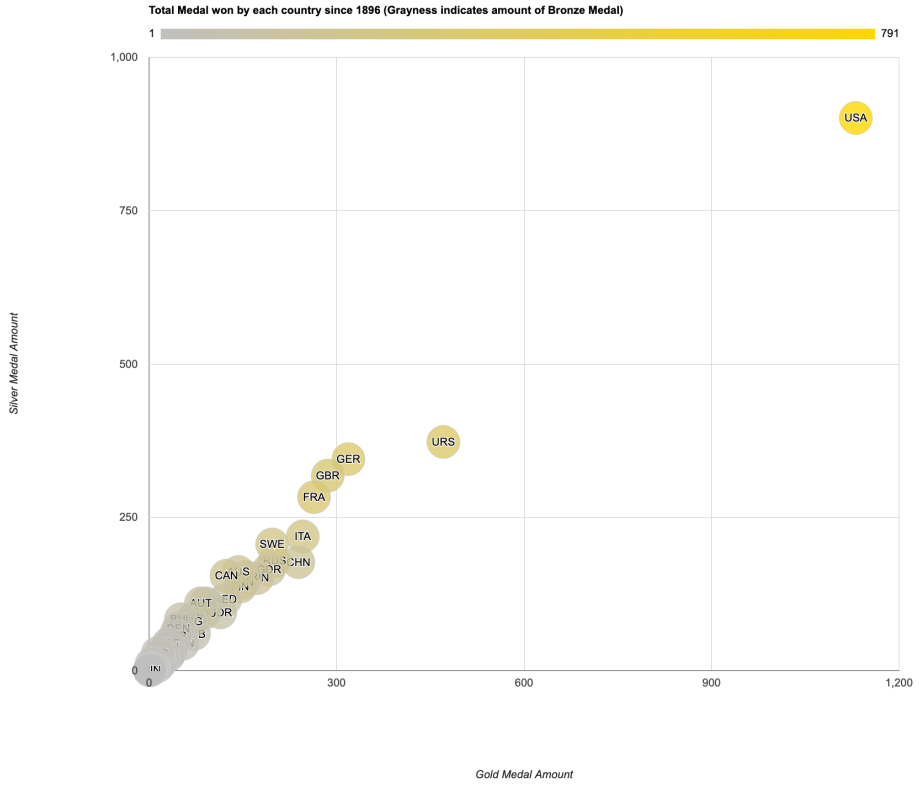
The following commend draws the Barchart base on "medal_vs_country"

```
%%chart columns --data medal_vs_country
title: Total Medal won by each country since 1896 - Gold, Silver, Bronze
height: 1500
width: 1950
hAxis:
  title: Country Code
vAxis:
  title: Medal Amount
isStacked:
  true
bar:
  { groupWidth: '75%' }
legend:
  {position: 'top', textStyle: {color: 'navy', fontSize: 16}}
```



The following commend draws the Bubblechart base on "medal_vs_country"

```
%%chart bubbles --data medal_vs_country
title: Total Medal won by each country since 1896 (Grayness indicates amount of Bronze Medal)
height: 1800
width: 2200
hAxis:
  title: Gold Medal Amount
vAxis:
  title: Silver Medal Amount
colorAxis:
  {colors: ['silver', 'gold']}
```



City VS Medal

The following table is the number of medals of each host city, listed under year.

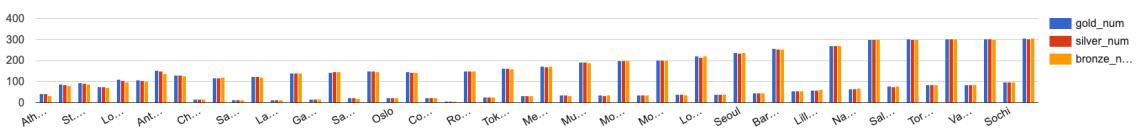
```
%%bq query
WITH Gold AS(
SELECT City, year, COUNT(Medal) AS gold_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.sorted_city`
WHERE Medal LIKE 'Gold')
GROUP BY City, year),
Silver AS(
SELECT City, year, COUNT(Medal) AS silver_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.sorted_city`
WHERE Medal LIKE 'Silver')
GROUP BY City, year),
Bronze AS(
SELECT City, year, COUNT(Medal) AS bronze_num
FROM (
SELECT * FROM `cohort-b-team-1.olympic_data.sorted_city`
WHERE Medal LIKE 'Bronze')
GROUP BY City, year)

SELECT Gold.City AS city, Gold.year AS year, Gold.gold_num AS gold_num, Silver.silver_num AS silver
_num, Bronze.bronze_num AS bronze_num
FROM Gold
LEFT JOIN Silver
USING (year,City)
LEFT JOIN Bronze
USING (year,City)
ORDER BY year
```

city	year	gold_num	silver_num	bronze_num
Athina	1896	43	41	32
Paris	1900	89	86	80
St. Louis	1904	95	92	88
Athina	1906	74	74	71
London	1908	109	103	97
Stockholm	1912	106	103	100
Antwerpen	1920	154	151	138
Paris	1924	129	130	128
Chamonix	1924	17	16	16
Amsterdam	1928	118	118	119
Sankt Moritz	1928	13	12	13
Los Angeles	1932	125	125	120
Lake Placid	1932	14	14	14
Berlin	1936	140	140	140
Garmisch-Partenkirchen	1936	17	17	17
London	1948	143	145	147
Sankt Moritz	1948	22	22	21
Helsinki	1952	149	149	146
Oslo	1952	22	22	22
Melbourne	1956	145	143	142
Cortina d'Ampezzo	1956	24	23	24
Stockholm	1956	6	6	6
Roma	1960	150	148	149
Squaw Valley	1960	27	26	27
Tokyo	1964	163	163	160

(rows: 52, time: 4.6s, 585KB processed, job: job_4DsgKX_mScwHfJUPNJW84Oqc6drQ)

```
%%bq query -n city_vs_medal
SELECT * EXCEPT (year)
FROM `cohort-b-team-1.olympic_data.City_VS_Medal`
```

```
%%chart columns --data city_vs_medal
```



As we can see from the bar chart, the number of gold, silver and bronze are almost equal among each Olympic game. Also, Summer Olypmic games' number of medals are much higher than Winter Olympic games in most case, which there has a increase trendency in the amount of medals if look separately.

Sex VS Medal

Description: The query below calculated the winning percentage by sex. For all female athletes who participated in Olympic, the rate of those who have medals is around 33%. For all male athletes, this number is about 28%.

Sex -> M as male athletes, F as female athletes Medal_ppl -> the sum number of athletes won the medal regardless of gold, silver, and bronze. Total_ppl -> the sum number of athletes who attends the Olympic Winning_Percentage -> (Medal_ppl / Total_ppl) *100

```
%%bq query
SELECT t1.Sex, count(Medal) AS Medal_ppl,total_ppl, ROUND((count(Medal) / total_ppl)*100,0) Winnin
g_Percentage
FROM `cohort-b-team-1.olympic_data.athlete_data` t1
LEFT JOIN
(SELECT Sex,COUNT(DISTINCT ID) AS total_ppl
FROM `cohort-b-team-1.olympic_data.athlete_data`
GROUP BY Sex) t2
on t1.Sex = t2.Sex
WHERE Medal != '0'
Group by Sex,total_ppl
Order by Sex DESC
```

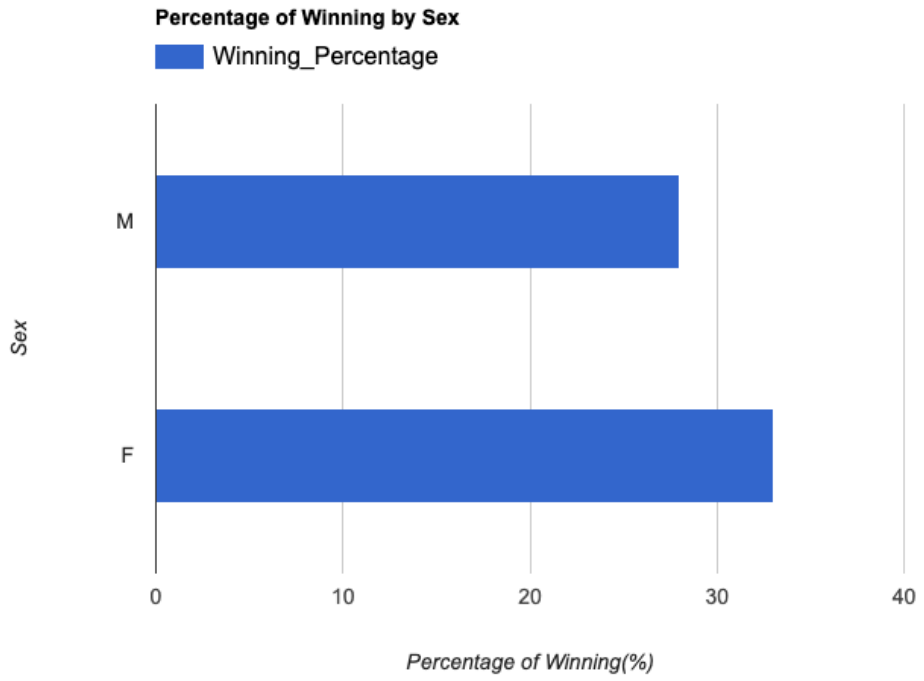
Sex	Medal_ppl	total_ppl	Winning_Percentage
M	28530	101590	28.0
F	11253	33981	33.0

(rows: 2, time: 1.6s, 3MB processed, job: job_ktnj2aqyfnKFI59lh4deBTvhlbJP)

```
%%bq query -n sex_Medal
SELECT t1.Sex, ROUND((count(Medal) / total_ppl)*100,0) Winning_Percentage
FROM `cohort-b-team-1.olympic_data.athlete_data` t1
LEFT JOIN
(SELECT Sex,COUNT(DISTINCT ID) AS total_ppl
FROM `cohort-b-team-1.olympic_data.athlete_data`
GROUP BY Sex) t2
on t1.Sex = t2.Sex
WHERE Medal != '0'
Group by t1.Sex,total_ppl
```

Percentage of Winning by Sex

```
%%chart bars --data sex_Medal
title: Percentage of Winning by Sex
height: 500
width: 800
legend:
  {position: 'top', textStyle: {color: 'black', fontSize: 16}}
hAxis:
  title: Percentage of Winning(%)
vAxis:
  title: Sex
bar:
  { groupWidth: '40%' }
isStacked: True
```



Age VS Medal

Description: Compared average medal-winning age by gender. The result below illustrates that in general, female medal-winner is younger than male medal-winner. However, the male medal winner also tends to have larger variations on the age range.

```
%%bq query
SELECT Sex, Medal, ROUND(AVG(Age),1) AS Average_Age, ROUND(STDDEV(Age),1) AS Standard_
Deviation
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Medal != '0'
Group by Sex, Medal
Order by Sex DESC, Medal;
```

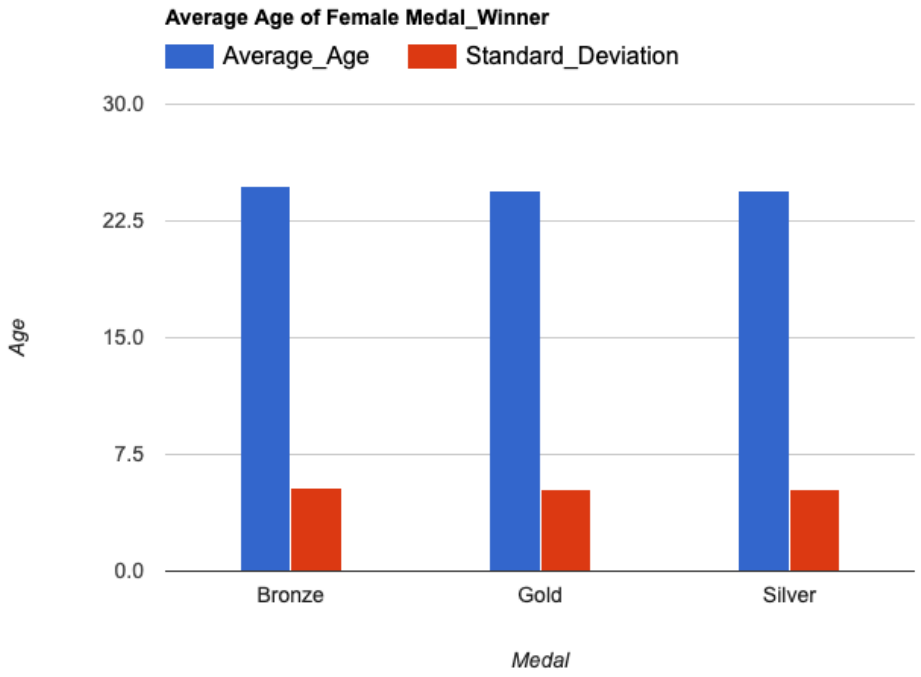
Sex	Medal	Average_Age	Standard_Deviation
M	Bronze	25.6	7.4
M	Gold	26.1	6.8
M	Silver	25.8	7.6
F	Bronze	24.7	5.4
F	Gold	24.4	5.3
F	Silver	24.4	5.3

(rows: 6, time: 0.7s, 3MB processed, job: job_GPd_Sgr-Njz5P5Eu8GjdV1iU__5b)

```
%%bq query -n Age_Medal_Female
SELECT Medal, ROUND(AVG(Age),1) AS Average_Age, ROUND(STDDEV(Age),1) AS Standard_Devia
tion
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Medal != '0'
Group by Sex, Medal
HAVING Sex ='F'
Order by Sex DESC, Medal
```

Average Age and Standard Deviation of Female Athlete Won the Medal

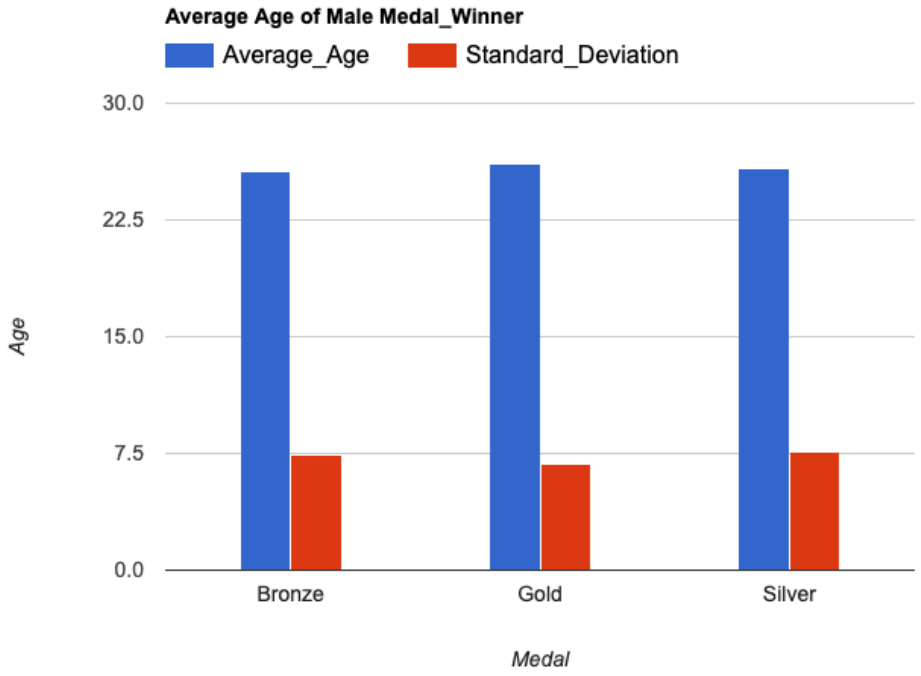
```
%%chart columns --data Age_Medal_Female
title: Average Age of Female Medal_Winner
height: 500
width: 800
legend:
  {position: 'top', textStyle: {color: 'black', fontSize: 16}}
hAxis:
  title: Medal
vAxis:
  title: Age
bar:
  { groupWidth: '40%' }
```




```
%%bq query -n Age_Medal_Male
SELECT Medal, ROUND(AVG(Age),1) AS Average_Age, ROUND(STDDEV(Age),1) AS Standard_Devia
tion
FROM `cohort-b-team-1.olympic_data.athlete_data`
WHERE Medal != '0'
Group by Sex, Medal
HAVING Sex ='M'
Order by Sex DESC, Medal
```

Average Age and Standard Deviation of Male Athlete Won the Medal

```
%%chart columns --data Age_Medal_Male
title: Average Age of Male Medal_Winner
height: 500
width: 800
legend:
  {position: 'top', textStyle: {color: 'black', fontSize: 16}}
hAxis:
  title: Medal
vAxis:
  title: Age
bar:
  { groupWidth: '40%' }
```



Summary

This exploratory research revealed a few interesting trends in among Olympic medal winners. Over time, the chances for winning olympic medals have increased dramatically, in both Winter and Summer. There are some differences between male and female medal winners, both physically and in among other personalistic traits. Additionally, the United States is unsurprisingly a clear winner in the total amount of medals won since 1896. Overall, this exploratory research is a good start into looking at what makes an Olympic medal winner, but lacks in it's predictive capacities (future work needed).