Data Visualization Project Proposal

Basic Info:

Project Title: OLYMPICS STATS

Names: Krunal Jain, Abhishek Lakshminarayana, Vinod Reddy Gooty

Emails : krunal.bitsy@gmail.com , abhishekl9426@gmail.com, gootyvinod12@gmail.com

UID: u1140629, u1145084, u1142935

Project Repository Link: https://github.com/krunaljain/dataviscourse-olympics_stats

Background and Motivation: All countries vouch for success at Olympic games, which are held once every 4 years. The performance of various countries over the years has shown some interesting trends. The aim of our project is to capture and visualize this trend and provide an insight into how different regions across the world fared in these games.

Project Objectives: Throughout the course, we studied different visualization techniques and their applications in specific use cases. Through this project, we wish to incorporate these visualizations to highlight the Olympic performances of different countries over the years.

Data: The dataset we used is available at
https://www.kaggle.com/hugomathien/soccer/data

Data Processing: We have two kind of views in our project:
1) Basic Views: For these views, we can directly use the data available in the CSV file and do projections to filter out columns depending on our requirement.
2) Aggregated views: For these type of views, the data is aggregated based on some parameter. So, we need to do data processing.

Visualization Design: We plan to use different visualization elements depending on the data we wish to present. Some of those elements are:
1) Map
2) Area Chart
3) Bar Chart
4) Stacked Bar Chart
5) Pie Chart

Must-Have Features:

1) Medal count by countries:
- Attributes Summary:
  - Individual country medal count for each year.
  - Distribution of medals for each country.
  - Distribution of different types of medals (Gold/Silver/Bronze) of particular country.

2) Aggregated stats for a country for a fixed time period:

    We fix a time period spanning multiple Olympics and compare the performance of different countries over that time period. This visualization helps us in comparing the performances of different countries over the years and study the trend in distribution of medals.

3) Variations in total medal counts for all Olympics:

    We study the variation in total medal count for all Olympics and try to reason about the trends observed.

Optional Features:

- Evaluation of performance of host nations across Olympics.

- Effect of GDP on medals count: We wish to analyze the effect of GDP on the country's performance in Olympics.

- Effect of sex ratio on medal count: We wish to analyze the effect of GDP on the country's performance in Olympics.

Project Schedule:

Week 1:

- Data preprocessing.
- Map showing medals by countries.

Week 2:

- Map showing the aggregated statistics.
- Top countries in each Olympics

Week 3:

- Medal distribution for top sports

Week 4:

- Total medal count over years
- Additional work

Comparison between alternative designs shown in picture below:

Medals count by country:

O Alternate 1: World map:

Pros:

It gives spatial context to our data unlike other designs. It's very intuitive for the audience to relate to particular country when shown through a map.

O Alternatives: We've explored many alternatives like Tables, BarCharts, etc for this. However, nothing makes data so easily discernible like World Maps.

Aggregate stats of each country over the years:

o Alternate 1: Map

Pros:

As discussed earlier, world maps have the added advantage of giving the audience spatial context.

O Alternatives: Tables or any other views lack the spatial context, and it's also too cumbersome to list so many countries in the available space.

Top countries in each Olympics:

o Alternate 1: Pie Chart

Pros:

Comparing fixed number of countries with three different medal categories looks good on pie chart.

o Alternate 2: Stacked Bar Chart

o Alternate 3: Bubble Chart

Variation in medal counts over the years:

o Alternate 1: Area Chart:

Pros:

Trend over time can be understood easily.

o Alternate 2: Stacked Bar Chart

o Alternate 3: Bubble Chart

Medal distribution for sports over years:

o Alternate 1: Stacked Bar Chart

o Alternate 2: Pie Chart

o Alternate 3: Progress Circle Chart

Evaluation of host nation on medal count:

o Alternate 1: Line Chart

o Alternate 2: Bubble Chart

o Alternate 3: Grouped Bar chart

We have planned on using Alternative 1 on most of the part unless we come up with anything interesting while working on project.
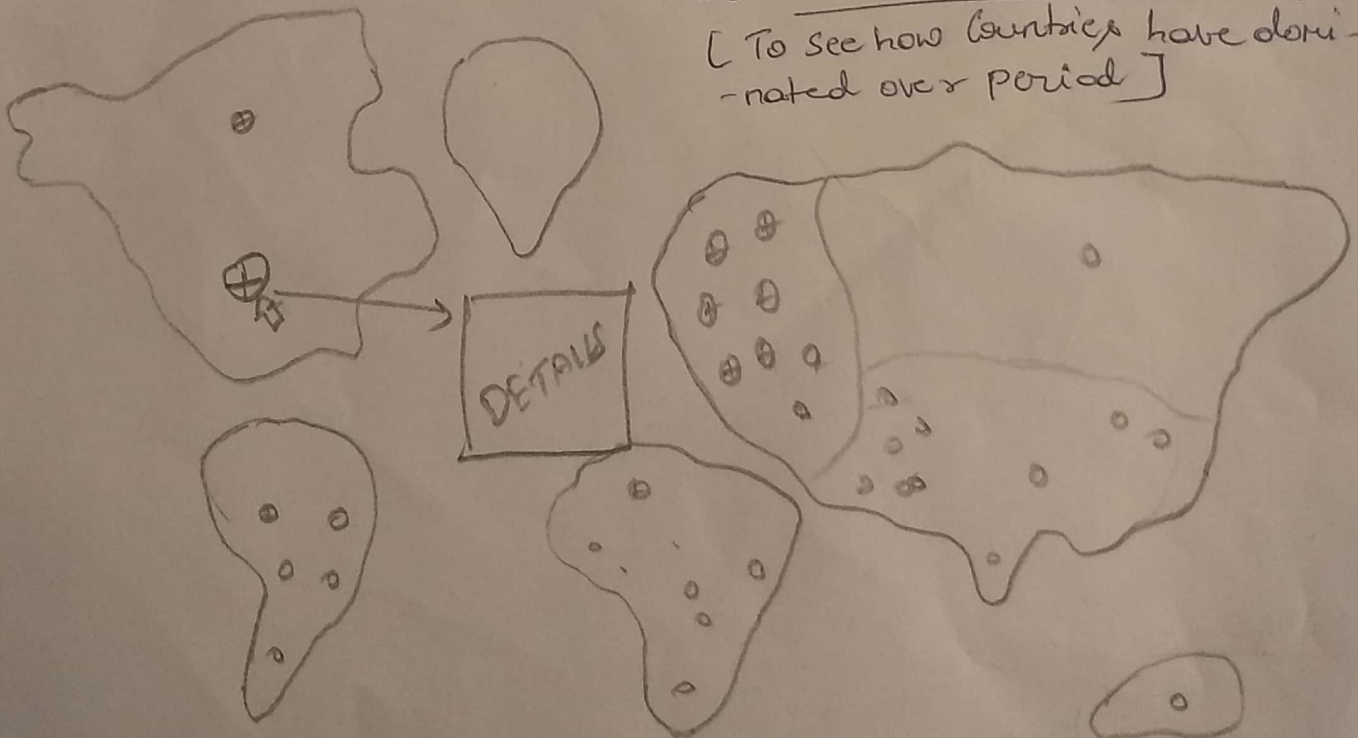
MEDAL DISTRIBUTION OVER YEARS

DETAILS

1896 ——————————————————— 2016

MEDAL DISTRIBUTION OVER QUARTER

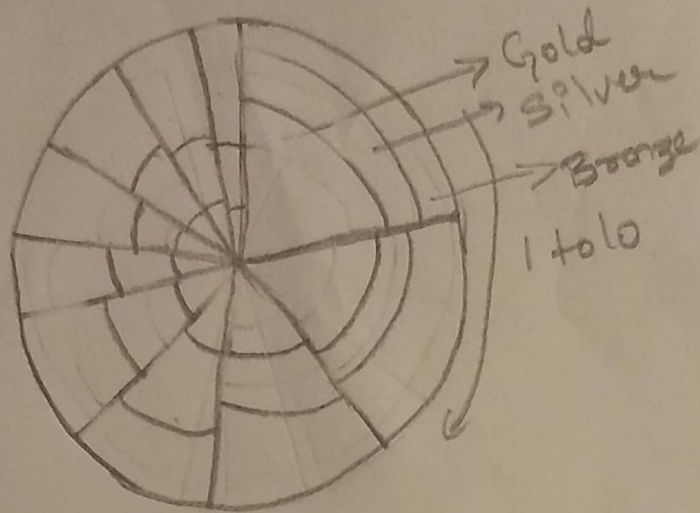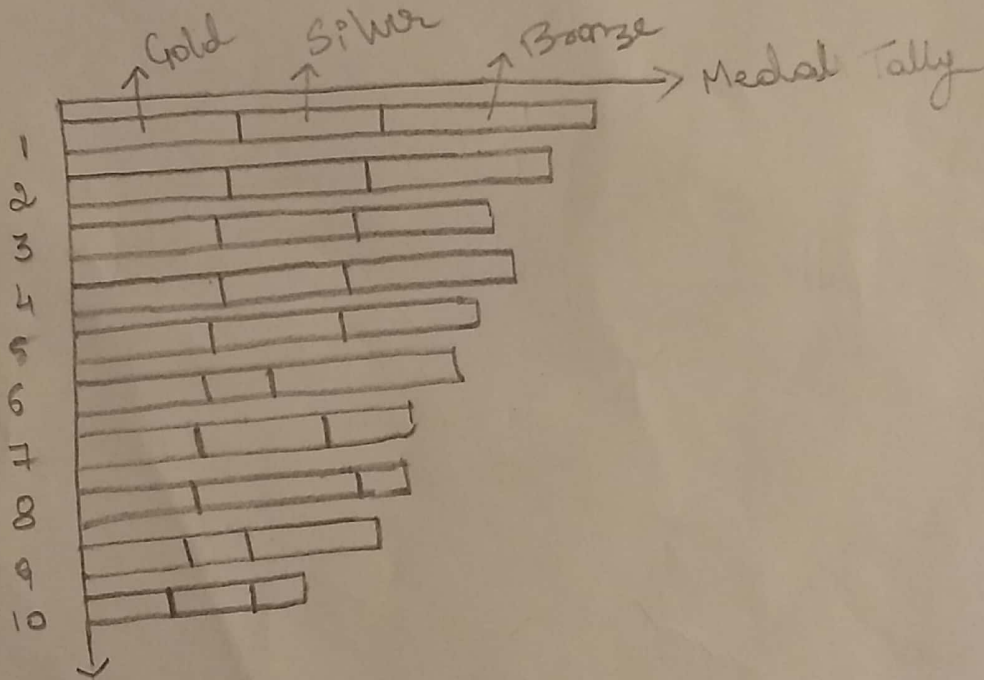[To see how countries have domi-
-nated over period]

DETAILS

Quarters
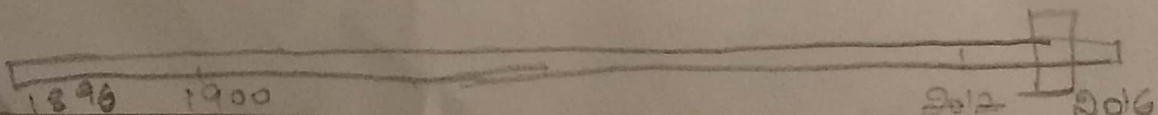[6]   1 2 3 4 5 6

TOP 10 COUNTRIES IN EACH OLYMPICS

1)

Gold
Silver
Bronze

1 to 10

2)

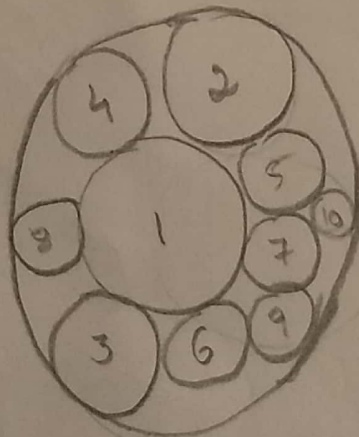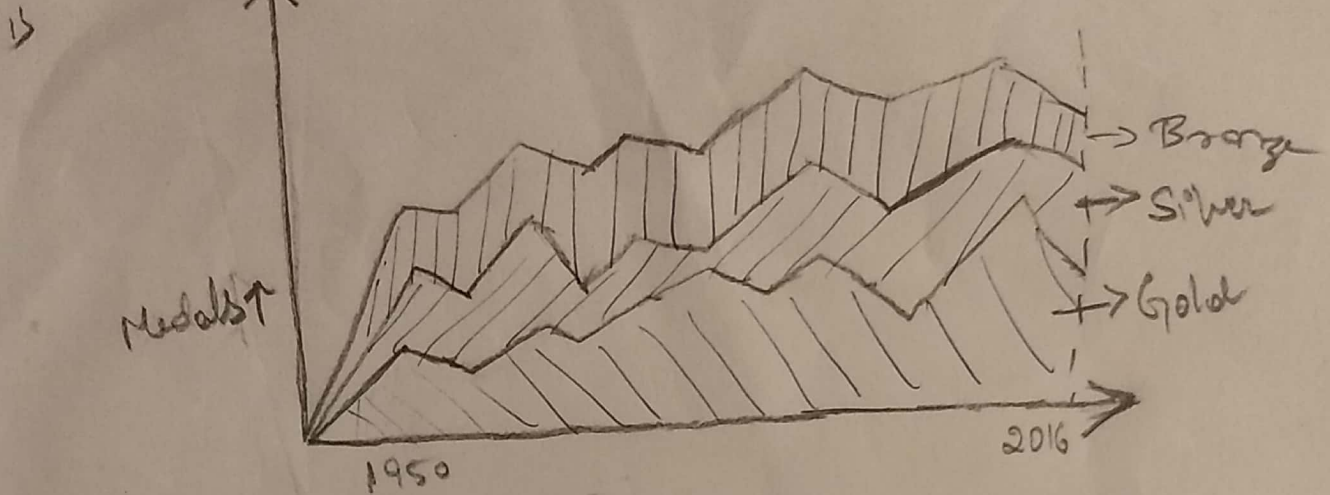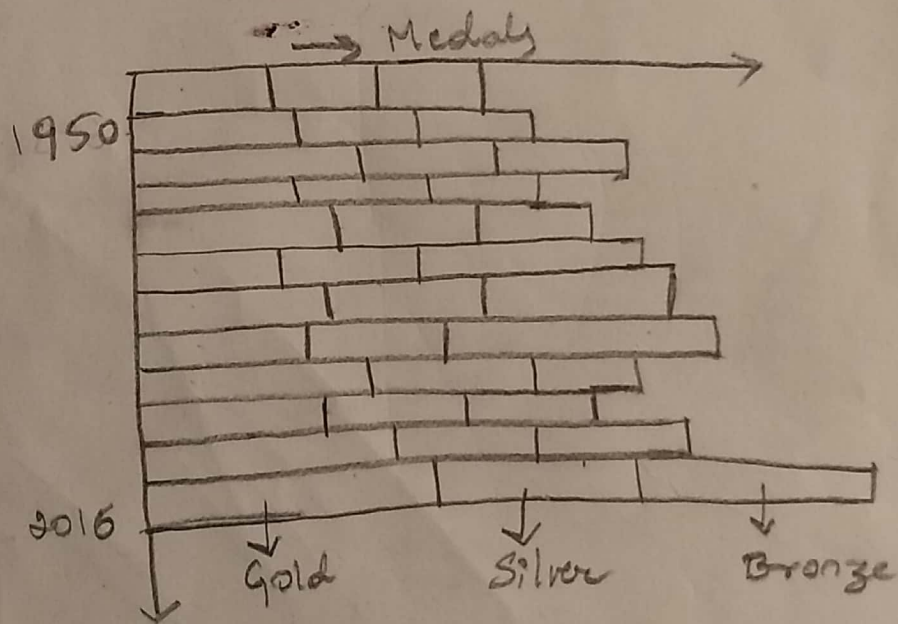Gold   Silver   Bronze   Medal Tally
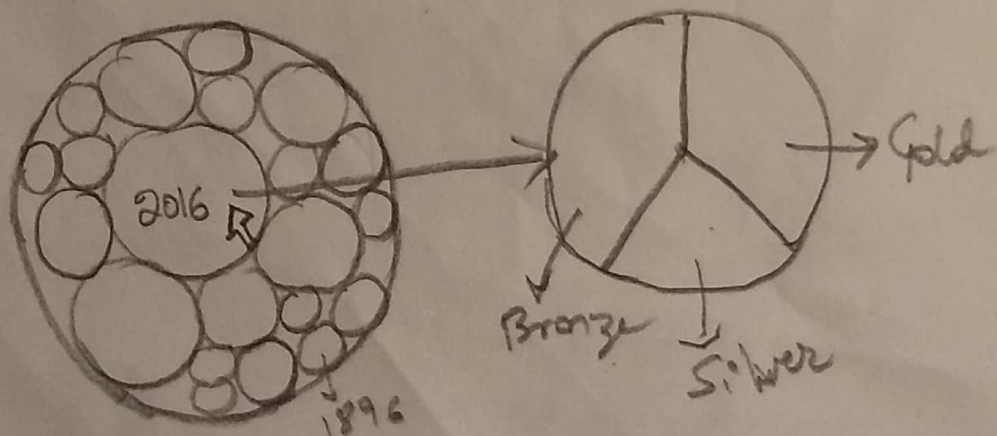
1
2
3
4
5
6
7
8
9
10

3)

1890   1900   2012   2016

TOTAL MEDAL COUNT OVER YEARS

1)

Medals ↑

1950

2016

→ Bronze
→ Silver
→ Gold

2)

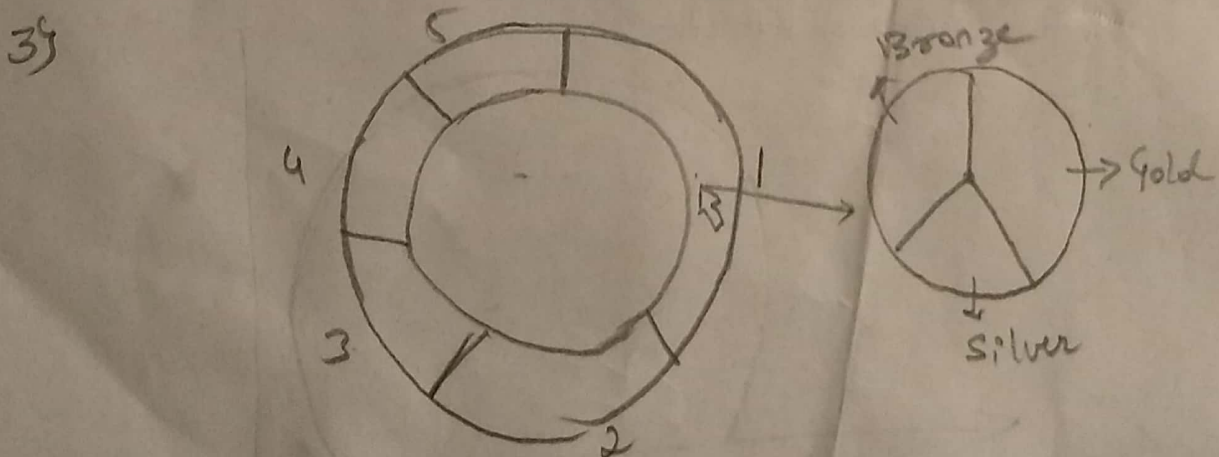

→ Medals

1950

2016

Gold   Silver   Bronze

3)

2016

1896

→ Gold

Bronze   Silver

MEDAL DISTRIBUTION FOR TOP 5 SPORTS OVER YEARS

1)

Medals →

Sports ↓

1
2
3
4
5

Gold   Silver   Bronze

2)

1 to 5

Bronze   Silver   Gold

3)

5
4
3
2
1

Bronze
Gold
Silver
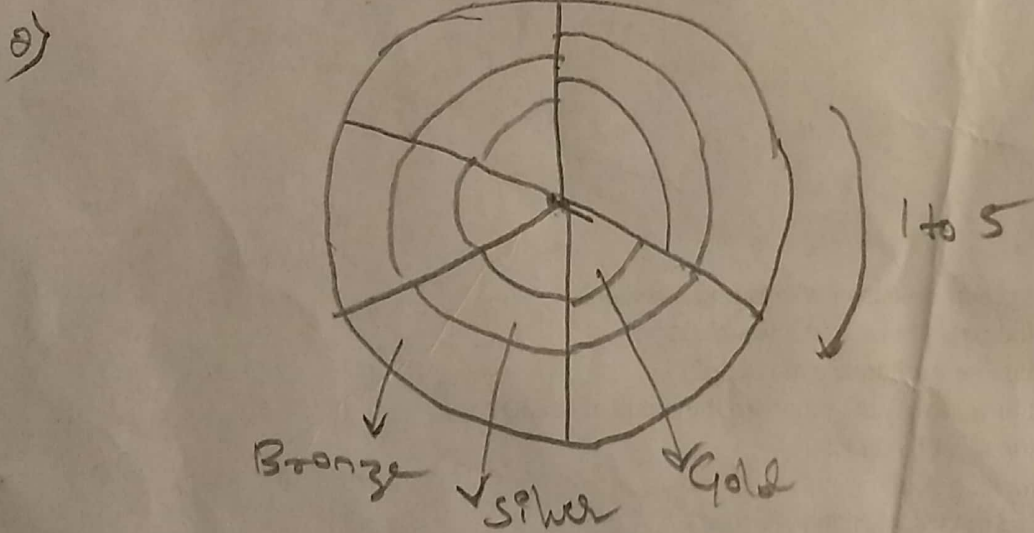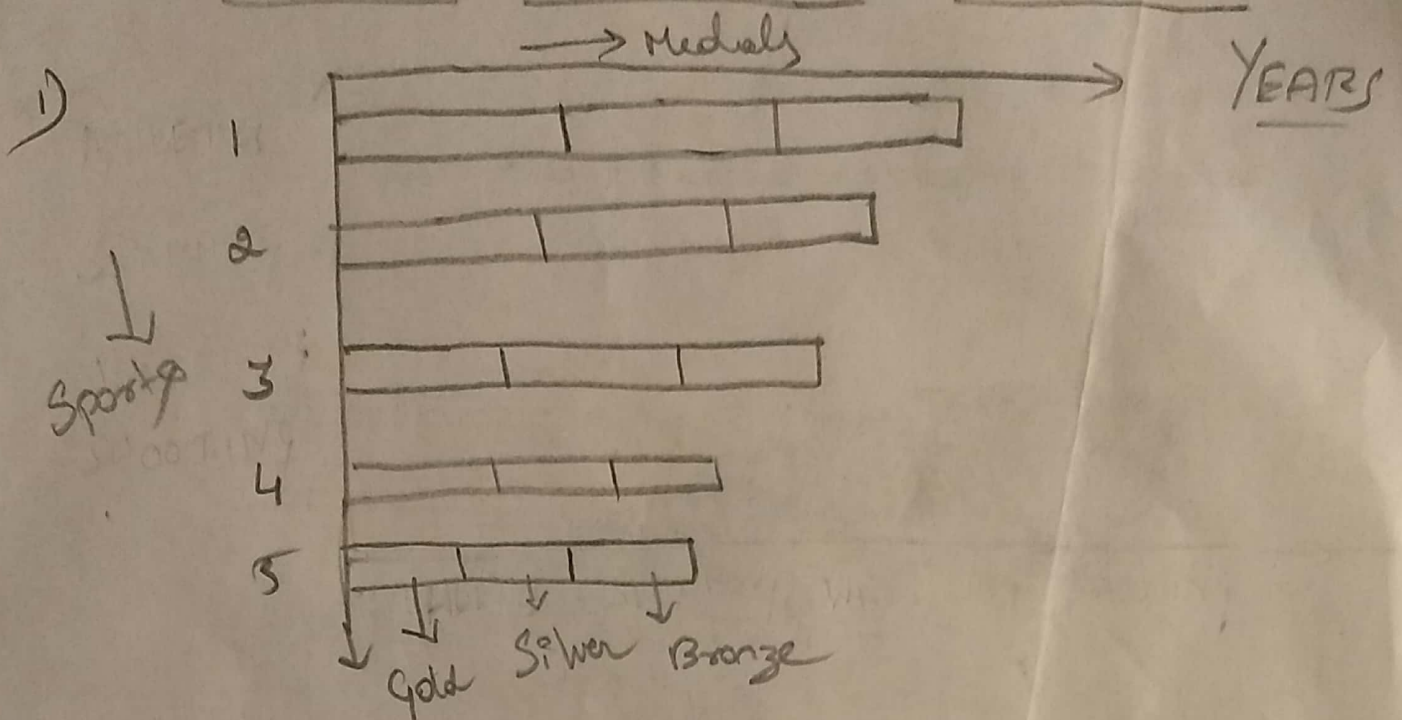
1896   2016

EFFECTIVE OF HOST NATION ON MEDAL COUNT

AUS — MEDALS — 300 / 200 / 100

CAN — 60 / 50 / 40

CHN — 900 / 800 / 700

ATHENS   ATLANTA   BARCELONA   SYDNEY.

OLYMPICS

ALTERNATIVE 1.

AUS

CAN

CHN

2001   YEAR

☐ AUS
☐ CAN
☐ CAN

ALTERNATIVE 2

Medals

ATHENS   BARCELONA   SYDNEY