Copy these 2 codes:

```
import kagglehub


path = kagglehub.dataset_download("heesoo37/120-years-of-olympic-history-athletes-and-results")

print("Path to dataset files:", path)
```

```
import pandas as pd

df = pd.read_csv(path + "/athlete_events.csv")
```

Dataset itself:

https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-results?resource=download

-- Step 0: Dataset overview

- Print dataset shape
- Print column names
- Display first 5 rows

-- Step 1: Olympic Games over time (line plot)

- Count number of unique Olympic Games per year
- Plot **number of events per year**
- X-axis: Year
- Y-axis: Number of events

-- Step 2: Athlete participation growth

- Count number of unique athletes per year
- Plot a **line chart** showing participation growth
- Add grid

-- Step 3: Gender participation comparison

- Count male vs female athletes
- Create a **bar chart**
- X-axis: Gender
- Y-axis: Number of athletes

-- Step 4: Top 10 countries by medals

- Count total medals per country (ignore NaN medals)
- Select top 10 countries
- Plot a **horizontal bar chart**

-- Step 5: Medal distribution

- Count number of Gold, Silver, Bronze medals
- Plot a **pie chart** showing medal proportions

-- Step 6: Athlete age distribution

- Remove rows with missing age
- Plot a **histogram** of athlete ages
- Choose an appropriate bin count

-- Step 7: Age vs medal (scatter plot)

- Select athletes who won medals
- Create a **scatter plot**:
  – X-axis: Age
  – Y-axis: Medal type (encoded numerically)

-- Step 8: Height vs weight by gender

- Remove rows with missing Height or Weight
- Create a **scatter plot**
- Use different markers for Male and Female

-- Step 9: Sports popularity

- Count number of athletes per sport
- Select top 15 sports
- Plot a **bar chart**

-- Step 10: Summer vs Winter Olympics

- Count number of athletes in Summer and Winter games
- Plot a **bar chart** comparison

-- Step 11: Country performance over time

- Select one country (e.g., USA)
- Count medals per year
- Plot a **line chart**

-- Step 12: Subplots dashboard

- Create **2×2 subplots**:

1. Athletes per year
2. Medal distribution
3. Gender count
4. Top sports

-- Step 13: Save visualization

- Save the dashboard as "olympics_analysis.png"
- Use tight layout