RA2211003010979_KOTIPALLI_SRIKESH_PBLIComponent_PracticeExpe

September 21, 2024

[1]:
```
!pip install pandas numpy matplotlib seaborn
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages
(2.1.4)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages
(1.26.4)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-
packages (3.7.1)
Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-
packages (0.13.1)
Requirement already satisfied: python-dateutil>=2.8.2 in
/usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-
packages (from pandas) (2024.2)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-
packages (from pandas) (2024.1)
Requirement already satisfied: contourpy>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-
packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.7)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-
packages (from matplotlib) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.4)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-
packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
```

[2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

[3]:
```
df=pd.read_csv('/content/Olympic_Medal_Tally_History.csv')
```

[4]:
```
df.head()
```

[4]:
```
                edition  edition_id  year        country country_noc  gold  \
0  1896 Summer Olympics           1  1896  United States         USA    11
1  1896 Summer Olympics           1  1896         Greece         GRE    10
2  1896 Summer Olympics           1  1896        Germany         GER     6
3  1896 Summer Olympics           1  1896         France         FRA     5
4  1896 Summer Olympics           1  1896  Great Britain         GBR     2

   silver  bronze  total
0       7       2     20
1      18      19     47
2       5       2     13
3       4       2     11
4       3       2      7
```

[6]:
```
df.tail()
```

[6]:
```
                edition  edition_id  year      country country_noc  gold  \
1802  1906 Intercalated           4  1906       Canada         CAN     1
1803  1906 Intercalated           4  1906       Norway         NOR     1
1804  1906 Intercalated           4  1906  Netherlands         NED     0
1805  1906 Intercalated           4  1906    Australia         AUS     0
1806  1906 Intercalated           4  1906      Bohemia         BOH     0

      silver  bronze  total
1802       1       0      2
1803       1       0      2
1804       1       2      3
1805       0       3      3
1806       0       2      2
```

[9]:
```
df.columns
```

[9]:
```
Index(['edition', 'edition_id', 'year', 'country', 'country_noc', 'gold',
       'silver', 'bronze', 'total'],
      dtype='object')
```

[10]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1807 entries, 0 to 1806
Data columns (total 9 columns):
 #   Column       Non-Null Count  Dtype
```

```
---  ------       --------------  -----
 0   edition      1807 non-null   object
 1   edition_id   1807 non-null   int64
 2   year         1807 non-null   int64
 3   country      1807 non-null   object
 4   country_noc  1807 non-null   object
 5   gold         1807 non-null   int64
 6   silver       1807 non-null   int64
 7   bronze       1807 non-null   int64
 8   total        1807 non-null   int64
dtypes: int64(6), object(3)
memory usage: 127.2+ KB
```

[11]:
```
df.describe()
```

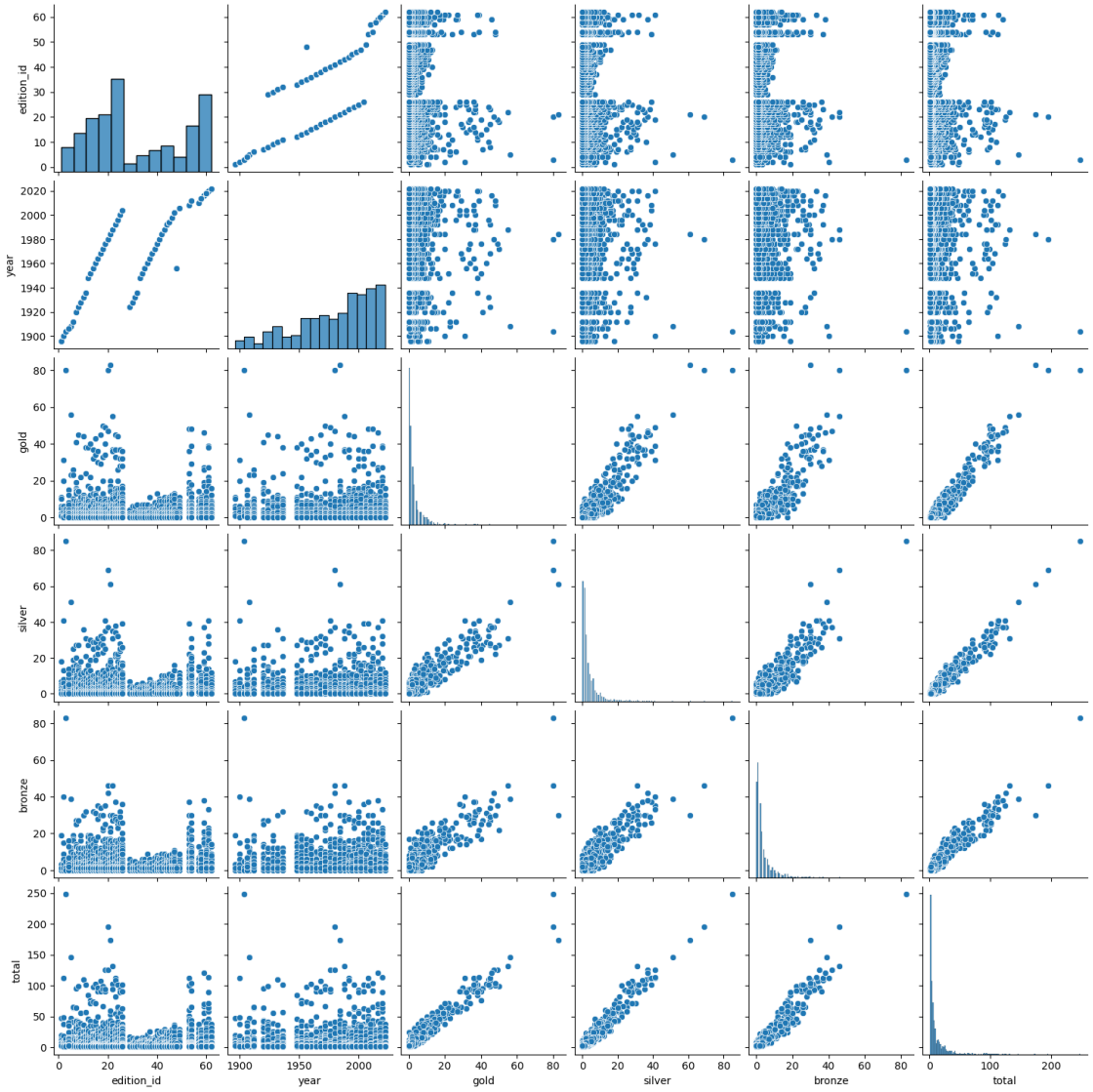
[11]:
```
         edition_id         year         gold       silver       bronze  \
count  1807.000000  1807.000000  1807.000000  1807.000000  1807.000000
mean     31.635307  1979.744328     3.737133     3.721638     3.971223
std      18.472012    32.726372     7.554092     6.411636     6.169554
min       1.000000  1896.000000     0.000000     0.000000     0.000000
25%      17.000000  1960.000000     0.000000     1.000000     1.000000
50%      25.000000  1988.000000     1.000000     2.000000     2.000000
75%      53.000000  2008.000000     4.000000     4.000000     5.000000
max      62.000000  2022.000000    83.000000    85.000000    83.000000

             total
count  1807.000000
mean     11.429994
std      19.423201
min       1.000000
25%       2.000000
50%       5.000000
75%      13.000000
max     248.000000
```

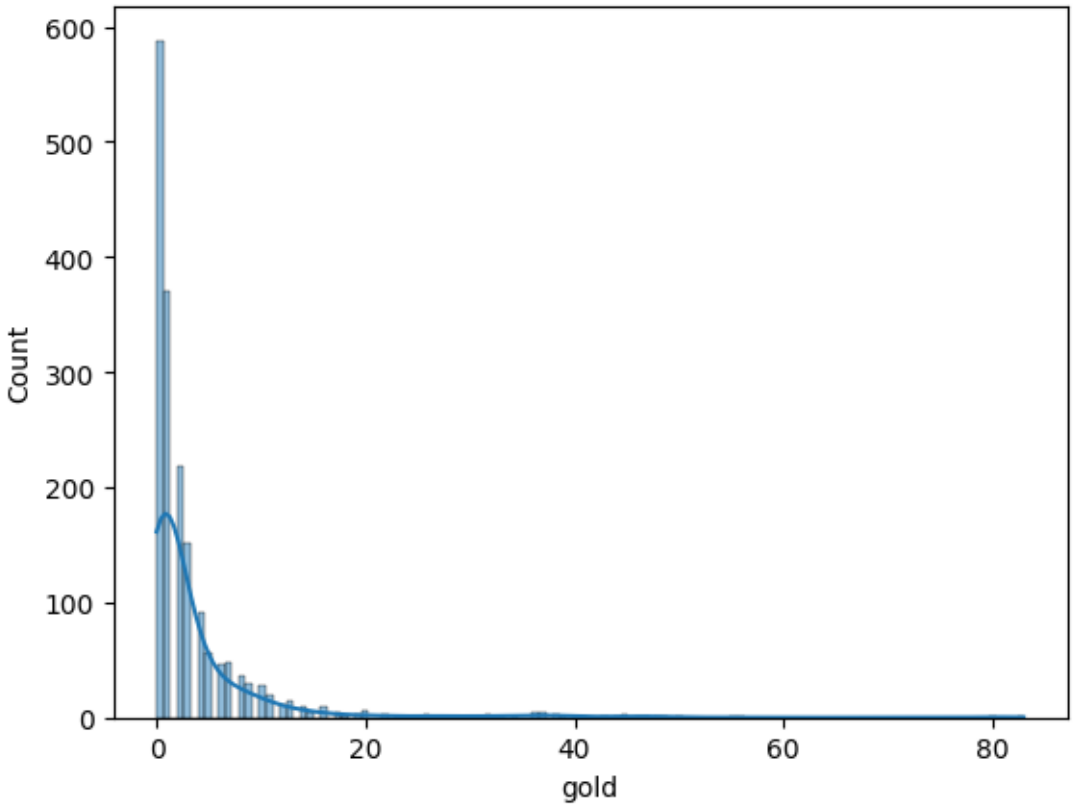
[12]:
```
df.isnull().sum()
```

[12]:
```
edition        0
edition_id     0
year           0
country        0
country_noc    0
gold           0
silver         0
bronze         0
total          0
dtype: int64
```

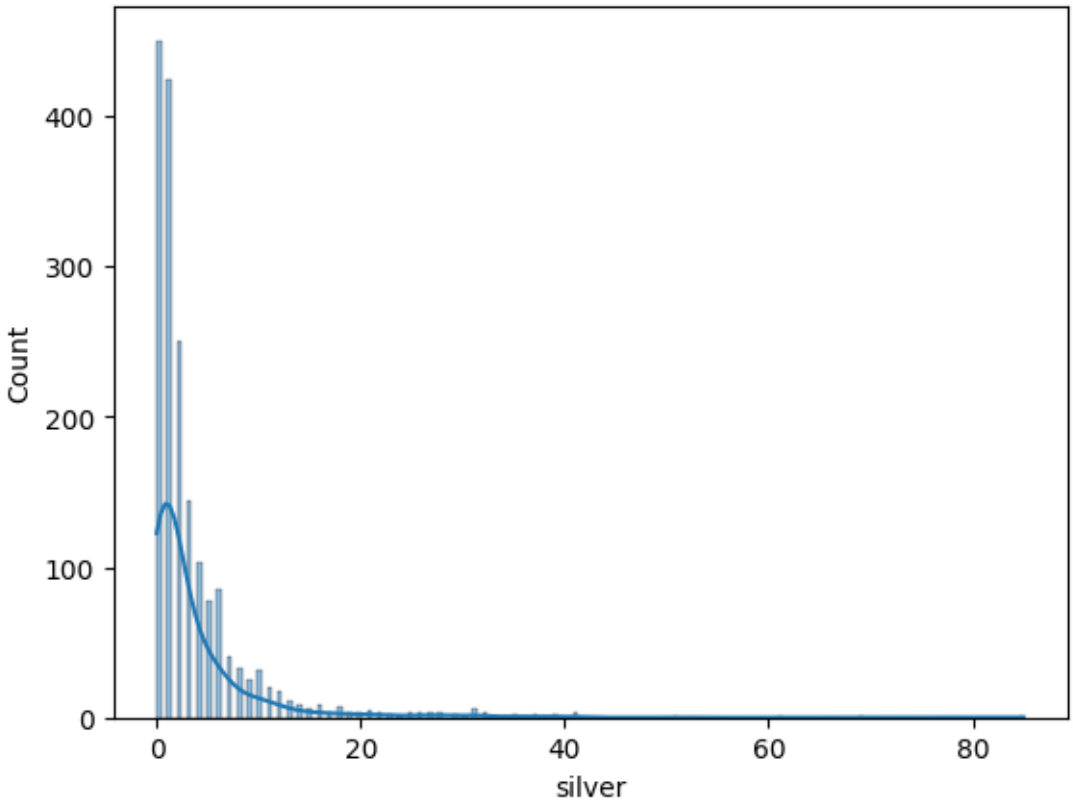
[23]:
```
sns.pairplot(df)
plt.show()
```



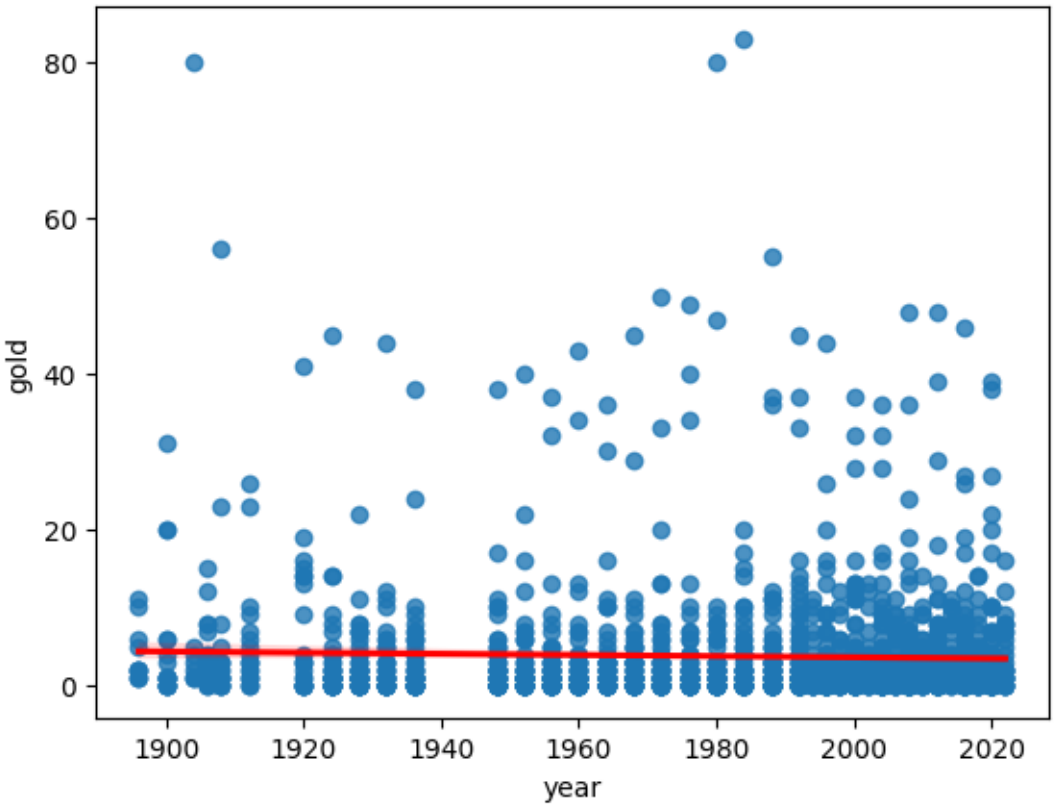
[30]:
```
sns.histplot(df, x="gold", kde=True)
plt.show()
```



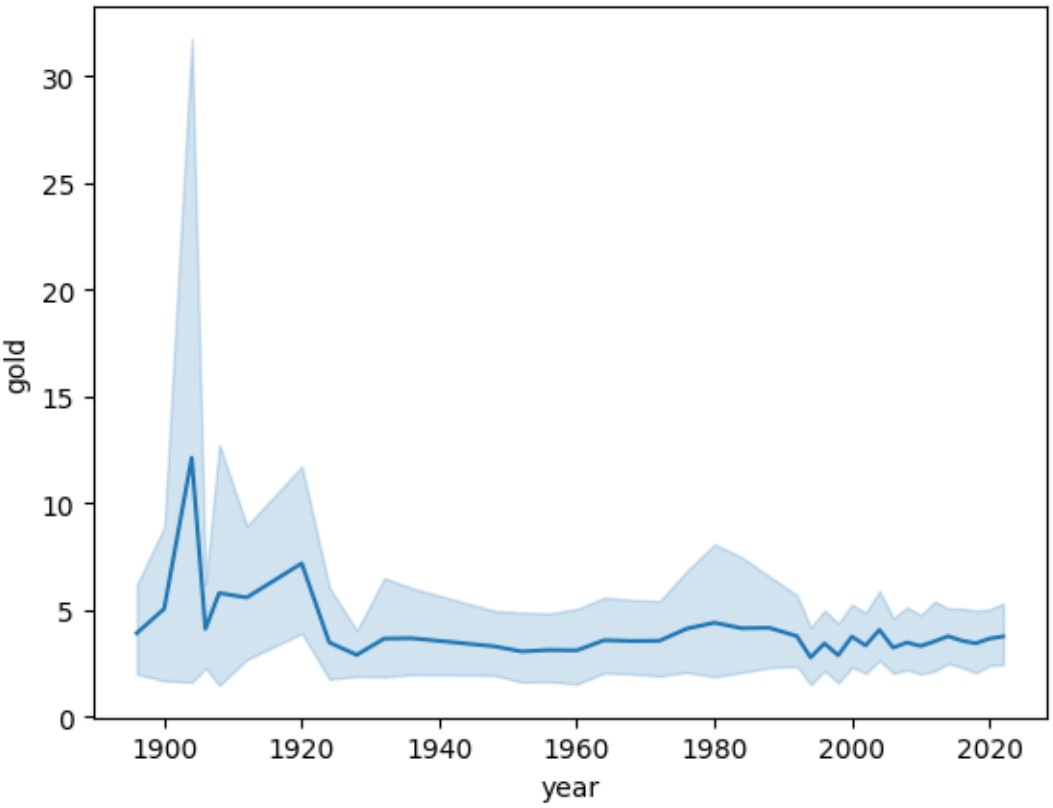
[31]:
```
sns.histplot(df, x="silver", kde=True)
plt.show()
```



[38]:
```
sns.regplot(x='year',y='gold',data=df,line_kws={"color": "red"})
plt.show()
```



[41]:
```
sns.lineplot(x='year',y='gold',data=df)
plt.show()
```


gold
year
0
5
10
15
20
25
30
1900
1920
1940
1960
1980
2000
2020

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:02 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect Gemini

```
!pip install pandas numpy matplotlib seaborn
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (2.1.4)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (1.26.4)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-packages (0.13.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.1)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.7)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.4)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
```

```
[ ] import pandas as pd
    import numpy as np
    import matplotlib.pyplot as plt
    import seaborn as sns
```

```
[ ] df=pd.read_csv('/content/Olympic_Medal_Tally_History.csv')
```

```
[ ] df.head()
```

	edition	edition_id	year	country	country_noc	gold	silver	bronze	total

```
[ ] df=pd.read_csv('/content/Olympic_Medal_Tally_History.csv')
```

```
df.head()
```

	edition	edition_id	year	country	country_noc	gold	silver	bronze	total
0	1896 Summer Olympics	1	1896	United States	USA	11	7	2	20
1	1896 Summer Olympics	1	1896	Greece	GRE	10	18	19	47
2	1896 Summer Olympics	1	1896	Germany	GER	6	5	2	13
3	1896 Summer Olympics	1	1896	France	FRA	5	4	2	11
4	1896 Summer Olympics	1	1896	Great Britain	GBR	2	3	2	7

```
[ ] df.tail()
```

	edition	edition_id	year	country	country_noc	gold	silver	bronze	total
1802	1906 Intercalated	4	1906	Canada	CAN	1	1	0	2
1803	1906 Intercalated	4	1906	Norway	NOR	1	1	0	2
1804	1906 Intercalated	4	1906	Netherlands	NED	0	1	2	3
1805	1906 Intercalated	4	1906	Australia	AUS	0	0	3	3
1806	1906 Intercalated	4	1906	Bohemia	BOH	0	0	2	2

```
[ ] df.columns
```

```
Index(['edition', 'edition_id', 'year', 'country', 'country_noc', 'gold',
       'silver', 'bronze', 'total'],
      dtype='object')
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1807 entries, 0 to 1806
Data columns (total 9 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   edition      1807 non-null   object
 1   edition_id   1807 non-null   int64
 2   year         1807 non-null   int64
 3   country      1807 non-null   object
 4   country_noc  1807 non-null   object
 5   gold         1807 non-null   int64
 6   silver       1807 non-null   int64
 7   bronze       1807 non-null   int64
 8   total        1807 non-null   int64
dtypes: int64(6), object(3)
memory usage: 127.2+ KB
```

```
[ ] df.describe()
```

	edition_id	year	gold	silver	bronze	total
count	1807.000000	1807.000000	1807.000000	1807.000000	1807.000000	1807.000000
mean	31.635307	1979.744328	3.737133	3.721638	3.971223	11.429994

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:02 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect Gemini

max	62.000000	2022.000000	83.000000	85.000000	83.000000	248.000000

```
df.isnull().sum()
```

	0
edition	0
edition_id	0
year	0
country	0
country_noc	0
gold	0
silver	0
bronze	0
total	0

dtype: int64

```
[ ] sns.pairplot(df)
    plt.show()
```

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:02 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

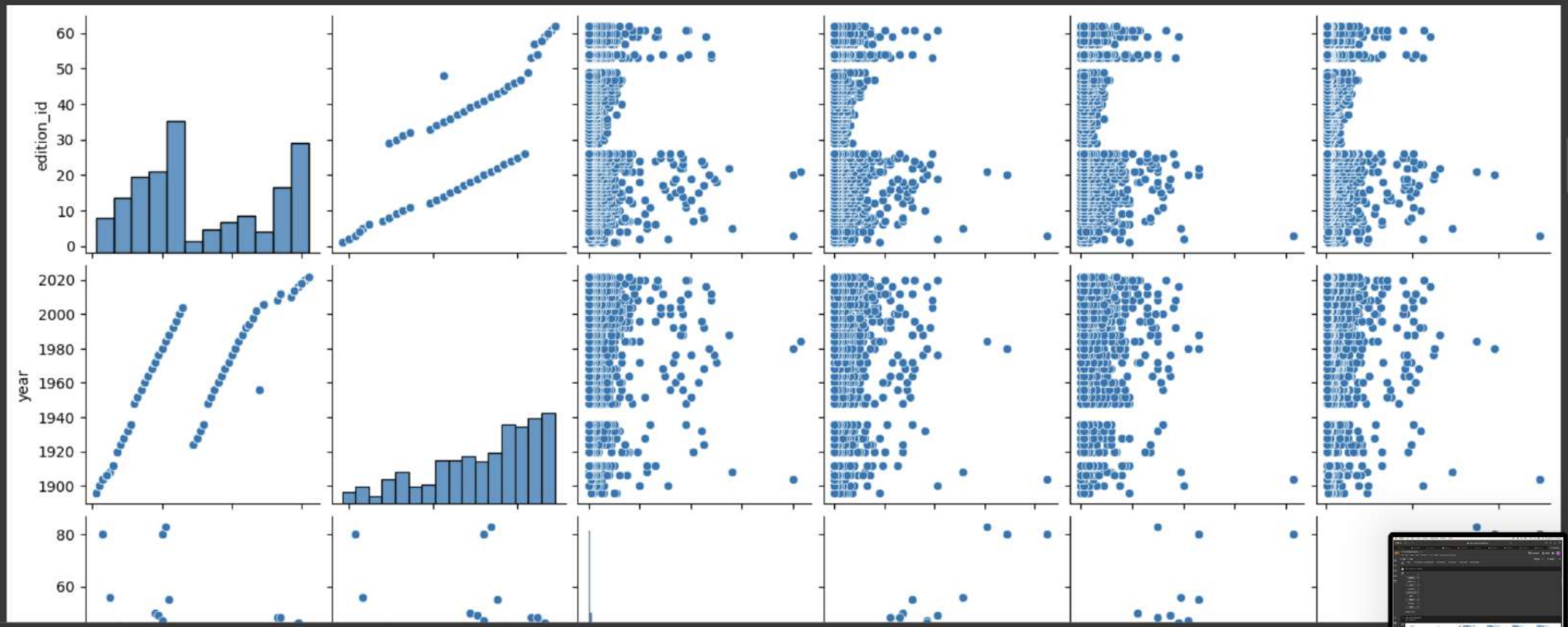
Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect Gemini

```
sns.pairplot(df)
plt.show()
```



Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share

+ Code + Text Connect Gemini

```
sns.distplot(df['year'])
```

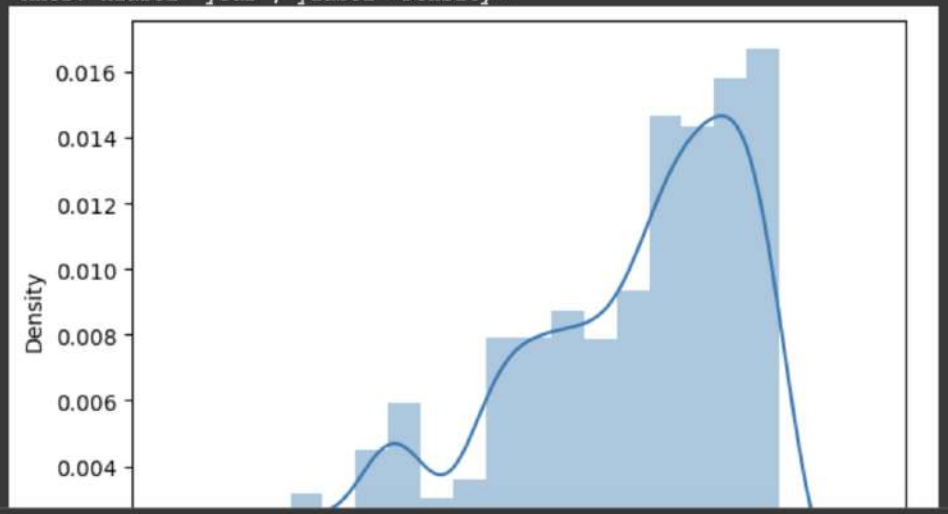
```
<ipython-input-16-ea1a7bf6e32c>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['year'])
<Axes: xlabel='year', ylabel='Density'>
```



Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect ✦ Gemini

```
sns.distplot(df['gold'])
```

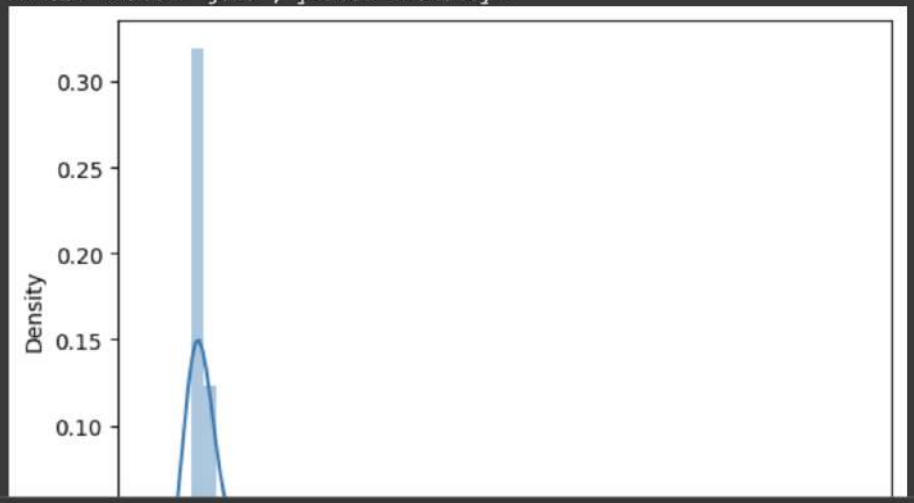
```
<ipython-input-26-6678dfa23346>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['gold'])
<Axes: xlabel='gold', ylabel='Density'>
```



Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

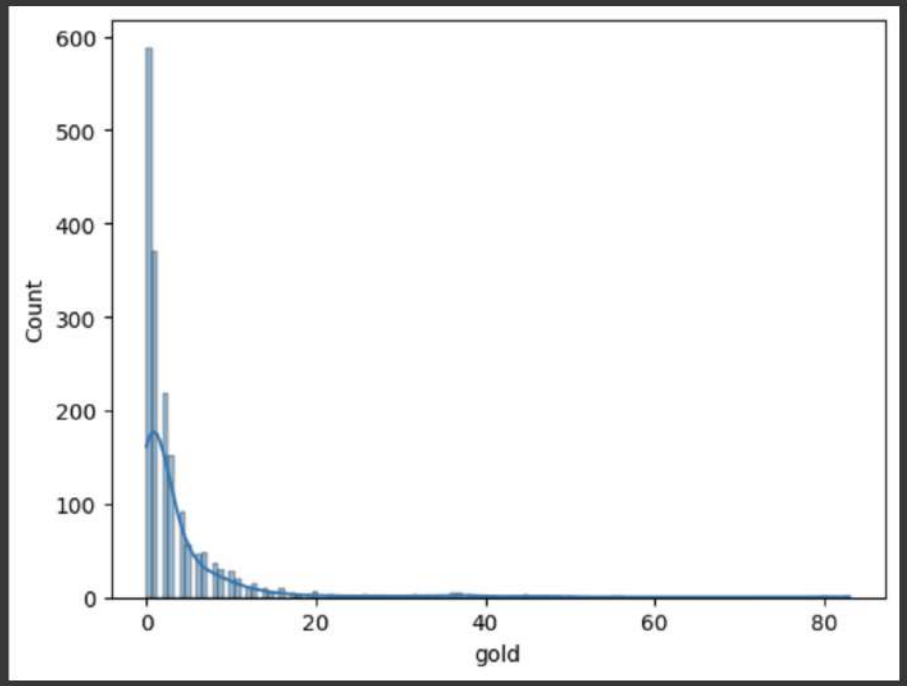
Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect ✦ Gemini

```
sns.histplot(df, x="gold", kde=True)
plt.show()
```




```
[ ] sns.histplot(df, x="silver", kde=True)
```

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

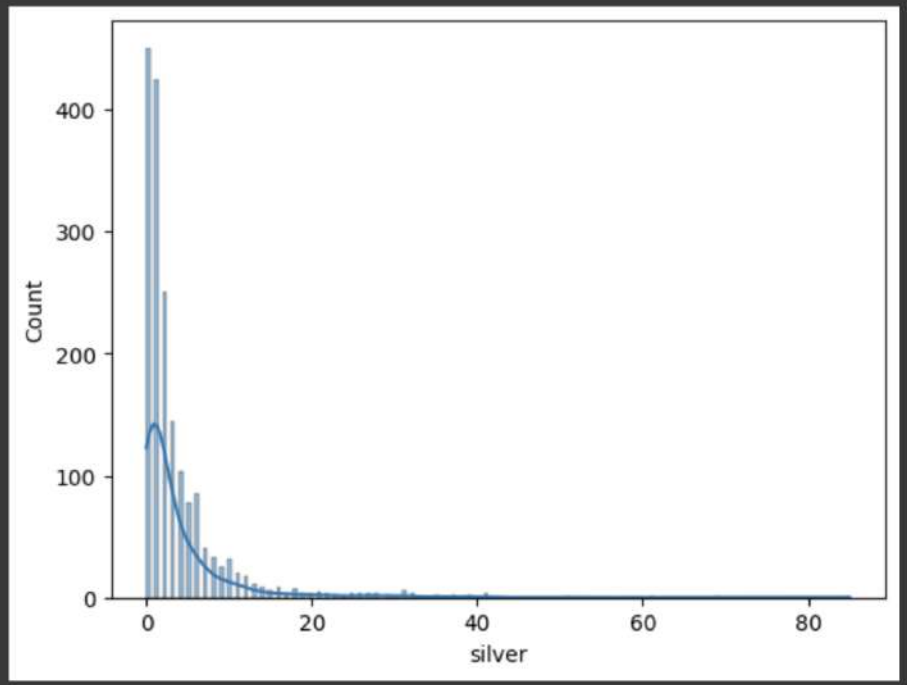
Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect Gemini

```
sns.histplot(df, x="silver", kde=True)
plt.show()
```




```
[ ] sns.regplot(x='year',y='gold',data=df,line_kws={"color": "red"})
    plt.show()
```

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

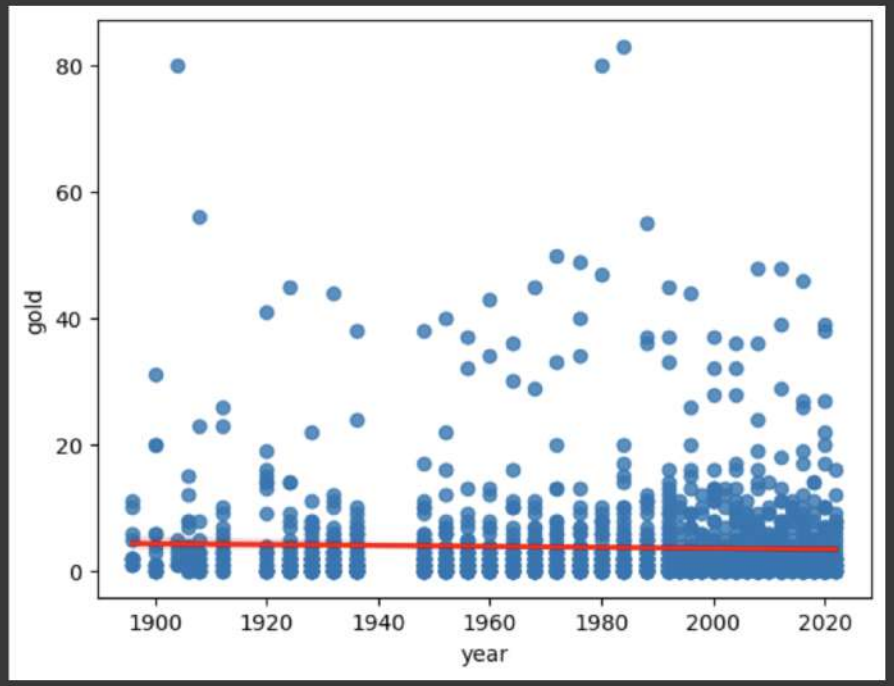
Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect Gemini

```
sns.regplot(x='year',y='gold',data=df,line_kws={"color": "red"})
plt.show()
```




```
[ ] sns.lineplot(x='year',y='gold',data=df)
```

Safari File Edit View History Bookmarks Window Help Sat Sep 21 11:03 AM

colab.research.google.com

Home | RA221100... | Jupyter S... | Manage y... | (2) YouTube | Home | RA221100... | RA221100... | Untitled5.i... | Welcome... | Untitled3.i...

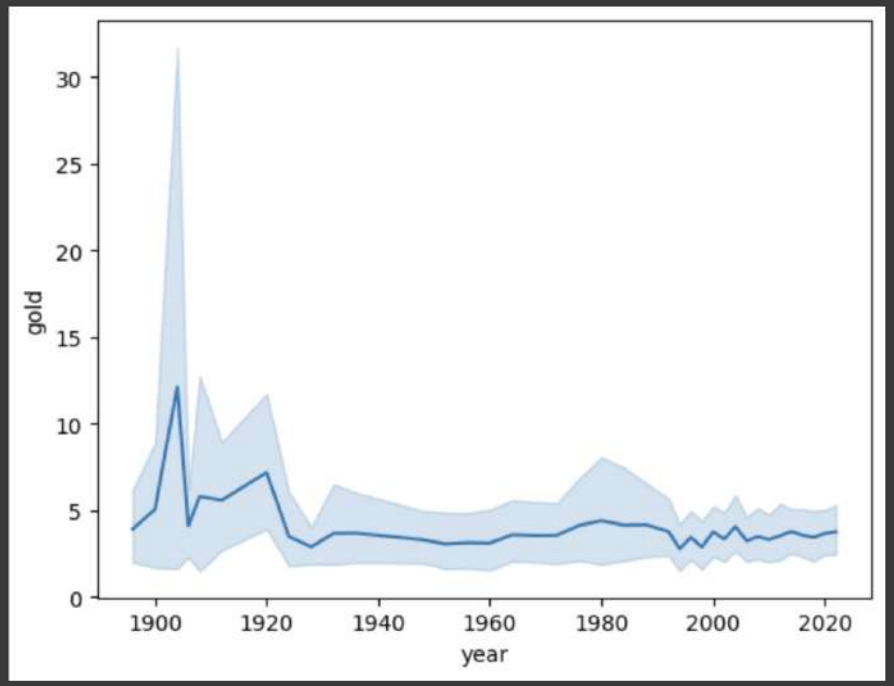
Untitled3.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:59 AM

Comment Share K

+ Code + Text Connect ✦ Gemini

```
sns.lineplot(x='year',y='gold',data=df)
plt.show()
```



Double-click (or enter) to edit