Task 1:

```python
# let's create a function that prints out the EDA actions I did previously
def eda(df):
    print("Shape:", df.shape)
    print(" ")
    print("Data Types:")
    for col in df.columns:
        print(col, df[col].dtype)
    print(" ")
    print("Missing Values:")
    print(df.isnull().sum())
    print(" ")
    print("Duplicates:")
    print(df.duplicated().sum())

# let call on that function
eda(games)
```

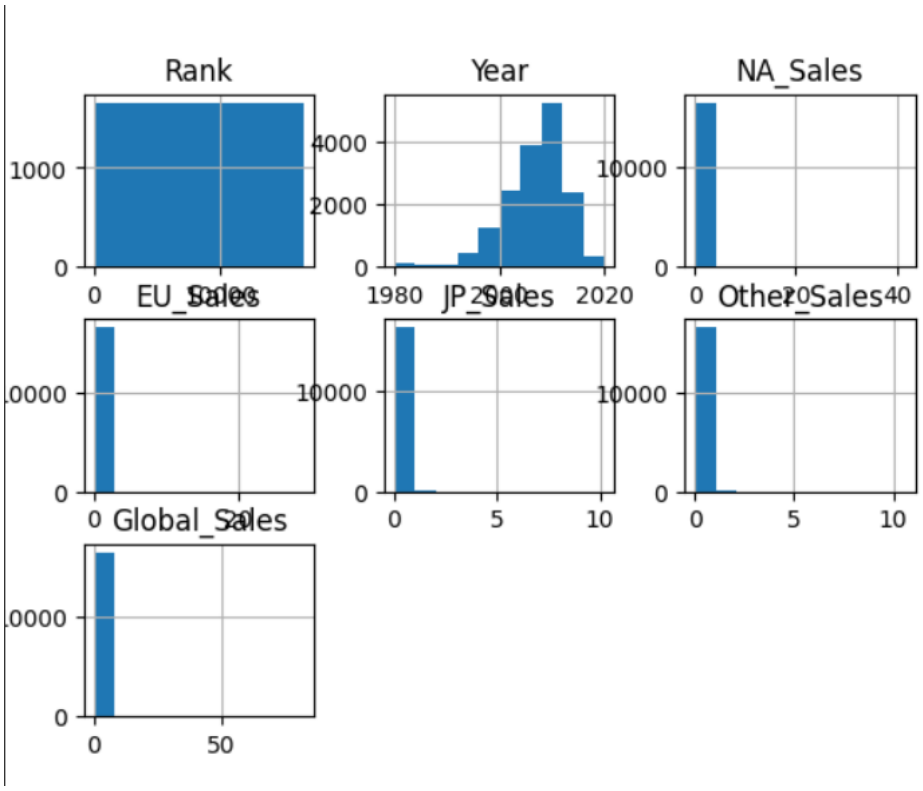
```
Shape: (16598, 11)

Data Types:
Rank int64
Name object
Platform object
Year float64
Genre object
Publisher object
NA_Sales float64
EU_Sales float64
JP_Sales float64
Other_Sales float64
Global_Sales float64

Missing Values:
Rank            0
Name            0
Platform        0
Year          271
Genre           0
Publisher      58
NA_Sales        0
EU_Sales        0
JP_Sales        0
Other_Sales     0
Global_Sales    0
dtype: int64

Duplicates:
0
```

The code defines a function called `eda()` that takes a pandas DataFrame as its input. This function performs some basic exploratory data analysis on the input DataFrame and prints out the results. The output shows the shape of the DataFrame, the data types of each column, the number of missing values in each column, and the number of duplicated rows in the DataFrame. By printing out this information, the function provides a quick and simple way to assess the quality of the data in the DataFrame. The output (on the right) is useful in identifying potential issues with the data, such as missing values, duplicated rows, or inconsistent data types. By using the `eda()` function, data analysts can easily gain insight into their data and identify areas that may require further investigation or cleaning.

After doing

```
# Histograms
games.hist()
```

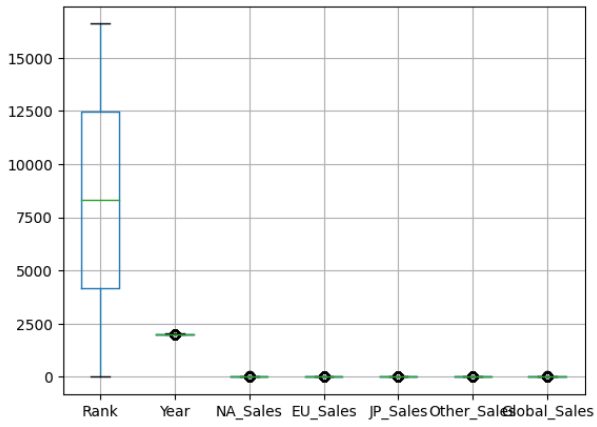
I got



After doing

```
# boxplots
games.boxplot()
```

I got


15000
12500
10000
7500
5000
2500
0
Rank
Year
NA_Sales
EU_Sales
JP_Sales
Other_Sales
Global_Sales