**Investigate the Films that had Zero Revenue**

Now that we've done some legwork on cleaning our data, we can investigate our data set more thoroughly.

Challenge 1

1. What is the average production budget of the films in the data set?
2. What is the average worldwide gross revenue of films?
3. What were the minimums for worldwide and domestic revenue?
4. Are the bottom 25% of films actually profitable or do they lose money?
5. What are the highest production budget and highest worldwide gross revenue of any film?
6. How much revenue did the lowest and highest budget films make?

Challenge 2

How many films grossed $0 domestically (i.e., in the United States)? What were the highest budget films that grossed nothing?

Challenge 3

How many films grossed $0 worldwide? What are the highest budget films that had no revenue internationally (i.e., the biggest flops)?

.

.

..

...

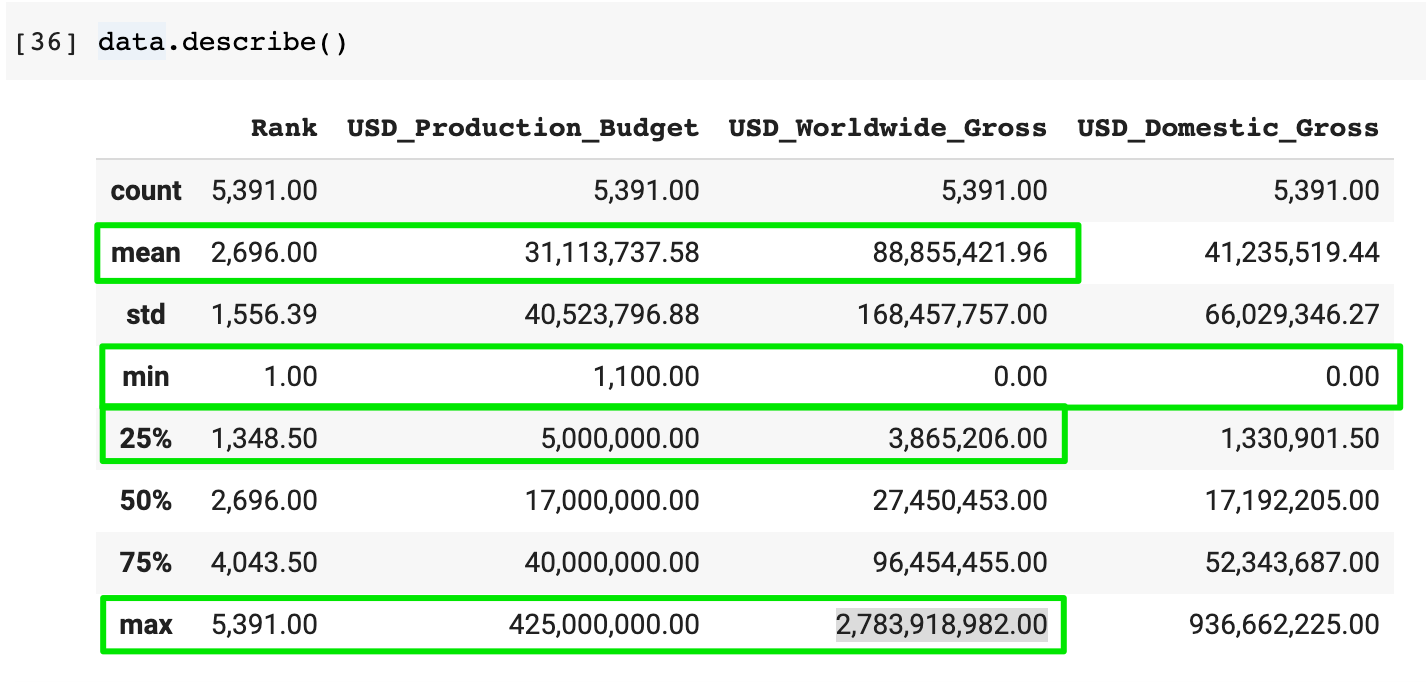
..

.

.

**Solution for Challenge 1**

We can answer many of the questions with a single command: .describe().



The average film costs about $31m to make and makes around 3x that (or ~$89m) in worldwide revenue. So that's encouraging.

But quite a lot of films lose money too. In fact, all the films in the bottom quartile lose money, since the average cost is $5 million and they only bring in $3.8m in worldwide revenue!

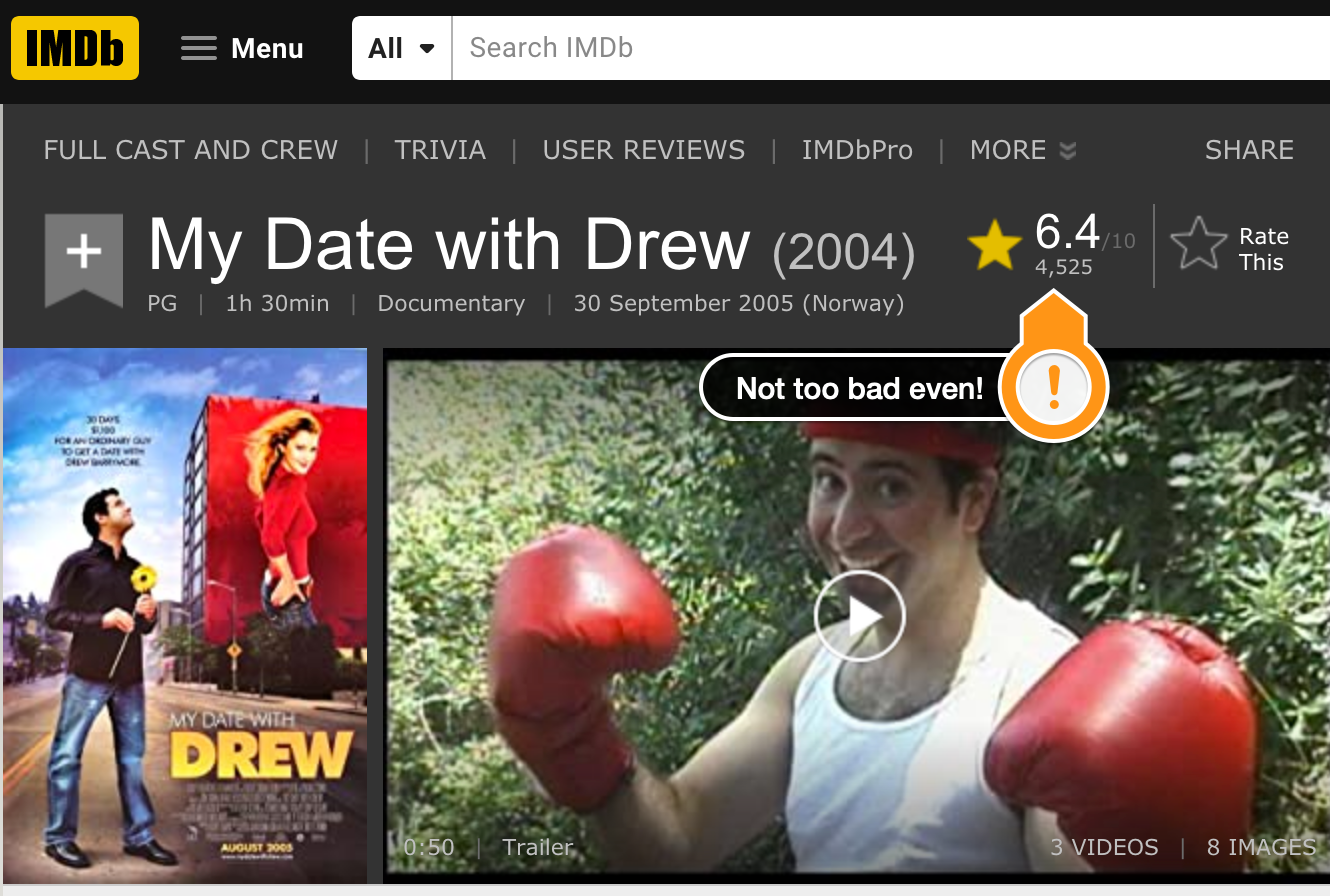
The minimum domestic and worldwide revenue is $0. That makes sense. If a film never gets screened or is cancelled, then this is the number we would expect to see here.

On the other hand, the highest production budget was $425,000,000 and the highest worldwide revenue was $2,783,918,982. $2.7 Billion revenue! Holy smokes.

So which film was the lowest budget film in the dataset?



I've ... never heard of this film. But it looks like a real money maker. It grossed $181,041 with a measly $1,100 budget. 😮 Wow. Talk about return on investment!



And the highest budget film in the dataset is:



Sigh, I remember watching this film in the cinema with 3D glasses 🤓 and not wanting the film to ever end! I would have been quite content living with those blue people. 😊

**Solution to Challenge 2: No domestic revenue**

We see that there are 512 films in the dataset that had no revenue in the United States. However, the highest budget films with no revenue have a release date AFTER the date on which the dataset was compiled (May 1st, 2018).

1. zero\_domestic = data[data.USD\_Domestic\_Gross == 0]
2. print(f'Number of films that grossed $0 domestically {len(zero\_domestic)}')
3. zero\_domestic.sort\_values('USD\_Production\_Budget', ascending=False)



**Solution to Challenge 3: No worldwide revenue**

When we check worldwide revenue instead, we see that there are 357 films that made no money internationally. Once again, some of the films have not been released yet at the time the data was compiled. However, 512 versus 357. Why is there a difference?

The reason some international films were never screened in the United States.  In fact, we can see an example of this in our previous screenshot. "Don Gato, el inicio de la pandilla" made about $4.5 million dollars in the box office, but nothing in the United States. Perhaps they should have screened it there too, considering it cost $80 million to make!

1. zero\_worldwide = data[data.USD\_Worldwide\_Gross == 0]
2. print(f'Number of films that grossed $0 worldwide {len(zero\_worldwide)}')
3. zero\_worldwide.sort\_values('USD\_Production\_Budget', ascending=False)

