Submission requirement:

1. PDF file answering 3 quesion plus your own

2. A .twbx file and a link to Tableau based visualization:

   - 4 dashboards
   - each dashboard answers a question with 2-3 plots
   - 1 narrative to further dig into one question: Qx Story
   - use figure types: line, scatter, bar, small multiple,
   - group by color, shape, size, good color palette
   - clear labels on all axes and legends
   - use filters, highlight, text for interactivity

Business scenario

As a business analyst consultant, you provide guidence for a new movie production company to understand movie trends for better decision. The client comes out 3 questions and you create your own.

step 1: Data Clean up and Attribute Selection

list attributes I plan to dive further:

- budget Adj, revenue Adj
- tagline, keywords
- Genres
- production company
- vote average
- release year

: Awesome: This is good to list down the attributes used during your analysis. With this, you can easily explain your answers.

Movie industry changes over time. Among the 10866 movies since 1960, **7168** was produced since 2000, or 66%. I will be focused on the data in 2000-2015.

step 2: Tableau visualizations

The link is here:
https://public.tableau.com/profile/yuchao.jiang#!/vizhome/vismoviedata/Story1

step 3: Questions

Q1: How have movie genres changed over time?

Movie usually comes with a hybrid type. For example, counts for the top 10 genres of raw format are:

```
 1  Drama                     521
 2  Comedy                    493
 3  Documentary               290
 4  Horror|Thriller           196
 5  Drama|Romance             181
 6  Comedy|Drama              181
 7  Comedy|Romance            177
 8  Horror                    170
 9  Comedy|Drama|Romance      139
10  Drama|Thriller            102
```

To have a more generalized answer, I split the genres and study the primary genre and secondary genre separately. As we see in the dashboard, both ways of ranking shows the same top 5 genres:

- drama
- comedy
- thriller
- horror
- action

They are all growing in the last 15 years. To my surprise, thriller and horror movies are growing much faster in recent 5 years. It implies a trend that the audience are more interested to see shocking stimulation. However, the number of drama produced annually is still at least 2 times of other types. A safte strategy is Drama plus other types.

Q 2: How do the attributes differ between Universal Pictures and Paramount Pictures?

Ans: A movie is usually co-produced by several companies. Now we only consider the first authors.

The movie counts for top 5 production companies are:

```
1  Universal Pictures                          200
2  Paramount Pictures                          162
3  Columbia Pictures                           149
4  Walt Disney Pictures                        140
5  New Line Cinema                             120
6  Twentieth Century Fox Film Corporation       96
```

Budget and revenue: Only 288 records have the valid data. The median value of budget is 60M for Paramount, and 47 M for Universal. The median value of revenue is 123M for Paramount, and 106M for Universal. So Universal has lower average cost and higher gross margin rate.

Annual release is about 10 movies for each company. It seems there is slightly decrease in revenue for the last two years but they still earn a lot of money. What's more, the production cycle for a good movie may be more than 1 year, so some fluctuations is normal.

Q3: How have movies based on novels performed relative to movies not based on novels?

Ans: First, I use `CONTAINS([Keywords], "novel")` in calculated field to identify whether a movie is based on novel or not. Then I filter out all the null value, 0 zero in budget/revenue to get my data ready.

Since 2000, there are 130 movies based on novels and 2260 not. So novel-based movies is only 6% of the sample. In terms of median budget and revenue, novel-based movies are only slightly higher than the other.

Q4: What are the best-selling movies

Ans: The top 5 best-selling movies accross all genres are

1. Avatar
2. Star Wars
3. Titanic
4. The Exorcist
5. Jaws

The avenue for each film ranges from 1.9 to 2.8 B US dollars.

: Awesome: You included a question as required. This is good.

: Suggestion: In the instructions, it was stated you need to say how you came up with this question and I don't see it in your answers. Please include because it is part of this question too.

Story

Story line is based on Q4 and loosely related to Q1.

As a new movie production company, we are not only interested in the overall performance, but also pay special attention dive into the more specific genres. If we decide to make certain type of movie, who should we learn from? How to position our movie to maximize audience's viewing experience?

For example, the best animation in market is Frozen, best drama is Titanic, best thriller is Da vinci Code.

We also search into the year by year market trend. The top gross movie is Hobbit in 2014 and Minions in 2015. These movies share some fantacy elements.