題目:分析1990年以來最popular movie/TV series genres timelapse

組員:107703035 朱宇方

107703008 蔡桂毓 107703019 林干婷

動機和預期任務:

電影或電視劇是每一個人都會接觸到的,因此影視產業的商機必然是很大的。然而,在不同的時空背景下,大家會喜歡的劇情種類可能會有些許改變,我們希望透過IMDB的大量數據,來分析歸納每一年觀眾喜歡的類型並加以視覺化呈現。

我們希望User透過視覺化可以更清楚的知道每一年熱門Top 5 種類,而且透過動態的呈現還可以看到變化的過程。另外在圖表呈現方面,會同時放上電視劇和電影兩者個別的Top 5熱門種類,讓User可以一目瞭然的比較兩者之間差別。

資料集:

IMDb movie datasets[1]

我們從資料集中選用需使用到的類別如:

- tconst (string) alphanumeric unique identifier of the title
- titleType (string) the type/format of the title (e.g. movie, short, tyseries, typisode, video, etc)
- startYear (YYYY) represents the release year of a title. In the case of TV Series, it is the series start year
- genres (string array) includes up to three genres associated with the title
- averageRating weighted average of all the individual user ratings
- numVotes number of votes the title has received

以genres分類,根據年份將討論度以及評分做出區別,並可透過types分別以 dynamic bar chart呈現。

相關研究:

對於先前研究關於資料Ranking在多維度(如時間)的呈現方法,除了用時間 Dynamic 看到變化以外,其實也有許多其他靜態方法達到相同的目的:如 Microsoft發表的RankExplorer[2]。一個工具延伸Theme River 來表示資料隨時間的變化。

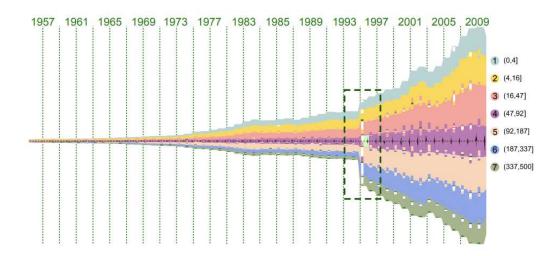
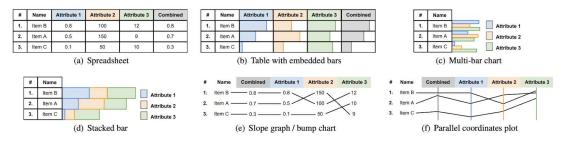


Fig. 14. US Fortune 500 data from 1955 to 2010. There is a great change in 1995 (in the dashed rectangle).

但是Theme River 的弱點是不會更新Ranking的排名而是視覺化不一樣性質的比例 變化。

另外一種視覺化分析可以看看Samuel Gratzld開發的LineUp工具[3]。LineUp是個彙整多種Ranking Visualization的工具。



LineUp的強項是在於分析多個Attributes影響一個Final Ranking 時可以使用的模式。例如調整一個Attribute比重等等。LineUp也可以觀察到Multi-Timeframe變化,利用如下Slope跟顏色變換來視覺畫呈現Ranking變化的劇烈程度。

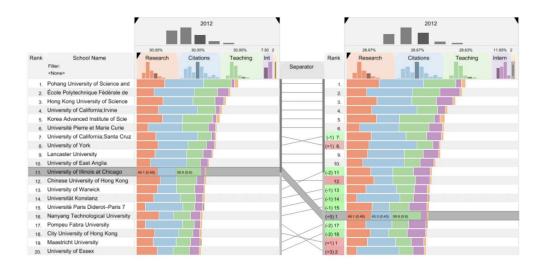


Fig. 6. Comparison between rankings. The concept of rank separators between score attribute columns makes it possible to use different orders on both sides and to relate them to each other by following the lines connecting them. Changes in either one of the rankings are immediately reflected in the visualization. This is visually supported with animations and also indicated by color changes of the rank label: green = item moved up, red = item moved down. The more intense the color, the more ranks were gained or lost.

還有另一種視覺化分析是由·浙江大學CAD&CG國家重點實驗室和阿里巴巴-浙江大學前沿技術聯合研究院所開發的SRVis[4]·他是一個支持大規模整合排序分析的視覺化方法。

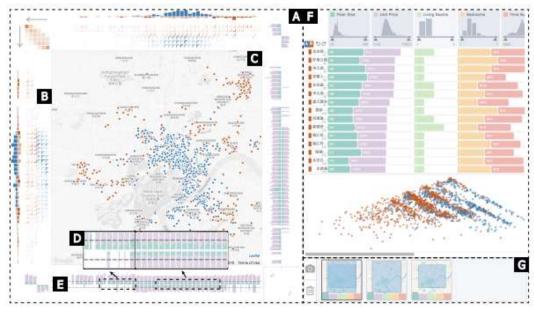


Fig. 1. The interface of SRVis. (A-E) The ranking view presents the rankings of alternatives and the cause of rankings with a matrix-based context-integrated visualization. Flexible spatial filtering features provided by B, C, and E enable users to conveniently explore and identify spatial patterns in the ranking datasets. (F) The inspector view adopts a table-based ranking technique to show all alternatives in the filtered region. A projection view is also utilized to assist users in finding similar alternatives based on their criteria. (G) The snapshot view allows users to save snapshots of rankings and criterion weights, such that users can compare these snapshots to find insights from the comparative analysis.

這個工具可以針對一些需要更為大量及詳細的數據分析,例如連鎖店需要於此地點的何處進行開設,可透過大量資料視覺化產生。

但是我們認為LineUp及SRVis的方式過於複雜化資料,雖然可以處理大量資料,但因為我們的資料集本身是較為單純的,所以用LineUp這種Multi-Attribute Analysis可能也是不會達到我們的目的,

另外也是沒有強調到單一資料比例的變化,最終仍然是以Dynamic Timeframe Data 變化來處理,較能達到呈現比例變化,以及Ranking變化最為清楚。

預期呈現方法:



可以按play:就會動態的從1990開始,一直跑到2021年的圖表也可以按年份:就會只出現那一年的統計結果

參考資料:

[1]: IMDb datasets: https://datasets.imdbws.com/

[2]: RankExplorer: Visualization of Ranking Changes in Large Time Series Data, 12/2016

[3]: LineUp: Visual Analysis of Multi-Attribute Rankings, Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister and Marc Streit, 2013 InfoViz

[4]: SRVis: Towards Better Spatial Integration in Ranking Visualization, Di Weng, Ran Chen, Zikun Deng, Feiran Wu, Jingmin Chen, and Yingcai Wu, 2018