

---

---

# Movie Critic

Final Presentation

---

---

Team Member: Melinda Haughey, Jessica Liu, Olivia Li

# Background

Reviews on movies are created by critics and have an impact on the success and awards of a given movie. These reviews include chunks of text that describe the movie and the review. We are interested in providing the user information about how a critics rate movies and how they impact the overall ratings so they can see if there is variation in rating and sentiment over the years. Also, we would like to extract the sentiment from their reviews and present everything in a visualized way.

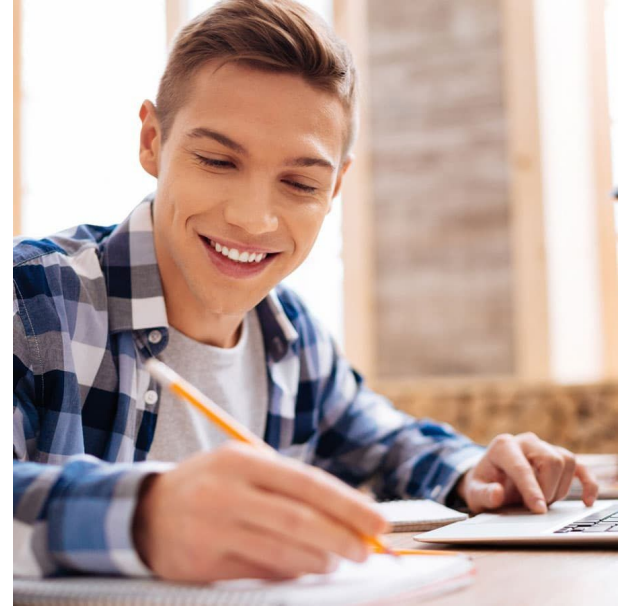
To achieve this goal we will need to use packages for data manipulation, visualization (for exploratory analysis) and learning.

# The story

John is a junior student majored in Cinema and Media Studies in University of Washington. Recently, he has an assignment which asks about certain critics on Rotten Tomatoes. Since this is a open topic, he wants to start from the overall rating trend to get a general idea of how their attitudes changes over given time period. He is also interested in who writes the most reviews.

After having a general understanding of it, he is interested in sentiment in those reviews and know the difference between critics. He wants to get a list of example words for each emotion to see what is behind that. Also, he would expect a clear visualization for the percentage of emotion in the reviews.

He doesn't know much about computer science and he just expects a pretty simple website that when he clicks and selects, all results will appear in just one second.



# Data Used



- (1) List of movies that came out from 1903 to 2011
- (2) List of reviews from Rotten Tomatoes from 1929 to 2013 including their score, the reviewer, and the reviewer contents

Limitation:

We only have the overall rating for each movie, but not from each critic. It's hard for us to analyze each critic's rating and we could just focus on their reviews instead.

Package :

Altair (Data visualization); sentiments\_nrc (Sentiment analysis)

# Data

critic	fresh	imdb	link	publication	quote	review_date	rtid	title
Derek Adam	fresh	114709	http://www.	Time Out	So ingenious	#####	9559	Toy Story
Richard Corli	fresh	114709	http://www.	TIME Magaz	The year's m	#####	9559	Toy Story
David Ansen	fresh	114709	http://www.	Newsweek	A winning ar	#####	9559	Toy Story
Leonard Klad	fresh	114709	http://www.	Variety	The film spo	6/9/08 0:00	9559	Toy Story
Jonathan Ro	fresh	114709	http://onfilm	Chicago Rea	An entertain	#####	9559	Toy Story
Michael Boo	fresh	114709	http://www.	Denver Post	As Lion King	5/3/07 0:00	9559	Toy Story
Geoff Andre	fresh	114709	http://www.	Time Out	The film will	#####	9559	Toy Story
Janet Maslin	fresh	114709	http://movie	New York Tir	Children will	#####	9559	Toy Story
Kenneth Tur	fresh	114709	http://www.	Los Angeles	Although its	#####	9559	Toy Story
Roger Ebert	fresh	114709	http://www.	Chicago Sun-	The result is	1/1/00 0:00	9559	Toy Story
Susan Stark	fresh	114709	http://data.c	Detroit New	Disney's first	1/1/00 0:00	9559	Toy Story
Peter Stack	fresh	114709	http://www.	San Francisco	The script, b	1/1/00 0:00	9559	Toy Story
James Berar	fresh	114709	http://www.	ReelViews	The one big i	1/1/00 0:00	9559	Toy Story
Rita Kemple	fresh	114709	http://www.	Washington	It's a nice ch	1/1/00 0:00	9559	Toy Story
	fresh	114709	http://www.	Entertainme	I can hardly i	#####	9559	Toy Story
Roger Moore	fresh	114709	http://www.	Orlando Sem	The great vo	#####	9559	Toy Story
Roger Ebert	rotten	113497	http://www.	Chicago Sun-	A gloomy spi	1/1/00 0:00	12436	Jumanji
Richard Schi	fresh	107050	http://www.	TIME Magaz	Walter Matt	#####	10498	Grumpy Old M
Derek Adam	rotten	107050	http://www.	Time Out	Mediocre, re	#####	10498	Grumpy Old M
Caryn James	fresh	107050	http://movie	New York Tir	Just don't ex	#####	10498	Grumpy Old M
James Berar	fresh	107050	http://www.	ReelViews	While it won	1/1/00 0:00	10498	Grumpy Old M
Roger Ebert	rotten	107050	http://www.	Chicago Sun-	The movie is	1/1/00 0:00	10498	Grumpy Old M
Desson Thon	fresh	107050	http://www.	Washington	If you poke t	1/1/00 0:00	10498	Grumpy Old M
Liam Lacey	rotten	114885	http://www.	Globe and M	Never escap	#####	16697	Waiting to Ex
Kenneth Tur	fresh	114885	http://www.	Los Angeles	A pleasant if	#####	16697	Waiting to Ex
Edward Gut	rotten	114885	http://www.	San Francisc	You want the	1/1/00 0:00	16697	Waiting to Ex
Desson Thon	fresh	114885	http://www.	Washington	This sister-ce	1/1/00 0:00	16697	Waiting to Ex
Joan Walsh	rotten	114885	http://www.	Salon.com	The male-ba	1/1/00 0:00	16697	Waiting to Ex
Joe Baltake	rotten	114885	http://www.	Sacramento	With one po	1/1/00 0:00	16697	Waiting to Ex
Susan Stark	rotten	114885	http://data.c	Detroit New	For all the pl	1/1/00 0:00	16697	Waiting to Ex
James Berar	rotten	114885	http://www.	ReelViews	With the exc	1/1/00 0:00	16697	Waiting to Ex
Roger Ebert	fresh	114885	http://www.	Chicago Sun-	An escapist f	1/1/00 0:00	16697	Waiting to Ex

10682 lines (10681 sloc) 510 KB

```

1 1::Toy Story (1995)::Adventure|Animation|Children|Comedy|Fantasy
2 2::Jumanji (1995)::Adventure|Children|Fantasy
3 3::Grumpier Old Men (1995)::Comedy|Romance
4 4::Waiting to Exhale (1995)::Comedy|Drama|Romance
5 5::Father of the Bride Part II (1995)::Comedy
6 6::Heat (1995)::Action|Crime|Thriller
7 7::Sabrina (1995)::Comedy|Romance
8 8::Tom and Huck (1995)::Adventure|Children
9 9::Sudden Death (1995)::Action
10 10::GoldenEye (1995)::Action|Adventure|Thriller
11 11::American President, The (1995)::Comedy|Drama|Romance
12 12::Dracula: Dead and Loving It (1995)::Comedy|Horror
13 13::Balto (1995)::Animation|Children
14 14::Nixon (1995)::Drama
15 15::Cutthroat Island (1995)::Action|Adventure|Romance
16 16::Casino (1995)::Crime|Drama
17 17::Sense and Sensibility (1995)::Comedy|Drama|Romance
18 18::Four Rooms (1995)::Comedy|Drama|Thriller
19 19::Ace Ventura: When Nature Calls (1995)::Comedy
20 20::Money Train (1995)::Action|Comedy|Crime|Drama|Thriller
21 21::Get Shorty (1995)::Action|Comedy|Drama
22 22::Copycat (1995)::Crime|Drama|Horror|Mystery|Thriller
23 23::Assassins (1995)::Action|Crime|Thriller
24 24::Powder (1995)::Drama|Sci-Fi
25 25::Leaving Las Vegas (1995)::Drama|Romance
26 26::Othello (1995)::Drama
27 27::Now and Then (1995)::Drama
28 28::Persuasion (1995)::Drama|Romance

```

# Use Cases



- USER INPUT:** User selects the year range that he/she is interested in
- SYSTEM:** Pulls movie review data and visualizes the rating distribution over time
- SYSTEM OUTPUT:** System shows visualization of rating distribution during time range to the user
- SYSTEM:** Select a specific year that the user is interested in
- 
- USER INPUT:** The Specific year
- SYSTEM OUTPUT:** The Top 5 critics having most quotes of that year and their sentiment analysis

# Use Cases

Hey!



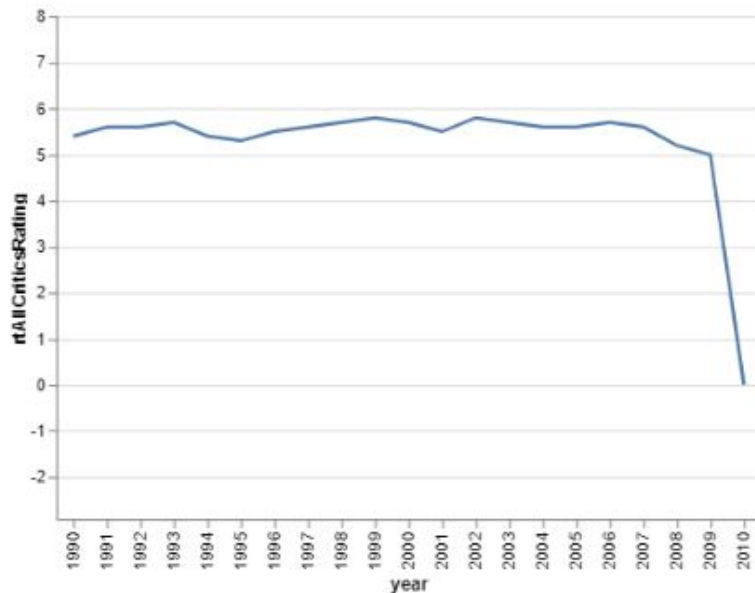
Welcome to our project. In order to see how a critic has scored movies over time, select a critic and time range below!

Type in a year range

Start Year

End Year

Submit



# Use Cases

```
Please input a year of interest to see critic activity that year:  
2009
```

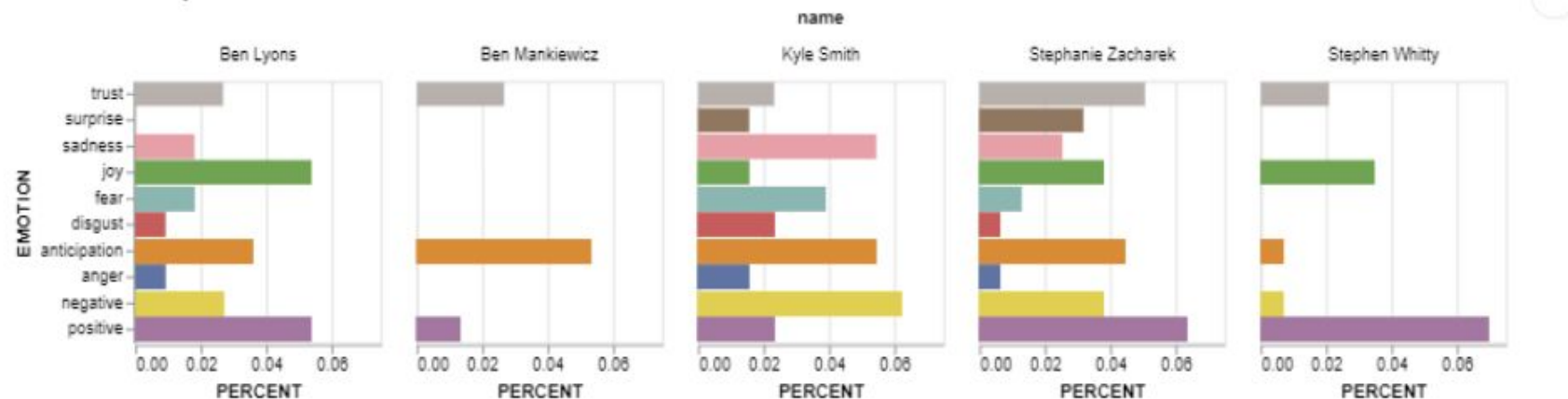


```
['Stephanie Zacharek', 'Kyle Smith', 'Ben Lyons', 'Ben Mankiewicz', 'Stephen Whitty']  
      critic                                     quote  
0      Ben Lyons  A terrific twist, but goes wasted in a movie t...  
1      Ben Mankiewicz  I just don't think there was enough drama in t...  
2      Kyle Smith  This movie knows exactly what it is: Gonzo sil...  
3      Stephanie Zacharek  Ninja Assassin lives in the moment, a visceral...  
4      Stephen Whitty  A Perfect Getaway is one of those very clever ...
```



# Use Cases

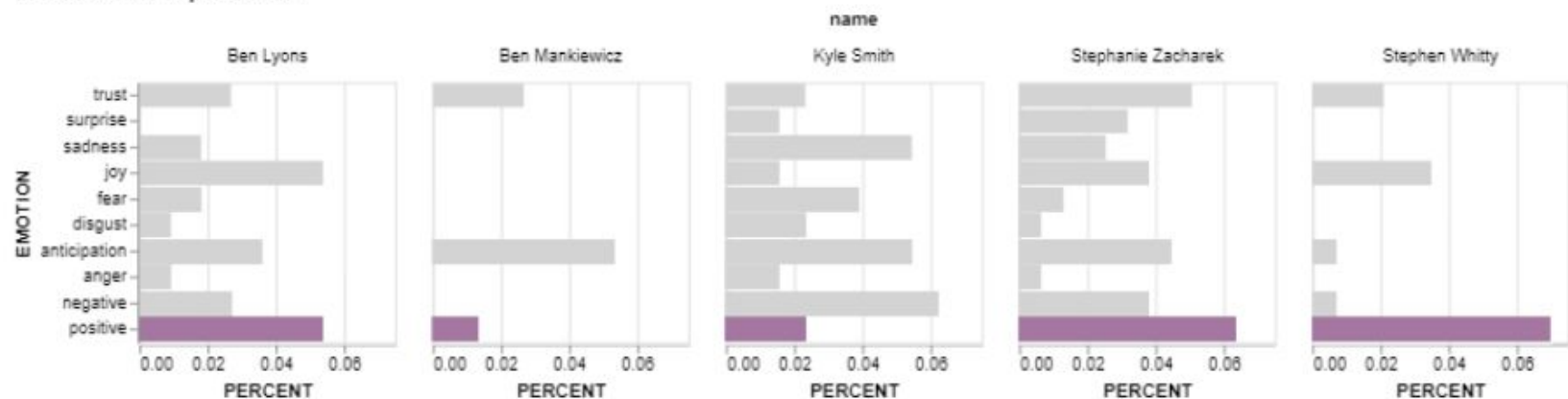
Sentiment for Top 5 Critics



Type\_of\_EMOTION positive ▼

# Use Cases

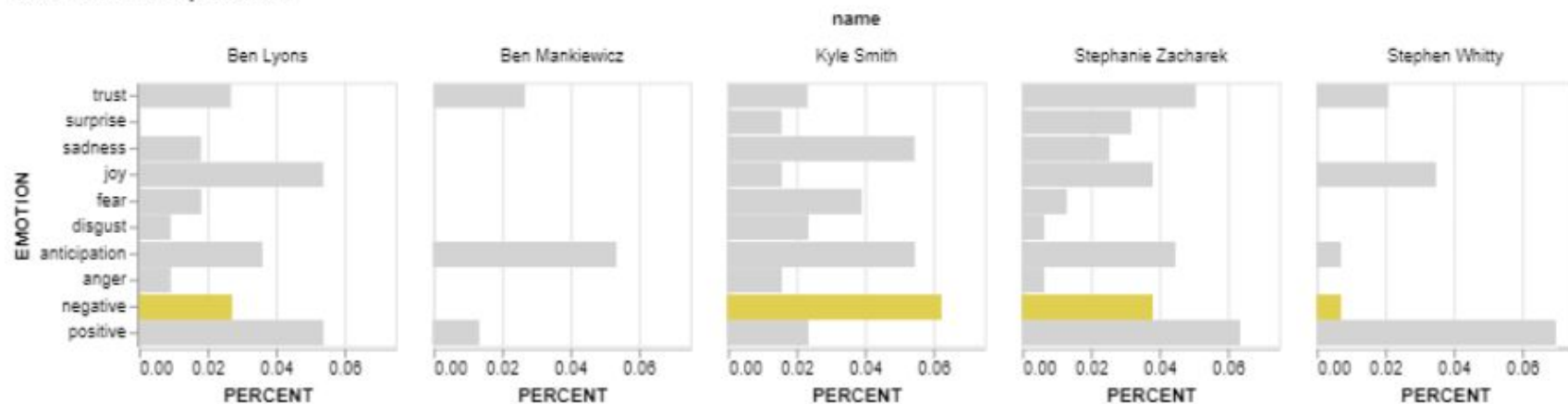
Sentiment for Top 5 Critics



Type\_of\_EMOTION

# Use Cases

Sentiment for Top 5 Critics

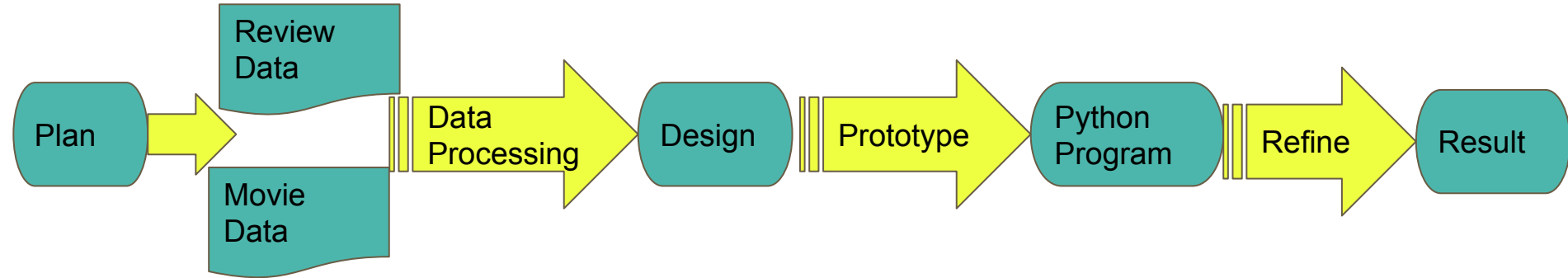


Type\_of\_EMOTION

# Use Cases



# Design



# Project Structure

```
Movie Rating Analysis/  
| - movieCriticsOverTime  
|   - __init__.py  
|   - main.py  
|   - clean_data.py  
|   - overall_rating.py  
|   - sentiment_setup.py  
|   - sentiment_analysis.py  
|   - sentiments_nrc.py  
|   - visualization.py  
|   - word_cloud.py  
|   - movieCriticsOverTime.ipynb  
| - example  
|   - notebook_showcase_html  
|   - notebook_showcase_pdf  
| - test  
|   - test_columns.py  
|   - test_type.py  
| - data  
|   - movies.dat  
|   - reviews.csv  
| - docs  
|   - .DS_Store  
|   - component_specs.md  
|   - functional_specs.md  
|   - technology_review_ppt_group5  
|   - final_presentation_ppt_group5  
|-travis.yml  
|-setup.py  
|- .gitignore  
|- README.md  
|- LICENSE
```

**DEMO**

# Lessons learned

- Need to understand data before trying some techniques.
- The goal can be changed slightly during exploration, not necessary be the original one.
- Take user experience into consideration; try to make interface more user-friendly.

# Future work

- Get each critic's rating from web scraping, which enables us to build a rating prediction based on their sentiment.
- Update our tool to more movie reviews sites.
- More interactive features which enable user to explore more by themselves.



**Thanks!**  
**Any questions?**