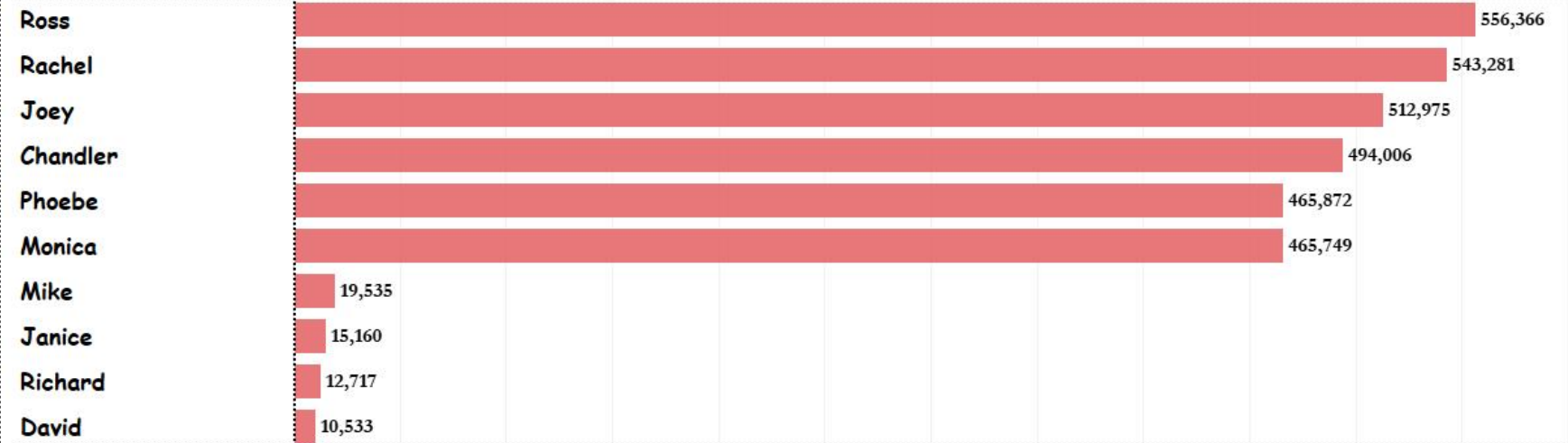


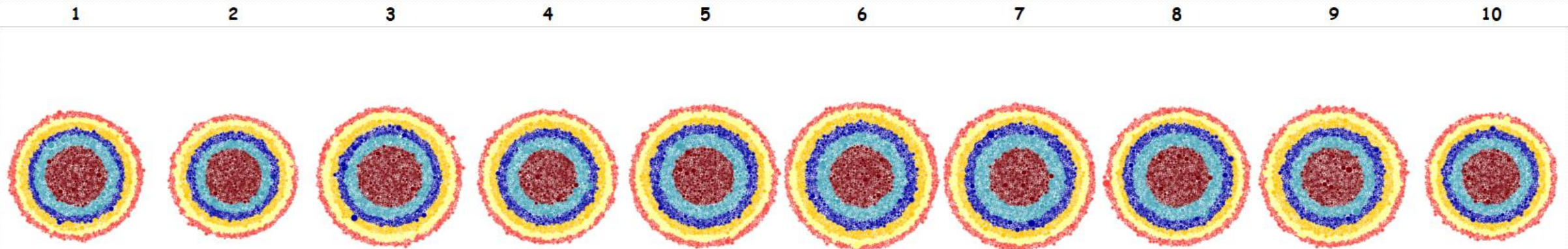
# Friends

1. Top 6 characters below account for **84% of total dialogue length** in the script
2. Overall, **Ross has the longest dialogues** in Friends, Rachel is a close second
3. There is relatively even distribution of dialogue between the six main characters in Friends, and no other character ever comes in the top 6

## Top Characters - by Dialogue Length



## Dialogue Distribution - by Top 6 Characters per Season



■ Ross

■ Rachel

■ Joey

■ Chandler

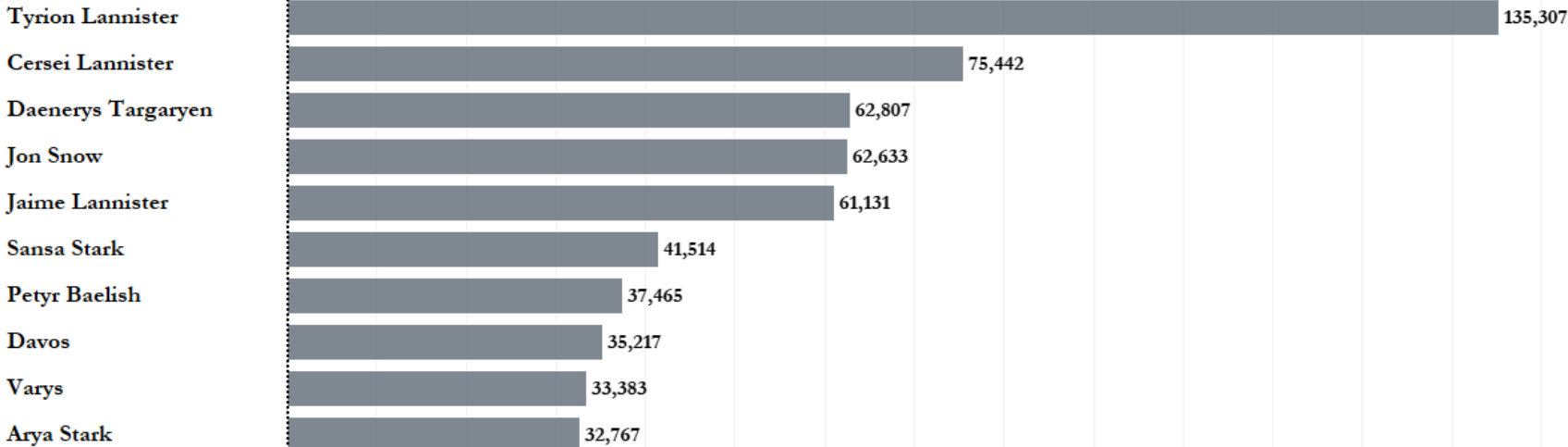
■ Phoebe

■ Monica

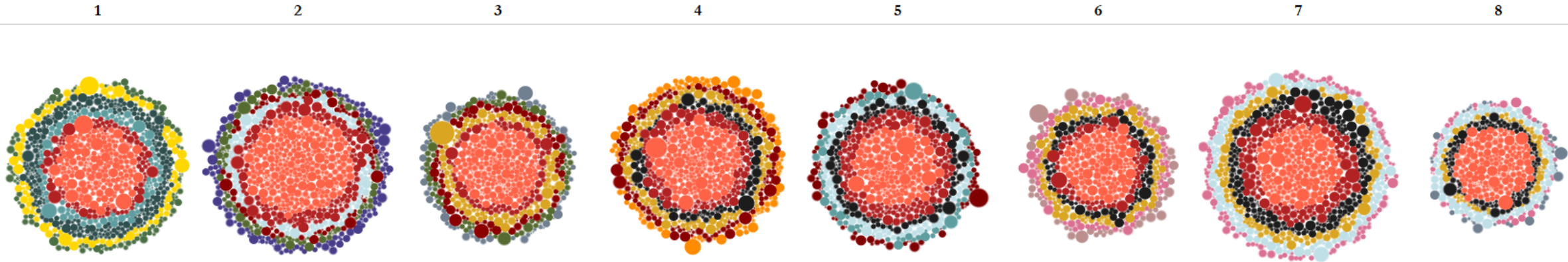
# Game of Thrones

1. Top 17 characters below account for just **34% of total dialogue length** in the script
2. Overall, **Tyrior** has the longest **dialogues** in Game of Thrones followed by Cersei
3. Rapid Changes in storyline, unequal character dominance throughout the series, explains the inconsistent distribution of dialogues below

Top Characters - by Dialogue Length



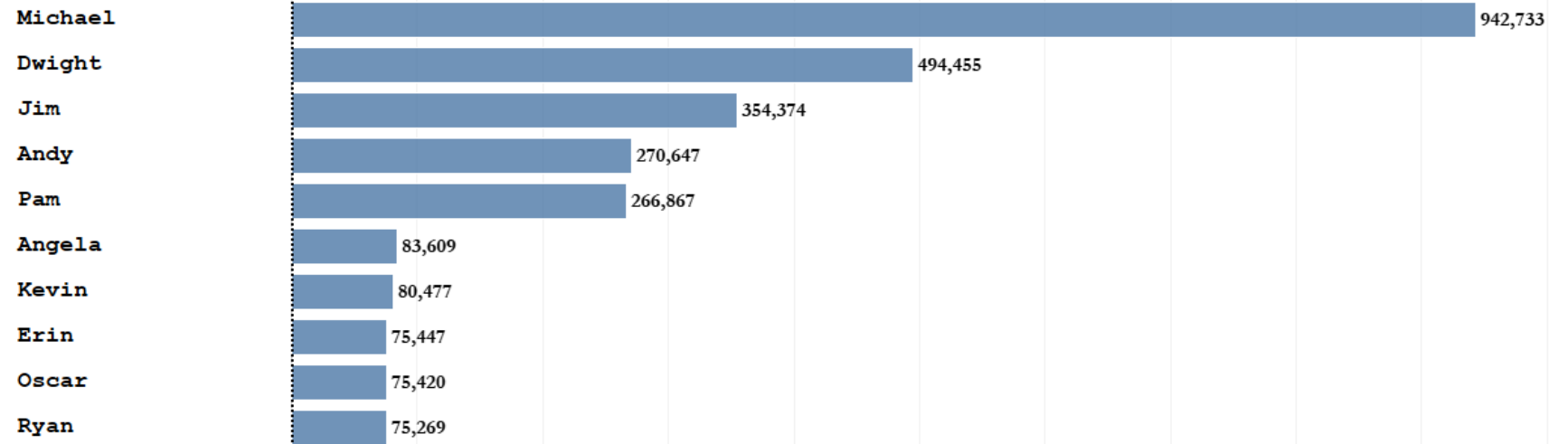
Dialogue Distribution – by Top 6 Characters per Season



# The Office

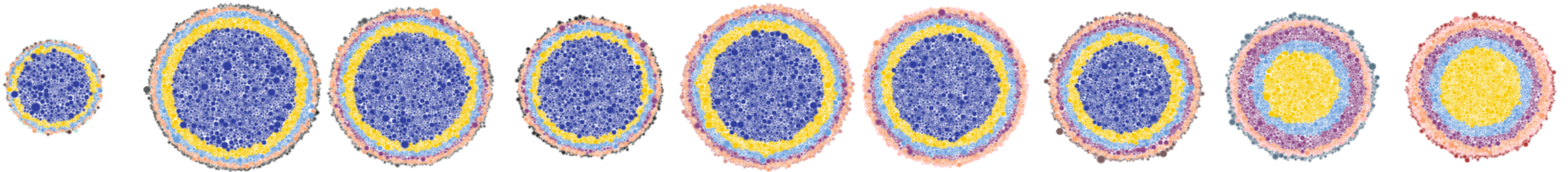
1. Top 14 characters below account for **71% of total dialogue length** in the script
2. Overall, **Michael has the longest dialogues**, holding over 26% of all the dialogues length
3. Dialogue distribution shows that “The Office” dialogue is dominated by few main characters. **Michael departed the show in season 7**

## Top Characters - by Dialogue Length



## Dialogue Distribution - by Top 6 Characters per Season

1 2 3 4 5 6 7 8 9



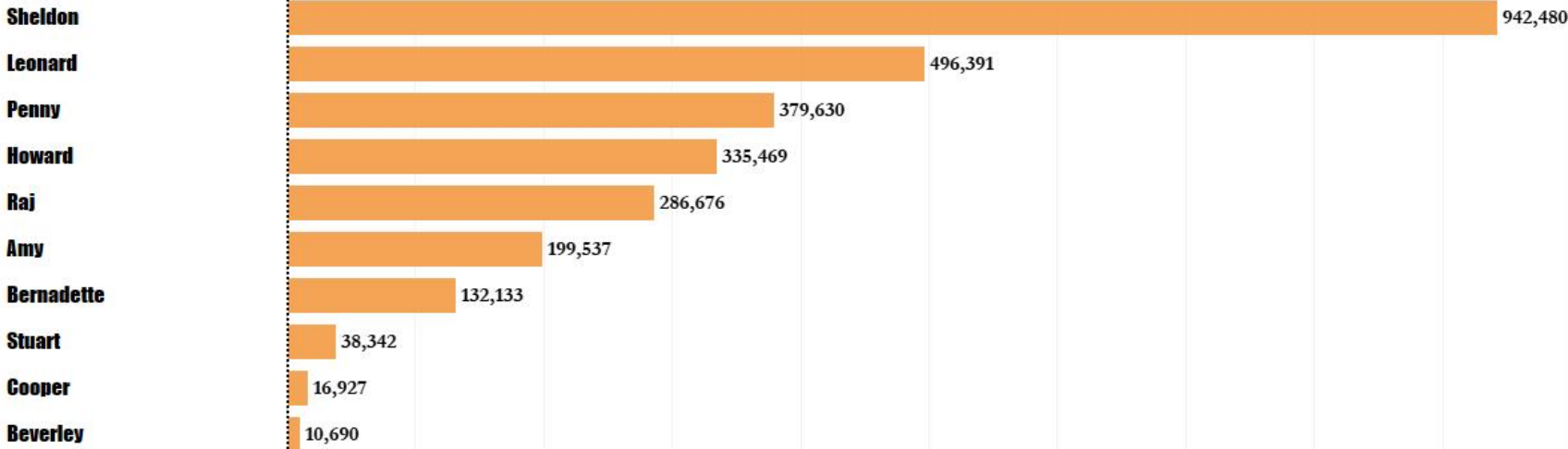
Michael Jim Pam Robert Ryan Nellie Mr. Brown  
Dwight Andy Erin Jan Deangelo Holly Oscar



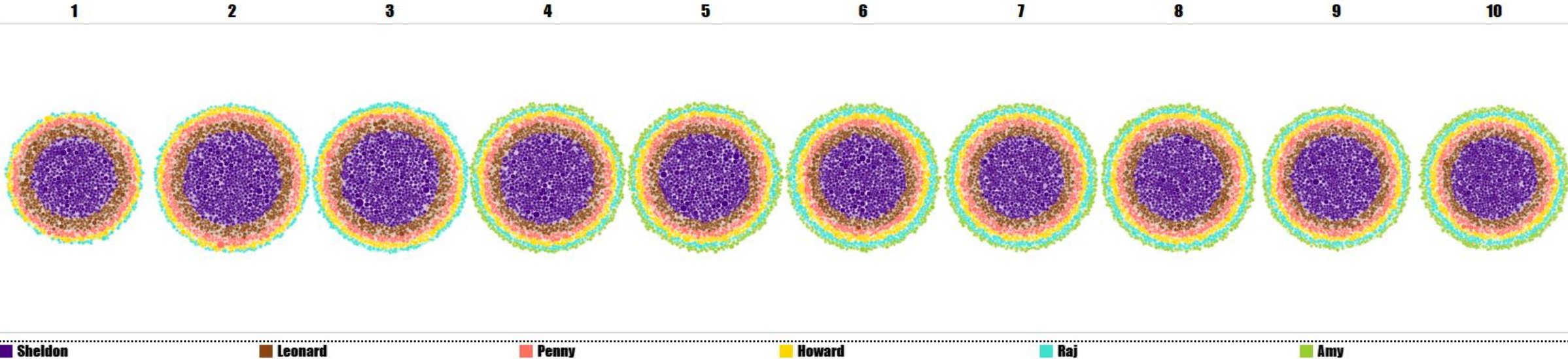
# The Big Bang Theory

1. Top 6 characters below account for **83% of total dialogue length** in the script, very similar to Friends
2. Overall, **Sheldon has the longest dialogues**, holding over 29% of all the dialogues length.
3. The Big Bang Theory is consistent in dialogue distribution, with **Amy** appearing as a prominent character from season 4 onwards

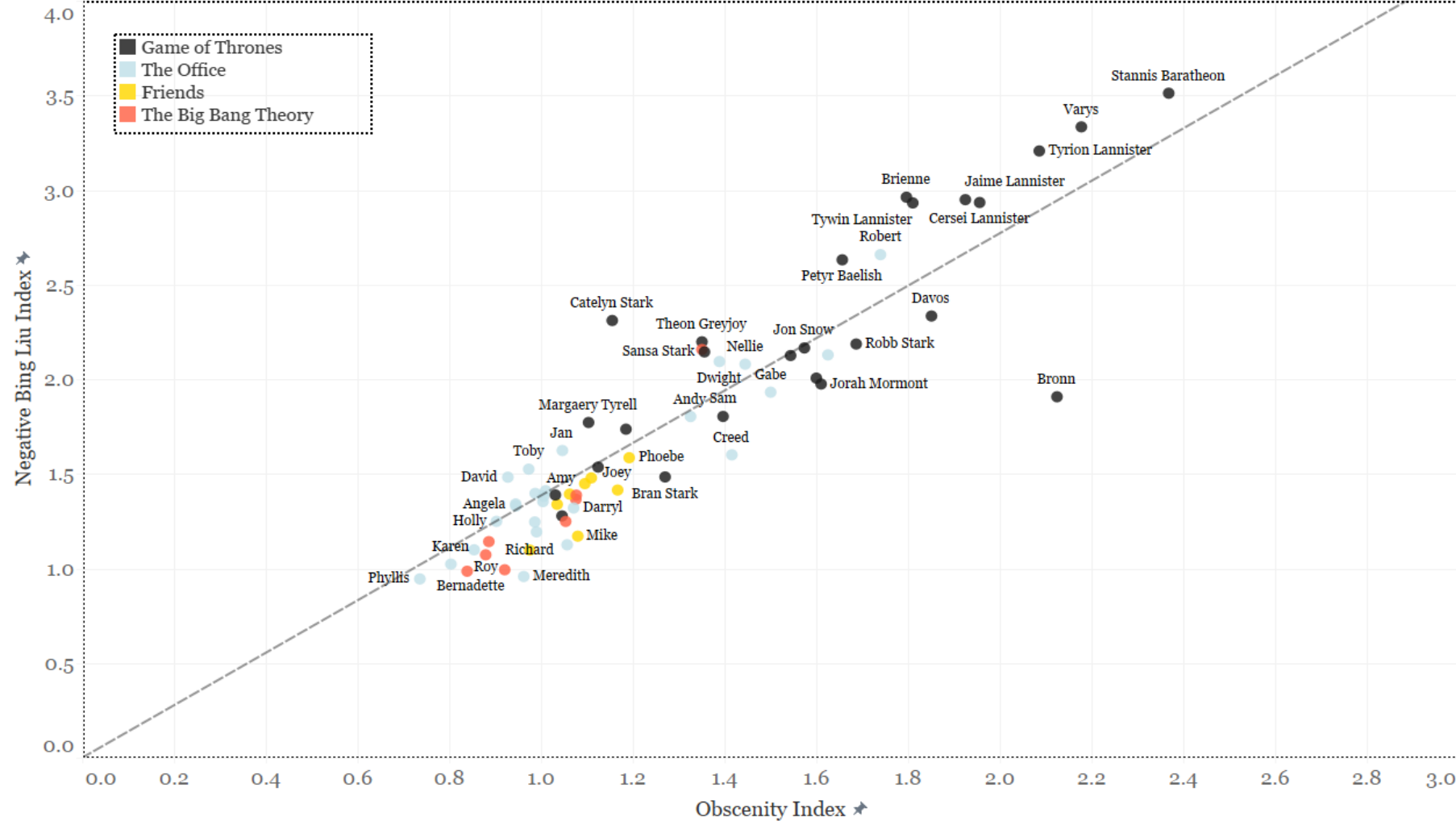
Top Characters - by Dialogue Length



Dialogue Distribution – by Top 6 Characters per Season



# Comparative Trend: Negative BL vs. Obscenity Score



1. There is a **linear trend** between Bing Liu's Negative Index and Obscenity index, showing the words with negative connotation in dialogues and obscene dialogues are interrelated
2. As expected, Game of Thrones has highest obscenity and negative BL index, which explains its **18+ only rating**
3. Characters from friends are similar in negative aspects as well, with all character being very close to each other on the plot. Overall, Friends also has the lowest value of these index

## Obscenity Index

Game of Thrones	2.334
The Office	1.545
The Big Bang Theory	1.144
Friends	1.123

## Negative BL Index

Game of Thrones	3.470
The Office	2.138
The Big Bang Theory	1.634
Friends	1.465

\* The Obscenity index is derived with help of a list of 1,300+ English terms that could be found offensive, located at <https://www.cs.cmu.edu/~biglou/resources/>

## Cosine Similarity among Top 6 Characters across Series

		Friends						Game of Thrones						The Big Bang Theory						The Office					
		Chandler	Joey	Monica	Phoebe	Rachel	Ross	Cersei Lannister	Daenerys Targaryen	Jaime Lannister	Jon Snow	Sansa Stark	Tyrion Lannister	Amy	Howard	Leonard	Penny	Raj	Sheldon	Andy	Angela	Dwight	Jim	Michael	Pam
Friends	Chandler		0.91	0.89	0.93	0.92	0.93	0.34	0.27	0.34	0.34	0.30	0.39	0.52	0.73	0.69	0.72	0.74	0.58	0.69	0.50	0.61	0.66	0.71	0.66
	Joey	0.91		0.89	0.89	0.87	0.93	0.30	0.24	0.31	0.30	0.26	0.35	0.47	0.71	0.67	0.70	0.73	0.54	0.69	0.46	0.58	0.65	0.67	0.65
	Monica	0.89	0.89		0.91	0.91	0.91	0.31	0.24	0.31	0.31	0.28	0.36	0.49	0.69	0.64	0.69	0.69	0.54	0.63	0.48	0.56	0.62	0.66	0.63
	Phoebe	0.93	0.89	0.91		0.95	0.92	0.30	0.23	0.30	0.30	0.27	0.36	0.51	0.73	0.69	0.76	0.75	0.59	0.69	0.50	0.61	0.68	0.72	0.70
	Rachel	0.92	0.87	0.91	0.95		0.89	0.28	0.22	0.28	0.29	0.26	0.34	0.49	0.69	0.65	0.72	0.70	0.56	0.65	0.49	0.57	0.64	0.67	0.65
	Ross	0.93	0.93	0.91	0.92	0.89		0.31	0.25	0.32	0.32	0.28	0.37	0.51	0.74	0.71	0.74	0.77	0.57	0.70	0.50	0.61	0.69	0.72	0.69
Game of Thrones	Cersei Lannister	0.34	0.30	0.31	0.30	0.28	0.31		0.54	0.73	0.56	0.59	0.76	0.34	0.41	0.36	0.32	0.41	0.40	0.37	0.31	0.35	0.30	0.38	0.31
	Daenerys Targaryen	0.27	0.24	0.24	0.23	0.22	0.25	0.54		0.53	0.52	0.46	0.64	0.27	0.32	0.29	0.25	0.32	0.32	0.29	0.25	0.28	0.24	0.31	0.24
	Jaime Lannister	0.34	0.31	0.31	0.30	0.28	0.32	0.73	0.53		0.63	0.58	0.76	0.35	0.42	0.38	0.32	0.41	0.41	0.36	0.31	0.36	0.31	0.38	0.32
	Jon Snow	0.34	0.30	0.31	0.30	0.29	0.32	0.56	0.52	0.63		0.61	0.66	0.34	0.40	0.37	0.32	0.39	0.39	0.35	0.30	0.34	0.31	0.38	0.31
	Sansa Stark	0.30	0.26	0.28	0.27	0.26	0.28	0.59	0.46	0.58	0.61		0.61	0.29	0.35	0.32	0.28	0.34	0.34	0.30	0.27	0.29	0.27	0.33	0.27
	Tyrion Lannister	0.39	0.35	0.36	0.36	0.34	0.37	0.76	0.64	0.76	0.66	0.61		0.39	0.46	0.42	0.37	0.47	0.47	0.41	0.35	0.40	0.35	0.44	0.36
The Big Bang Theory	Amy	0.52	0.47	0.49	0.51	0.49	0.51	0.34	0.27	0.35	0.34	0.29	0.39		0.80	0.89	0.81	0.79	0.67	0.52	0.41	0.46	0.48	0.54	0.49
	Howard	0.73	0.71	0.69	0.73	0.69	0.74	0.41	0.32	0.42	0.40	0.35	0.46	0.80		0.91	0.90	0.91	0.76	0.72	0.52	0.63	0.66	0.73	0.68
	Leonard	0.69	0.67	0.64	0.69	0.65	0.71	0.36	0.29	0.38	0.37	0.32	0.42	0.89	0.91		0.90	0.90	0.73	0.67	0.49	0.59	0.64	0.69	0.65
	Penny	0.72	0.70	0.69	0.76	0.72	0.74	0.32	0.25	0.32	0.32	0.28	0.37	0.81	0.90	0.90		0.89	0.77	0.67	0.49	0.59	0.65	0.70	0.67
	Raj	0.74	0.73	0.69	0.75	0.70	0.77	0.41	0.32	0.41	0.39	0.34	0.47	0.79	0.91	0.90	0.89		0.77	0.73	0.53	0.64	0.67	0.74	0.69
	Sheldon	0.58	0.54	0.54	0.59	0.56	0.57	0.40	0.32	0.41	0.39	0.34	0.47	0.67	0.76	0.73	0.77	0.77		0.59	0.47	0.54	0.54	0.61	0.55
The Office	Andy	0.69	0.69	0.63	0.69	0.65	0.70	0.37	0.29	0.36	0.35	0.30	0.41	0.52	0.72	0.67	0.67	0.73	0.59		0.72	0.81	0.84	0.87	0.83
	Angela	0.50	0.46	0.48	0.50	0.49	0.50	0.31	0.25	0.31	0.30	0.27	0.35	0.41	0.52	0.49	0.49	0.53	0.47	0.72		0.75	0.80	0.79	0.73
	Dwight	0.61	0.58	0.56	0.61	0.57	0.61	0.35	0.28	0.36	0.34	0.29	0.40	0.46	0.63	0.59	0.59	0.64	0.54	0.81	0.75		0.85	0.88	0.90
	Jim	0.66	0.65	0.62	0.68	0.64	0.69	0.30	0.24	0.31	0.31	0.27	0.35	0.48	0.66	0.64	0.65	0.67	0.54	0.84	0.80	0.85		0.92	0.89
	Michael	0.71	0.67	0.66	0.72	0.67	0.72	0.38	0.31	0.38	0.38	0.33	0.44	0.54	0.73	0.69	0.70	0.74	0.61	0.87	0.79	0.88	0.92		0.89
	Pam	0.66	0.65	0.63	0.70	0.65	0.69	0.31	0.24	0.32	0.31	0.27	0.36	0.49	0.68	0.65	0.67	0.69	0.55	0.83	0.73	0.90	0.89	0.89	

- Very high similarity among all Friends characters, further its characters are also like TBBT and The Office characters, maybe due to Sitcom nature of all these shows; even TBBT and The Office characters have high similarity among its characters
- Low similarity between Game of Thrones characters excluding family members (Lannister's)
- Notable cross show similarities include Penny and Phoebe, along with Ross and Raj

0.22  0.95

*\* Cosine Similarity works well the longer documents, here dialogues are first grouped at Series x Top 6 level, then similarity is evaluated for each pair*



## Same Words, Different Voices: Characters Who Speak Alike



Cosine Similarity  
0.74 1.00

Hey, do uh, do you have a minute?	Hey, you got a minute?	0.75
I can't catch a break!	Man, I cannot catch a break.	0.89
I don't know what to say.	I don't know what to say.	1.00
I don't know, but I ah, I have the feeling that my being there will do it. ...	Give me back that juice.	0.74
I need juice! People need juice!!	Give me back that juice.	0.82
Oh my God! You did that yourself?	Oh, my God, what did you do?	1.00
Okay, guys, wish me luck.	Anyway, wish me luck.	0.85
There was never a good time.	Really not a good time.	0.83
	This is not a good time.	1.00
What the hell are you doing?	What the hell are you doing?	1.00
What, oh my God. What did you do?	Oh, my God, what did you do?	1.00



Cosine Similarity  
0.71 1.00

Here, kitty-kitty. Here kitty-kitty. Where did you go, little kitty-kitty-kitty? Here kitty-kitty-kitty-kitty...	Oh, hon... Uh... (singing) Soft kitty, warm kitty...	0.71
	Soft kitty, warm kitty, little ball of fur.	0.73
	Soft kitty, warm kitty, little ball of fur. Happy kitty, um...	0.82
	Wait, wait. Let's sing it as a round. I'll start. Soft kitty, warm kitty, little ball of fur... See..	0.73
Honey that sounds like fun.	That sounds like fun.	0.83
Joey, what-what are you doing going into my bedroom?!	What are you doing in my bedroom?	0.79
My God! What happened to you?	Oh, my God, what happened?	0.92
Oh my God! What happened?	Oh, my God! What happened?	1.00
Oh wow, it's so beautiful...	Oh, that's so beautiful.	0.81
Oh wow. Are you in trouble.	You're in trouble.	0.76
Okay, let me just change.	Yeah, I just have to change.	0.71
So Mac and C.H.E.E.S.E.	Want some mac and cheese?	0.73
Yes, but you can not tell Ross 'cause I want to surprise him.	'Cause I want to surprise him.	0.77

1. All positive sentiment showcase a negative correlation with IMDB ratings of Game of Thrones, which could imply that **negative content do help in driving the ratings** of the show
2. IMDB plot with Positive Vader index show that ratings were high at 3 of the least positive episodes (based on Positive Vader)

## Correlation – by IMDB & Sentiment

	Friends	Game of Thrones	The Big Bang Theory	The Office
Text Blob	0.01	-0.26	-0.07	-0.03
Positive LM	-0.06	-0.21	0.14	0.11
Positive Vader	-0.01	-0.17	-0.21	-0.15
Positive BL	-0.13	-0.07	0.06	0.07
Negative Vader	0.00	-0.02	-0.01	0.08
Obscenity	-0.11	0.10	0.17	0.12
Negative LM	-0.08	0.12	0.19	0.13
Negative BL	-0.17	0.15	0.24	0.18

-0.26 0.24

## Games of Thrones IMDB Vs. Positive Sentiment – by Season x Episode

