

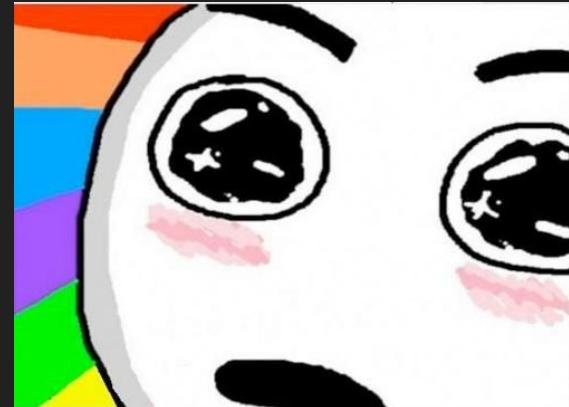
# Meme Review: Dead or Alive?

*Predicting Reader Response to Timely Memes using Machine Learning*

Phase 2

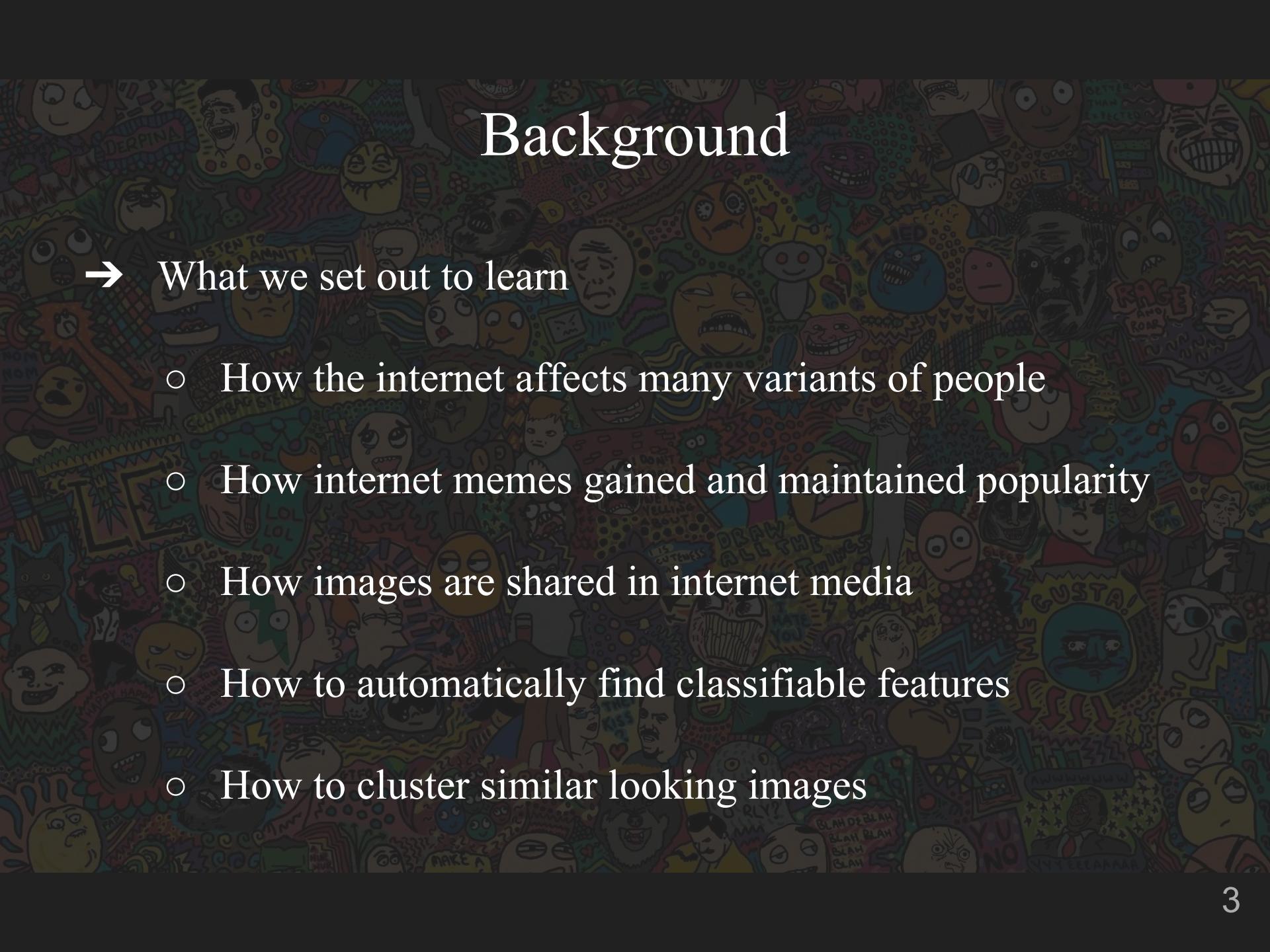
# Motivation & Problem Definition Recap

- 2 out 5 Facebook posts are memes
  - Memes are a great **marketing** tool
  - Millennials LOVE memes
- 
- Online Businesses and **Influencers**
  - Massive Data dump



When corporations try to use memes  
in their social media marketing





# Background

→ What we set out to learn

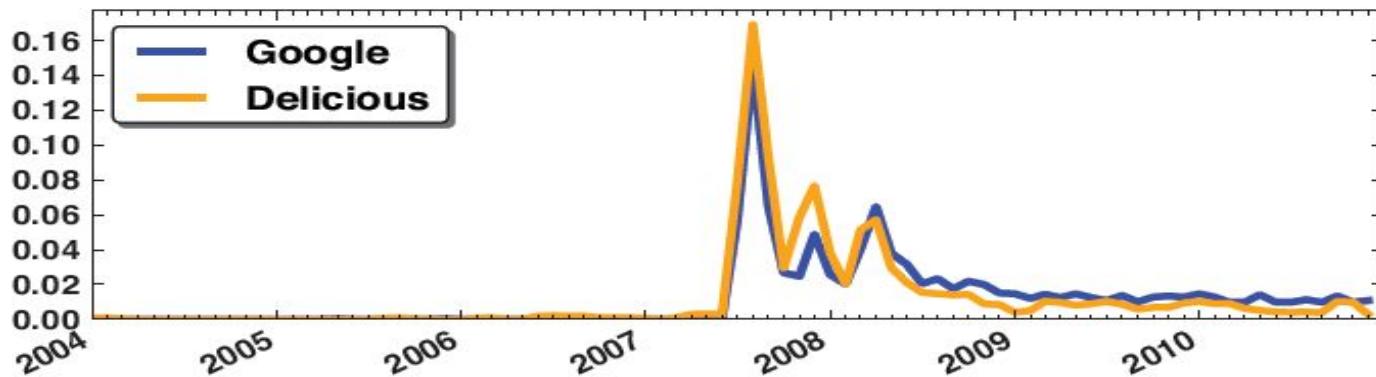
- How the internet affects many variants of people
- How internet memes gained and maintained popularity
- How images are shared in internet media
- How to automatically find classifiable features
- How to cluster similar looking images

# “Insights into Internet Memes”

- It is commonly assumed that Internet memes spread virally but scientific evidence as to this assumption is scarce.
- Investigate the epidemic dynamics of 150 famous Internet memes.
- The analysis is based on time series data that were collected from Google Insights, Delicious, Digg, and StumbleUpon.

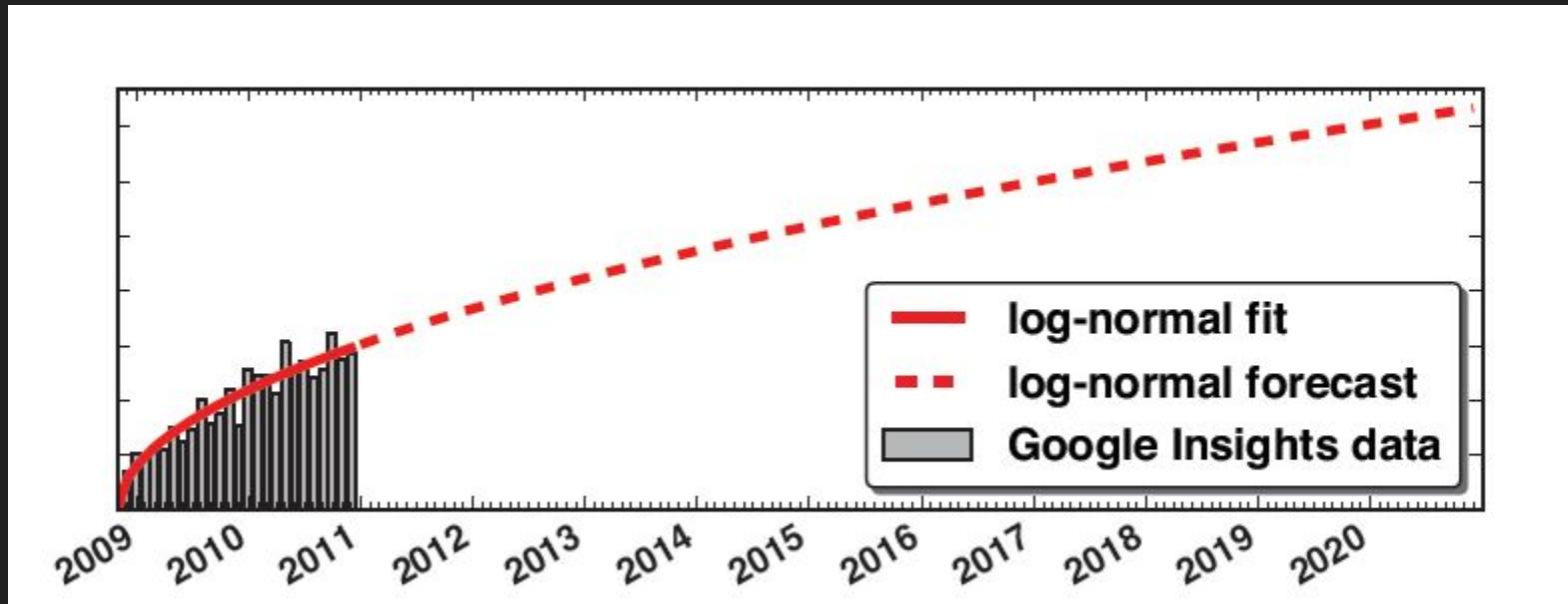


(a) instances of the “chocolate rain” meme



(b) two time series (retrieved from Google Insights and Delicious) reflecting the rise and decline in popularity of this Internet meme

- Found that differential equation models from mathematical epidemiology as well as simple log-normal distributions give a good account of the growth and decline of memes.



- Traditional compartment models might be good but lack flexibility.
- Log-normal distributions are more practical.
- Thus, concluding by saying memes might be viral but only on social media platforms and not the rest of the internet.

# “Predicting the Content Diffusion Path in Online Social Networks”

# “Marketing with Twitter: Challenges and Opportunities”

- 255 million monthly active users
- Specific features may have more overall effect in the final decision of a user response
- Upvotes, Author, Dates

# “A Study of Meme Propagation”

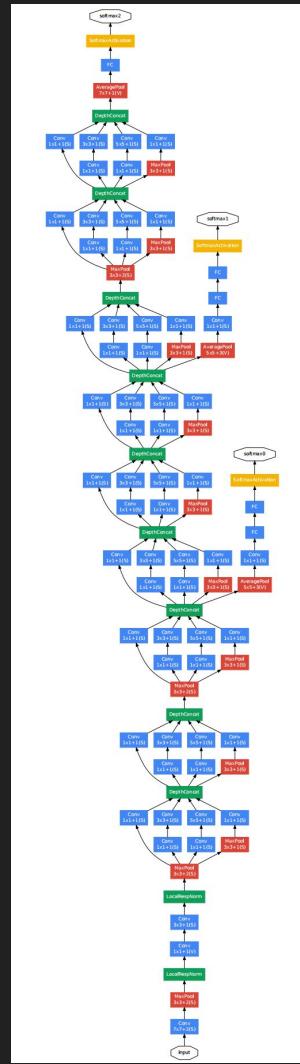
- Memes gain momentum and then lose their popularity over time
- Categorized the Memes, i.e. Political memes vs General Memes

# “Going Deeper with Convolutions”

- Google's own research paper proposing the Deep Convolutional **Inception** Neural Net
  - By C. Szegedy & 8 others from Google Inc, University of North Carolina and
  - Published to Cornell's ArXiv on the 17th of September 2014
- ImageNet Large-Scale Visual Recognition Challenge 2014 **1st Place Winner** for Classification and Detection.

# Confusing Convolutions

type	patch size/ stride	output size	depth
convolution	$7 \times 7 / 2$	$112 \times 112 \times 64$	1
max pool	$3 \times 3 / 2$	$56 \times 56 \times 64$	0
convolution	$3 \times 3 / 1$	$56 \times 56 \times 192$	2
max pool	$3 \times 3 / 2$	$28 \times 28 \times 192$	0
inception (3a)		$28 \times 28 \times 256$	2
inception (3b)		$28 \times 28 \times 480$	2
max pool	$3 \times 3 / 2$	$14 \times 14 \times 480$	0
inception (4a)		$14 \times 14 \times 512$	2
inception (4b)		$14 \times 14 \times 512$	2
inception (4c)		$14 \times 14 \times 512$	2
inception (4d)		$14 \times 14 \times 528$	2
inception (4e)		$14 \times 14 \times 832$	2
max pool	$3 \times 3 / 2$	$7 \times 7 \times 832$	0
inception (5a)		$7 \times 7 \times 832$	2
inception (5b)		$7 \times 7 \times 1024$	2
avg pool	$7 \times 7 / 1$	$1 \times 1 \times 1024$	0
dropout (40%)		$1 \times 1 \times 1024$	0
linear		$1 \times 1 \times 1000$	1
softmax		$1 \times 1 \times 1000$	0



# Stats and Scores

Team	Year	Place	Error (top-5)	Uses external data
SuperVision	2012	1st	16.4%	no
SuperVision	2012	1st	15.3%	Imagenet 22k
Clarifai	2013	1st	11.7%	no
Clarifai	2013	1st	11.2%	Imagenet 22k
MSRA	2014	3rd	7.35%	no
VGG	2014	2nd	7.32%	no
GoogLeNet	2014	1st	6.67%	no

Table 2: Classification performance.

Team	Year	Place	mAP	external data	approach
UvA-Euvision	2013	1st	22.6%	none	Fisher vectors
Deep Insight	2014	3rd	40.5%	ImageNet 1k	CNN
CUHK DeepID-Net	2014	2nd	40.7%	ImageNet 1k	CNN
GoogLeNet	2014	1st	43.9%	ImageNet 1k	CNN

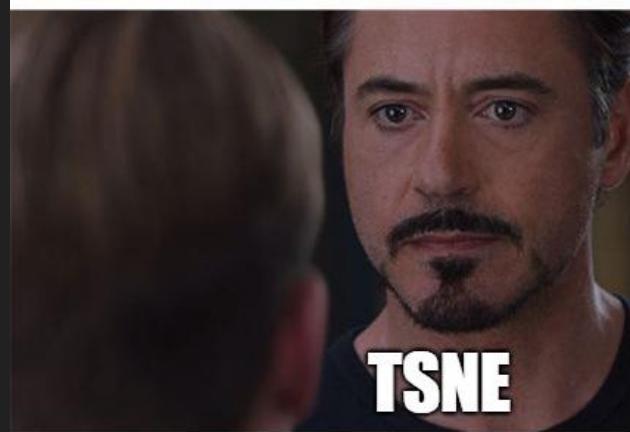
Table 4: Comparison of detection performances.

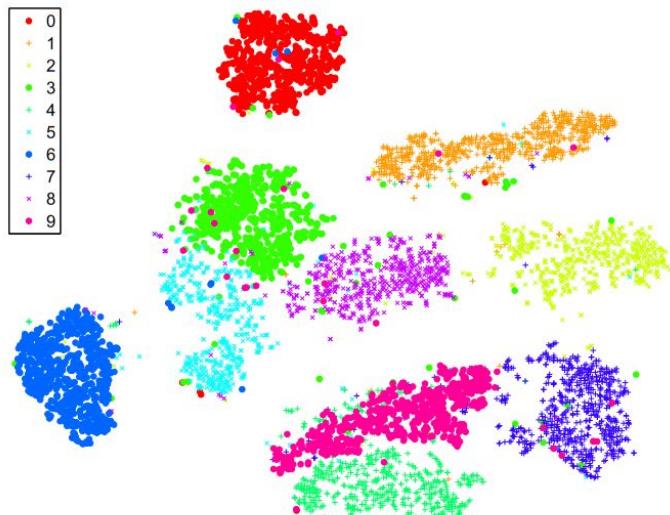
# “Visualizing Data using TSNE”

- T-Distributed Stochastic Neighbor Embedding
- Published by Laurens van der Maaten of Tilburg University and Geoffrey Hinton of University of Toronto
- First published to The Journal of Machine Learning Research on the 9th of November 2008

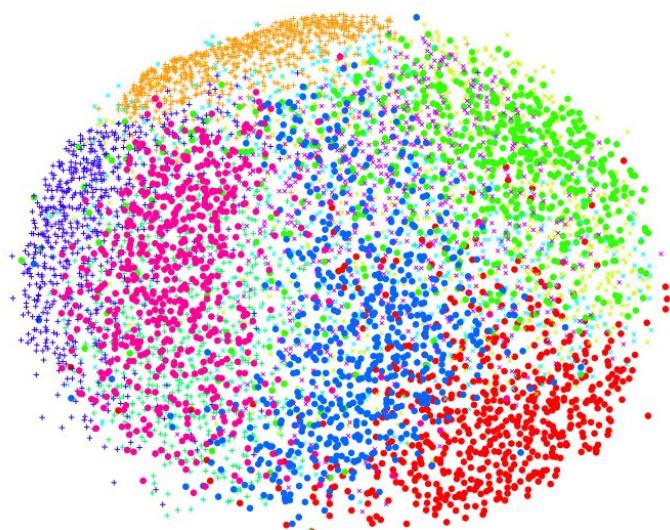
# Dimension Warping

- TSNE is a Dimensionality Reduction technique
- TSNE Projection can map and cluster High Dimensional Data to a 2D or 3D Space by representing noticeable features as vectors





(a) Visualization by t-SNE.



(b) Visualization by Sammon mapping.

Figure 2: Visualizations of 6,000 handwritten digits from the MNIST data set.

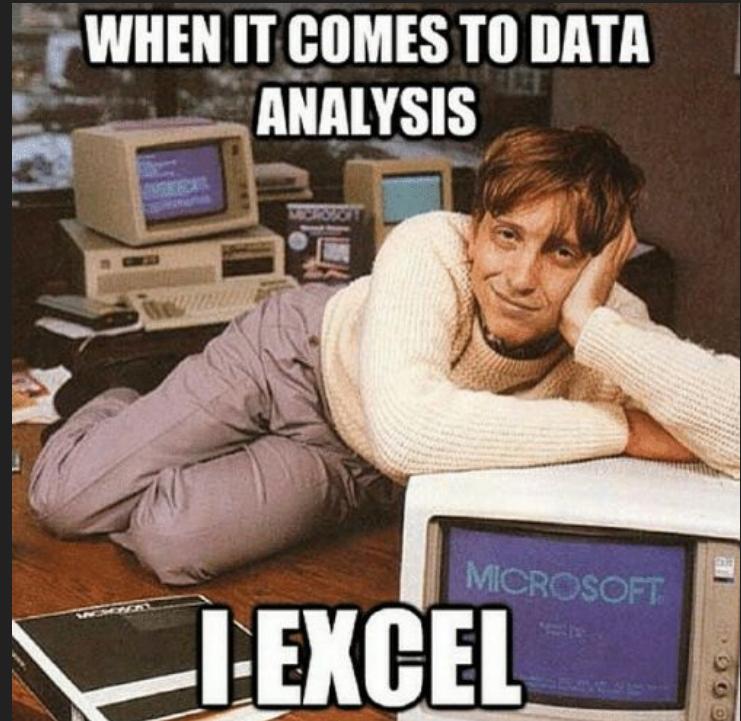
- Easily Separable Classifications
- Optimized for learning and preserving non-linear patterns in HD datasets

Back to  
Meme Review

What have We done so far?

# Our Data Source & Preprocessing

- Kaggle
- Reddit meme dump
- JSON metadata
  
- Image Preprocessing
  - ◆ Resize Resolution
  - ◆ Convert All images to JPG
- Process JSON metadata
  - ◆ Messy JSON Organization
  - ◆ UNIX-based UTC Timestamp



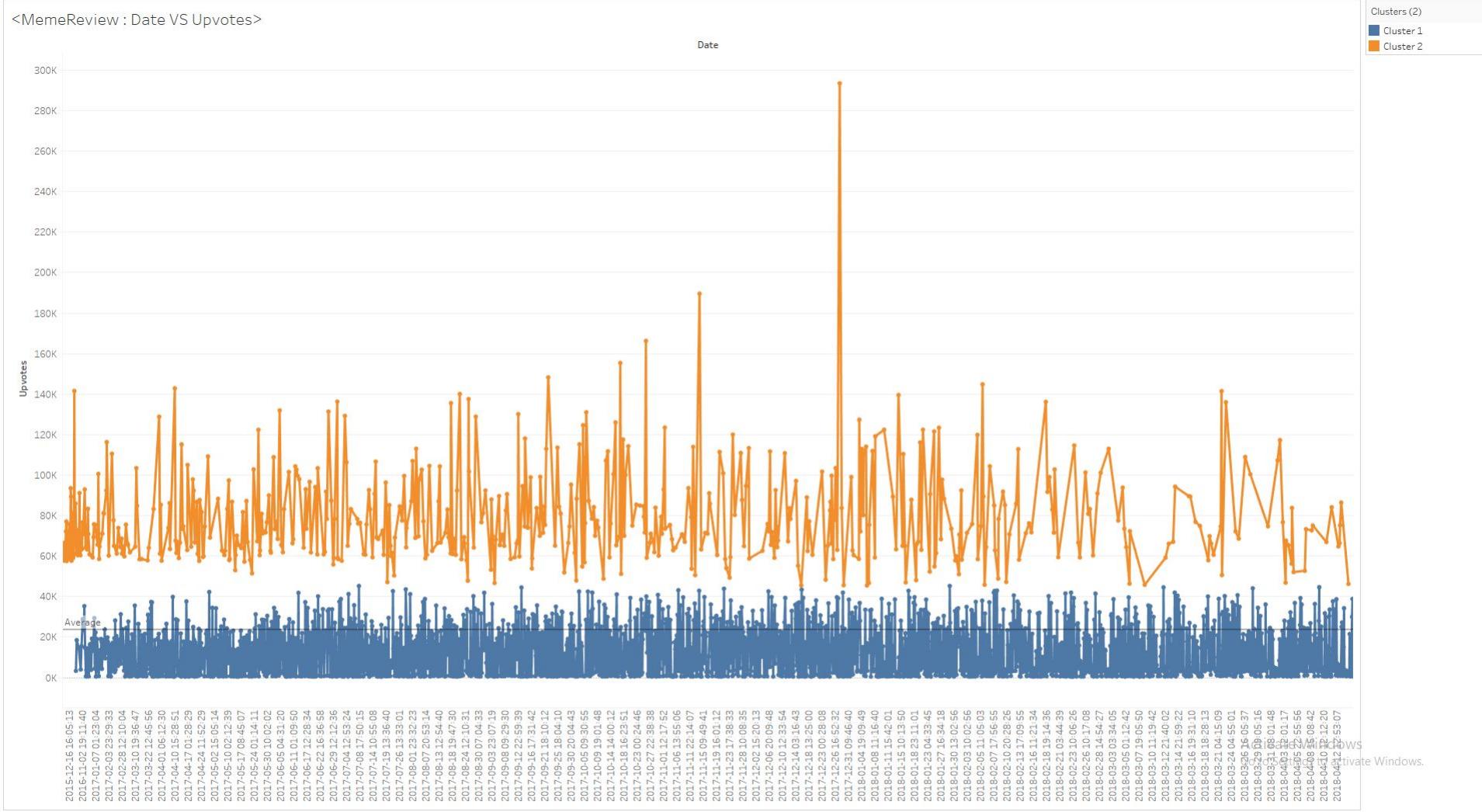
# Neural Networks

- Inception
  - Used before the TSNE Cluster to produce classifiable features
- AlexNet
- VGGnet



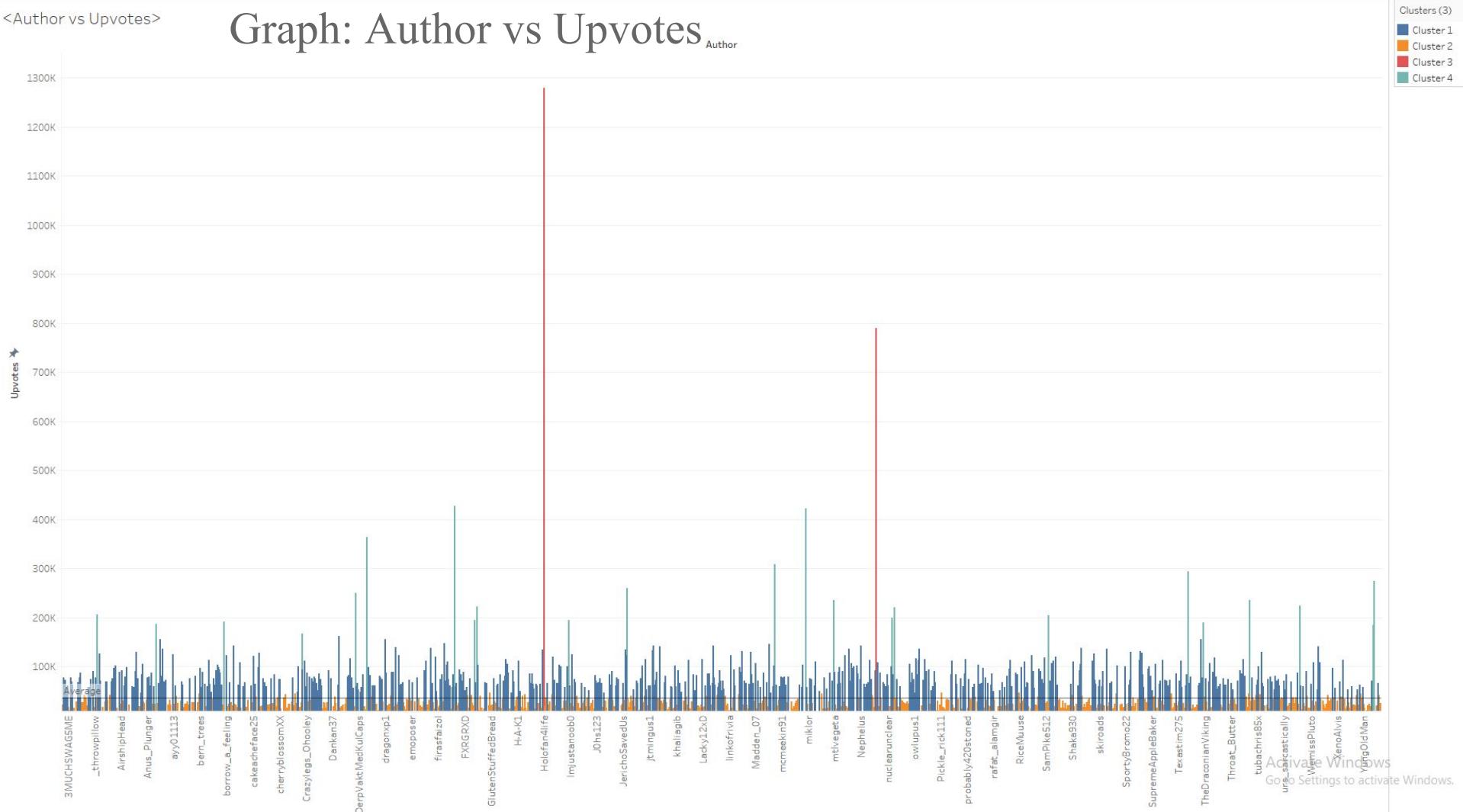
# Data Visualization of Correlations

<MemeReview : Date VS Upvotes>

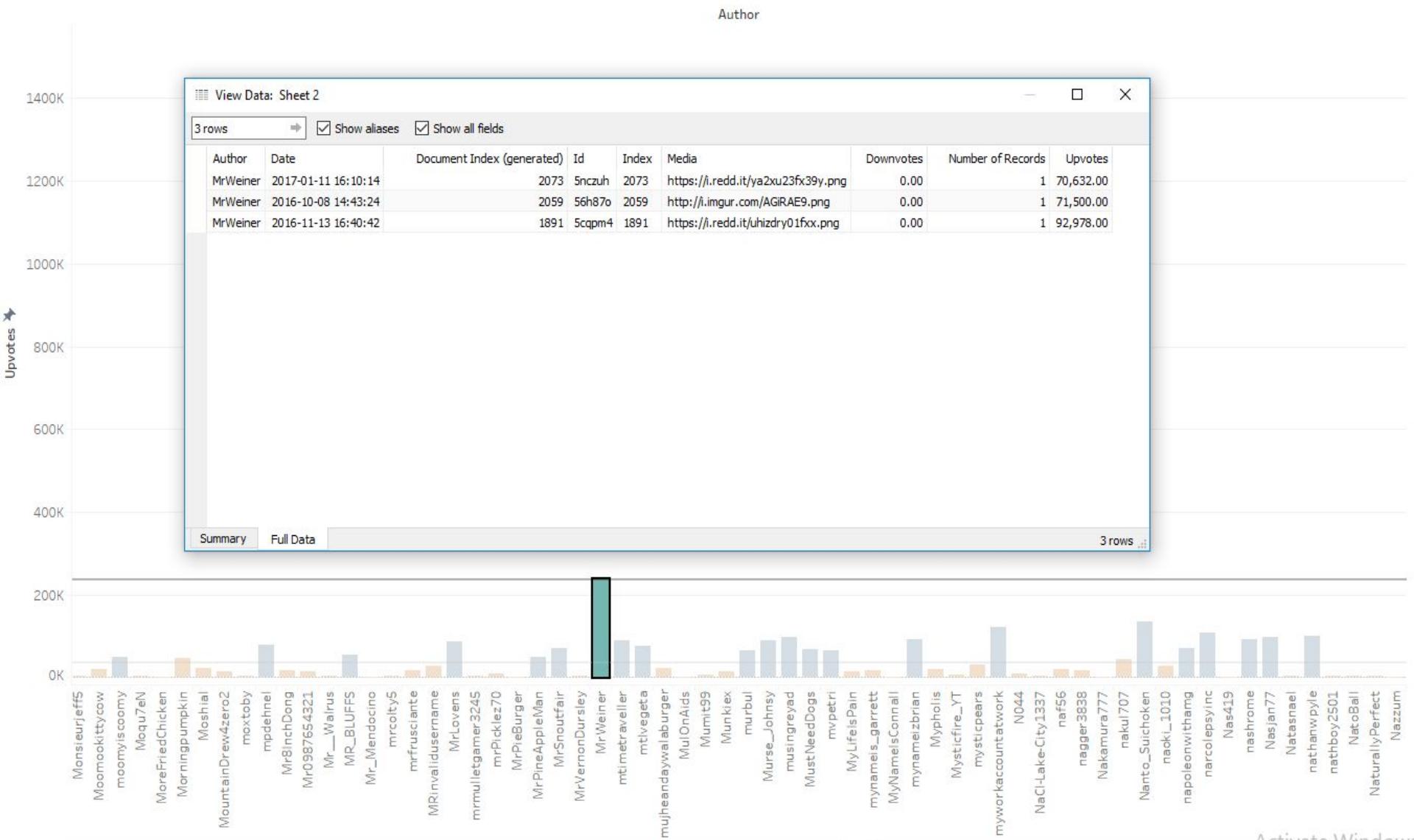


Graph: Date vs Upvotes

# Data Visualization of Correlations



## <Author vs Upvotes>



Activate Windows  
Go to Settings to activate



MrWeiner

u/MrWeiner

I draw comics at smbc-comics.com, and other stuff.

Karma

1,153,632

Cake day

April 17, 2009

FOLLOW

MORE OPTIONS

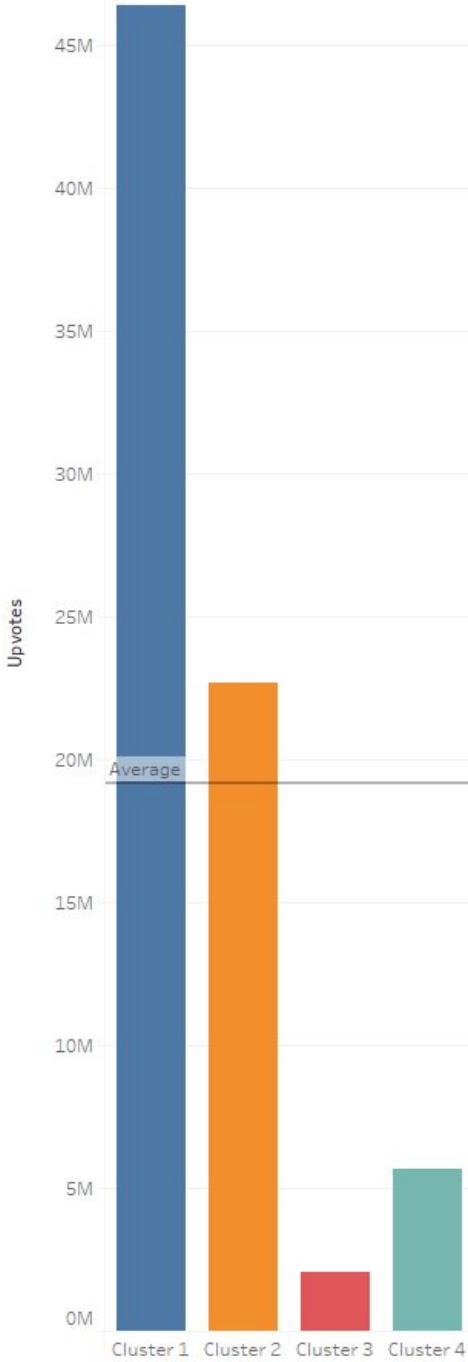
#### CONNECTED ACCOUNTS



Zach Weinersmith

Get exclusive Patreon flair by becoming a patron





# Author Clusters

4 Types  
of  
“Meme-creators”

# See in Numbers

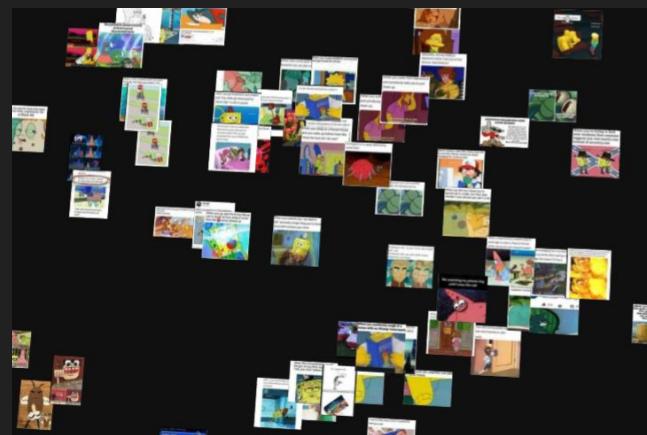
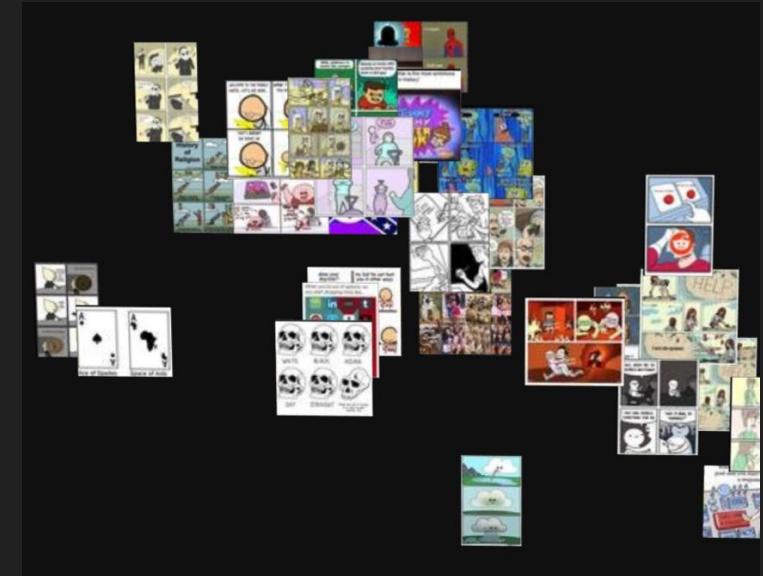
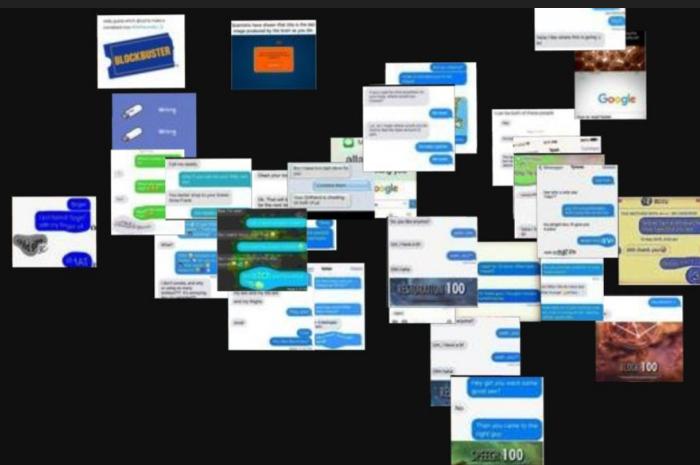
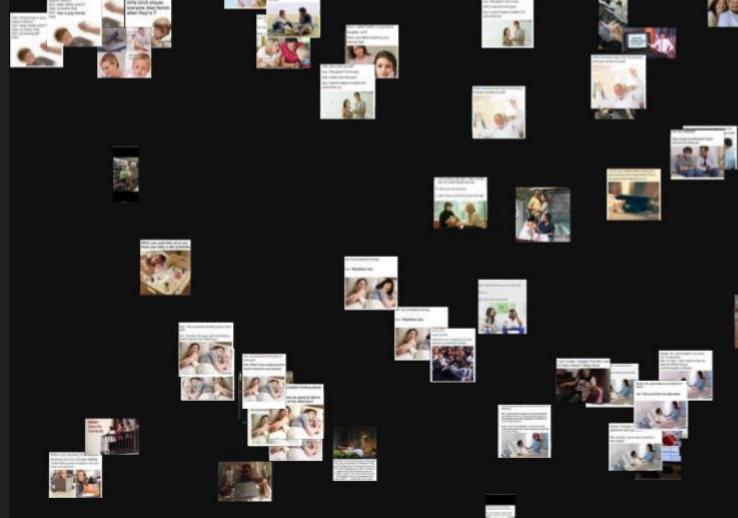
- Produce Image Vectors from Inception Object Detection Net node weights using TF
- Use previous Vectors for Clustering UMAP, which is very similar to TSNE, to project onto an HTML Canvas running D3.js



# A UMAP Map of Memes



# Clusters of Memes

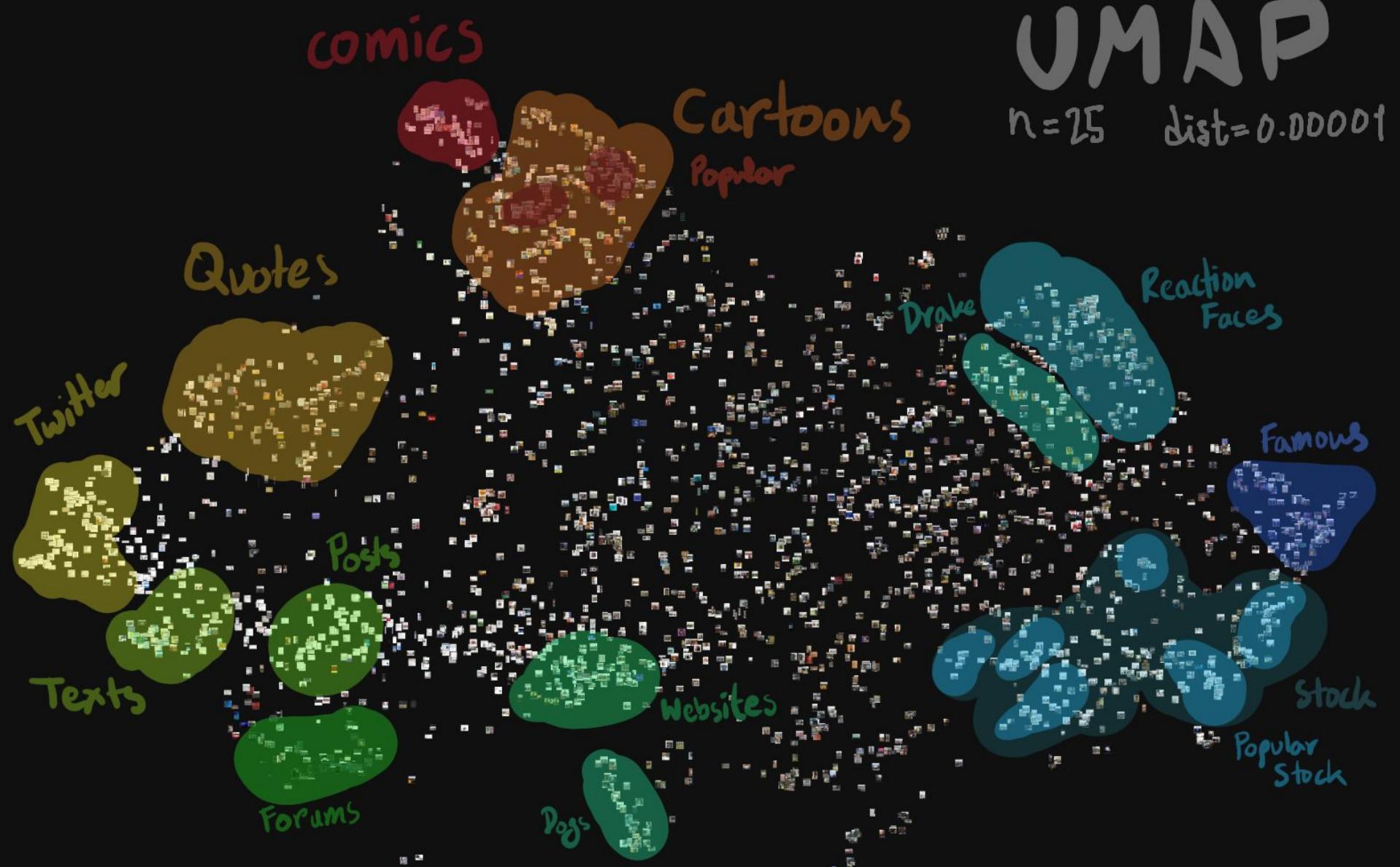




# UMAP

n=25

dist=0.00001



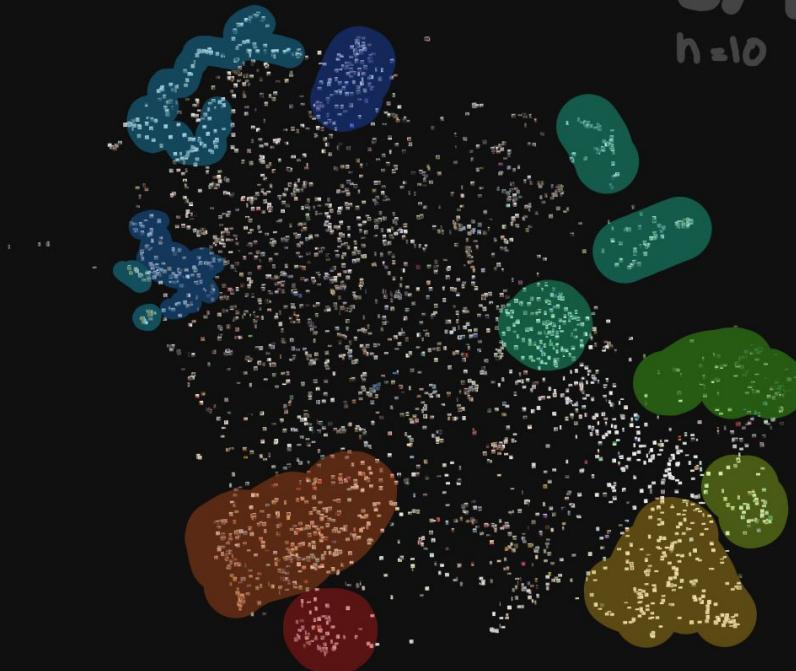
# UMAP

n=10 dist=0.00001



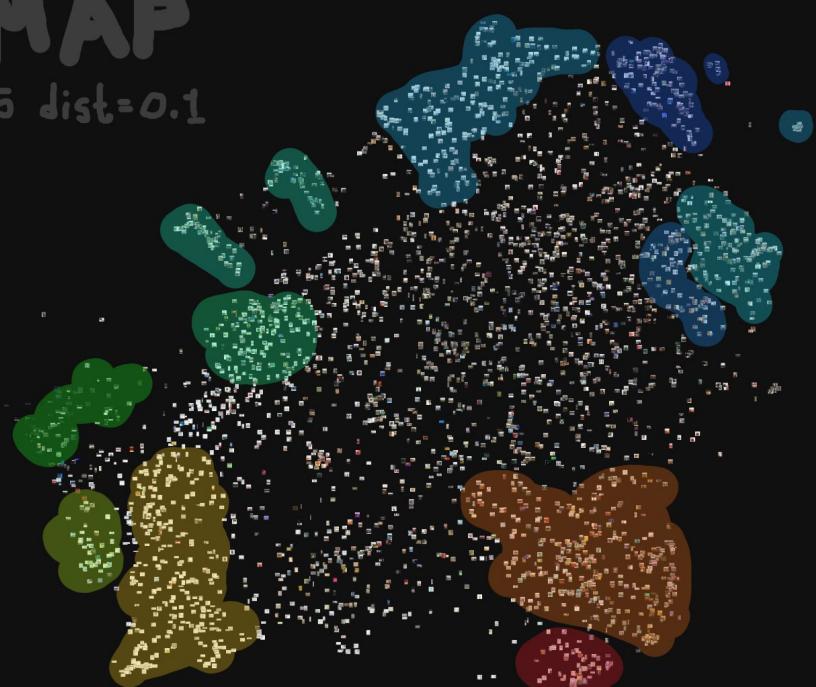
# UMAP

n=10 dist=0.1

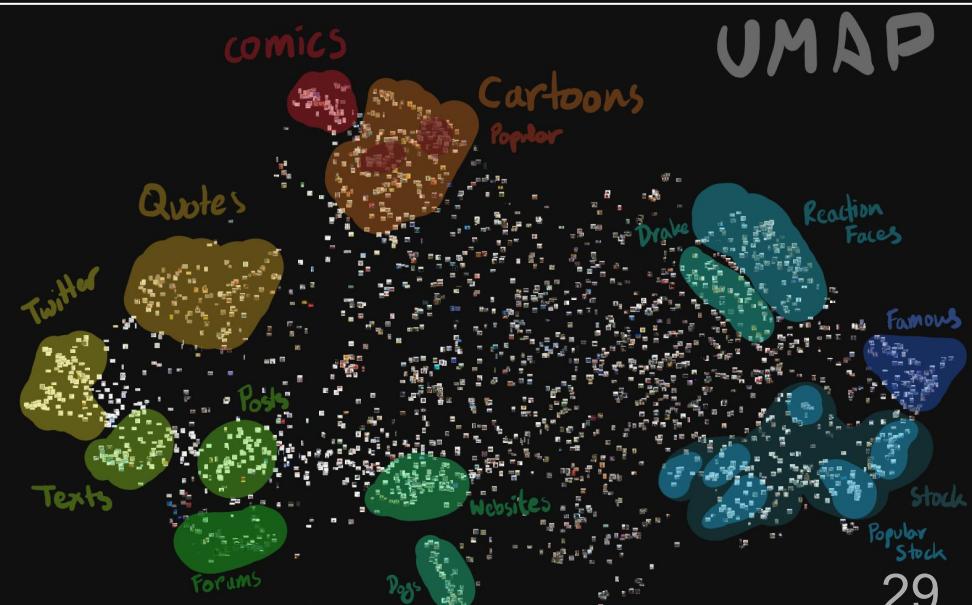


# UMAP

n=25 dist=0.1



# UMAP



# Contributors

- Abir Rahman
  - *Marketing with Twitter*
  - *A Study of Meme Propagation*
- Jawad Aziz Khan
  - *Going Deeper with Convolutions*
  - *Visualizing Data with TSNE*
- Taufiq Rahman
  - *Insights into Internet Memes*
  - *Predicting the Content Diffusion Path in Online Social Networks*

# Any Questions?

