AMT Second Round

- First round
  - Video topic expressed by Freebase
  - Perception of quality assessed by whether user is willing to share content, whether she likes it and whether she believes it will become popular
  - Videos with at least 5000 views. In a pair, one video has 50% more views than the other
- Second round (differences only)
  - Videos from the same month (April 2012)
  - 3 videos in each pop groups
    - [10, 100); [1000, 10,000); [100,000, 1,000,000)
  - All pairs (36) comparison following round robin tournament
  - **Each video potentially watched by 8*8 users**

Results in 3 metrics

- Do users agree among themselves?
  - Kappa Score
  - Does not have to agree with YouTube
- Do users agree with YouTube?
  - Measured only when users agree
  - Binomial CI
- Is there a correlation between user perception and popularity?
  - Kendall or Spearman

Kappa Results

- Are users collective perception of quality biased towards one video of the pair?
  - **Does not have to be the video with more views!**

	Which video do you like the most?	Which video would you share with Friends?	Which video do you predict will become more popular?
P-val <= 0.05	A fraction of **0.30** pairs pass the test	**0.19**	**0.5**
P-val <= 0.01	**0.25**	**0.11**	**0.41**
P-val <= 0.001	**0.22**	**0.11**	**0.33**

Similar to previous results. Prediction perception is stronger than others

Kappa Results

- **Baseball**

	Which video do you like the most?	Which video would you share with Friends?	Which video do you predict will become more popular?
P-val <= 0.05	A fraction of **0.30** pairs pass the test	**0.19**	**0.5**
P-val <= 0.01	**0.25**	**0.11**	**0.41**
P-val <= 0.001	**0.22**	**0.11**	**0.33**

- **Music Videos**

	Which video do you like the most?	Which video would you share with Friends?	Which video do you predict will become more popular?
P-val <= 0.05	A fraction of **0.13** pairs pass the test	**0.05**	**0.19**
P-val <= 0.01	**0.13**	**0.05**	**0.16**
P-val <= 0.001	**0.11**	**0.05**	**0.08**

Binomial results

- Ok. So, when users **agree** do they get the most popular video right?

Testing views with user pred
Testing with 36 pairs
Binomal Proportion CI of agreements: 0.889 +- 0.236
Binomial Test for 0.5 (random chance): p = 0.001

Testing views with user share
Testing with 14 pairs
Binomal Proportion CI of agreements: 1.000 +- 0.410
Binomial Test for 0.5 (random chance): p = 0.016

Testing views with user like
Testing with 22 pairs
Binomal Proportion CI of agreements: 1.000 +- 0.285
Binomial Test for 0.5 (random chance): p = 0.001

Binomial results

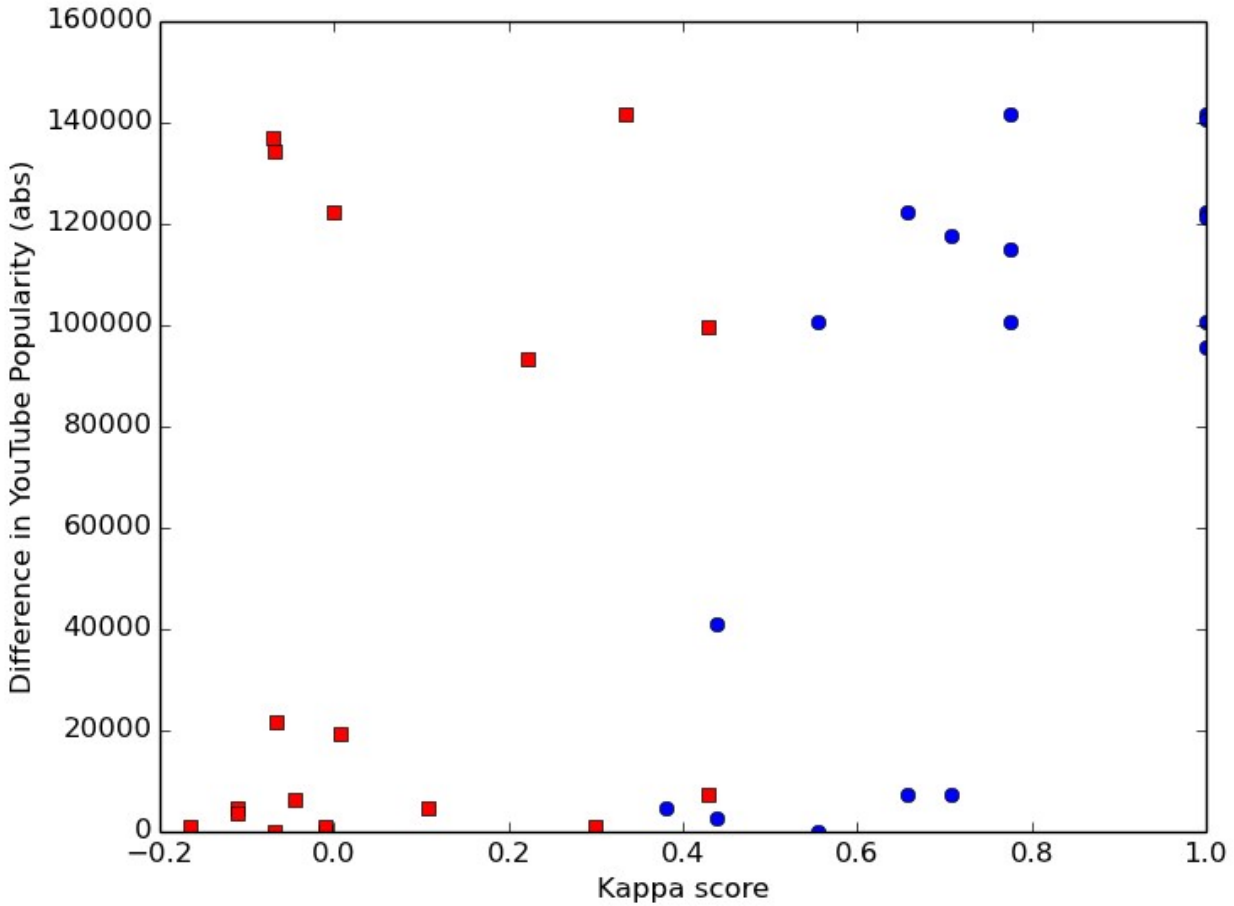
- Music

Testing views with user pred
Testing with 14 pairs
Binomal Proportion CI of agreements: 1.000 +- 0.410
Binomial Test for 0.5 (random chance): p = 0.016

Testing views with user share
Testing with 4 pairs
Binomal Proportion CI of agreements: 0.500 +- 0.487
Binomial Test for 0.5 (random chance): p = 1.000

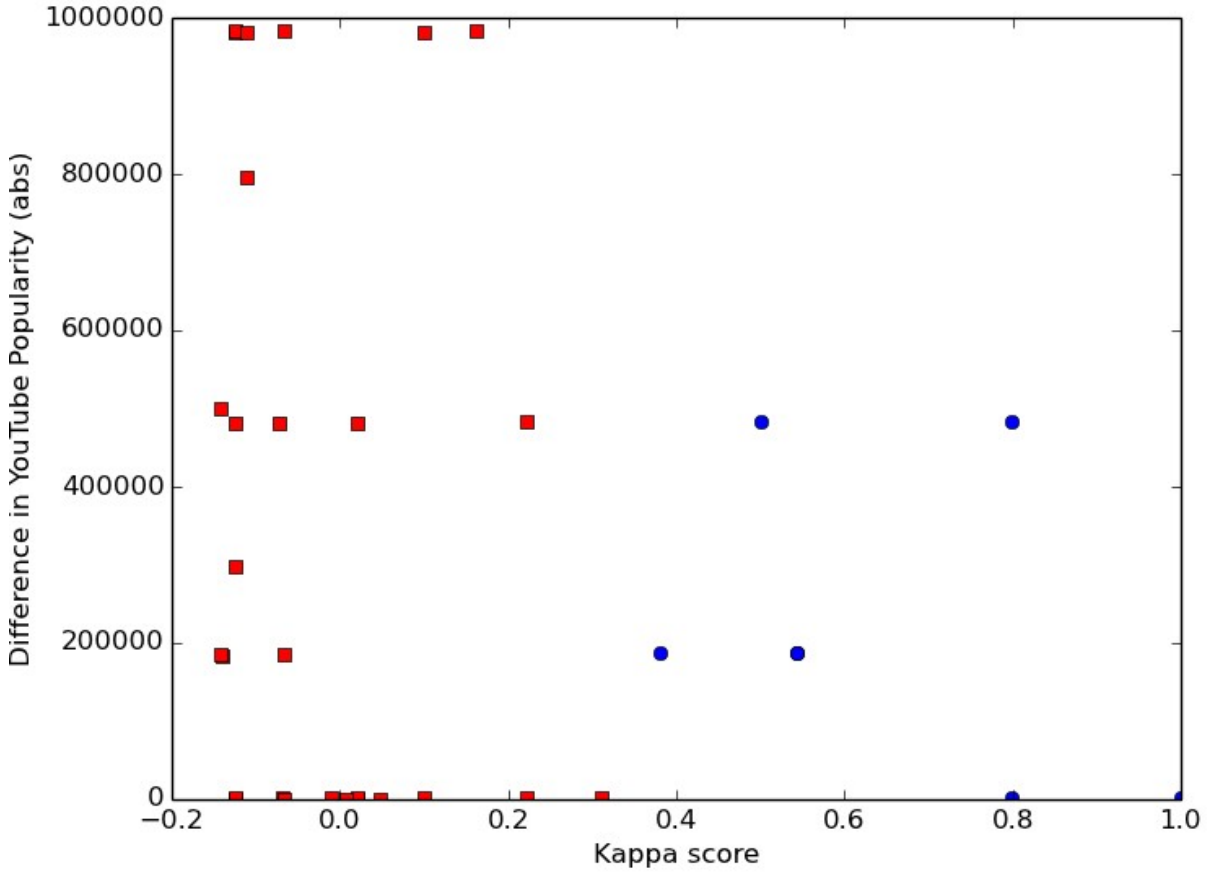
Testing views with user like
Testing with 10 pairs
Binomal Proportion CI of agreements: 0.800 +- 0.516
Binomial Test for 0.5 (random chance): p = 0.375

Correlation of Pairs By Kappa Score with Difference in Views



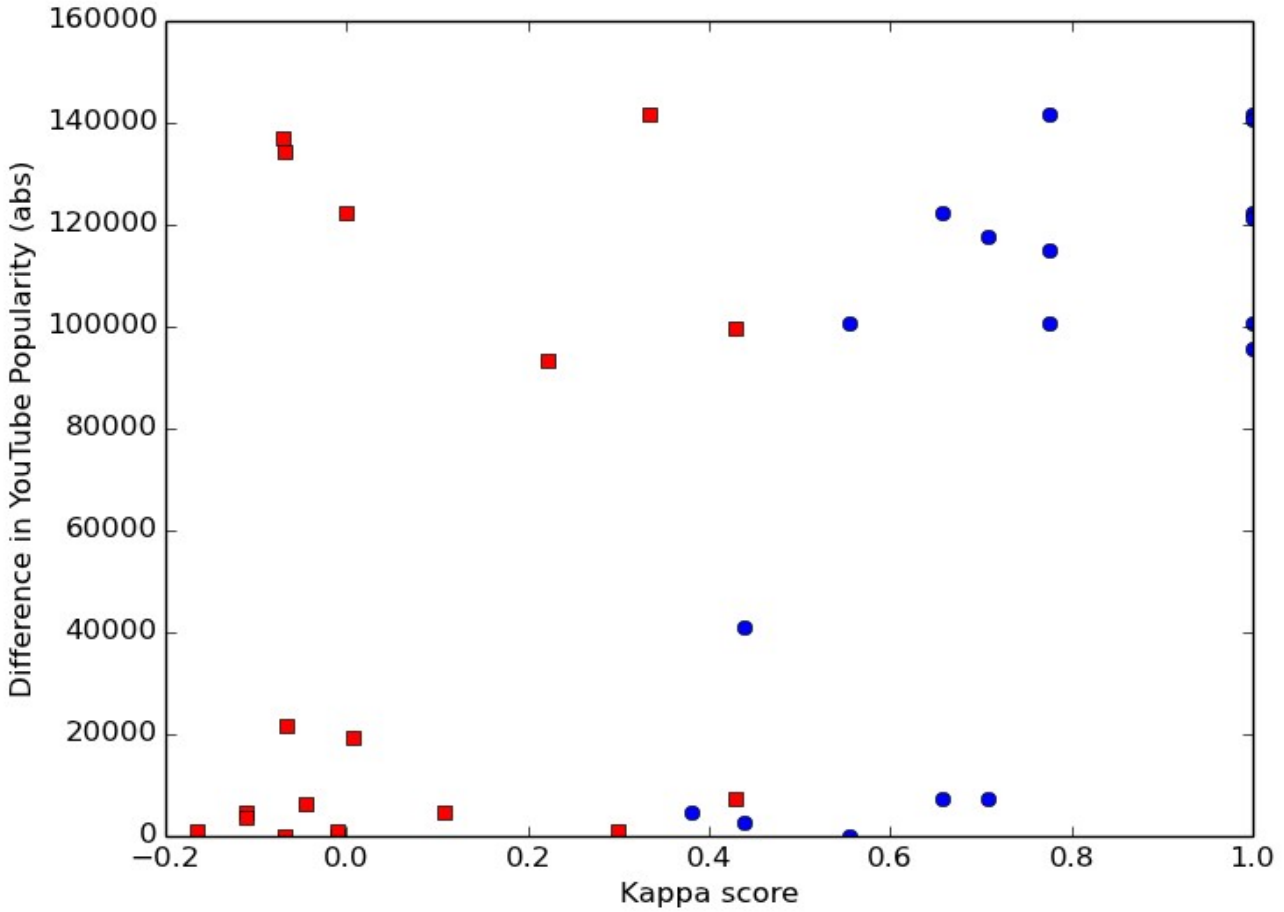
Baseball with predictions. Spearman correlation of 0.65, Kendall of 0.47. Both with p-val < 0.01 for blue dots (agreements)

Correlation of Pairs By Kappa Score with Difference in Views



Music dataset with predictions. Spearman correlation of -0.46, Kendall of -0.41. Both with p-val < 0.28 for blue dots (agreements). Bad correlations

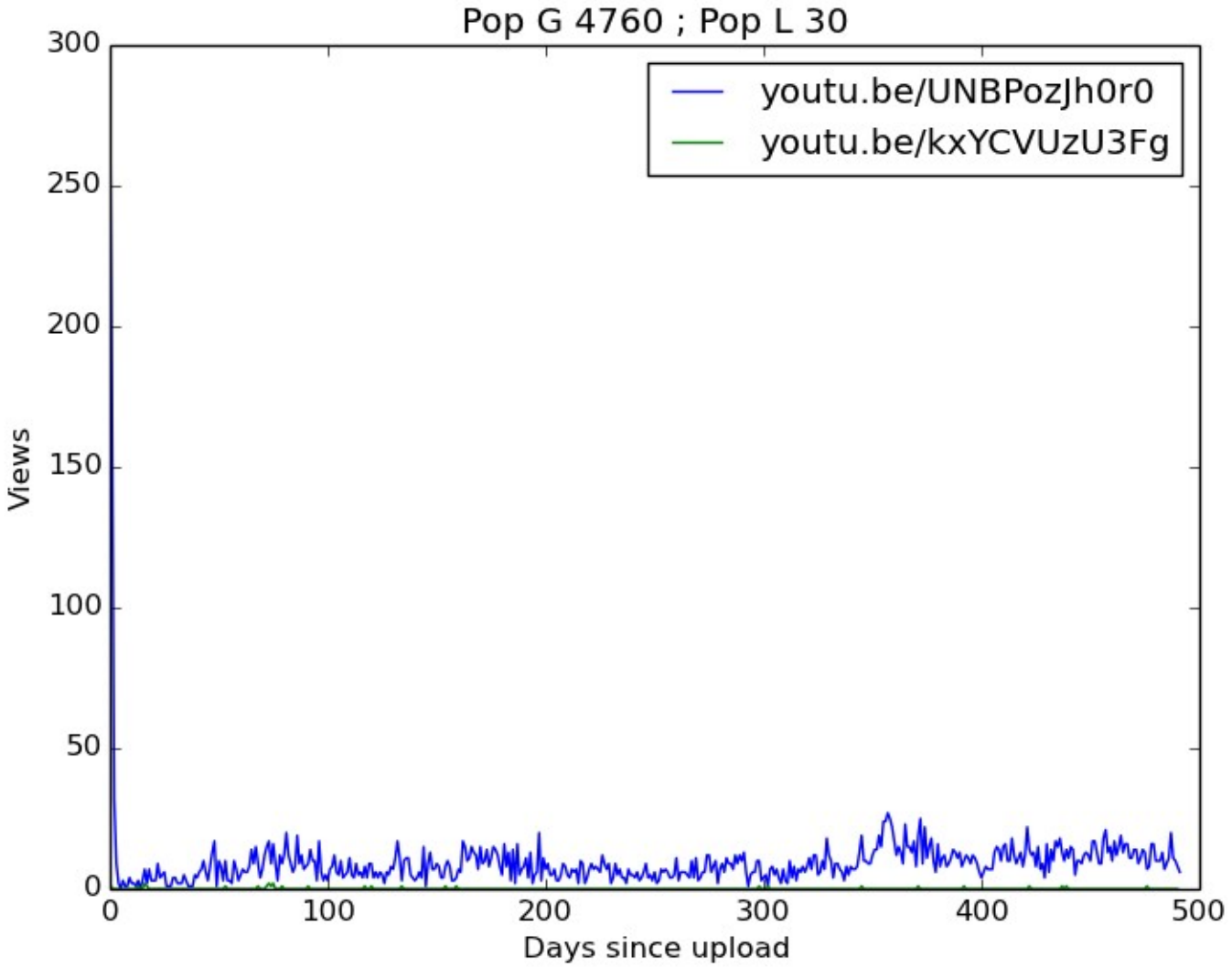
Correlation of User Perception Aggregated with Views



Baseball with predictions. Spearman correlation of 0.65, Kendall of 0.47. Both with p-val < 0.01 for blue dots (agreements)

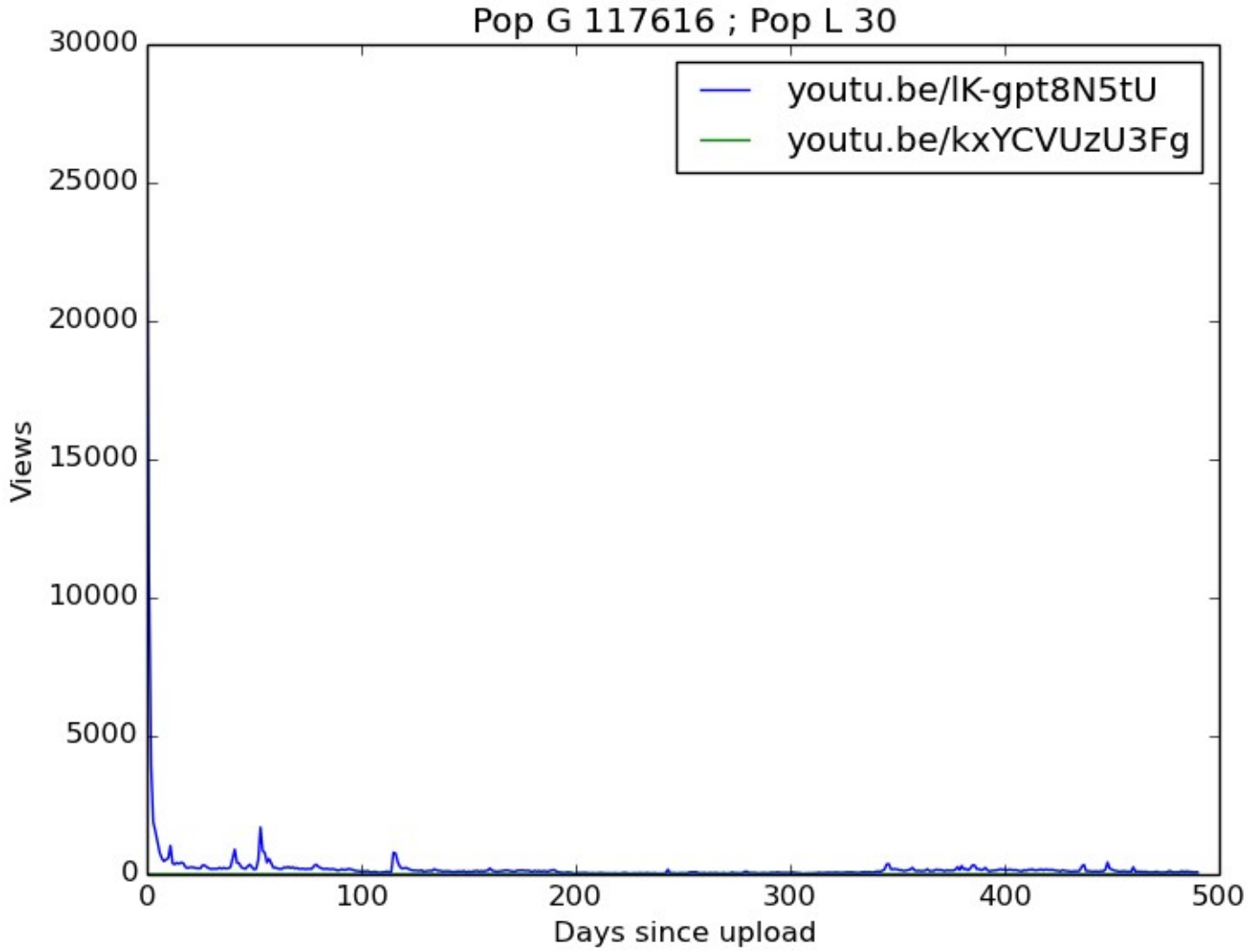
Lets have a look at cases when users disagree!

- **Baseball**



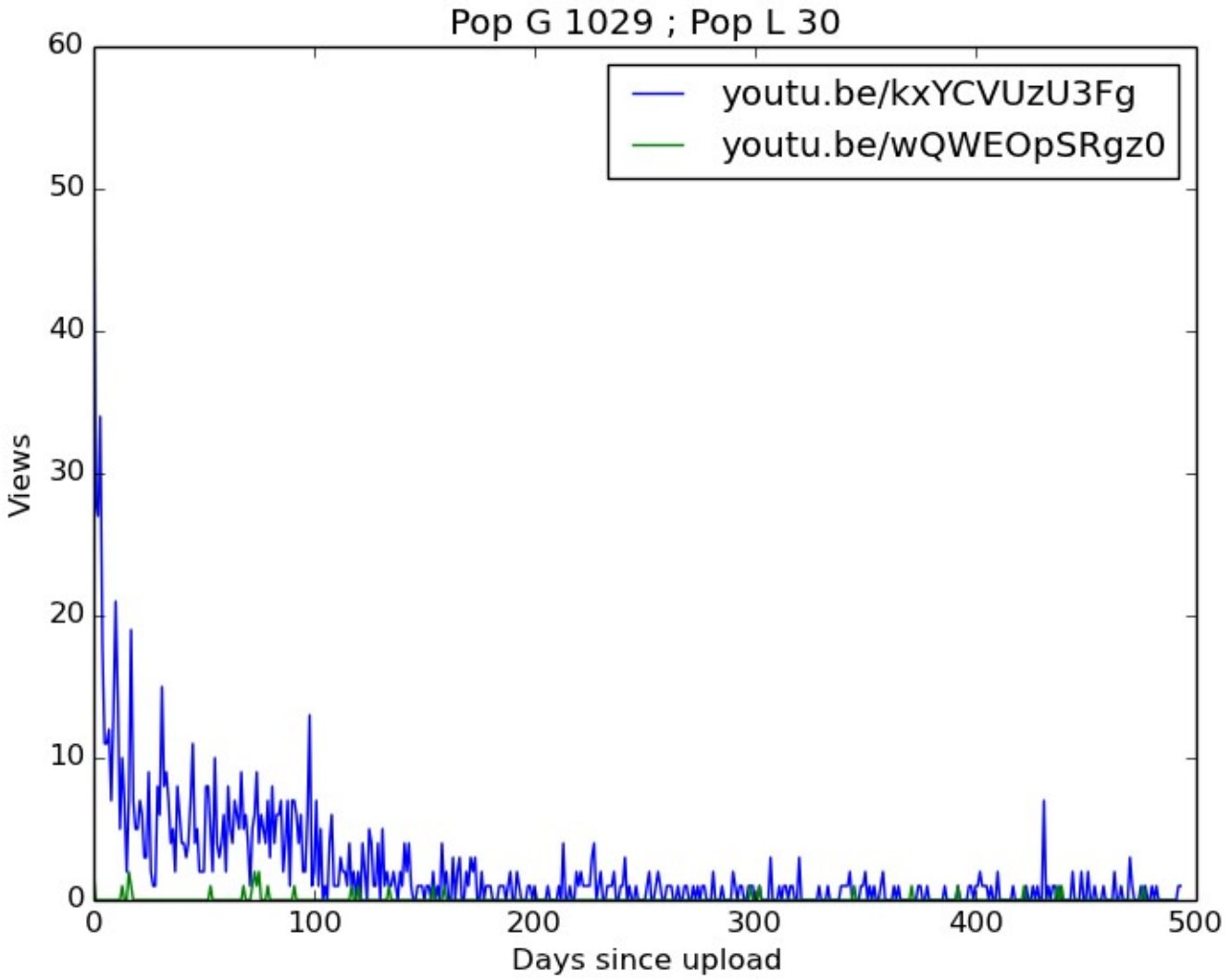
Lets have a look at cases when users disagree!

- **Baseball**



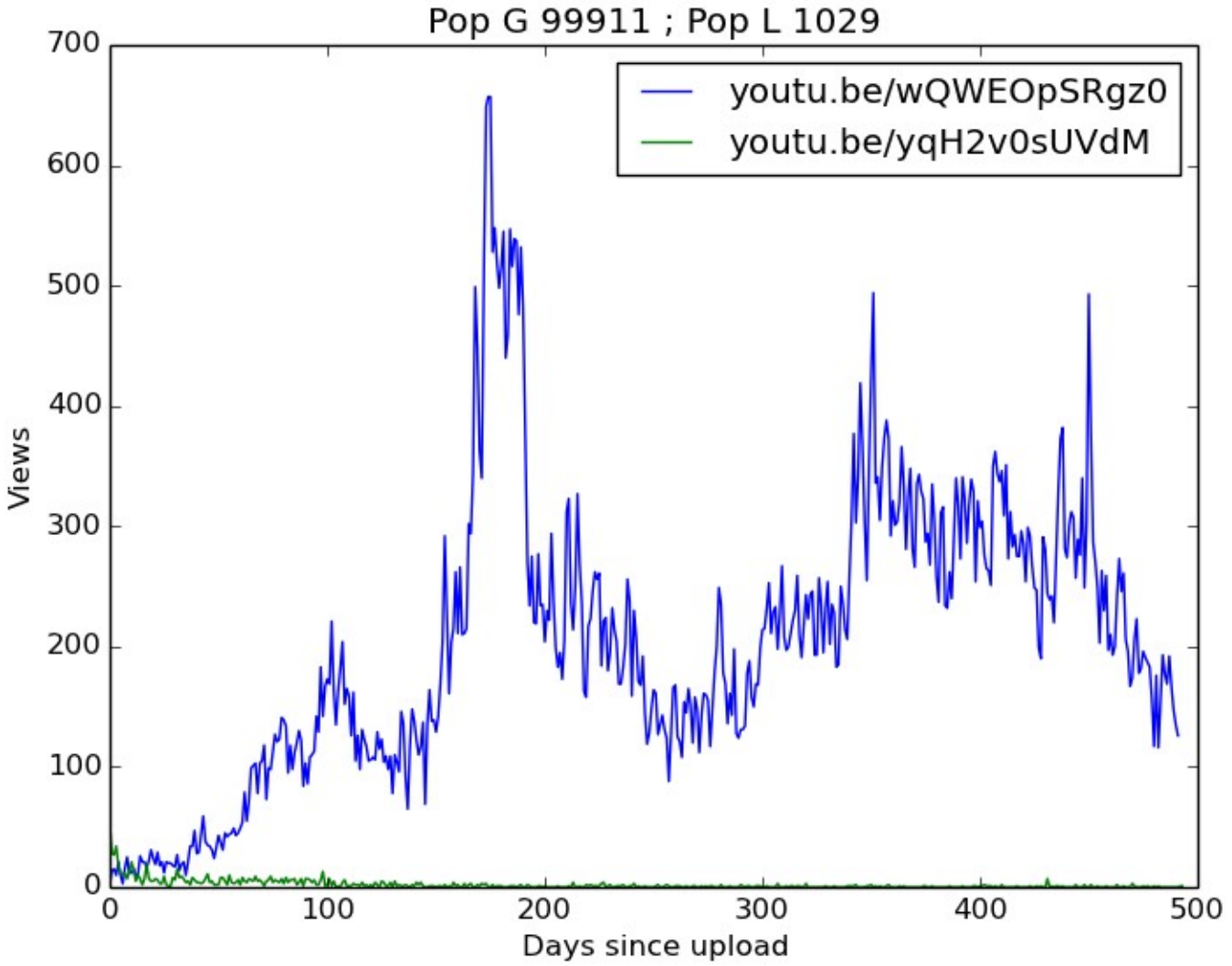
Lets have a look at cases when users disagree!

- **Baseball**



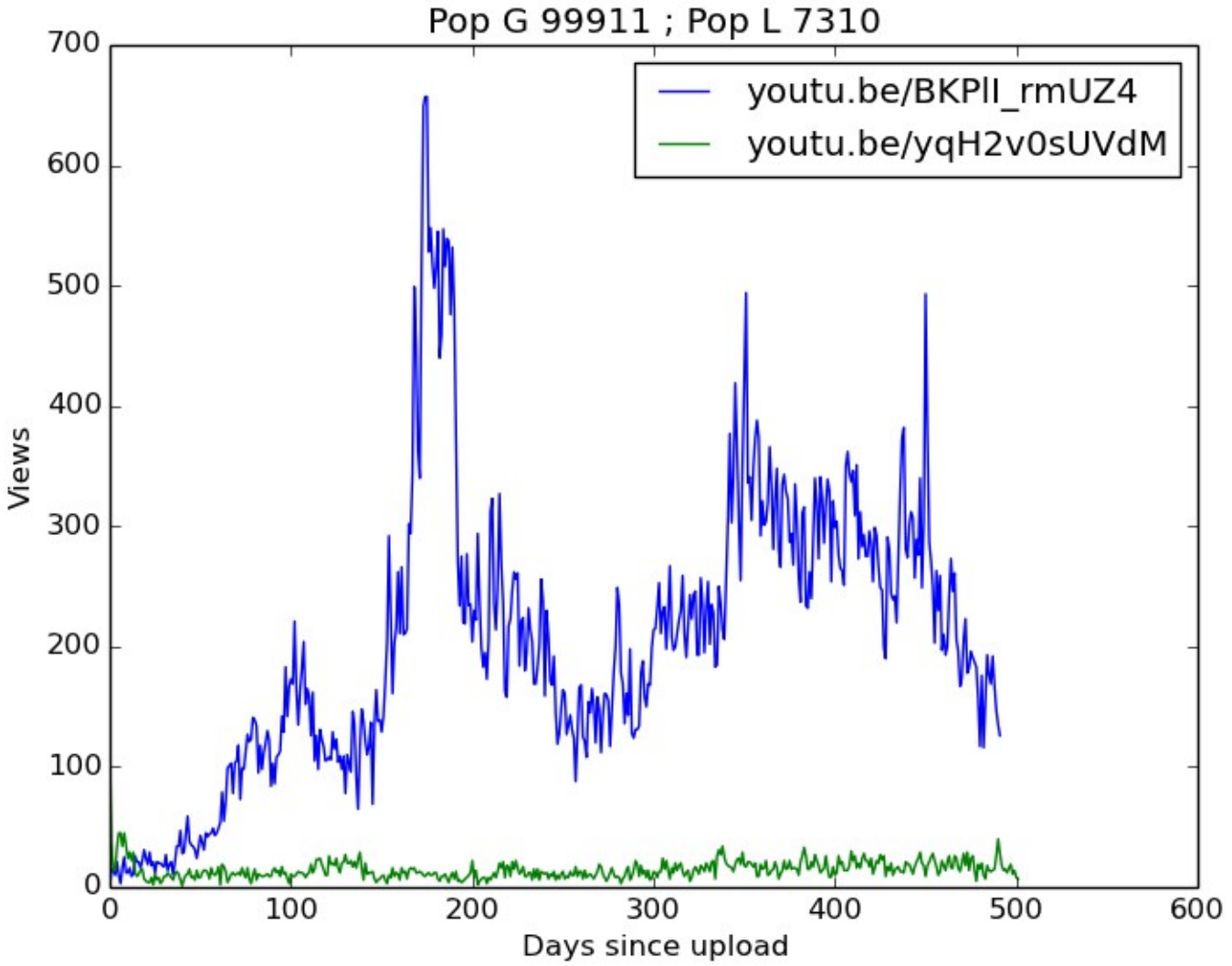
Lets have a look at cases when users disagree!

- **Baseball**



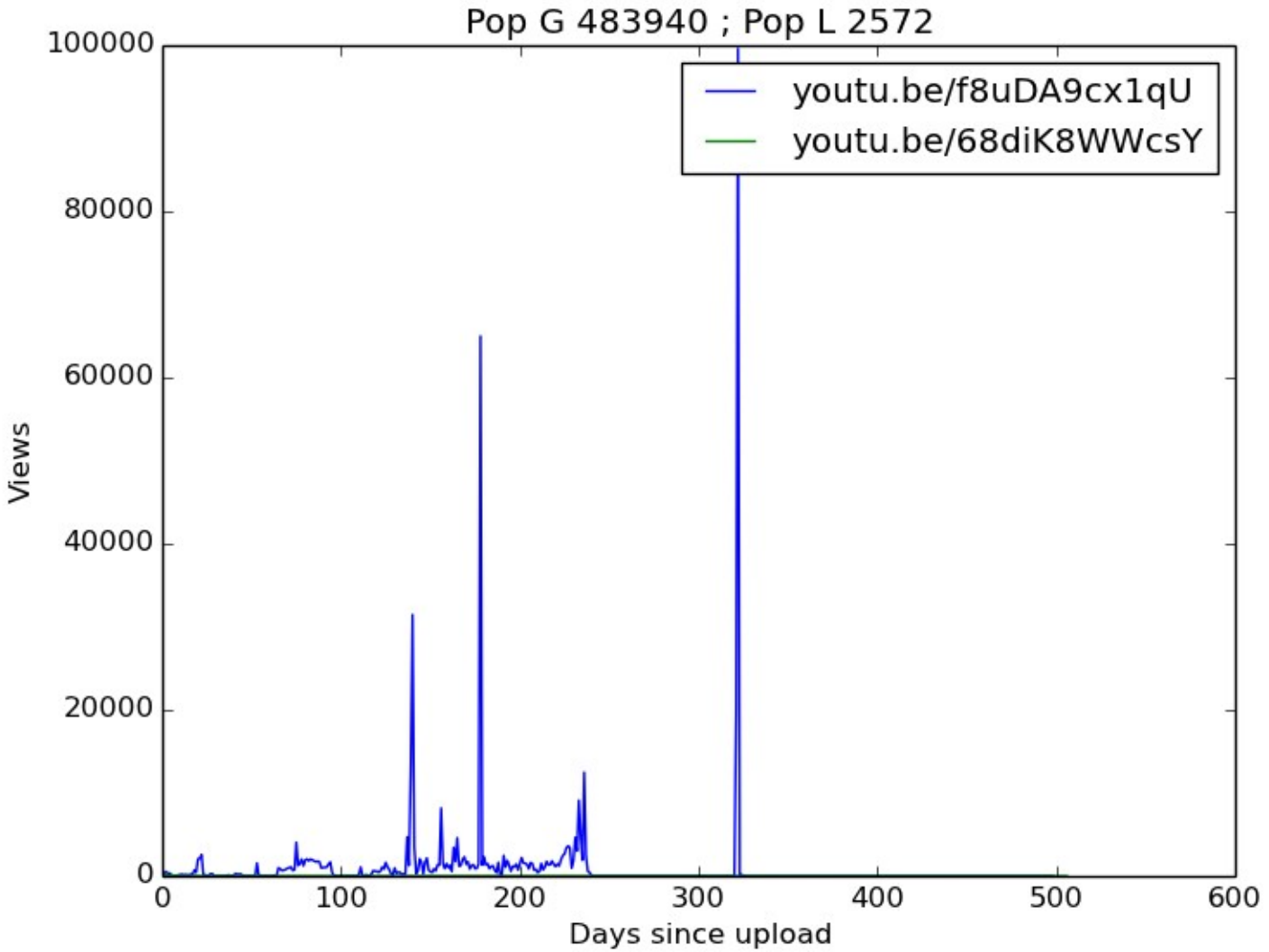
Lets have a look at cases when users disagree!

- **Baseball**



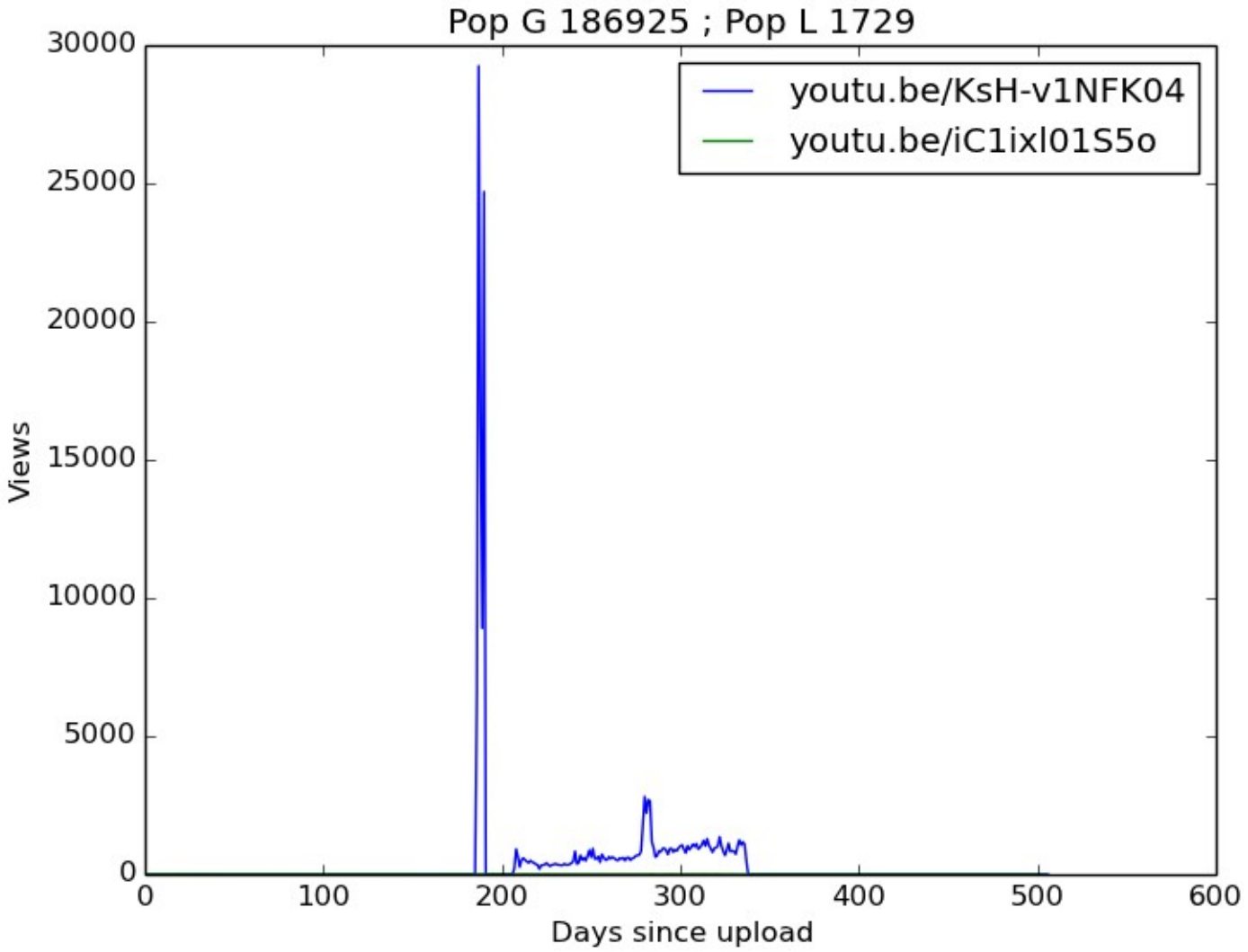
Lets have a look at cases when users disagree!

- **Music**



Lets have a look at cases when users disagree!

- **Music**



Lets have a look at cases when users disagree!

- **Music**