

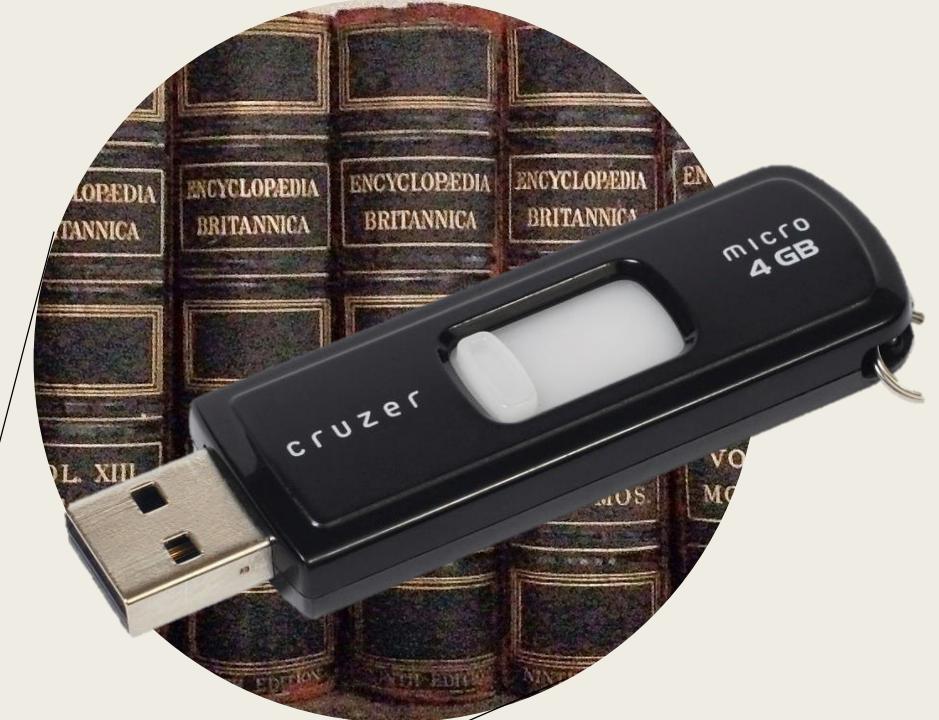
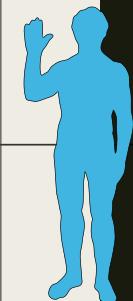


MINIPEDIA

Looking for the meaning
in a haystack.

Inspiration

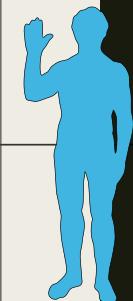
Wikipedia collection as of 2010, in Encyclopedias Britanicas



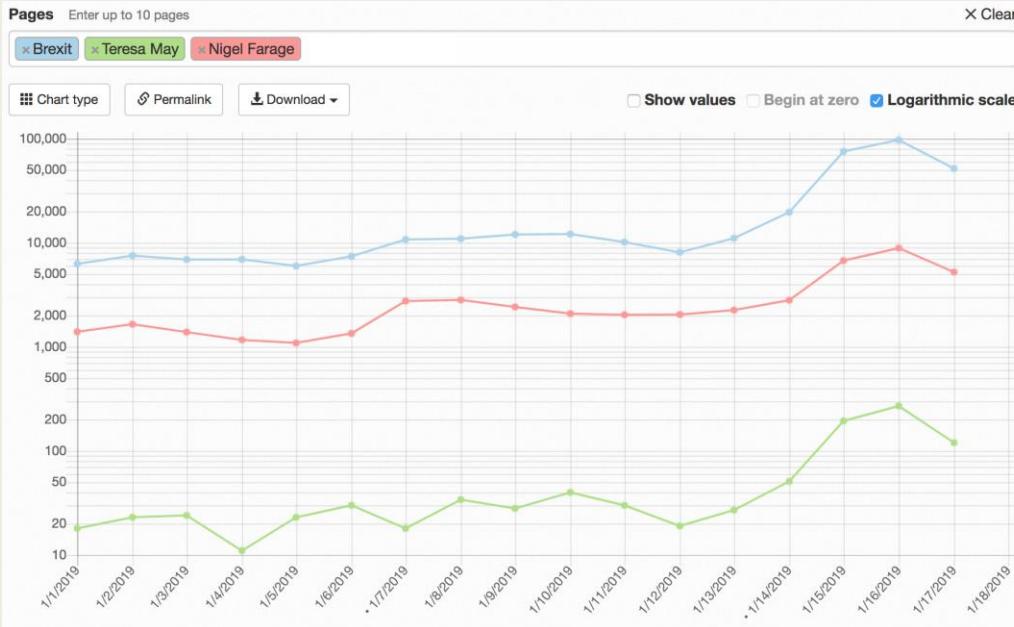
Inspiration



Wikipedia collection as of 2010, in Encyclopedias Britanicas



Knowledge graph building strategies



What are the users
telling us?

Mouvement des Gilets jaunes en France

Pour un article plus général, voir [Mouvement des Gilets jaunes dans le monde](#).



Cet article concerne un événement récent ou en cours. Ces informations peuvent manquer de recul, ne pas prendre en compte les derniers développements ou changer à mesure que l'événement progresse. Le titre lui-même peut être provisoire. N'hésitez pas à l'améliorer en la modifiant.

→ Sur Wikinews, voir aussi : [Catégorie:Mouvement des gilets jaunes](#).

Le mouvement des Gilets jaunes en France est un mouvement social spontané apparu en octobre 2018 et s'étant étendu dans de moindres proportions à d'autres pays. Il trouve son origine dans la diffusion sur les réseaux sociaux d'appels à manifester.

La mobilisation a pour motivation de départ le rejet de l'augmentation de la taxe intérieure de consommation sur les produits énergétiques (TICPE). Elle s'élargit rapidement à d'autres revendications fiscales et sociales (hausse du pouvoir d'achat des classes moyennes et populaires, maintien des services publics, taxation du kérosène et des fioul maritimes, rétablissement de l'ISF...), ainsi que politiques (amélioration de la démocratie représentative, notamment par le référendum d'initiative citoyenne, démission du président de la République, Emmanuel Macron...).

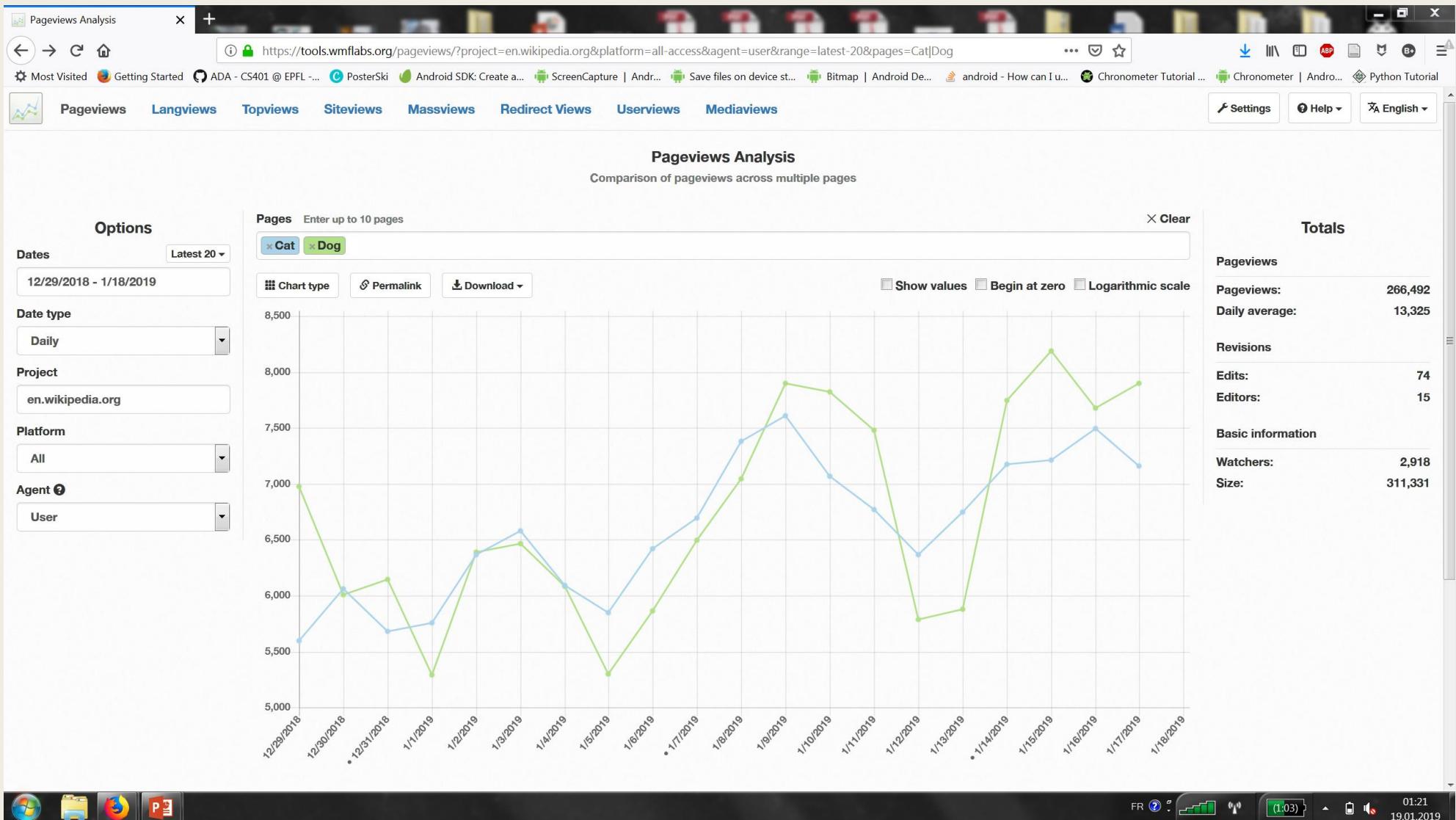
À partir du 17 novembre 2018, le mouvement s'organise autour de blocages de routes et ronds-points et de manifestations nationales chaque samedi. Majoritairement soutenues par l'opinion publique, les protestations trouvent un fort écho dans les zones rurales et périurbaines, mais s'étendent aussi aux grandes villes, où des violences lors des manifestations hebdomadaires provoquent d'importants dégâts matériels et donnent lieu à de nombreuses arrestations. À l'occasion de rassemblements locaux, dix personnes trouvent la mort et plusieurs centaines d'autres sont blessées, plus ou moins grièvement, notamment chez les manifestants.

Initialement hostile à toute revendication du mouvement, le pouvoir exécutif renonce finalement à la hausse de la TICPE et annonce des mesures, qui, jugées insuffisantes par les manifestants, ne mettent pas un terme à la contestation.

Sommaire [masquer]

Do we have to use all
the hyperlinks on a
page ?

Adventures in web scrapping



Adventures in web scrapping

The screenshot captures a multi-tab browser interface focused on 'Pageviews Analysis'. The top tab is 'Pageviews Analysis' and the bottom tab is 'Pageviews'. Both tabs are from the URL <https://tools.wmflabs.org/pageviews/?project=en.wikipedia.org&platform=all-access&agent=user&range=latest-20&pages=Cat|Dog>. The main content area displays a chart titled 'Pageviews Analysis' comparing pageviews across multiple pages for the categories 'Cat' and 'Dog'. The chart shows a peak for 'Cat' around December 29, 2018. The 'Totals' section on the right provides summary statistics: Pageviews: 266,492, Daily average: 13,325, Revisions: 74. The bottom of the screen shows the browser's developer tools, including the DOM tree, style editor, and element inspector.

Pageviews Analysis
Comparison of pageviews across multiple pages

Options

Dates: Latest 20
Pages: Enter up to 10 pages
Categories: Cat, Dog

Totals

Pageviews: 266,492
Daily average: 13,325
Revisions: 74

Pageviews Analysis
Comparison of pageviews across multiple pages

Options

Dates: Latest 20
Pages: Enter up to 10 pages
Categories: Cat, Dog

Totals

Pageviews: 266,492
Daily average: 13,325
Revisions: 74

Pageviews

Pageviews: 266,492
Daily average: 13,325
Revisions: 74

Inspector

DOM Tree:

```
<!--Pageviews Analysis tool-->
<!--Copyright 2016-2018 MusikAnimal-->
<!DOCTYPE html>
<html> [event]
  <head> [event]
    <body class="pageviews">
      <div class="text-center site-notice-wrapper"> [event]
        <nav class="top-nav navbar"> [event]
        <header class="site-header"> [event]
        <aside class="col-lg-2 col-md-2 page-selector"> [event]
          <div class="chart-buttons">
            <button class="btn btn-default btn-sm btn-chart-type chart-specific js-test-change-chart" data-target="#chart-type-modal" data-toggle="modal"> [event]
              [event]
            <a class="btn btn-default btn-sm permalink" href="?project=en.wikipedia.org&platform=all-access&agent=user&start=2018-12-29&end=2019-01-18&pages=Cat%7CDog"> [event]
              [event]
            <input class="permalink-copy">
          </div>
        </aside>
      </div>
      <div class="col-lg-12 data-links">
        <div class="chart-buttons">
          <button class="btn btn-default btn-sm dropdown-toggle" type="button" data-toggle="dropdown" aria-haspopup="true" aria-expanded="false"> [event]
            [event]
        </div>
      </div>
    </body>
  </html>
```

Style Editor:

Computed:

Animations:

Layout:

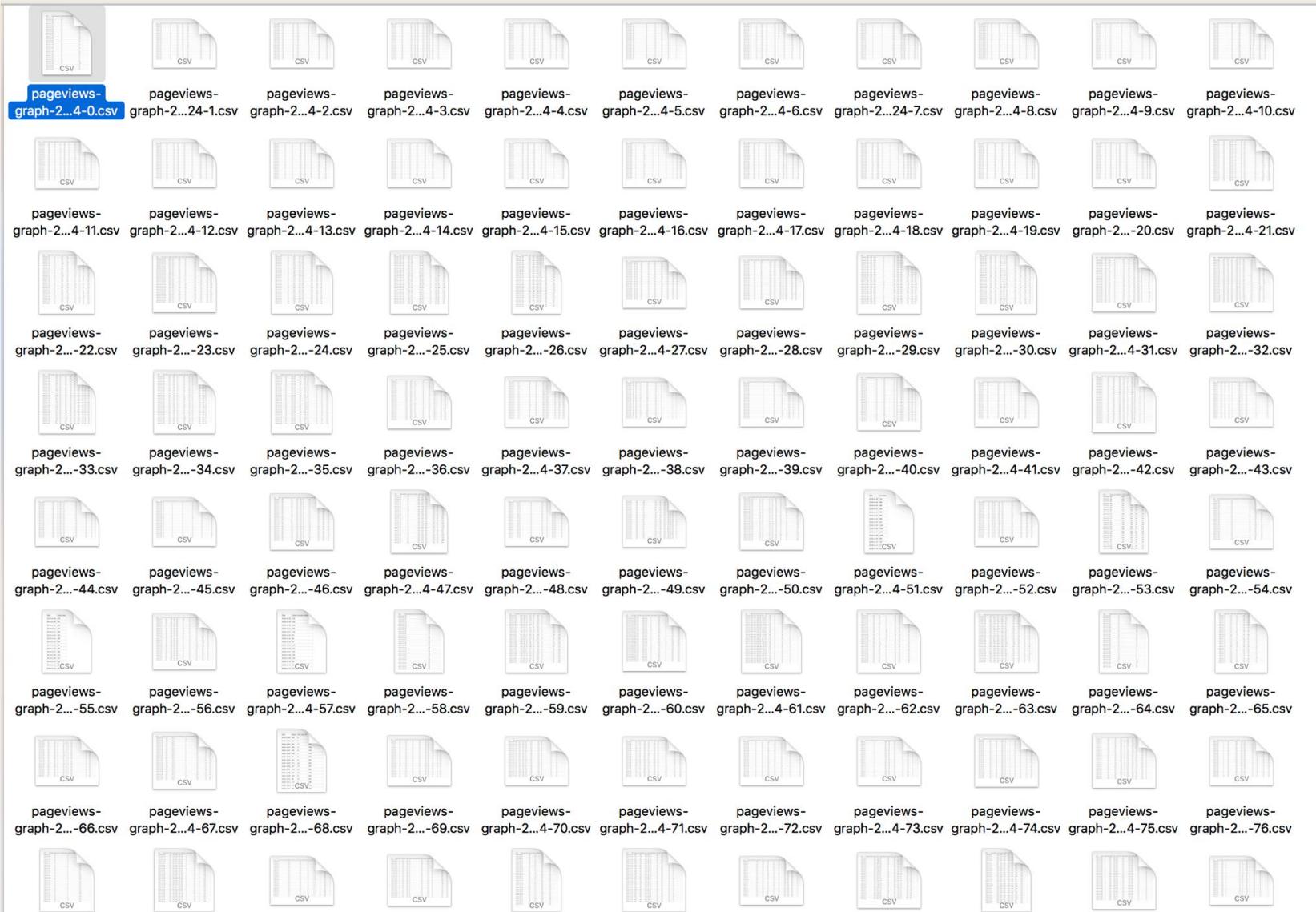
Box Model:

Position: 0
Margin: 0
Border: 0.666667px solid black
Padding: 5px
Width: 84.437px
Height: 18px

Box Model Properties:

border-box
display: block

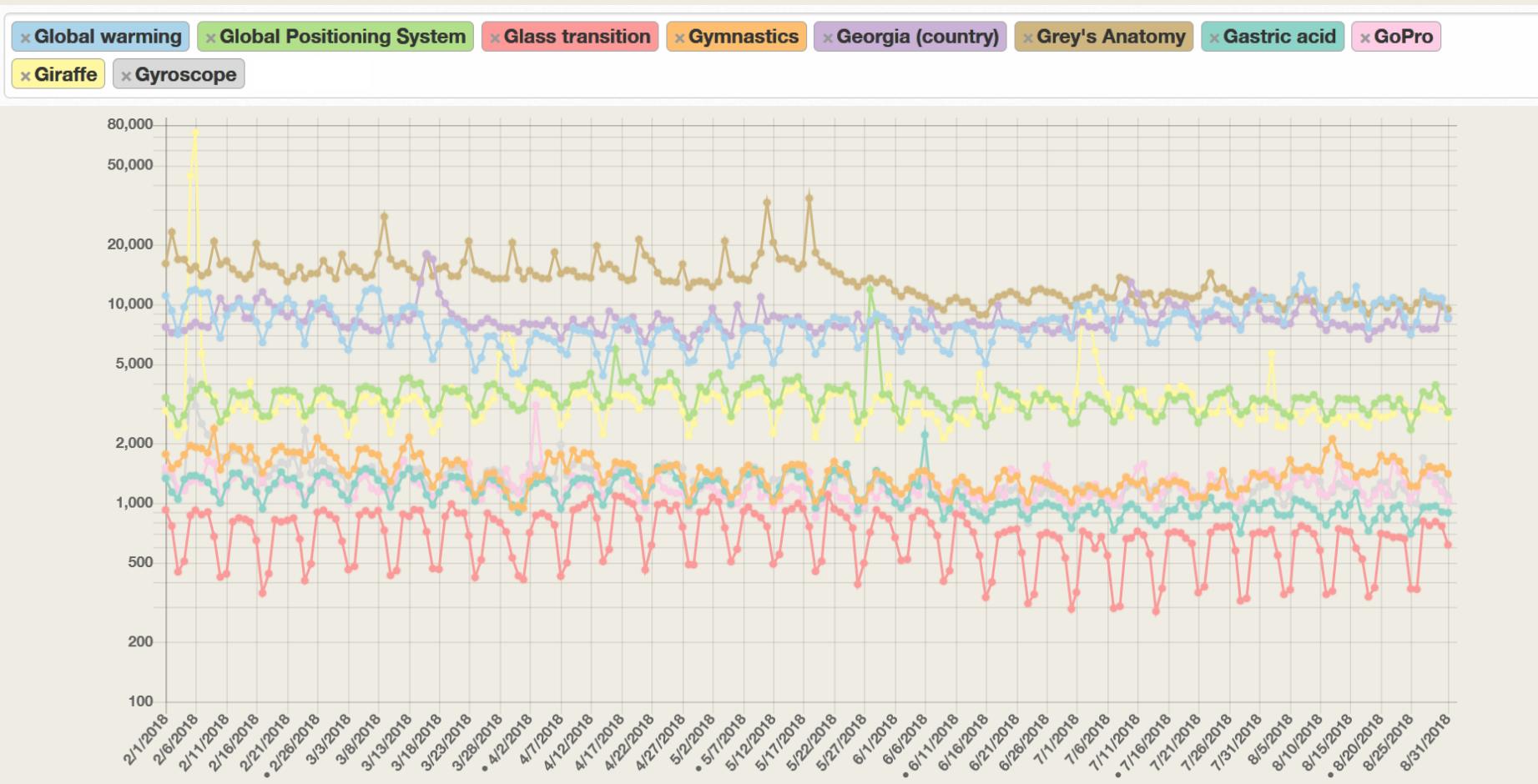
Data



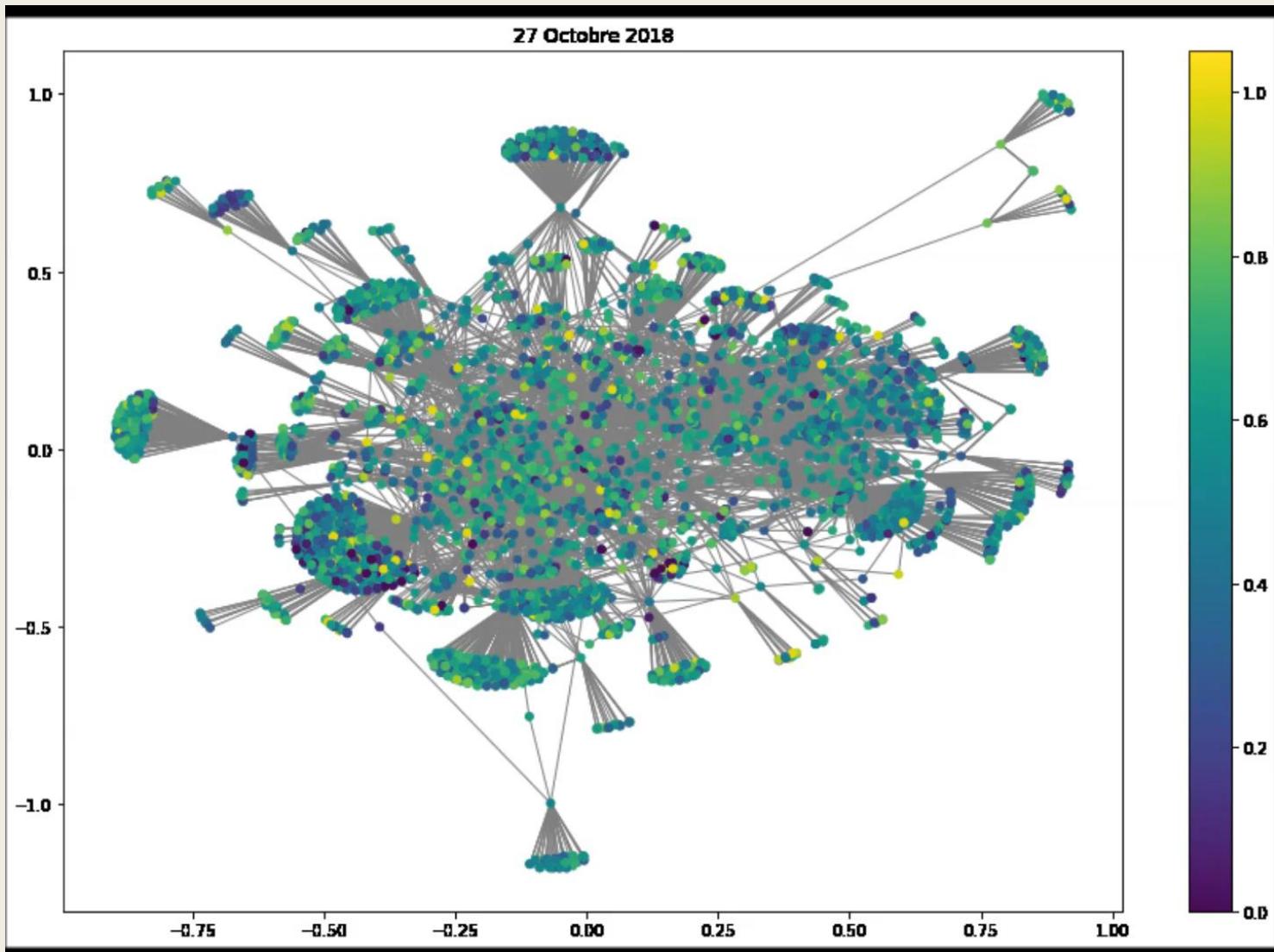
Data

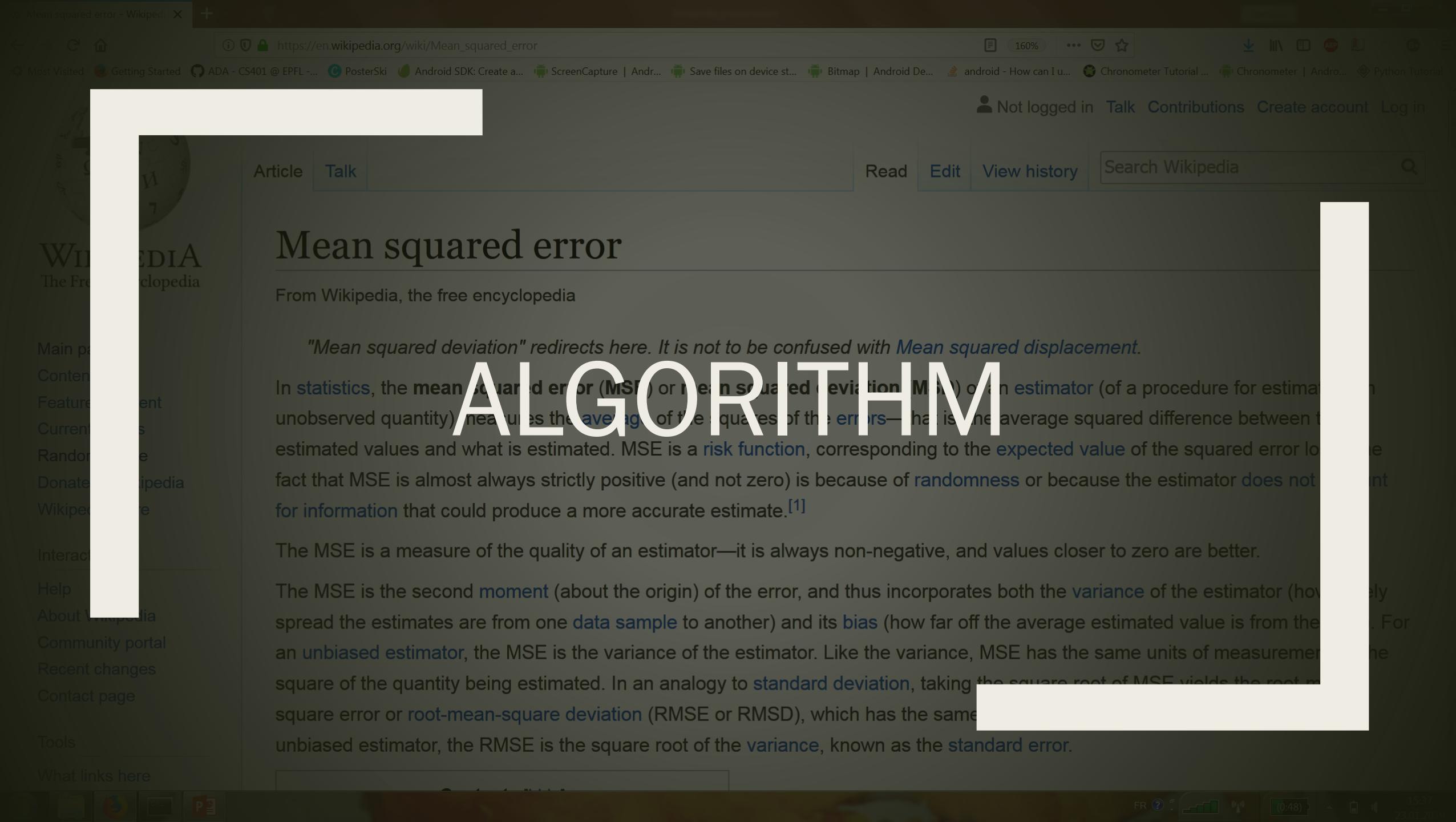
Title	2018-12-28	2018-12-29	2018-12-30	2018-12-31	2019-01-01	2019-01-02	2019-01-03	2019-01-04	2019-01-05	2019-01-06	2019-01-07	2019-01-08	2019-01-09	2019-01-10	2019-01-11	2019-01-12
Donald Trump	41187	37606	37877	42329	44247	44605	48776	49674	47385	45924	47476	47440	64928	57276	53199	53199
Barack Obama	23923	21101	20520	21760	21493	23586	25675	24383	22957	25937	24726	24985	25341	26366	24390	24390
List of Presidents of the United States	21826	21722	20183	18058	18270	19033	20081	22413	21240	21382	21369	22195	23116	22502	22967	22967
New York City	17372	16283	16720	15600	19340	19250	17724	16989	17207	17318	18774	17114	19072	18871	16842	16842
Melania Trump	18418	15078	14294	13428	13160	12243	11647	12787	13021	12123	10692	10609	12943	11987	11603	11603
President of the United States	9036	9165	8635	8489	8571	9029	10003	10924	9791	9782	10479	10388	11348	11066	10385	10385
Republican Party (United States)	6253	7199	5779	5911	6102	7400	10613	13167	9815	8695	8807	9812	10821	9740	9003	9003
Democratic Party (United States)	5712	5699	5554	5987	6386	7447	11919	13063	10226	9374	9285	9311	11056	10046	9856	9856
2016 United States presidential election	5980	5715	5591	6859	6419	13740	9975	10097	8710	8148	8542	8147	9885	10000	9458	9458
Ivanka Trump	6440	5810	5880	5618	5950	5874	5981	5943	7988	7035	5896	5594	6562	7127	7244	7244
Mike Pence	4770	4642	4652	4649	4300	5104	7191	10946	12939	8049	6951	9321	10583	8631	6113	6113
Ivana Trump	6945	5667	5582	5414	5247	5689	5448	5691	6480	6325	5412	5268	6784	6757	6231	6231
Fred Trump	5949	5304	4720	4244	4544	4976	4635	4543	4557	4639	4809	4751	8386	7693	6920	6920
Marla Maples	4774	4381	4254	4161	4155	4473	4224	4675	6678	5880	4714	4192	5541	5430	4892	4892
Tiffany Trump	4479	4027	3956	3899	4008	4115	4308	4374	8216	6413	5074	4300	4765	5615	4688	4688
Donald Trump Jr.	4917	3815	3943	5398	6042	4552	4411	4025	4331	4484	4320	5079	5594	6073	4687	4687
Presidency of Donald Trump	2329	2129	2268	2606	3896	3139	3990	4894	3469	3478	5586	4194	5023	5420	5636	5636
Vice President of the United States	2658	2434	2338	2299	2402	3078	3955	4248	3885	3448	4209	3850	4002	3630	3292	3292
Economics	2726	2762	2551	2647	2735	4164	3539	3394	3095	3274	4148	4721	4802	4867	4027	4027
Family of Donald Trump	3706	3357	3260	3153	3548	3372	3324	3384	3896	3871	3486	3215	4416	4221	4092	4092
White House	3281	3154	3062	2853	2902	3760	3406	3910	3486	3457	3560	3685	4764	3986	3643	3643
Barron Trump	3478	3074	2936	2727	2734	2828	2865	2931	3250	3237	2778	2708	3413	3210	3026	3026
Eric Trump	2801	2531	2584	2897	2948	2706	2731	2707	3291	4271	3172	2840	3119	3257	3031	3031
Mary Anne MacLeod Trump	3116	2908	2526	2182	2307	2166	2270	2472	2460	2402	2175	2245	6697	3649	3005	3005
Incumbent	1780	1619	1661	1642	1870	2138	2479	2843	2469	2359	2510	2414	2657	2618	2558	2558
Cabinet of Donald Trump	1889	1444	1561	1695	1741	2753	3371	3171	2188	2299	2626	2584	2762	2697	2391	2391
Bachelor of Science	1719	1648	1534	1607	1638	2277	2439	2288	2170	2110	2516	2715	2824	2742	2508	2508
Mar-a-Lago	3092	3425	2532	2609	2773	2690	2246	1994	2040	2084	1869	1666	2201	2191	1955	1955
Independent politician	1686	1672	1773	1768	1740	2112	2589	2688	2184	2006	2134	2172	2801	2420	2166	2166

Page view signals



Page view signals (movie)

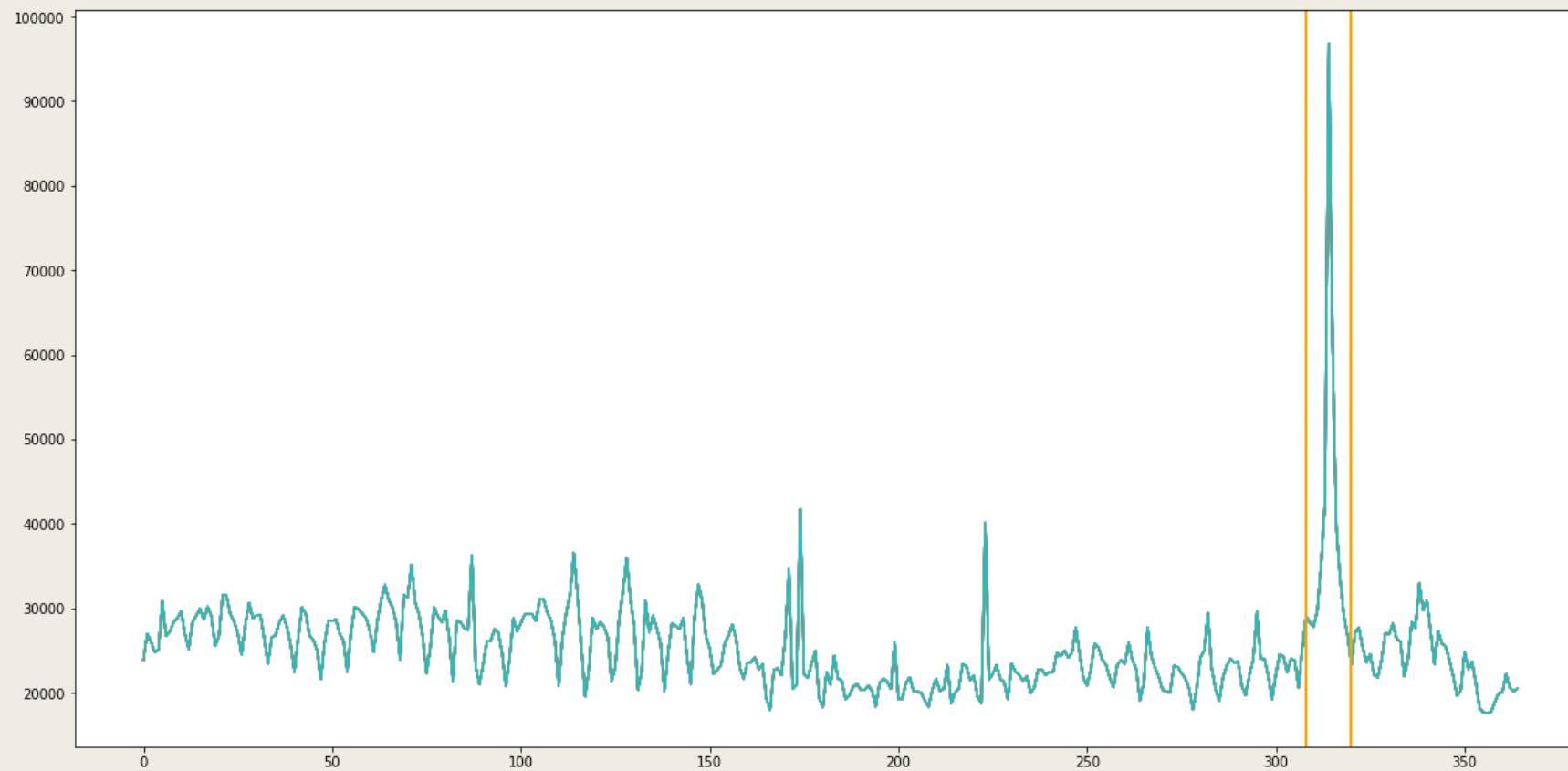




Algorithm

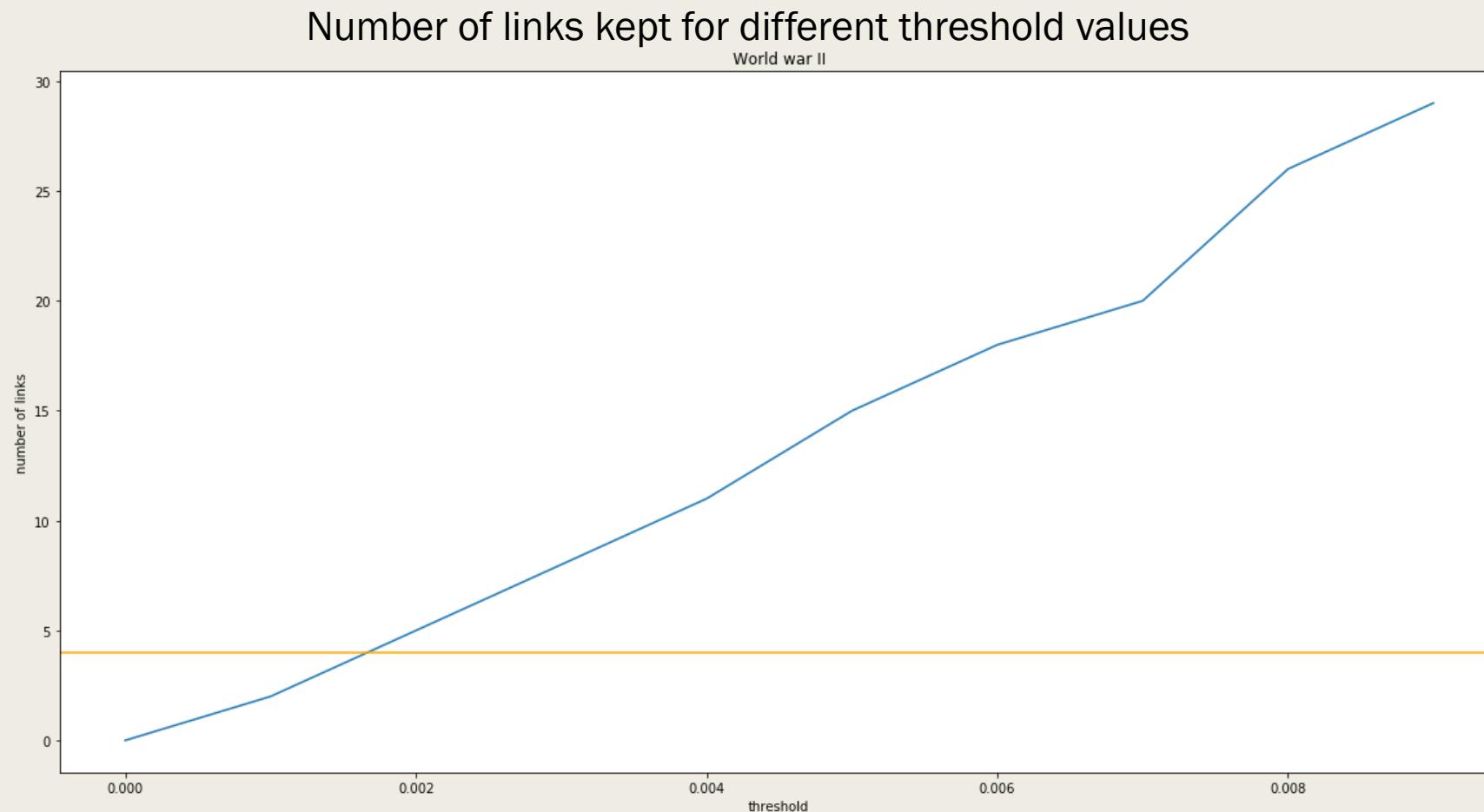
Step 1: Find peak activity

“WWII” Wikipedia page daily views for 2018



Algorithm

Step 2: Select most important nodes



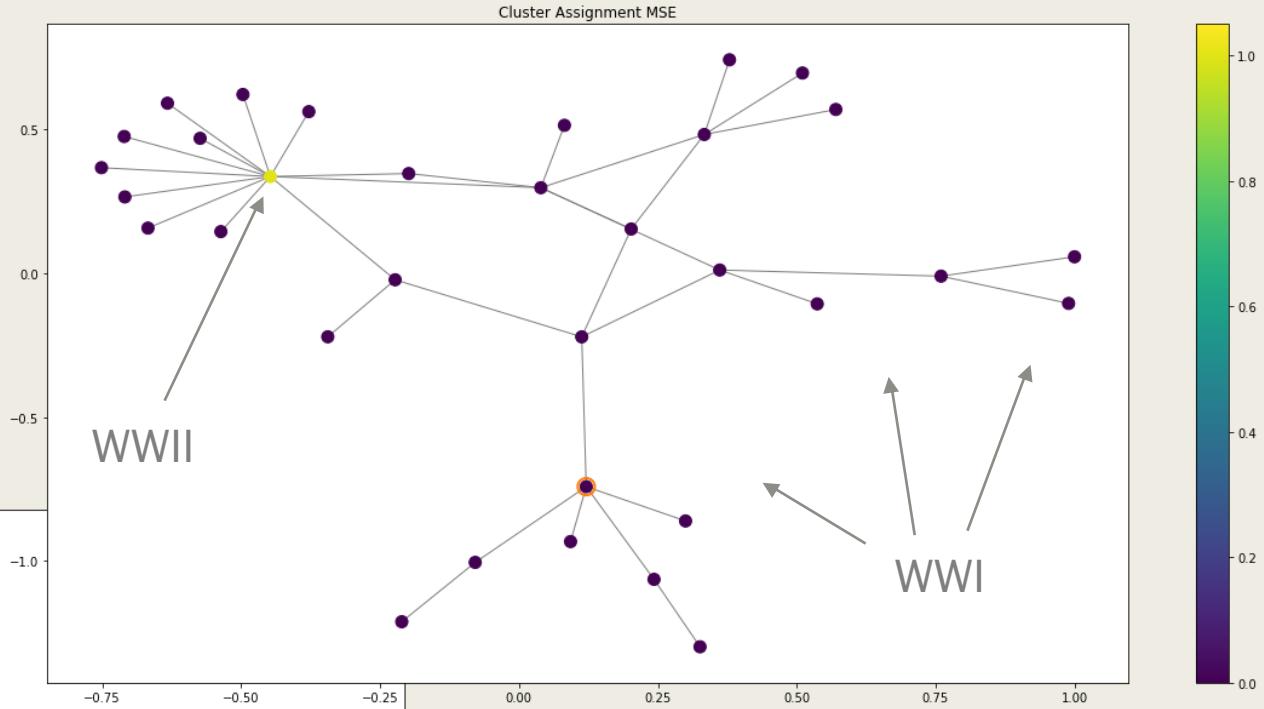
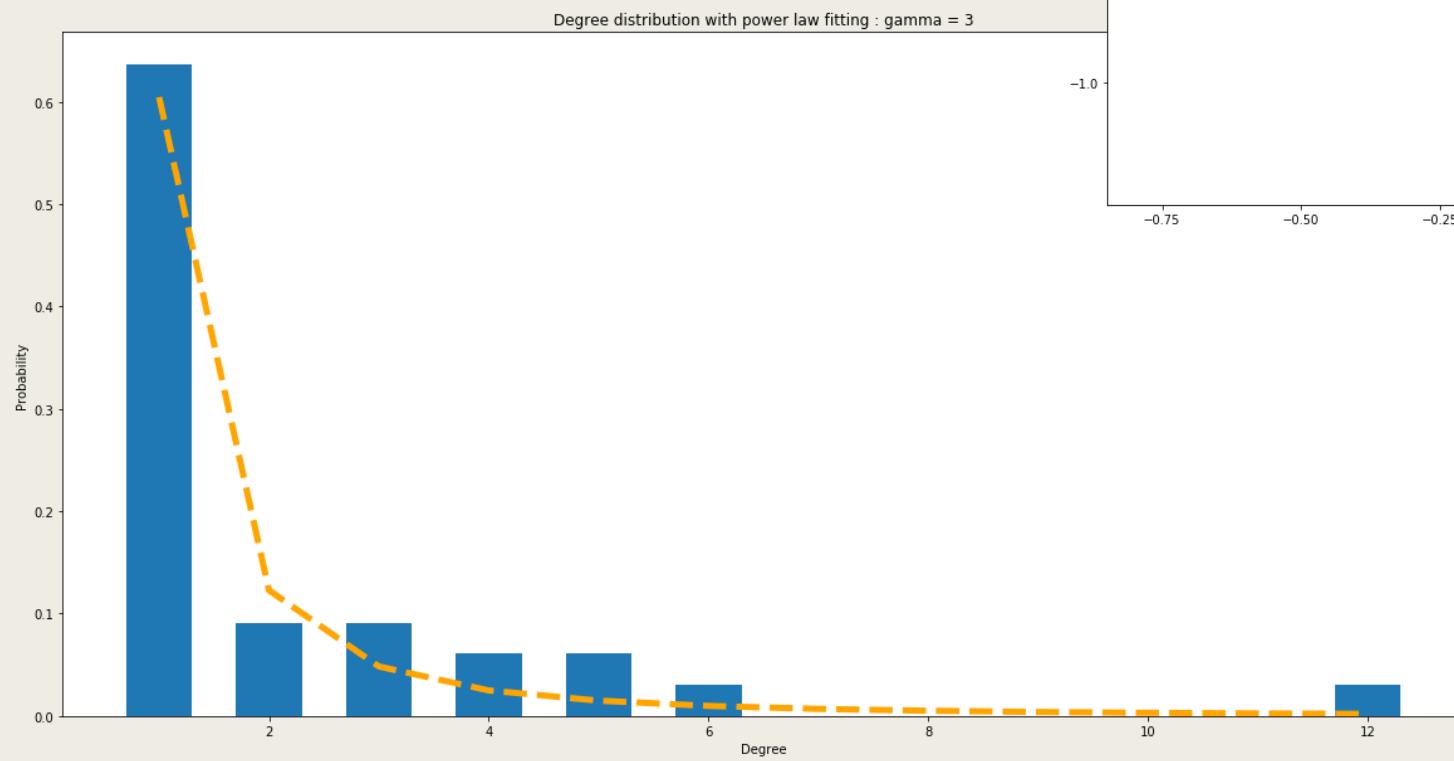
Algorithm

Step 3: Repeat for each link kept until desired length is reached

“WWII” fifth most visited Wikipedia article

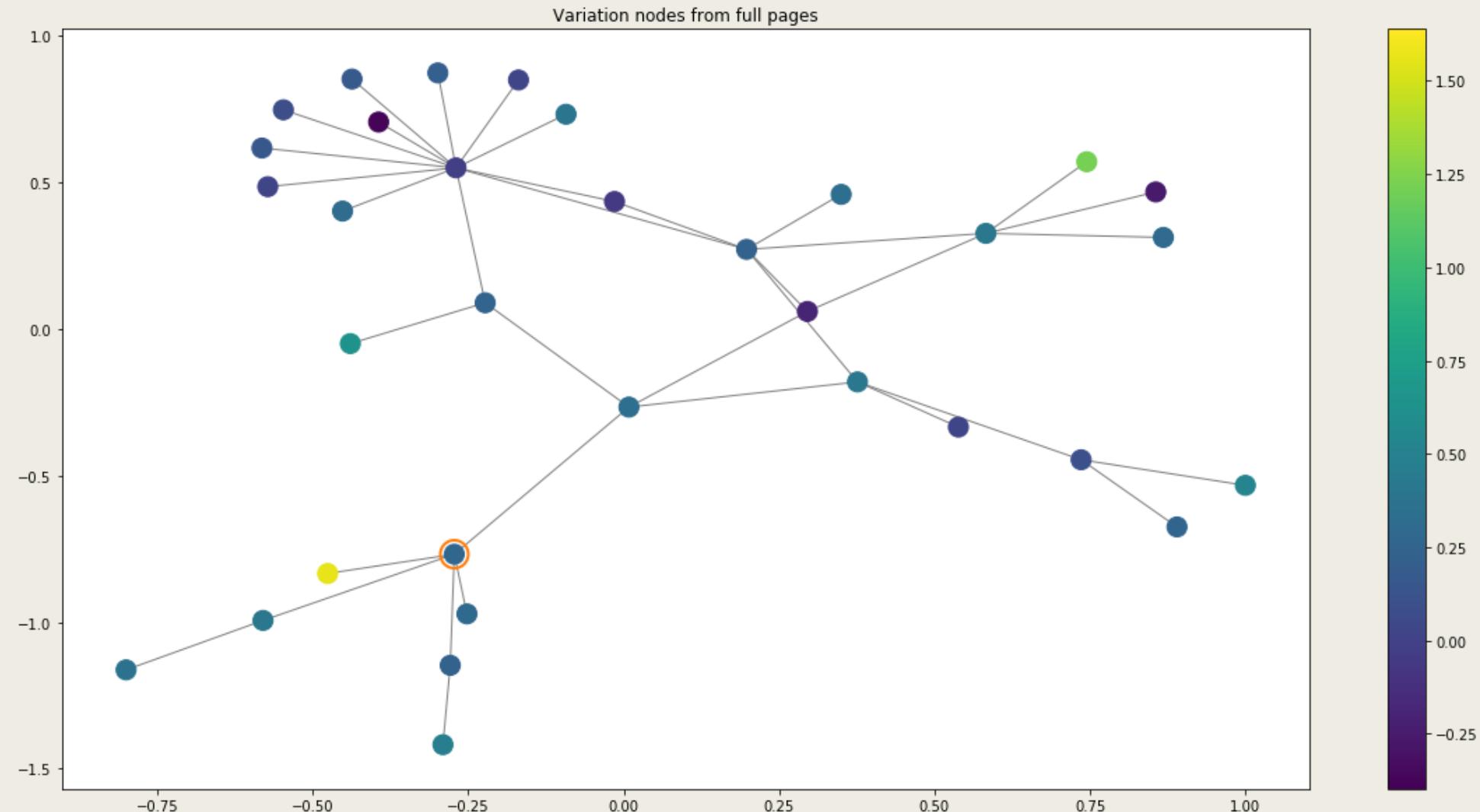


Network characteristics



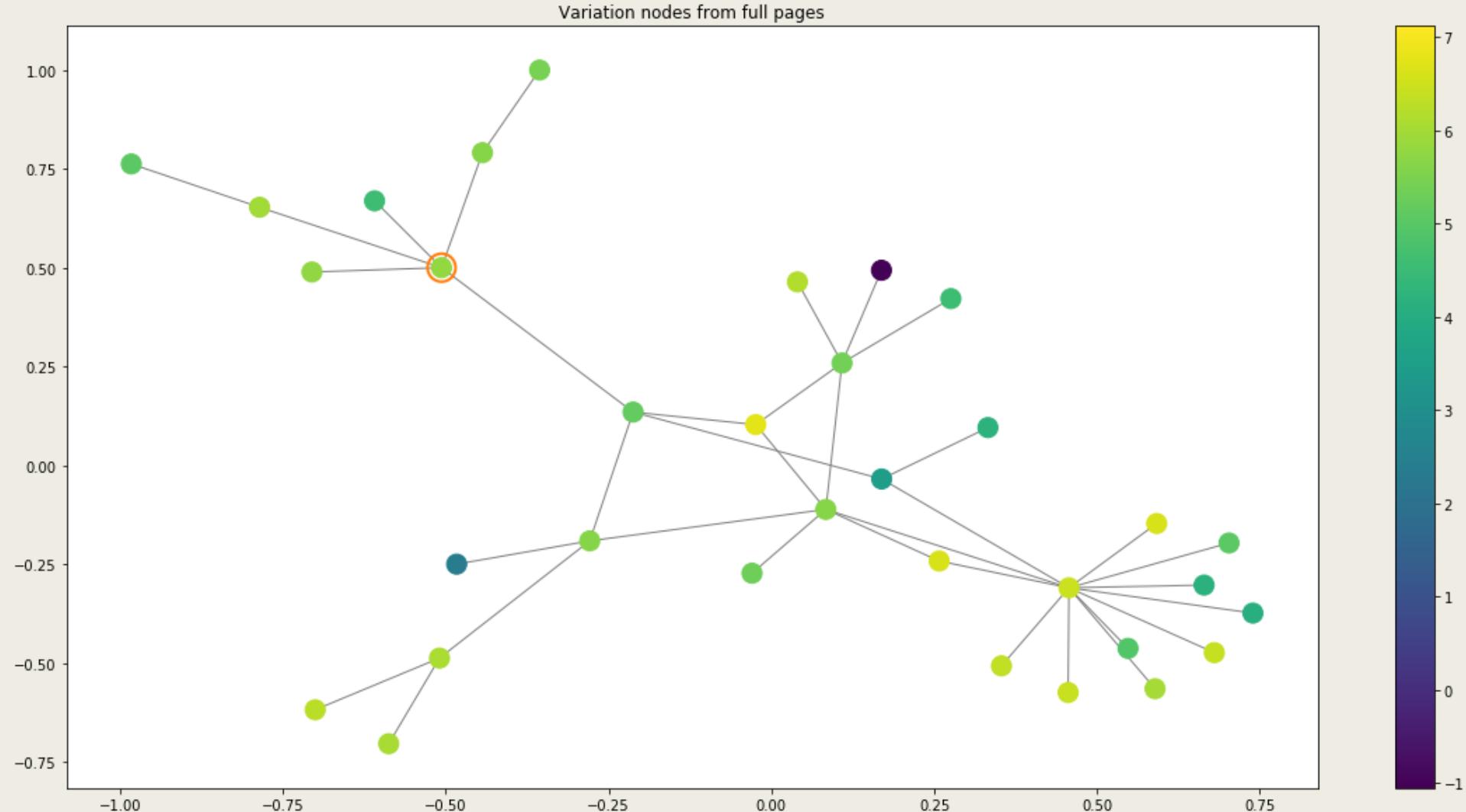
Evaluation

Drawback: computing time
~1h for 35 nodes and 36 edges



Evaluation

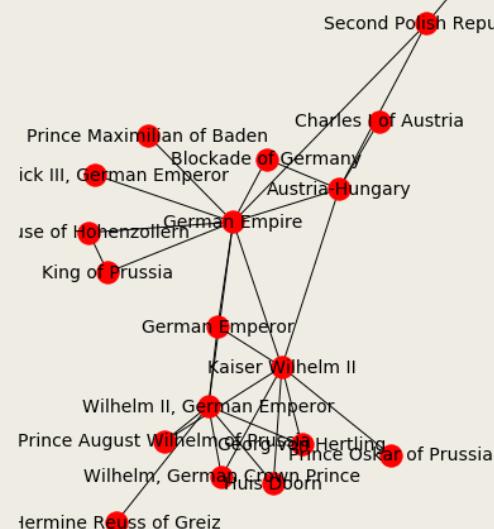
Drawback: computing time
~1h for 35 nodes and 36 edges



Restricted algorithm

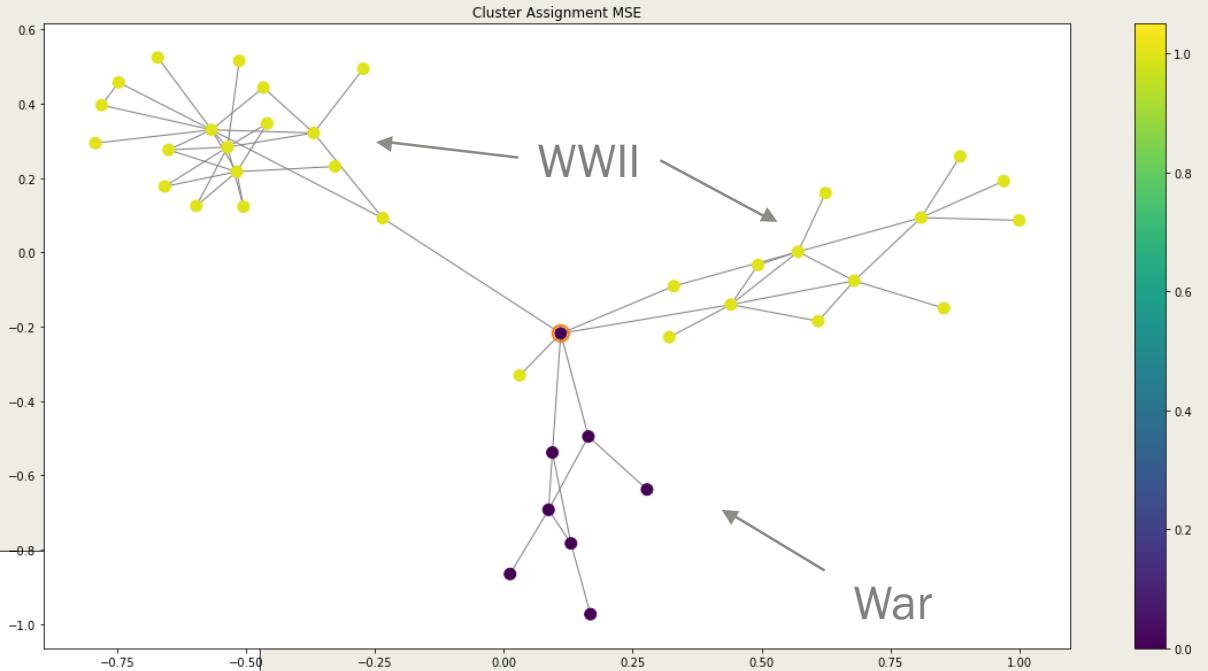
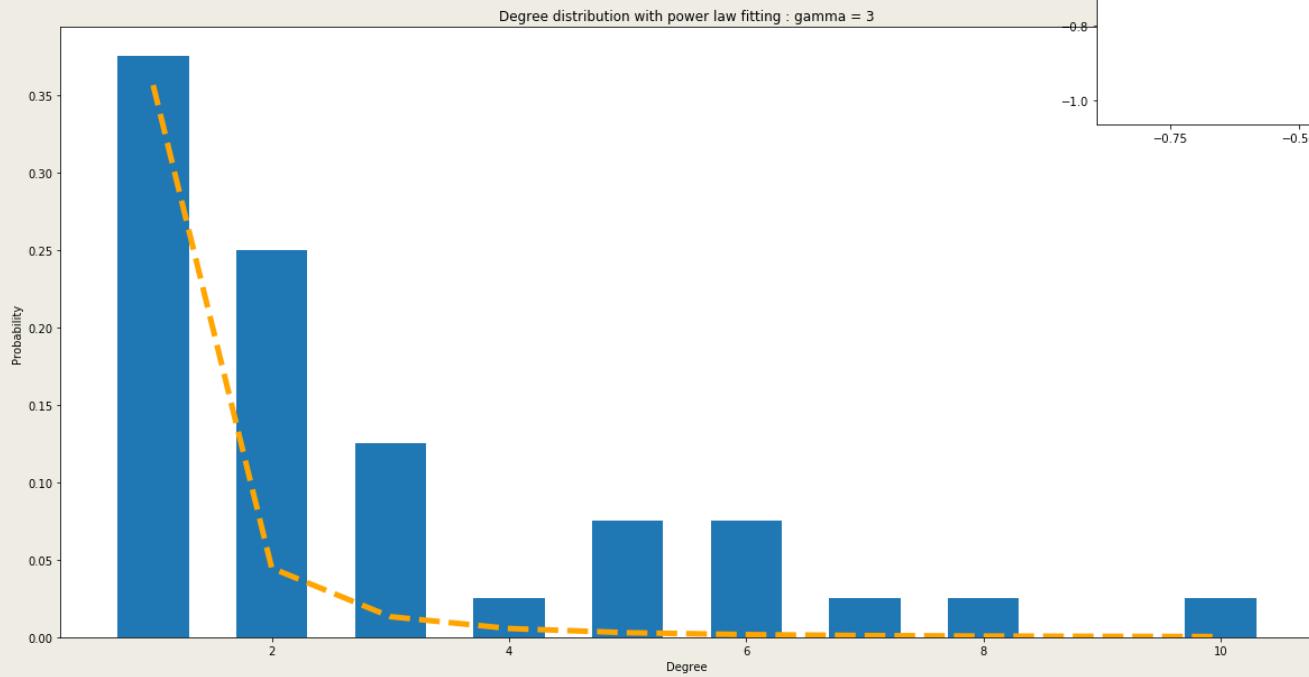
Hypothesis: page structure will influence the user

Improvement: computing time
~10min for 40 nodes and 56 edges



Graph of categories rather than individual articles

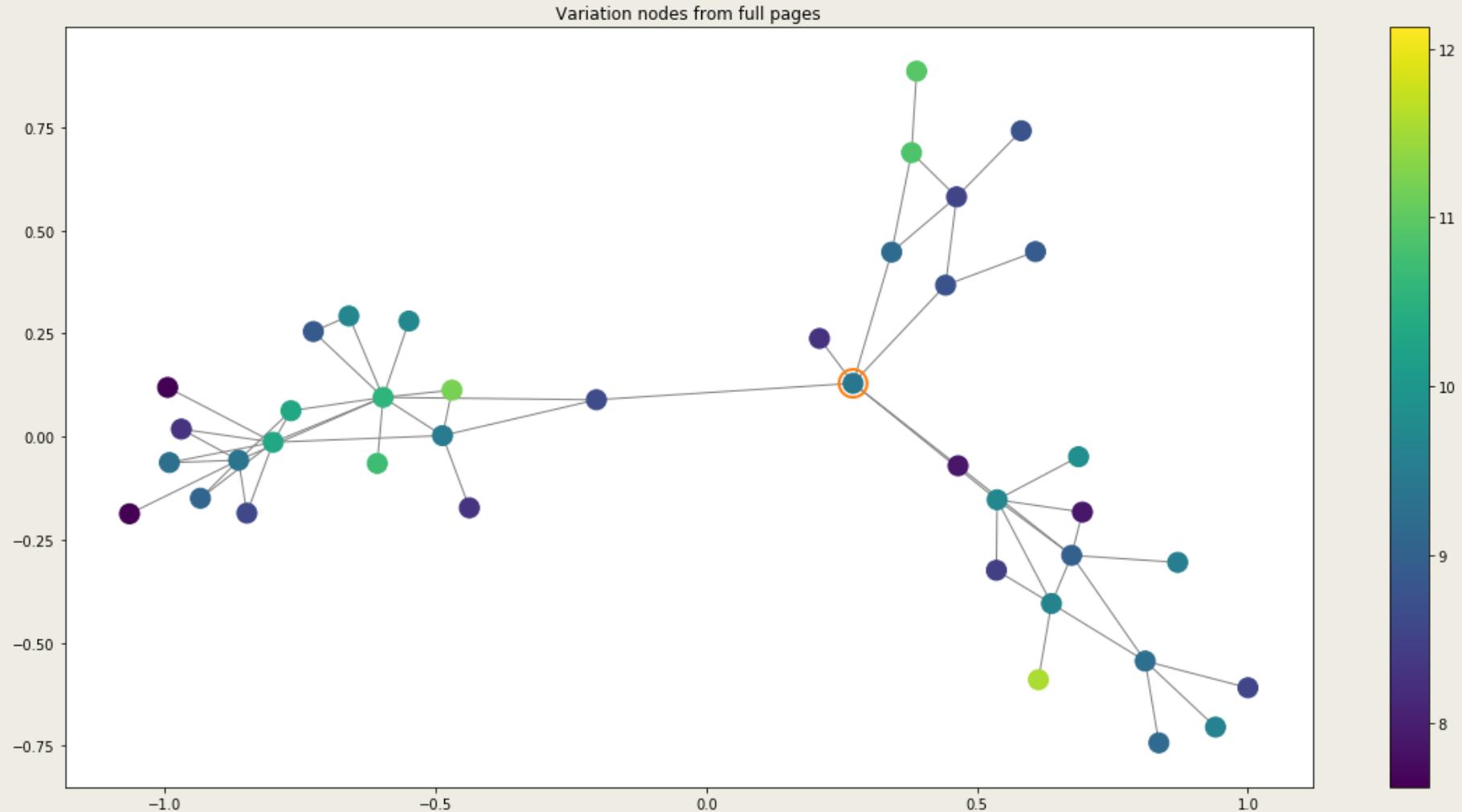
Network characteristics



We had to increase
the threshold to have a
comparable graph size

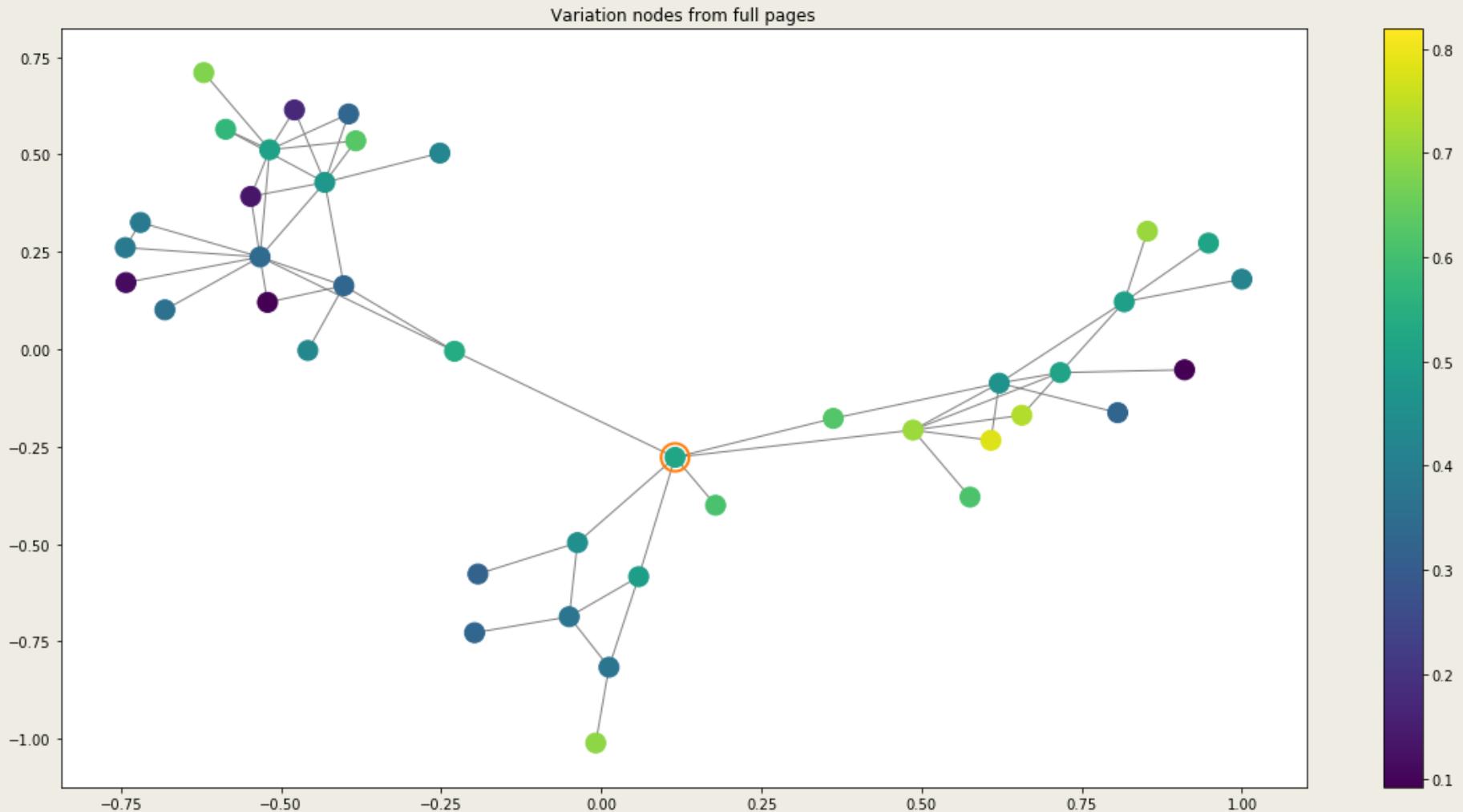
Evaluation

```
for day in range(len(signal)):  
    Var_signal[day,:] = (signal[day,:]-np.mean(signal,0))/np.std(signal,0)
```

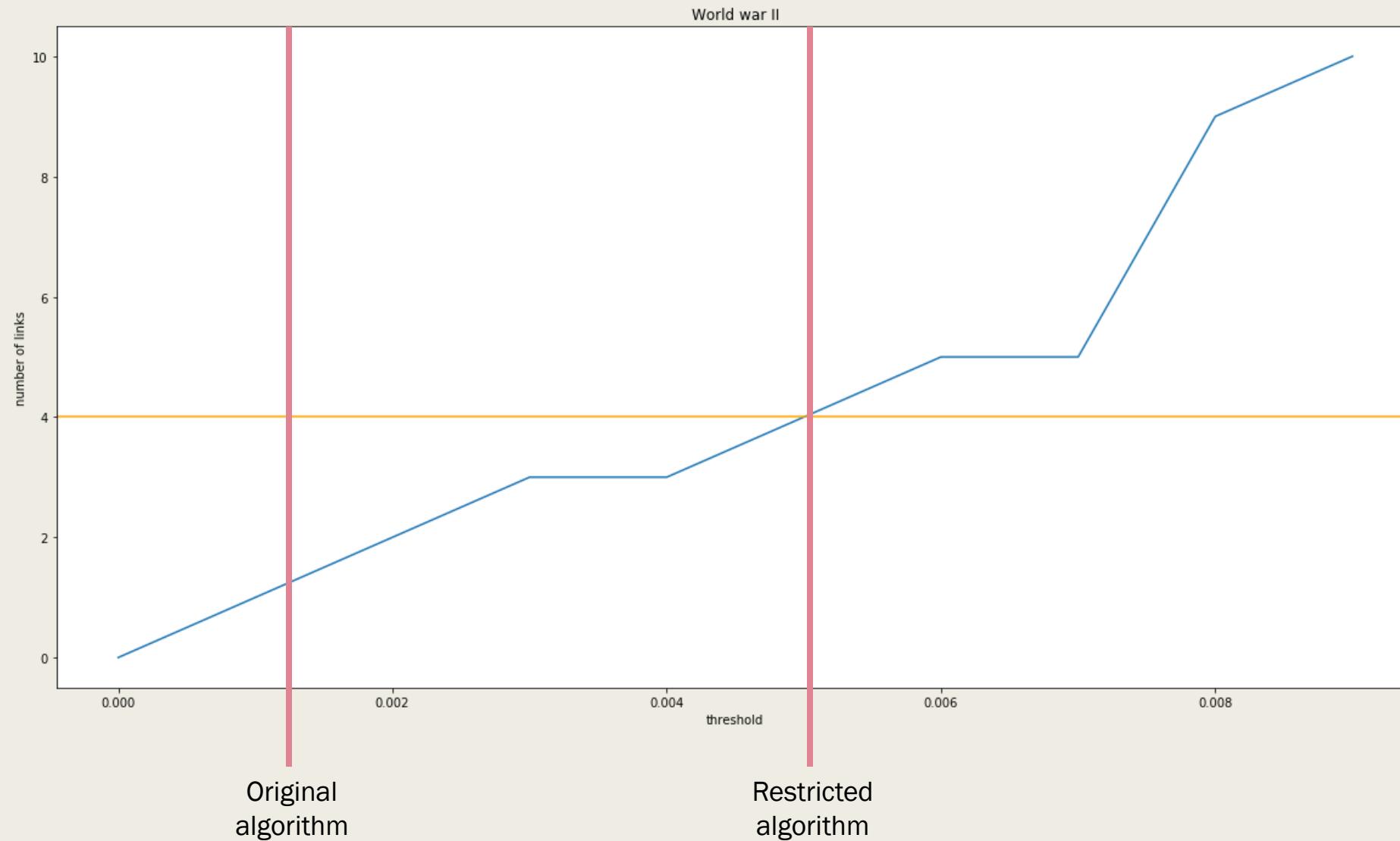


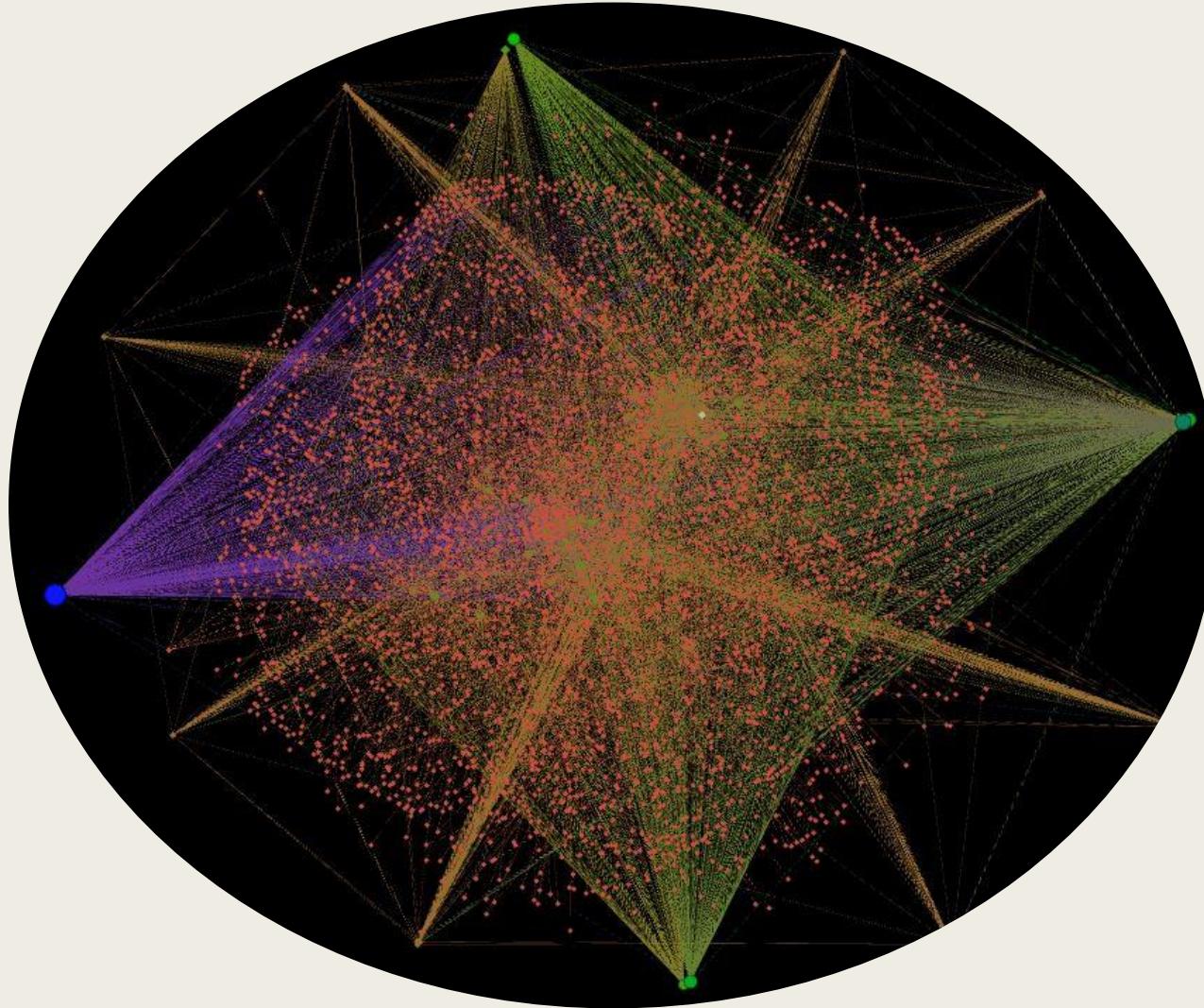
Evaluation

```
for day in range(len(signal)):  
    Var_signal[day,:] = (signal[day,:]-np.mean(signal,0))/np.std(signal,0)
```



Evaluation





Conclusion