* Dredze, M., Cheng, R., Paul, M.J. and Broniatowski, D., 2014, June. HealthTweets. org: a platform for public health surveillance using Twitter. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.
* Lampos, V., Moura, S., Yom-Tov, E., Cox, I.J., McKendry, R. and Edelstein, M., 2020. Tracking COVID-19 using online search. *arXiv preprint arXiv:2003.08086*.
* Liautaud, P., Huybers, P. and Santillana, M., 2020. Fever and mobility data indicate social distancing has reduced incidence of communicable disease in the United States. *arXiv preprint arXiv:2004.09911*.
* Liu, D., Clemente, L., Poirier, C., Ding, X., Chinazzi, M., Davis, J.T., Vespignani, A. and Santillana, M., 2020. A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models. *arXiv preprint arXiv:2004.04019*.
* Lu, F.S., Hattab, M.W., Clemente, C.L., Biggerstaff, M. and Santillana, M., 2019. Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches. *Nature communications*, *10*(1), pp.1-10.
* Menkir, T.F., Chin, T., Lipsitch, M., Niehus, R., Hay, J.A., Surface, E., de Salazar, P.M., Buckee, C., Mina, M.J., Khan, K. and Watts, A., 2020. Estimating the number of undetected COVID-19 cases exported internationally from all of China. *medRxiv*. <https://www.medrxiv.org/content/10.1101/2020.03.23.20038331v2.full.pdf>

2018 US census: SOURCE??

Death and number cases, John Hopkins University: <https://raw.githubusercontent.com/CSSEGISandData/COVID-19/>

Google Mobility Data: [htts://www.google.com/covid19/mobility/](https://www.google.com/covid19/mobility/)

[p](https://www.google.com/covid19/mobility/)

High Dimension Clustering w/ t-SNE & DBSCAN, available at: <https://towardsdatascience.com/high-dimension-clustering-w-t-sne-dbscan-dcec77e6a39b>