DevFest Hackathon Notes

## The interests of Judges:

* Orphans and children: Derek
* Domestic violence: Wangzhi (George)
* Human trafficking: Yuki
* Search other ideas: Amaury

## Notes

### Orphans and Children

### Domestic Violence

What is Domestic Violence:

<https://assets.speakcdn.com/assets/2497/domestic_violence2.pdf>

Domestic violence is the willful intimidation, physical assault, battery, sexual assault, and/or other abusive behavior as part of a systematic pattern of power and control perpetrated by one intimate partner against another. It includes physical violence, sexual violence, threats, and emotional/psychological abuse.

1. **Family Violence Related Snapshots: New York City Community Board Districts Data set (Col:8, Row:138)**

[**https://data.cityofnewyork.us/Public-Safety/Family-Violence-Related-Snapshots-New-York-City-Co/a35y-93e7**](https://data.cityofnewyork.us/Public-Safety/Family-Violence-Related-Snapshots-New-York-City-Co/a35y-93e7)

1. **Data Collection: National Crime Victimization Survey (NCVS)**

[**https://www.icpsr.umich.edu/icpsrweb/NACJD/series/95**](https://www.icpsr.umich.edu/icpsrweb/NACJD/series/95)

### Human Trafficking

**-**Trafficking involves transporting someone into a situation of exploitation. This can include forced labor, marriage, prostitution, and organ removal. This kind of exploitation is known by a few different names -- “human trafficking,” “trafficking of persons,” and “modern slavery” are the ones accepted by the US Department of State.

-internationally there are between 20 million and 40 million people in modern slavery today

- Estimates suggest that, internationally, only about .04% survivors of human trafficking cases are identified, meaning that the vast majority of cases of human trafficking go undetected

- Globally, an estimated 71% of enslaved people are women and girls, while men and boys account for 29%

- Estimates suggest that about 50,000 people are trafficked into the US each year, most often from Mexico and the Philippines

-Advocates report a growing trend of traffickers using online social media platforms to recruit and advertise targets of human trafficking : "2019 Trafficking in Persons Report.” United States Department of State. Accessed July 31, 2019. https://www.state.gov/wp-content/uploads/2019/06/2019-Trafficking-in-Persons-Report.pdf. [↩︎](https://www.dosomething.org/us/facts/11-facts-about-human-trafficking#fnref9)

-In 2018, The National Human Trafficking Hotline received more calls from California than any other state in the US, followed by Texas and Florida, respectively.

- Statistics provided by the National Human Trafficking Hotline: <https://humantraffickinghotline.org/sites/default/files/Polaris_National_Hotline_2018_Statistics_Fact_Sheet.pdf>

- The database contains over 50,000 individual cases, with approximately 5,000 new cases added each year. IOM currently assists between 7,000 and 9,000 victims annually, collecting a unique source of data on victims of trafficking that is international in scope. Data captured include information about the victims’ backgrounds, trafficking locations and routes, how people fall into the trafficking process, associated forms of exploitation and abuse, sectors of exploitation, means by which victims are controlled, and some information on perpetrators.

: <https://www.ctdatacollaborative.org/download-global-dataset>

The above data is 13MB with

### Other topics:

1. World poverty
   1. <https://www.kaggle.com/theworldbank/poverty-and-equity-database>
   2. <https://www.kaggle.com/ophi/mpi>
2. Australian Fire
3. Coronavirus Prediction
   1. <https://www.kaggle.com/brendaso/2019-coronavirus-dataset-01212020-01262020>
4. Safety of Students - shooting, crime in NYC
5. Helicopter Safety

### Choices

1. Domestic Violence
2. Human Trafficking o
3. World poverty $
4. Australian fire o
5. **Coronavirus Prediction o$oo**
6. Student Safety o

## About Coronavirus Prediction

### About the dataset

* Kaggle Dataset (Update):

<https://www.kaggle.com/de5d5fe61fcaa6ad7a66/coronavirus-dataset-update-0206>

* Google Drive : <https://drive.google.com/drive/folders/1nzETKi4oOlJWenAI4KoPcPWHz1K4Bs87?usp=sharing>
* WHO 2019nCoV statistics: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>
* Government 2019n-CoV Data visualization

<http://2019ncov.chinacdc.cn/2019-nCoV/>

* Baidu Index

<http://index.baidu.com/v2/main/index.html#/trend/%E6%96%B0%E5%9E%8B%E5%86%A0%E7%8A%B6%E7%97%85%E6%AF%92?words=%E6%96%B0%E5%9E%8B%E5%86%A0%E7%8A%B6%E7%97%85%E6%AF%92>

* WHO SARS statistics

<https://www.who.int/csr/sars/country/en/>

* Kaggle-SARS (total)

<https://www.kaggle.com/zhongtr0n/sars-who-data>

### Other infectious diseases

### 2019 n-COV

SARS:

1. <https://www.kaggle.com/zhongtr0n/sars-who-data>
2. <https://www.who.int/csr/sars/archive/en/>
3. <https://www.kaggle.com/nattay/who-sars-cumulative-reported-cases#2003_04_04.csv>

### Objective: How will the Corona Virus spread?

* Relate the spread of Corona Virus with other infectious diseases
  + Rate of increase (Infections/Day, Infection percentage of population/Day)
    - Can be used for predicting using time series regression
  + What kind of people are vulnerable to Corona Virus?
    - Which people have higher death rate or higher infection rate
  + Which countries will most likely be affected?
    - Prediction using past data
  + How they spread (Human contact, transportation, through air, exported goods, pests, etc)

Kernels:

1.data visualization:

1.1 Geographical propagation analysis:

<https://www.kaggle.com/grebublin/coronavirus-propagation-visualization-forecast>

<https://www.kaggle.com/parulpandey/wuhan-coronavirus-a-geographical-analysis>

1.2 general visualization:

<https://www.kaggle.com/therealcyberlord/corona-virus-visualization>

<https://www.kaggle.com/vikassingh1996/coronavirus-an-exploratory-study-w-detail-report> (including geographical propagation, case study, and bar chart analysis)

1.3 special region analysis:

<https://www.kaggle.com/gatunnopvp/coronavirus-growth-in-2019-2020-updated-02-07-20> (focus on a particular region)

ANALYSIS ON PLOTS

1. World map
2. SARS vs. 2019-nCoV
3. Search Trend

References:

1.Barnaghi, Peiman, Parsa Ghaffari, and John G. Breslin. "Opinion mining and sentiment polarity on twitter and correlation between events and sentiment." 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService). IEEE, 2016.Limited # in Tweets that can be collected

2.https://www.nytimes.com/2020/02/09/world/asia/coronavirus-china.htmlAmaury lost his voice

3.https://zhuanlan.zhihu.com/p/70153413

4.https://zhuanlan.zhihu.com/p/30061051

5.https://www.rollingstone.com/culture/culture-news/coronavirus-wuhan-china-virus-quarantine-death-toll-941417/

6.https://news.163.com/special/epidemic/

7.https://towardsdatascience.com/twitter-sentiment-analysis-classification-using-nltk-python-fa912578614c

8.Ji, Xiang, Soon Ae Chun, and James Geller. "Monitoring public health concerns using twitter sentiment classifications." 2013 IEEE International Conference on Healthcare Informatics. IEEE, 2013.

9.Pollacci, Laura, et al. "Sentiment spreading: an epidemic model for lexicon-based sentiment analysis on twitter." Conference of the Italian Association for Artificial Intelligence. Springer, Cham, 2017.

10.https://github.com/jonbakerfish/TweetScraper