

# Global News Feeds

---

Sunday 21/11/2021

Newsletter

Edition: 001

By: Samaher AL-Hadi, Asrar Abdulrahman



# Introduction

---

Sunday 21/11/2021

Newsletter

Edition: 001

Global News is a collection of the most current news around the world, include news websites, government agencies, tech journals, company websites, etc. In this project we aim to build a model that predict the most topic published over the news.



# Problem statement

---

Sunday 21/11/2021

Newsletter

Edition: 001



Given a global news in different languages , we are challenged to discover the topics that occur in a collections of articles.

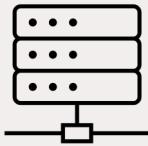
# DATA

---

Sunday 21/11/2021

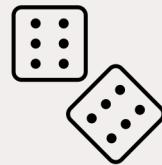
Newsletter

Edition: 001



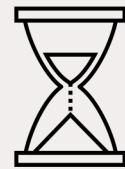
## DATASETS

Global News Feeds  
&  
Language Identification  
dataset



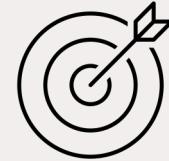
## SIZE

3,305,325 rows  
&  
22,000 rows



## TIME

The two dataset  
between at  
2017 - 2018



## SOURCE

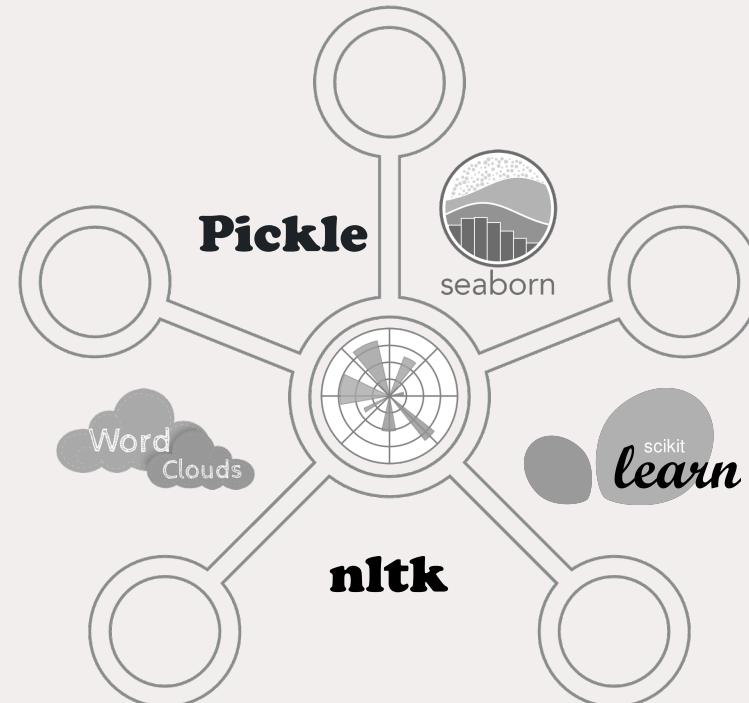
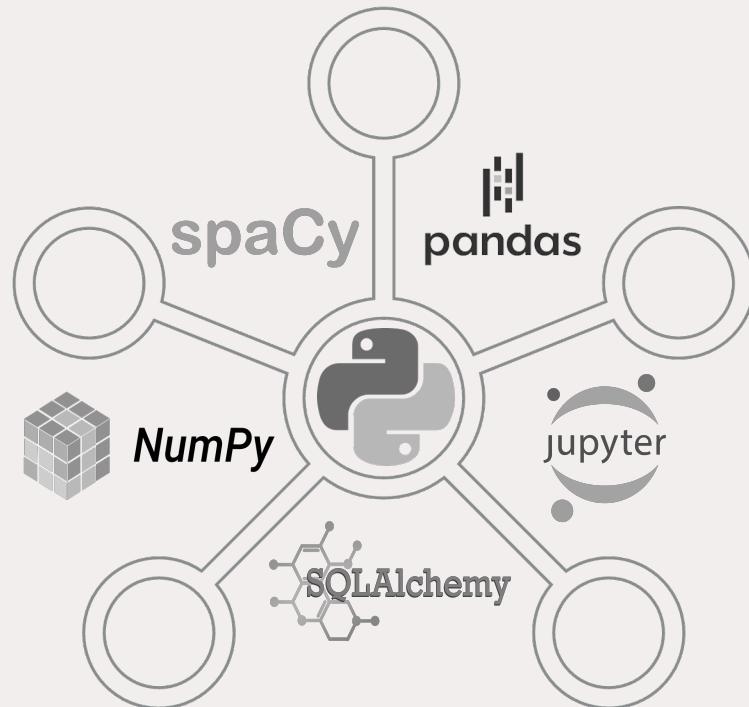
[Kaggle](#)

# Technologies and libraries

Sunday 21/11/2021

Newsletter

Edition: 001



# Workflow

Sunday 21/11/2021

Newsletter

Edition: 001



# Workflow

Sunday 21/11/2021

Newsletter

Edition: 001



## DATA Preprocess

- Drop Unnecessary Columns
- Rename Columns
- Dataset Concat

## EDA

- Drop Duplicates
- Drop Unnecessary Columns
- Check Null
- Text Preprocessing
- Apply Spacy Lemmatization

# Dimensions Reduction

Sunday 21/11/2021

Newsletter

Edition: 001

## Principal Component Analysis (PCA)

### Term Frequency-Inverse Document Frequency (TF-IDF)



# Topic Modeling & Clustering

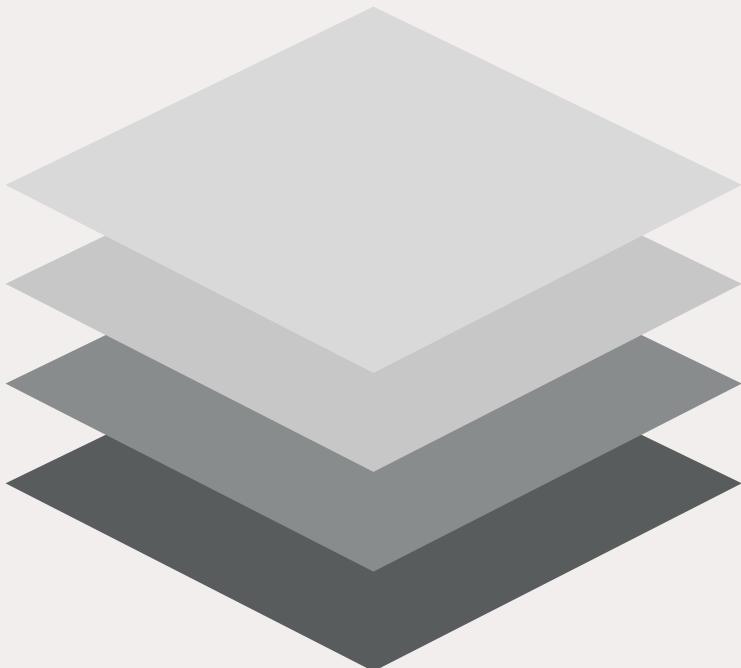
---

Sunday 21/11/2021

Newsletter

Edition: 001

## k-means



Political.

Culture and Society.

Cases and Accidents.

Humanities.

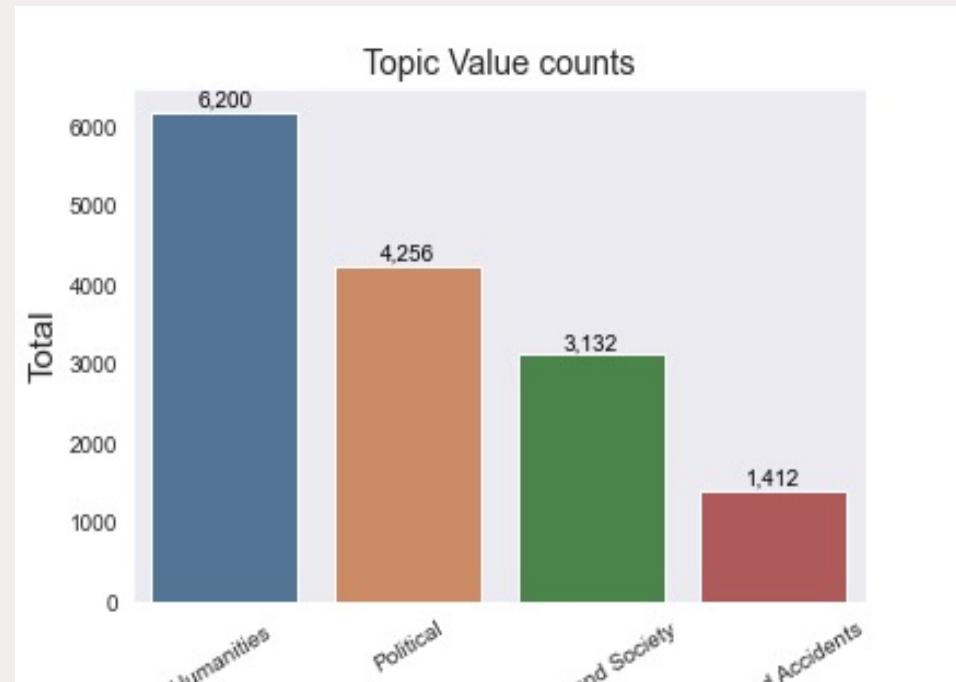
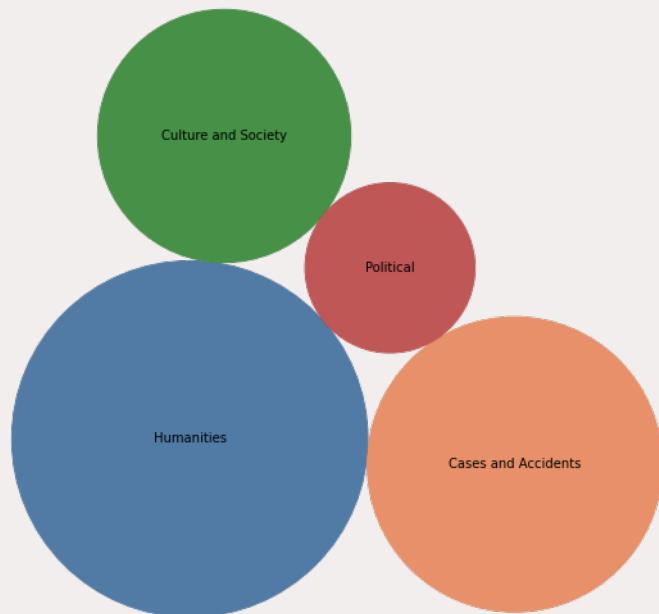
# Global News Topics

Sunday 21/11/2021

Newsletter

Edition: 001

## Most Discussed Topic by PCA



# Conclusion

---

Sunday 21/11/2021

Newsletter

Edition: 001



Based on our global news articles, most of the topics fall under Humanities, Cases - Accidents , Culture - Society and Political.

## Future Work

Accurately classify global news on several different topics.

# Global News Feed

Sunday 21/11/2021

Newsletter

Edition: 001

## THANKE YOU !

