

×

X

Victor QIAN Martin RAMPONT



#### Landslide after Rainfall



×

×



#### Thousands deaths Billions \$ damages

/Each year

# Use cases of Landslide Analysis



#### Landslide Suceptibility

Identify areas that are prone to landslides + map in real time





## Early warning system

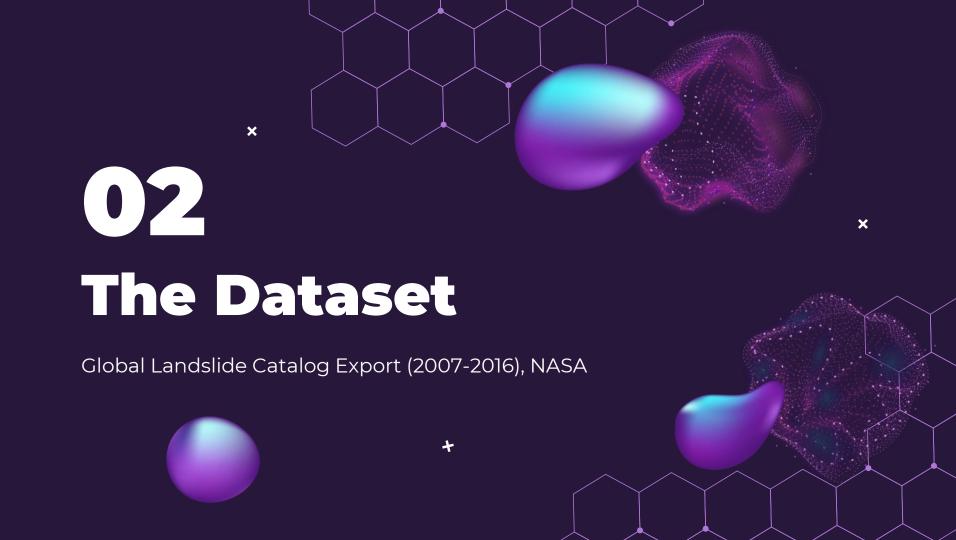
Provide timely warnings to residents and authorities.



X

#### Impact Assessment

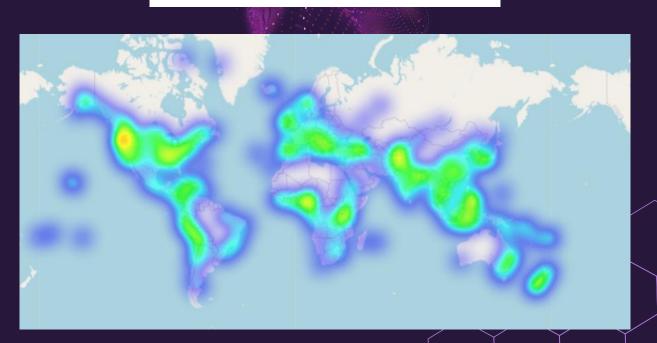
Assess the extent of damage and plan for recovery (Computer Vision)

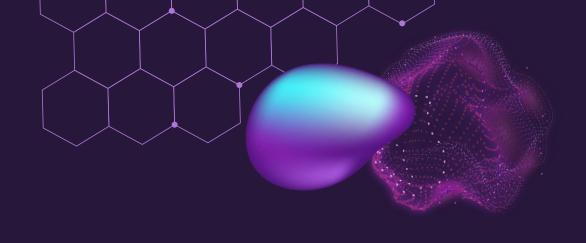


#### Global Landslide Catalog Export (2007-2016), NASA

X

Rows Columns Each row is a Landslide





**03 EDA** 

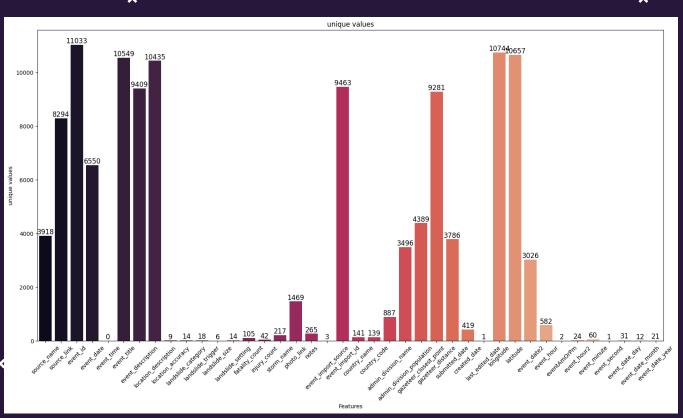
ASL Recognition



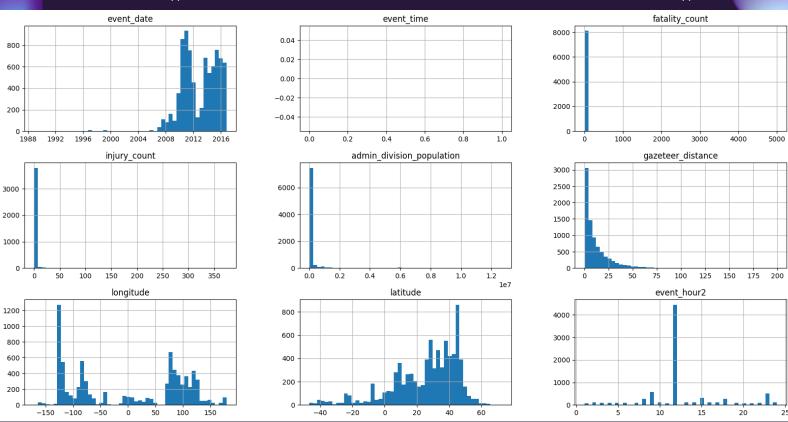
×



## Missing values

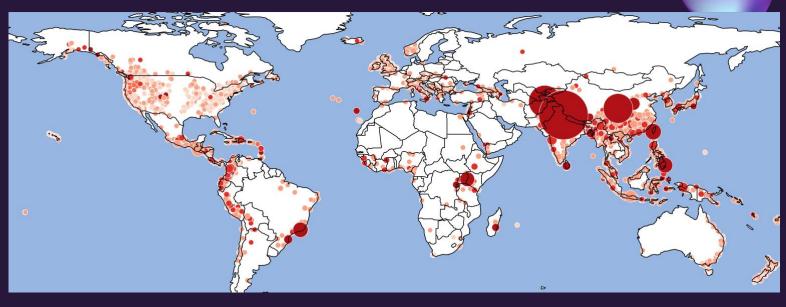


#### Distribution / outliers



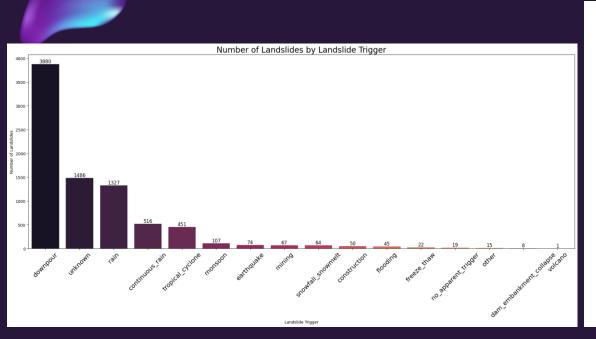
## **Map of Fatality Count**

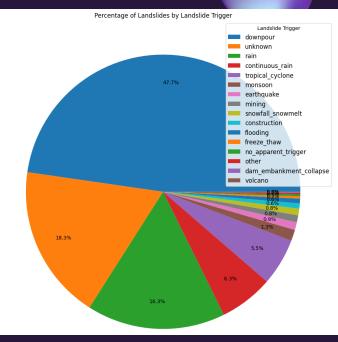
- landslide\_size small
- medium
- unknown
- large
- very\_large fatality\_count
- •





## **Triggers**





X

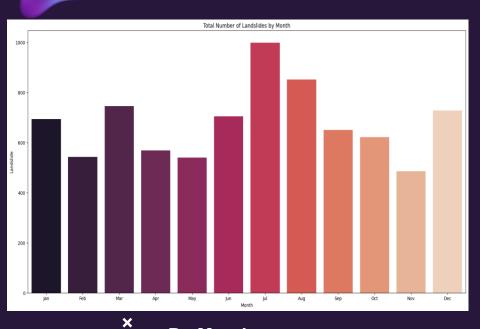
**×** By Month

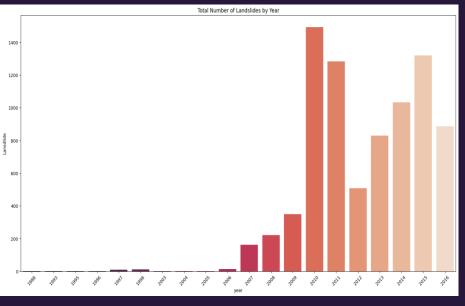
By Year

#### **Time Seasonality**



×

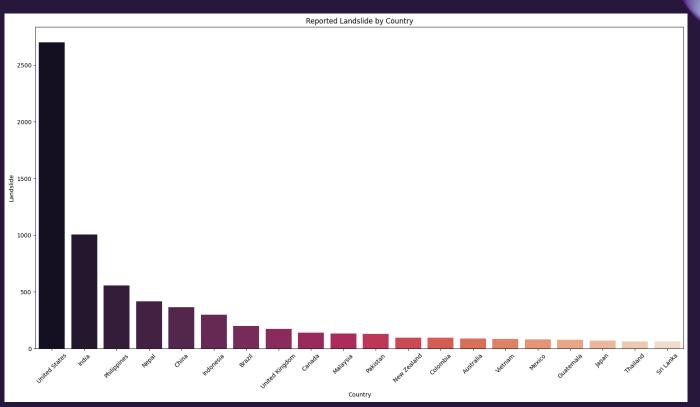




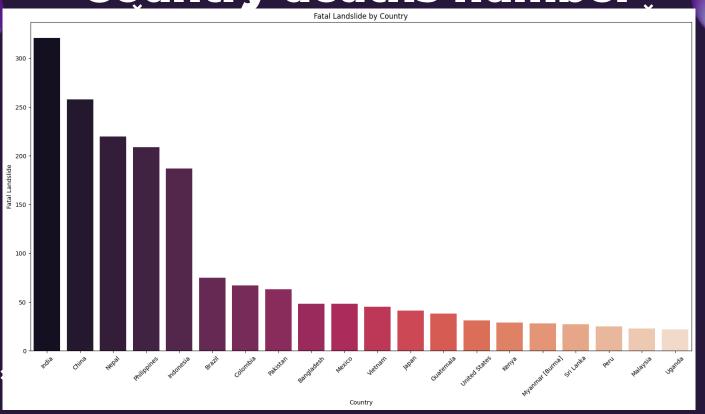
By Month

By Year

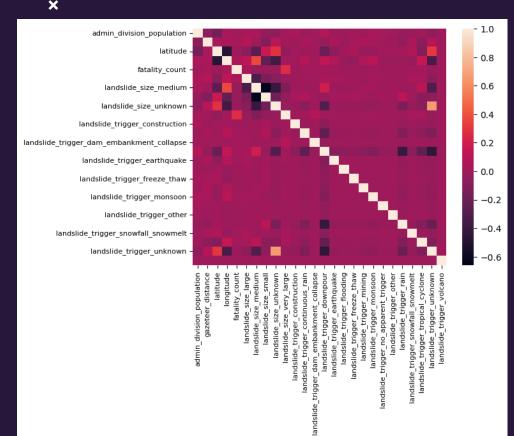
## Country Landslide Number



#### **Country deaths number**



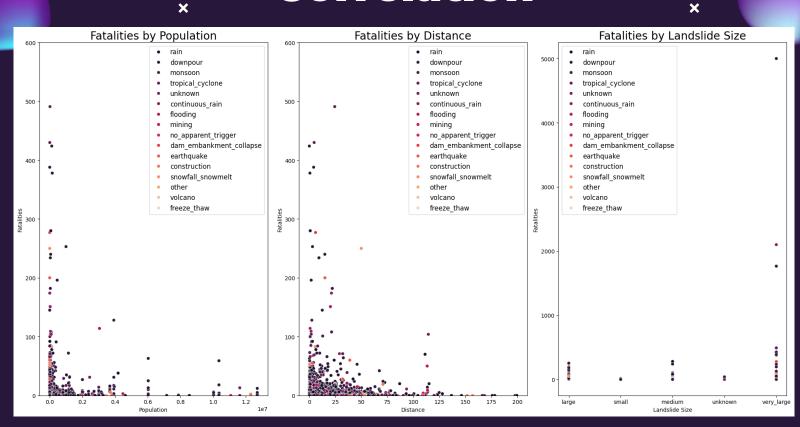
#### Correlation







#### Correlation





#### **Data Preprocessing**

Incorrect data types and values

#### Values Formating



Deals with missing values

×

#### Final output

#### Normalization



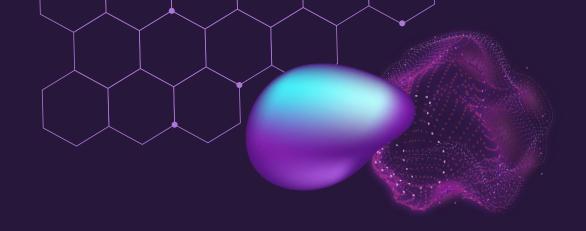
Time data splitting



×

#### Types of trained models

Al model	× Advantages	Drawbacks
Random Forest	Handles non-linearity and interactions well	Can be prone to overfitting, especially on noisy data
Random Forest Tuned (Grid Search)	Improved performance through hyperparameter tuning	Increased computational cost due to tuning
SVR	Effective in high-dimensional spaces	Sensitive to noise and may require careful preprocessing
SVR Tuned	Better generalization with optimized parameters	Computationally expensive, especially with grid search
Linear Regression	Simple and interpretable	Assumes a linear relationship, may not capture complex patterns
Neural Network	Powerful for complex, non-linear relationships	Requires large amounts of data and computation resources
X Tensorflow	Comprehensive deep learning framework	Steeper learning curve, especially for beginners
Xgboost	Handles missing values and outliers well	Prone to overfitting, may require tuning



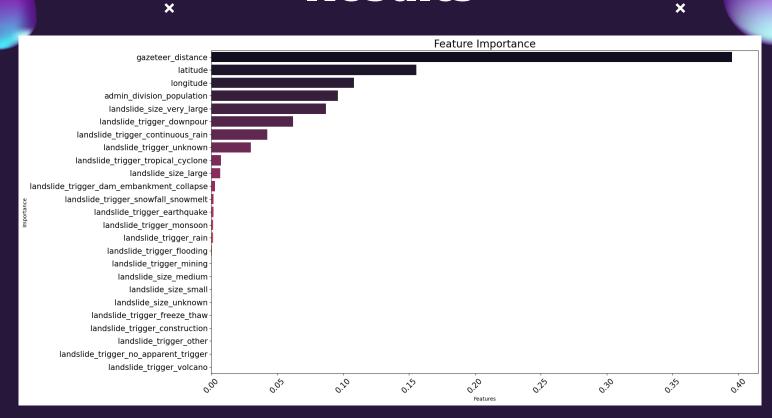
# 05 Results

ML Results

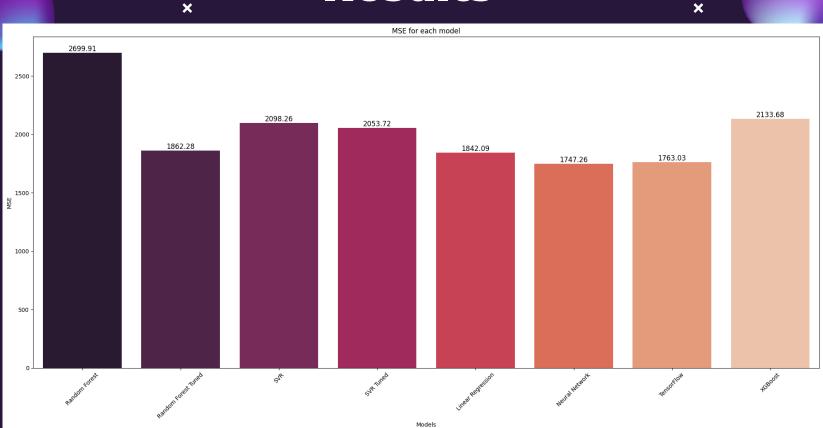


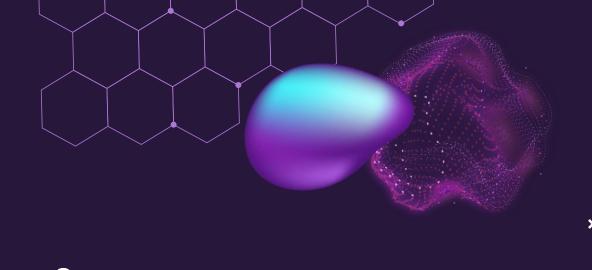


#### Results



#### Results





# 06

## **Streamlit App**

Project summary





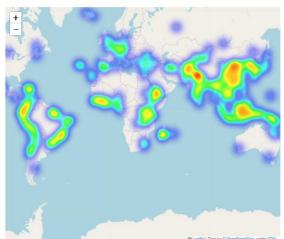


#### **Streamlit App**

#### Landslide

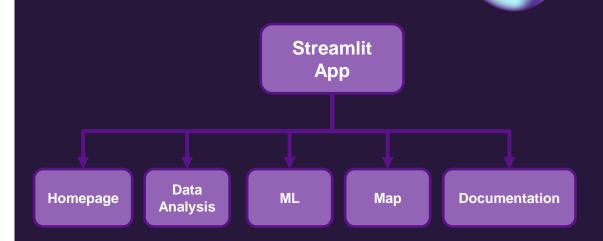
#### MAP

Fatality Count due to Landslides from 2007 to 2016



The map shows the number of fatalities due to landslides from 2007 to 2016. We can see that the number of fatalities is higher in Asia and South America (developping countries), even if the number of landslides is higher in Europe and North America (developped countries). This might be due to the fact that developped countries have better infrastructure and are more prepared to face natural disasters.

Moreover the number of fatalities is higher in the mountainous regions such as Pakistan and China. Finally the density of population in mountainous regions is higher in thoose countries.



#### Python for Data Analysis

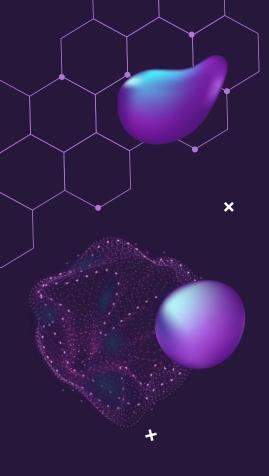
- Documentation

Landslide after Rainfall

QIAN Victor - RAMPONT Martin



×



## THANKS!



DO YOU HAVE ANY QUESTIONS?

