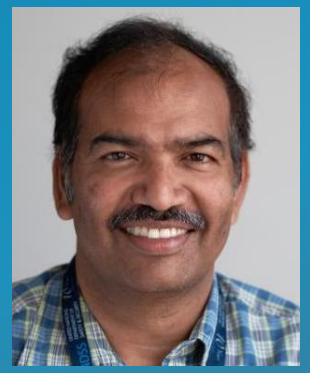
UN environment programme

50 1972-2022

A 20-minute journey through the technological aspects and a view from above

Meet your Hack Mentors and Experts

UN environment programme | 50 1972-2022



Muralee Thummarukudy

Paula Padrino Vilela

Matheus Ferreira Gois Fontes

Thomas Matheickal

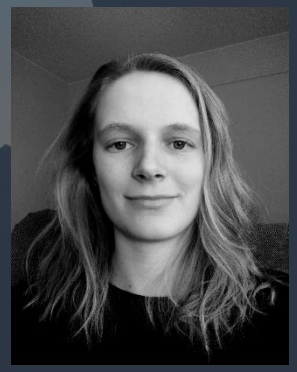
Maxime Croft

Melissa Puerto Aguayo

Dominik Bisslich

Arielle Aberdeen

Tlamelo Makati



Eike Hinrichsen

Abhijith Mohanan

Aman KC

Darius Moruri

HKUST and Zindi are our Partners for this Hack

UN environment programme 50 1972-2022

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

ZINDI

UN environment programme 50 1972-2022

Characteristics about Hong Kong and landslides *by UNEP*

Introduction to the features of the dataset *by HKUST*

First steps with the dataset and the Starter Notebook *by Zindi*

Join us in our mission to revolutionize disaster management *by UNEP*

UN environment programme 50 1972-2022

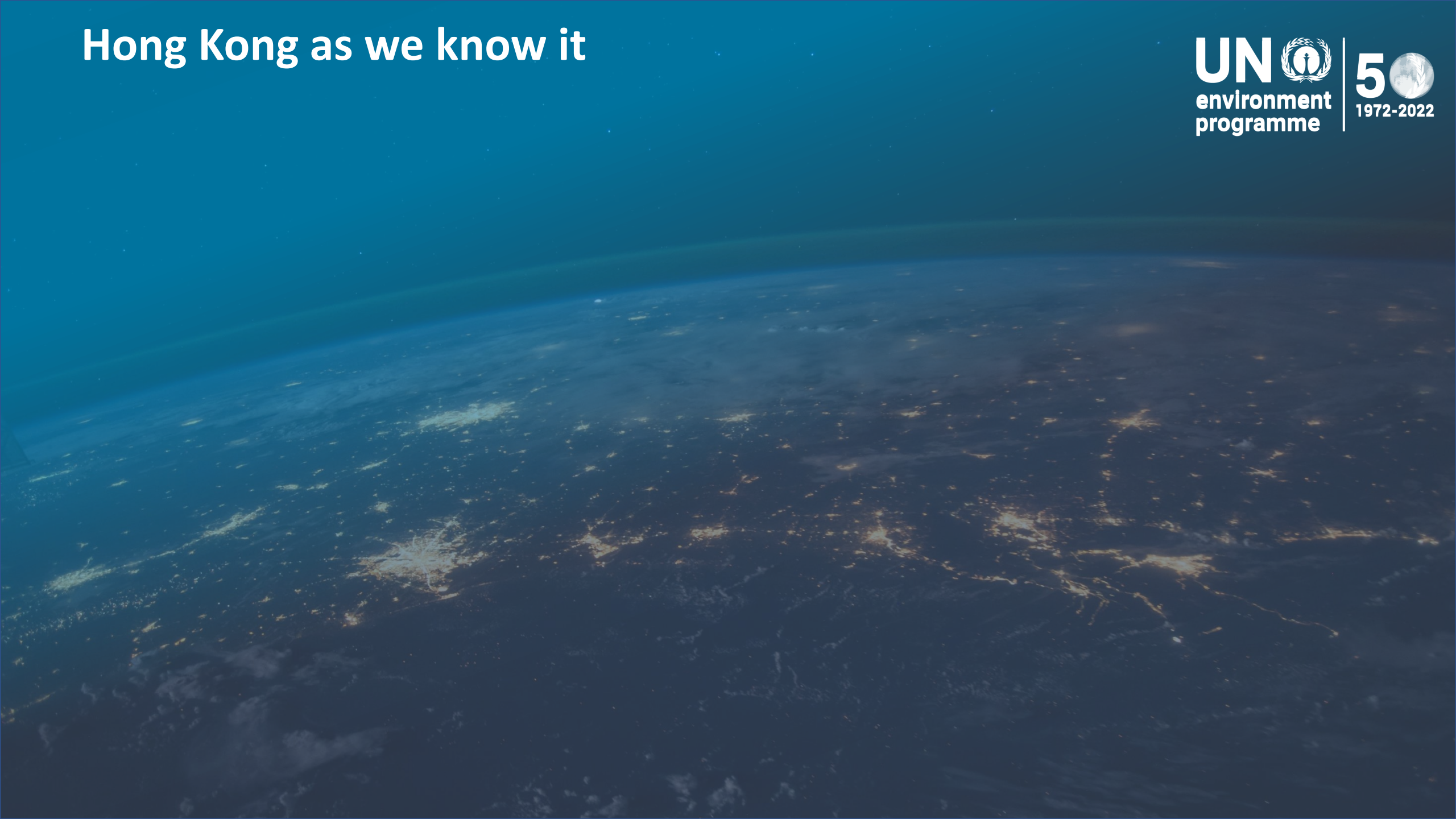
Characteristics about Hong Kong and landslides *by UNEP*

Introduction to the features of the dataset *by HKUST*

First steps with the dataset and the Starter Notebook *by Zindi*

Join us in our mission to revolutionize disaster management *by UNEP*

Hong Kong as we know it




UN environment programme
50
1972-2022
Po Shan Road
67 fatalities
1972
Image by Hong Kong Government

UN environment programme 50 1972–2022

Sau Mau Ping
18 fatalities
1976

Image by Hong Kong Government


UN environment programme
50
1972-2022
Tai O
< 1000 people stranded
2008
Image by Hong Kong Government

Our current technique is visual inspection, but...

UN environment programme | 50 1972-2022

Visual Inspection that is labor-intensive, expensive, sometimes difficult and cannot be repeated often

Modern Technologies like Satellite images, drone footage and Machine Learning/Deep Learning are paving the way

We expect an improved model for the technical part

UN environment programme 50 1972-2022

1010
1010

Binary Classification of Landslides with a robust model

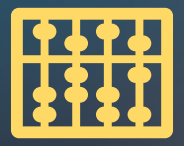
Improve your model performance with cross-validation, feature engineering, different modelling techniques…

Use your model and gathered insights to envision a new way of disaster management

We use the F1-Metric for technical evaluation

UN environment programme 50 1972-2022

For evaluating your technical sophistication we use F1-Score, a combination of Precision and Recall



Your model will account for 60 %, the other 40 % are evaluated based on your entrepreneurial spirit

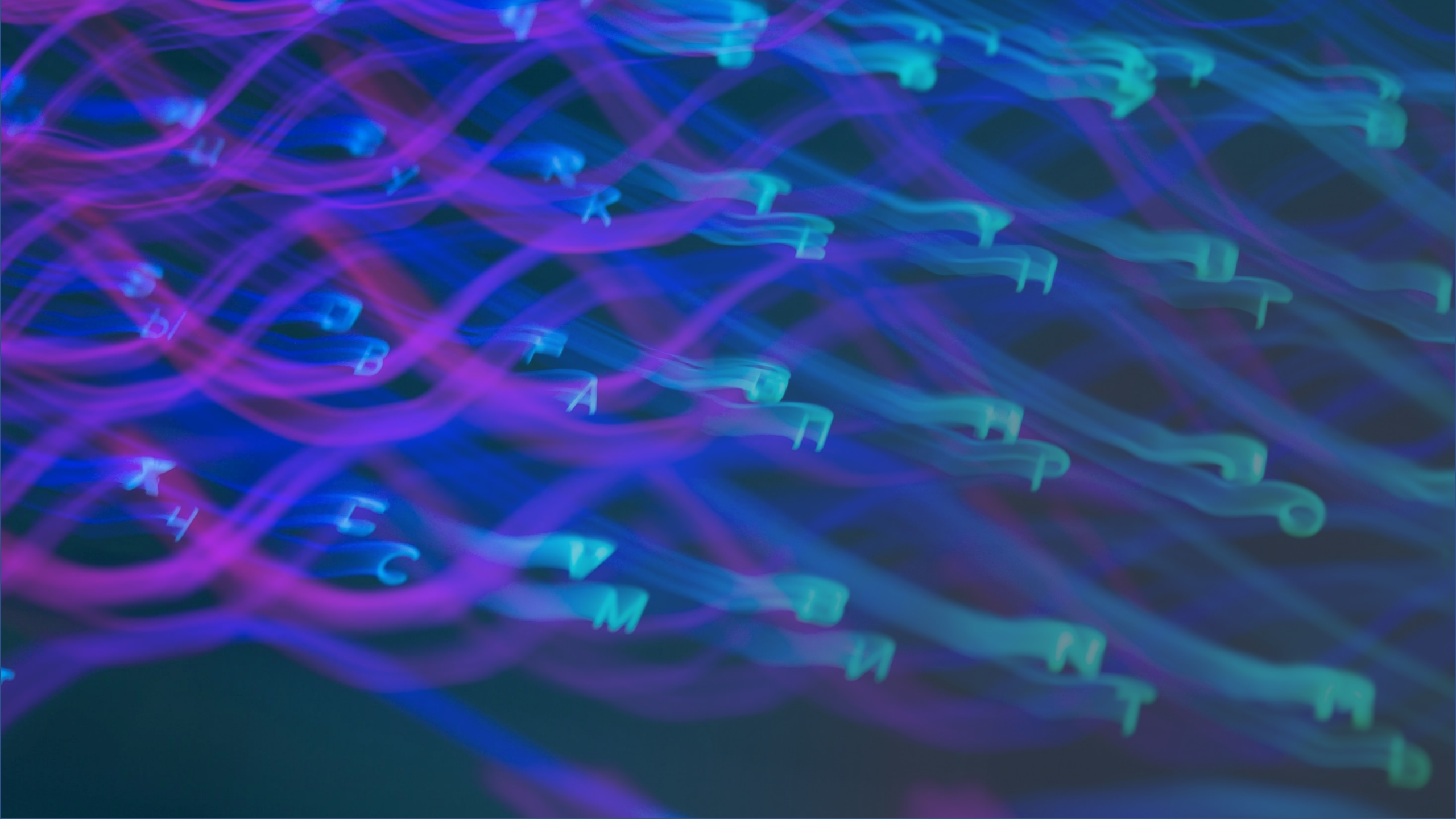
UN environment programme 50 1972-2022

Characteristics about Hong Kong and landslides *by UNEP*

Introduction to the features of the dataset *by HKUST*

First steps with the dataset and the Starter Notebook *by Zindi*

Join us in our mission to revolutionize disaster management *by UNEP*



UN environment programme 50 1972-2022

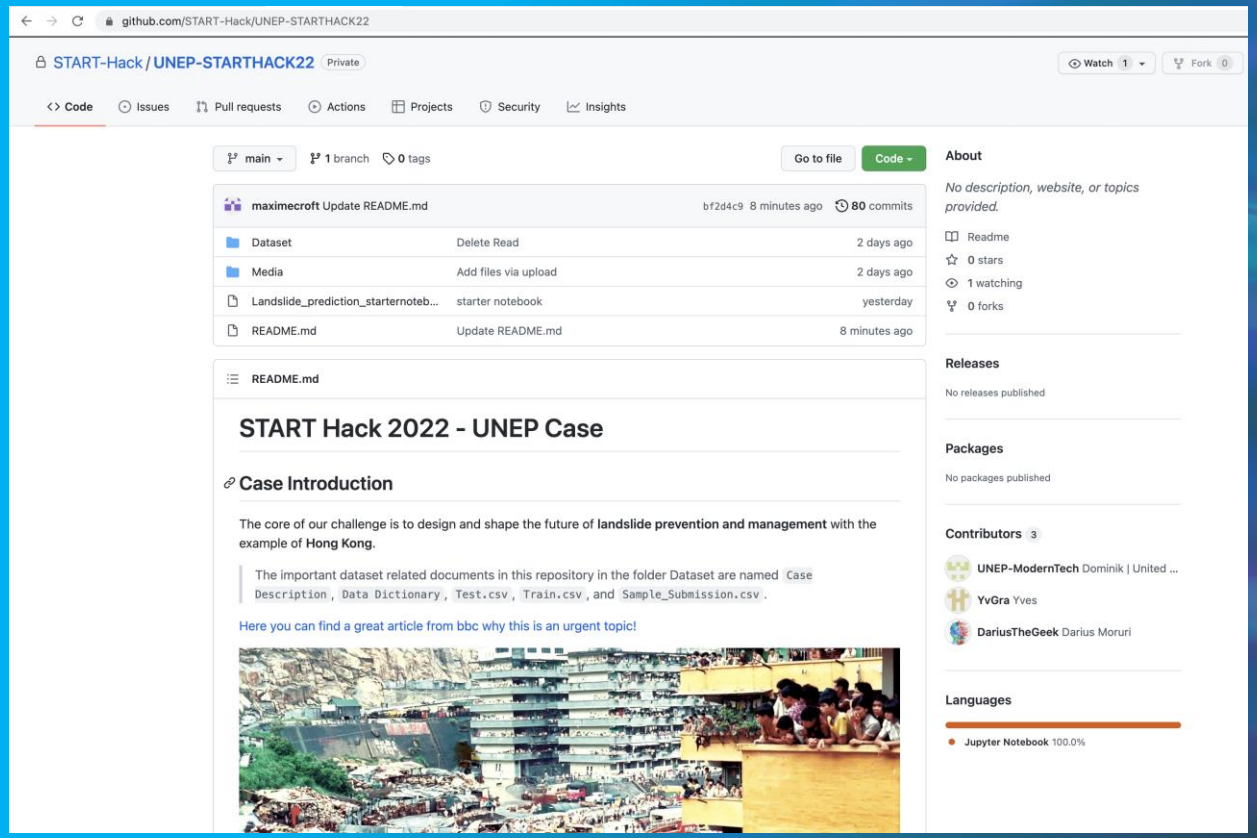
Characteristics about Hong Kong and landslides *by UNEP*

Introduction to the features of the dataset *by HKUST*

First steps with the dataset and the Starter Notebook *by Zindi*

Join us in our mission to revolutionize disaster management *by UNEP*

In the Notebook you will find everything you need



Content

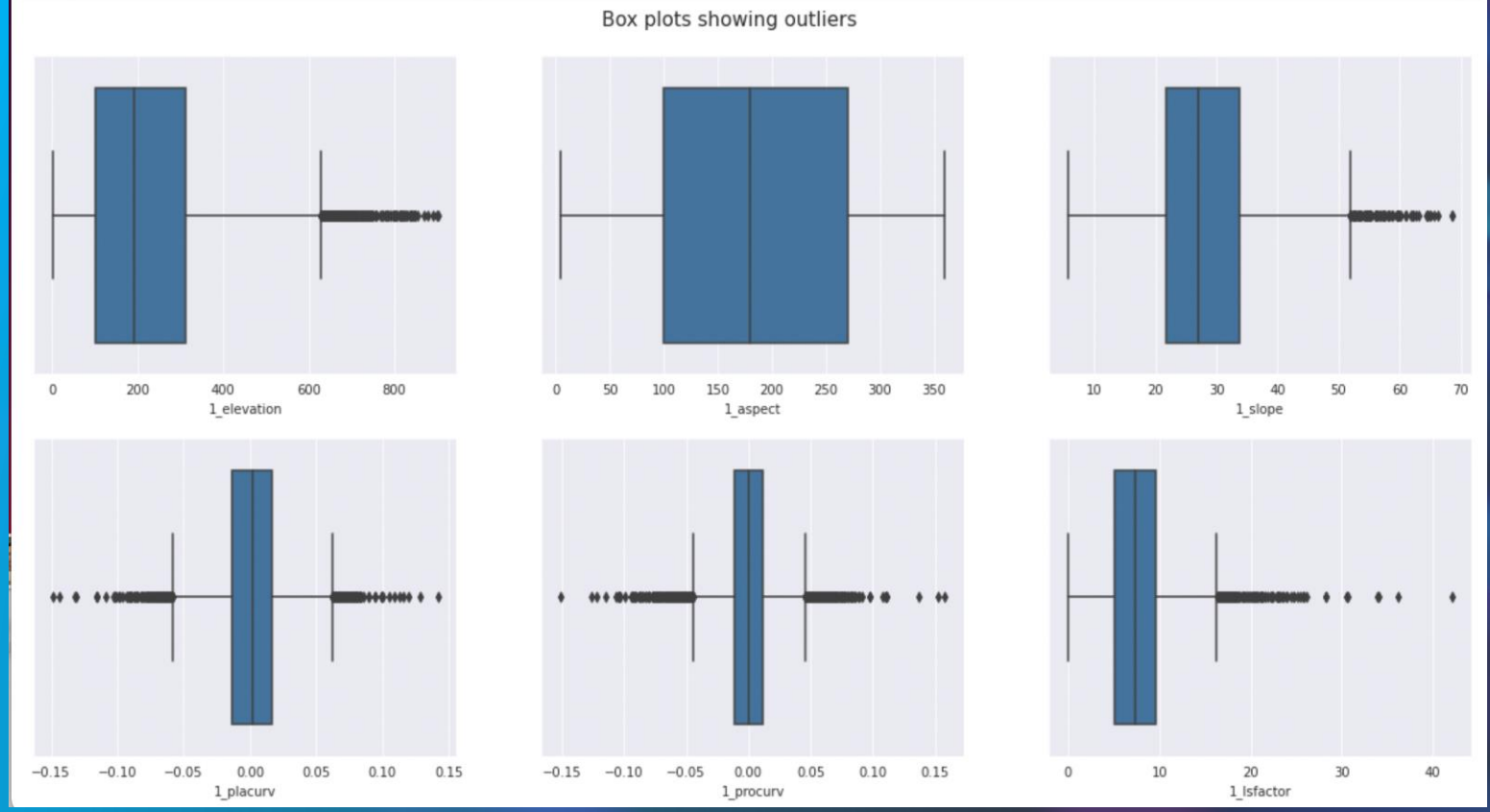
1. Import relevant libraries
2. Load files
3. Preview files
4. Data dictionary
5. Data-exploration
6. Target distribution
7. Outliers
8. Correlations
9. Model training
10. Test data predictions
11. Creating submission file
12. Tips to improve model performance

The outliers hint, transform them before modelling

```
# Exploring some features for cell 1
explore_cols =  ['1_elevation', '1_aspect', '1_slope', '1_placurv', '1_procurv', '1_lsfactor']
explore_cols
```

```
['1_elevation', '1_aspect', '1_slope', '1_placurv', '1_procurv', '1_lsfactor']
```

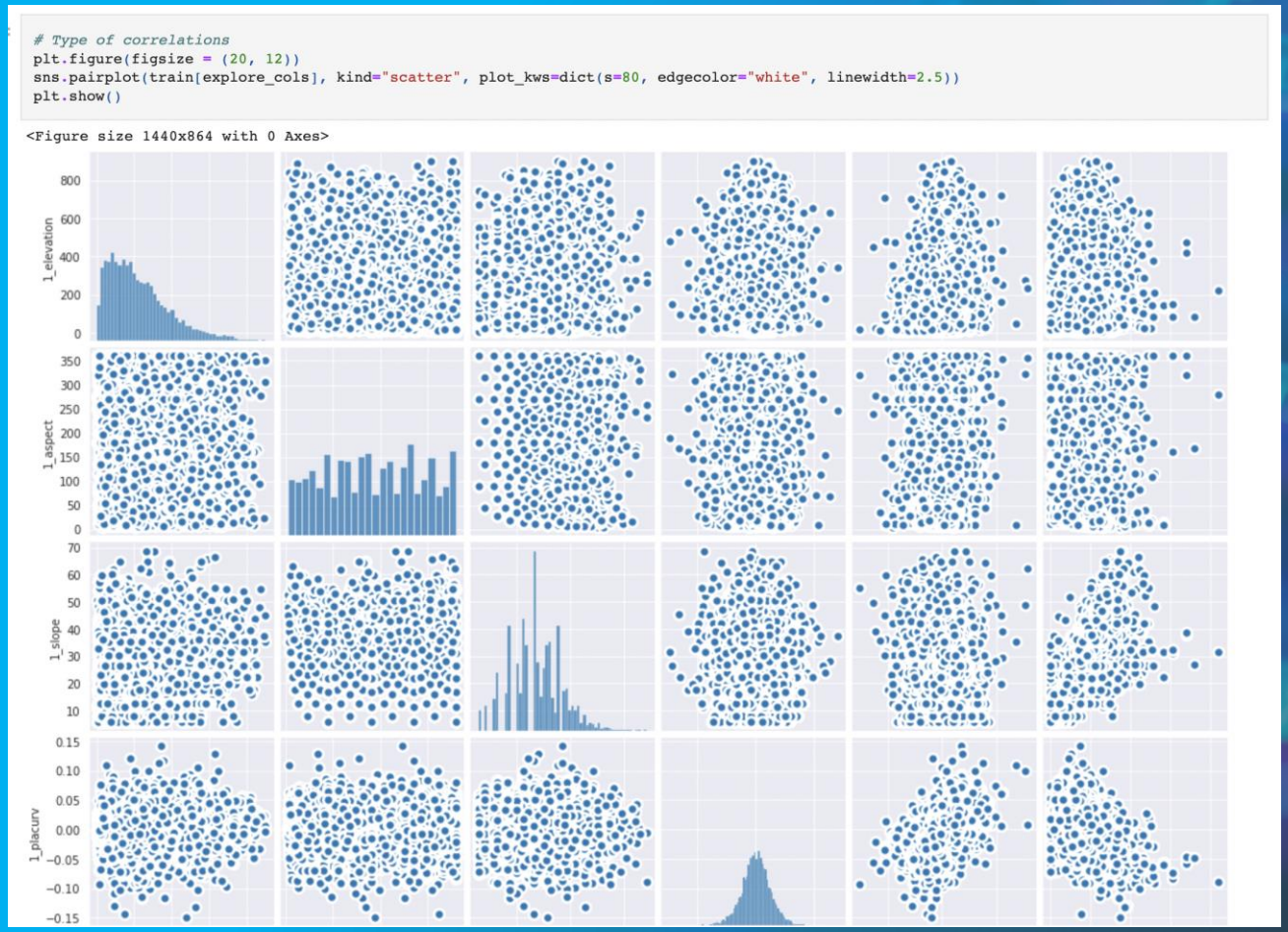
```
# Plotting boxplots for each of the numerical columns
fig, axes = plt.subplots(nrows = 2, ncols = 3, figsize = (20, 10))
fig.suptitle('Box plots showing outliers', y= 0.93, fontsize = 15)

for ax, data, name in zip(axes.flatten(), train, explore_cols):
  sns.boxplot(train[name], ax = ax)
```

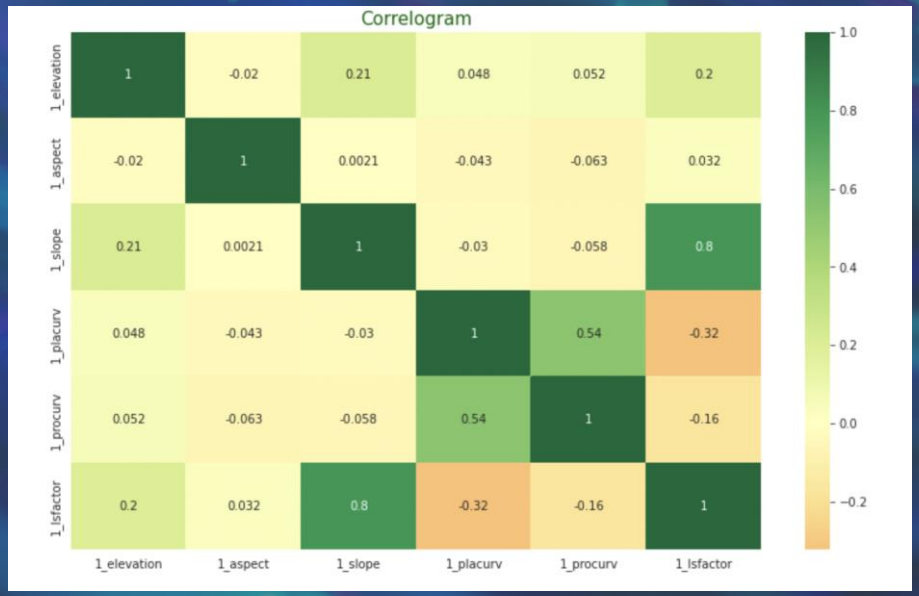


The correlation analysis shows xy

```
# Type of correlations
plt.figure(figsize = (20, 12))
sns.pairplot(train[explore_cols], kind="scatter", plot_kws=dict(s=80, edgecolor="white", linewidth=2.5))
plt.show()
```

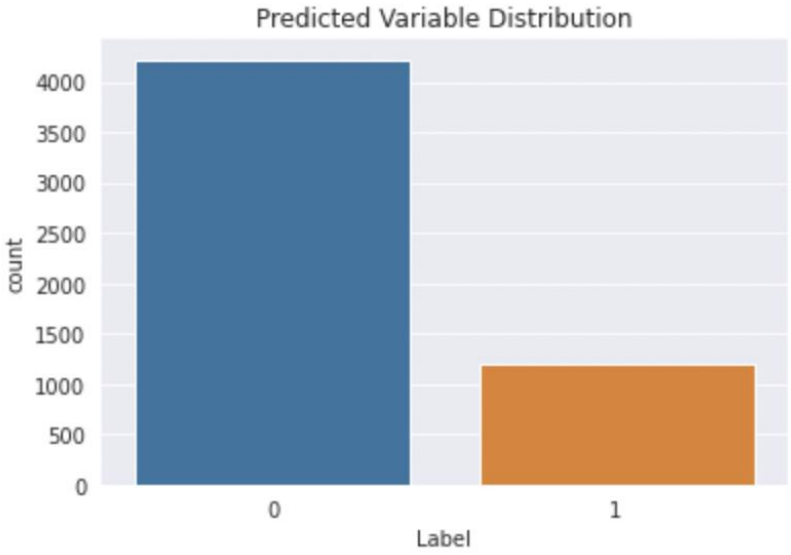
```
<Figure size 1440x864 with 0 Axes>
```





We also built the first model to start with

```
# Make prediction on the test set
test_df = test[main_cols]
predictions = model.predict(test_df)

# Create a submission file
sub_file = pd.DataFrame({'Sample_ID': test.Sample_ID, 'Label': predictions})

# Check the distribution of your predictions
sns.countplot(x = sub_file.Label)
plt.title('Predicted Variable Distribution');
```



Our tips

- Use cross-validation techniques
- Feature engineering
- Handle class imbalance
- Try different modeling techniques
- Data transformations
- Feature selection tech.
- Domain knowledge

Zindi is the leading data competition platform in Africa

ZINDI

Zindi is the first data science competition platform in Africa

Hosting an entire data science ecosystem on solving Africa's most pressing problems

UN environment programme 50 1972-2022

Characteristics about Hong Kong and landslides *by UNEP*

Introduction to the features of the dataset *by HKUST*

First steps with the dataset and the Starter Notebook *by Zindi*

Join us in our mission to revolutionize disaster management *by UNEP*

Join us in our mission!

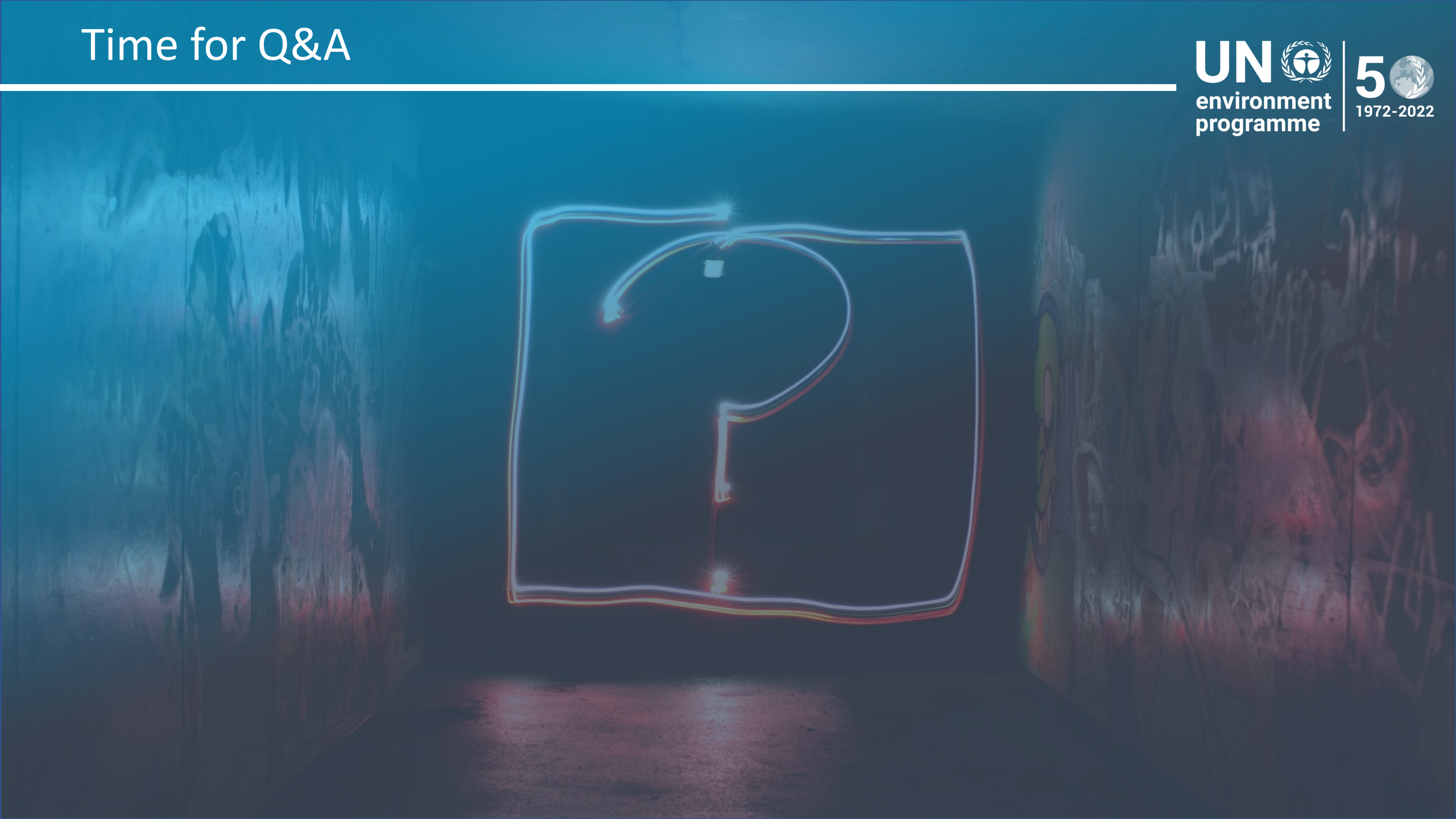
UN environment programme 50 1972-2022

Use Real-time Ranking and the expertise from zindi

Work with a dataset from the real world

Prevent, mitigate, and respond to disasters to protect lives

Time for Q&A



UN environment programme

50 1972-2022

Enjoy our Hack and you can always visit us at the Sustainability Lounge