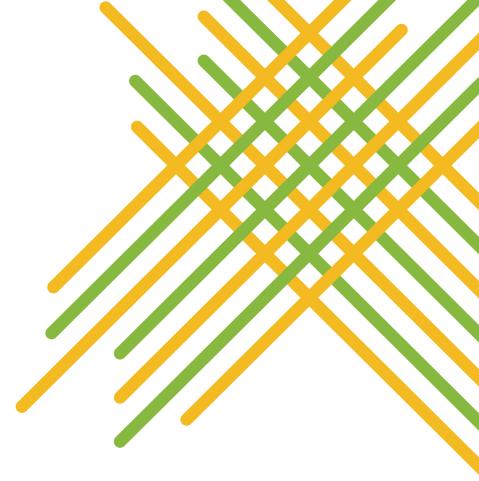




AKADEMIYA



ReSAKSS DATA CHALLENGE

GUIDELINES

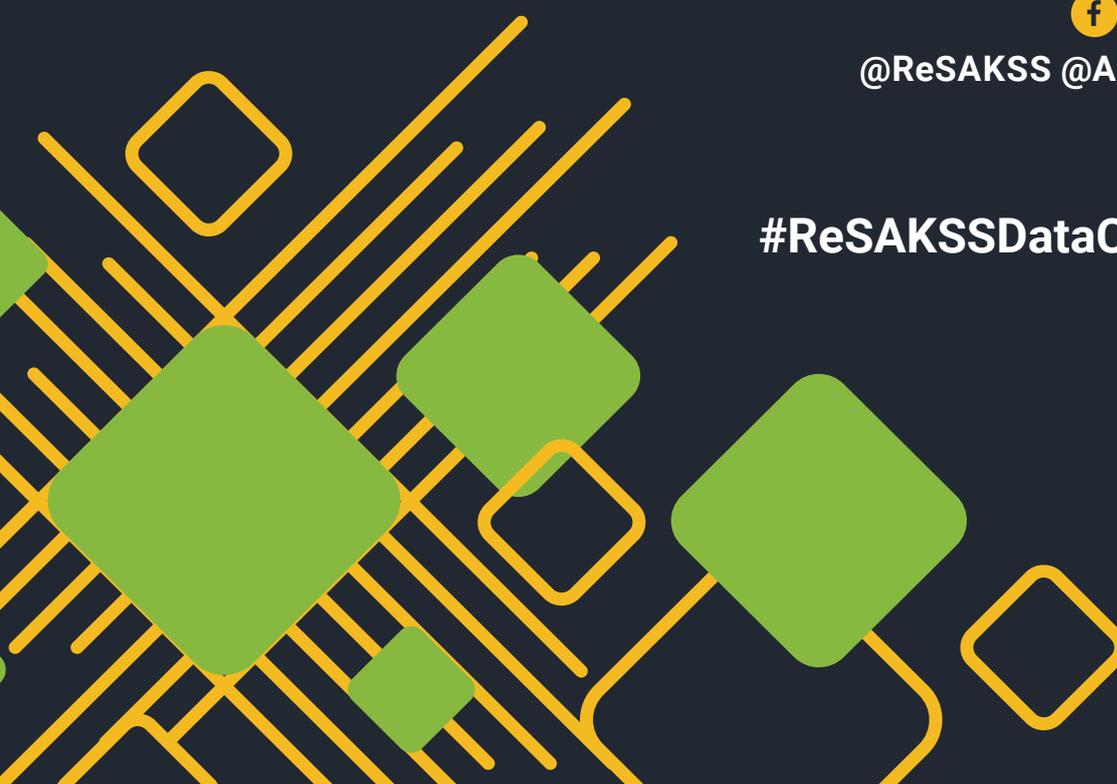
IT Predictive Modeling

Join the conversation:



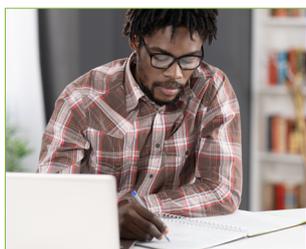
@ReSAKSS @AKADEMIYA2063

#ReSAKSSDataChallenge22



1. Official Rules

The ReSAKSS Data Challenge edition 2022, is a competition for enthusiastic data users to showcase their creativity and analytical skills using data and other resources on the ReSAKSS website to develop an innovative knowledge product or project that addresses a development challenge faced by Africa in one of four categories:



i) Essay or Report

ii) Visual Art

iii) ¹Information Technology (IT) Product or Service /IT Predictive Modeling

iv) ²High School Project

2. Eligibility

The challenge is open to researchers, high school and university students, graphic designers, IT specialists, analysts, and anyone else in Africa, passionate about creatively using data available on the [ReSAKSS](#), [eAtlas](#), and [AAgWa](#) platforms to communicate and address development challenges faced by Africa. The challenge is open to all Africans within the continent and in the diaspora. All individuals under the age of 13 must have a parent or legal guardian submit the application on their behalf. All participants under the age of 18 must have permission from a parent or legal guardian to enter this challenge. The challenge organizers reserve the rights to verify an individual's eligibility to participate in the challenge.

Previous winners are not eligible to apply.

Eligible participants are invited to submit their products or projects through the ReSAKSS Data Challenge website to compete for prizes:

- **Website:** <https://data-challenge.resakss.org/>
- **Email inquiries:** Kindly submit any questions via: data-challenge@resakss.org

Each candidate's submission along with all materials provided in connection with the challenge is referred to herein as an "application". The challenge will be comprised of an open application phase, selection phase, announcement of winners, and a presentation of awards to the most innovative projects from each competition category during the ReSAKSS Annual Conference to be held in October 2022.

3. Dates and Times

Applications open on July 5, 2022, and close on September 5, 2022. A summary of the schedule for the data challenge is presented below:

Activity	Timeline
Eligible participants submit their application	July 5 to September 5, 2022
Jury selects most innovative projects and winners	September 30
Announcement of winners	October 5
Presentation of awards to winners during the 2022 ReSAKSS Annual Conference	October 27

¹ The IT Product & Service Category guideline is available on the IT Product & Service page on the challenge website

² The High School Project category is open to high school students only. Submitted projects in this category can be: i) Essay or Report, ii) Visual Art, iii) Information Technology (IT) Product or Service or a project in any other field of their choice as long as they use ReSAKSS resources.

4. How to Participate

- To enter the competition, submit one application either as an individual or as a team.
- A maximum 4 individuals per team is permitted.
- You may participate in **ONLY ONE (1)** category, whether as an individual or a team. If you participate in multiple category applications, only the first application you submit will be considered.
- Applications not submitted in accordance with the instructions provided on the ReSAKSS Data Challenge Site and these official rules may be disqualified at the sole discretion of the challenge organizers.
- **Inquiries:** Kindly submit any questions via: data-challenge@resakss.org.

5. Conditions and Requirements

- Digital platforms: resources to be exploited include:



ReSAKSS: www.resakss.org

eAtlas: <https://eatlas.resakss.org/>

AAgWa: www.aagwa.org

Each submitted knowledge product or project must be innovative and clearly demonstrate how ReSAKSS data or other resources have been used to address a development challenge faced by Africa.



Remember to cite any ReSAKSS data you use using the following suggested citation: "ReSAKSS, 2022."



Fill in the application form and send it along with a clear and detailed description of your work in PDF and submit a link to your work if it is a video, a photo, application, or other heavy material.



The application form is accessible directly from the ReSAKSS Data Challenge website:

<https://data-challenge.resakss.org/>



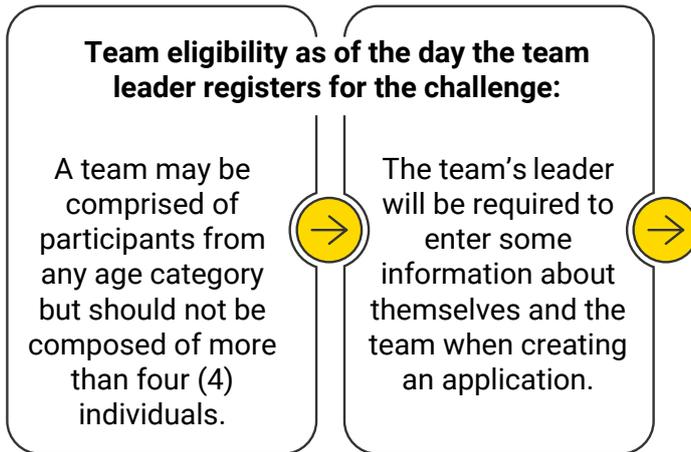
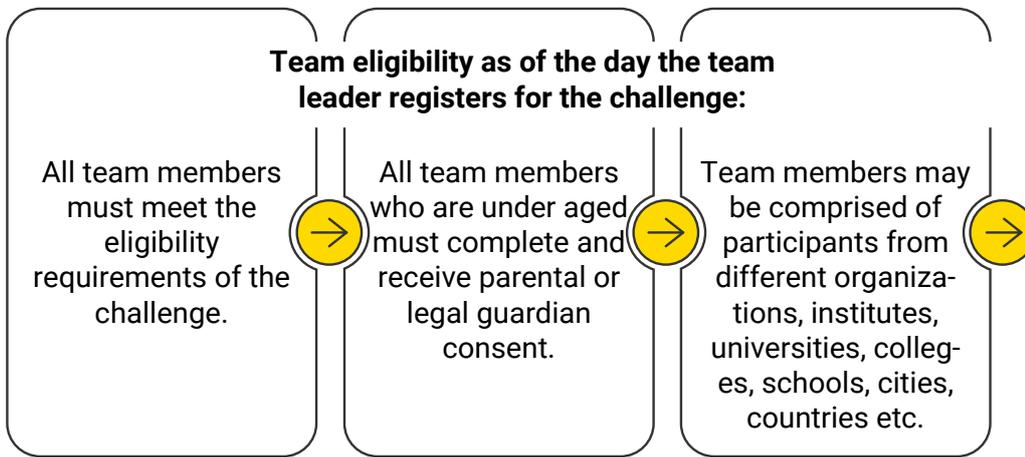
Creating your Application (as an individual or a team)

The submitted documents must follow the technical, creative, and legal entry requirements set out in these official rules.



You will be required to complete all sections marked with an asterisk (*) on the application form on the Data Challenge website.





6. Judging Criteria and Process

The ReSAKSS Data Challenge jury members are experts in their respective fields: university professors, senior researchers, visual designers, journalists, and IT specialists.

Category 3: (IT Contest 2) Predictive Modeling

Description: The world is experiencing an increase in wheat prices since March 2021. The main reason is the Russia-Ukraine conflict and both countries are in the pool of the global top-10 producers of wheat and count for 30% of the global wheat trade. In 2020, Russia was the third largest wheat producer in the world with near 86 million tons produced in the country and Ukraine was at the eighth place with near 26 million of tons. The conflict is causing an increase in wheat price due to Ukraine likely inability to produce as much wheat as before and the fact that other countries are reducing or shutting down their trade with Russia. Such dynamic suggests most exposed countries to scrutinize more their locally produced wheat to at least reduce the gap in wheat availability. A first step towards that goal would be to assess how much wheat production can be expected from the current season and it is the purpose of this challenge.

The challenge is about predicting – at a pixel level – wheat production in Zimbabwe for the year 2022 using machine learning techniques. The model you will develop uses remote sensing data such as the Normalized Difference Vegetation Index (NDVI), daytime Land Surface Temperature (LST), Rainfall (R), and evapotranspiration (ET) as explanatory variables, and historical wheat production values as response variable. Both parameters, explanatory and response, are shared in csv files with latitude and longitude coordinates. You are asked to predict wheat production for each pixel. A functional form of the task is to build a model F such as:

$$\text{Wheat Production} = F(\text{NDVI}, \text{LST}, \text{R}, \text{ET})$$

The choice of the machine learning technique should be the participants' choices as well as the addition of explanatory variables available on the ReSAKSS, e-Atlas, and AAgWa platforms for better accuracy performances. However, any addition of explanatory variables should be relevant to the prediction task and explainable.

The provided datasets are on an annual basis aggregated for each year's growing season. Your prediction, therefore, will be for the 2022 season.

Evaluation

The evaluation metrics is the Root Mean-Squared Error. Your submission file should contain predicted production values for wheat, for each pixel and for the year 2022. We isolated a part of the dataset for accuracy evaluation of your model's outputs. You should use the same production unit as the one used in the provided dataset.

Your submission should look like:

Pixel ID	Latitude	Longitude	2022 wheat production (predicted values).
00001	20
00002
...			

Information about the dataset parameters

- NDVI stands for Normalized Difference Vegetation Index. The NDVI is retrieved from multispectral satellite images and can help to assess how healthy a vegetation is. For more information, please visit this website: https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php.
- LST stands for Land Surface Temperature and is in Celsius Degree for day and night-time. It is the temperature of the air at one meter above ground.
- The rainfall dataset was retrieved from the Climate Hazards Group InfraRed Precipitations with Station data (CHIRPS) databases. For more information, please visit <https://www.chc.ucsb.edu/data/chirps>.
- The evapotranspiration was retrieved from the MOD16A2 version 006 from the Land Processes Distributed Active Archive Center (LP DAAC) product. For more information, please visit <https://lpdaac.usgs.gov/products/mod16a2v006/>.

Rules

The solution must use publicly available and open-source packages.

- Only submission files containing prediction values for all pixels will be considered.
- Individuals can form teams to participate in this contest. However, you cannot submit your work as an individual while you are a member of a participating team.
- Only one submission is allowed. However, you can request a second submission and the ReSAKSS Data Challenge team will review it.
- If you are the winner, the ReSAKSS Data Challenge team will request you to submit your model for copyright and predictive values verification.
- If two solutions have the same scoring, date and time of submissions will be considered to assess the winner.
- You agree to legally assign ownership of all rights of copyright of the winning solution code to the ReSAKSS Data Challenge, while reserving the right to solution code for non-commercial purposes, while crediting the ReSAKSS Data Challenge.
- The ReSAKSS Data Challenge can change the rules at any time.
- Each candidate can only participate in one of the two IT contests.



Prize: Only the 1st prize will be attributed for this IT Predictive Modeling contest.

NOTE: Any document submitted explaining the concept and/or realization of a project should not exceed 2,500 WORDS.

**BY SUBMISSION OF YOUR MATERIALS,
YOU INDICATE YOUR ACCEPTANCE OF THE GUIDELINES ABOVE**

-  AKADEMIYA2063 | Kicukiro/Niboye KK 341 St 22 | 1855 Kigali-Rwanda
-  +221 77 761 73 02 | +250 788 315 318 |
-  hq-office@akademiya2063.org
-  www.akademiya2063.org