EMERGENT

DATA 4 GOOD CHALLENGE

DOCUMENTATION

McKinsey & Company

CRONOS LEUVEN

KBC

Agilytic

Table of Contents

EMERGENT

DATA 4 GOOD CHALLENGE

1. Introduction

What is Emergent? What is the D4GC?

Emergent, the Data Science & AI student community

We believe that Data Science & AI will change every aspect of our future society. We aim to be the leading student community in this field, welcoming people from all backgrounds. We prepare students to reach their full potential and effect positive change with Data Science & AI throughout their lives.

One of our ways of achieving this vision is by letting students apply their skills to real-life problems through competitions like the D4GC.

The Data 4 Good Challenge, our flagship event

The D4GC allows students to solve a real case. Participants work together in multidisciplinary teams. Together, they are tasked with solving a socio-economic problem through the use of data. In a truly Emergent fashion, participants will have to think through all aspects of their solution, not only considering the economical effects but also the social and ethical consequences of their plan.

Since the challenge is two-fold - produce insights from data, then develop a strategy based on the insights - an interdisciplinary collaboration is the best way to tackle this challenge successfully.

EMERGENT

DATA 4 GOOD CHALLENGE

2. The Challenge

Adaptation, the solution to climate change?

No corner of the globe is protected from the devastating impact of climate change. Rising temperatures are fueling extreme events, agricultural degradation, economic crises and rising sea levels. The World Bank estimates that, in the absence of action, more than 140 million people in Sub-Saharan Africa, Latin America, and South Asia will be forced to migrate within their regions by 2050.

Despite increased climate ambitions and net-zero emissions commitments, governments still plan to produce more than double the amount of fossil fuels in 2030 than what would be consistent with limiting global warming to 1.5°C (2021 Production Gap Report). As we speak, leaders from around the world have gathered for one of the most important climate summits in years, the COP26 in Glasgow, Scotland. Political obstacles are mounting, developing countries face high costs and barriers to participate, and many world leaders of major emitters have declined their attendance.

As politicians all over the world remain short-sighted, a possible scenario is that solving the climate problem will largely come down to adaptation, rather than mitigation. Adapting to climate change means taking action to prepare for and adjust to both the current effects of climate change and the predicted impacts in the future.

During this challenge, you will work under the assumption that emissions follow a 'business-as-usual' trajectory. Your target is to assess the risks that climate change pose under this scenario and figure out how we can adapt to them.

3. The Data

Dig in & gain some insights!

You are provided with 5 datasets. Datasets 1 and 2 will help you in your analysis of climate change risk. Dataset 3 provides general economic and demographic data Datasets 4 and 5 provide background information around the climate change issue that you can use to build your storyline.

The datasets are complementary to each other. You can get limited insights from the individual datasets, but to properly tackle the challenge, you should combine some of them.

Dataset 1: Climate Change

This dataset contains historical and projected data on variables related to various aspects of weather and climate. Note: you should take this data as given, and assume that there is little uncertainty around the accuracy of the predictions and the scenarios they are based on.

Variable Name	Variable Description	Possible Risks
country iso_alpha3 region sub_region	Name of Country, alphabetically sorted ISO 3-letter country code Region country belongs to Sub-region country belongs to	
yearmonth	Month of measurement or prediction (Jan 1960 - Dec 2100[1])	
tas	Near surface air temperature: Temperature of air at 2m above the surface of land, sea or inland waters, in degrees Celsius (Range: Jan 1960 - Dec 2100)	Destruction of flora and fauna; altered precipitation; change in sea levels (Source)
scfwind	Near surface wind speed: Magnitude of the two dimensional horizontal air velocity near the surface, in m/s (Range: Jan 1960 - Dec 2100)	Signals climate disruptions as the main cause of wind is differential heating (Source)
pr	Precipitation: The sum of rain and snow that falls to the Earth's surface, in kg/m^2s (Range: Jan 1960 - Dec 2100)	Exposure to flood events; overwhelmed drainage infrastructure; decrease in water quality; disruption of human security (Source)
prsn	Snowfall flux: Mass of water in the form of snow precipitating per unit area, in kg/m^2s (Range: Jan 1960 - Dec 2100)	Increase in the amount of solar radiation absorbed by the earth due to decreasing coverage of snow and ice, thereby enhancing the global warming process (Source)

[1] The data from Jan 1960 - Dec 2014 are historical data (measured); The data from Jan 2015 - Dec 2100 are predictions

EMERGENT

DATA 4 GOOD CHALLENGE

3. The Data

Dig in & gain some insights!

Dataset 1: Climate Change (cont.)

This dataset contains historical and projected data on variables related to various aspects of weather and climate. Note: you should take this data as given, and assume that there is little uncertainty around the accuracy of the predictions and the scenarios they are based on.

Variable Name	Variable Description	Possible Risks
evspsbl	Evaporation including sublimation and transpiration, in kg/m^2s (Range: Jan 2015 - Dec 2100)	Increased risk of droughts; insufficient or irregular water supply for people and industry causing severe harm and economic impact; loss of agricultural productivity; risk of food security (Source)
tos	Sea surface temperature: Temperature of sea water measured near the surface, in degrees Celsius (Range: Jan 1960 - Dec 2100)	Destruction of ecosystems; Increased susceptibility of costal and small island developing states (SIDS) fishing communities who depend on these ecosystems (Source)
sos	Average annual sea surface salinity: salt concentration close to the ocean's surface, in PSU (Range: Jan 1960 - Dec 2100)	Along with the water temperature, it determines water density which drive circulation. Circulation plays a key role in transporting heat and maintaining the Earth's climate (Source)
sithick	Sea ice thickness: Vertical extent of ocean sea ice, in m (Range: Jan 2015 - Dec 2100)	Sea ice forms bright surfaces which reflect sunlight into the atmosphere instead of the ocean preventing the increase of ocean temperature; Impacts circulation by affecting salinity (Source)

3. The Data

Dig in & gain some insights!

Dataset 2: Sea Level Change

This dataset contains historical and projected data on variables related to sea level and coastal populations. Note: you should take this data as given, and assume that there is little uncertainty around the accuracy of the predictions and the scenarios they are based on.

Variable Name	Variable Description	Possible Risks
country	Name of Country, alphabetically sorted	
iso_alpha3	ISO 3-letter country code	
region	Region country belongs to	
sub_region	Sub-region country belongs to	
year	Year of measurement or prediction (2020 - 2150, 10 year periodicity)	
avg_sea_level_change_rate	Average rate of sea-level change in country, in mm per year (Range: 2020-2150)	Impacts costal habitats; Increased risk of erosion, flooding, soil contamination with salt which impacts agricultural productivity (Source)
avg_sea_level_change_value	Average value or level of sea-level change in country, in mm (Range: 2020-2150)	
pop_hi_tide	Population in Areas Below Projected High Tide Line[2], in millions (Range: 2020, 2050, 2100)	
pop_coast_flood	Population in Areas Below Projected Annual Average Coastal Flood Levels, in millions (Range: 2020, 2050, 2100)	

[2]The line of intersection of the land with the water's surface at the maximum height reached by a rising tide

EMERGENT

DATA 4 GOOD CHALLENGE

3. The Data

Dataset 3: Economic & Demographic Indicators

This dataset contains general data like macro-economic and demographic indicators for a number of countries and years. It can be useful to combine with climate change data for the risk assessment, as well as to determine adaptation strategies.

Variable Name	Variable Description
country	Name of Country, alphabetically sorted
year	Year of measurement or prediction (Range: 1960-2100)
population	Population of country (Range: 1960 - 2100)
real_gdp	Real annual GDP, in millions of USD (Range: 1960 - 2020, 2030, 2035, 2040, 2050)
gdp_per_capita	GDP per capita, in USD (Range: 1960 - 2020, 2030, 2035, 2040, 2050)
agri_percent_gdp	Agriculture, forestry and fishing, in percentage of GDP (Range: 1960 - 2020)
exp_percent_gdp	Exports of goods and services, in percentage of GDP (Range: 1960 - 2020)
ind_percent_gdp	Industry (including construction), in percentage of GDP (Range: 1960 - 2020)
manu_percent_gdp	Manufacturing, in percentage of GDP (Range: 1960 - 2020)
serv_percent_gdp	Services, in percentage of GDP (Range: 1960 - 2020)

3. The Data

Dataset 4 : Emissions & Energy Consumption

This dataset includes data on CO2 emissions (annual, per capita, cumulative and consumption-based), other greenhouse gases, energy mix, and other relevant metrics for a number of countries and years. This is based on territorial emissions, which do not account for emissions embedded in traded goods.

Variable Name	Variable Description	Possible Risks
country	Name of Country, alphabetically sorted	
iso_alpha3	ISO 3-letter country code	
region	Region country belongs to	
subregion	Sub-region country belongs to	
year	Year of measurement (1960-2019)	
co2	Annual production-based emissions of carbon dioxide (CO2), measured in million tonnes	
co2_growth_abs	Annual growth in production-based emissions of carbon dioxide (CO2), measured in million tonnes	
co2_per_capita	Annual production-based emissions of carbon dioxide (CO2) from other industry sources, measured in tonnes per person	Decrease of the ocean's pH (raising its acidity); energy imbalance causing rising temperatures (Source)
share_global_co2	Annual production-based emissions of carbon dioxide (CO2), measured as a percentage of global production-based emissions of CO2 in the same year	
co2_per_gdp	Annual production-based emissions of carbon dioxide (CO2), measured in kilograms per USD of GDP	
total_ghg	Total greenhouse gas emissions including land use change and forestry, measured in million tonnes of carbon dioxide-equivalents	Tipping the Earth's energy budget out of balance, trapping additional heat and raising the Earth's average temperature (Source)
ghg_per_capita	Total greenhouse gas emissions including land use change and forestry, measured in tonnes of carbon dioxide-equivalents per capita	
primary_energy_consumption	Primary energy consumption, measured in terawatt-hours per year	
energy_per_capita	Primary energy consumption per capita, measured in kilowatt-hours per year	Rising global environmental temperatures (Source)
energy_per_gdp	Primary energy consumption per unit of gross domestic product, measured in kilowatt-hours per USD	

3. The Data

Dataset 5: ISSP Survey

The ISSP Environment III survey (2010) measures respondents attitudes towards environmental protection and their preferred government measures for environmental protection.

Variable Name	Variable Description
country	A respondent's country
exag_eng	Many of the environmental issues are exaggerated. 3 (Strongly Disagree) - 15 (Strongly Agree)
climate_crisis	Climate issues caused by modern lifestyle are extremely dangerous for the environment. 7 (Strongly Disagree) - 35 (Strongly Agree)
protect_env	My daily behavior is environmentally friendly (e.g. no use of plastics). 3 (Strongly Disagree) - 12 (Strongly Agree)
sex	A respondent's sex 0 (Male), 1 (Female)
age	A respondent's age
educyrs	A respondent's years of education
degree	Whether a respondent has a university degree or not Uni (Yes), noUni (No)
rank	A respondent's social status in their country
gov_peo	Government or ordinary people: who decides how to protect environment? 0 (People should decide for themselves how to protect the environment), 1 (Government should regulate laws to make them protect the environment)
gov_bus	Government or business: who decides how to protect environment? 0 (Businesses should decide for themselves for environmental protection), 1 (Government should regulate laws to make them protect environment)
how_bus	How should we get businesses to protect environment? 0 (Heavy fines for damaging the environment), 1 (Tax benefits to reward businesses), 2 (More education about benefits of environmental protection)
how_peo	How should we get people to protect the environment? 0 (Heavy fines for damaging the environment), 1 (Tax benefits to reward people), 2 (More education about benefits of environmental protection)
mem_env	Are you a member of any group aiming for environmental protections? 0 (No), 1 (Yes)
care_env	Overall, do you care about environmental issues? 0 (No), 1 (Yes)

4. Inspiration

Where to start?

To successfully tackle the challenge, you should combine solid data analysis with strategic insight. On which areas you focus is completely up to your team. Here, we have listed a couple of starting points to solve the challenge. You can also find a code sample in the drive to help you get started.

Note that these are just guidelines. You can solve the challenge any way you want.

Focus on one stakeholder

Your team can consider any of these (and other!) stakeholders and build a solid case for them. Think about what is important for these stakeholders. Why should they care about climate change, how are they impacted by it and what potential do they have to develop a response? Once you identified these points, start looking for answers in the data we provided. Don't forget to clearly communicate your stakeholder to the judges!

Households

People all over the world will be affected by climate change. How can people become more resilient and prepare for the future?

Policymakers

National and global policymakers have a responsibility to protect people from the adverse effects of climate change. What adaptation policies should they implement?

Nonprofit organisations

Many NGO's fight to save people, animals and ecosystems. Can you recommend them areas to focus on and determine which adaptation initiatives to take?

Industry

The economic effects of climate change could be devastating (or opportune) for certain industries. Which player(s) will be most affected and how should they respond?

Farmers

Agriculture is a sector that will be heavily affected. What investments should farmers make, how can they protect their livelihoods, and which crops should they grow?

Someone else

Feel free to come up with your own stakeholder!

EMERGENT

DATA 4 GOOD CHALLENGE

4. Inspiration

How to approach the problem?

Structure your approach

Instead of analyzing all aspects of the problem, it often makes more sense to dive deep into one or two crucial aspects of the problem. What follows is a non-exhaustive list of aspects of the problem coupled with some of the ways you can tackle that aspect of the problem.

Note that how you implement these aspects might differ based on the stakeholder you have chosen. Always remember to link back your research to the end user of your recommendations!

1. Climate change risk assessment

A first part of the challenge is assessing climate change risk for countries, in terms of temperature, weather conditions, sea level rise, and more. Which characteristics (demographic, geographic, economic...) make countries vulnerable? Which countries are plagued by which risks? How do these risks affect the stakeholder you have chosen?

Suggestion for quantifying effects: make a distinction between transitional risks and physical risks. Transitional risks are costs associated to the adaptation to climate change like for example the costs for farmers to start growing new types of crops due to less precipitation in a region. Physical risks are costs associated to the physical destruction of assets like for example the collapse of buildings due to floods. Note that you should make a lot of simplifying assumptions to be able to quantify these effects. See page 13 for more inspiration on quantifying effe

At this point, it might be a good idea to narrow down the scope of your research (e.g. limit yourself to a certain region or a set of countries with similar characteristics/risks)

4. Inspiration

How to approach the problem?

2. Adaptation strategies

Now that you have a set of countries that are at risk, and the characteristics that make those countries vulnerable, it is up to you to propose strategies to adapt to climate change. This could be in terms of building infrastructure, migration, investment in technology, or whatever creative solutions you can come up with.

The key is in linking your adaptation strategies back to the data and to the risk assessment you made in the previous step.

You can start from the available body of research surrounding adaptation. Note: you have limited time, so choose wisely which one you will consult:

- What do adaptation to climate change and climate resilience mean?
- AR5 Climate Change 2014: Impacts, Adaptation and Vulnerability, summary for policymakers
- Climate Change Adaptation Strategies for Local Impact - Key Messages for UNFCCC Negotiators
- Various approaches to long-term adaptation planning
- Databases - Adaptation and resilience
- National Adaptation Plans
- Climate Change Laws of the World

4. Inspiration

How to approach the problem?

Frameworks for quantifying effects

These frameworks could be a structural guide to help you quantify climate change risks based on the data provided to you and to structure your adaptation response.

A five-step process for quantifying transitional and physical risks

Observed phenomena	Variables X and Y in the datasets change in direction Z
Direct effects on the climate	Describe what immediate effect these changes cause (e.g. in terms of weather changes, biodiversity, living conditions)
Indirect effects	Describe what indirect effects these changes cause (e.g. on economic activity, on people, on migration)
Type of risks	Try to categorize and structure all the identified risks (e.g. technological risks, market risks, demographic risks)
Response	How can we adapt to and plan for these risks?

Three drivers of physical risks

Hazard (acute/chronic)	What is the probability of dangerous climate change events? (e.g. floods)
Exposure	Risk associated to each facility of being exposed
Vulnerability	The sensitivity to the hazard in terms of physical, societal and economic factors which includes actions taken to reduce or adapt to the hazard; this could be referred to as a net risk value

5. Submission

What is the output? What are the rules?

Create a pitch deck

You will use a slide deck to pitch your solution to the jury. You are free to use any software you want (PowerPoint, Canva...) to create the presentation, but the result must be a file in .pdf format.

Your pitch can be a maximum of 6 minutes long. An Emergent member will be present to signal the time you have left and to cut you off if you go over the time limit.

After the pitch, the jury members will have some time to ask questions. It's a good idea to think about which points of discussion might come up (e.g. limitations of your analysis).

Submit your work by 6:00 PM

At 18:00, you will need to submit everything you worked on. Please collect all your work, put it in a zip file, and upload it. The following needs to be submitted (if applicable):

- Your slide deck/presentation (in .pdf format)
- Your python/R/other code (notebooks, scripts, or other types of files, as long as they are editable with freely available software)
- Your BI tool files (Tableau, Power BI...)
- If you used cloud software for which you can not download the underlying analysis (e.g. Google Data Studio), or code that can not be opened with freely available software, please take screenshots of everything and upload them as well.

You will not be able to change anything about your submission (e.g. correcting typos, adding/removing slides, fixing bugs) after uploading. The slide deck you submitted will be the one used for your pitch, and it will be pre-downloaded to the pitch room computer.

CLICK TO SUBMIT

Rules & anti-cheating measures

Breach of the following rules may result in disqualification from the challenge:

- When using additional resources, figures or data, you must specify a source.
- Every statistic or figure that can be calculated based on the provided data has to be created by the team. You may not copy (partial) solutions from the internet.
- You may not use any data visualisation found online (All data visualisations or analyses should be fully yours)

Anti-cheating measures have been implemented during the creation of the challenge. After months of working on this challenge, we have gotten a very good idea about what can and cannot be found online. Cheating will not be tolerated and results in immediate disqualification from the D4GC and all future Emergent events.

6. Evaluation & Prizes

Which prizes can we win? How do we win them?

The solutions that you present will be judged on many different aspects. The exact details about how you will be judged will not be shared with the participants. Nonetheless, we have some guidelines that can help you win one of these prizes!

Important notes: the jury consists of both technical and non-technical people. Make sure you communicate in a way that both understand. Consider also that the jury might have only a limited understanding of the topic of the challenge.

Best Data Visualisation - €250

Convince the jury by creating stunning visualizations that contribute to the case you are trying to build. Make sure to present the insights you gained in a clear and understandable way, and link it to impact.

Best Pitch - €250

Focus on delivering a pitch that can convince stakeholders of your solution. Clearly communicate in a structured and comprehensive way. When one of the judges asks questions, answer them in an insightful way.

Best Technical Solution - €250

Your solution should be a technical solution that uses the most appropriate techniques to successfully tackle this challenge. You should not be afraid of using more advanced methods as long as they are the right tool for this problem.

Best Strategy - €250

Develop a solution that is feasible, valuable and tackles the problem or opportunity you identified in a complete way. Don't forget to clearly communicate what your solution tries to achieve and why it is valuable.

Best Overall - €1000

Find a balance between all the points above and convince the jury that your solution is the solution your stakeholder(s) need(s).

7. FAQ

Frequently asked questions.

Do we have to use all of the available datasets?

No, it's up to you to decide what data to work with. You can also decide to work only with a specific region/country/countries.

Do we have to use a particular software?

No, you can use whatever software(s) you would like to.

Can we use data from the internet?

Yes, in this case you must provide the source of the dataset(s) you used.

Can we use information from the internet?

Yes, in this case you must provide the source of the information you used.

Do we need to prepare a slideshow presentation?

Yes, you need to prepare a slideshow presentation and upload a PDF version of it along with your code/BI tool files. The submission link can be also found at dataforgood.be/overview.

Does everyone from the team need to present to the jury?

No, you may decide among yourselves, but at least one person should present.

How much time do we have to present?

You have 6 minutes to present your strategy and 3 minutes to answer the questions from the jury.

Do we have to do a technical analysis?

No, it is not necessary. That is up to you.
You can opt for a purely strategic approach, of course grounded in data.
If you want to include a technical approach, you can. But, keep in mind that your presentation time is limited to 6 minutes only.

8. References

Costal Population at Risk. Dataset accessed [2021-10-25] at http://sealevel.climatecentral.org/

FAO. 2018. The future of food and agriculture – Alternative pathways to 2050. Rome.

Fox-Kemper et al., (2021), Ocean, Cryosphere and Sea Level Change. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. In press.

Gapminder (v6), HYDE (v3.2), United Nations Population Division (2019)

Garner, G. et al., (2021). IPCC AR6 Sea-Level Rise Projections. Version 20210809. PO.DAAC, CA, USA. Dataset accessed [2021-10-25] at https://podaac.jpl.nasa.gov/announcements/2021-08-09-Sea-level-projections-from-the-IPCC-6th-Assessment-Report.

Garner, G. et al., in prep. Framework for Assessing Changes To Sea-level (FACTS). Geoscientific Model Development.

GESIS, ed. 2019. ISSP 2010 Environment III, Variable Report: Documentation release 2019/06/13, related to the international dataset Archive-Study-No. ZA5500 Version 3.0.0, Variable Reports 2019|05. Cologne: GESIS.

Hannah Ritchie and Max Roser (2020) - "CO_2 and Greenhouse Gas Emissions". Retrieved from: 'https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions'

Kulp, S.A., Strauss, B.H. New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. Nat Commun 10, 4844 (2019). https://doi.org/10.1038/s41467-019-12808-z

OECD (2018), GDP long-term forecast (indicator). doi: 10.1787/d927bc18-en (Accessed on 29 October 2021)

OECD (2021), Access to computers from home (indicator). doi: 10.1787/a70b8a9f-en (Accessed on 28 October 2021)

Swart et al., (2020). CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP. Version 20150325. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.1317

World Development Indicators, The World Bank

https://github.com/owid/co2-data/blob/master/owid-co2-codebook.csv

EMERGENT

DATA 4 GOOD CHALLENGE