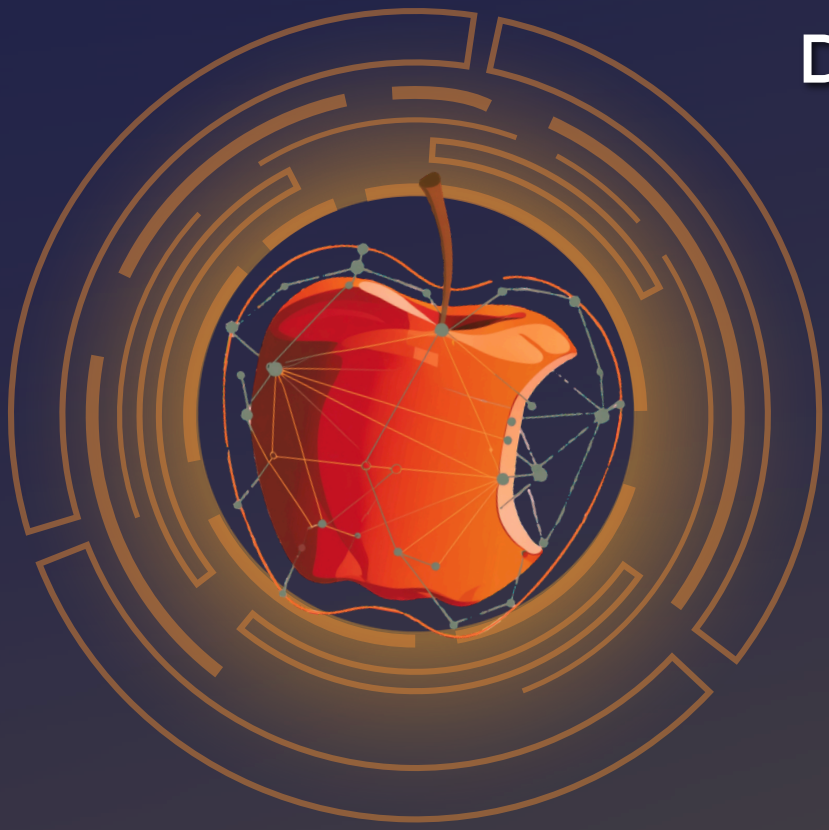
EMERGENT

NOVEMBER 20th

DATA 4 GOOD CHALLENGE

DOCUMENTATION



aws KBC McKinsey & Company RISK CONCILE

FOOD SAVING Since 2016 Leuven editx leuven

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

Crunching data: Shaping a Sustainable Food Future

In today's world, addressing food waste is not just a choice—it's a necessity. Environmental and societal challenges call for innovative solutions, and many organizations and volunteer initiatives have risen to meet this demand. These groups work tirelessly to reduce food waste and redistribute surplus food to underserved communities. However, many still rely on manual processes and lack access to data science and automation tools that could streamline their efforts and reveal process inefficiencies.

This year, we are excited to collaborate with two key organizations—Food Saving Leuven and Leuven Stad —under one mission: "Saving Food". Together, we aim to optimize food-saving initiatives by integrating data-driven solutions to maximize impact. This collaboration enables the winning teams to further develop their innovative ideas alongside these organizations, amplifying their impact beyond the event.

The focus is clear: Develop a data-driven strategy to identify and address bottlenecks in the processes of the Food Saving Group, providing actionable solutions and uncovering new opportunities to optimize their food-saving initiatives. How will you achieve this mission? By harnessing the power of data storytelling and creating a state-of-the-art solution that drives meaningful change.

Data is your most potent tool. You have access to a wealth of valuable resources, including five years' worth of data on food pick-ups in Leuven, as well as the latest comprehensive list of all restaurants, supermarkets, and bakeries in the area. Through our collaboration with Leuven Stad, you also have access to key metrics and indicators on demographics, environmental factors, and wealth distribution across Leuven.

Your objective is to craft a problem statement that highlights and addresses critical aspects such as streamlining food pick-up procedures, encouraging the growth of partnerships and volunteers through market analysis, and even identify potential hubs for food drop-off points. With data at your fingertips, you will explain your findings, explore innovative strategies, and provide a roadmap for change.

Join us in "Crunching data: Shaping a Sustainable Food Future". Together, we will pioneer innovative solutions that reduce food waste, optimize distribution, and create a sustainable and organized system that ensures no food goes to waste.

3. About our Challenge Partners

Meet our challenge partners for this D4GC edition:

Food Saving Leuven

Food Saving Leuven (FSL) is a volunteer-driven organization dedicated to reducing food waste in Leuven. Founded by a group of friends in 2016, FSL has since grown to over 100 active members. These volunteers collect surplus food from six partner supermarkets on an almost daily basis—food that would otherwise be discarded. In 2024 alone, FSL has already saved more than 2,800 kilograms of food. The recovered food is then distributed at designated "drop-off" points, where it is freely available to everyone, not just FSL members.

Leuven Stadt

Stad Leuven is the governing body of Leuven, dedicated to the sustainable development and digital transformation of the city. As a forward-thinking municipality, it spearheads various initiatives aimed at enhancing the quality of life for its citizens through innovation and data-driven solutions.

Leuven Smart City

Leuven's Smart City initiative is designed to leverage technology to create a more connected, efficient, and sustainable urban environment. The project integrates intelligent systems and data management to optimize city services, enhance mobility, and promote citizen engagement. More information can be found here.

Data and GIS

Stad Leuven's Data and GIS project supports city services by providing integrated file handling and geospatial data management. Following a 2019-2020 feasibility study, the system now connects all applications through a geo app for both citizens and employees, using data from authentic sources. It also provides visualizations and dashboards through Power BI to enhance decision-making.

4. The Data

Dig in & gain some insights!

In the Data 4 Good Challenge, we have curated a diverse and comprehensive collection of data sources to empower participants in developing a data-driven waste management strategy. Our data resources are divided into three categories, each offering unique insights and perspectives on food waste and demographics across Leuven. This multifaceted data environment will serve as the foundation for participants to craft compelling narratives and innovative solutions.

Dataset 1: Food Pickup Trends in FSL Platform (2020-2024)

This dataset, scraped from the Food Saving Leuven platform, comprises five years of pickup history from various stores in Leuven and detailed descriptions of food saved at each pickup. It also includes anonymized information about the members activity and a "trust score" given to each within the community.

Dataset 2: Spatial Dataset of Amenities in Leuven (2024)

The dataset, scraped from OpenStreetMap (an open-source community-created geospatial project), contains an up-to-date list of restaurants, bakeries, coffee shops, grocery stores, and supermarkets within an 8 km radius of Leuven-Centrum. For each amenity, the dataset provides details about its type, as well as additional attributes such as cuisine and specialty. It also includes comprehensive information on the amenities' addresses, coordinates, and statistical sectors.

Dataset 3: Key Metrics and Indicators in Leuven (2021/2023)

Metrics on demographics, income, commercial spaces, and housing across statistical sectors in Leuven were gathered from the Leuven In cijfers platform.

5. 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 to the problem coupled with some of the ways you can tackle that aspect of the problem.

Begin by organizing your team's efforts and responsibilities. Allocate tasks based on individual strengths and expertise to ensure a cohesive approach to the challenge.

1. Problem Statement

Clearly define the specific food management issue you aim to address. This could be related to food pickup procedures, partnership growth and engagement, or potential hubs for food pickups. A well-articulated problem statement sets the foundation for your solution.

Your statement should be formulated concisely. You should deliver your "solution" to the jury which consists of a mix of different profiles and hence it should be understood easily even by non-technical background personnel.

2. Decide a direction

After having built your problem statement, it might be a good idea to narrow down the scope of your research. Consider the data at your disposal and decide on the direction of your analysis. Are you focusing on predictive modeling, benchmarking, or a combination of both? Define a clear path to guide your research.

Don't forget that you are given only 4 hours to not only build your solution, but also prepare a slide deck for a presentation.

3. Develop a DATA-DRIVEN solution

We recommend spending around 1-2 hours completing the first two steps. Now, you should carefully analyze the given data to support (or modify if needed) your statement or prove whether your hypothesis is correct or not. Utilize the insights gathered from your analysis to develop a DATA-DRIVEN waste management strategy. Leverage the power of visualization, and storytelling to create an impactful solution that addresses the identified problem and engages users effectively.

5. 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. We have prepared a code sample (both in Python and R).

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

Use case 1: Identifying bottlenecks in the food pickup procedure

What challenges can be identified from the history of food pickups? What are the key parameters that contribute to a successful pickup? Specifically, how do factors such as the weight of the saved food, the time of day, the day of the week, and the number of available slots influence the success of a food pickup? What strategies can Food Saving Leuven adopt to hold volunteers accountable for their roles in food collection, drop-off, and distribution?

Use case 2: Finding new hubs for food drop-off

Currently, the Food Saving Leuven group operates a single hub located at the Agora Learning Center, Edward van Evenstraat 4, Leuven. How can a cross-analysis of Leuven's list of amenities, demographics, and income data help identify underserved regions? Can this approach help identify potentially new places that could allow the set up of new hubs based on their attributes (size, location, type) to serve their neighborhood?

Use case 3: Identifying possible new places to partner with

Can the food pick-up data reveal untapped regions in Leuven that could benefit from food saving initiatives? Can combining this data with the amenities data and demographics data reveal insights on the disparities across Leuven, and identify new places with potentially higher risk of food waste?

5. Inspiration

Interesting sources to look at

Belgium second largest food waster in Europe

The Belgian food industry wastes 161 kg of food per inhabitant per year, compared to 59 kg in the Netherlands, 29 kg in France and 19 kg in Germany. Belgium is almost at the bottom of the European ranking with Cyprus only doing even worse.

<LINK>

How One Belgian City Used Data Analytics to Reduce its Food Waste by 30%

The City of Roeselare, with support from the province of West Flanders, has made significant strides in reducing food waste, achieving a 30% reduction. Key participants, including Hospital AZ Delta and residential care centers Sint Henricus and De Hovenier, collectively saved 34,500 kg of food and over €70,000 in costs within a year.
Their success involved adapting meal offerings to better suit diverse patient needs and utilizing data analytics to manage meal cancellations for fasting patients.

<LINK>

FOODWIN

EMERGENT

DATA 4 GOOD CHALLENGE

6. 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
- Your BI tool files (Tableau, Power BI...)
- If you used cloud software for which you can not download the underlying analysis (e.g. AWS), 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.

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.

EMERGENT

DATA 4 GOOD CHALLENGE

7. 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 - €500

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 Business Strategy - €500

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).

EMERGENT

DATA 4 GOOD CHALLENGE

8. FAQ

Frequently asked questions by participants.

Do we have to use all of the available datasets?

No, it's up to you to decide what data to work with.

Do we have to use a particular software / product?

No, you can choose either Python or R for your analysis. You can of course use both if needed. Additionally you're allowed to use whatever other software you find necessary or useful

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.

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 4 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.