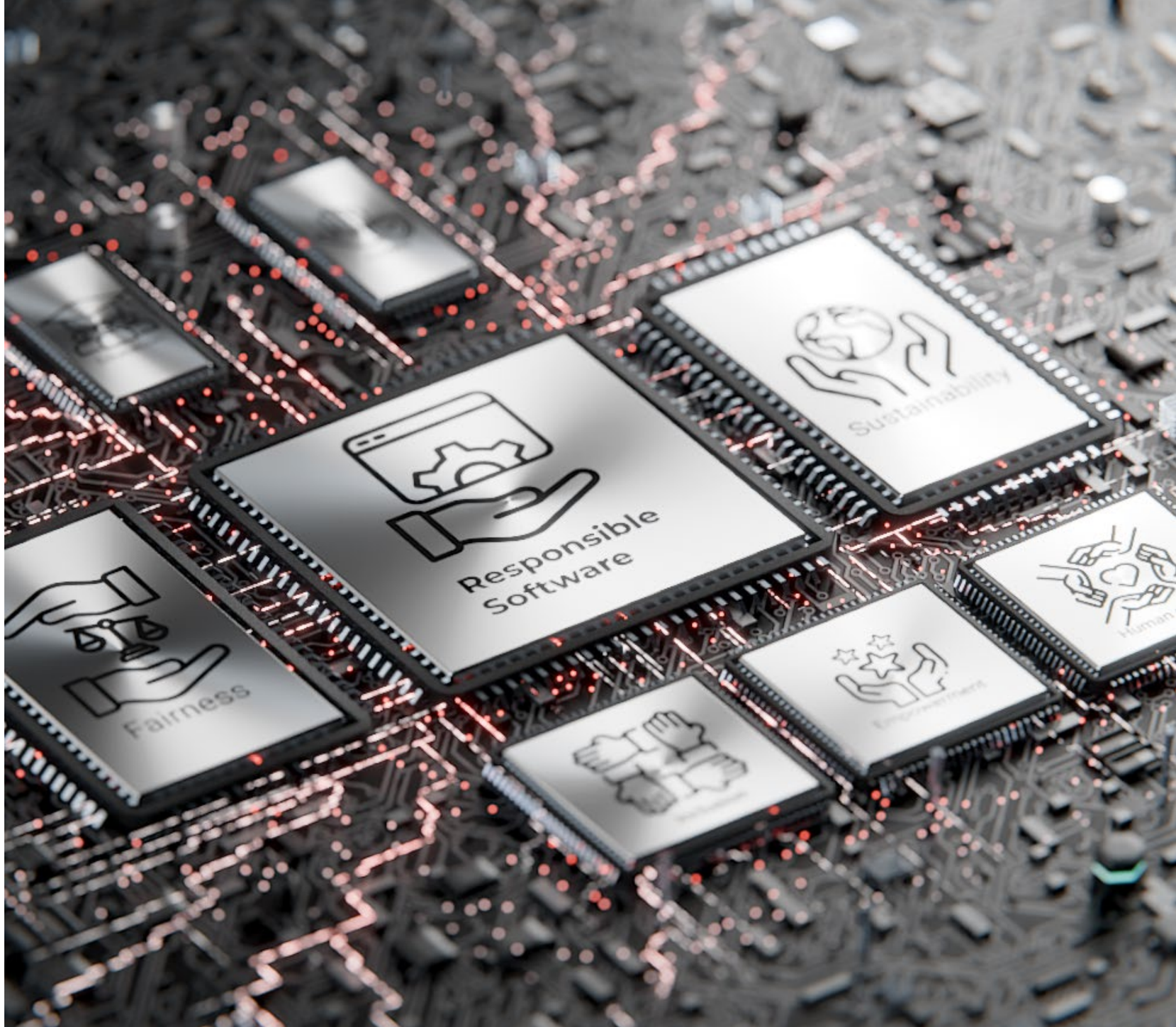
EPFL

Graded 2 Debriefing

1 dec.

Cécile Hardebolle

Responsible Software



Agenda for today

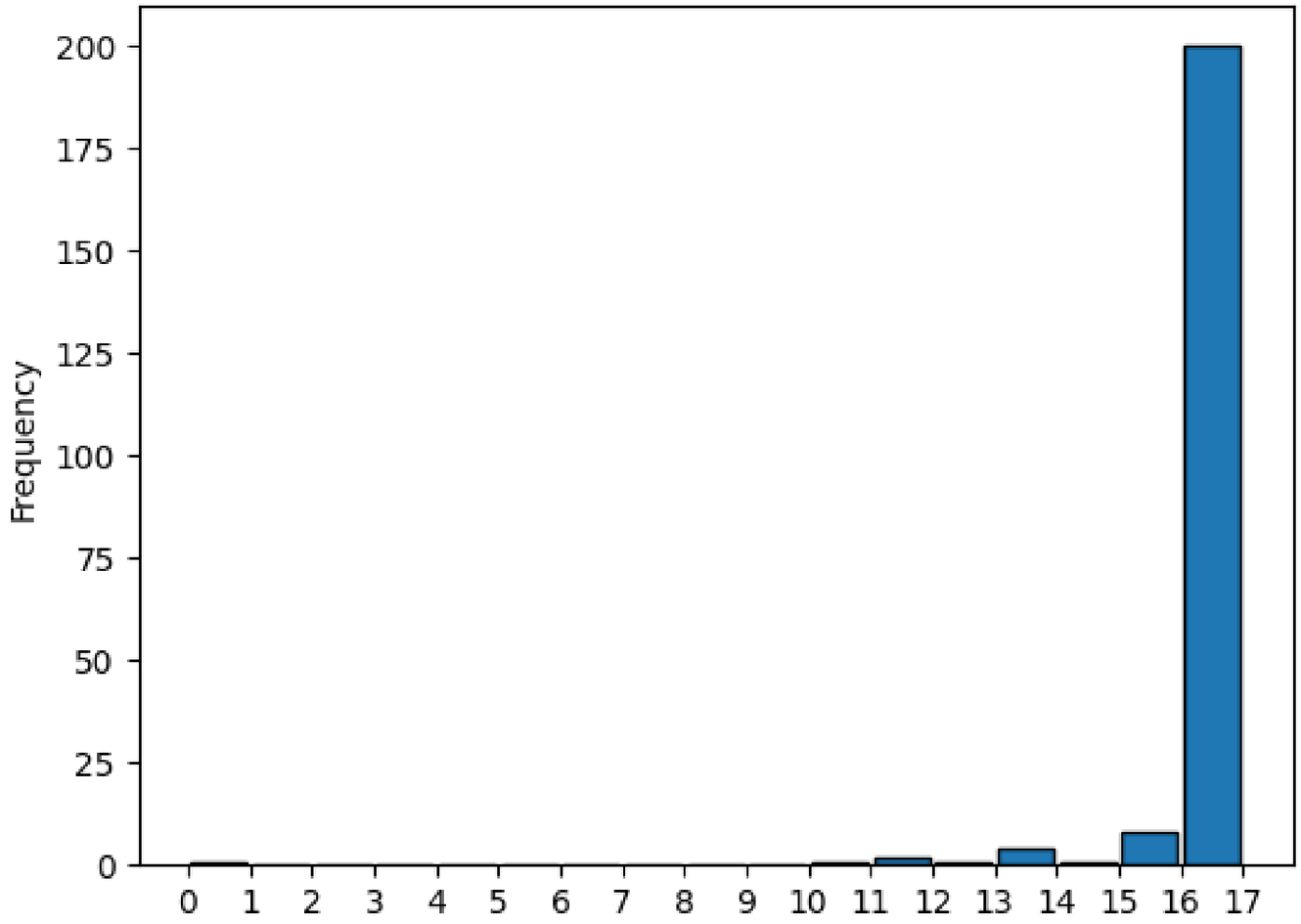
1. Groups for the Graded Case
2. Feedback on the Graded 2 notebook
3. Next dates

Graded case: groups

- For those who have not yet signed up for their groups:
  - If you don’t know with whom to work:
    We will try to do a random assignment tomorrow
  - If you know already with whom to work:
    Send a private message on Ed,
    category Graded Assignments > Graded Case
    Indicate the names of people in your group

Feedback on Graded Notebook 2

Programming questions



218 submissions

Maximum possible:
17 points

Mean: 16.6 points
Median: 17 points
(std: 1.5 points)

Grading of open reflection questions in progress

Learning goals

- Practice with **estimating** the scale of **two different types of environmental impacts** from ML models:
  - Water footprint
  - Carbon footprint

- Identify:
  - **Most significant factors** in both footprints
  - **Scaling factors** for the carbon footprint of LLMs (but applies to water…)
  - **Mitigation options** and the stakeholders who can implement them

Feedback on Graded Notebook 2

Exercise 1:

The Water Footprint of AI in Africa

1.1.1 Countries and cities

Compute the following values:

- **How many different countries are there in the dataset?**
- **How many different cities are there in the dataset?**
- **How many unique timestamps are there in the dataset?**

Typical oversights:

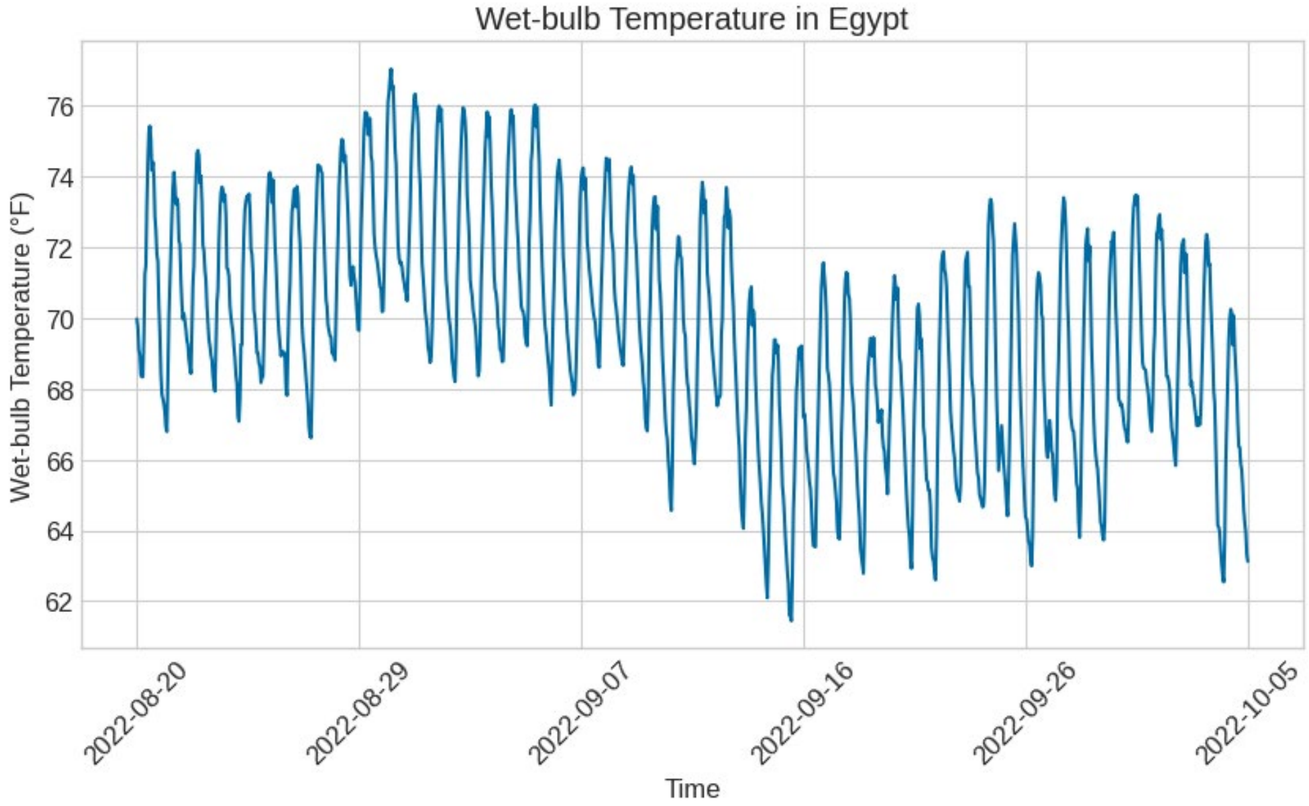
- Hard coded (wrong!) values
- Typos

1.1.2 Wet-bulb temperature

The dataset contains a measure of the wet-bulb temperature for each timestamp in each city.
We would like to average the wet-bulb temperature (wetbulb_temperature) at the level of a given country for each timestamp.

Complete the function below.

Typical oversights: none!



Wet-bulb temperature

URL: ttpoll.eu
Session ID: cs290

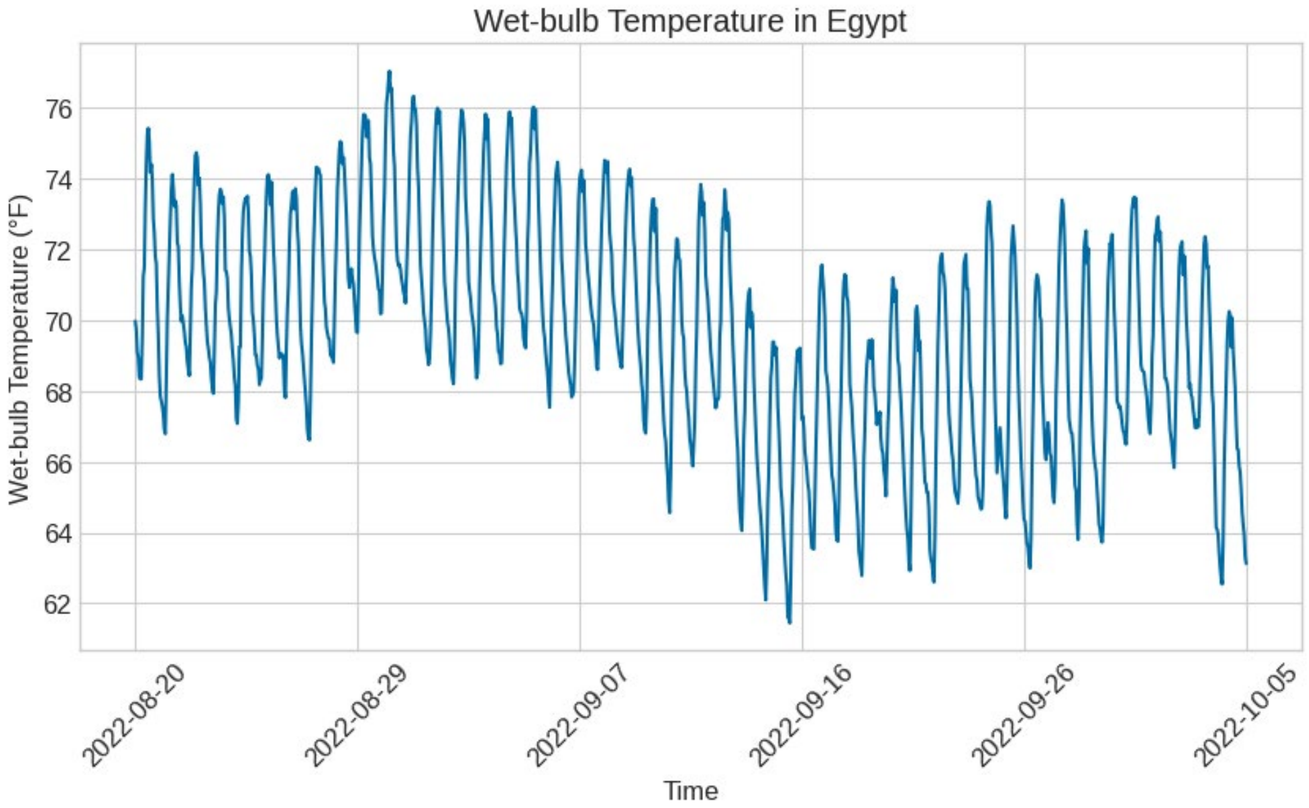
Which factors can affect the wet-bulb temperature as observed on the graph? (Select all that apply)

(✅ %) a. Air pressure

✅ 5% b. Humidity

✅ 8% c. Sunlight

✅ 3% d. Wind



- Main factors: sunlight, humidity, wind
- Air pressure plays a role (as it affects condensation/evaporation) but way smaller

1.2.1 Calculate water consumption

Implement the calculate_water_consumption function, using the equations above.

- Onsite water footprint: $W_{\text{on}} = \gamma_{\text{on}} \cdot E$
- Offsite water footprint: $W_{\text{off}} = \gamma_{\text{off}} \cdot \rho \cdot E$
- Total water footprint: $W_{\text{total}} = W_{\text{on}} + W_{\text{off}}$

Where:

- γ_{on} is the onsite Water Usage Effectiveness (`wue_onsite`) in L/kWh.
- γ_{off} is the offsite Water Usage Effectiveness (`wue_offsite`) in L/kWh.
- ρ is the Power Usage Effectiveness (PUE) of the data center, a dimensionless ratio.
- E is the server energy required for the specific computing task in kWh.

Typical oversights: none!

1.3.1 Water footprint of training - function

Complete the function below that returns the total water consumption for a computing task given the electricity consumed by that task in a given city.

You will need to compute:

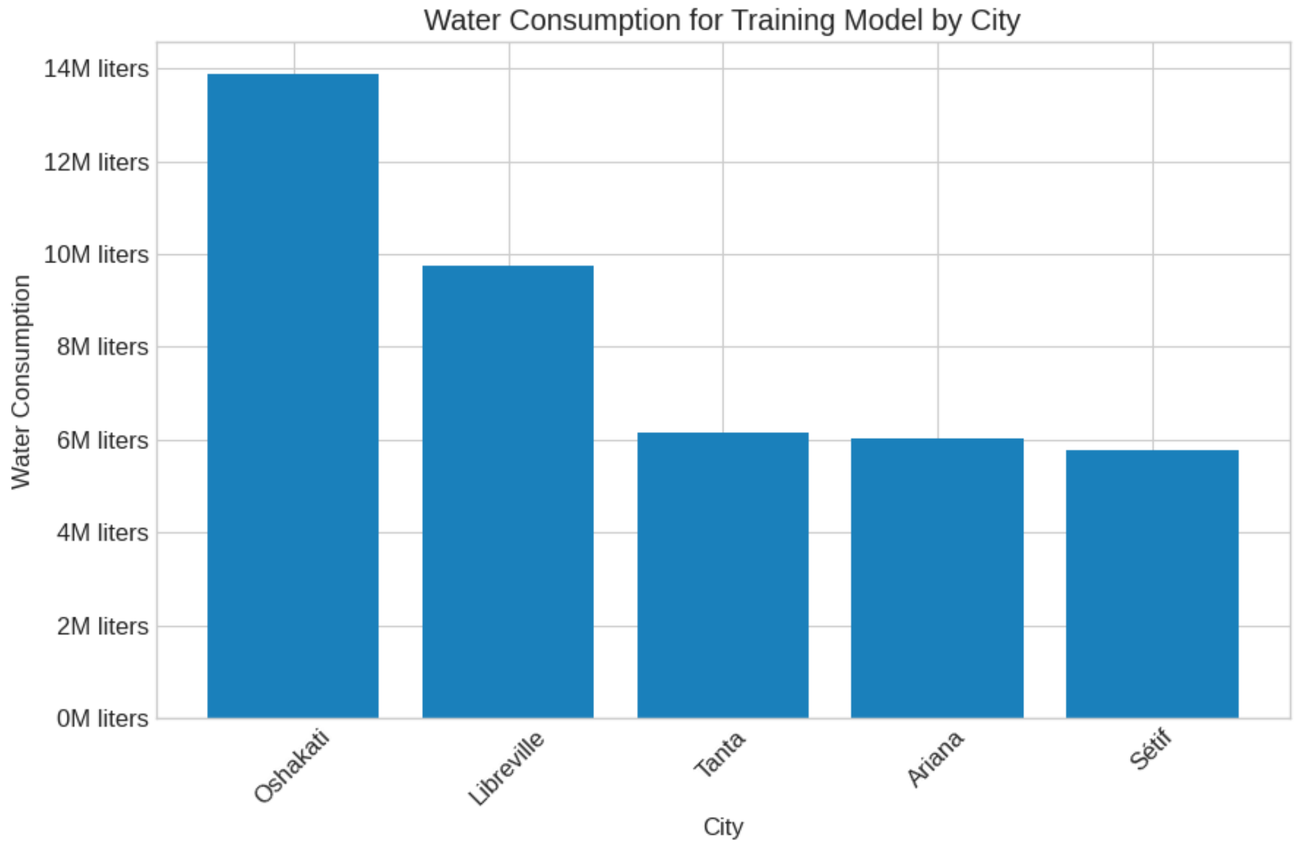
- **The average onsite WUE (L/kWh) for this city (computed over all timestamps),**
- **The average offsite WUE (L/kWh) for this city (computed over all timestamps),**
- **The PUE for this city.**

Based on this data, you will compute the total water consumption (in liters) corresponding to the electricity consumed by the task.
NB: of course you can reuse the function you have created in the previous section.

Typical oversights: none!

1.3.2 Water footprint of training - cities

Complete the code below to compute the water consumed for training this model in the cities considered above.



Typical oversights:

- Using 1kWh instead of the electricity consumed for the task

1.3.3 Stakeholders

After presenting your results, Tam tells you that they still decided to build the new data centers in Oshakati, as the cost of construction in Namibia is lower than in the other countries.

Identify 2 different stakeholders affected by this decision and for each (1 point / stakeholder):

- **Name the stakeholder**
- **Identify 1 benefit for this stakeholder (1 sentence, 0.5 point)**
- **Identify 1 harm for this stakeholder (1 sentence, 0.5 point)**

Typical oversights:

- Not explicitly relating the benefit/harm to the stakeholder

  Example: Stakeholder: LockedAI
  Harm: Higher water use and environmental impact.

The water cycle

URL: ttpoll.eu
Session ID: cs290

How long does it take on average for evaporated water to come back to the ground as rain?

Water scarcity

URL: ttpoll.eu
Session ID: cs290

Which of the following is a cause of increasing water scarcity in many regions around the world?

❌ 0% a. The natural increase in the Earth's total water volume

✅ 76% b. Uneven demand-supply ratios for freshwater resources

❌ 18% c. A global decrease in rainfall across all climate zones

❌ 6% d. The melting of polar ice caps increasing ocean levels

+ Large impact from climate as warmer air can store larger quantities of water in vapor form -> terrestrial water storage diminishes

Water scarcity

(Bengtsson, 2010; Mekonnen & Hoekstra, 2016; Livingstone, 2023)

‘It’s pillage’: thirsty Uruguayans decry Google’s plan to exploit water supply

Country suffering its worst drought in 74 years, with government even mixing saltwater into drinking supply

People take part in a protest amid a shortage of drinking water reserves in Montevideo on 31 May 2023. Photograph: Eitan Abramovich/AFP/Getty Images

(Livingstone, 2023)

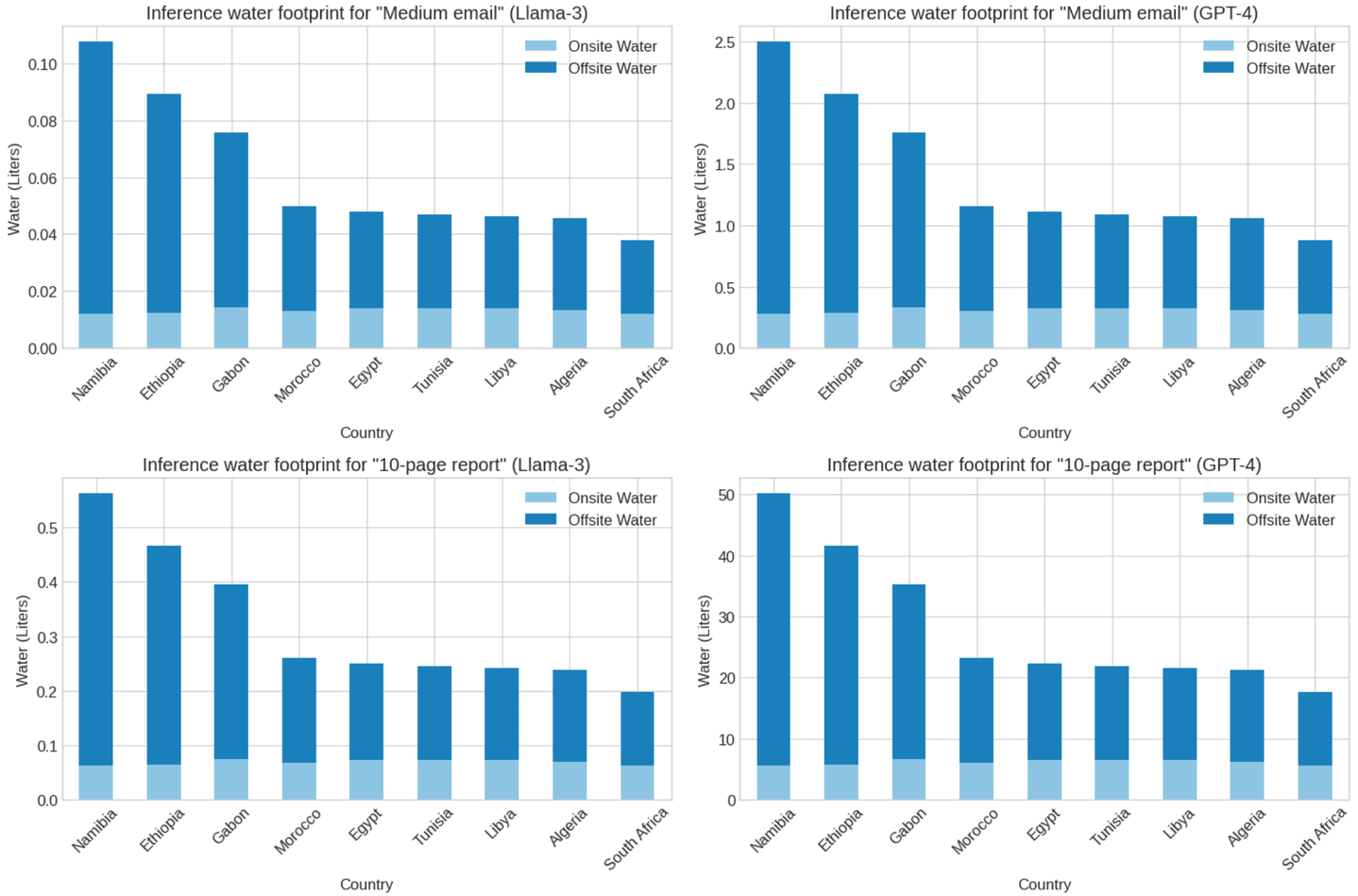
Water is never really “lost”…
…But location and time matter!

“At the global level and on an annual basis, enough freshwater is available to meet such demand, but **spatial and temporal variations of water demand and availability are large**, leading to **water scarcity in several parts of the world during specific times of the year**.”

(Mekonnen & Hoekstra, 2016)

1.4.1 Water footprint of inference - countries

Complete the code below to calculate the water consumption for each country, model, and task. Store , , and in columns named w_on_L, w_off_L, and w_total_L respectively.



Typical oversights: none!

1.4.2 Main factor in water consumption

Based on the results you just obtained, which of the two types of water consumption plays the most important role in the total water footprint of the models for these inference tasks? (1 sentence, 1 points)

What could explain the large differences between the total water footprints of the two models? Cite 1 reason and describe your reasoning (1 to 2 sentences, 1 points)

Typical oversights:

- Explaining why electricity production consumes water (types of plants) instead of explaining the difference between the two models

Feedback on Graded Notebook 2

Exercise 2:

The Carbon Footprint of ChatGPT

Feedback on Graded Notebook 2

Exercise 2:
The Carbon Footprint of ChatGPT

2.1 - Carbon footprint per token per region

2.1.1 Carbon footprint per token

Write a function that computes the carbon footprint per token of using ChatGPT in g CO2e.
Complete the cell below.

Pay attention to your units.

Typical oversights:

- Hard coded values (e.g. model speed)
- Forgot to convert to hours

2.1.2 Footprint per datacenter

For each datacenter in the datacenter_efficiency.csv dataset, compute the corresponding carbon footprint per token and add the result as a new column Footprint gCO2e/token.
Complete the cell below.

Typical oversights:

- Power of 1 GPU **in W instead of kW**
- Typos

2.1.3 Location

From the datacenter_eff dataframe, identify which geography (country name) has the lowest and highest footprint per token and provide the value of the footprint for each.
Complete the cell below.

Typical oversights:

- Give the *index* instead of the *name* for the geography

Feedback on Graded Notebook 2

Exercise 2:
The Carbon Footprint of ChatGPT

2.1 - Carbon footprint per token per region
2.2 - ChatGPT usage for an individual student

2.2.1 Queries & tokens

Your goal is to:

- **Create a totalTokens column keeping track of the whole number of tokens used for each interaction.**
- **Create a queries column that indicates if the message corresponds to a query done by the user, i.e. queries = 1 when the message is an input and 0 otherwise**

Complete the cell below.

Typical oversights:

- Incorrect sum (summing the whole columns instead of summing by row)

2.2.2 Usage analysis

Let's compute some statistics:

- **total_number_queries : total number of queries made by the student during the time period**
- **total_number_token : total number of tokens used in the interactions during the time period (input, output and reasoning)**
- **days_with_activity : number of days the student used ChatGPT**
  (Hint: some dates appear multiple times, and not every day is represented so you might want to check pandas documentation to find unique days.)
- **average_query_day : average number of queries per day the student was active**
- **average_query_size : average number of tokens used per query**
- **total_reasoning_tokens : total number of extra tokens coming from model**
- **reasoning total_reasoning_tokens_prop : proportion in percentage (value between 0 and 1 - watchout) that the tokens generated for reasoning represent over the total number of tokens**

Complete the cell below.

Confusion on average_query_size
-> **not counted** i.e. 6 variables instead of 7 to get 1 point

Typical oversights:

- Incorrect calculation for days with activity
- Incorrect calculation for average query per day

2.2.3 Total carbon footprint

Compute the total carbon footprint for this student (in gCO2e):
- **for the geography with the lowest footprint per token**
- **for the geography with the highest footprint per token**

Complete the cell below.

Typical oversights:

- **Hard-coded numerical values** for footprint factors (lowest/highest)
  ⚠️ ⚠️ ⚠️ multiplied by huge number of tokens -> **large error**

2.2.4 Temporal usage patterns

In the code cell below, we add a new column is_weekend to the dataset to indicate for each row if it corresponds to a weekend (True) or to week day.
Compute the total number of tokens used by our student on weekdays and on weekends.

Typical oversights:

- Wrong column name

Confusion on the form of the table
-> evaluation only on the column with sum of tokens

2.2.5 Temporal patterns

Compute the carbon footprint corresponding to the number of tokens used on weekdays and on weekends:

- **use the highest carbon footprint per token obtained earlier (highest_footprint)**
- **make sure to account for the lower carbon footprint of electricity on weekends (90% of that of weekdays)**
- **store the results in a new column Footprint gCO2e of the dataframe total_tokens_weekdays_weekends obtained earlier.**

Typical oversights:

- Forgetting the 90% footprint for weekends (half point)
- Hard coded value for footprint -> large error

Confusion on the form of the table
-> evaluation only on the footprint column
-> independent from access method

Feedback on Graded Notebook 2

Exercise 2:
The Carbon Footprint of ChatGPT

2.1 - Carbon footprint per token per region
2.2 - ChatGPT usage for an individual student
2.3 - ChatGPT usage at scale

2.3.1 Extrapolating to one year

We provide you with the number of days of the period where we have analyzed the student's usage of ChatGPT.
Compute the number of tokens annual_number_tokens_1student that the student would use over a complete year (assume a year has 365 days).
Complete the cell below (remember that you can use the usage stats we have computed earlier).

Typical oversights:

- None

2.3.2 EPFL adoption scenarios

For each of the adoption rates:
- **Compute the corresponding number of users nb_users_rate**
- **Compute the corresponding total number of tokens total_tokens_rate**
- **Compute the corresponding annual carbon footprint for the lowest_footprint and highest_footprint factors, in kgCO2e**

Complete the cell below.
Pay attention to your units!

Error in the number of students at EPFL:
14 000 instead of 18 000
(no impact on grading)

Typical oversights:
- Incorrect conversion to kg (absent or erroneous)
- Use total_number_tokens instead of annual_number_tokens_1student
- Incorrect calculation of nb of users
- (lots of typos)

Mitigation options

You have analyzed:

- The carbon footprint per token per region
- The ChatGPT usage for an individual student
- ChatGPT usage at scale for different adoption rates at EPFL

Identify mitigations that could be put in place by different stakeholders to reduce the carbon footprint in our scenario.

👉 1 post = 1 mitigation option

See on SpeakUp.
Important: link with analysis (data evidence).

Post your ideas:
https://speakup.epfl.ch
Room key: **22855**

“Reasoning” tokens

- Models with a “reasoning mode”:
  - Perform **additional steps** before generating the answer
  - Some models do this **systematically** (e.g. Qwen Thinking) whereas for ChatGPT the mode is activated based on the query (API gives option to limit this behavior)
  - Some models **provide the reasoning traces in output**, while others provide only summaries or hide the intermediate outputs

- Dilemma:
  - Higher quality in the response (not in all cases though)
  - Higher electricity consumption (more tokens, more steps)

Feedback on Graded Notebook 2

Exercise 2:
The Carbon Footprint of ChatGPT

2.1 - Carbon footprint per token per region
2.2 - ChatGPT usage for an individual student
2.3 - ChatGPT usage at scale
2.4 - Introducing a "No-AI Day" policy

2.4.1 Simulating a No-AI-Day

Simulate this reduction and compute the annual emissions (low and high) under this restriction.
Starting from your adoption_df from the previous question, add two new columns:
- **annual_low_noAI (kgCO_2e with 1 no-AI day/week)**
- **annual_high_noAI (kgCO_2e with 1 no-AI day/week)**

💡 **Hint: If ChatGPT is blocked 1 day per week, usage is reduced to 6/7 of the original.**
Complete the cell below.

Typical oversights:

- Question not done

2.4.2 Side effects

What are the potential ethical issues with implementing this "No-AI Day" policy?
Identify 2 different issues (1 point / issue) and for each:

- **indicate the type of ethical issue among safety, fairness, sustainability and empowerment (0.25 points)**
- **describe who is impacted and what is the impact (1 sentence, 0.75 points)**

👉 1 post = 1 ethical issue:

- Type of issue (safety, fairness, sustain., empow.)
- Brief description

See on SpeakUp.
Important: link between type of issue and argument.

Post your ideas:
https://speakup.epfl.ch
Room key: **39351**

2.4.3 Another possibility

Imagine that, instead of implementing a "No-AI Day", EPFL decides to deploy a lighter and more efficient version of ChatGPT on campus, reducing the per-query energy consumption by 40%.
What could be the drawbacks from such a scenario from a sustainability point of view? Identify 2 different sustainability drawbacks and describe them (1 sentence & 1 point / drawback).

👉 1 post = 1 sustainability drawback

See on SpeakUp.

Post your ideas:
https://speakup.epfl.ch
Room key: **20588**

Other info on Graded Notebook 2

Graded 2: grade release

- Release of **score** on **the week before holidays (around Dec. 17)**
- Corresponding **grades** will be released afterwards
  - Computation:
    - Max points = 28 = grade of 6
    - If points <= 14, grade = 1 + 3/14 * points
    - If points > 14, grade = 2 + 2/14 * points
  - Then grade x 0,08 to get the proportion of final grade
    - A grade of 6 gives 0,48 in the final grade
- Questions on your score:
  - Before **Sunday 11 January at 23h59** at the latest
  - Post on Ed:
    - **Private** message
    - In Questions on Grades > Questions on Graded 2

What’s next?

Next dates

	Monday (STCC Cloud C)	Tuesday (INF1 & CO5)
1 Dec – 7 Dec	Debriefing Graded 2	Empowerment 2 notebook
8 Dec – 14 Dec	Empowerment 2 cases	Graded Case
15 Dec – 21 Dec	Conclusion cases + Q&A	---

- Empowerment 2:
  - 1 notebook
  - Only **2 video** + quizzes
  - Review cases (bad actors, ethical speculation, datasheet) + review quiz
- Conclusion:
  - Review cases (digital ethics canvas, ethics canvas) + review quiz
  - Q&A 👉 Questions to post on SpeakUp

Conclusion Q&A

You can ask any question you want on the course content, the exam, etc. before 23h59 on Dec. 14:

👉 1 post = 1 question

Post your ideas:
https://speakup.epfl.ch
Room key: **53228**