DELIVERABLES :

- ☑ ~~Logbook (This doc)~~
- ☑ ~~Google slides presentation~~
  - ☑ ~~Part 1 : why the project and aspects to consider (the big picture not too detailed i guess ?)~~
  - ☑ ~~Part 2 : audit a software and present the results~~
- ☑ ~~GitHub repository containing a reproducible workflow for the estimation of the carbon footprint of your chosen software~~
- ☑ ~~CV~~
- ☑ ~~Statement of Motivation~~

Readings before the ppt

- ☑ ~~https://everse.software/RSQKit/improving_environmental_sustainability~~
- ☑ ~~Ten simple rules to make your computing more environmentally sustainable~~
- ☑ ~~Green Algorithms: Quantifying the Carbon Footprint of Computation - Lannelongue - 2021 - Advanced Science - Wiley Online Library~~

SLIDE 1

WHY environmentally sustainable computing ?

The impact of computing on our environment is not that significant compared to industries . But does that make it insignificant ? It's not.

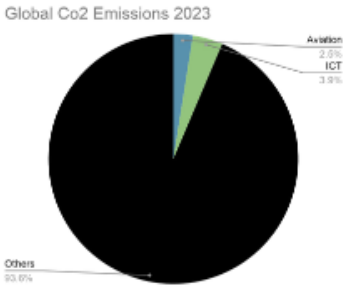
For a comparison
Aviation industry Carbon emission :

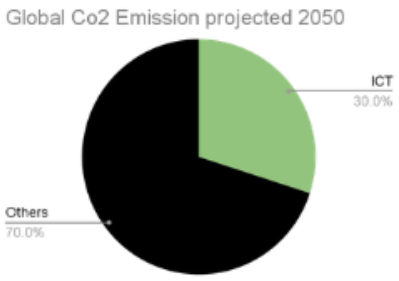
- 2023 : 2.5% of total global carbon emissions [2]

ICT (Information and Communication Technology)

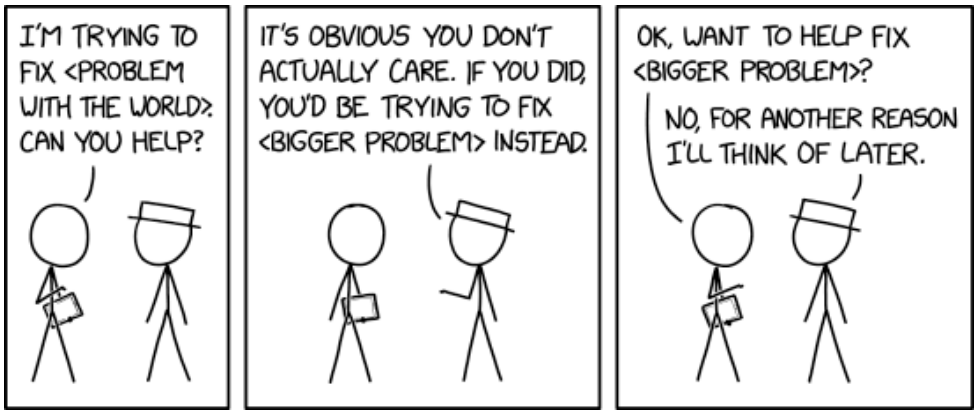
- Includes consumer devices, data centers and telecommunication services
- 2022: 1.7 - 4 % of total global carbon emissions [3]

This is expected to be increased to 30% by 2050 , imagine that
Now this seems like a real deal right ?
Plotting this on a pie chart





Here is a xkcd comics that comes to my mind



- 2022: 1.7% of total global carbon emissions [3] by ICT industry
- And it's ever increasing .
- Rise of AI and cryptocurrencies causes exponential growth of energy usage by computing .

  Annualized Total Bitcoin Footprints [4]
  Carbon Footprint
  114.03 Mt Co2
  Comparable to the carbon footprint of the Czech **Republic**.

  Electrical Energy
  204.44 TWh
  Comparable to the power consumption of **Thailand**.

  Electronic Waste
  22.59 kt
  Comparable to the small IT equipment waste of **the Netherlands**.

  Fresh Water Consumption
  3222 GL

REFERENCES:
[1] https://arxiv.org/abs/2506.14365
[2] https://www.iea.org/energy-system/transport/aviation

[3] https://openknowledge.worldbank.org/entities/publication/a8feecdd-5f79-412a-834f-6f45903e18a5
[4] https://digiconomist.net/bitcoin-energy-consumption

[5] https://docs.er.kcl.ac.uk/green-computing/

https://pollution.sustainability-directory.com/area/core-principles-sustainable-computing/resource/8/

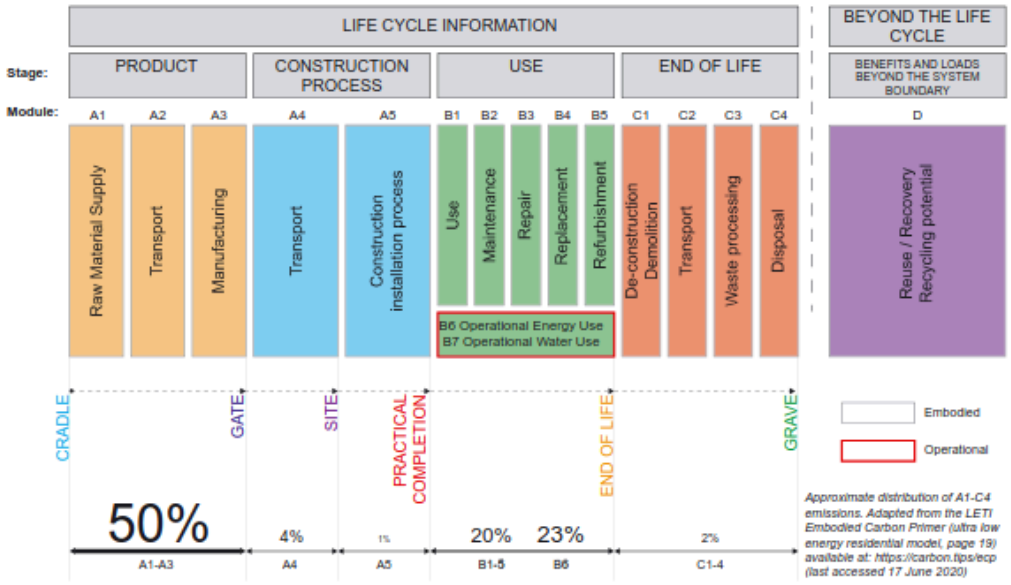
https://www.researchgate.net/publication/381307691_Beyond_Efficiency_Scaling_AI_Sustainably
https://www.sigarch.org/reducing-embodied-carbon-is-important/

SLIDE 2

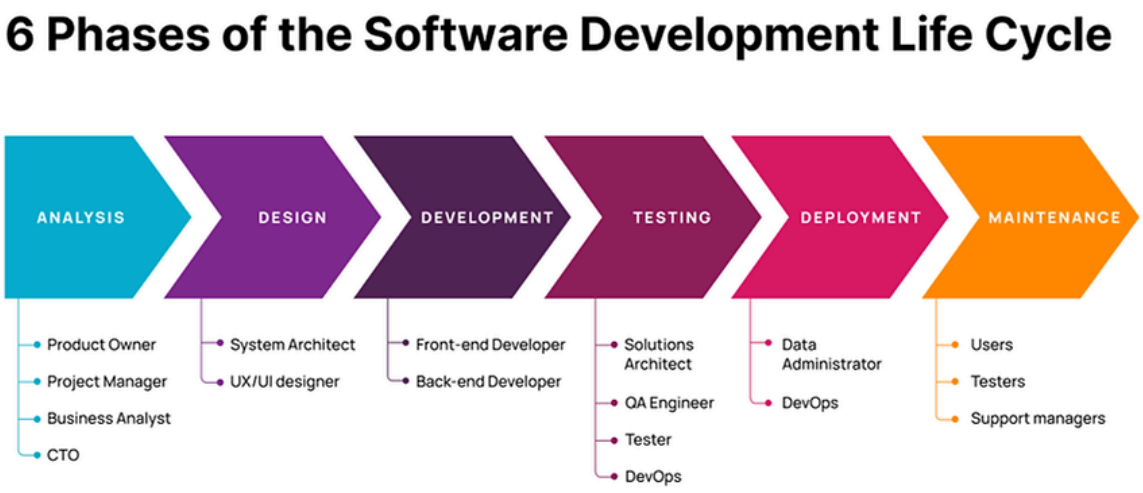
How computing affects the environment

During the lifecycle of usage and development of a software , what steps are taken and how they impact our environment.

- Need to do a LCA (life cycle assessment of scientific computing software)
- Two major carbon emission sources [1]
    - embodied carbon of the hardware (including manufacturing, transport and end-of-life disposal)
        - software effects the hardware and decreases its lifespan
    - runtime emissions (when the software runs)
        - Electricity consumption (major source)
    - my own additions
        - Development of the software
        - Testing of software (CI/CD pipelines)
        - deployment of the software
        - sharing of the software (pre build binaries ? , need to build from source)
    - Too many factors to consider , not possible to cover all



COMPLETE LIFECYCLE OF HARDWARE WHICH ACCOUNTS FOR EMBODIED CARBON
https://www.istructe.org/IStructE/media/Public/TSE-Archive/2020/A-brief-guide-to-calculating-embodied-carbon.pdf



SOFTWARE DEVELOPMENT LIFECYCLE

- **how computing contributes to emissions (details)**

SLIDE 3

SOME PRINCIPLES FOLLOWED WHILE DEVELOPING AND USING SOFTWARE THAT LEADS TO LESS CARBON FOOTPRINT

- **Energy efficiency**: consume the least amount of electricity possible [5]
- **Hardware efficiency**: use the least amount of embodied carbon possible [5]
- **Carbon awareness**: do more when electricity supplies are coming from renewable sources, and less at times when electricity is generated from carbon-intensive sources [5]
- More detailed points here
  https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202100707

  Reduces E Waste
  HOW
  62 billion kg of e-waste was generated globally in 2022
  https://ewastemonitor.info/the-global-e-waste-monitor-2024/

  Reduces Energy consumption
  https://www.sciencedirect.com/science/article/abs/pii/S0308596123002124

  Better Resource Utilization

  Increase hardware life
  https://circularcomputing.com/news/carbon-footprint-laptop/

  Lower Costs

  address climate change.

Problems with sustainability of scientific computing

Scientific computing

- Focuses on accuracy and performance
- Is highly resource and data intensive
- Is repetitive in nature

As a result

- No standard
- No transparency

SUSTAINABILITY AUDIT

WHAT IS CONSIDERED AS RESEARCH SOFTWARE
https://zenodo.org/records/5504016

- Software in Research v/s Research Software

IMPORTANT METRICS TO CONSIDER

- Carbon footprint : should also consider

(i) the carbon footprint of powering the computers during the task itself; (ii) the impact of long-term data storage; and (iii) the life cycle footprint of the hardware.

all aspects are taken into account, from extraction of raw materials to transportation, manufacturing, and disposal

On what kind of devices does the software run ?

- Is Linux compatible ?
- Is ARM compatible ?

Geographical location also matters

- Some countries has more green energy source than others

We tend to improve hardware rather than optimizing code

code profilers such as RStudio's built-in tool (and the "profvis" visualisation library) or the equivalent "cProfile" package in Python can be used to find sections that would benefit most from optimisation.

Green algorithms calculator :

Complex tasks can be broken down into simple quantifiable things :
number of cores, memory size and usage factor

do not perform a life cycle assessment and therefore, do not consider the full environmental and social impact of manufacturing, maintaining, and disposing of the hardware used

GHG is measured in terms of CO2 emissions , other gases like methane are converted to CO2 equivalent

We should also consider the energy usage during development of a software

- CI/CD pipelines during code push , runs on VM takes up energy

- Developers debugging codes and so

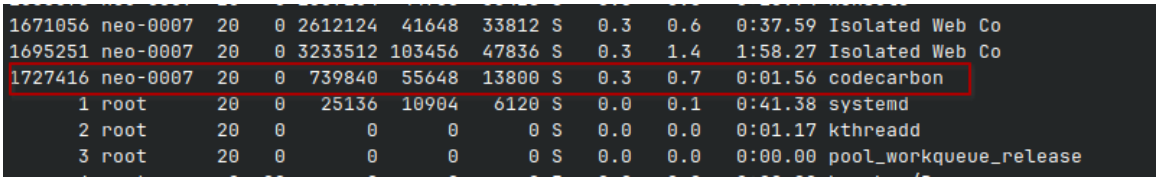
We should consider the whole lifecycle of a software and hardware it runs on

Is the software only available on Windows (proprietary license)?

Choose the tool that will fit us best :
- CodeCarbon (runtime)
- Scaphandre (runtime)
- CarbonTracker (runtime)
- Green Algorithms calculator (modelling based , post-hoc estimation without execution)
- ML CO_2 Impact calculator (modelling based , post-hoc estimation without execution)

Runtime monitoring will be suitable for our software
Choices : CodeCarbon, Scaphandre, CarbonTracker

Brief about each:
- Codecarbon
  - Python package , easy installation and setup
  - offline/ online modes , offline will be good for our use case
  - Can monitor process-level emissions
  - Allows also code level monitoring (emissions of a specific block of code.)
  - Mostly aimed towards ml tasks (well root can be used for ML tasks , but i have not used that , can i use a existing example ?)
  - Should I run root macros as subprocesses ?
  - Less overhead (at least that's what the docs say)
  - Good documentation
  - GREAT : CodeCarbon supports: - Linux (primary) - macOS (Intel and Apple Silicon) - Windows (experimental)
    Overhead:



    0.3% CPU and 0.7% Memory usage (8gb RAM system)

- Scaphandre
  - Installation and setup can be complex
  - Good visualization but too many dependencies
  - Can monitor usage by vms but that will be too heavy for this task
  - Too much features , confusing
- CarbonTracker
  - Easy setup and installation via pip

- As written directly *“Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models”*
- Can we still use that for root
- Good compatibility (macOS, Linux and Windows running on VMs , Containers, HPC clusters etc.)

All these tools only take account of the runtime effect of the code , life-cycle emissions of computing infrastructure are neglected , can we quantify that ?

Understanding the Scientific Computation Workflow

- General : https://www.youtube.com/watch?v=CchqE-9JZHk
- Task specific (Bioinformatics) :
  - https://www.youtube.com/watch?v=EwJYaVdKA0Q
  -
- ROOT :
  - Ramtools are tools that can do analysis of genome sequence data stored as .root files
  - Currently ramtools is in its early stage with some simple sets of features
  - Can build a simple data analysis pipeline with simple blocks to monitor energy usage in the whole workflow

REPOSITORY : THIS CODECARBON ANALYSIS ACCOUNTS FOR FOR **Operational Impact**

setup.sh : installs all the necessary software and dependencies

- Root (packaged binary)
- ramtools (must build from source , no prebuild binaries)
  - ROOT 6.26+
  - C++17 compatible compiler
  - CMake 3.16+
- Python , pip , python-venv
- Codecarbon

FOR SETUP SCRIPT , I cannot support all OS it will be a mess and will take time , so for now i am adding support for Ubuntu 22 , 24 and 25 and MacOS

- MacOS installation is easy : can use homebrew
- Ubuntu installation is tricky : no apt release , only snap release but can we use that for our ramtools ? , i was not able to , so i will follow the ci pipeline used to build ramtools , from here https://github.com/compiler-research/ramtools/blob/develop/.github/workflows/ci.yml

- This should do the job
- Found some decent size files to analyze , but they are in bam format
- So add a new step to convert bam to sam
- https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/pilot_data/data/NA06994/alignment/
- Pilot dataset but will do the work

Workflow:

- Download the data to analyze (SAM files)
- Convert to RAM files using ramtools
- Split into different chromosome files
- Query specific regions

FIND A SAMFILE :

- Most of the sam files i found are too big , 50+gb , what should i do
- Small sam files are there but they are too small , in kbs and are not real .
- Should I use them since I need them for testing only ?
- Will go with the example samfile in the github repo for now https://github.com/GeneROOT/ramtools/blob/develop/samexample.sam

analyzer.py

- Main analyzer that uses the codecarbon to start measurement and also run different steps of the workflow
- Download a SAM file (small for now , < 100mbs)
- Run a root macro to convert SAM-> RAM as a subprocess
    - ramtosam.C
- Run split chromosome root macro as subprocess
- Run multiple queries subprocess
    - ramntupleview.C
- Monitor the whole workflow and measure each step individually. Will generate multiple csv files for different steps of the workflow (will be helpful to analyze hotspots)

report.py
Make a combined report with our required fields analyzed
Now the values obtained from this source will be used in GreenMetaData

REFERENCES

https://www.green-coding.io/about/
https://docs.codecarbon.io/latest/
https://greensoftware.foundation/articles/how-to-accurately-measure-the-energy-consumption-of-application-software

ANALYSIS OF DATA

DATA CLEAN (original outputs)

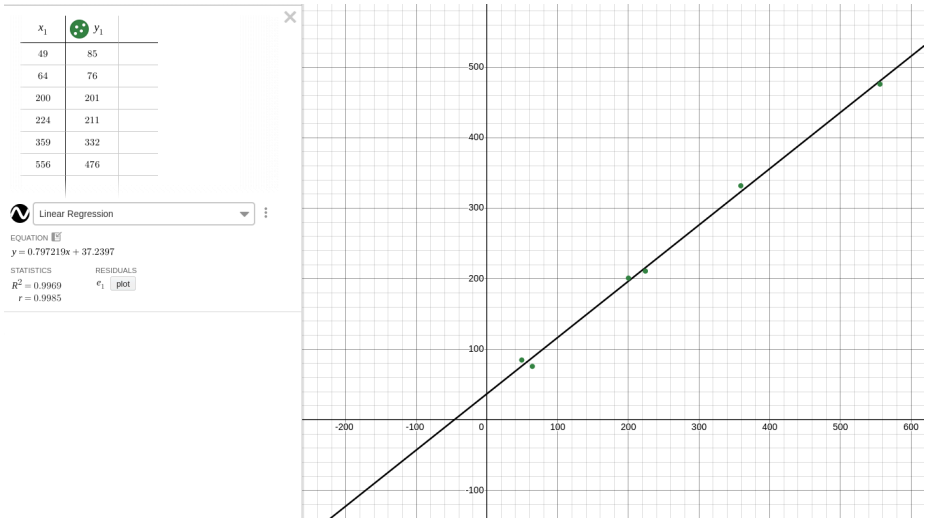
OVERALL :

Dataset Size (MB)	Total Time (s)	Total Energy (kWh)	Total Emissions (kg)
49	85	0.0002746	0.0001696
64	76	0.000287	0.000183
200	201	0.0006347	0.0003634
224	211	0.0006788	0.0003916
359	332	0.0010349	0.0003357
556	476	0.0015173	0.0010825

For calculating the per mb rates

- Use linear regression
- Used desmos for plotting and all finding out the equations
- https://www.desmos.com/calculator/

Time taken per MB



Equation :
y=0.797219x+37.2397
Rate = 0.797 s / MB

Energy used per MB

Equation:

y=0.00000247977x+0.00013778

Rate = 0.00000247977 kWh / mb

~ 2.48 × 10^{-6} kWh / MB

Average Emissions per MB

x_1	y_1
49	0.0001696
64	0.000183
200	0.0003634
224	0.0003916
359	0.0003357
556	0.0010825

Equation
y=0.00000159058x+0.0000360455
Rate = 0.00000159058 kg /MB
~ 1.5 x 10^-6 kg / MB

STEP WISE :

Dataset Size (MB)	Download Time (s)	BAM to SAM Time (s)	Chromosome Split Time (s)	Region Query Time (s)	SAM to RAM Time (s)
49	40.34	2.25	7.78	23.13	11.58
64	31.81	2.32	7.17	22.96	11.26
200	95.68	4.22	32.5	27.6	40.53
224	93.13	4.67	37.85	27.67	47.89
359	134.42	10.1	7.12	100.89	79.76
556	226.77	7.9	10.79	140.26	89.77

Dataset Size (MB)	Download Energy (kWh)	BAM to SAM Energy (kWh)	Chromosome Split Energy (kWh)	Region Query Energy (kWh)	SAM to RAM Energy (kWh)

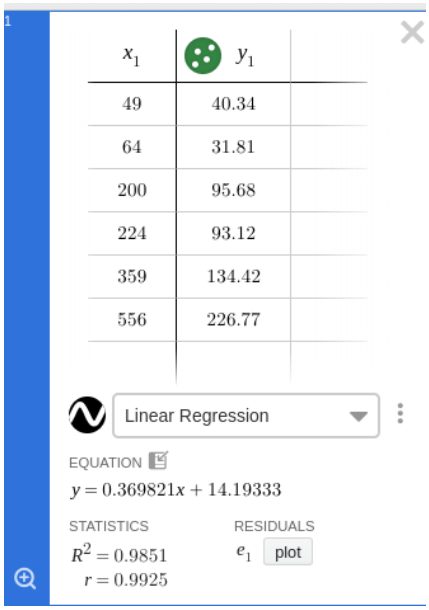
49	0.000124257	0.000005601	0.000037037	0.000071774	0.000035912
64	0.000144831	0.000005796	0.000030641	0.000070934	0.000034765
200	0.00029461	0.000012148	0.000111762	0.000090216	0.000125933
224	0.000286458	0.000013606	0.000130645	0.000100471	0.000147588
359	0.000415156	0.000030988	0.000021252	0.000315903	0.000251602
556	0.000740592	0.000024111	0.000033065	0.000440115	0.000279371

Dataset Size (MB)	Download Emissions (kg)	BAM→SAM Emissions (kg)	Chromosome Split Emissions (kg)	Region Query Emissions (kg)	SAM→RAM Emissions (kg)
49	0.00008865	0.000003996	0.000000146	0.000051206	0.000025621
64	0.000103329	0.000004135	0.000000121	0.000050607	0.000024803
200	0.000210187	0.000008667	0.000079736	0.000064364	0.000000496
224	0.000204371	0.000009707	0.000000514	0.00007168	0.000105295
359	0.000296189	0.000022108	0.000015162	0.000001244	0.000000991
556	0.000528369	0.000017202	0.00002359	0.000313996	0.000199315

For calculating the rates :

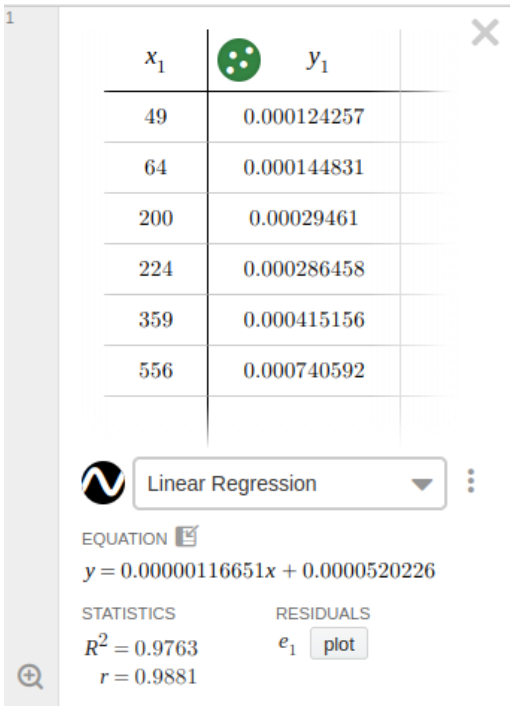
- Use linear regression
- Used desmos for plotting and all finding out the equations
- https://www.desmos.com/calculator/

DOWNLOAD DATASET :
Time avg



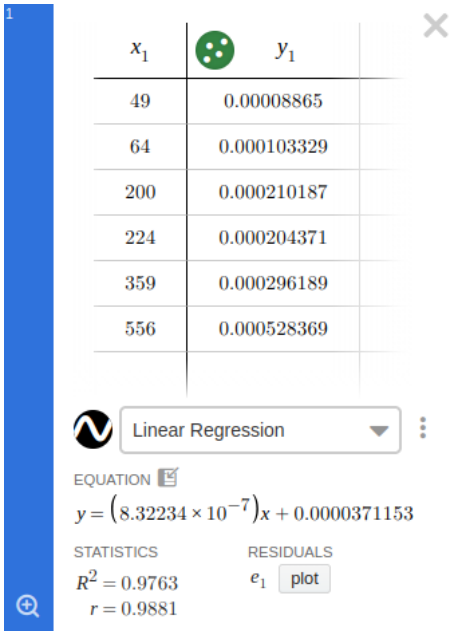
Equation :
y=0.369821x+14.19333
Rate : 0.369821 s/MB

Energy avg



Equation : y = 0.00000116651x+0.0000520226
= 0.00000116651 kWh / MB

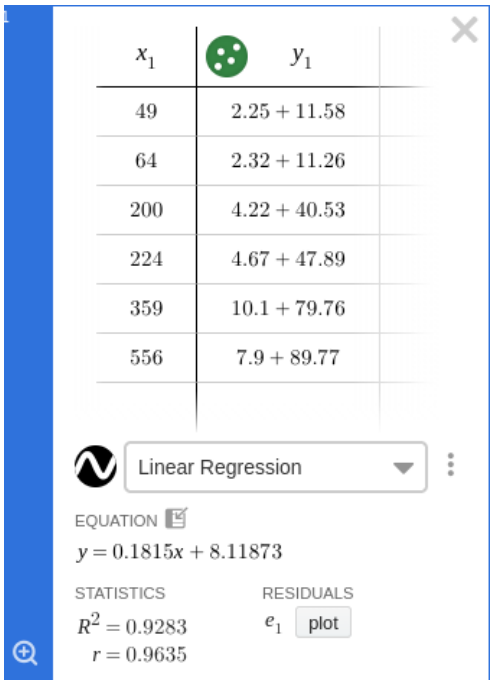
Emissions avg



0.0000009856026171 kg / MB
0.986 × 10^{-6} kg CO_2 / MB

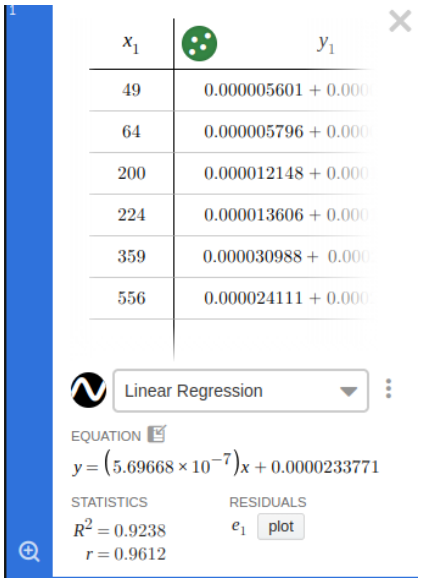
FORMAT CONVERSION (bam to sam and sam to ram) :

Time avg



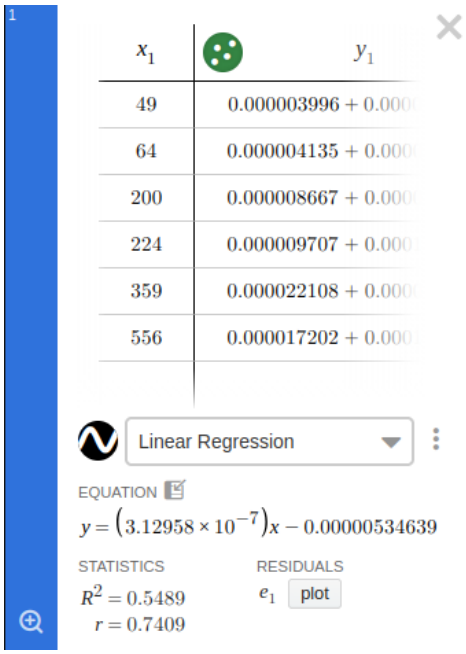
Equation :
y=0.1815x+8.11873
Rate : 0.1815 s/MB

Energy avg



Eqn : y= 5.69668 x 10^-7x+0.0000233771
Rate : 5.69668 x 10^-7

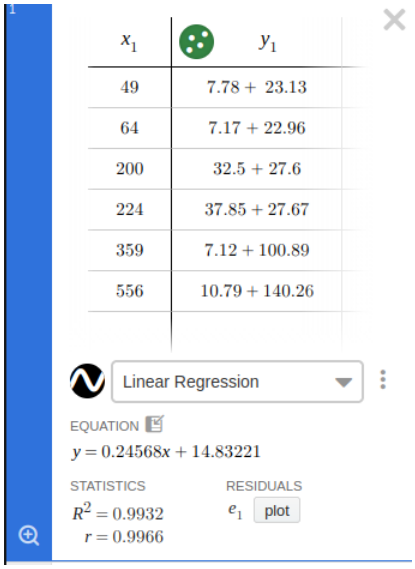
Emission avg



Eqn : y= 3.12958 x 10^-7x - 0.00000534639
Rate : 3.12958 x 10^-7

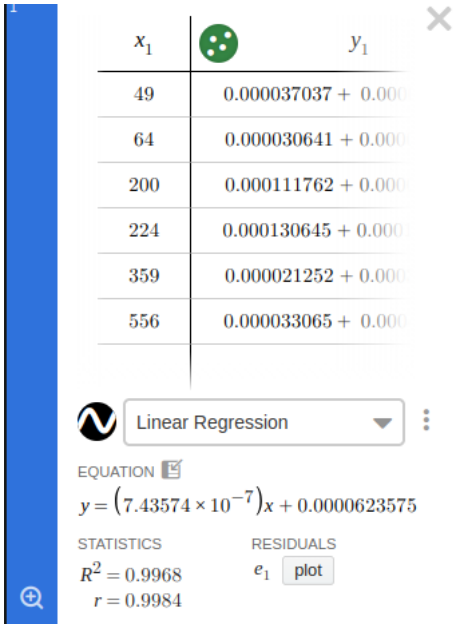
ANALYSIS (chromosome split and region query)

Time avg



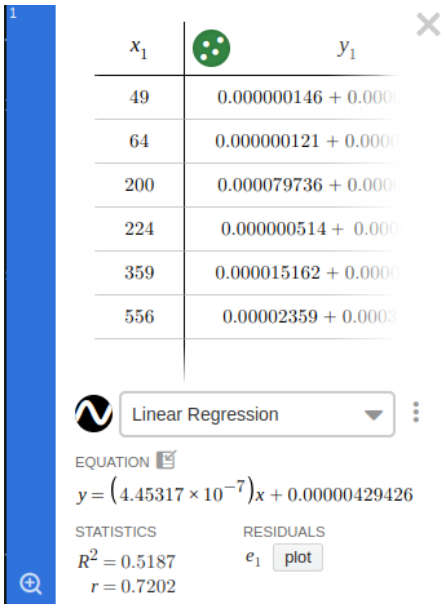
Equation : y=0.24568x+14.83221
Rate : 0.24568

Energy avg



Eqn : y=7.43574 x 10^-7x+0.0000623575
Rate : 7.43574 x 10^-7

Emission avg



Eqn : y= 4.45317 x 10^-7x+0.00000429426
Rate : 4.45317 x 10^-7

TO ANALYZE THE 37.7x COVERAGE HUMAN GENOME USING THIS PIPELINE (82GB)

Source of size of human genome
https://www.strand-ngs.com/support/ngs-data-storage-requirements

Energy consumption (in kWh):
y=0.00000247977x+0.00013778

Energy for 82 gb = 0.00000247977 (82 x 1024) +0.00013778
= 0.2083591074 kWh

Carbon footprint (in gCO2e):
y=0.00000159058x+0.0000360455
Emission for 82 gb = 0.00000159058 (82 x 1024) + 0.0000360455
= 0.1335938669 Kg
= 133.5938669 g

Time taken :
y=0.797219x+37.2397
Time = 0.797219 (82 x 1024) +37.2397
= 18.60 hours

CURRENT GREENMETADATA FIELDS

Operational Impact
Energy consumption (in kWh): 0.2083591074 kWh
Carbon footprint (in gCO2e): 133.5938669 g
Water consumption (in liters): 0 (assuming no data center used)

Methods and Scope

Software boundaries - Stages included
Use

Hardware boundaries - Components included
CPUs
Memory
Storage
Networking

Tool stages included
Use

Details on hardware used:

Product Name: VivoBook_ASUSLaptop X421IAY_M413IA
Manufacturer: ASUSTeK Computers Inc.
Processors: 6 × AMD Ryzen 5 4500U with Radeon Graphics
Memory: 7.2 GiB of RAM

Infrastructure

Infrastructure elements included
Idle machines

Power Usage Effectiveness (PUE) Value: 1
PUE Estimation method used: local computation , no overhead of data centers and hpc

Water Usage Effectiveness (WUE) Value (in L/kWh): 1
WUE Estimation method used: local computation , no overhead of data centers and hpc

Electrical Carbon Intensity Value (in gCO2e/kWh): 641.18
Electrical Carbon Intensity Source: codecarbon (uses https://app.electricitymaps.com/map/live/fifteen_minutes)

Embodied Impact
Total lifetime of the laptop : 4 years = 35040 hours
Our computing run : 18.6 hours

Percentage out of whole lifecycle for our computation :
(18.6 / 35040) x 100
= 0.0530822 %

Lifetime fraction used:
18.6 / 35040
= 0.000530822

Total embodied carbon of a Asus vivobook 14 : **≈ 191 kg CO_2e during its whole lifecycle**
https://esg.asus.com/files/carbon_footprint/Product-Carbon-Footprint-Report_X1404V.pdf

Embodied carbon allocated for our computation :

0.000530822 x 191
≈ 0.101387 kg CO_2e

which equals **≈ 101.387 g CO_2e**

Taken as a reference : https://esg.asus.com/english/file/Notebook_LCA_Verification_Report.pdf
Different model from the same manufacturer

Product Name	ASUS ExpertBook B9 B9400						
Verification Unit	1 Set (Packaging material included)						
Functional Unit	1 Set (Packaging material included)						
Life Cycle Emissions							
Environmental Impact Indicator	Unit	Target Product Environmental Impact Indicator Emissions						Functional Unit Emissions
		Material	Manufacture	Distribution	Use	Disposal	Total	
Acidification (fate not incl.)	kg SO2 eq	1.119	0.001	0.032	0.171	2.909E-04	1.325	1.325
Eutrophication	kg PO4--- eq	0.181	2.348E-04	0.005	0.020	5.243E-05	0.207	0.207
Global warming (GWP100a)	kg CO2 eq	170.001	0.567	6.649	41.413	0.230	218.861	218.861
Photochemical oxidation	kg NMVOC	0.660	0.001	0.037	0.096	3.740E-04	0.795	0.795
Abiotic depletion, elements	kg Sb eq	0.053	1.251E-04	3.238E-06	2.301E-04	1.345E-07	0.053	0.053
Abiotic depletion, fossil fuels	MJ	1911.385	3.167	93.070	494.377	0.624	2502.624	2502.624
Water scarcity	m3 eq	44.845	0.125	0.098	6.807	0.006	51.882	51.882
Ozone layer depletion (ODP) (optional)	kg CFC-11 eq	1.693E-05	5.916E-08	1.187E-06	1.991E-06	7.407E-09	2.017E-05	2.017E-05

Carbon Footprint

Carbon footprint (gCO2e): 101.387
Source and method (carbon footprint): above

Depletion of Abiotic Resources (Minerals, Metals)

Total : 0.053
For our computation : 0.000530822 x 0.053
= 2.8133566 x 10^-5
= 0.000028133566
Depletion of Abiotic Resources (kg Sb-eq):
Source and method (depletion):

Particulate Matter Emissions

Particulate Matter Emissions (disease incidence):
Source and method (particulate matter):

Acidification potential (kg mol H+):
Total : 1.325
For our computation :
= 1.325 x 0.000530822
= 7.0333915 x 10^-4
= 0.0007033915

Source and method (acidification):

Ionising Radiation Related to Human Health

Ionising Radiation (kBq U-235):
Source and method (ionising radiation):

Photochemical Ozone Formation

Photochemical Ozone Formation (kg NMVOC-eq):
Source and method (ozone):
Total : 0.795
For our computation : 0.000530822 x 0.795
= 4.2200348 x 10^-4
= 0.00042200348

Abiotic Depletion Potential (Fossil Fuels)

Total : 2502.624
For our computation : 0.000530822 x 2502.624
= 1.328447877
Abiotic Depletion Potential (MJ):
Source and method (fossil fuels):

Freshwater Eco-Toxicity Potential

Freshwater Eco-Toxicity Potential (CTUe):
Source and method (ecotoxicity):

CODECARBON OUTPUTS RAW:

49MB DATA
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/pilot_data/data/NA06994/alignment/NA06994.454.MOSAIK.SRP000033.2009_11.bam

SYSTEM INFORMATION
cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query
duration_seconds: 23.131968240995775
energy_consumed_kwh: 7.177354984704599e-05
emissions_kg: 5.1206193176426336e-05
cpu_energy_kwh: 1.0304132191502985e-05
ram_energy_kwh: 6.146941765554301e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 73.1952380952381

STEP: sam_to_ram_conversion
duration_seconds: 11.58325966099801
energy_consumed_kwh: 3.5912251448558404e-05
emissions_kg: 2.5621272585710955e-05
cpu_energy_kwh: 5.1312540569040684e-06
ram_energy_kwh: 3.0780997391654334e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 71.8

STEP: bam_to_sam_conversion
duration_seconds: 2.2542785440018633

energy_consumed_kwh: 5.601173379889029e-06
emissions_kg: 3.996106737321408e-06
cpu_energy_kwh: 7.323371132261427e-07
ram_energy_kwh: 4.868836266662886e-06
cpu_utilization_percent: 0
ram_utilization_percent: 0

STEP: download_dataset
duration_seconds: 40.34384045400657
energy_consumed_kwh: 0.00012425687901396068
emissions_kg: 8.864995202059913e-05
cpu_energy_kwh: 1.637105033339327e-05
ram_energy_kwh: 0.00010788582868056742
cpu_utilization_percent: 5.2631578947368425
ram_utilization_percent: 71.23684210526316

STEP: chromosome_split
duration_seconds: 7.778511659998912
energy_consumed_kwh: 3.703682048474098e-05
emissions_kg: 1.4582711130477826e-07
cpu_energy_kwh: 1.6838479404176168e-05
ram_energy_kwh: 2.0198341080564812e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 75.56666666666666

TOTALS
total_duration_seconds: 85.09185856000113
total_energy_kwh: 0.0002745806741741951
total_emissions_kg: 0.0001696193516313626

359MB
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/phase1/data/HG00118/alignment/HG00118.unmapped.ILLUMINA.bwa.GBR.low_coverage.20101123.bam

SYSTEM INFORMATION
cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query

duration_seconds: 100.89319506000902
energy_consumed_kwh: 0.00031590340408602275
emissions_kg: 1.2438238559973126e-06
cpu_energy_kwh: 4.539102883878329e-05
ram_energy_kwh: 0.0002705123752472395
cpu_utilization_percent: 2.0202020202020203
ram_utilization_percent: 63.57676767676768

STEP: sam_to_ram_conversion
duration_seconds: 79.7576220260089
energy_consumed_kwh: 0.0002516019484089277
emissions_kg: 9.906461962695762e-07
cpu_energy_kwh: 3.8470071233962945e-05
ram_energy_kwh: 0.00021313187717496475
cpu_utilization_percent: 0.0
ram_utilization_percent: 72.312

STEP: bam_to_sam_conversion
duration_seconds: 10.0956556500023
energy_consumed_kwh: 3.098766732516754e-05
emissions_kg: 2.2107872364134852e-05
cpu_energy_kwh: 4.34065846124467e-06
ram_energy_kwh: 2.6647008863922868e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 68.6375

STEP: download_dataset
duration_seconds: 134.4236039410025
energy_consumed_kwh: 0.0004151560005264536
emissions_kg: 0.00029618931217159357
cpu_energy_kwh: 5.4304374779200865e-05
ram_energy_kwh: 0.0003608516257472527
cpu_utilization_percent: 4.545454545454546
ram_utilization_percent: 72.07954545454545

STEP: chromosome_split
duration_seconds: 7.117093718989054
energy_consumed_kwh: 2.1251681545695182e-05
emissions_kg: 1.5161820933642316e-05
cpu_energy_kwh: 2.876011284586478e-06
ram_energy_kwh: 1.8375670261108706e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 64.0

TOTALS
total_duration_seconds: 332.2871703960118
total_energy_kwh: 0.0010349007018922668
total_emissions_kg: 0.0003356934755216376

64MB
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/pilot_data/data/NA06994/alignment/NA06994.454.ssaha.SRP000033.2009_10.bam

SYSTEM INFORMATION
cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query
duration_seconds: 22.958602225000504
energy_consumed_kwh: 7.093398606250086e-05
emissions_kg: 5.060721395041668e-05
cpu_energy_kwh: 9.945713654166956e-06
ram_energy_kwh: 6.098827240833391e-05
cpu_utilization_percent: 14.285714285714286
ram_utilization_percent: 65.08095238095238

STEP: sam_to_ram_conversion
duration_seconds: 11.256827389996033
energy_consumed_kwh: 3.476540697354811e-05
emissions_kg: 2.480306671661514e-05
cpu_energy_kwh: 4.890090898579619e-06
ram_energy_kwh: 2.9875316074968496e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 64.37777777777778

STEP: bam_to_sam_conversion
duration_seconds: 2.3157779640023364
energy_consumed_kwh: 5.7963137486876495e-06
emissions_kg: 4.1353278771774655e-06
cpu_energy_kwh: 7.5730479591177e-07
ram_energy_kwh: 5.0390089527758795e-06
cpu_utilization_percent: 0
ram_utilization_percent: 0

STEP: download_dataset

duration_seconds: 31.806857741001295
energy_consumed_kwh: 0.00014483130145724008
emissions_kg: 0.00010332858854295482
cpu_energy_kwh: 1.8993026307270997e-05
ram_energy_kwh: 0.00012583827514996908
cpu_utilization_percent: 6.666666666666667
ram_utilization_percent: 57.24333333333333

STEP: chromosome_split
duration_seconds: 7.165571454999736
energy_consumed_kwh: 3.0640971644611176e-05
emissions_kg: 1.206443837247364e-07
cpu_energy_kwh: 1.2160377205756378e-05
ram_energy_kwh: 1.8480594438854798e-05
cpu_utilization_percent: 20.0
ram_utilization_percent: 68.62

TOTALS
total_duration_seconds: 75.5036367749999
total_energy_kwh: 0.0002869679798865879
total_emissions_kg: 0.00018299484147088883

200MB
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/phase1/data/HG00118/alignment/HG00118.chrom20.ILLUMINA.bwa.GBR.low_coverage.20101123.bam

SYSTEM INFORMATION
cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query
duration_seconds: 27.603855327994097
energy_consumed_kwh: 9.021611554855039e-05
emissions_kg: 6.436387569307333e-05
cpu_energy_kwh: 1.6326182026359574e-05
ram_energy_kwh: 7.38899335221908e-05
cpu_utilization_percent: 7.6923076923076925
ram_utilization_percent: 57.75769230769231

STEP: sam_to_ram_conversion
duration_seconds: 40.53148161500576

energy_consumed_kwh: 0.00012593336196842496
emissions_kg: 4.958443557626804e-07
cpu_energy_kwh: 1.7524167096140493e-05
ram_energy_kwh: 0.00010840919487228449
cpu_utilization_percent: 2.5641025641025643
ram_utilization_percent: 63.620512820512815

STEP: bam_to_sam_conversion
duration_seconds: 4.216679463002947
energy_consumed_kwh: 1.2148029754787214e-05
emissions_kg: 8.666902496285146e-06
cpu_energy_kwh: 1.8295114297685306e-06
ram_energy_kwh: 1.0318518325018684e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 62.2

STEP: download_dataset
duration_seconds: 95.68139385399991
energy_consumed_kwh: 0.000294609732222245
emissions_kg: 0.0002101866619663707
cpu_energy_kwh: 3.858268493614552e-05
ram_energy_kwh: 0.00025602704728609955
cpu_utilization_percent: 5.319148936170213
ram_utilization_percent: 63.35531914893617

STEP: chromosome_split
duration_seconds: 32.49913514200307
energy_consumed_kwh: 0.00011176198079885052
emissions_kg: 7.973557934311272e-05
cpu_energy_kwh: 2.5747079940520826e-05
ram_energy_kwh: 8.601490085832969e-05
cpu_utilization_percent: 3.3333333333333335
ram_utilization_percent: 78.48666666666666

TOTALS
total_duration_seconds: 200.53254540200578
total_energy_kwh: 0.0006346692202928581
total_emissions_kg: 0.00036344886385460454

224MB
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/phase1/data/HG00114/alignment/HG00114.chrom20.ILLUMINA.bwa.GBR.low_coverage.20101123.bam

SYSTEM INFORMATION

cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query
duration_seconds: 27.667660192993935
energy_consumed_kwh: 0.00010047090988770917
emissions_kg: 7.168006642119712e-05
cpu_energy_kwh: 2.6451718815480528e-05
ram_energy_kwh: 7.401919107222865e-05
cpu_utilization_percent: 7.6923076923076925
ram_utilization_percent: 61.03461538461538

STEP: sam_to_ram_conversion
duration_seconds: 47.88916581599915
energy_consumed_kwh: 0.00014758807255694313
emissions_kg: 0.00010529538207309807
cpu_energy_kwh: 2.0133335895849938e-05
ram_energy_kwh: 0.00012745473666109317
cpu_utilization_percent: 0.0
ram_utilization_percent: 59.64565217391304

STEP: bam_to_sam_conversion
duration_seconds: 4.67397983900446
energy_consumed_kwh: 1.360620921938236e-05
emissions_kg: 9.70722751168537e-06
cpu_energy_kwh: 2.017881094391493e-06
ram_energy_kwh: 1.1588328124990867e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 58.4

STEP: download_dataset
duration_seconds: 93.12706886000524
energy_consumed_kwh: 0.00028645772850131586
emissions_kg: 0.00020437068827970729
cpu_energy_kwh: 3.75276496651547e-05
ram_energy_kwh: 0.00024893007883616114
cpu_utilization_percent: 2.197802197802198
ram_utilization_percent: 59.48021978021978

STEP: chromosome_split
duration_seconds: 37.853025393007556
energy_consumed_kwh: 0.0001306452823167888

emissions_kg: 5.14396859031237e-07
cpu_energy_kwh: 2.9683030477916533e-05
ram_energy_kwh: 0.00010096225183887225
cpu_utilization_percent: 2.7777777777777777
ram_utilization_percent: 76.71944444444445

TOTALS
total_duration_seconds: 211.21090010101034
total_energy_kwh: 0.0006787682024821393
total_emissions_kg: 0.0003915677611447191

556MB :
https://ftp-trace.ncbi.nih.gov/1000genomes/ftp/phase1/data/HG00114/alignment/HG00114.unmapped.ILLUMINA.bwa.GBR.low_coverage.20101123.bam

SYSTEM INFORMATION
cpu_model: AMD Ryzen 5 4500U with Radeon Graphics
cpu_count: 6
ram_total_size_gb: 7.184230804443359

STEP: region_query
duration_seconds: 140.25935417800792
energy_consumed_kwh: 0.0004401148807918602
emissions_kg: 0.00023199600066702555
cpu_energy_kwh: 6.443264156691188e-05
ram_energy_kwh: 0.00037568223922494835
cpu_utilization_percent: 3.6231884057971016
ram_utilization_percent: 72.2913043478261

STEP: sam_to_ram_conversion
duration_seconds: 89.76993465599662
energy_consumed_kwh: 0.00027937084318797114
emissions_kg: 0.00019931461373486932
cpu_energy_kwh: 3.836536423801577e-05
ram_energy_kwh: 0.0002410054789499554
cpu_utilization_percent: 6.818181818181818
ram_utilization_percent: 72.62840909090909

STEP: bam_to_sam_conversion
duration_seconds: 7.903546944988193
energy_consumed_kwh: 2.411130751439165e-05
emissions_kg: 1.7201995344375094e-05
cpu_energy_kwh: 3.5511870727237466e-06

ram_energy_kwh: 2.0560120441667903e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 70.43333333333334

STEP: download_dataset
duration_seconds: 226.77417134999996
energy_consumed_kwh: 0.0007405921075465603
emissions_kg: 0.0005283687738001255
cpu_energy_kwh: 9.130175306878824e-05
ram_energy_kwh: 0.0006492903544777721
cpu_utilization_percent: 2.6785714285714284
ram_utilization_percent: 74.42589285714287

STEP: chromosome_split
duration_seconds: 10.7924131359905
energy_consumed_kwh: 3.30651935774635e-05
emissions_kg: 2.359006477109914e-05
cpu_energy_kwh: 4.48040601915524e-06
ram_energy_kwh: 2.858478755830826e-05
cpu_utilization_percent: 0.0
ram_utilization_percent: 70.88888888888889

TOTALS
total_duration_seconds: 475.4994202649832
total_energy_kwh: 0.0015172543326182467
total_emissions_kg: 0.0010824714483174945