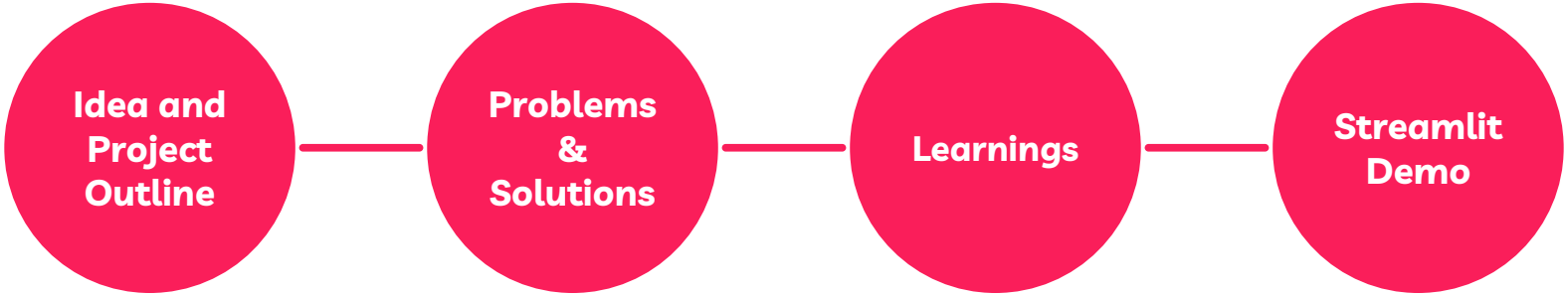
Served with Microplastics

Uncovering the Hidden Contaminants on Our Plates

TechLabs Düsseldorf
Summer Term 2025
Data Science Track – Group 2

Today's Agenda



PUREPLATE

Initial Planning

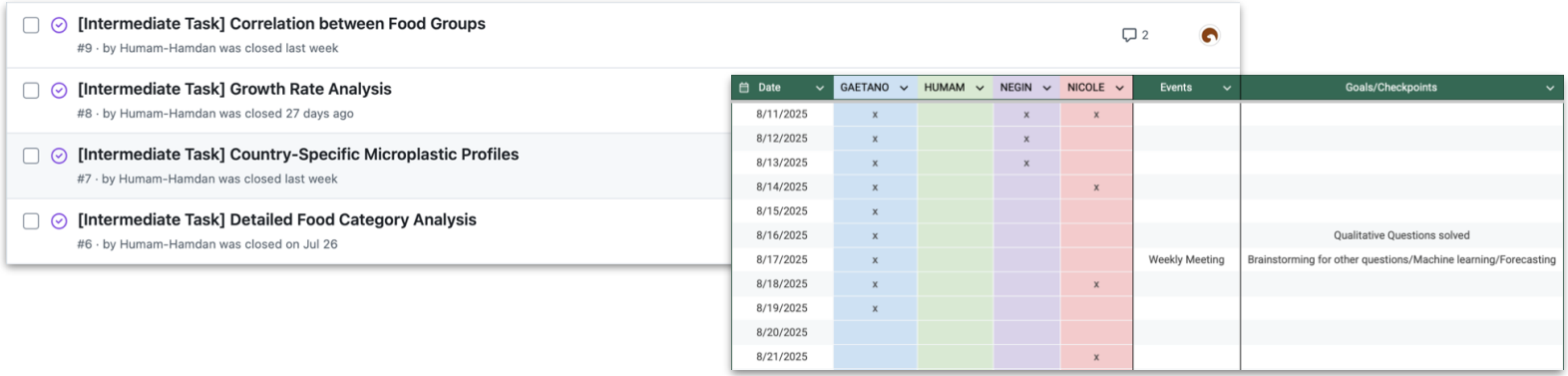
- Global non-profit dedicated to food safety and advocacy
- Mission: raise awareness on emerging contaminants in diets
- Focus on microplastics and their impact on human health

Project Goals

- Analyze global dataset on estimated microplastic consumption
- Pinpoint high-risk food categories
- Identify trends over time and geographical variations
- Translate findings into public health campaigns & policy guidance

Our initial planning...

- Group events, deadlines as well as work capabilities were documented
- We created forks for each contributor
- The questions of the INSTRUCTIONS.md were translated issues and assigned



[Intermediate Task] Correlation between Food Groups
#9 · by Humam-Hamdan was closed last week
2

[Intermediate Task] Growth Rate Analysis
#8 · by Humam-Hamdan was closed 27 days ago

[Intermediate Task] Country-Specific Microplastic Profiles
#7 · by Humam-Hamdan was closed last week

[Intermediate Task] Detailed Food Category Analysis
#6 · by Humam-Hamdan was closed on Jul 26

Date	GAETANO	HUMAM	NEGIN	NICOLE	Events	Goals/Checkpoints
8/11/2025	x		x	x		
8/12/2025	x		x			
8/13/2025	x		x			
8/14/2025	x			x		
8/15/2025	x					
8/16/2025	x					Qualitative Questions solved
8/17/2025	x				Weekly Meeting	Brainstorming for other questions/Machine learning/Forecasting
8/18/2025	x			x		
8/19/2025	x					
8/20/2025						
8/21/2025				x		

...vs how it actually went

- We kept it flexible: if deadlines couldn't be met as planned, we discussed possible solutions
- Questions suggested from the mentors were used as a baseline
- After getting the first interesting results, we expanded our analysis with our personal ideas

Packages & Frameworks

PYTHON NUMPY PANDAS MATPLOTLIB SEABORN STREAMLIT

Version control

GitHub

Communication

slack

Of course, this project didn't come without challenges!

Navigating version control

```
Auto-merging src/colors.txt
CONFLICT (content): Merge conflict in src/colors.txt
Automatic merge failed; fix conflicts and then commit the result.
```

Deciphering pandas errors

```
File /srv/conda/envs/notebook/lib/python3.11/site-packages/pandas/core/indexes/base.py:6373, in Index
   6370 assert kind in ["getitem", "iloc"]
   6372 if key is not None and not is_integer(key):
-> 6373     self._raise_invalid_indexer(form, key)

File /srv/conda/envs/notebook/lib/python3.11/site-packages/pandas/core/indexes/base.py:4151, in Index
   4149 if reraise is not lib.no_default:
   4150     raise TypeError(msg) from reraise
-> 4151 raise TypeError(msg)

TypeError: cannot do slice indexing on RangeIndex with these indexers [fruit] of type str
```

Making a storyline out of our data

Top contributors in top 10 most concentrated countries:

	top_category	share	value
Congo	refined_grains	0.579	811.49
Angola	refined_grains	0.516	685.7
Togo	refined_grains	0.508	544.06
Zimbabwe	refined_grains	0.486	390.04
Madagascar	refined_grains	0.461	479.69
Cote D'Ivoire	other_starchy_vegetables	0.45	696.76
Tanzania	refined_grains	0.445	583.53
Cambodia	refined_grains	0.442	379.9
Ethiopia	refined_grains	0.425	333.84
Niger	refined_grains	0.421	445.42

What did we learn?

- Using pandas for grouping, filtering, and aggregations
- Creating different visualizations (line plots, stacked area, bar plots)
- Applying metrics like CAGR, rank shifts, and concentration measures
- Version control with Git/GitHub, resolving merge conflicts
- Streamlit for turning analysis into an interactive app

But more importantly...

- How to divide tasks in a data science team
- Importance of clear communication at every step
- Documenting steps so teammates can follow
- Not panicking when merge conflicts show up!

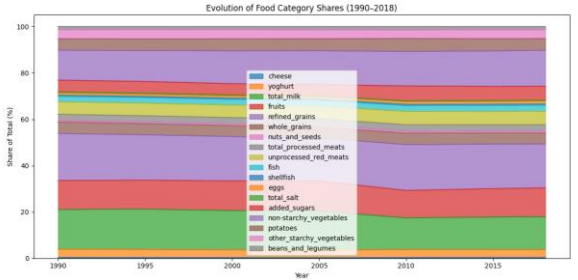
The Project Phase

- Challenging but very practical:

  not just coding, but also collaborating like in a “real” data science team

- Fun to discover unexpected patterns

  (like contamination being very stable across decades)



Next time we will... Plan visualizations earlier, so the storyline emerges sooner.

Streamlit Demo

Which categories are overtaking the other?

	1990 Share	2018 Share	Rank 1990	Rank 2018	Rank Change	CAGR %
eggs	0.989805	1.233199	14.0	13.0	1.0	1.568054
non-starchy_vegetables	12.959423	15.582782	3.0	2.0	1.0	1.439307
added_sugars	5.026419	6.020226	6.0	5.0	1.0	1.425083
nuts_and_seeds	0.699539	0.824895	15.0	15.0	0.0	1.368652
cheese	0.383982	0.424147	18.0	17.0	1.0	1.132337
fish	2.377785	2.602444	12.0	12.0	0.0	1.099108
unprocessed_red_meats	5.428887	5.877132	5.0	6.0	-1.0	1.059591
total_processed_meats	2.645601	2.744104	11.0	11.0	0.0	0.905310
shellfish	0.455409	0.467749	16.0	16.0	0.0	0.869919
potatoes	4.886678	4.970289	8.0	7.0	1.0	0.834733
beans_and_legumes	1.191264	1.187611	13.0	14.0	-1.0	0.762601
fruits	12.467458	12.414322	4.0	4.0	0.0	0.758285
other_starchy_vegetables	4.055513	3.974451	9.0	9.0	0.0	0.701015
whole_grains	4.891093	4.669130	7.0	8.0	-1.0	0.606643
yoghurt	3.550047	3.323837	10.0	10.0	0.0	0.536968
refined_grains	20.355815	19.025627	1.0	1.0	0.0	0.530725
total_salt	0.451478	0.411997	17.0	18.0	-1.0	0.444842
total_milk	17.183804	14.246059	2.0	3.0	-1.0	0.101133

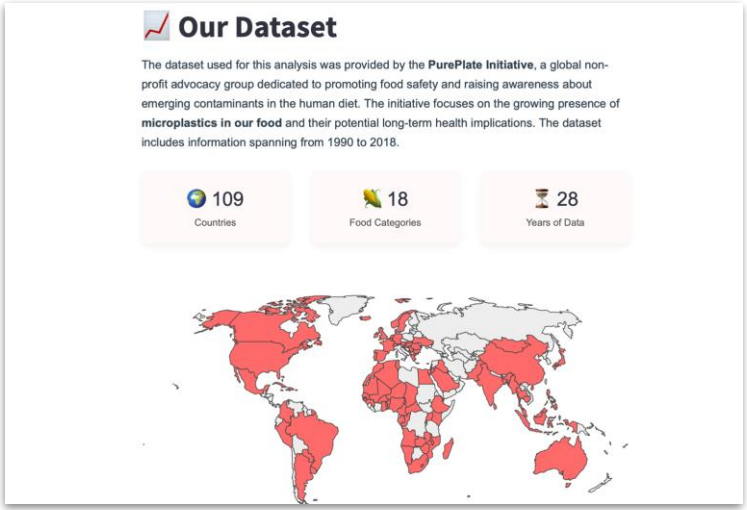
Top countries by growth rate:

	country	starting_year	first_value	finishing_year	last_value	period_years	growth_rate
25	Croatia	2010	1861.635563	2018	2582.073733	8	0.041740
52	Laos	1990	713.832073	2018	1923.902724	28	0.036044
9	Belgium	2010	1734.369945	2018	2280.741383	8	0.034825
66	Myanmar	1990	493.477704	2018	1232.894174	28	0.033242
106	Vietnam	1990	639.514688	2018	1449.618314	28	0.029658

Originally…

Analysis stored in main.py (prints & calculations)

However, results were difficult to interpret visually



So…

We built a Streamlit web app for clear, interactive insights

PurePlate Initiative

Group 3

Humam Hamdan
Gaetano Manzone
Negin Doostdar
Nicole Mikula