

SDMX Workflows with Ethical AI: Early Experiments from the Pacific

Giulio Valentino Dalla Riva

SDMX conf 2025

Monday, September 29, 2025 – Rome

We want LLM for SDMX

We want LLM for SDMX
to improve
user and developer
experience

01 Challenges

LLMs* are error prone

* and unavoidably so

LLMs* are expensive

* big, commercial

LLMs* are risky

* Cloud hosted ones

SDMX* is info rich

* even too much

02

Design

Provide help, not answers

Because LLMs are error prone

Prefer local LLMs

Because LLMs are expensive

Avoid sending (micro)data

Because LLMs are risky

Simplify and summarise

SDMX is info rich

03

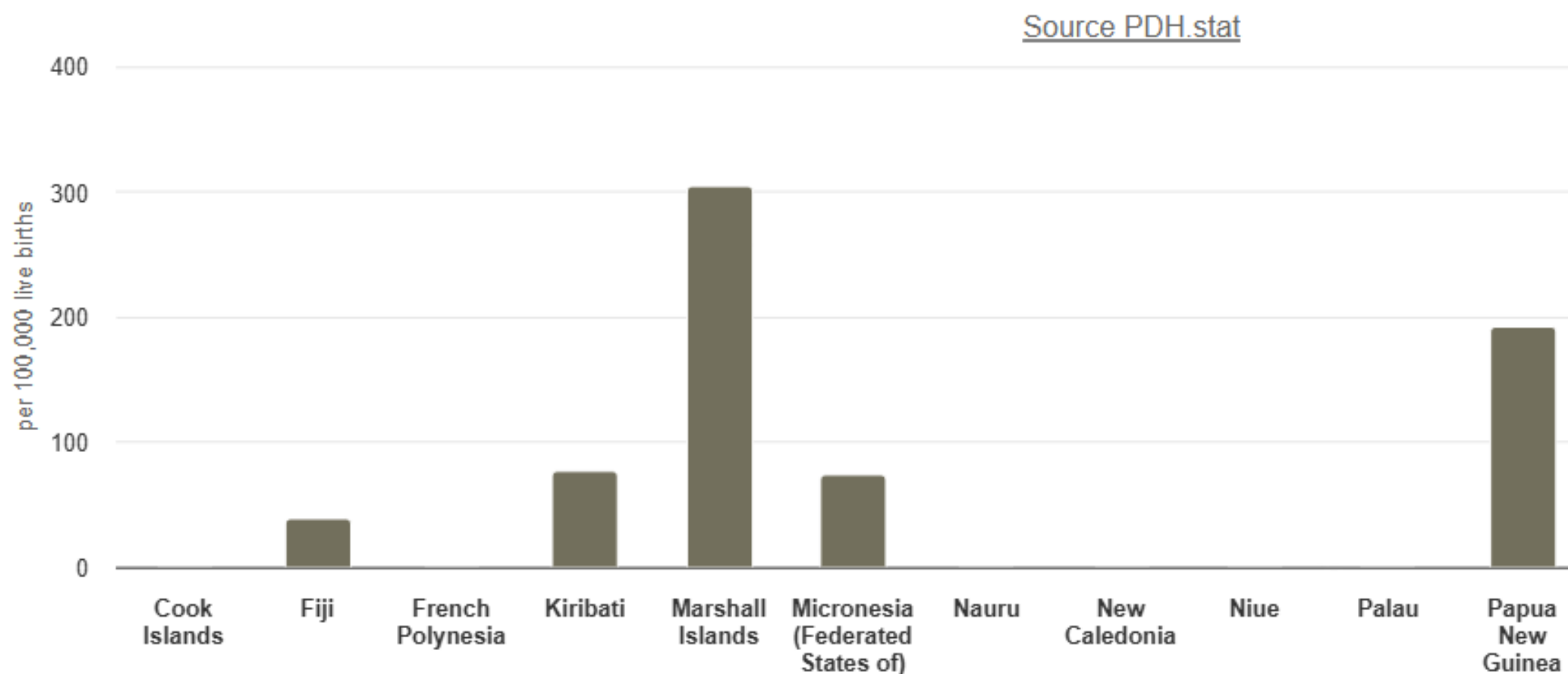
Case studies

03.1

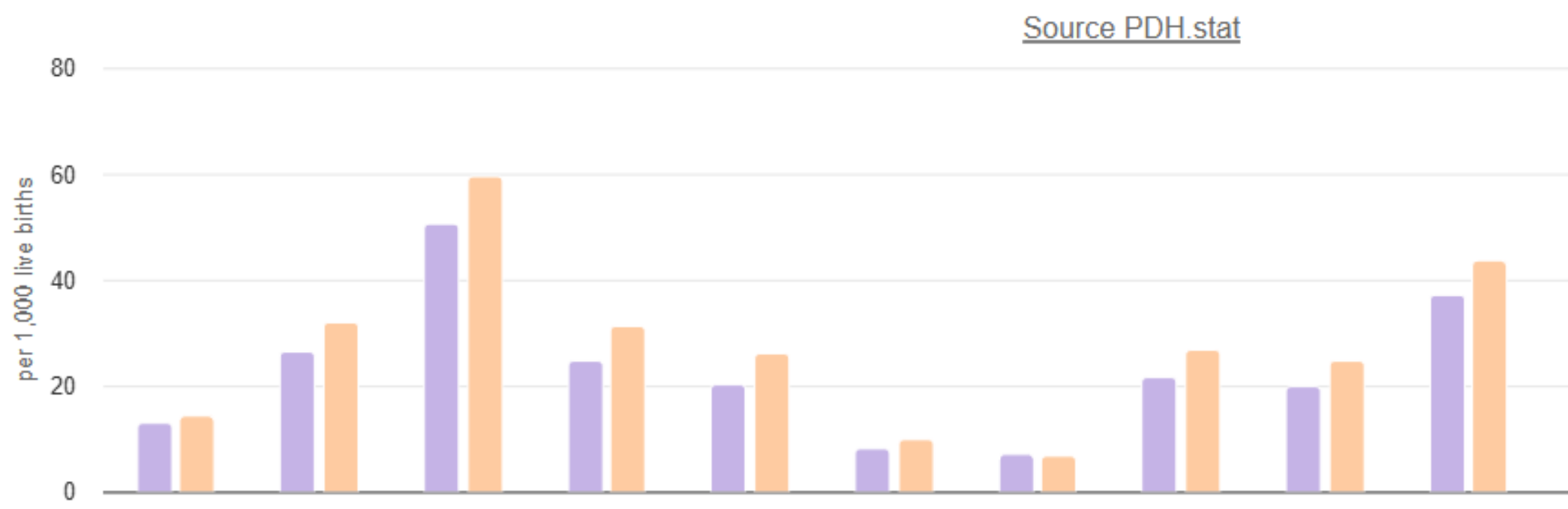
Searching for Indicators in Blue Pacific 2050

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached zero in the graph below.



Under-five mortality rate



Blue Pacific 2050

~ 100 indicators

A nested hierarchy

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as

Word match search

User need to know how an
indicator is phrased

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as

LLM based RAG search

AI used to “guess” semantic

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as

LLM based RAG search

It “works”

Quality and completeness of
structural metadata is KEY

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as



Pacific
Community
Communauté
du Pacifique

LLM based
RAG search

sub-second answers

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as

LLM based RAG search

Runs on a very small server

Search Indicators

eco

Found 7 indicators

People Centred Development

2.2 Strengthened Pacific education curriculum and assessment systems to meet employment needs

Completion rate, upper **secondary**

3.1 Strengthened institutional capacity and capability in Pacific island countries to promote equality, equity, and social inclusion

Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement) and **economic** benefits

Existence of implementation plan for the different UN Treaty Body **recommendations** and UN **recommendations** which are fully or partially resourced

Resource and Economic Development

1.5 Improved collaboration and policies for the planning, development, and management of

Implementation of standard accounting tools to monitor the **economic** and environmental as

LLM based RAG search

Cost is just* the server

* plus software development

The AI is *embarrassingly* simple

```
using Tidier
using RAGTools

labels = read_csv("../Data/Out/Labels_clean.csv")

indicator_index = build_index(
    labels.label;
    chunker_kwargs = (; sources = labels.code)
)

answer = retrieve(indicator_index, "Proportion project goals new buildings")

candidates = labels[answer.reranked_candidates.positions,:]

print(candidates[1,2])
```

Prototype
Code
in
Julia

The software-engineering less so

```
using Tidier
using RAGTools

labels = read_csv("../Data/Out/Labels_clean.csv")

indicator_index = build_index(
    labels.label;
    chunker_kwargs = (; sources = labels.code)
)

answer = retrieve(indicator_index, "Proportion project goals new buildings")

candidates = labels[answer.reranked_candidates.positions,:]

print(candidates[1,2])
```

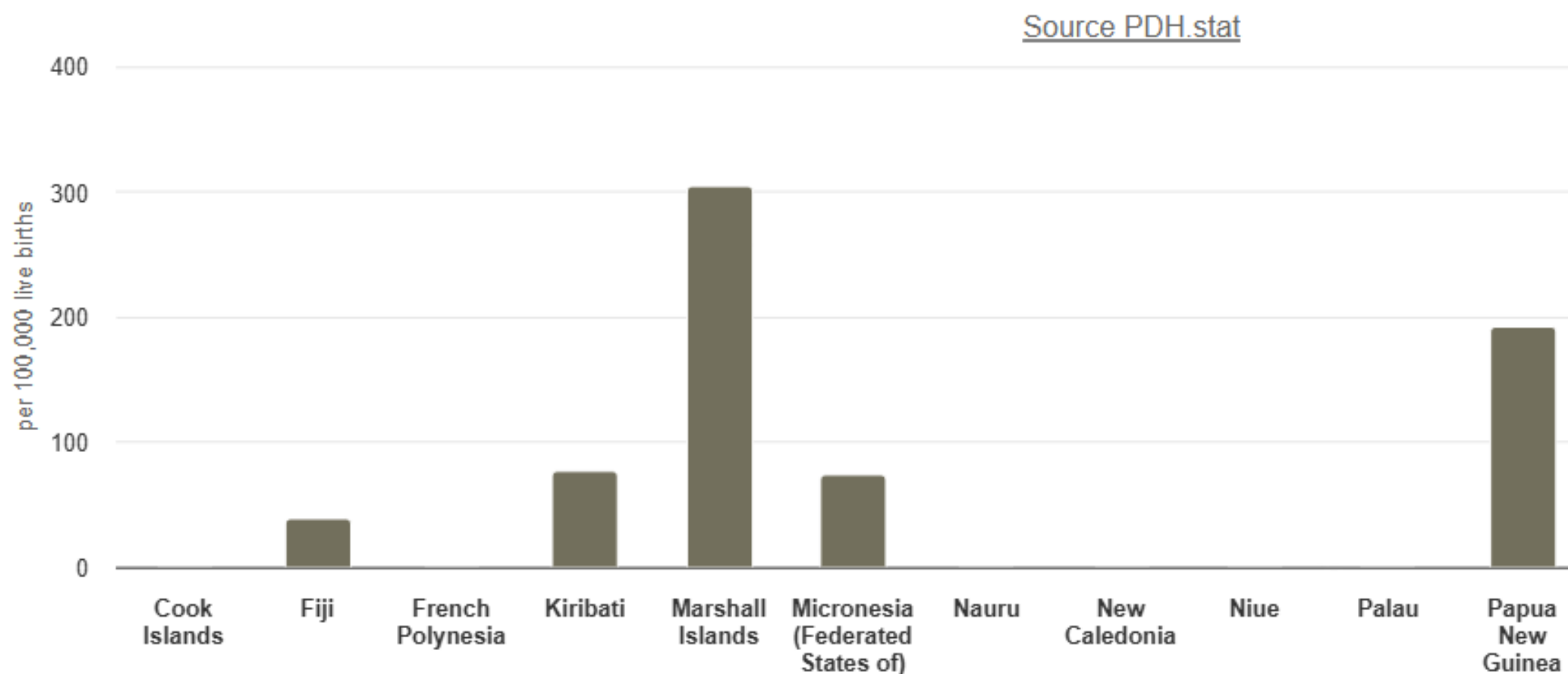
APIs
Authorization
Cybersecurity
Maintenance
Scheduling
Caching
...

03.2

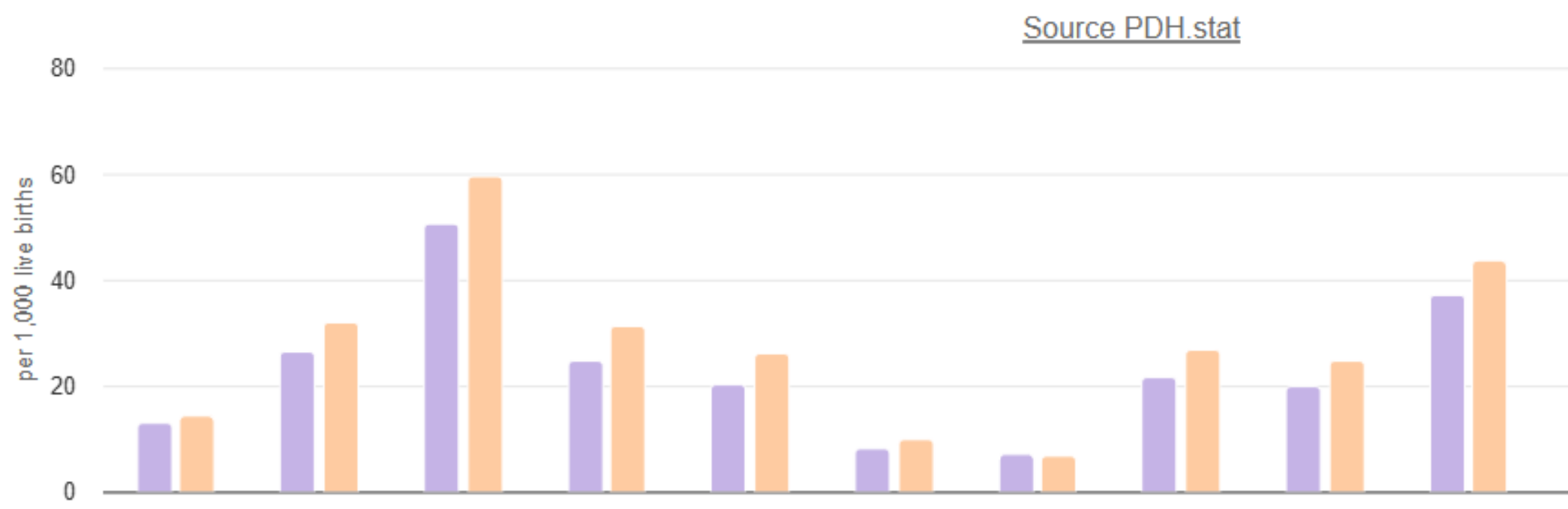
Teaching LLMs to talk with SDMX .Stat

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached this target as zero in the graph below.



Under-five mortality rate



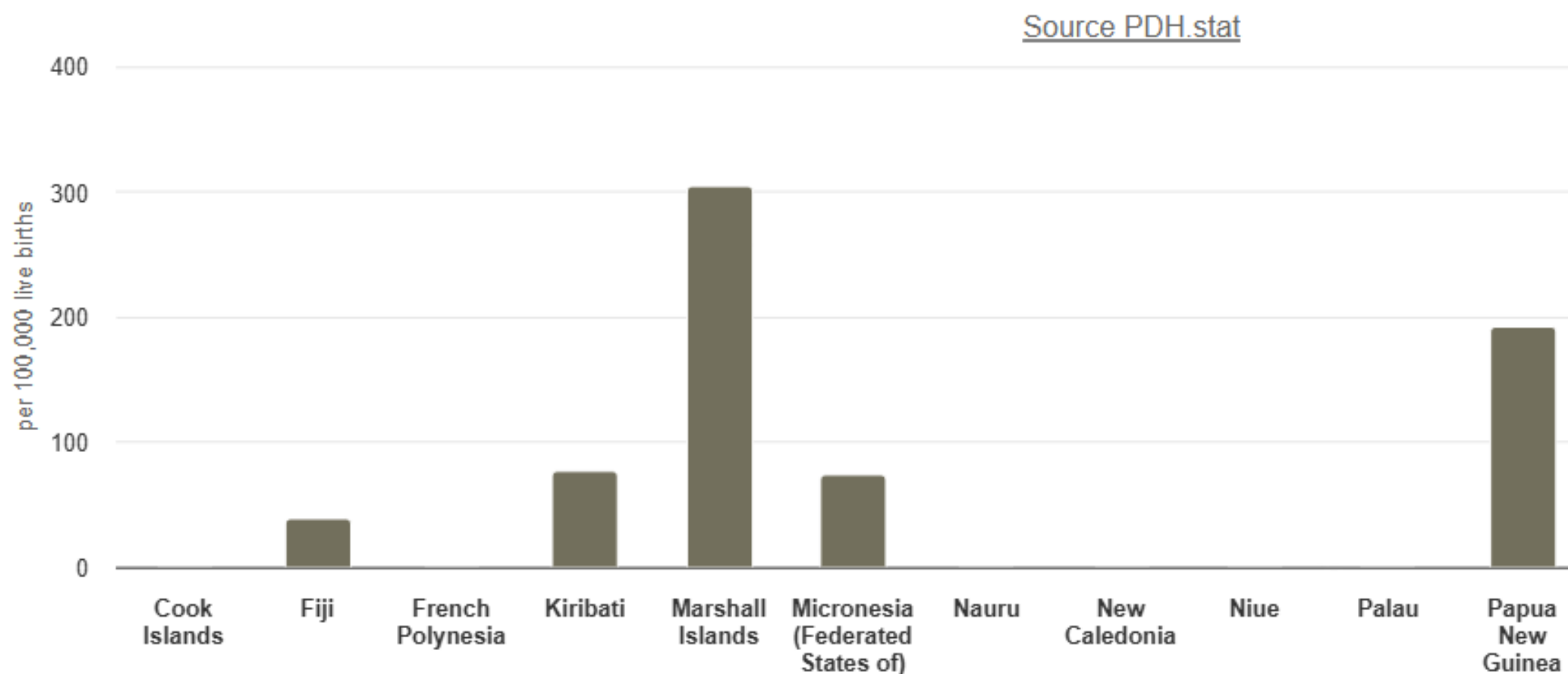
SDMX MCP Gateway

progressive discovery of data

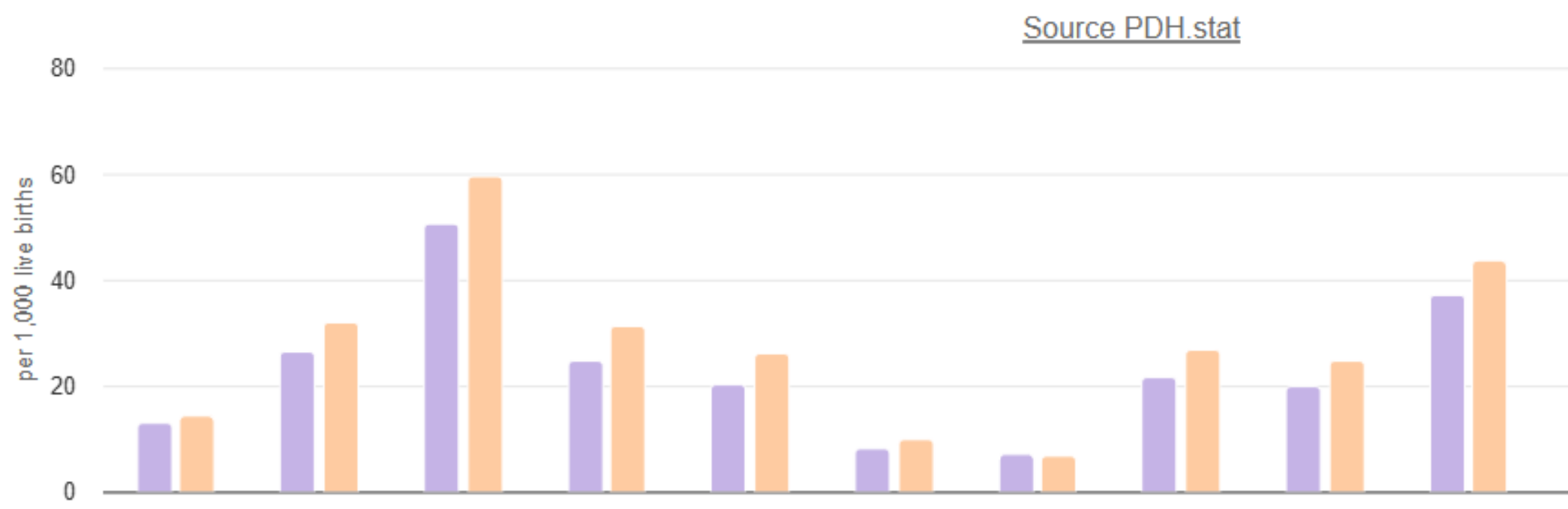
a standard for LLM tools

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached zero in the graph below.



Under-five mortality rate



SDMX MCP Gateway

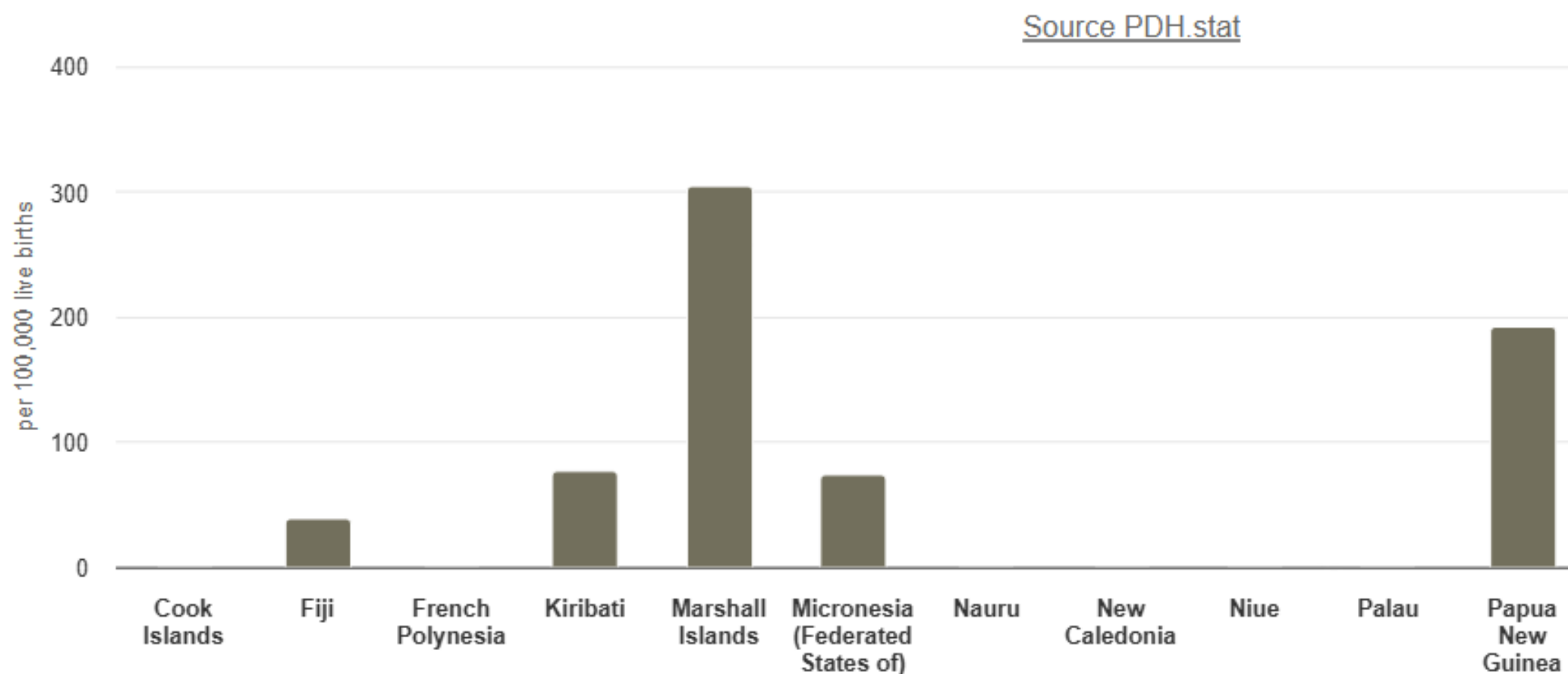
NOT AI: a *tool* for AI clients

Fast
Cheap
Light

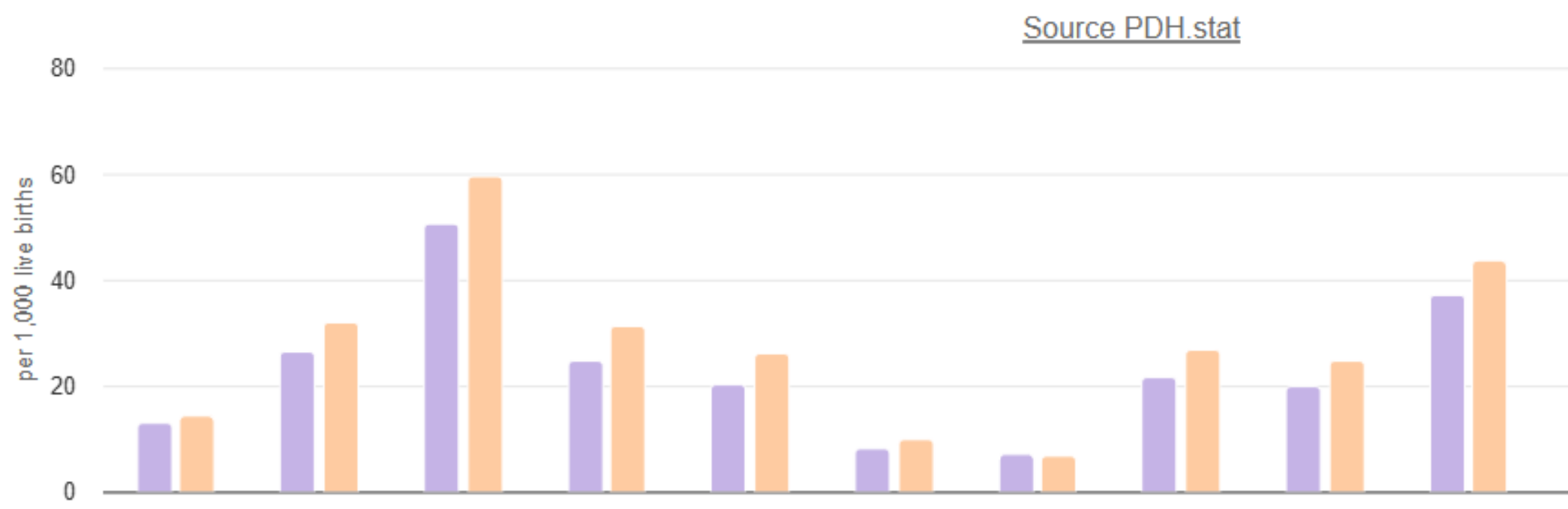
...

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached zero in the graph below.



Under-five mortality rate



SDMX MCP Gateway

Overview (~300 bytes) - Find dataflows

Structure (~1KB) - Understand structure

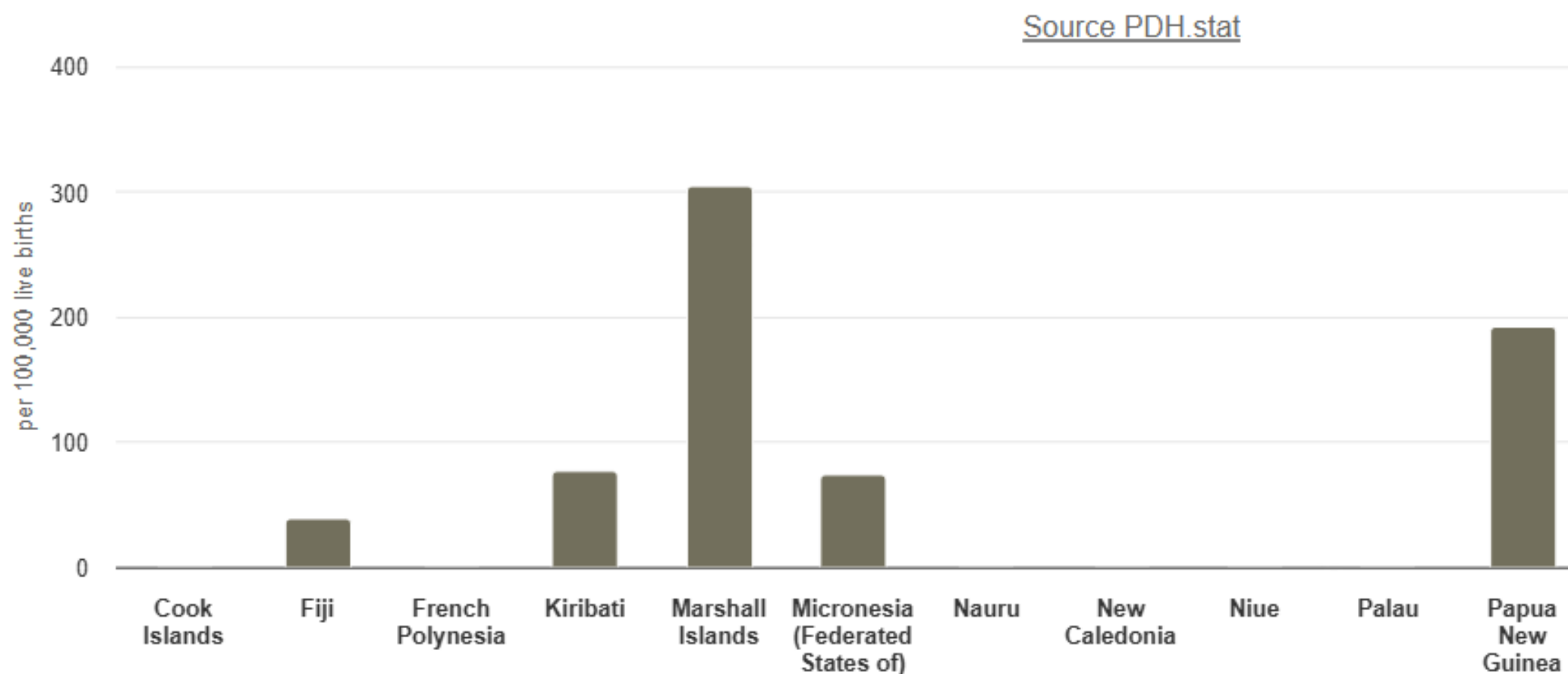
Drill-down (~500 bytes) - Get codes

Availability (~700 bytes) - Check data

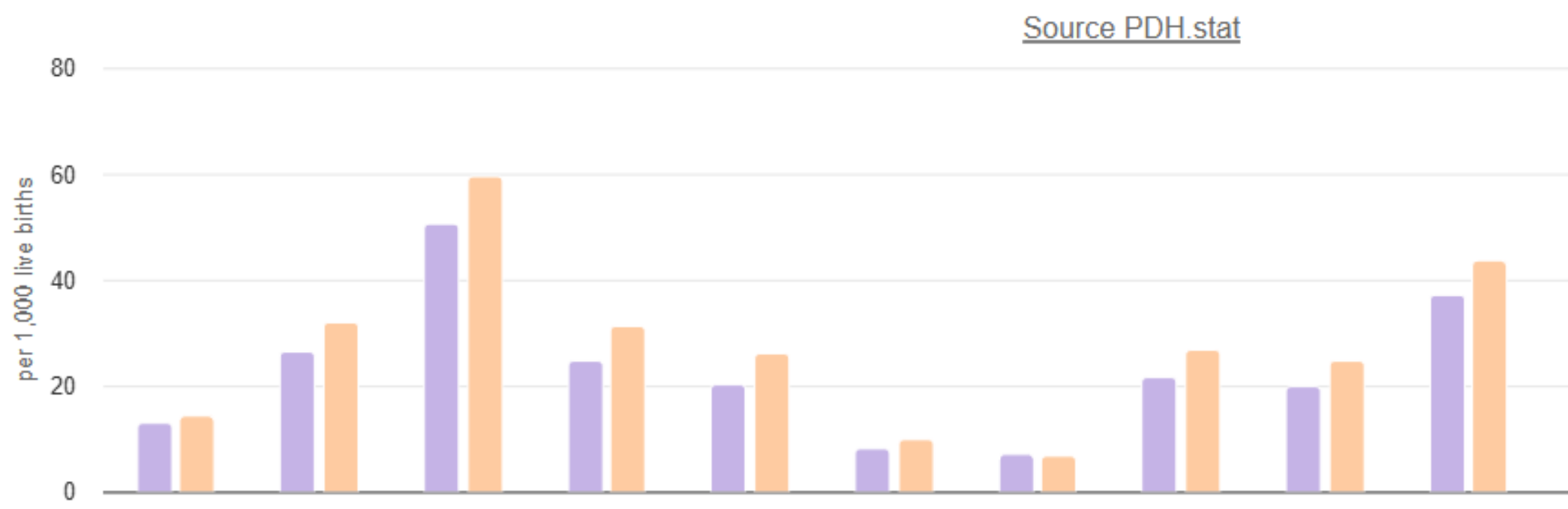
Query (~200 bytes) - Build query URL

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached zero in the graph below.



Under-five mortality rate



SDMX MCP Gateway

It uses the public SDMX .Stat API

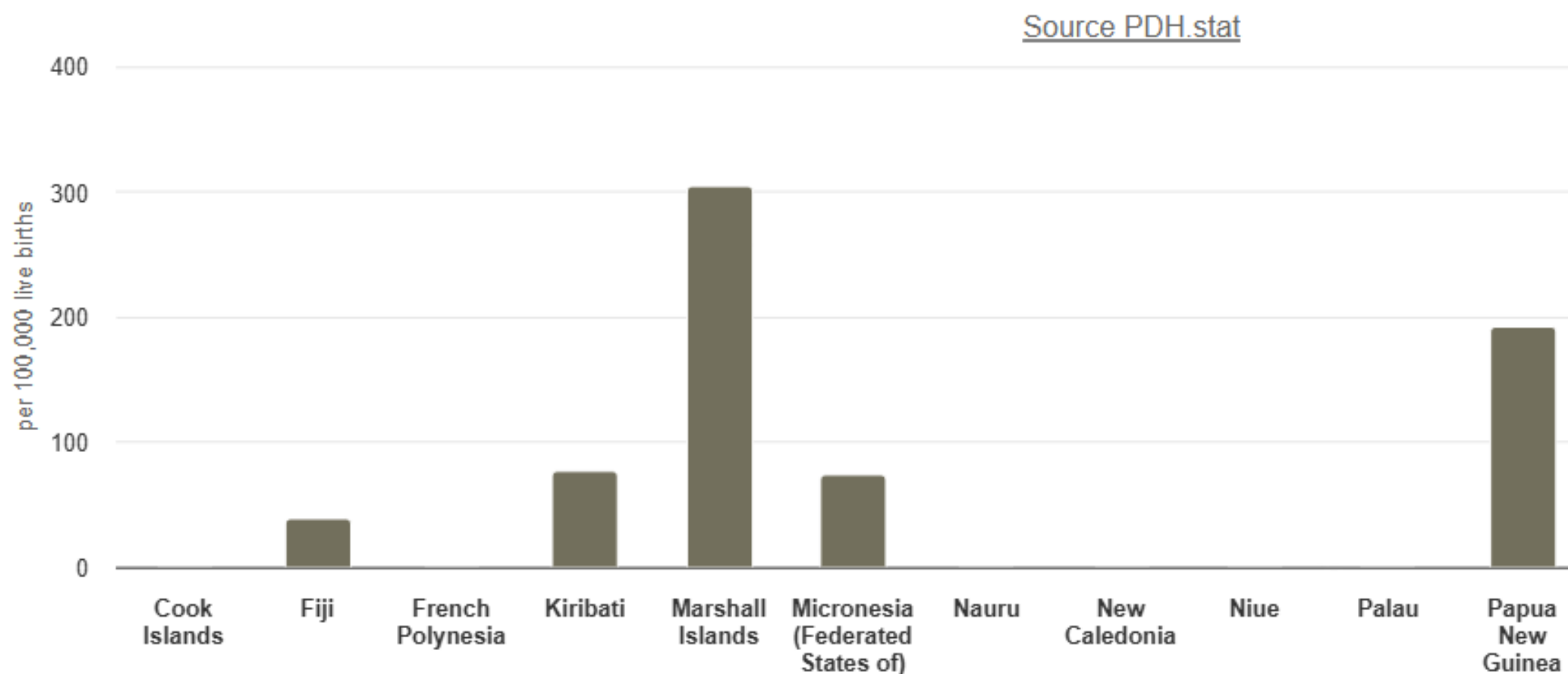
Light configuration (default endpoint, ...)

Doesn't store data (nor structures)

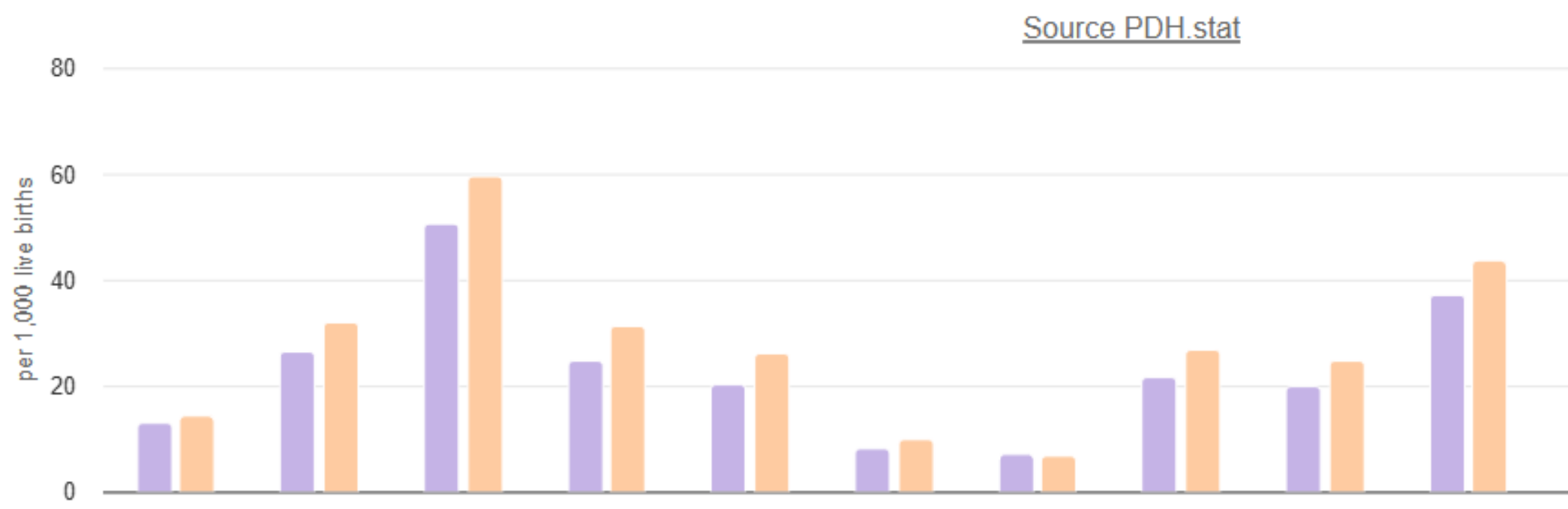
Provide help, not structure

Maternal mortality ratio

The 2030 target for maternal mortality is less than 70 deaths per 100,000 live births. A number of countries have reached zero in the graph below.



Under-five mortality rate



SDMX MCP Gateway

Way more code to wrap the SDMX API

Some personal choices:
relies heavily on .Stat “developer” links

03.3

**Wrangle raw data
to SDMX-CSV ready
with LLMs**

“Can you draft a data transformation script to take this messy Excel into an SDMX-CSV for this dataflow?”

Early prototypes in Julia / R

04

Lessons learnt

Small is cool:

- Let big labs develop big GPTs
- Write lean tools for LLMs
- Are you sure you want to pay AI for everyone?
- Maintain sovereignty over data and tools
- Most challenges are “classic” dev issues

Small is cool:

- Let big labs develop big GPTs
- Write lean tools for LLMs
- Are you sure you want to pay AI for everyone?
- Maintain sovereignty over data and tools
- Most challenges are “classic” dev issues

Thank you so much!

Giulio Valentino Dalla Riva

giuliodr@spc.int – github.com/gvdr