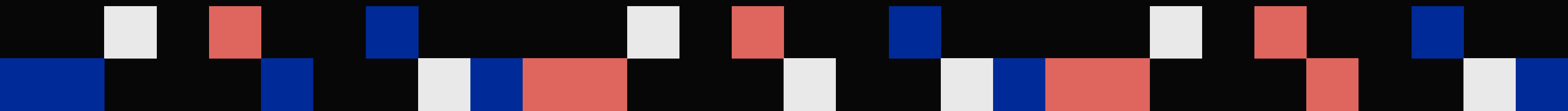


DATA PROJECT 1

DATA JOBS AROUND EUROPE

IRONHACK DATA PT MADRID

by Marina Fdez. Banda



How are data jobs distributed?

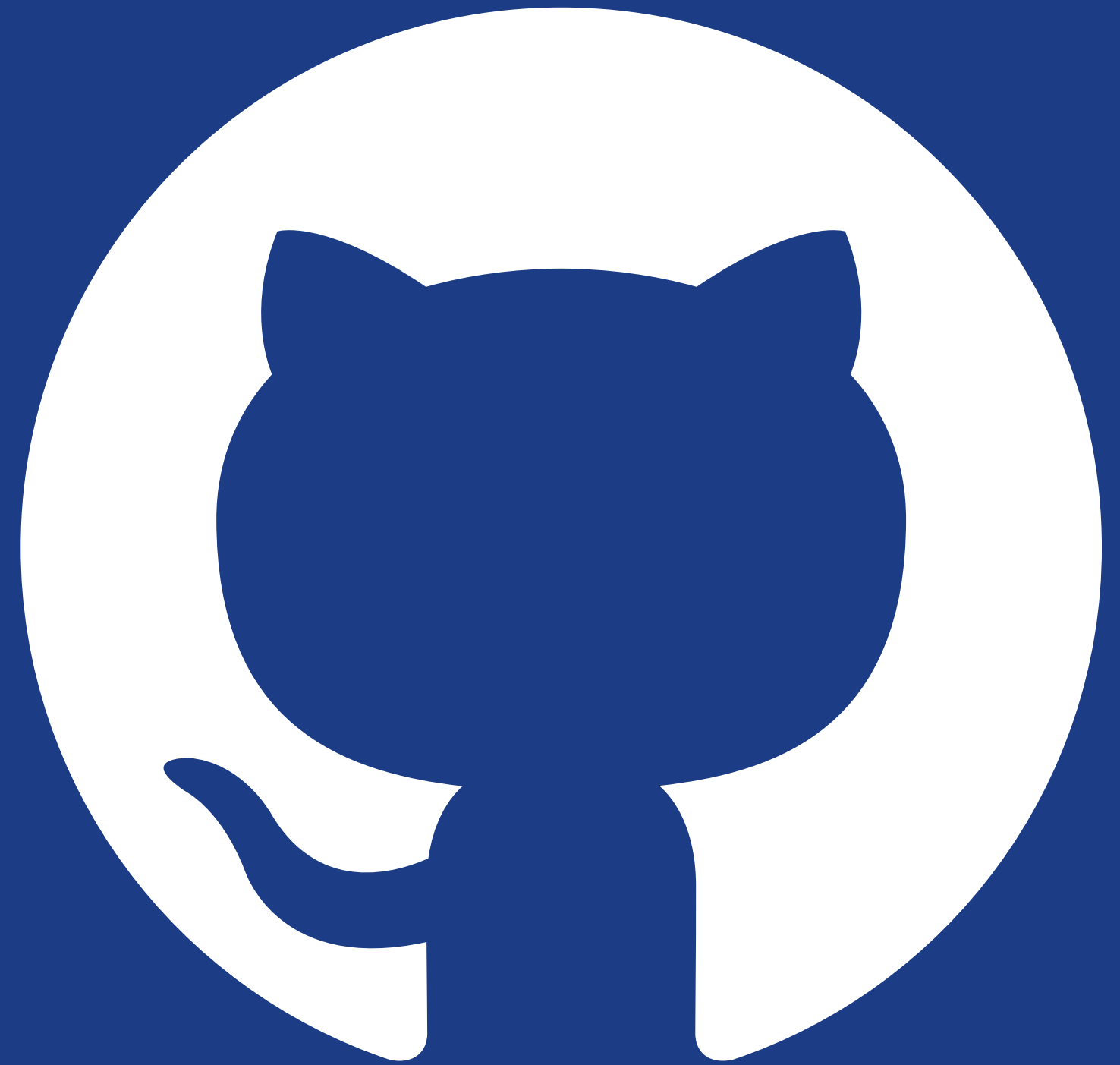
Data-related jobs are growing exponentially in recent years.



How are data jobs distributed?

Data-related jobs are growing exponentially in recent years.

This has led to an increase in demand and more and more people are interested in pursuing a future in this area.

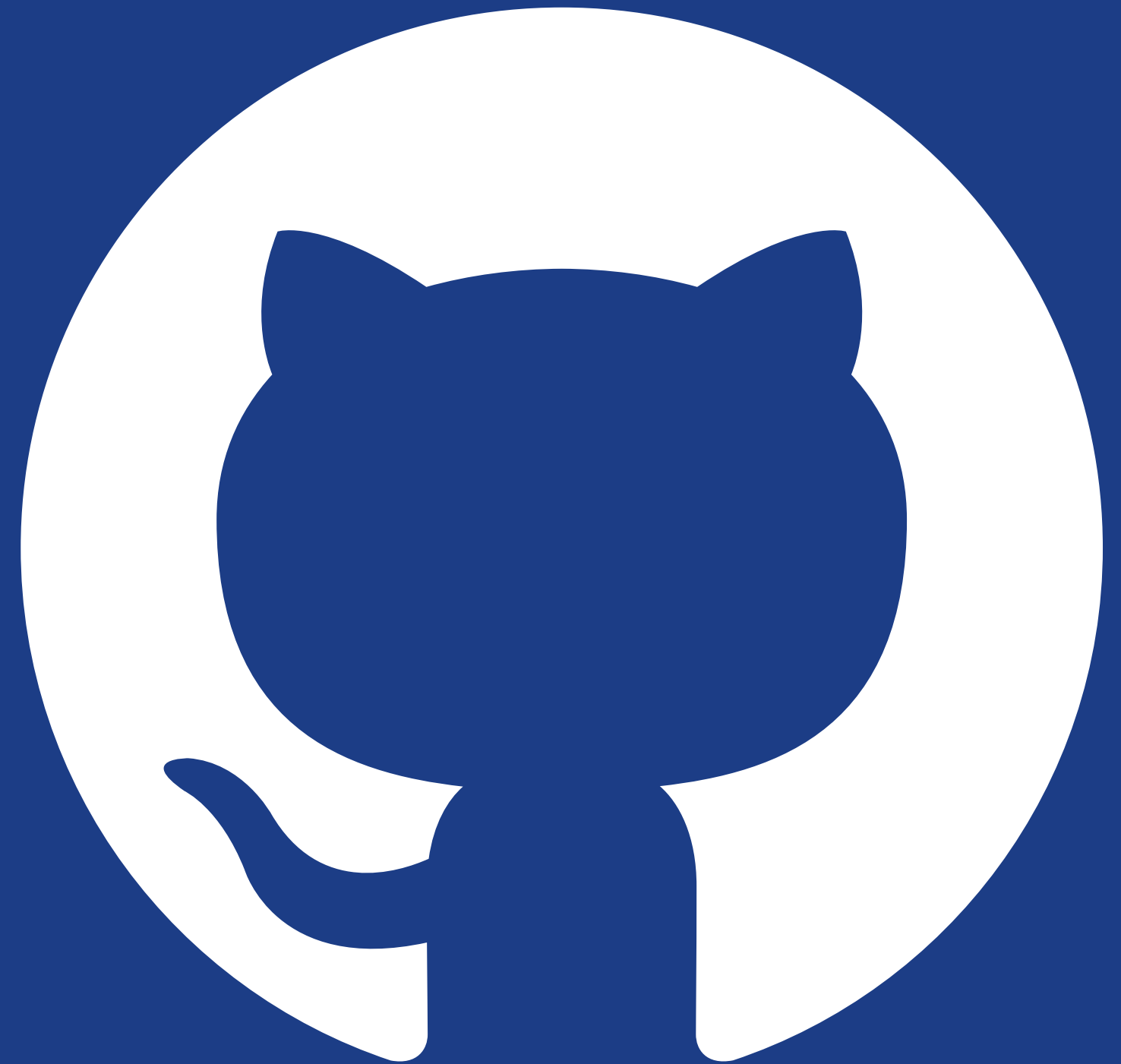


How are data jobs distributed?

Data-related jobs are growing exponentially in recent years.

This has led to an increase in demand and more and more people are interested in pursuing a future in this area.

But... how well distributed are these jobs in Europe?



How are data jobs distributed?

Data Sources

Tables(.db)

API

Web Scraping

First approach...

How it started



How it's going



Py . . . Charming?



How are data jobs distributed?

Data Sources

Tables(.db)

API

Web Scraping

Methodology

Export

Clean

Merge

Analyze

How are data jobs distributed?

Data Sources

Tables(.db)
API
Web Scraping

Methodology

Export
Clean
Merge
Analyze

Stack

Python
sqlalchemy
Pandas
Numpy
Requests
Matplotlib
Argparse

How are data jobs distributed?

Data Sources

Tables(.db)
API
Web Scraping

Methodology

Export
Clean
Merge
Analyze

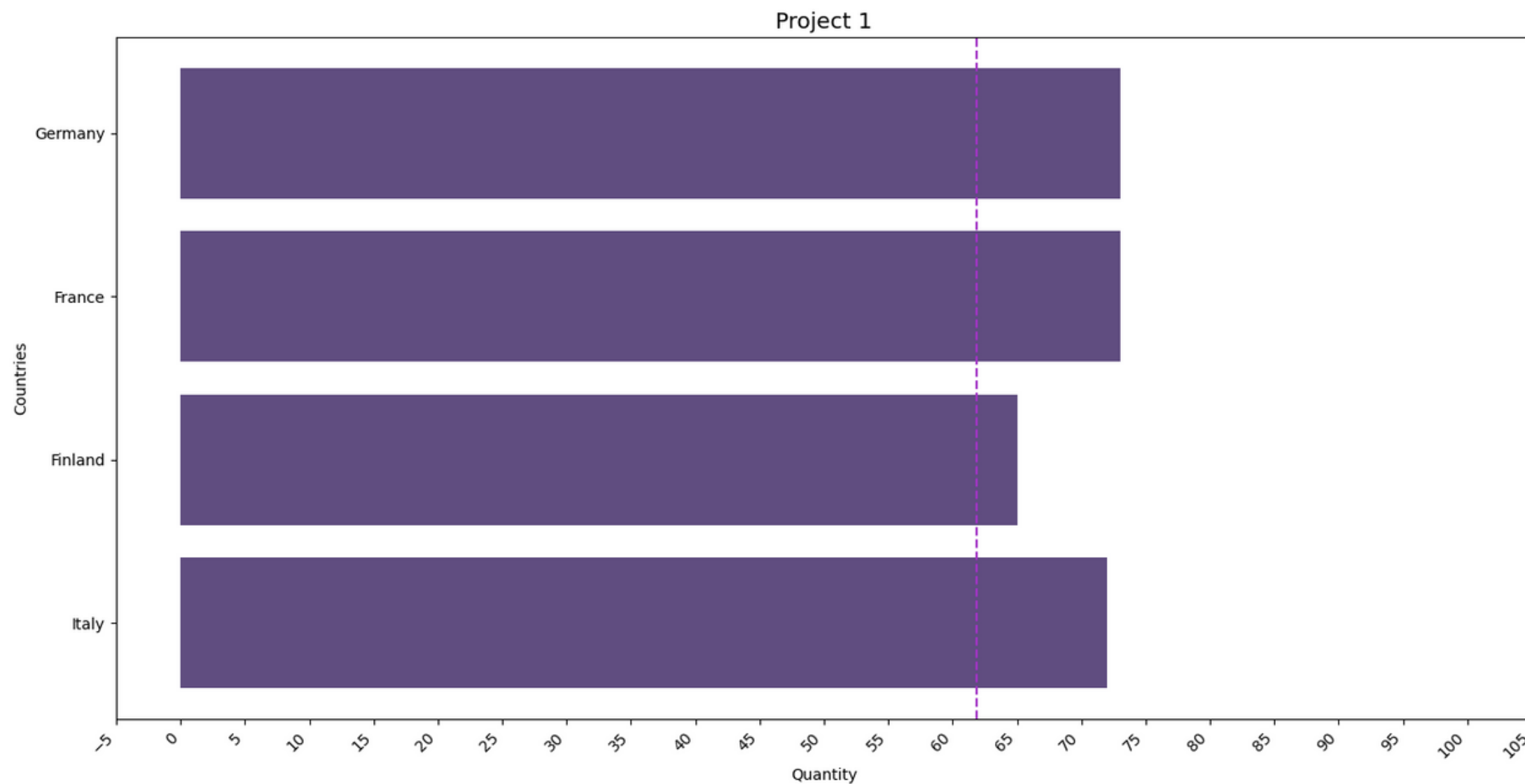
Stack

Python
sqlalchemy
Pandas
Numpy
Requests
Matplotlib
Argparse

Goal

Analyze the
distribution
of data-
related jobs
across
Europe.

How are data jobs distributed?

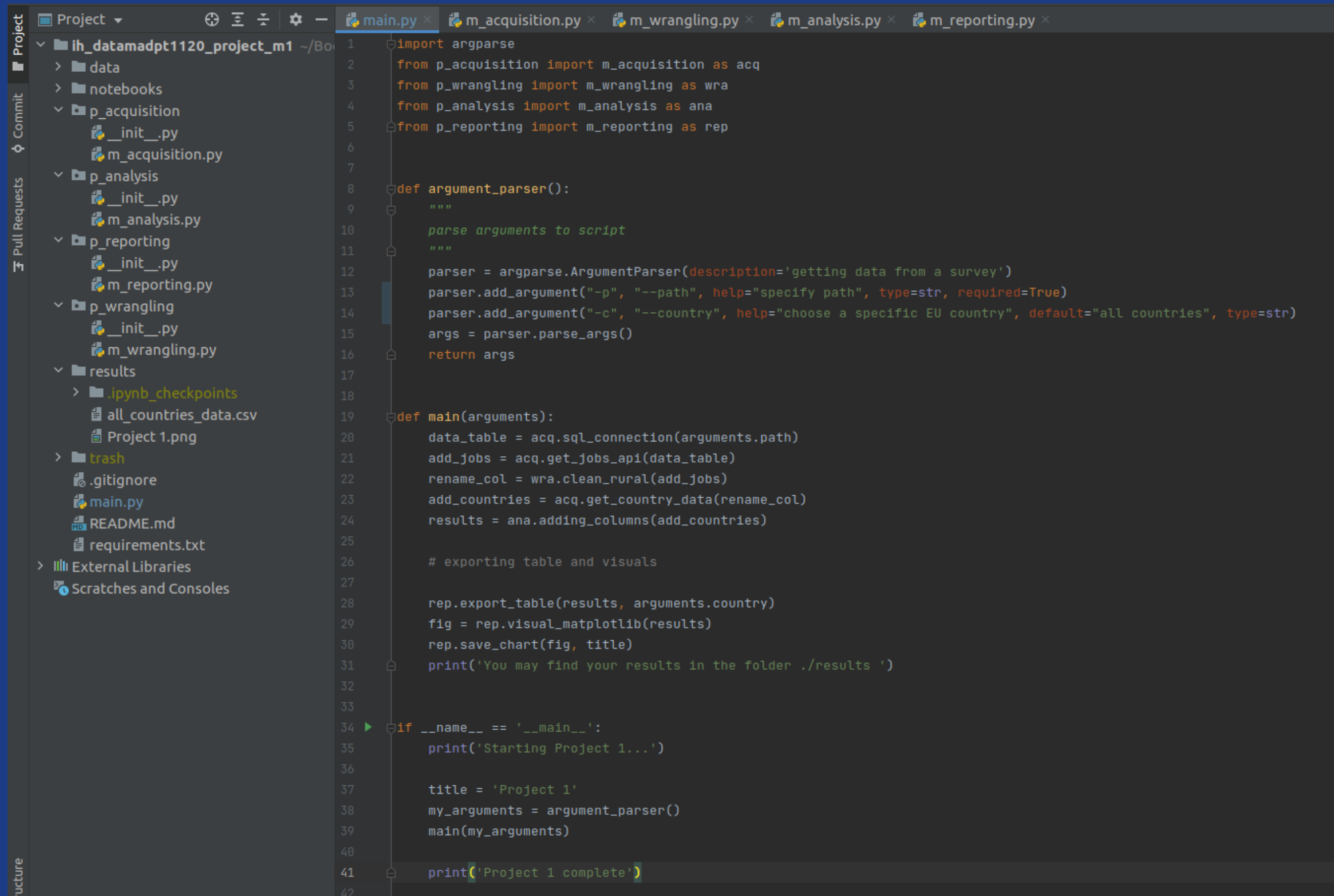


Those countries
have a high
volume of data
experts...

Dare to look for
another country
where you can
excel!



A sample of the code used in the project



The screenshot displays a JupyterLab environment. On the left, the 'Project' sidebar shows a file tree for 'ih_datamadpt1120_project_m1'. The tree includes folders for 'data', 'notebooks', and 'p_acquisition', 'p_analysis', 'p_reporting', and 'p_wrangling'. Each of these folders contains an '__init__.py' file and a corresponding module file (e.g., 'm_acquisition.py', 'm_analysis.py', etc.). There are also files like 'all_countries_data.csv', 'Project 1.png', '.gitignore', 'main.py', 'README.md', and 'requirements.txt' in the root. The 'External Libraries' and 'Scratches and Consoles' sections are also visible.

The main editor area shows the 'main.py' file with the following Python code:

```
1 import argparse
2 from p_acquisition import m_acquisition as acq
3 from p_wrangling import m_wrangling as wra
4 from p_analysis import m_analysis as ana
5 from p_reporting import m_reporting as rep
6
7
8 def argument_parser():
9     """
10     parse arguments to script
11     """
12     parser = argparse.ArgumentParser(description='getting data from a survey')
13     parser.add_argument("-p", "--path", help="specify path", type=str, required=True)
14     parser.add_argument("-c", "--country", help="choose a specific EU country", default="all countries", type=str)
15     args = parser.parse_args()
16     return args
17
18
19 def main(arguments):
20     data_table = acq.sql_connection(arguments.path)
21     add_jobs = acq.get_jobs_api(data_table)
22     rename_col = wra.clean_rural(add_jobs)
23     add_countries = acq.get_country_data(rename_col)
24     results = ana.adding_columns(add_countries)
25
26     # exporting table and visuals
27
28     rep.export_table(results, arguments.country)
29     fig = rep.visual_matplotlib(results)
30     rep.save_chart(fig, title)
31     print('You may find your results in the folder ./results ')
32
33
34 if __name__ == '__main__':
35     print('Starting Project 1...')
36
37     title = 'Project 1'
38     my_arguments = argument_parser()
39     main(my_arguments)
40
41     print('Project 1 complete')
```

```
print( 'Thank you! ' )
```