Global assessment of Big Data for Official Statistics

Ronald Jansen

Assistant Director

United Nations Statistics Division

Overview

- 2015 global assessment of Big Data for Official Statistics
  - Structure of the Questionnaire
  - Management of Big Data
  - Overall results
- Big Data general responses
- Big Data Projects

Structure of the Questionnaire

Part I

Management of Big Data (Questions 1 - 7)
Advocacy and communication (Question 8)
Linking Big Data and SDGs (Questions 9 - 11)
Access, privacy and confidentiality (Questions 12 - 13)
Skills and training (Questions 14 - 17)

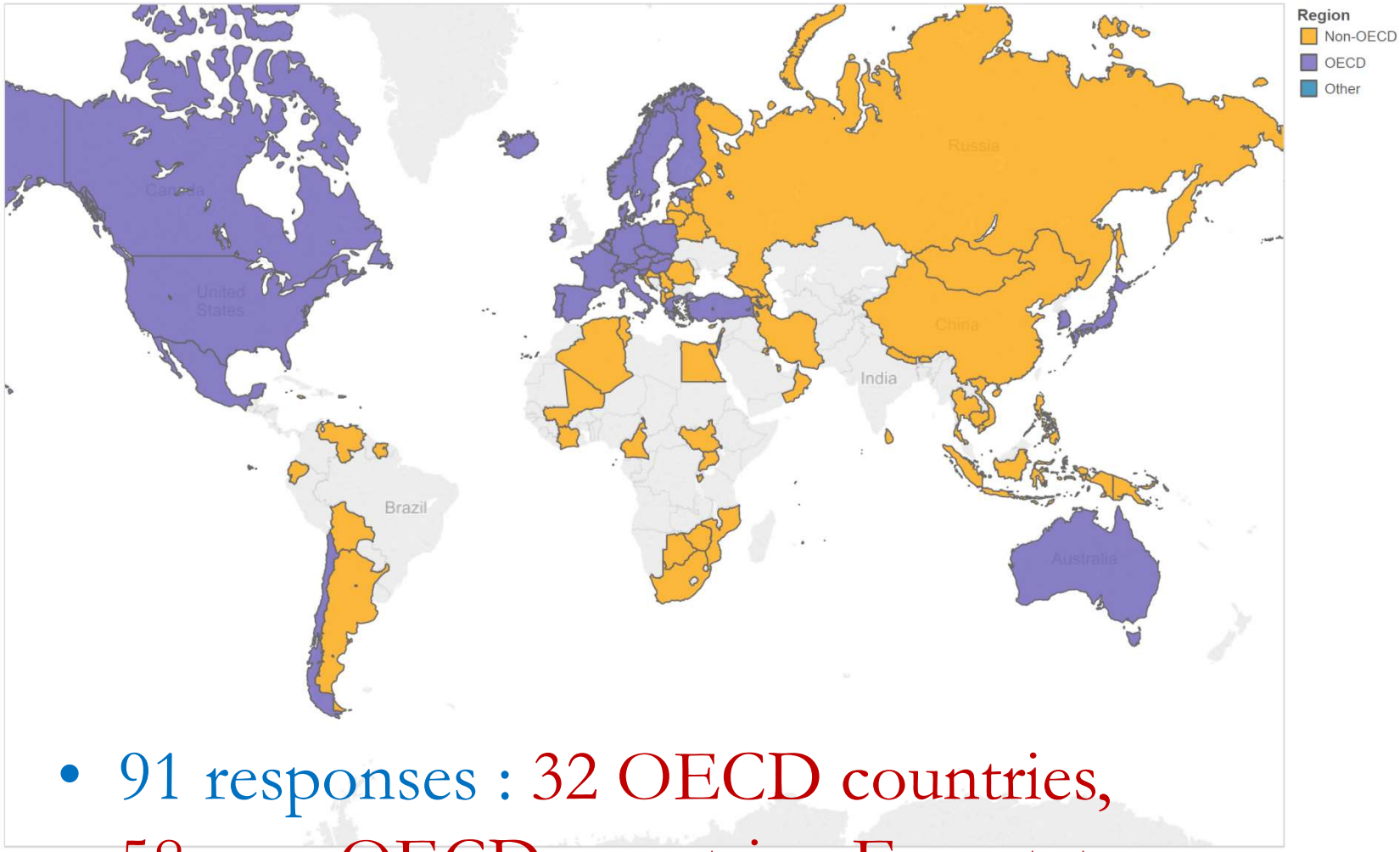
Part II

Needs assessment (Question 18)

Part III

Project availability (Question 19)
Project 1 (Questions 20 - 43)
Project 2 (Questions 44 - 67)
Project 3 (Questions 68 - 91)
Project 4 (Questions 92 - 115)
Project 5 (Questions 116 - 138)

Responses



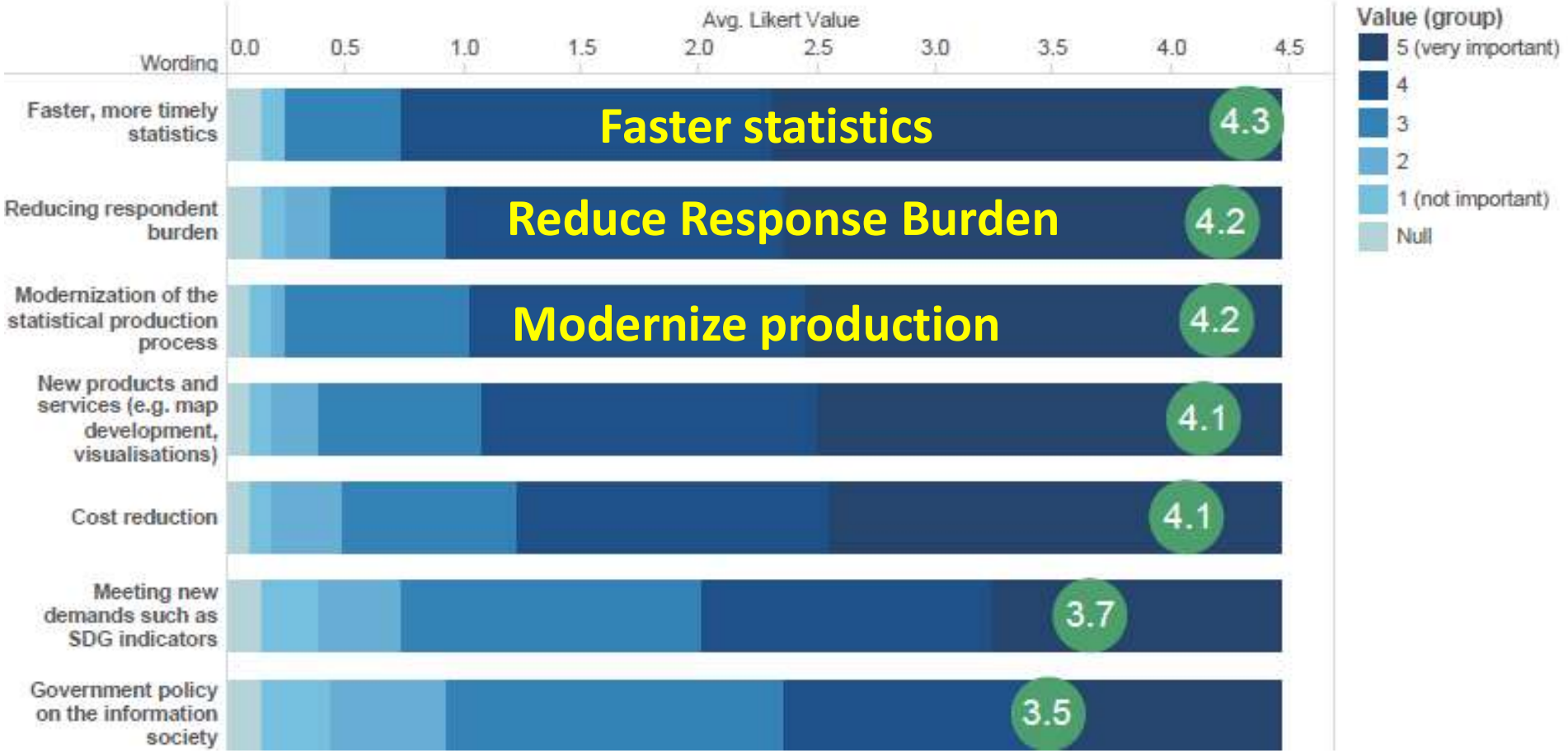
- 91 responses : 32 OECD countries, 58 non-OECD countries, Eurostat

Responses

- 115 Big Data projects were submitted:
  - 89 from OECD countries
  - 22 from non-OECD countries
  - 4 from regional organization

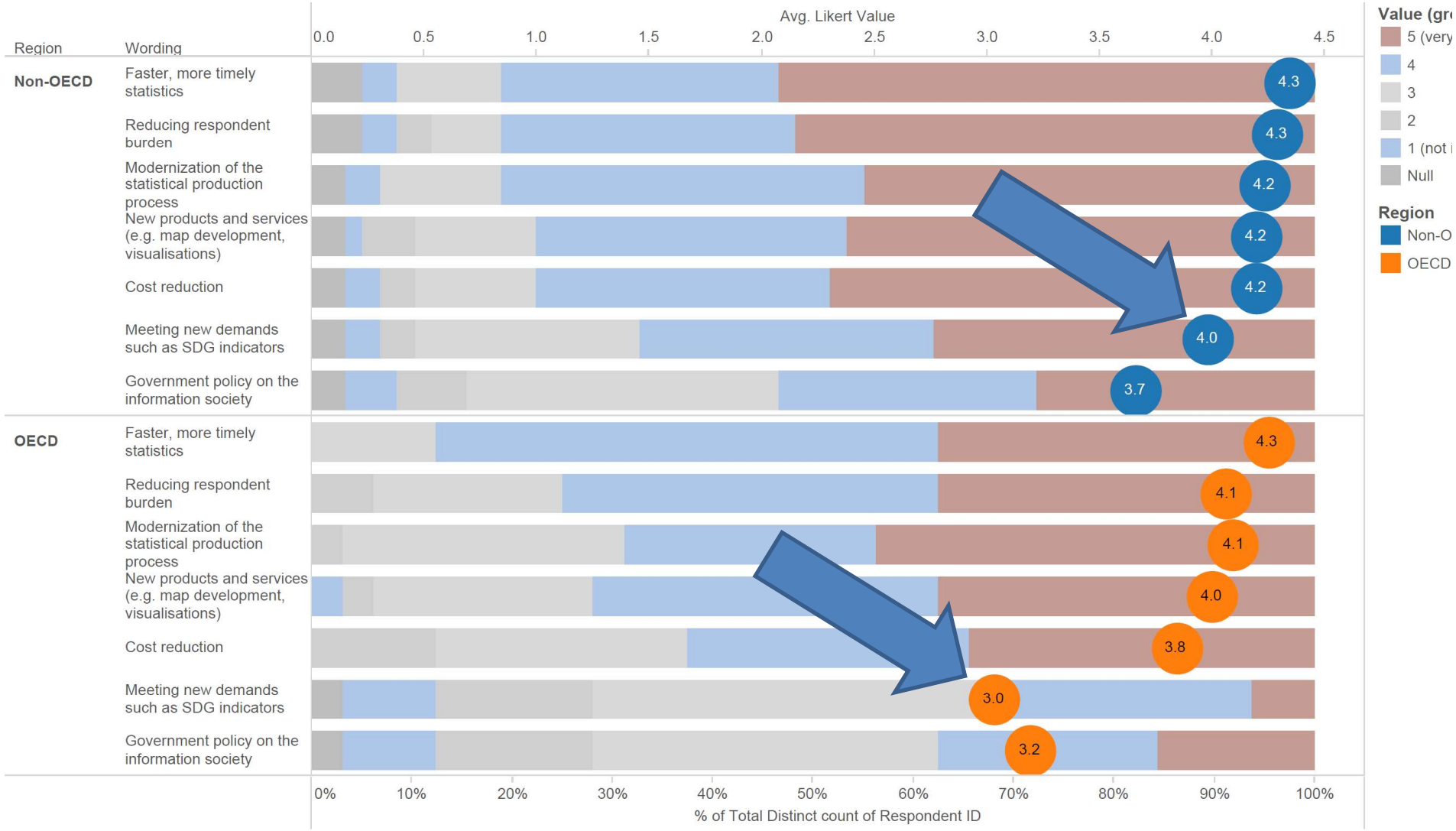
Management of Big Data

What do you see as the main reasons or business benefits for using Big Data in your institute?



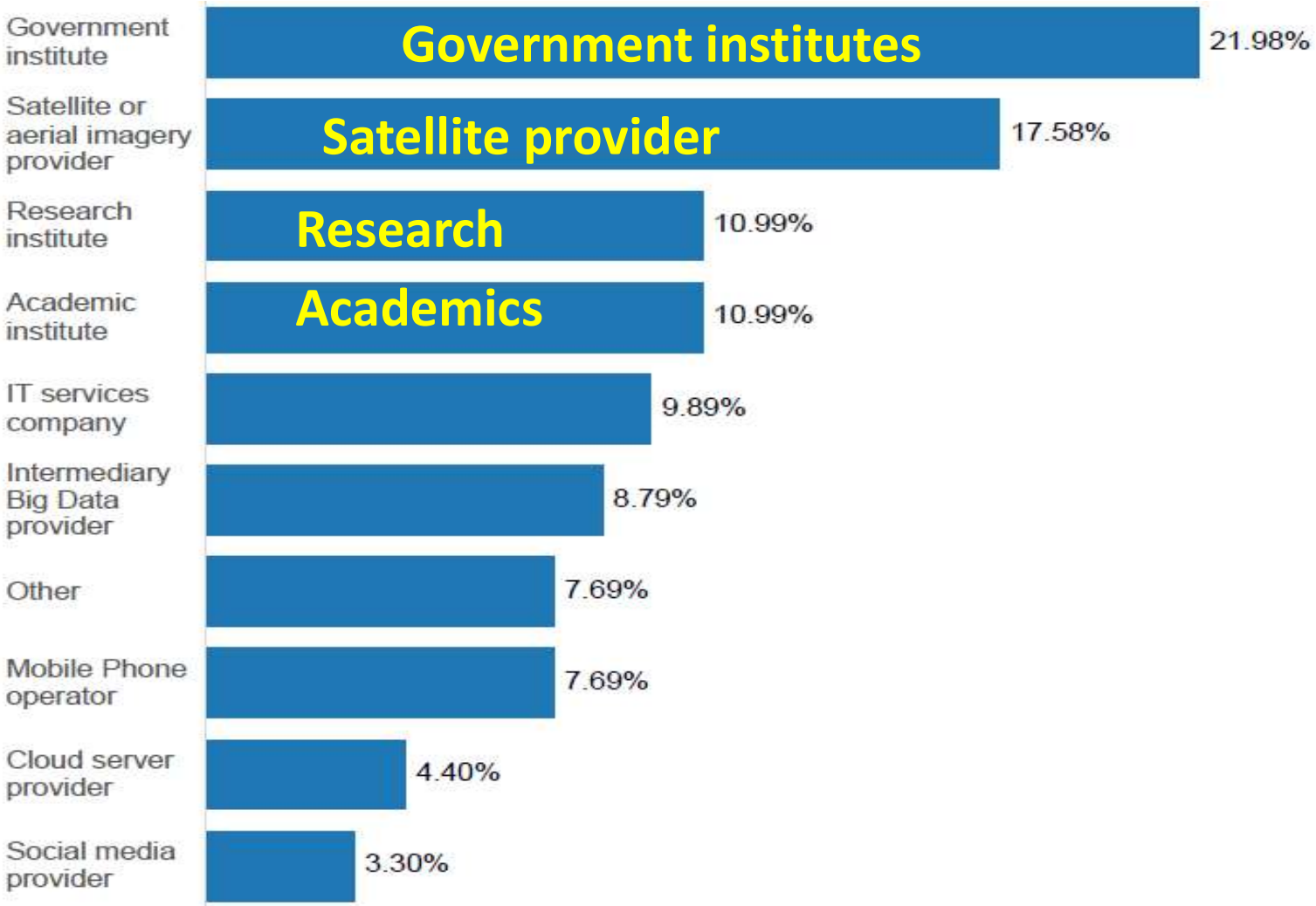
Big Data - General responses

Area of Statistics	Partnership	Sources	Sources - vertical presentation	Guidance	Activities needed	Skills needed	Busine



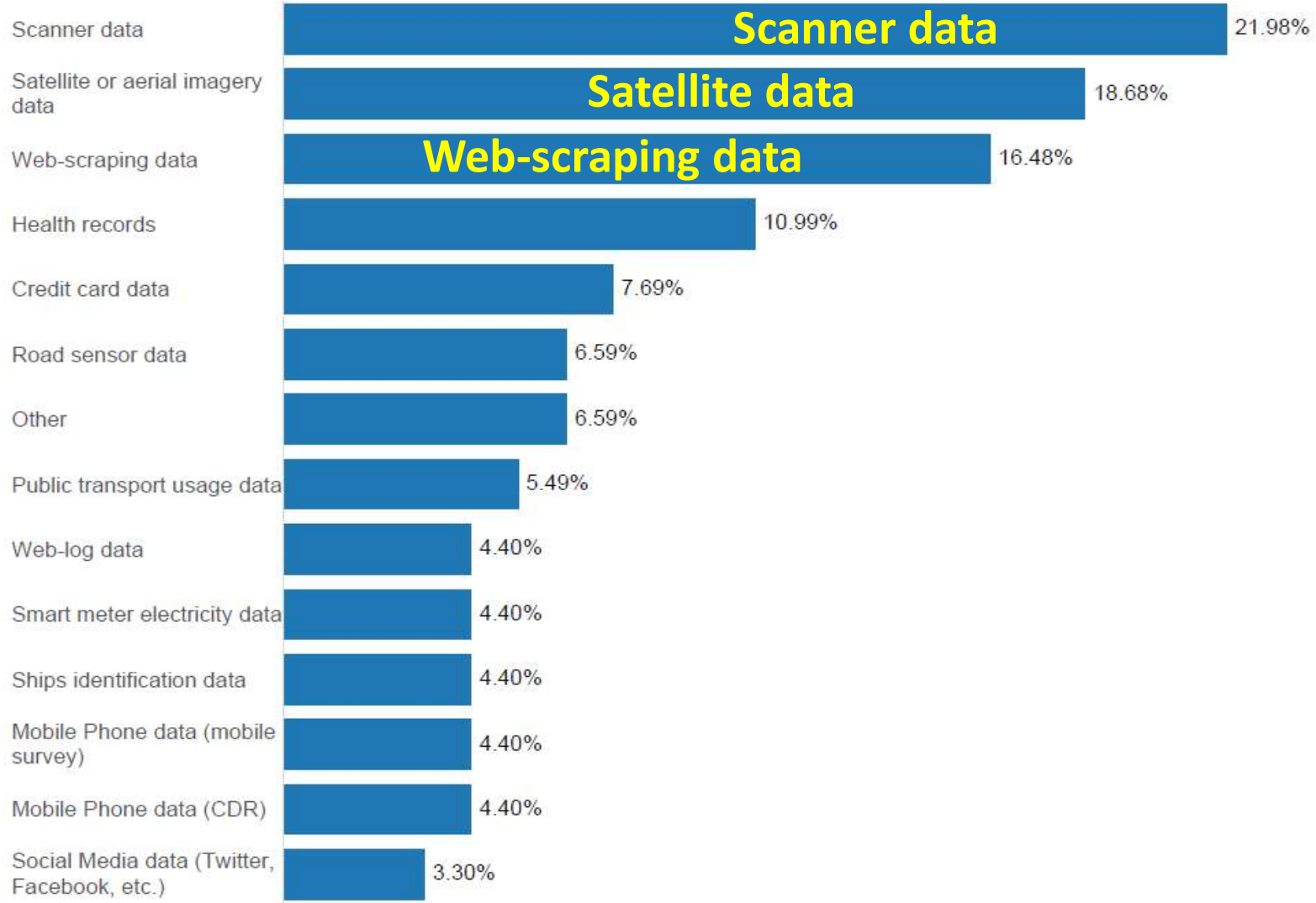
Management of Big Data

Which specific Big Data **partnerships** have you established?

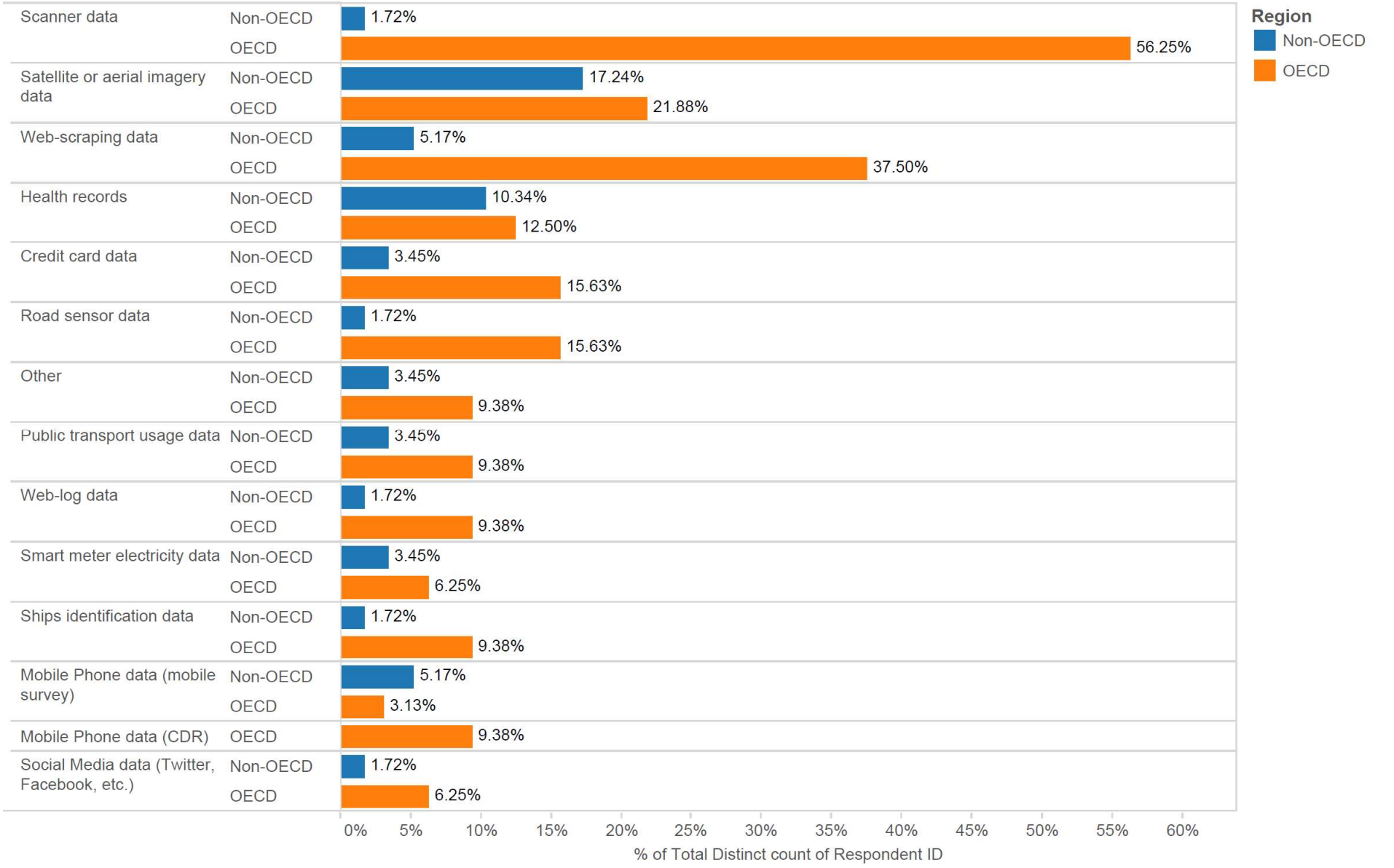


Management of Big Data

Which specific Big Data sources have you used ?

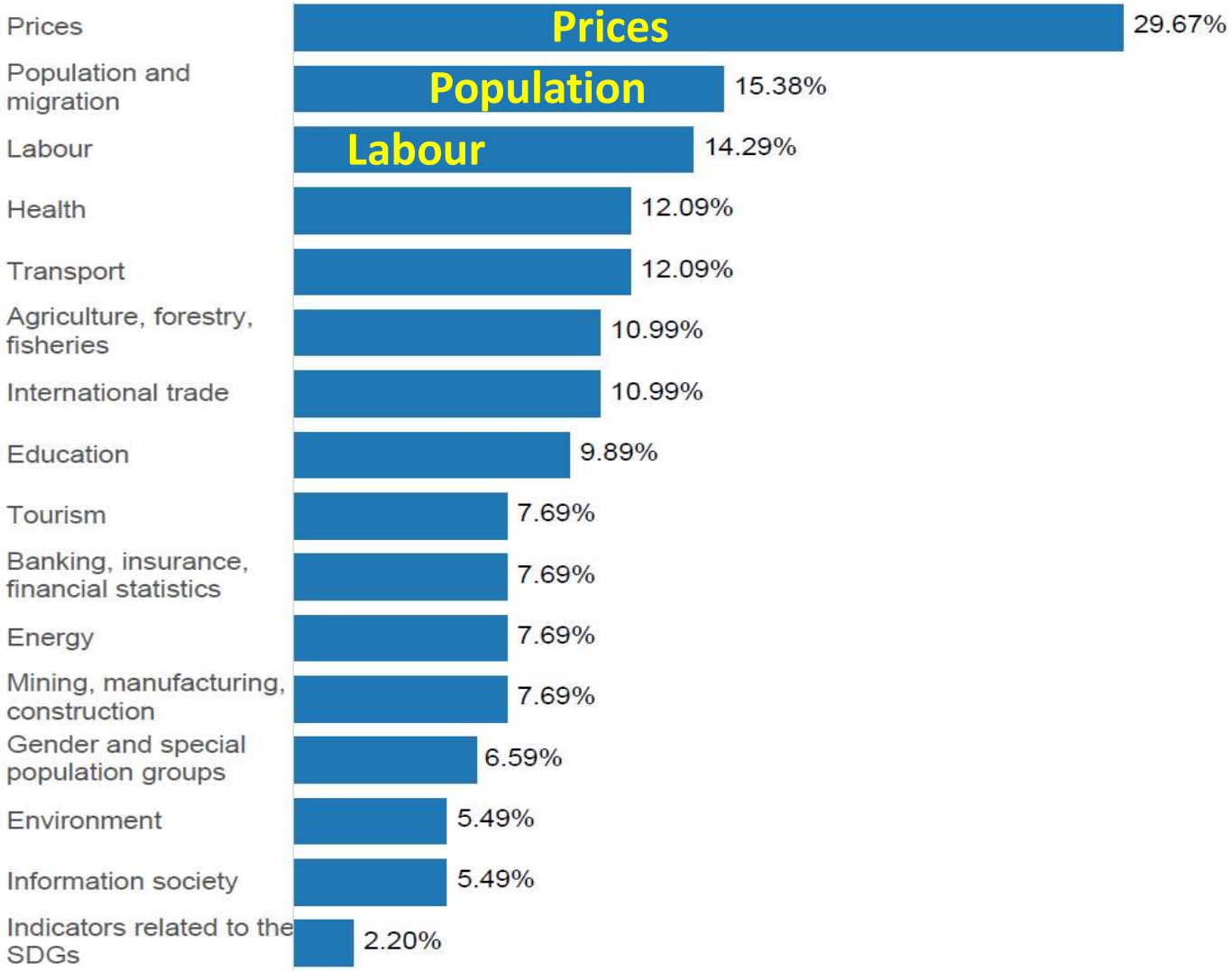


Big Data Sources

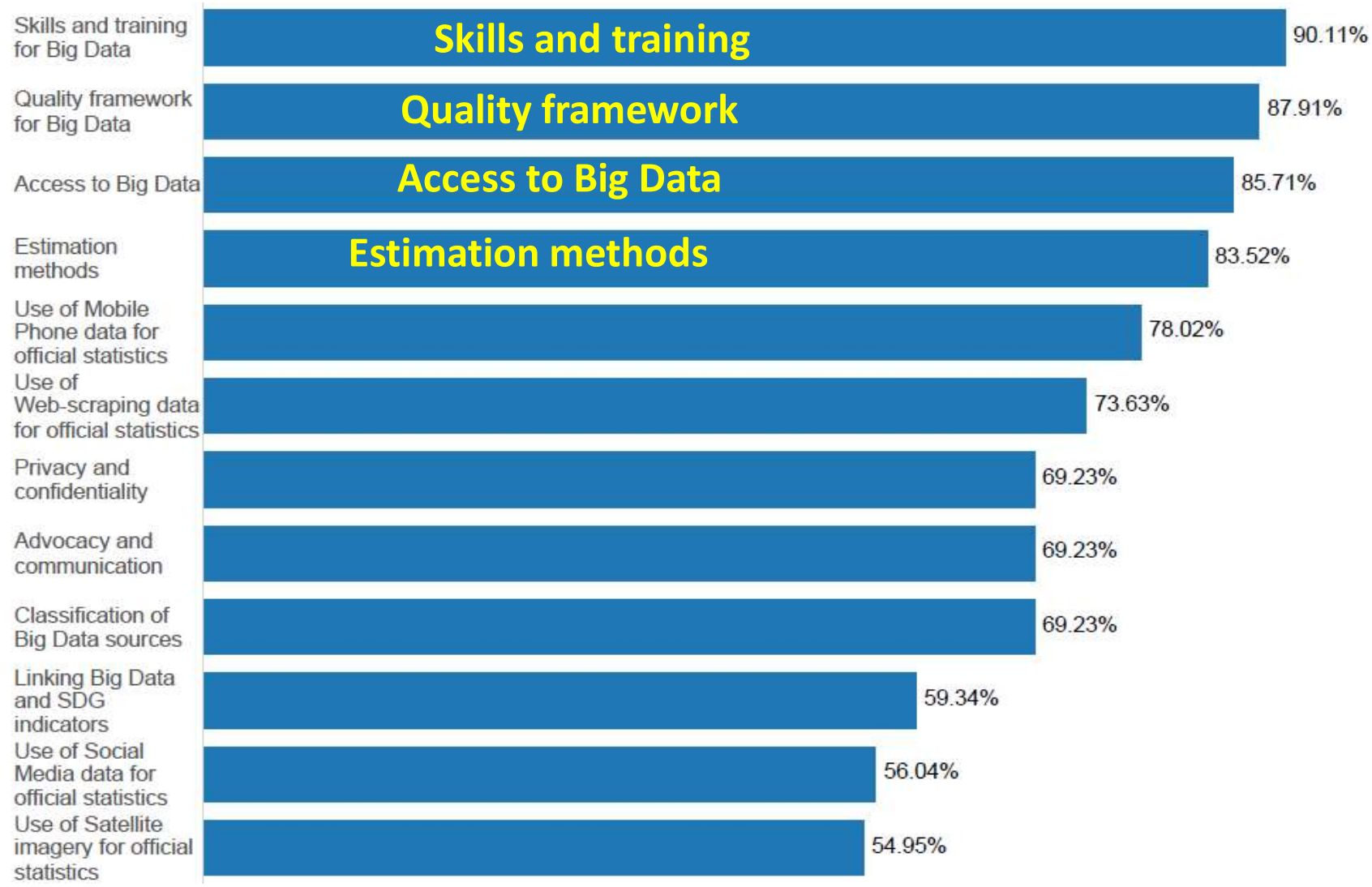


Management of Big Data

In which statistical domains did you use Big data ?



Needs assessment



On which topics do you see an urgent need for guidance ?

Objective(s)

Exploration

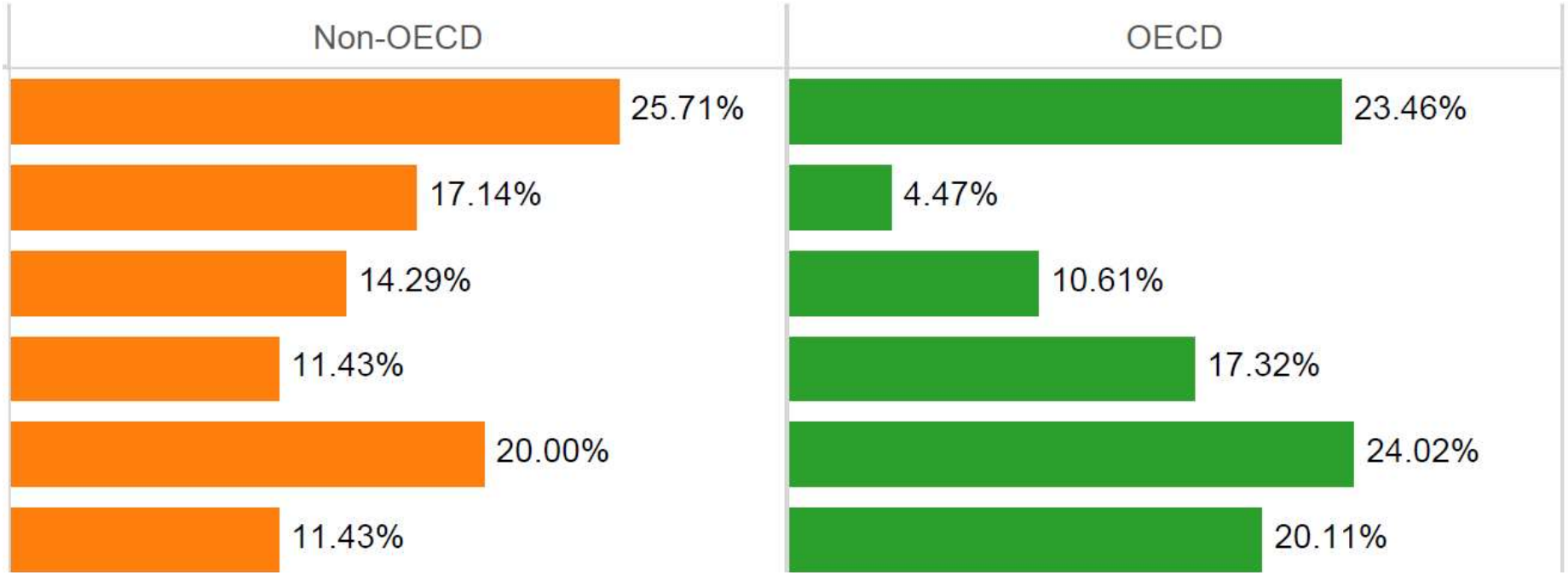
Other

Pilot intended to go to production to improve timeliness

Pilot intended to go to production to replace existing data

Pilot intended to go to production to supplement existing data

Scientific / research



Summary of main responses

- Sources
  - **The most used sources are scanner data, satellite imagery and web-scraping data.**
  - **Social media and mobile phone data are much less used.**
- Partnership
  - **The most common partnership is with other government institutes, followed by satellite imagery providers.**
- Area of statistics
  - **Big Data has been mostly applied to price statistics, followed by population/migration and labour statistics.**
  - **Big Data is not commonly used for the indicators related to the SDGs.**

Summary of main responses

Business benefits

- The two main reasons to use Big Data for official statistics are to provide faster, more timely statistics and to reduce the response burden.
- Meeting new demands related to the SDGs is deemed less important.

Guidance

- The statistical community expressed the need for guidance in the areas of skills and training, quality frameworks for Big Data and access to Big Data.
- There is also demand for guidance on the usage of specific data sources such as web-scarping and mobile phone data.

Thank you