STATBEL
Belgium in figures

Accessing privately held data in Belgium: Insights & challenges

Marie Vandresse, DG Statistics Belgium

10-04-2025

Content

1. Experiences with the use of privately held data: two examples
2. Challenges
3. Implementation of the new 223/2009

STATBEL

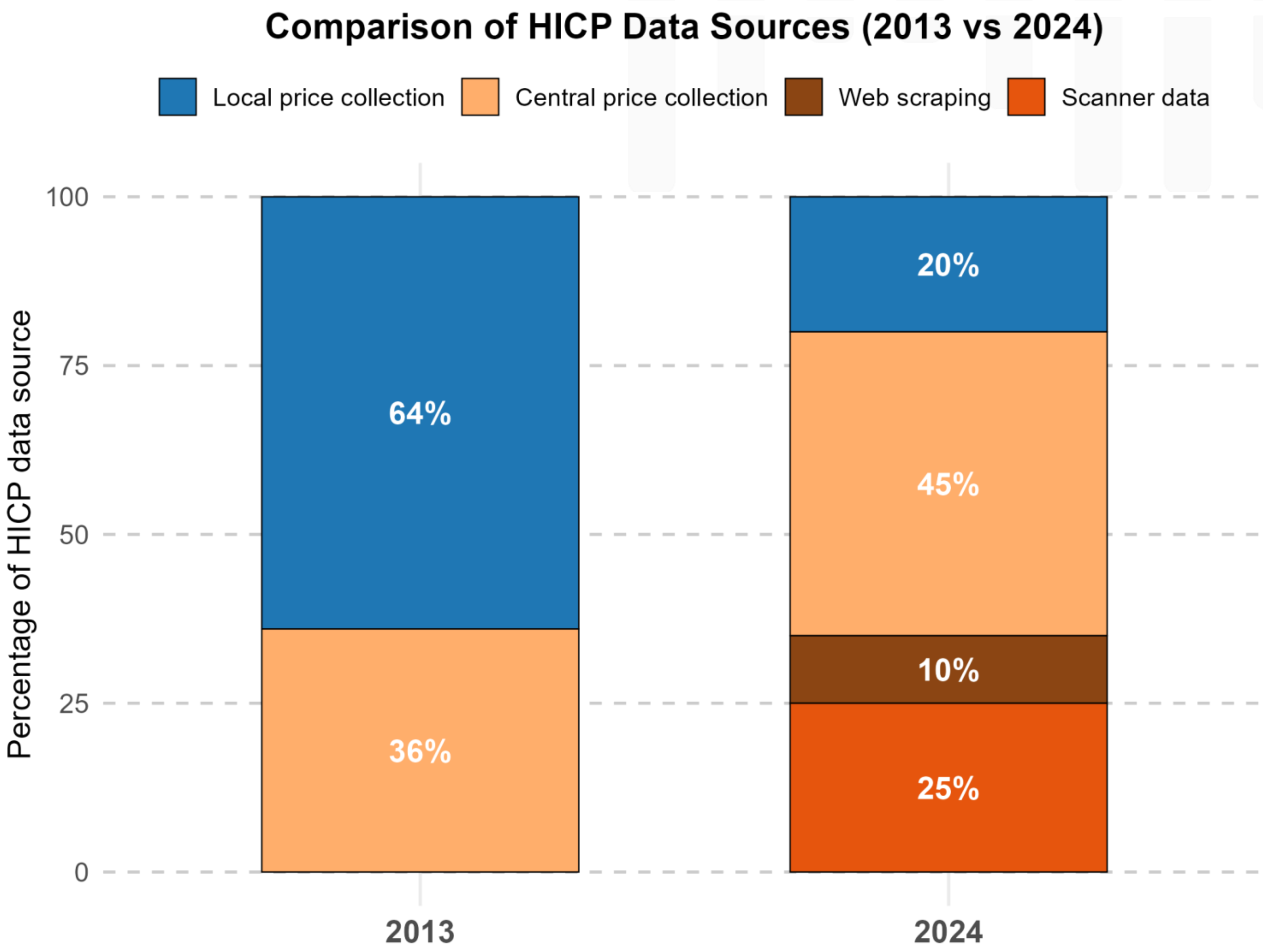
01

Experiences

- CPI/HICP
- Transport

STATBEL

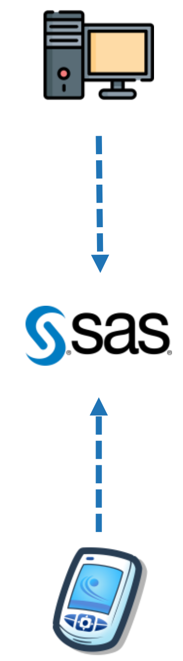
1. The challenge of modernizing the CPI/HICP



Example 1: CPI/HICP

STATBEL

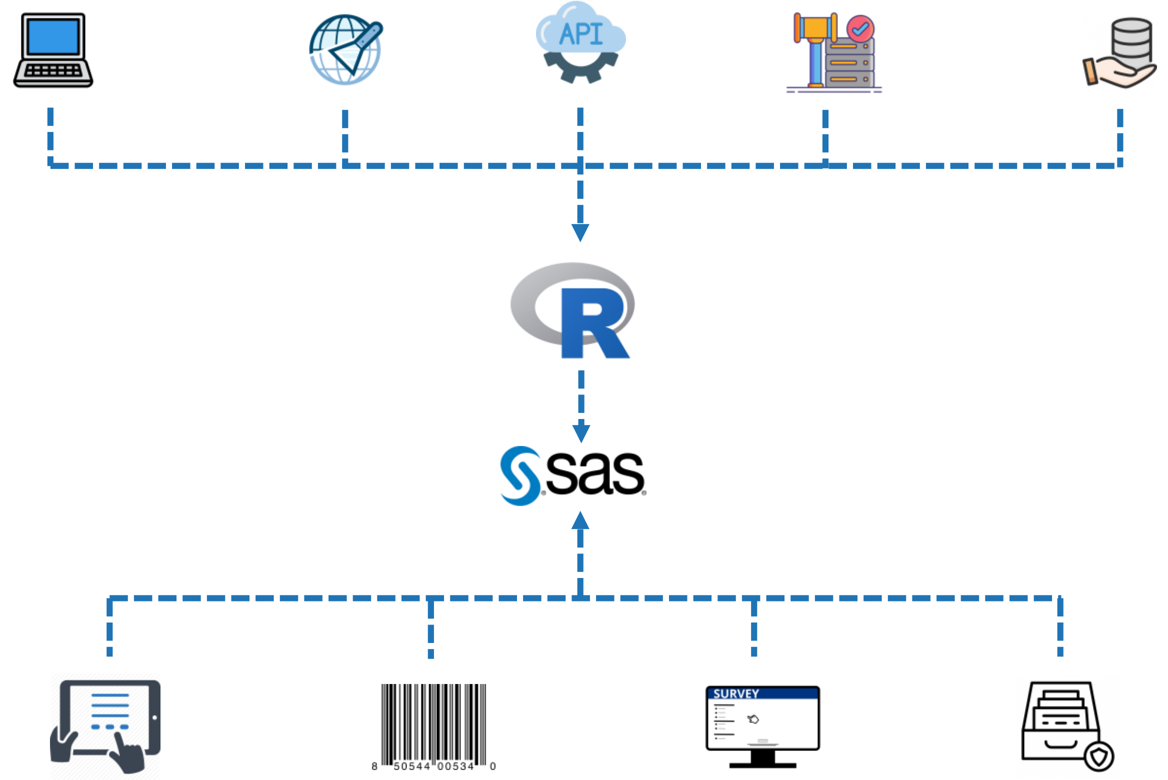
CPI/HICP in 2013



Desk research

Local price collection

CPI/HICP in 2024



Desk research

Web scraping

API

Regulatory data

Privately held data

Local price collection

Scanner data

Web survey

Administrative data

STATBEL

Example of **shrinkflation -** price changes less than package size:

- **New barcodes assigned,** making relaunches hard to track in scanner data
- Using unique barcodes in scannerdata **risks missing these changes**
- **New methods were needed** to detect and adjust for shrinkflation

Cat-astrophe? Whiskas pet food criticised over shrinking servings

Cat food brand caught in shrinkflation storm as owners say new pouches and recipe leaves pets hungry

(The Guardian, 2023)

STATBEL

Identified challenges and opportunities

New data sources ≠ automatically better or easier to use:

- **Challenges in Data access:** despite legal obligations (years of negotiation, no guarantees)
- **Methodological hurdles:** new methods are complex and hard to explain
- **Classification matters:** millions of products require automatic classification (ML/text mining)
- **Evolving production systems:** shift from SAS to R, GIT version control integration, …
- **Growing need for data science skills** and highly educated staff
- **Contingency plans essential:** less survey data means greater reliance on data providers

STATBEL

Identified challenges and opportunities

Legislation

- While not enough on its own, article 5, §4 of the HICP regulation (2016) mandates access to electronic data such as scanner data for NSIs:

  "Upon the request of the national bodies responsible for compiling the harmonised indices, the statistical units shall provide, where available, electronic records of transactions, such as scanner data, and at the level of detail necessary in order to produce harmonised indices and to evaluate compliance with the comparability requirements and the quality of the harmonised indices."

  ⇨ lack of an enforcement mechanism (possibility to use new 223?)

  ⇨ Data providers' collaboration is essential to understand: what data they have; how data connections can be established and how to efficiently set up data flows

  ⇨ Practical challenges: every dataset received within a sector differs in structure and format; variations occur due to differences in underlying data systems

- Highlighting legal enforcement in discussions can increase voluntary data transmission agreements

2. Transport Statistics

Freight type: ………………………

Weight: ……………………………

Date: ………………………………

Loading place: ……………………

Unloading place: …………………

Distance: …………………………

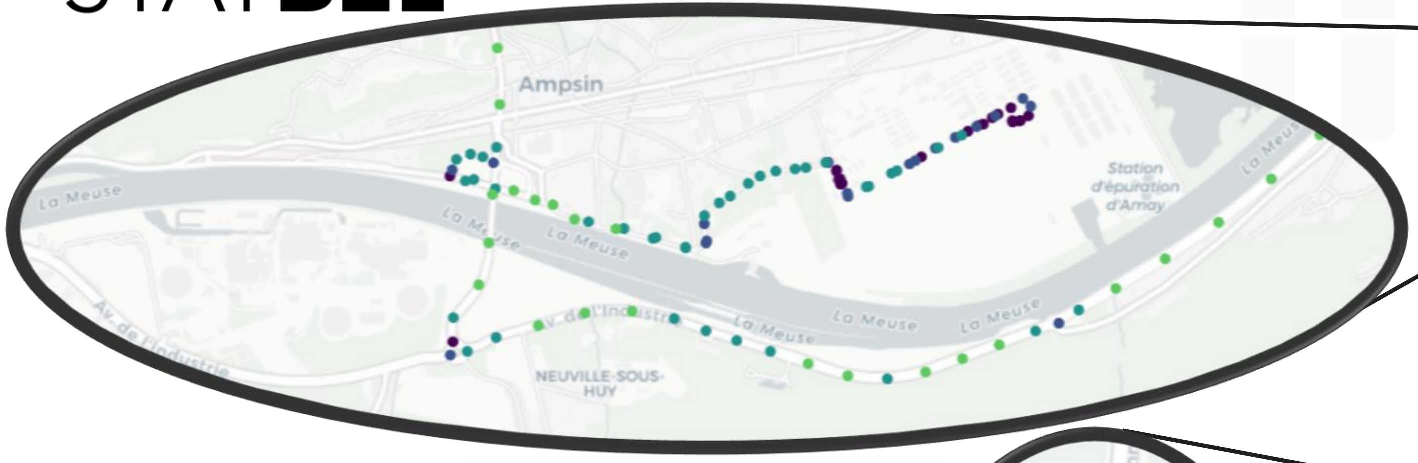
Opportunity:
Remove or simplify questions that represent a significant burden

NEW

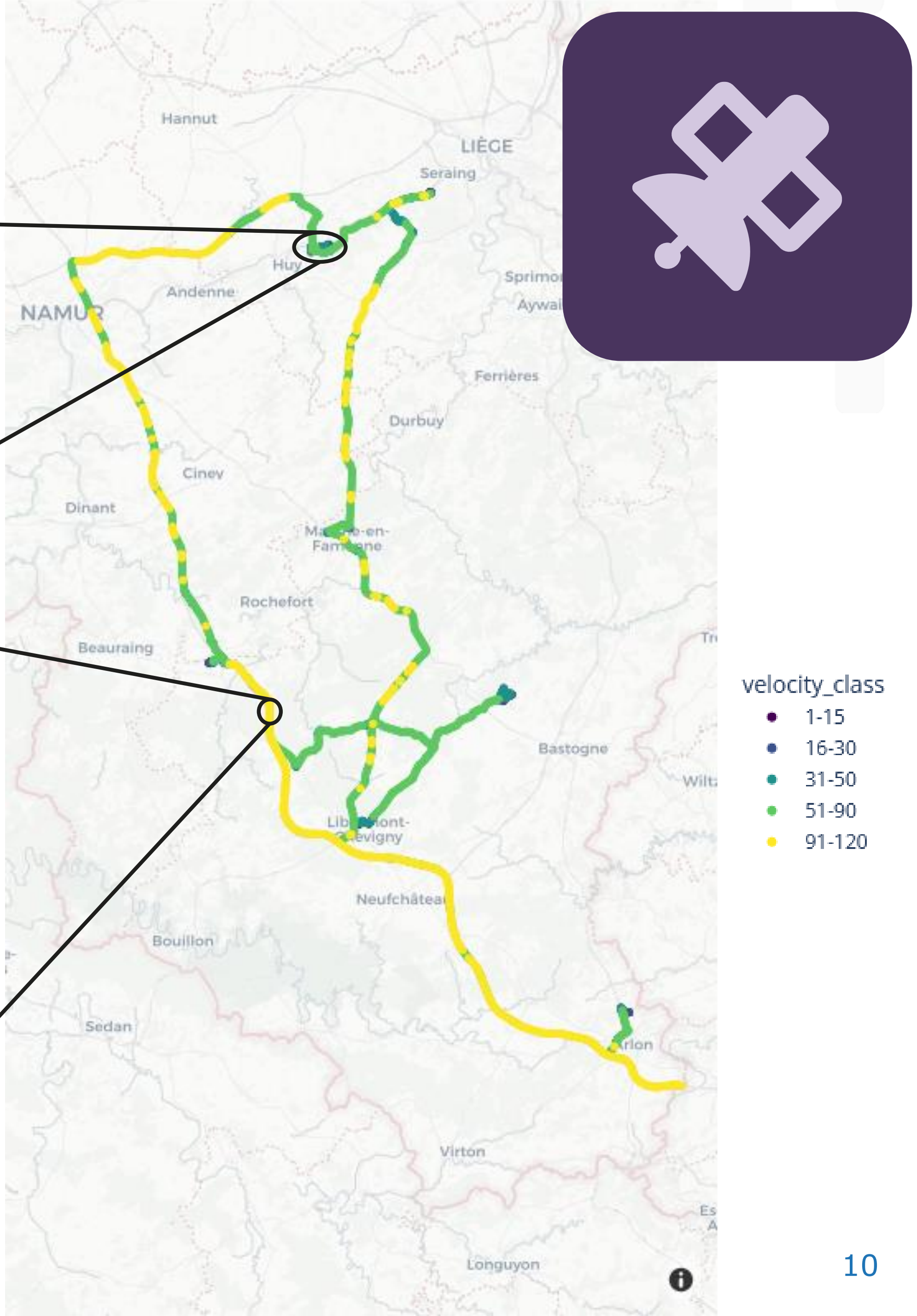
Mandatory use of GPS tracking for road taxation since 2016

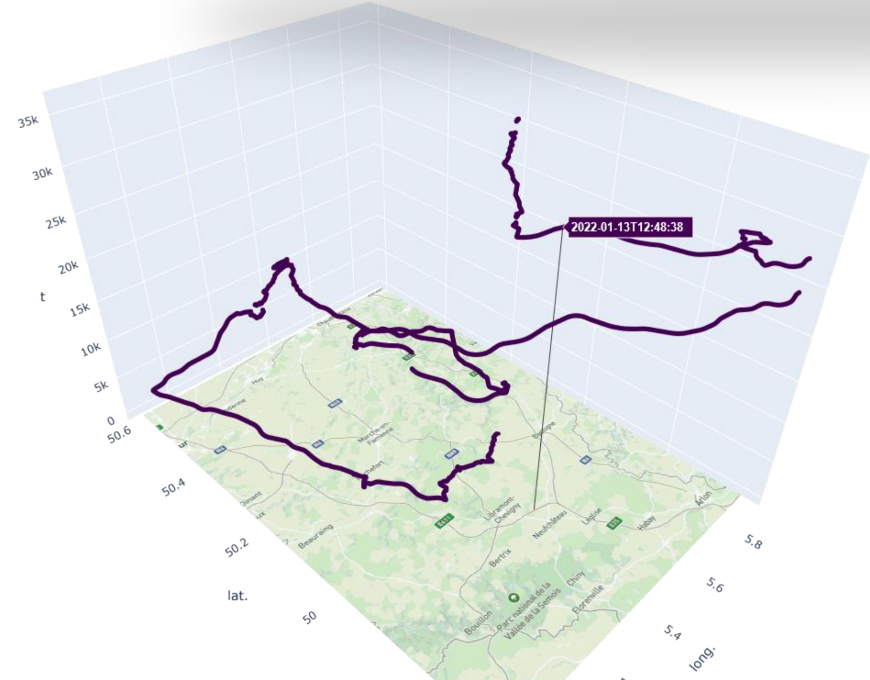
High resolution OBU data

STATBEL


Ampsin
La Meuse
La Meuse
La Meuse
La Meuse
La Meuse
La Meuse
Station d'épuration d'Amay
Av. de l'Industrie
Av. de l'Industrie
NEUVILLE-SOUS-HUY

Autoroute des Ardennes
Autoroute des Ardennes


Hannut
LIÈGE
Seraing
Huy
Andenne
NAMUR
Sprimont
Aywaille
Ferrières
Durbuy
Ciney
Dinant
Marche-en-Famenne
Rochefort
Beauraing
Bastogne
Wiltz
Libramont-Chevigny
Neufchâteau
Bouillon
Sedan
Arlon
Virton
Longuyon
velocity_class
● 1-15
● 16-30
● 31-50
● 51-90
● 91-120

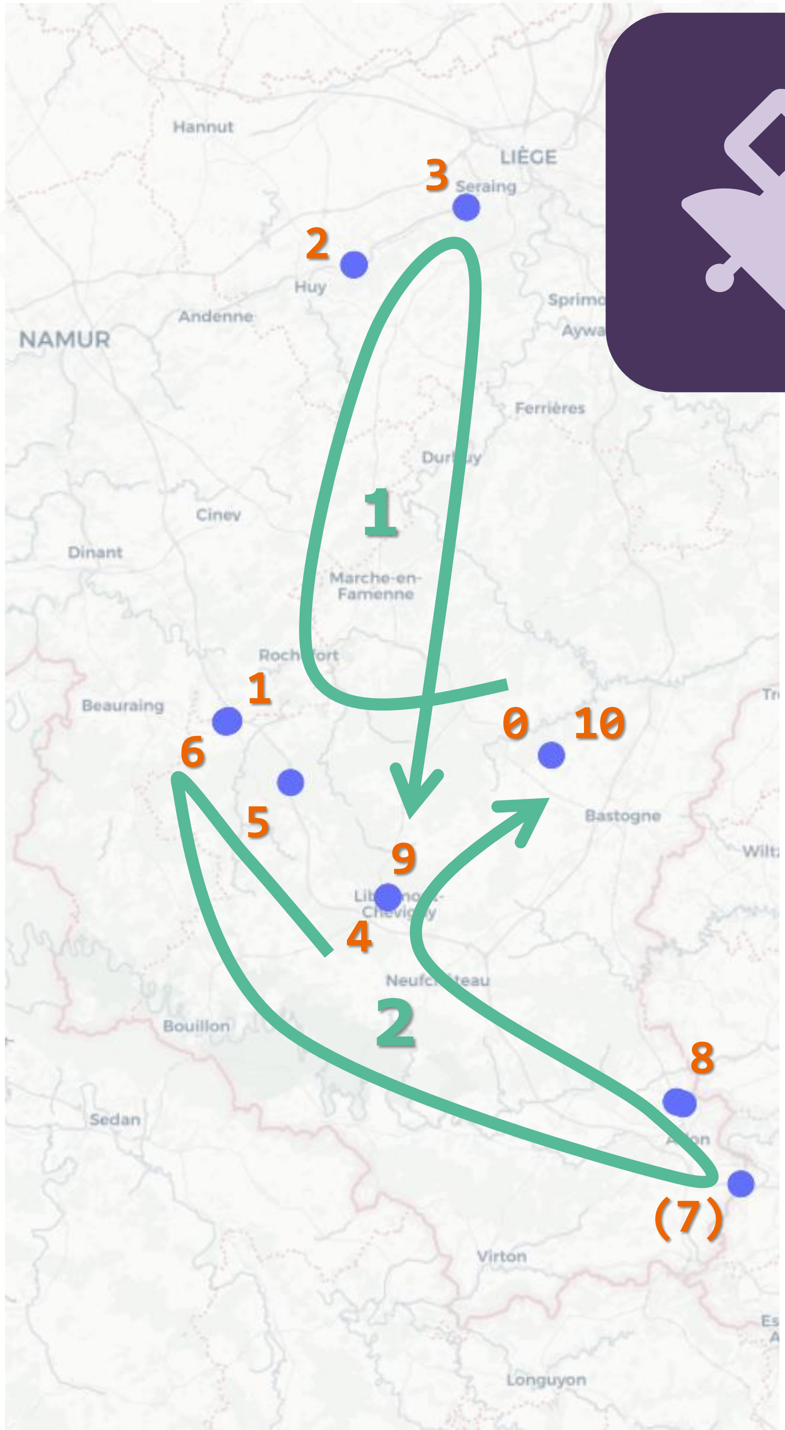


New analyses

STATBEL

Criterium: halt for at least 1 minute

ID journey	Origin	Destination	Length	Duration
1	Givry	Wellin	53 km	53 min
2	Wellin	Huy	93 km	1 h 18 min
3	Huy	Flémalle	20 km	20 min
4	Flémalle	Libramont	96 km	1 h 22 min
5	Libramont	Area of Tellin	24 km	17 min
6	Area of Tellin	Wellin	11 km	9 min
7	Wellin	(Luxembourg)	89 km	1 h 13 min
?	*(Luxembourg)*	*(Luxembourg)*	*?*	*?*
8	(Luxembourg)	Arlon	18 km	17 min
9	Arlon	Libramont	57 km	45 min
10	Libramont	Givry	27 km	30 min
TOTAL	**Givry**	**Givry**	**488 km**	**7 h 04 min**


Hannut
LIÈGE
Seraing
3
2
Huy
Andenne
NAMUR
Sprimo
Aywa
Ferrières
Durbuy
Ciney
Dinant
1
Marche-en-Famenne
Rochefort
Beauraing
1
6
0
10
5
Bastogne
9
Libramont-Chevigny
4
Neufchâteau
2
Bouillon
8
Sedan
Arlon
(7)
Virton
Longuyon

To conclude

Opportunities

- Reduced response burden
- More accurate official statistics (high spatial and temporal resolutions)
- New relevant statistics (traffic, pollutant exposure, gas emissions, etc.)

Challenges

- Service providers own OBU data
- Software development skills
- Multi-modal data merging

STATBEL

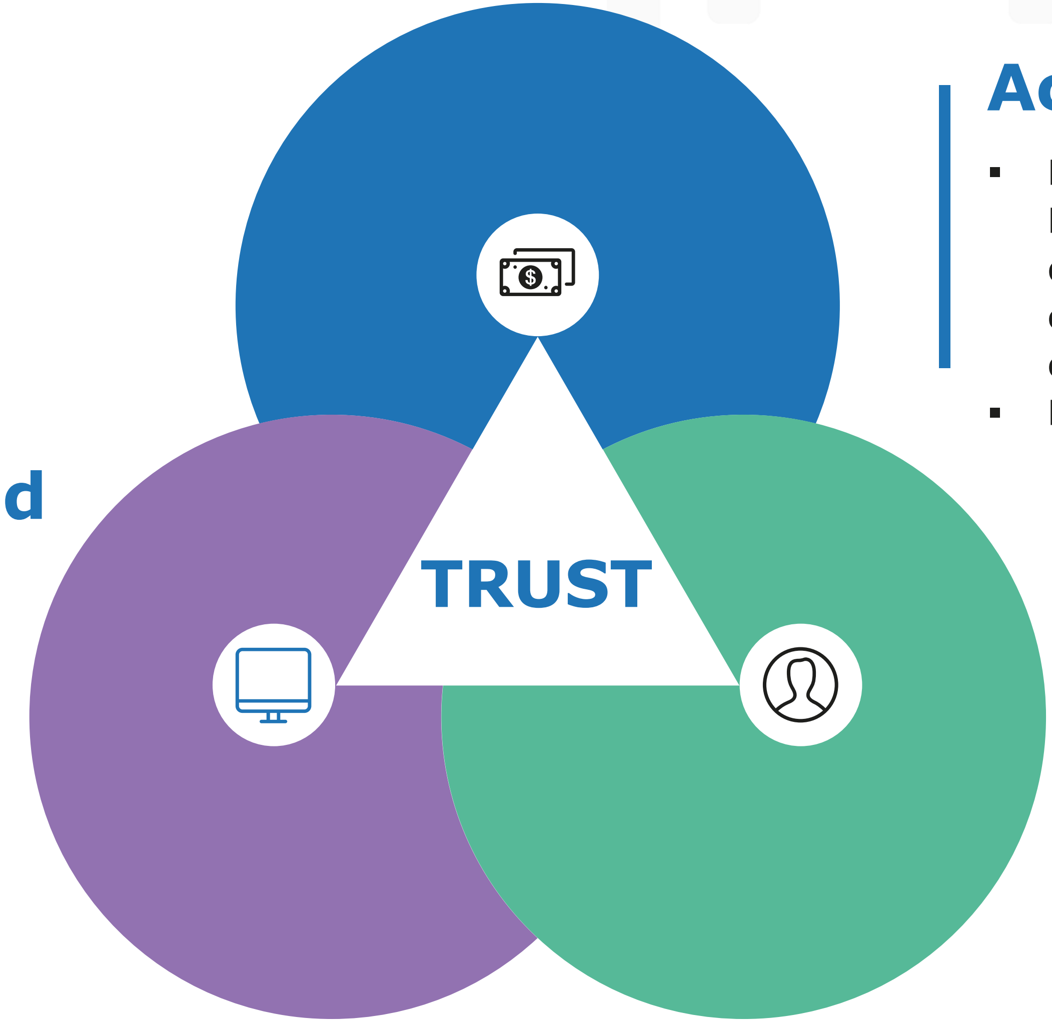
02

Challenges:

why businesses accept’ versus “why businesses refuse”

STATBEL

Overall Challenges



Access

- Reluctance from Private Entities: Businesses may be hesitant to share data due to concerns over confidentiality, competitive advantage, or lack of clear incentives.
- PHD and existing regulations

Technology and Methodology

- Our limited Technology Stack.
- Data Integration and Standardization

Strategy

- Key : Building Trust with Data Providers!
- Data stewardship: bridging between external providers, our data portfolio, and the statistical system.
- Leveraging soft skills!
- High Level engagement

Challenges

STATBEL

Building Trust with Data Providers:

- Move beyond transactional data requests to long-term, trust-based partnerships.
- Understand business concerns (privacy, competition, ethics) and co-design solutions.
- Communicate clearly how data contributes to public value.

Data stewardship: Statbel acts as a central (data) connector within the statistical system – this gives us a unique vantage point to align private data flows with public statistical needs.

Soft skills:

- Key soft skills: listening, negotiating, co-creating, and being credible.
- Even in an era of AI, trust remains something profoundly human — built through empathy, credibility, and dialogue.

Challenges

STATBEL

High-Level engagement and active support

- Contacts with business development are necessary but not sufficient
- Collaboration is effective during research/testing/pilot stages but not always beyond that
- Transitioning to regular production requires decisions at the highest level for both NSI and private data holders

The challenge of PHD and existing statistical regulations:

- Existing regulations are designed for traditional surveys
- NSIs can ensure timeliness and frequency for these surveys (e.g. monthly)
- PHD can improve quality but may reduce timeliness or frequency
- Example: Telecom and insurance PHD follow quarterly accounting cycles
- Trade-off: Maintain costly, lower-quality surveys or accept delays for better-quality and cheaper PHD

STATBEL

03

Implementation of Art. 17, b, c, and e of Regulation (EC) No 223/2009

STATBEL

Early impressions

- **'Light' approach** to PHD access:
  - Broad guidelines based on new 223/2009 provisions in general legal framework (Belgian Statistics Act)
  - Details handled in contracts with data holders rather than rigid legal provisions
  - Prioritising 'negotiation' over 'obligation': Instead of mandatory relationships, adopt a more negotiation-based approach.
  - Legal safeguard already in place – last resort approach
  - Will allow sufficient flexibility:
    - Contractual flexibility: Agreements can be more creative within contractual margins.
      - Best possible implementation would be to create Win-win situation: Balancing interests of all parties involved
  - Challenge: need for broader access to PHD (223/2009 only allows for access within the scope of european statistics)