

WG on Big Data and Open Data

Item 4– *Data Validity*

16-17 March 2017- PARIS

The Statistical Validity of Big Data: Overview

- I. Defining Big Data
- II. Adapting OECD Data Quality Standards
- III. The Main Quality Challenges of Big Data
- IV. Comments from Statistical Meeting
- V. The Importance of Using Big Data For Transport Research



I. Big Data, Big Potential ... Big Challenges

- **The Six V's:**
 - Volume
 - Variety
 - Velocity
 - Value
 - Veracity
 - Variability

II. Adapting OECD Data Quality Standards to Big Data

Standard	Big Data Benefits	Big Data Challenges
Relevance	Allows the study of new subjects	Can lack information needed to respond to objectives
Accuracy	More detail	Measures of statistical quality of results have to be modified
Credibility	---	Can lack methods of scientific rigor and objectivity, although work is being done to improve this
Timeliness	More data at faster rates	Data becomes out-of-date faster, requiring new methods of analysis
Accessibility	---	Metadata and collection methods can sometimes be confidential and costly to obtain, assuming proper documentation has been recorded.
Interpretability	---	Can lack variables of interest or those needed for extrapolation of findings
Coherence	---	Lacks consistent collection and documentation procedures

III. The Main Quality Challenges for Big Data

- 1) Inauthentic data collection
- 2) Information incompleteness and noise of big data
- 3) Unrepresentativeness
- 4) Consistency and reliability
- 5) Interpreting and visualizing big data: the dawn of new analytics

IV. Comments from Statistical Meeting

- Sustainability
- Accessibility
- Timeliness
- Data processing
 - Developing data science methods
- New Sources of Big Data



V. The Importance of Using Big Data For Transport Research

- Detailed concerning spatial and temporal units
 - Particularly important for transport policy
- New forms + New volumes = New understandings of the world
- Timeliness: More data at quicker rates than ever before!
 - Potential real-time impact on public
- And much more to come ...

Thank you

Ashley ACKER
T +33 (0)1 45 24 47 71
E ashley.acker@itf-oecd.org

Postal address
2 rue Andre Pascal
75775 Paris Cedex 16