PROJECT MEETING ON MEASURING HUMAN MOBILITY

27th – 29th March 2019, Tbilisi (Georgia)

Session 3: Other Human Mobility Statistics - Tourism and event statistics

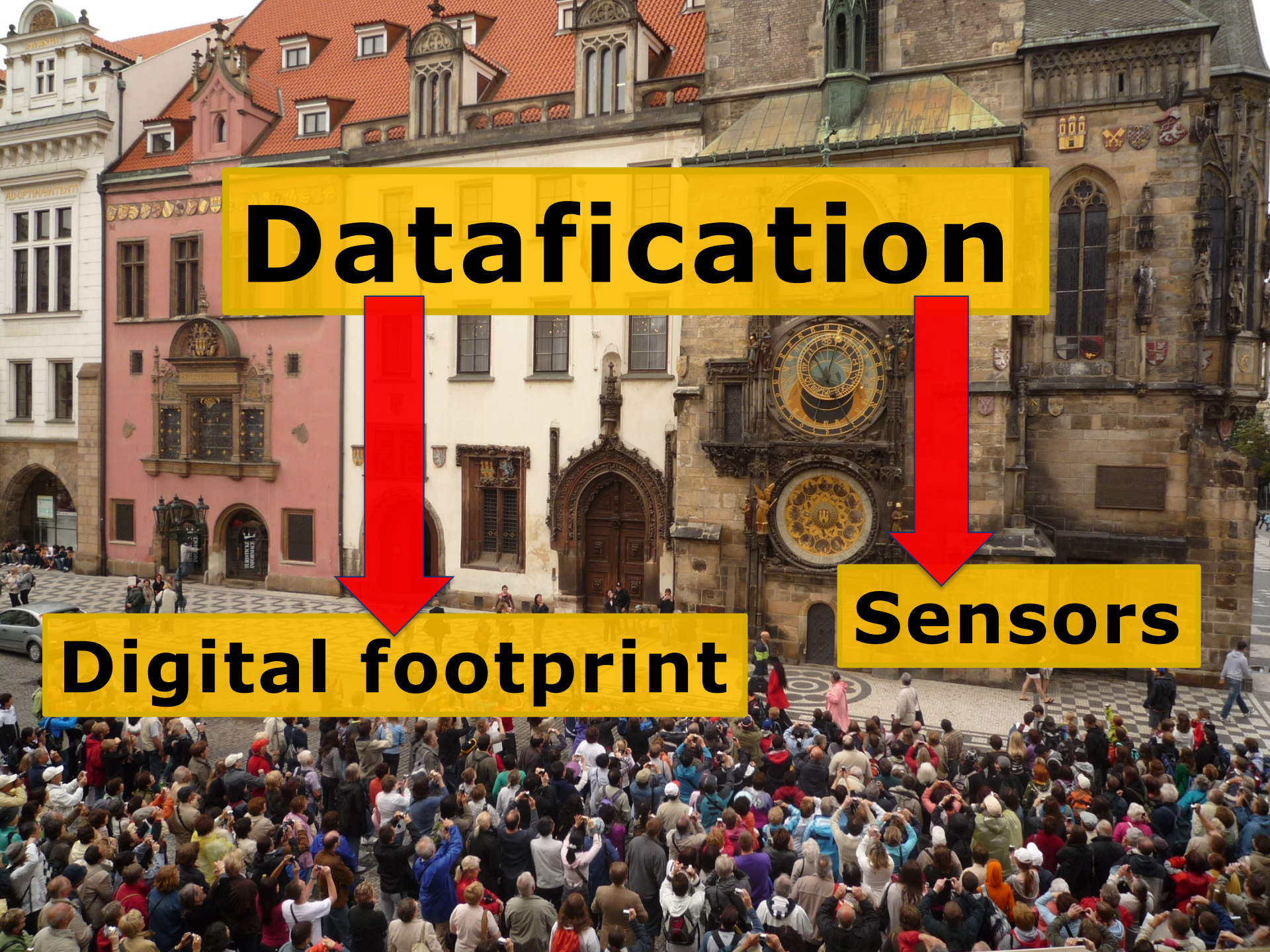
Presentation by Eurostat

Christophe Demunter
European Commission
DG EUROSTAT – Unit G-2 'European Businesses'

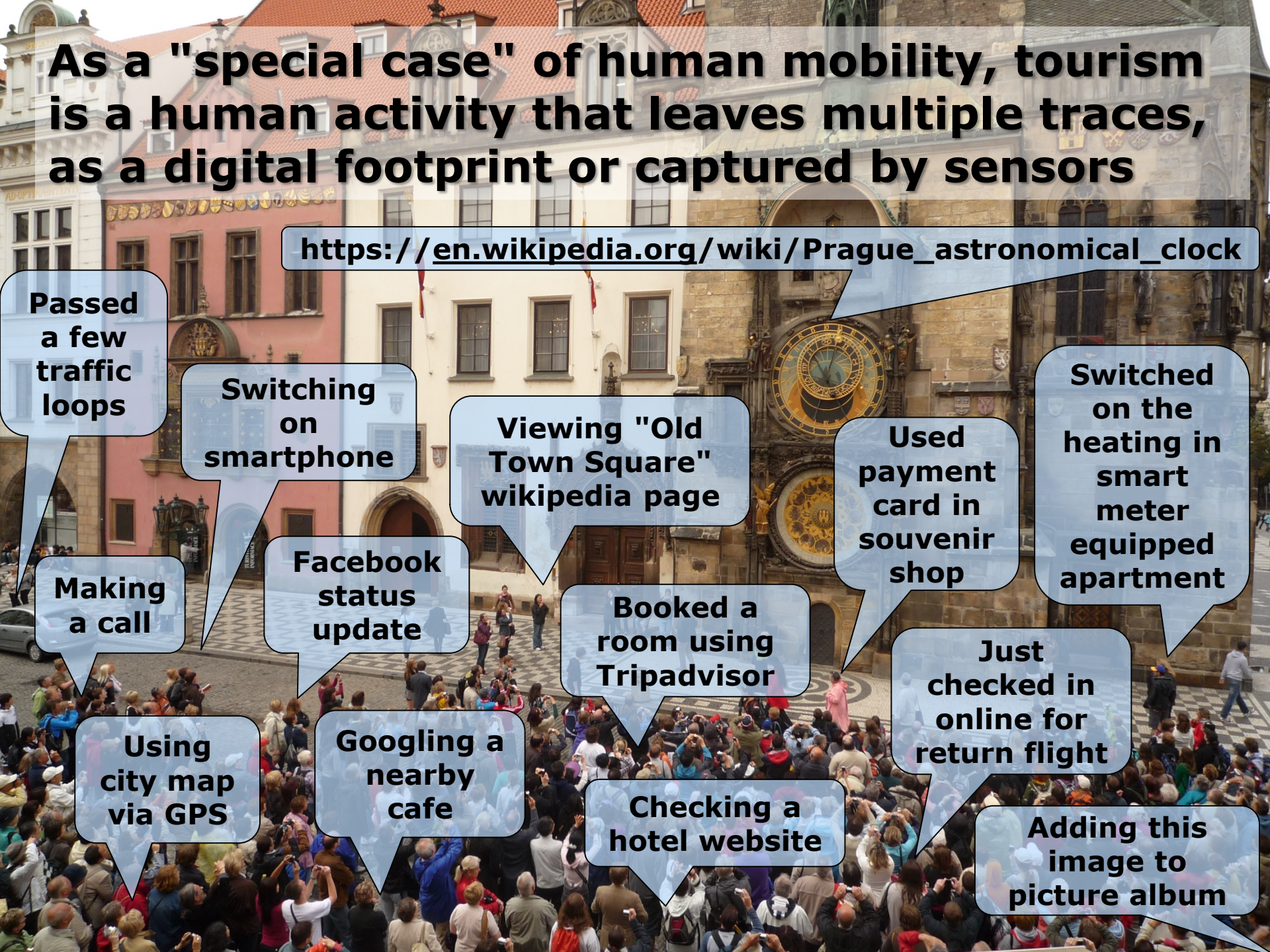
eurostat

Outline of the presentation

- ⇨ **Tourists never leave without a souvenir**
- ⇨ **Better addressing the user needs…**
- ⇨ **The example of MNO data**
- ⇨ **Tourism-related big data projects**
- ⇨ **Take-aways**


Datafication
Digital footprint
Sensors

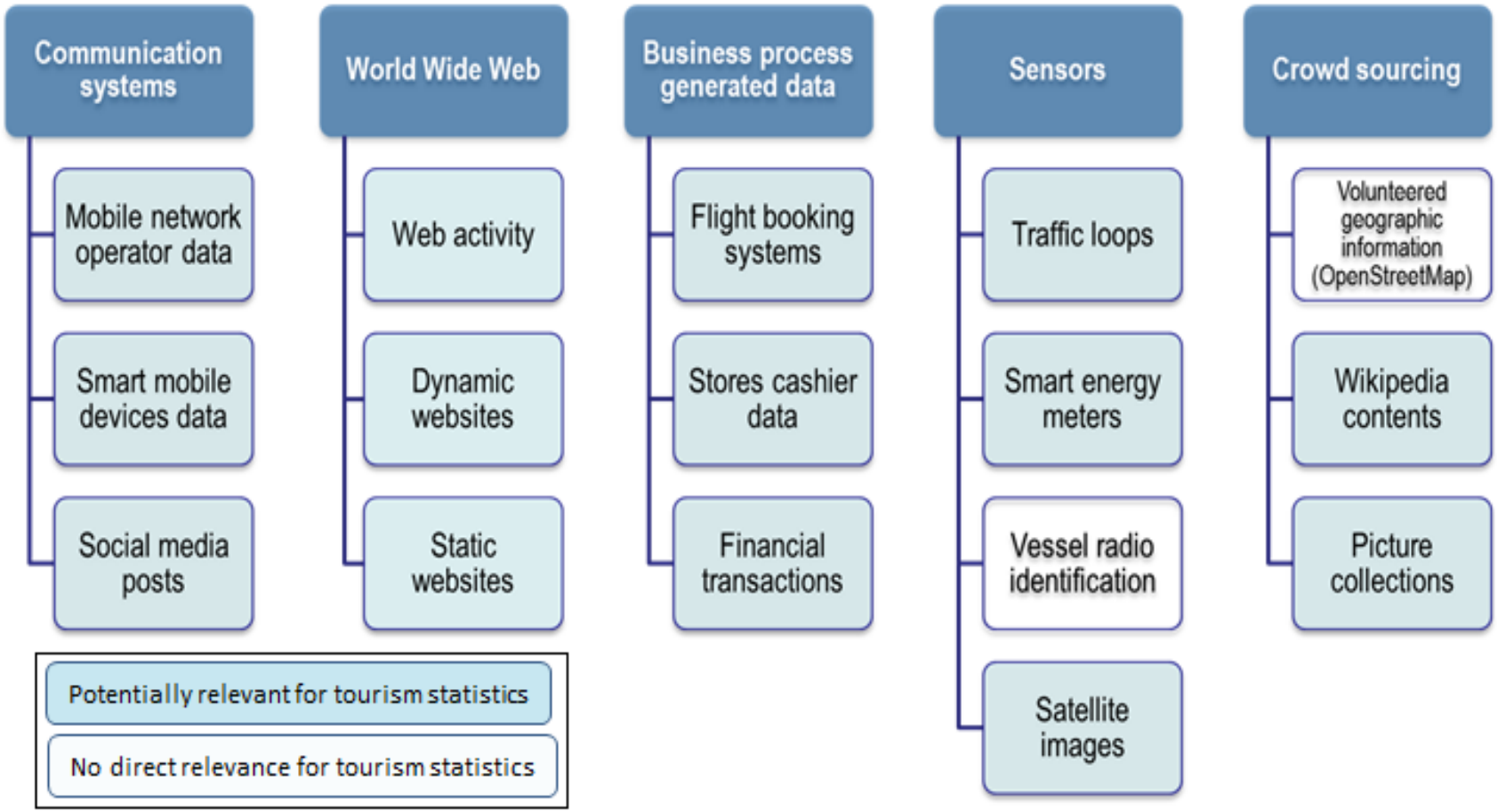
As a "special case" of human mobility, tourism is a human activity that leaves multiple traces, as a digital footprint or captured by sensors



Tourism statistics: early adopters of big data?

⇨ **Eurostat paper** (2017)

Taxonomy of big data sources (Eurostat, 2017)



Meeting the user needs …

Case 1: What do users need to assess sustainability?

- ⇨ **Destination level data, even parts of a destination**
- ⇨ **Daily, weekly data**
- ⇨ **Data on flows, transport, energy use, waste, etc.**

What we can offer to measure sustainability?

- ⇨ **Annual regional occupancy/arrivals data** (NUTS2)
- ⇨ **Monthly national occupancy/arrivals data**
- ⇨ **Data on trips by means of transport, duration, etc.**

Meeting the user needs …

MEP Tonino PICULA

- ⇨ **"current official statistics are not enough for planning and managing public resources in communities with a high number of tourists, daily visitors and summer residents"**
- ⇨ **"developed new calculation method, not counting trips but person-days"**
- ⇨ **"better estimate for pressure on sewage, health care, waste, water"**

Meeting the user needs …

Case 2: How to measure same-day visits?

- ⇨ Very difficult using traditional sources (recall bias, grey area for delineating, …)
- ⇨ Unique potential of new data sources & methods
- ⇨ Same-day visits for tourism purpose as a special case of human mobility

Potential of innovative measurement approaches?

- ⇨ Geographical and temporal granularity; timeliness
- ⇨ Algorithm based; harmonisation & comparability

Meeting the user needs …

Newspaper article on tourism at the Belgian coast during the long Easter weekend, released one day after the weekend:

"150 000 same-day visitors on Easter Sunday, 400 000 during the entire three-day weekend"

Monitoring by the regional tourism board, in cooperation with a mobile network operator & the road infrastructure administration.

In comparison:

Current official statistics	Innovative approach
Entire quarter	Easter weekend
End June of following year	Next day
Entire country	Coastal strip of NUTS2 region

Kust tevreden met verlengd paasweekend

Vandaag om 11:17 door llo | Bron: BELGA

0 | 12 | 1 | Mail | Print

Tweeten | Aanbevelen

Foto: Kurt Desplenter

De toeristische sector aan de kust is tevreden over het verlengde openingsweekend van de paasvakantie. Dat blijkt dinsdag uit een monitoring van provinciebedrijf Westtoer in samenwerking met Proximus en het Agentschap Wegen en Verkeer.

Met 150.000 dagtoeristen op Pasen, en 400.000 dagtoeristen verspreid over het hele paasweekend, blikt de kust tevreden terug op het startschot van de paasvakantie.

'De vele evenementen in de kustgemeenten hebben tijdens de voorbije dagen gezorgd voor aangename vakantiesfeer', aldus Franky De Block, gedeputeerde en voorzitter Westtoer.

Ook voor het verblijfstoerisme schoot de paasvakantie goed uit de startblokken. Hotels noteerden een bezettingsgraad van 90 procent of meer. Bij de verhuur van vakantiewoningen worden degelijke cijfers genoteerd met een gemiddelde bezetting van 75 procent en iets meer reservaties in de eerste week, gaf Westtoer vorige week al mee, maar met de aankondiging van het uitstekende lenteweer in de loop van de week verwacht de sector nog tal van last-minuteboekingen.

Meeting the user needs …

Evolution or revolution?

1. **Big data as auxiliary info for surveys**

   Validation, calibration, deeper breakdowns

2. **Surveys as one of the sources**

   Increasing weight of MNO data and other new sources

   Flash estimates based on MNO data?

3. **New sources gradually replace surveys**

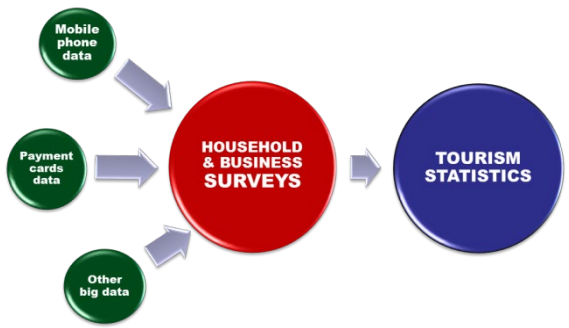
   Partially!

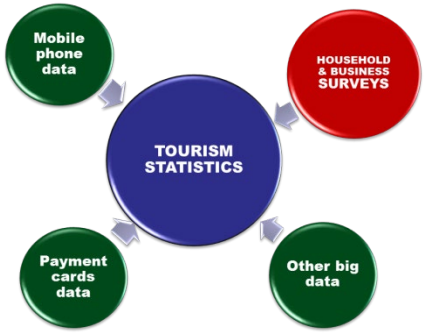
4. **Rethink the system of tourism statistics?**

Meeting the user needs …

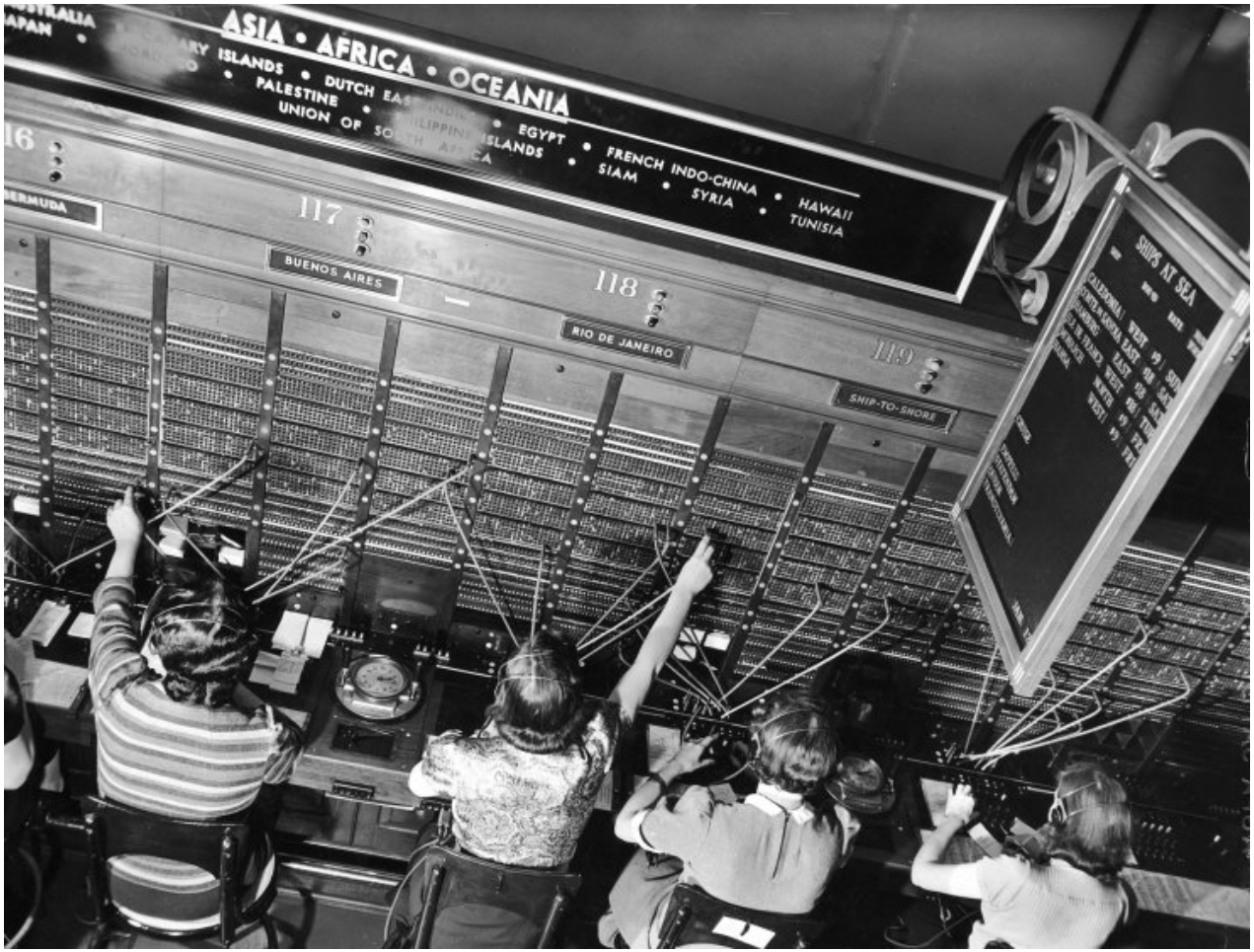
Need to re-think our system of tourism statistics?

- ⇨ **Now: we answer user needs bound by the constraints of the methods of the previous century**
- ⇨ **Future: integrate many sources (MNO, smart surveys,…) and better answer the user needs ; multi-source & multi-purpose**





Use of MNO data … a slow process



Use of MNO data … a slow process

Barriers include:

- **Data held externally**
- **Institutional barriers inside the organisation**
- **Mentality barriers in the organisation** (open to change?)
- **Public opinion** ("efficient modern organisation" or "big brother"?)
- **Skills**

☞ **From 'full control' to being a 'user'**

- MNO data held by private companies
- Need for partnerships & business model: development/pilot phase vs. long-term collaboration
- Quality control?
- Data sources were not conceived for making statistics

Example: MNO data for Belgium *(project 2015-2017)*

Partnership between mobile network operator and statistical offices

- Explore partnerships & business models; cooperate on concrete pilot projects

Project suddenly ended, but some preliminary results

- Focus on **outbound trips** made by residents of Belgium
- **Comparison** of estimates based on official statistics and estimates based on mobile phone data (April – September 2015/2016)
- Somehow promising (**the data makes sense**) but lots of open issues

[More information: paper & presentation for the NTTS conference 2017]

Example: MNO data for Belgium (project 2015-2017)

Data sources for the project

- **Mobile phone data**
  - ⇨ Data from one operator in Belgium
  - ⇨ **Signaling data (not Call detail records)**
  - ⇨ Better **temporal** (and geographical) **granularity**

- **Official tourism statistics**
  - ⇨ **Survey** based data on trips made by residents of Belgium (tourism demand micro data for Regulation (EU) 692/2011)
  - ⇨ Quarterly interviews, annual **sample ± 10000 trips** (domestic + outbound trips with overnight stays)

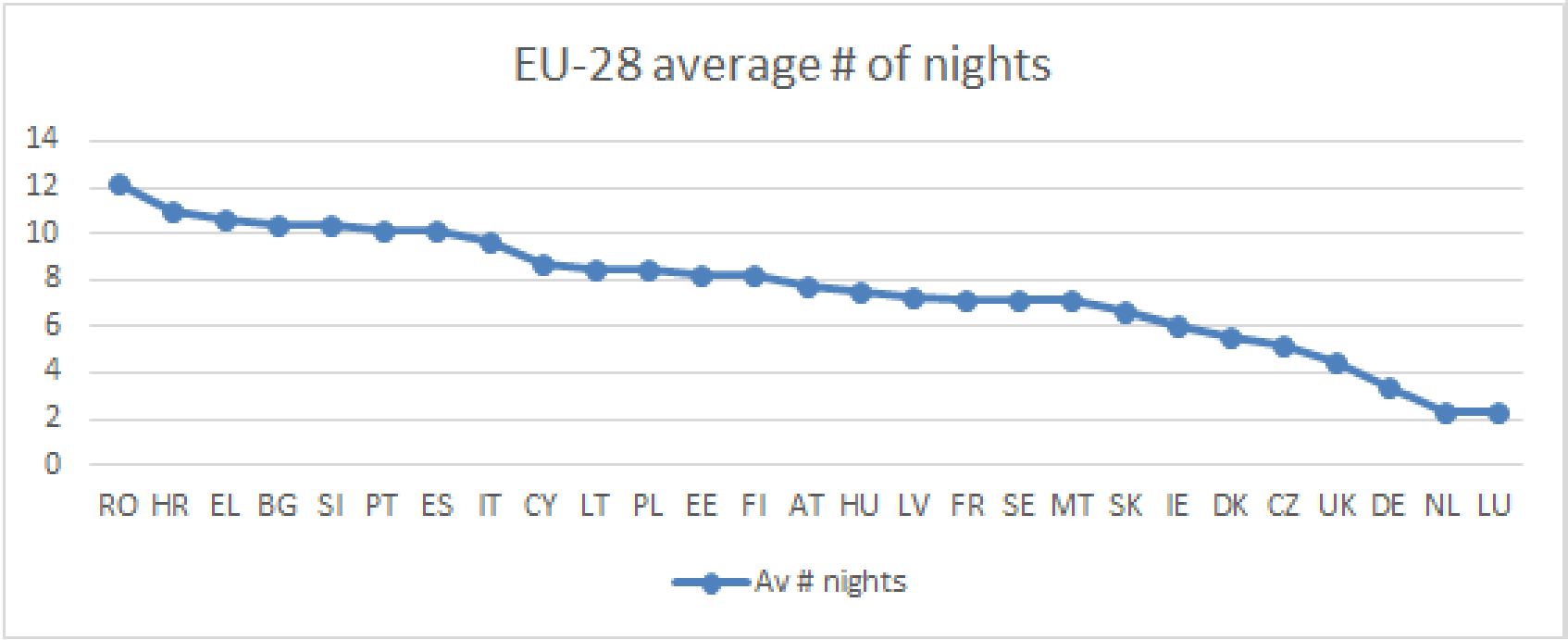
Example: MNO data for Belgium (project 2015-2017)

Scope & definitions

- **Focus on outbound trips**
  - ⇨ Mobile phone data: trips made **April – Sept 2016**
  - ⇨ Official tourism statistics: trips made **April – Sept 2015**
- **Definition of an outbound trip**
  - ⇨ From leaving the home network to returning
  - ⇨ Number of nights: number of hours divided by 24
  - ⇨ Overnight stay: minimum 10 hours and return after 4am
- **Usual environment**
  - ⇨ Duration (min. 10hrs + incl. 4am), border crossing (outbound)
  - ⇨ Filtering of frequent trips to the same destination during a given reference period (250 days) → threshold = **5** (arbitrary)

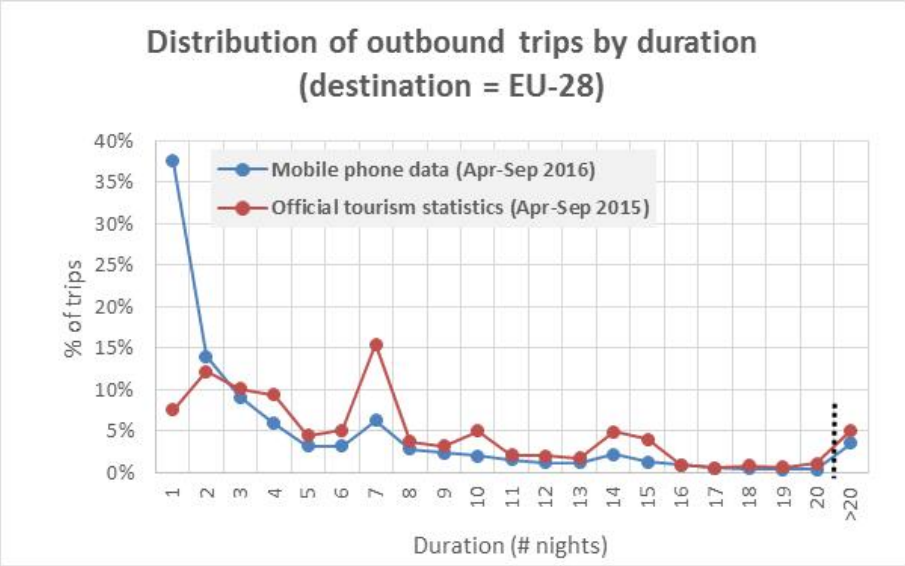
a. Ranking of destinations

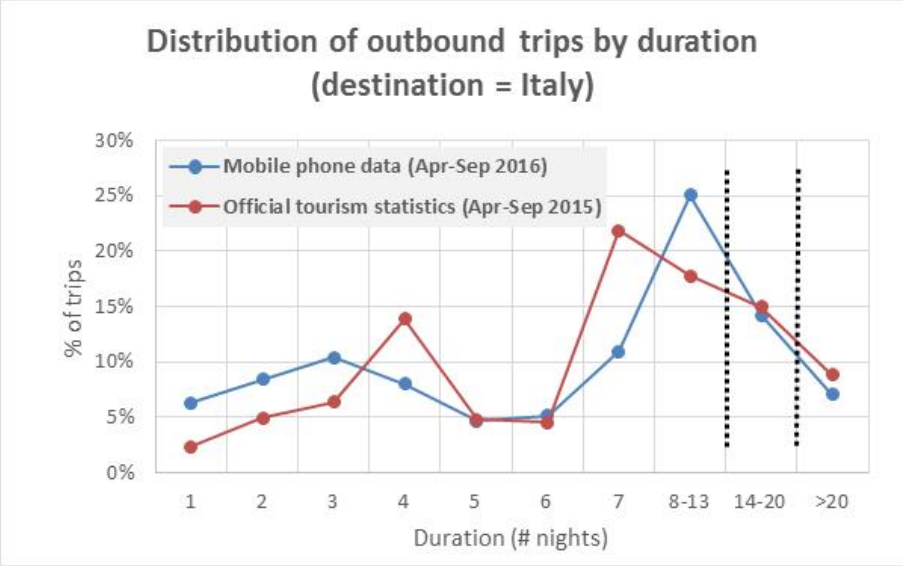
Ranking of EU-28 countries as destination for Belgian outbound trips (Mobile network operator data)



b. Outbound trips by duration: comparison

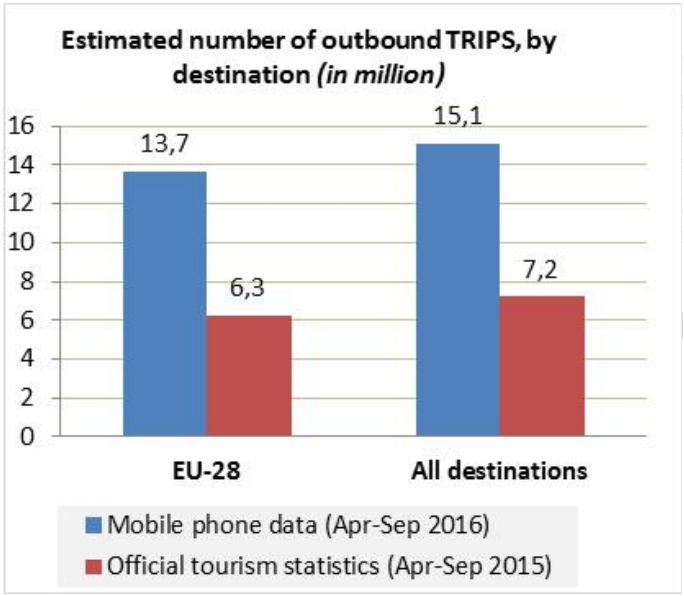
Comparison of the distribution of outbound trips to EU-28 and to Italy, by duration of the trips

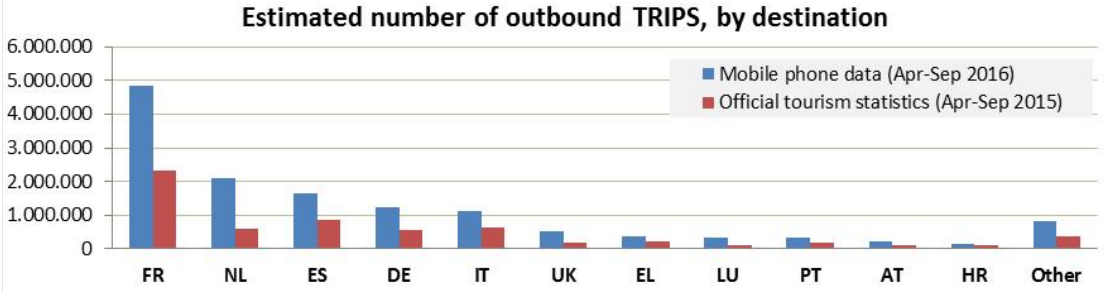


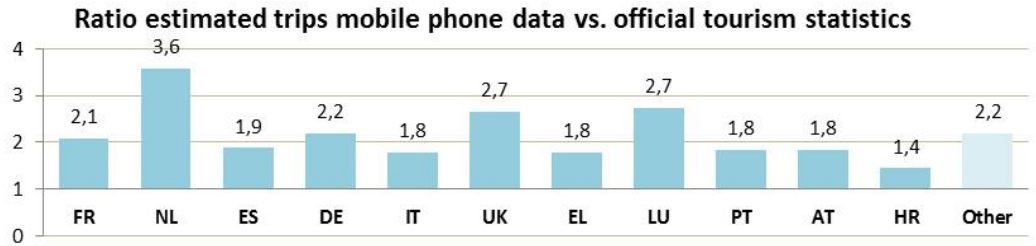


c. Volume of trips and nights: comparison

Comparison of estimated number of outbound trips, by destination







- **Observations:** Big differences between the sources, but of a systematic nature
- **Understanding (and solving…) the deviations:** differences in scope (e.g. age), parameter setting, selectivity bias, recall bias and non-response in surveys
- To be continued… (??)

Known weaknesses

Mobile phone data	Official tourism statistics
Selectivity bias	**Non-response, non-contact,…**
• **Extrapolation** (inverse of market share?)**; roaming between different operators**	**Recall bias, memory effect**
• **Socio-demographic composition of subscribers**	**Respondent burden**
• **Intensity of use; new SIM card in country visited; etc.**	**Timeliness**
Entirely algorithm based (choice of parameters?)	**Entirely respondent based ('subjective opinion')**
Continuity; independence	
Privacy; public opinion	

Some lessons learnt

- **Positive & fruitful experience with the partnership**
  - ⇨ Joining forces (statisticians, data holders, data scientists)
  - ⇨ Search for a win-win
  - ⇨ However… continuity is fragile (pilot was stopped)
- **Promising results, but lots of homework**
  - ⇨ Mobile phone data can capture tourism concepts/definitions
  - ⇨ Currently: satisfactory for trends, not for estimating volumes
  - ⇨ How to make the series/sources converge to the unknown true values?
  - ⇨ Extension to domestic tourism, to same-day visits
- **Further research to be encouraged** (other countries?)

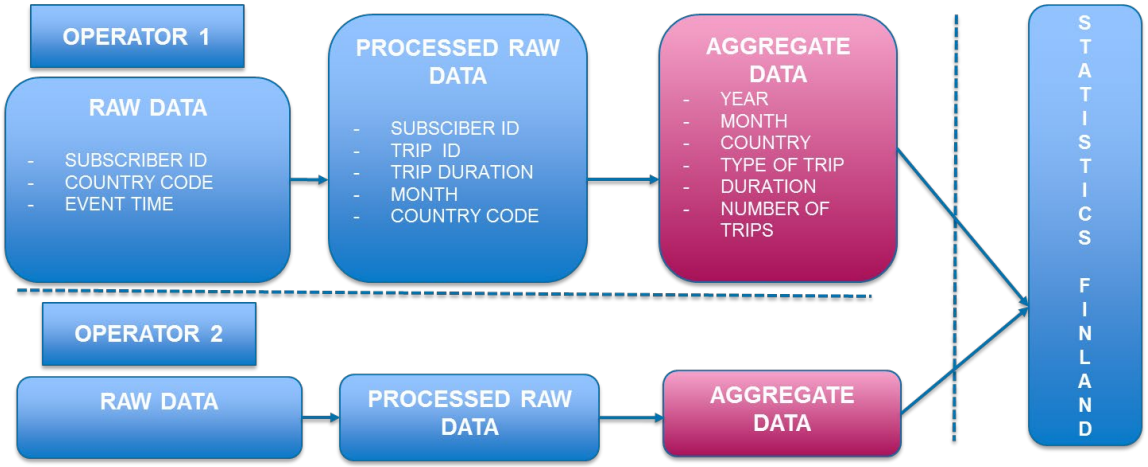
Auxiliary MNO data; roaming ⇨ Statistics Finland

[taken from Statistics Finland paper presented at the 15th Global Forum on Tourism Statistics, Cusco Peru, Nov 2018]

Problem statement:

- Alternative/complementary sources are needed because sample sizes in surveys are small and survey response rates are declining
- Only MNOs are allowed to process their data using automatic means in the current legislation

Data process:



Auxiliary MNO data; roaming ⇨ Statistics Finland

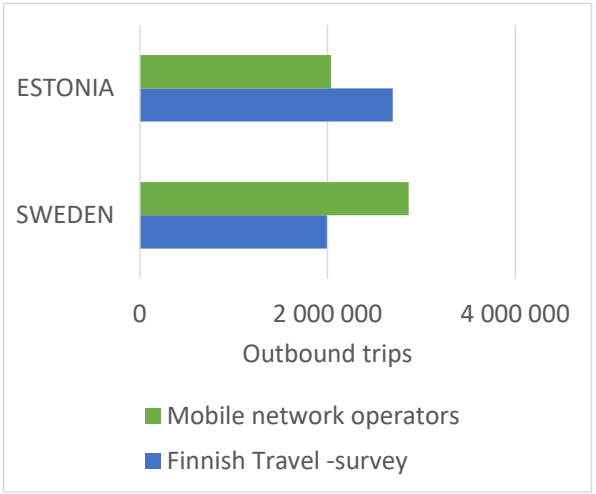
[taken from Statistics Finland paper presented at the 15th Global Forum on Tourism Statistics, Cusco Peru, Nov 2018]

Sources:

	Travel survey	Data from 2 MNO's
Sample size	28,500 persons	< 70% of population
Outbound trips observed	< 3,000 per year	< 7 million per year
Average weight of one outbound trip	< 4,000	< 1.3
Total number of outbound trips (2017)	10.5 million	10.5 million

Mismatches:

- EE & SE : nearly 50% of outbound tourism
- 24% less trips to EE; 44% more trips to SE
  - → # trips in MNO data biased depending on country
  - → ? non-tourism trips, border noise, phone off / flight modus, sample bias (one MNO missing)

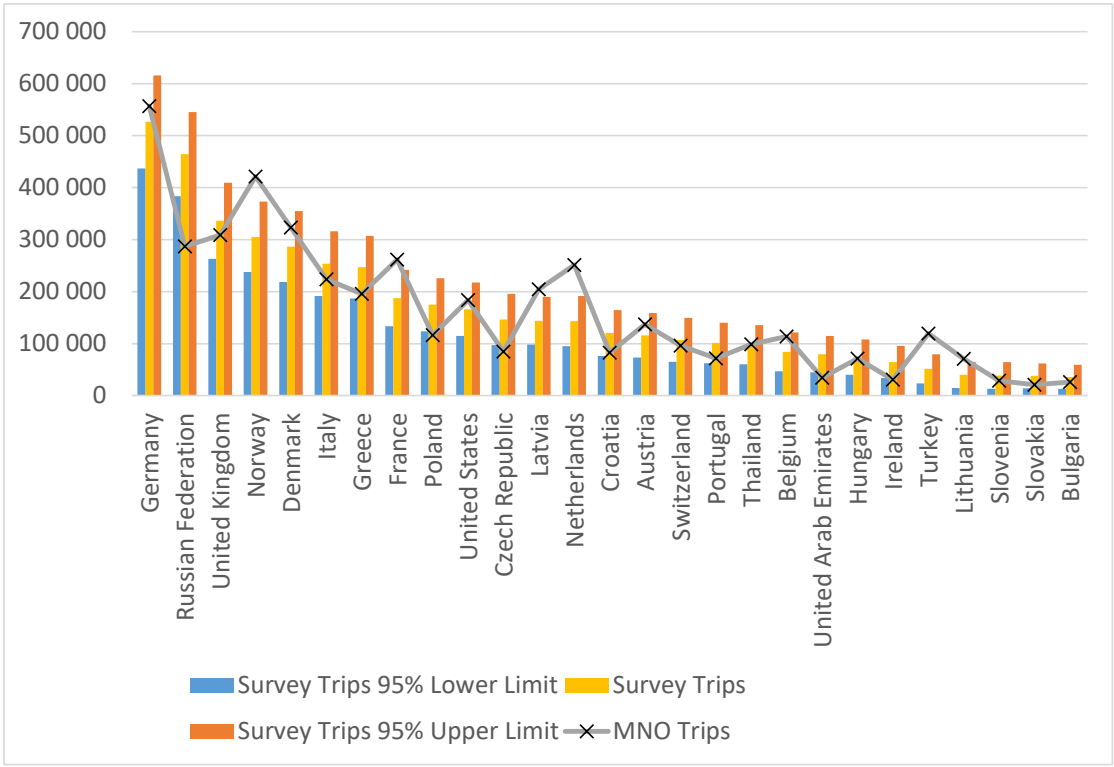


Auxiliary MNO data; roaming ⇨ Statistics Finland

[taken from Statistics Finland paper presented at the 15th Global Forum on Tourism Statistics, Cusco Peru, Nov 2018]

Finnish method to improve survey data using MNO data:

- Use as the frame the total number of annual outbound trips from the travel survey
- Determine the (95%) confidence intervals for each destination country
- Select the more reliable source data (survey or MNO data) for each destination country
- Calculate a coefficient factor for trips to those countries that will be based on MNO data
- Apply a monthly seasonality trend based on MNO data separately for each country



Auxiliary MNO data; roaming ⇨ Statistics Finland

Finnish method to improve survey data using MNO data:

Top 10 countries →

Small countries (top 30 - 40) →

Country	Finnish Travel -survey Outbound Trips (000)	95% CI Lower Limit	95% CI Upper Limit	MNO Top-down Outbound Trips (000)	Selected Source	Final Trips	Monthly Seasonality (from MNO)
TOTAL OUTBOUND TRIPS	**10 484**	**10 135**	**10 834**	**10 484**		**10 484**	
ESTONIA	2 695	2 519	2 871	2 036	SURVEY	2 695	
SWEDEN	1 991	1 833	2 150	2 865	SURVEY	1 991	
SPAIN	822	714	930	638	SURVEY	822	
GERMANY	526	437	616	557	SURVEY	526	
RUSSIAN FEDERATION	464	384	545	287	SURVEY	464	
UNITED KINGDOM	336	263	409	309	SURVEY	336	
NORWAY	305	238	373	421	SURVEY	305	
DENMARK	287	218	355	323	SURVEY	287	
ITALY	254	192	316	224	SURVEY	254	
NETHERLANDS	144	95	192	251	SURVEY	144	
GREECE	247	187	307	196	SURVEY	247	
CROATIA	121	77	164	83	MNO	88	
PORTUGAL	101	63	140	72	MNO	76	
HUNGARY	74	40	108	71	MNO	76	
LITHUANIA	40	15	65	71	SURVEY	40	
CANADA	15	0	30	46	SURVEY	15	
UNITED ARAB EMIRATES	80	45	115	34	SURVEY	80	
IRELAND	65	34	96	31	SURVEY	65	
ICELAND	23	2	43	29	MNO	31	
SLOVENIA	39	13	65	29	MNO	31	
BULGARIA	36	13	59	26	MNO	28	
JAPAN	17	0	34	26	MNO	27	
CHINA	22	4	40	23	MNO	25	

Auxiliary MNO data; roaming ⇨ Statistics Finland

[taken from Statistics Finland paper presented at the 15th Global Forum on Tourism Statistics, Cusco Peru, Nov 2018]

Finnish method to improve survey data using MNO data: summary of the results for 2017

	Estimated based on survey data	Estimated based on MNO data	Total
# trips by country	9.0 million	1.5 million	10.5 million
# destination countries	24	129	153
Average # trips per country	380 000	12 000	69 000

↳ **MNO data** used for **14% of trips** and for **84% of destinations**

Tourism-related big data projects in Eurostat

Horizontal projects:

- **ESSnet involving 20+ countries; Work Packages include:**
  - ⇨ Building a system of tourism statistics based on new sources/methods
  - ⇨ Financial transactions data (use case: collaborative economy)
  - ⇨ Mobile network operator data (use case: human mobility, tourism)
- **Projects on household budget survey & time use survey**
  - ⇨ Focus on new methods (smart apps), not only new sources
  - ⇨ Tourism as a special case of household expenditure and time use

Tourism-specific projects:

- **Small grants with 7 countries (duration 13 months, till end 2019)**
- **Development of a mid-term methodological blueprint on innovative tourism statistics (horizon: 2025)**
- **Experimental statistics**

Tourism-related big data projects in Eurostat

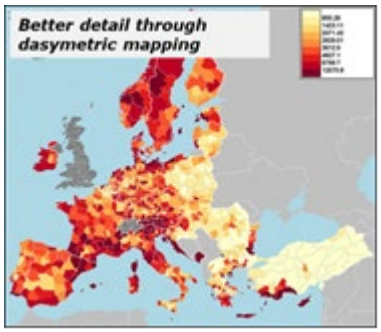
☞ Experimental statistics (new)

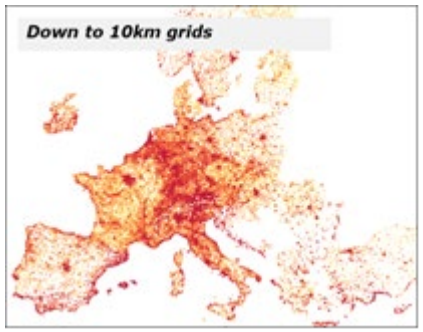
experimental

"Experimental statistics use new data sources & methods in an effort to better respond to our users' needs"

⇨ For tourism statistics:

- **Dissemination of accommodation statistics at NUTS3 (tables) and for 10km grids (maps)**
- **Roll-out: 2019**





Some take-aways

- **RE-THINK methods to serve users**
  "zero-base budgeting"
- **Build TRUST (internal & external)**
  harmonisation (algorithms vs. humans), quality, continuity
- **Invest in SKILLS**
  from data collector to data connector
- **Don't ignore GOVERNANCE**
  partnerships, resources (parallel work), ...

Thank you for your attention !

Eurostat data & publications are available free of charge from the Eurostat website

Tourism statistics leaflet