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

Presentation by Eurostat

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

https://en.wikipedia.org/wiki/Prague_astronomical_clock
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:

<table>
<thead>
<tr>
<th>Current official statistics</th>
<th>Innovative approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire quarter</td>
<td>Easter weekend</td>
</tr>
<tr>
<td>End June of following year</td>
<td>Next day</td>
</tr>
<tr>
<td>Entire country</td>
<td>Coastal strip of NUTS2 region</td>
</tr>
</tbody>
</table>
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

![Distribution of outbound trips by duration (destination = EU-28)](image1)

![Distribution of outbound trips by duration (destination = Italy)](image2)
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

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


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:

<table>
<thead>
<tr>
<th></th>
<th>Travel survey</th>
<th>Data from 2 MNO's</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>28,500 persons</td>
<td>&lt; 70% of population</td>
</tr>
<tr>
<td>Outbound trips observed</td>
<td>&lt; 3,000 per year</td>
<td>&lt; 7 million per year</td>
</tr>
<tr>
<td>Average weight of one outbound trip</td>
<td>&lt; 4,000</td>
<td>&lt; 1.3</td>
</tr>
<tr>
<td>Total number of outbound trips (2017)</td>
<td>10.5 million</td>
<td>10.5 million</td>
</tr>
</tbody>
</table>

Mismatch:

- 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


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)

<table>
<thead>
<tr>
<th>Country</th>
<th>Finnish Travel-survey</th>
<th>MNO</th>
<th>Monthly Seasonality (from MNO)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outbound Trips (000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95% CI Lower Limit</td>
<td>95% CI Upper Limit</td>
<td></td>
</tr>
<tr>
<td>TOTAL OUTBOUND TRIPS</td>
<td>10 484</td>
<td>10 135</td>
<td>10 834</td>
</tr>
<tr>
<td>ESTONIA</td>
<td>2 695</td>
<td>2 519</td>
<td>2 871</td>
</tr>
<tr>
<td>SWEDEN</td>
<td>1 991</td>
<td>1 833</td>
<td>2 150</td>
</tr>
<tr>
<td>SPAIN</td>
<td>822</td>
<td>714</td>
<td>930</td>
</tr>
<tr>
<td>GERMANY</td>
<td>526</td>
<td>437</td>
<td>616</td>
</tr>
<tr>
<td>RUSSIAN FEDERATION</td>
<td>464</td>
<td>384</td>
<td>545</td>
</tr>
<tr>
<td>UNITED KINGDOM</td>
<td>336</td>
<td>263</td>
<td>409</td>
</tr>
<tr>
<td>NORWAY</td>
<td>305</td>
<td>238</td>
<td>373</td>
</tr>
<tr>
<td>DENMARK</td>
<td>287</td>
<td>218</td>
<td>355</td>
</tr>
<tr>
<td>ITALY</td>
<td>254</td>
<td>192</td>
<td>316</td>
</tr>
<tr>
<td>NETHERLANDS</td>
<td>144</td>
<td>95</td>
<td>192</td>
</tr>
<tr>
<td>GREECE</td>
<td>247</td>
<td>187</td>
<td>307</td>
</tr>
<tr>
<td>CROATIA</td>
<td>121</td>
<td>77</td>
<td>164</td>
</tr>
<tr>
<td>PORTUGAL</td>
<td>101</td>
<td>63</td>
<td>140</td>
</tr>
<tr>
<td>HUNGARY</td>
<td>74</td>
<td>40</td>
<td>108</td>
</tr>
<tr>
<td>LITHUANIA</td>
<td>40</td>
<td>15</td>
<td>65</td>
</tr>
<tr>
<td>CANADA</td>
<td>15</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>UNITED ARAB EMIRATES</td>
<td>80</td>
<td>45</td>
<td>115</td>
</tr>
<tr>
<td>IRELAND</td>
<td>65</td>
<td>34</td>
<td>96</td>
</tr>
<tr>
<td>ICELAND</td>
<td>23</td>
<td>2</td>
<td>43</td>
</tr>
<tr>
<td>SLOVENIA</td>
<td>39</td>
<td>13</td>
<td>65</td>
</tr>
<tr>
<td>BULGARIA</td>
<td>36</td>
<td>13</td>
<td>59</td>
</tr>
<tr>
<td>JAPAN</td>
<td>17</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>CHINA</td>
<td>22</td>
<td>4</td>
<td>40</td>
</tr>
</tbody>
</table>
Finnish method to improve survey data using MNO data: summary of the results for 2017

<table>
<thead>
<tr>
<th></th>
<th>Estimated based on survey data</th>
<th>Estimated based on MNO data</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># trips by country</td>
<td>9.0 million</td>
<td>1.5 million</td>
<td>10.5 million</td>
</tr>
<tr>
<td># destination countries</td>
<td>24</td>
<td>129</td>
<td>153</td>
</tr>
<tr>
<td>Average # trips per country</td>
<td>380 000</td>
<td>12 000</td>
<td>69 000</td>
</tr>
</tbody>
</table>

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 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](#).