CESSDA Expert Seminar,
Cooperation between CESSDA archives and NSIs,
Gothenburg, 16-17 September, 2010

Czech Sociological Data Services and
Cooperation with the Czech Statistical Office

Yana Leontiyeva, SDA
Sociological Data Archive (SDA)

- SDA is a department of the Institute of Sociology, Academy of Sciences of the Czech Republic
- services since 1998, CESSDA since 2001
- access to Czech social science data for purposes of academic research and higher education
  - open access, on-line, free of charge
- promotion of data dissemination and secondary analyses
  - *Data a výzkum - SDA Info* (Czech language peer-reviewed journal)
  - methodological research (data quality, harmonization of data...)
  - Web directory, information services
- support for special research projects (ISSP, ESS)
- international co-operation
  - CESSDA (Council of European Social Science Data Archives)
  - REGIO (Central European Data Archives Network)
SDA current projects

**COMPDAT (2009-2011)**

- Data sources, research on data quality, standards, and methods of data harmonisation for the purposes of international comparative social research and integration into the CESSDA network
- Funded by Ministry of Education, Youth and Sports
- The aim is to strengthen the database and methodological platform of cross-national empirical research in the social sciences in the Czech Republic.
- Forthcoming workshop on harmonization of social survey data for cross-national comparison, October 19th, 2010

**Laboratory of Social Science Research – LSVV (2009-2011)**

- EU Structural Funds, coordinated by the Palacky University of Olomouc
- The aim is to upgrade educational programmes of sociology, political science and European studies.
- In 2009 SDA organised three workshops on data management, sources of research data and data quality and participated in creation of the Laboratory.

**Research on foreigners’ incomes, expenditures and remittances (2010-2011)**

- Funded by Czech Statistical Office (Eurostat)
- Survey of economically active foreigners from the third countries
- The aim is to improve estimations of remittances sent from the Czech Republic and to explore the structure of migrant workers’ expenditures.
SDA current projects

Memorandum of Understanding CESSDA-ERIC was signed

New CESSDA related project on the way…

- Construction of the Czech node of CESSDA-ERIC and its integration into this large-scale pan-European research infrastructure which aims to provide data services for socio-economic research
- 2010?-2015, supported by MILSA, head of the project is PhD. J. Krejčí
- CESSDA has been selected by the ESFRI Roadmap process to be upgraded as one of the central research infrastructures of the European Research Area.
- Construction of CESSDA-ERIC node at SDA, upgrade of and improvement data services and applying common CESSDA standards, hiring new staff (6FTE all together)
- Operational costs are fully covered from 2012
- Among other activities archive should also facilitate the combination of multiple data sources within research projects and contribute to the creation of new empirical material by linking and merging databases.
Czech Statistical Office (CZSO)

- Consulting and advisory activities
- Official excerpts on inflation rate, wages reached in individual regions or etc.
- Preparation and processing of customized information according to individual requirements.
- Selection of data from databases of individual statistics - demography, branch statistics (industry, agriculture, foreign trade, etc.), from the surveys carried out by the CZSO (household budgets statistics, price statistics, labour force sample survey) and last but not least extracts from the Business Register, regional databases.
- Providing and procuring of election statistics and population census results.

- Central Statistical Library is out of order after the floods
CZSO Microdata:

- Population Census or Population Register
- Household Surveys (e.g. Labour Force Survey, Household Income or Household Expenditure Surveys, Household Panel Surveys, Youth Cohort or Education Longitudinal Studies, etc.), EU-SILC,
- Individual Data from administrative registries (e.g. health data, social and welfare schemes),
- Surveys/Census of Businesses, or Business Register data
- other economic and financial data
- CZSO merges data from different sources but rarely on request
CZSO Data services:

- Public use files and tabulations, scientific use files/subsets - direct access on the internet
- Other form of access (on-site, mail, other on demand): all stated before plus extracts or subsets, special tabulations and raw data --- fees paid by the end user
- No access from a data service
- No special license for the institutional usage (named users, named project)
- Every end user (researcher), have to sign the contract, i.e. the declaration of confidentiality.
- For foreign researchers and residents of other countries the additional requirements are applied.
primary level data – mainly database of aggregated statistical data

secondary level data – statistical outputs (statistical tables, maps, graphs)
Public Database of the CZSO (VDB) is developed as a fundamental and unified data source for presentation of statistical data designed mainly for the public.

VDB contains solely aggregated statistical data covering all observed areas of statistics. It uses results of statistical data processing in the CZSO as well as statistical data from external and administrative sources especially from other workplaces of the state statistical service. It does not focus only on data covering the Czech Republic, but, in addition, it provides data for territorial administrative units of the CR (regions, districts, municipalities and cities, etc.) and also data from abroad.

VDB is linked to the statistical metainformation system. The data is identified by means of metainformation specifying the subject-matter, time and space dimension. Each data is equipped with information on data quality, confidentiality, source, contact person and wide range of other supplementary information.

Public Database secures confidentiality of presented data.

VDB is based on a single figure principle. The data may appear in a different context in different outputs (tables, maps, graphs, etc.).

VDB aims to provide an interactive user friendly access to statistical information including relevant metainformation, also to users with impaired vision and in multi-language regime.

VDB as well as other information systems:

Use
- interactive access to information for
- immediate provision of partial information
- export and print of displayed information
- preparation of more demanding (special) orders
- preparation of standard and non-standard e-publications
- preparation of standard and non-standard e-publications

VDB structure
VDB is created as a data mart drawing data from the following areas:

VDB includes the following:
- primary level – containing mainly data
- secondary level – containing statistical outputs (statistical tables, maps, graphs)
- interface – applications securing transformation and input of data to primary database (including universal XML interface) and applications enabling access to data and outputs

The core of VDB is a database with statistical indicators and their values. The structure of this database is closely linked to the metainformation system of the CZSO and mainly to the database of nomenclatures and classifications.
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Počet obyvatel v oblastech, krajích a okresech České republiky k 1. 1. 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Population of territorial units of the Czech republic, 1 January 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Kód území</td>
<td>Oblast, kraj, okres</td>
<td>Počet obyvatel</td>
<td>Population</td>
<td>Průměrný věk</td>
<td>Average age</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Code of</td>
<td>NUTS 2, NUTS 3, LAU 1</td>
<td>celkem</td>
<td>Males</td>
<td>Females</td>
<td>Total</td>
<td>Males</td>
</tr>
<tr>
<td>6</td>
<td>NUTS LAU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CZ0</td>
<td>Česká republika</td>
<td>10506813</td>
<td>5157197</td>
<td>5349616</td>
<td>40,6</td>
<td>39,1</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Oblast (Areas - NUTS 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>CZ31</td>
<td>Praha</td>
<td>1249026</td>
<td>608316</td>
<td>640710</td>
<td>41,6</td>
<td>39,8</td>
</tr>
<tr>
<td>11</td>
<td>CZ32</td>
<td>Středočeský</td>
<td>1247533</td>
<td>616058</td>
<td>631475</td>
<td>40,0</td>
<td>38,7</td>
</tr>
<tr>
<td>12</td>
<td>CZ33</td>
<td>Jihozápad</td>
<td>1209506</td>
<td>597391</td>
<td>612115</td>
<td>40,8</td>
<td>39,5</td>
</tr>
<tr>
<td>13</td>
<td>CZ34</td>
<td>Severozápad</td>
<td>1143034</td>
<td>563675</td>
<td>590169</td>
<td>39,9</td>
<td>38,4</td>
</tr>
<tr>
<td>14</td>
<td>CZ35</td>
<td>Severovýchod</td>
<td>1506758</td>
<td>741278</td>
<td>765480</td>
<td>40,6</td>
<td>39,1</td>
</tr>
<tr>
<td>15</td>
<td>CZ36</td>
<td>Jihovýchod</td>
<td>1666700</td>
<td>817509</td>
<td>84991</td>
<td>40,8</td>
<td>39,2</td>
</tr>
<tr>
<td>16</td>
<td>CZ37</td>
<td>StředoMorava</td>
<td>1233083</td>
<td>601973</td>
<td>631110</td>
<td>40,8</td>
<td>39,2</td>
</tr>
<tr>
<td>17</td>
<td>CZ38</td>
<td>Moravskoslezský</td>
<td>1247373</td>
<td>610997</td>
<td>636376</td>
<td>40,4</td>
<td>38,8</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>Kraje (Regions - NUTS 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>CZ510</td>
<td>Hlavní město Praha</td>
<td>1240626</td>
<td>603816</td>
<td>636710</td>
<td>41,6</td>
<td>39,8</td>
</tr>
<tr>
<td>21</td>
<td>CZ520</td>
<td>Středočeský</td>
<td>1247533</td>
<td>616058</td>
<td>631475</td>
<td>40,0</td>
<td>38,7</td>
</tr>
<tr>
<td>22</td>
<td>CZ531</td>
<td>Jihočeský</td>
<td>637643</td>
<td>314543</td>
<td>323000</td>
<td>40,7</td>
<td>39,3</td>
</tr>
<tr>
<td>23</td>
<td>CZ532</td>
<td>Plzeňský</td>
<td>571863</td>
<td>268248</td>
<td>268901</td>
<td>41,0</td>
<td>39,6</td>
</tr>
<tr>
<td>24</td>
<td>CZ541</td>
<td>Karlovarský</td>
<td>307636</td>
<td>151056</td>
<td>156580</td>
<td>40,1</td>
<td>38,7</td>
</tr>
<tr>
<td>25</td>
<td>CZ542</td>
<td>Ústecký</td>
<td>836190</td>
<td>412619</td>
<td>423579</td>
<td>39,6</td>
<td>38,3</td>
</tr>
<tr>
<td>26</td>
<td>CZ551</td>
<td>Liberecký</td>
<td>430027</td>
<td>215000</td>
<td>224027</td>
<td>40,1</td>
<td>38,7</td>
</tr>
<tr>
<td>27</td>
<td>CZ552</td>
<td>Královéhradecký</td>
<td>554402</td>
<td>271733</td>
<td>282669</td>
<td>41,1</td>
<td>39,5</td>
</tr>
<tr>
<td>28</td>
<td>CZ553</td>
<td>Pardubický</td>
<td>516329</td>
<td>254545</td>
<td>261874</td>
<td>40,6</td>
<td>39,0</td>
</tr>
<tr>
<td>29</td>
<td>CZ563</td>
<td>Vysocina</td>
<td>514992</td>
<td>255492</td>
<td>259000</td>
<td>40,5</td>
<td>39,1</td>
</tr>
<tr>
<td>30</td>
<td>CZ564</td>
<td>Jihomoravský</td>
<td>1151708</td>
<td>562017</td>
<td>589691</td>
<td>40,9</td>
<td>39,2</td>
</tr>
</tbody>
</table>
CZSO Pricing Policy:

- Price list is determined by CZSO on the basis of:
  - legal framework
  - agreements with governmental bodies
  - cas-by-case basis.

- **Justification**: extra work and data delivery, not for data itself (?)

Cooperation experiences

- The Institute is treated as an end user (National Accounts Statistics, Family Budget Survey, etc.)

- Experience in collecting data for CZSO (SDA), providing the expertise and consultations (for example Department of Economic Sociology and Socioeconomics of Housing Department), discussing data harmonisation (forthcoming workshop)

- SDA and other departments have good experiences with CZSO based on the personal communication

- Differentiated approach to the clients (SDA as a “gatekeeper”)

CZSO attitude towards cooperation according to CESSDA PPP NSI’s survey:

Q25. Whether or not your organisation already has an agreement with a data archive for distribution of its microdata, do you believe that such an agreement benefits/would benefit your organisation in its dissemination activities? - NO
Thank you for your attention!

Yana.Leontiyeva@soc.cas.cz