

gesis

Leibniz Institute
for the Social Sciences



PIAAC Data Analysis in Stata: A practical guide

Video 2: Data Access and First Steps in Stata
GESIS – Leibniz Institute for the Social Sciences

In a Nutshell

Goal:

- Practical guide for three existing tools to perform PIAAC data analysis in Stata

Target Group:

- Researchers with some experience in Stata, but little or no experience on analysis of PIAAC data in Stata

Structure:

- General overview of how the three tools work, as well as several exemplary analyses with the PIAAC data



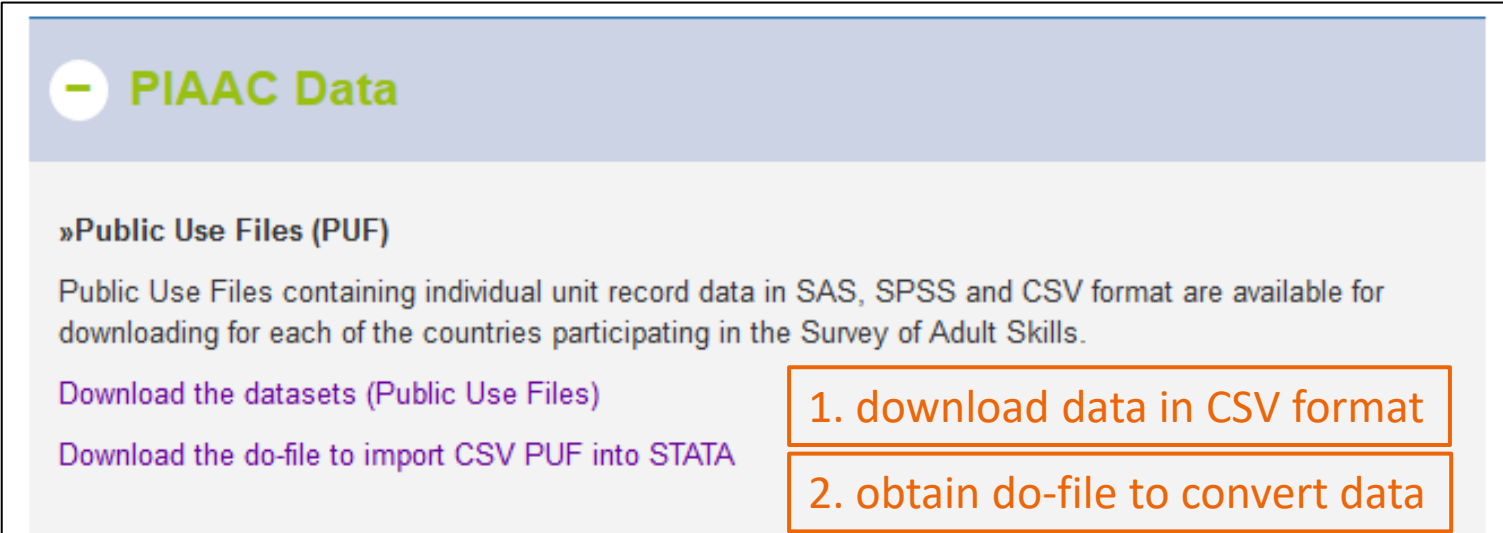
*Video 2:
Data Access and First Steps in Stata*

Data Access I

Public Use Files (PUFs) for most countries can be accessed via the OECD webpage:

<http://www.oecd.org/skills/piaac/data/>

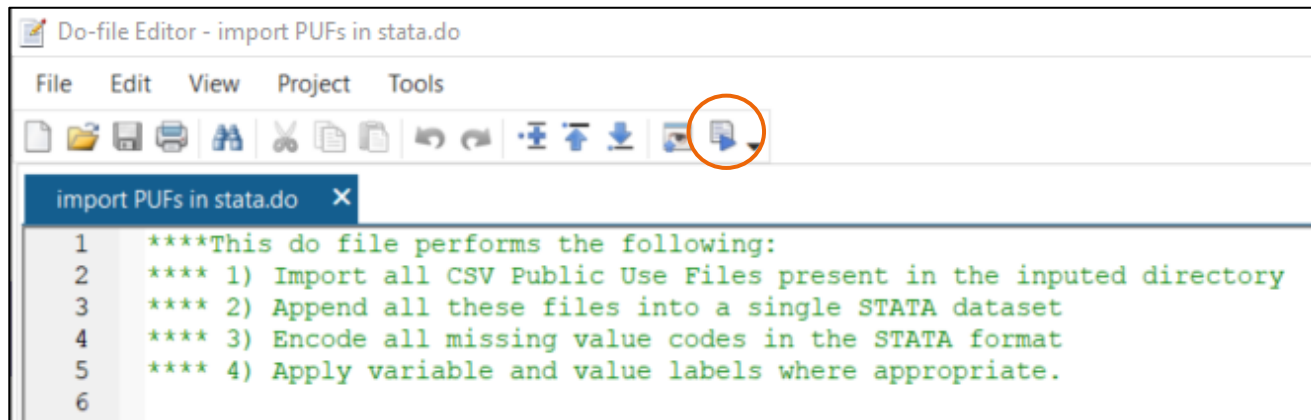
1. Download data in CSV format
2. Download do-file to convert data into Stata format



The screenshot shows the 'PIAAC Data' section of the OECD website. It features a header with a minus sign icon and the text 'PIAAC Data'. Below this, there is a sub-section titled '»Public Use Files (PUF)'. The text under this section states: 'Public Use Files containing individual unit record data in SAS, SPSS and CSV format are available for downloading for each of the countries participating in the Survey of Adult Skills.' There are two links: 'Download the datasets (Public Use Files)' and 'Download the do-file to import CSV PUF into STATA'. Two orange boxes are overlaid on the right side of the screenshot, containing the text: '1. download data in CSV format' and '2. obtain do-file to convert data'.

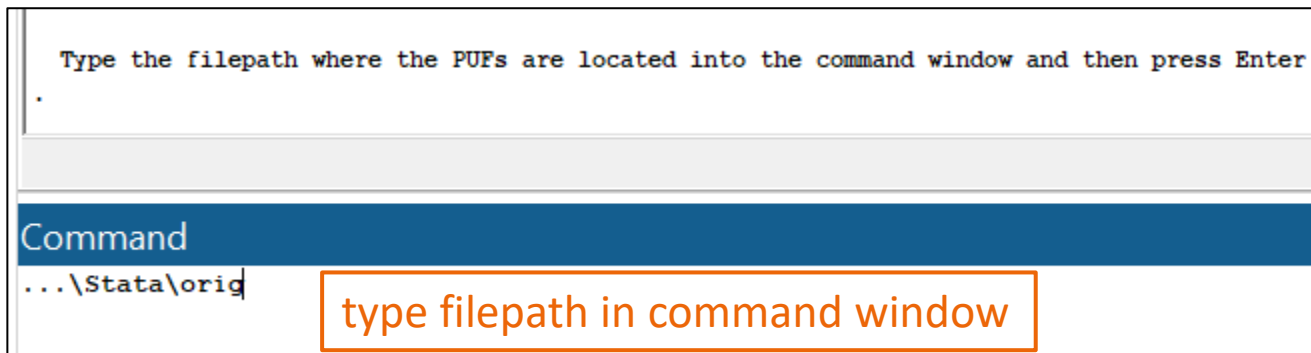
Data Access II

3. Run do-file



```
Do-file Editor - import PUFs in stata.do
File Edit View Project Tools
import PUFs in stata.do x
1 ****This do file performs the following:
2 **** 1) Import all CSV Public Use Files present in the inputed directory
3 **** 2) Append all these files into a single STATA dataset
4 **** 3) Encode all missing value codes in the STATA format
5 **** 4) Apply variable and value labels where appropriate.
6
```

4. Stata will request filepath where CSV files are saved:



```
Type the filepath where the PUFs are located into the command window and then press Enter
.
Command
...\Stata\orig
type filepath in command window
```

Data Access III

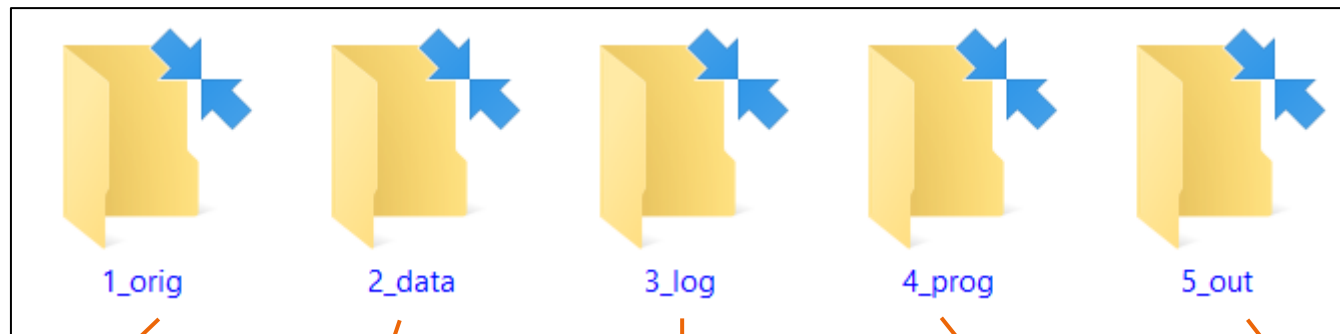
5. Save data file with all PUFs

```
Your PIAAC dataset is now ready to be used in Stata. Please consider saving this file.  
  
.end of do-file  
  
.  
  
Command  
save ...\Stata\data\PIAAC_All_PUFs.dta
```

Note: Country-specific Scientific Use Files (SUFs) may contain further variables and may be obtained from national data providers,
e.g. for Germany: Research Data Center PIAAC,
<https://www.gesis.org/en/piaac/rdc>

Data Management

- Suggested folder structure for data analysis:



original data

prepared data

log files

syntax files

results

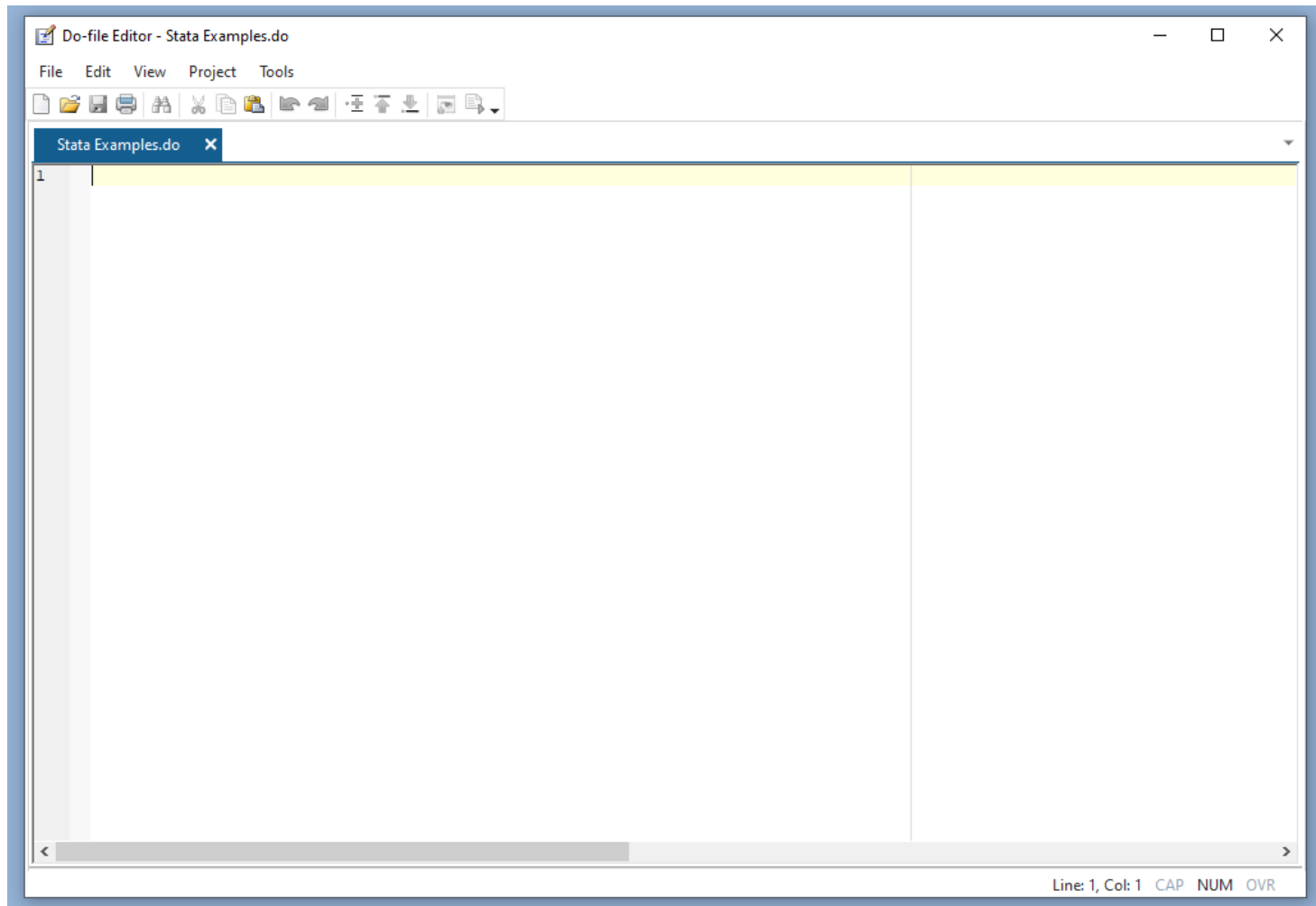
Variables in PIAAC

1. Respondents' answers to background questionnaire (BQ)
 - e.g., job satisfaction: `d_q014`
2. Variables derived from respondents' answers to BQ
 - e.g., current employment status: `c_d05`
3. Respondents' answers to cognitive items
 - e.g., scored response on “orchestra ticket”: `e664001s`
4. Ten plausible values (PVs) for each skill domain
 - e.g., `pvlit1`, `pvlit2`, ..., `pvlit10` for literacy
5. Interviewers' answers to observation module
 - e.g., assessment taking too long: `zz5`
6. Auxiliary variables on survey workflow and design
 - e.g., paper-based vs. computer-based assessment: `pbroute`

First Steps in Stata

- Tabulate and label country IDs
 - ▶ `cntryid`
- Tabulate plausible values for all skill domains:
 - ▶ `pvlit*`
 - ▶ `pvnum*`
 - ▶ `pvpsl*`
- Tabulate replicate weights:
 - ▶ `spfw*`
- Save smaller data set only including Germany, Greece, Singapore, Sweden, and the US
 - we will use this data set for the following sample analyses

Let's go to Stata



gesis

Leibniz Institute
for the Social Sciences

Leibniz
Leibniz
Association