UNIVERSITÄT OSNABRÜCK

Work and Organizational Psychology • Emphasis Cross-Cultural Business Psychology

STATA FOR PRACTITIONERS:

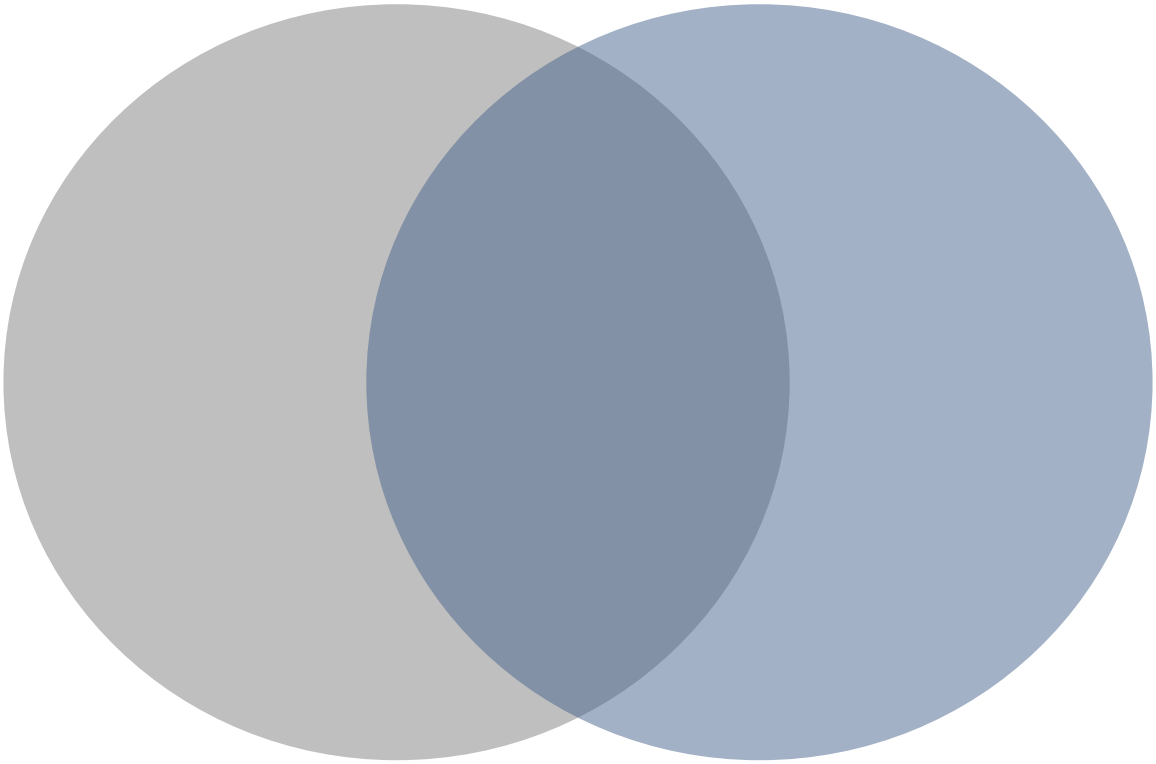
CREATING AUTOMATED REPORTS IN EXCEL

Dipl.-Soz.Wiss.
SVEN-OLIVERSPIESS

Stata well-established among academics

However, even in those applied areas which rely heavily on quantitative data (e.g. polling, market research, etc.) Stata arguably less popular

Some fields, however, reliant on both domains



Being able to use same tools across settings really handy

- In release 12 many useful features to make Stata more useful in even more settings
- One important yet somewhat neglected aspect is built-in compatibility with MS Excel data format

However, mixed feelings from both academics as well as practitioners

Objectives of this talk:

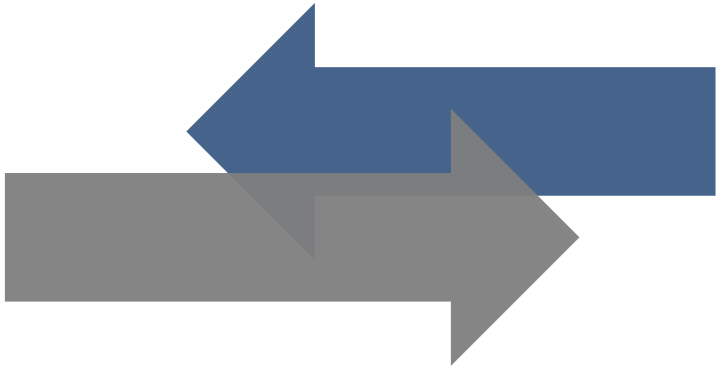
1. Illustrate some of the new possibilities deriving from the compatibility of Stata 12 with MS Excel

2. Make the case that Stata is actually also quite useful for practitioners

3. Advance the idea that the use of Stata outside of academia can promote scientific standards (e.g. reproducibility) not impede them

The problem: How to get from data to reports?

- Getting information frequently tit for tat: data for results

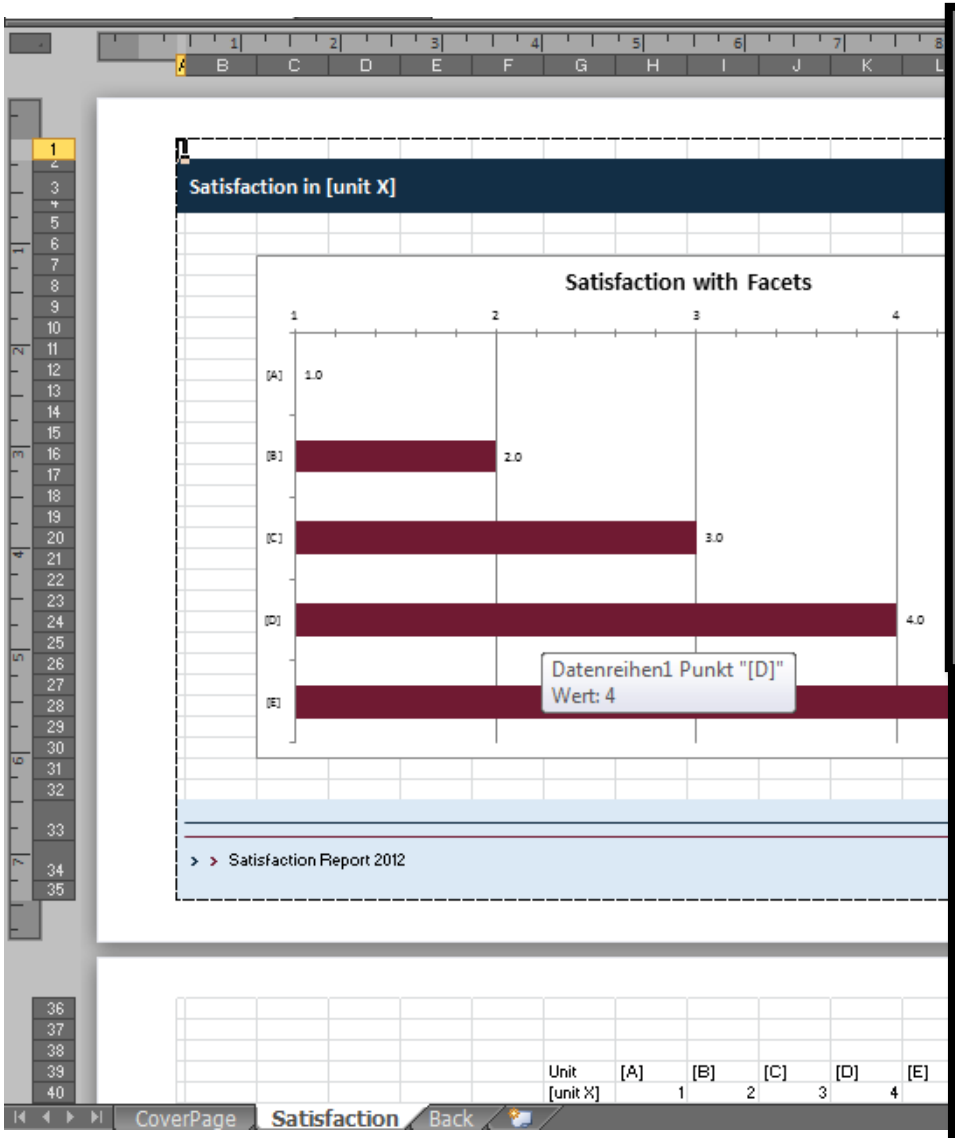


- Examples:
  - Student performance in schools
  - **Job satisfaction among employees**

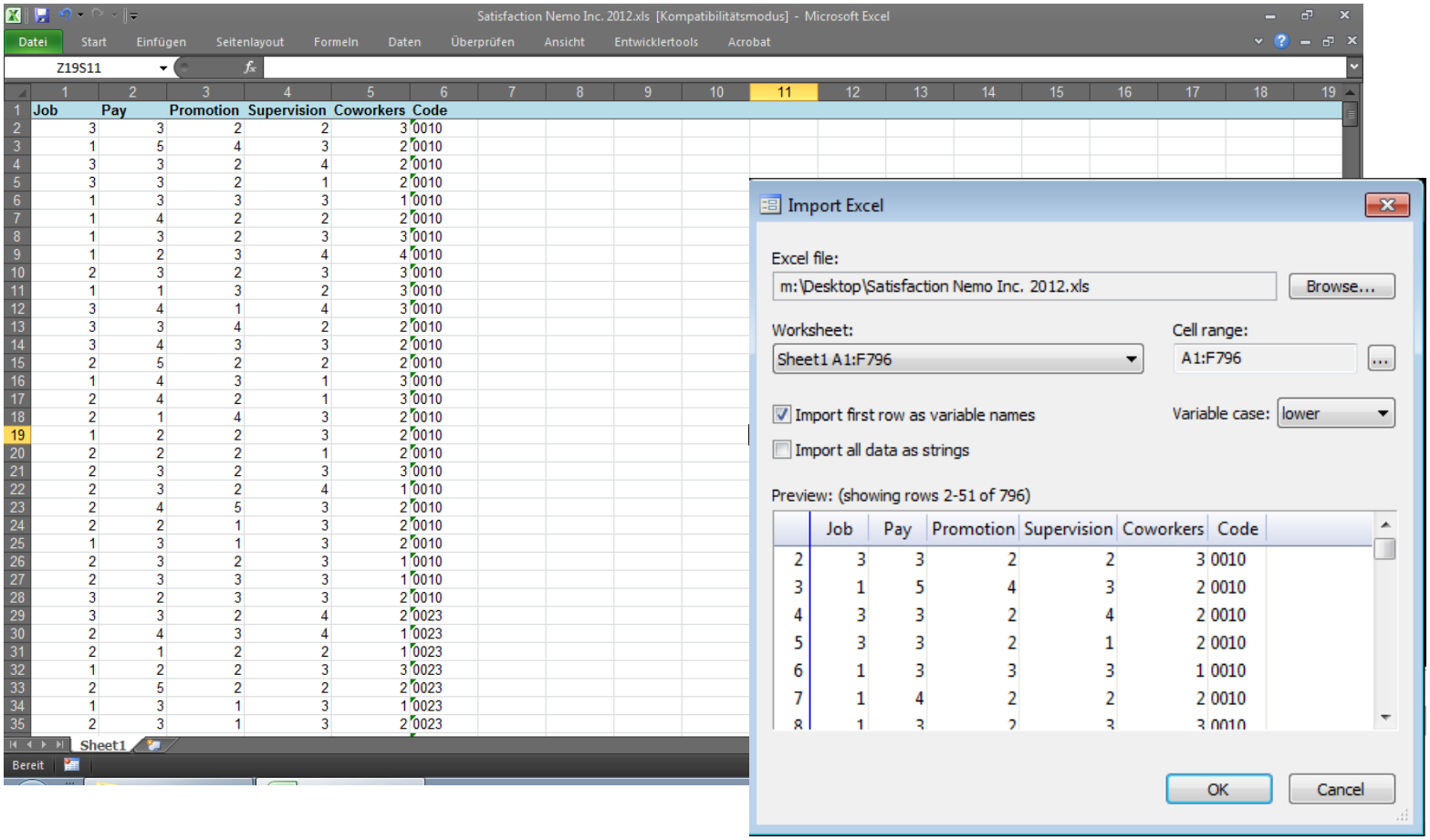
The steps

1. Create Excel report template
2. Import raw data
3. Clean data and compute statistics
4. Loop over units to save results for each unit into a new copy of the template

Step 1 – Creation of report template



Step 2 – Raw data import




```
import excel "Satisfaction data Nemo Inc 2012.xls", ///
      sheet("Sheet1") firstrow
```

Step 3 – Data cleaning and computation of needed results

- In this basic example simply unit means for the five satisfaction facets:

```
collapse Job Pay Promotion Supervision Coworkers,
by(Unit)
```

- Especially in more complex cases big advantage that all data manipulation and computation steps are documented and fully reproducible if needed

- If any errors occur they can be easily corrected locally without the need to start all over again

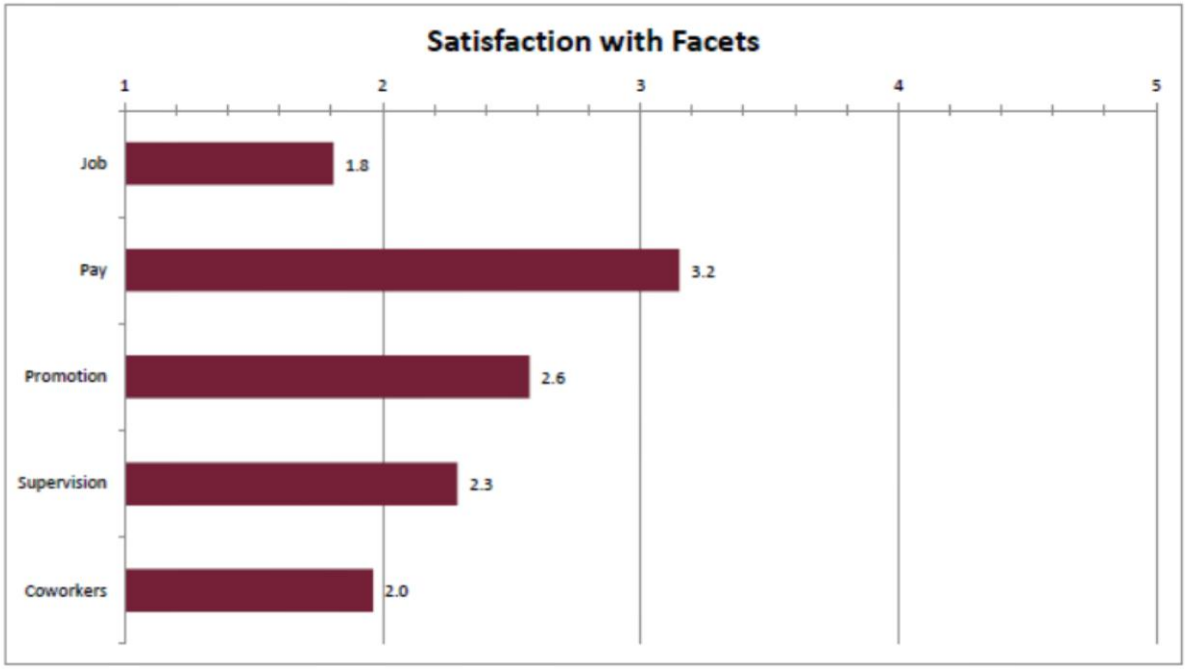
Step 4 – Save results for each unit into a new copy of the template

- Run loop which creates report for each unit by (a) deleting other units, (b) copying template for each unit with respective name, and (c) export data into template:

```
levelsof Unit, local(units)
preserve      // make changes to data reversible

foreach u of local units {
   keep if Unit==`u'

   copy "report_template.xlsx" "reports/unit_`u'.xlsx", replace

   export excel using "reports\unit_`u'.xlsx", sheet("Satisfaction")
   sheetmodify cell(G39) firstrow(variables)

   restore, preserve  // restore full dataset for next iteration
 }
```

Example of a final report

Satisfaction in Branch 23



Outlook

- This was just a very basic example

- Additional possibilities:
  - Averages of all or selected units for comparison
  - Different languages in international projects
  - Time trends
  - Plotting advanced graphs
  - Creating PDF documents
  - …

Cons

- Most useful only for highly standardized reports
- For some simple cases with only very few reports initial setup might take longer than producing each report individually

Pros

- Added value increases the more reports need to be created and the more complex the structure (e.g. multilingual reporting, time trends, etc.)
- Using syntax for the entire process of data management and analysis makes all steps reproducible and simultaneously easy to adjust
- Use of wide-spread Excel format facilitates collaboration with practitioners which can enhance *both* good science and good practice

➔ Overall, compatibility with MS Excel format opened many exciting possibilities when dealing with quantitative data

Thank you!

Dipl.-Soz.Wiss.
SVEN-OLIVERSPIESS