

# BFS R package

Search and download data from the Federal Statistical Office

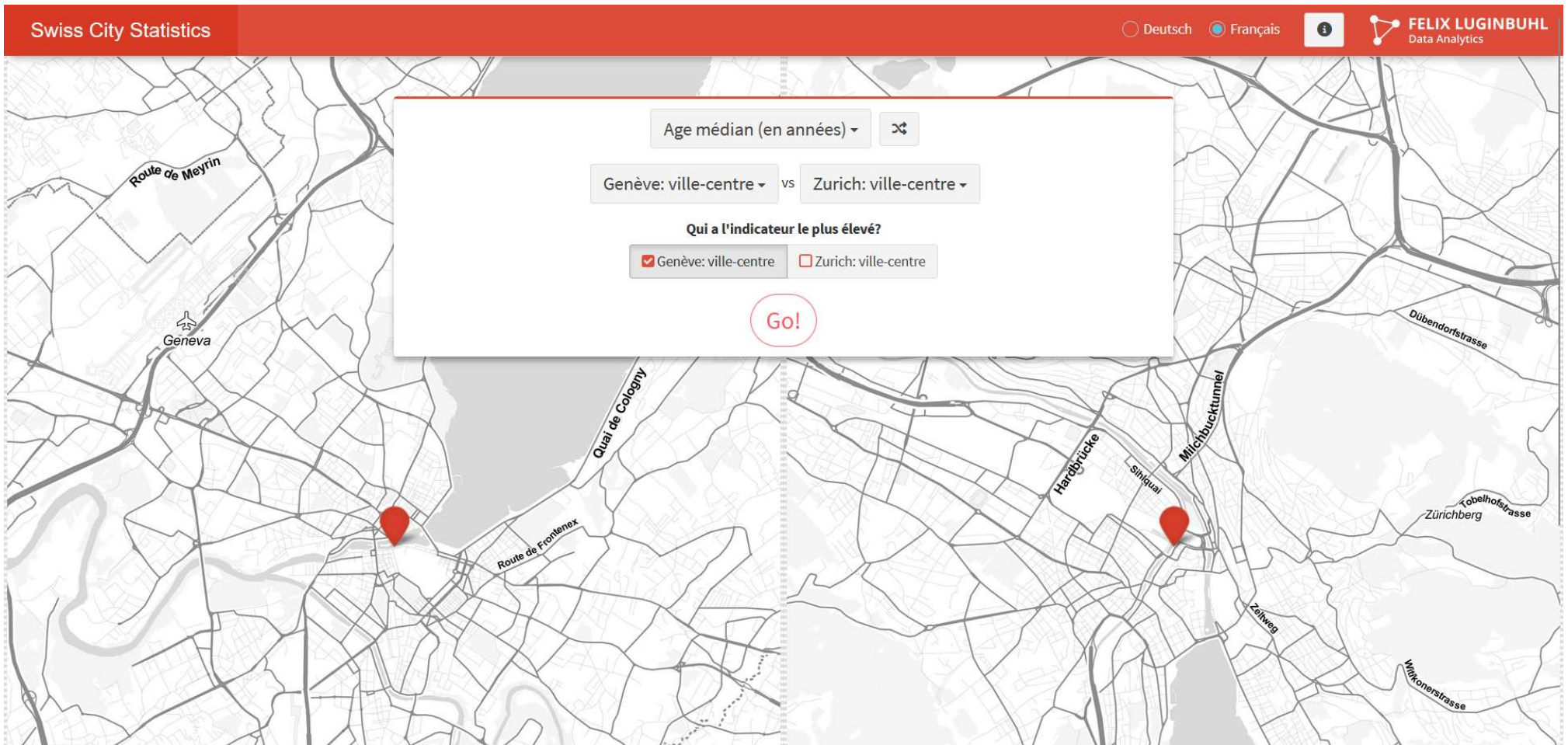
Félix Luginbühl

# Plan

1. A use case example.
  2. Motivations behind the BFS package
  3. Exploring the data catalog with BFS
  4. All you need is `bfs_get_data()`
- ... and your questions

# 1. A use case example

- Swiss City Statistics app (city.felixluginbuhl.com)



## 2. Motivations behind the BFS package

# Motivations behind the BFS package

- leveraging the full R ecosystem
- Data reproducibility and transparency
- Speed: get BFS data with 1 line of code

# 3. Exploring the data catalog with BFS

# Exploring the data catalog manually

## STAT-TAB - Interactive tables

The screenshot shows the Federal Statistical Office (FSO) website. The header includes the FSO logo and name in multiple languages, a search bar, and navigation links for Cart, Media, Jobs, and Contact. The main navigation bar lists various services like News, Look for statistics, Services, Basics and Surveys, Registers, National data management NaDB, Data Science Competence Center, and The FSO. The breadcrumb trail indicates the path: Federal Statistical Office > Services > Research > STAT-TAB - Interactive tables. The left sidebar shows a 'Research' section with 'STAT-TAB - Interactive tables' selected. The main content area is titled 'STAT-TAB - Interactive tables' and contains a description of the interactive database, its features, and a link to the direct access page. A contact information box is visible on the right side of the page.

Federal Administration > FDHA > FSO

Cart Media Jobs Contact DE FR IT **EN** RM

Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

**Federal Statistical Office**

Search...

News Look for statistics **Services** Basics and Surveys Registers National data management NaDB Data Science Competence Center The FSO

Federal Statistical Office > Services > Research > STAT-TAB - Interactive tables

< Services

**Research**

**STAT-TAB - Interactive tables**

Methodological Reports

Customized data exploitations

Experimental statistics


### STAT-TAB - Interactive tables

STAT-TAB is the FSO's interactive database, containing detailed statistical data and enabling simple, individual data queries.

Data are presented in 19 different topics, in the form of multi-dimensional data cubes. Data cubes can be selected either directly in the application or by using the filter search (see below).


Each data cube combines several dimensions (variables). By selecting the desired variables/values, an online table is created that, in addition to the selected data, also contains metadata and explanatory notes. Users can either continue to work on this table in STAT-TAB or export it in a standard data format and work on it offline.


Further useful tips on STAT-TAB can be found in the guide below.

 **STAT-TAB - Direct access**  
Interactive database

**Contact**

Federal Statistical Office  
Espace de l'Europe 10  
CH-2010 Neuchâtel  
Switzerland

 **Contact**



# Exploring the data catalog with BFS

```
1 #install.packages("BFS")  
2 library(BFS)
```



# Exploring the data catalog with BFS

- Get available datasets (from official [BFS RSS feed](#))

```
1 library(BFS) #install.packages("BFS")
2
3 # get the data catalogue
4 catalog_data_en <- bfs_get_catalog_data(language = "en")
5
6 catalog_data_en
```

## # A tibble: 179 × 5

| ##       | title  | langu... <sup>1</sup> | publi... <sup>2</sup> | url_bfs   |
|----------|--|-----------------------|-----------------------|-----------|
| url_px   |  |                       |                       |           |
| ##       | <chr>  | <chr>                 | <chr>                 | <chr>     |
| <chr>    |  |                       |                       |           |
| ##       | 1 Businesses by difficulties in recruiting staf... | en                    | Busine...             | https:... |
| https... |  |                       |                       |           |
| ##       | 2 Businesses by difficulties in recruiting staf... | en                    | Busine...             | https:... |
| https... |  |                       |                       |           |
| ##       | 3 Businesses by employment prospects and econom... | en                    | Busine...             | https:... |
| https... |  |                       |                       |           |
| ##       | 4 Businesses by employment prospects and major ... | en                    | Busine...             | https:... |
| https... |  |                       |                       |           |

<https://felixluginbuhl.com/BFS>

# Exploring the data catalog with BFS

- Choose a dataset using `filter()` from dplyr.

```
1 library(BFS) #install.packages("BFS")
2 library(dplyr) #install.packages("dplyr")
3
4 catalog_data_en <- bfs_get_catalog_data(language = "en")
5
6 # search for a recent dataset
7 catalog_data_uni <- catalog_data_en %>%
8   filter(title == "University students by year, ISCED field, sex and level
9
10 catalog_data_uni
```

```
## # A tibble: 1 × 5
```

```
##   title                                                                 langu...1 publi...2 url_bfs
url_px
##   <chr>                                                                 <chr>      <chr>      <chr>
<chr>
```

```
## 1 University students by year, ISCED field, sex ... en          Univer... https:...
https...
```

```
## # ... with abbreviated variable names 1language, 2published
```

<https://felixluginbuhl.com/BFS>

# Exploring the data catalog with BFS

- Get the BFS dataset with `bfs_get_data()`.

```
1 library(BFS) #install.packages("BFS")
2 library(dplyr) #install.packages("dplyr")
3
4 catalog_data_en <- bfs_get_catalog_data(language = "en")
5
6 catalog_data_uni <- catalog_data_en %>%
7   filter(title == "University students by year, ISCED field, sex and level
8
9 # get the data
10 df_uni <- bfs_get_data(url_bfs = catalog_data_uni$url_bfs, language = "en")
11
12 df_uni
```

```
## # A tibble: 17,640 × 5
```

```
##   Year   `ISCED Field`   Sex   `Level of study`
```

```
Unive...1
```

```
##   <chr>   <chr>           <chr>   <chr>
```

```
<dbl>
```

```
##   1 1980/81 Education science Male   First university degree or diploma
```

```
545
```

<https://felixluginbuhl.com/BFS>

|    |   |         |                   |      |  |
|----|---|---------|-------------------|------|--|
| ## | 2 | 1980/81 | Education science | Male | Bachelor                               |
| 0  |   |         |                   |      |  |
| ## | 3 | 1980/81 | Education science | Male | Master                                 |
| 0  |   |         |                   |      |  |
| ## | 4 | 1980/81 | Education science | Male | Doctorate                              |
| 93 |   |         |                   |      |  |
| ## | 5 | 1980/81 | Education science | Male | Further education, advanced studies... |
| 13 |   |         |                   |      |  |

# Exploring the data catalog with BFS

- get additional footnotes information

```
1 library(BFS) #install.packages("BFS")
2 library(dplyr) #install.packages("dplyr")
3
4 catalog_data_en <- bfs_get_catalog_data(language = "en")
5
6 catalog_data_uni <- catalog_data_en %>%
7   filter(title == "University students by year, ISCED field, sex and level
8
9 df_uni <- bfs_get_data(url_bfs = catalog_data_uni$url_bfs, language = "en")
10
11 # get data comments
12 comments <- bfs_get_data_comments(url_bfs = catalog_data_uni$url_bfs, langu
13
14 comments$comment
```

[1] "To ensure that the presentations from cubes containing the 'level of studies' variable are easy to understand, all post-graduate studies are included under the heading 'Continuing and further education', which are additionally published under the following headings:\r\n- Continuing education\r\n- Specialised and <http://he.europa.eu/data-catalogue/BFS>\r\n- Postgraduate studies

4. All you need is  
`bfs_get_data()`

# All you need is `bfs_get_data()`

- Using `pxweb` R package functions under the hood to query the Swiss Federal Statistical Office PXWEB API.<sup>1</sup>



1. PXWEB is an API structure developed by Statistics Sweden and other national statistical institutions to disseminate public statistics in a structured way.  
<https://felixluginbuhl.com/BFS>

# All you need is `bfs_get_data()`

- Better reproducibility (and stability) using ‘number\_bfs’.

```
1 # open webpage
2 browseURL("https://www.bfs.admin.ch/content/bfs/en/home/statistiken/katalog
```

Catalogues and Databases

Tables

**Data**

Graphs

Infographics

Maps

Publications


Press Releases

Surveys


Classifications

Definitions

University students by year, ISCED field, sex and level of study



Link STAT-TAB

 Download PX  
(PX, 77 kB)

Language

Type

Published on

Observation period

Published by

Terms of Use

Theme

FSO number

Surviv. statistics

DE/EN/FR/IT

Table

26.03.2021

1980-2020

Federal Statistical Office

[OPEN-BY-ASK](#)

Education and science

[px-x-1502040100\\_131](#)

[Students and degrees of higher education](#)

Remark

Our English pages offer only a limited range of information on our statistical production. For our full range please consult our pages in French and German (top right hand screen).

<https://felixluginbuhl.com/BFS>



# All you need is `bfs_get_data()`

- Better reproducibility (and stability) with `number_bfs`.

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "en"  
4 )
```

```
# A tibble: 18,060 × 5
```

|   | Year    | `ISCED Field`     | Sex   | `Level of study`   | `University students` |
|---|---------|-------------------|-------|--------------------|-----------------------|
|   | <chr>   | <chr>             | <chr> | <chr>              | <dbl>                 |
| 1 | 1980/81 | Education scie... | Male  | First universit... | 545                   |
| 2 | 1980/81 | Education scie... | Male  | Bachelor           | 0                     |
| 3 | 1980/81 | Education scie... | Male  | Master             | 0                     |
| 4 | 1980/81 | Education scie... | Male  | Doctorate          | 93                    |
| 5 | 1980/81 | Education scie... | Male  | Further educati... | 13                    |

```
# i 18,050 more rows
```

# All you need is `bfs_get_data()`

- Change the `language`

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "de"  
4 )
```

```
# A tibble: 18,060 × 5  
  Jahr      `ISCED Fach` Geschlecht Studienstufe Studierende an den u...1  
  <chr>    <chr>          <chr>      <chr>                <dbl>  
1 1980/81 Erziehungswi... Mann      Lizenziat/D...      545  
2 1980/81 Erziehungswi... Mann      Bachelor           0  
3 1980/81 Erziehungswi... Mann      Master              0  
4 1980/81 Erziehungswi... Mann      Doktorat            93  
5 1980/81 Erziehungswi... Mann      Weiterbildu...     13  
# i 18,050 more rows  
# i abbreviated name: 1 `Studierende an den universitären Hochschulen`
```

# All you need is `bfs_get_data()`

- Clean names with `janitor::clean_names()` (snake case)

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "de",  
4   clean_names = TRUE  
5 )
```

```
# A tibble: 18,060 × 5
```

|   | jahr    | isced_fach      | geschlecht | studienstufe   | studierende_an_den_u... <sup>1</sup> |
|---|---------|-----------------|------------|----------------|--------------------------------------|
|   | <chr>   | <chr>           | <chr>      | <chr>          | <dbl>                                |
| 1 | 1980/81 | Erziehungswi... | Mann       | Lizenziat/D... | 545                                  |
| 2 | 1980/81 | Erziehungswi... | Mann       | Bachelor       | 0                                    |
| 3 | 1980/81 | Erziehungswi... | Mann       | Master         | 0                                    |
| 4 | 1980/81 | Erziehungswi... | Mann       | Doktorat       | 93                                   |
| 5 | 1980/81 | Erziehungswi... | Mann       | Weiterbildu... | 13                                   |

```
# i 18,050 more rows
```

```
# i abbreviated name: 1studierende_an_den_universitaeren_hochschulen
```

# All you need is `bfs_get_data()`

- Query specific categories with `query`.

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "de",  
4   clean_names = TRUE,  
5   query = "all"  
6 )
```

# All you need is `bfs_get_data()`

- Use `bfs_get_metadata()` to get query `code` and `values` categories.

```
1 metadata <- BFS::bfs_get_metadata(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "de"  
4 )  
5 str(metadata)
```

```
tibble [4 × 6] (S3: tbl_df/tbl/data.frame)  
$ code      : chr [1:4] "Jahr" "ISCED Fach" "Geschlecht" "Studienstufe"  
$ text      : chr [1:4] "Jahr" "ISCED Fach" "Geschlecht" "Studienstufe"  
$ values    :List of 4  
..$ : chr [1:43] "0" "1" "2" "3" ...  
..$ : chr [1:42] "0" "1" "2" "3" ...  
..$ : chr [1:2] "0" "1"  
..$ : chr [1:5] "0" "1" "2" "3" ...  
$ valueTexts :List of 4  
..$ : chr [1:43] "1980/81" "1981/82" "1982/83" "1983/84" ...  
..$ : chr [1:42] "Erziehungswissenschaft" "Ausbildung von Lehrkräften ohne  
Fachspezialisierung" "Ausbildung von Lehrkräften mit Fachspezialisierung"
```

```
"Bildende Kunst" ...  
  ..$ : chr [1:2] "Mann" "Frau"
```

# All you need is `bfs_get_data()`

- Manually create BFS query dimensions.

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "en",  
4   query = list(  
5     "Jahr" = c("40", "41"),  
6     "ISCED Fach" = c("0"),  
7     "Geschlecht" = c("*"), # Use "*" to select all  
8     "Studienstufe" = c("2", "3")  
9   ))
```

# A tibble: 8 × 5

|   | Year    | `ISCED Field`      | Sex     | `Level of study` | `University students` |
|---|---------|--------------------|---------|------------------|-----------------------|
|   | <chr>   | <chr>              | <chr>   | <chr>            | <dbl>                 |
| 1 | 2020/21 | Education scien... | Male    | Master           | 151                   |
| 2 | 2020/21 | Education scien... | Male    | Doctorate        | 121                   |
| 3 | 2020/21 | Education scien... | Fema... | Master           | 555                   |
| 4 | 2020/21 | Education scien... | Fema... | Doctorate        | 306                   |
| 5 | 2021/22 | Education scien... | Male    | Master           | 143                   |

# All you need is `bfs_get_data()`

- Query the `code` variables and value types.

```
1 BFS::bfs_get_data(  
2   number_bfs = "px-x-1502040100_131",  
3   language = "en",  
4   query = list(  
5     "Jahr" = c("40", "41"),  
6     "ISCED Fach" = c("0"),  
7     "Geschlecht" = c("*"), # Use "*" to select all  
8     "Studienstufe" = c("2", "3")  
9   ),  
10  column_name_type = "code", # "text" by default  
11  variable_value_type = "code") # "text" by default
```

# A tibble: 8 × 5

|   | Jahr  | `ISCED Fach` | Geschlecht | Studienstufe | `University students` |
|---|-------|--------------|------------|--------------|-----------------------|
|   | <chr> | <chr>        | <chr>      | <chr>        | <dbl>                 |
| 1 | 40    | 0            | 0          | 2            | 151                   |
| 2 | 40    | 0            | 0          | 3            | 121                   |
| 3 | 40    | 0            | 1          | 2            | 555                   |
| 4 | 40    | 0            | 1          | 3            | 306                   |
| 5 | 41    | 0            | 0          |              | 143                   |

<https://felixluginbuhl.com/BFS>



# All you need is `bfs_get_data()`

- Documentation: [www.felixluginbuhl.com/BFS](http://www.felixluginbuhl.com/BFS)
- Source code: [www.github.com/lgnbhl/BFS](http://www.github.com/lgnbhl/BFS)

```
1 # open function documentation in R
2 ?bfs_get_data()
```

# 5. Questions

Thank you for your attention!

- BFS documentation: [felix.luginbuhl.com/BFS](https://felixluginbuhl.com/BFS)
- Swiss City Statistics app: [city.felixluginbuhl.com](https://city.felixluginbuhl.com)
- LinkedIn: [linkedin.com/in/felixluginbuhl/](https://linkedin.com/in/felixluginbuhl/)

