BFS R package

Search and download data from the Federal Statistical Office

Félix Luginbühl

https://felixluginbuhl.com/BFS

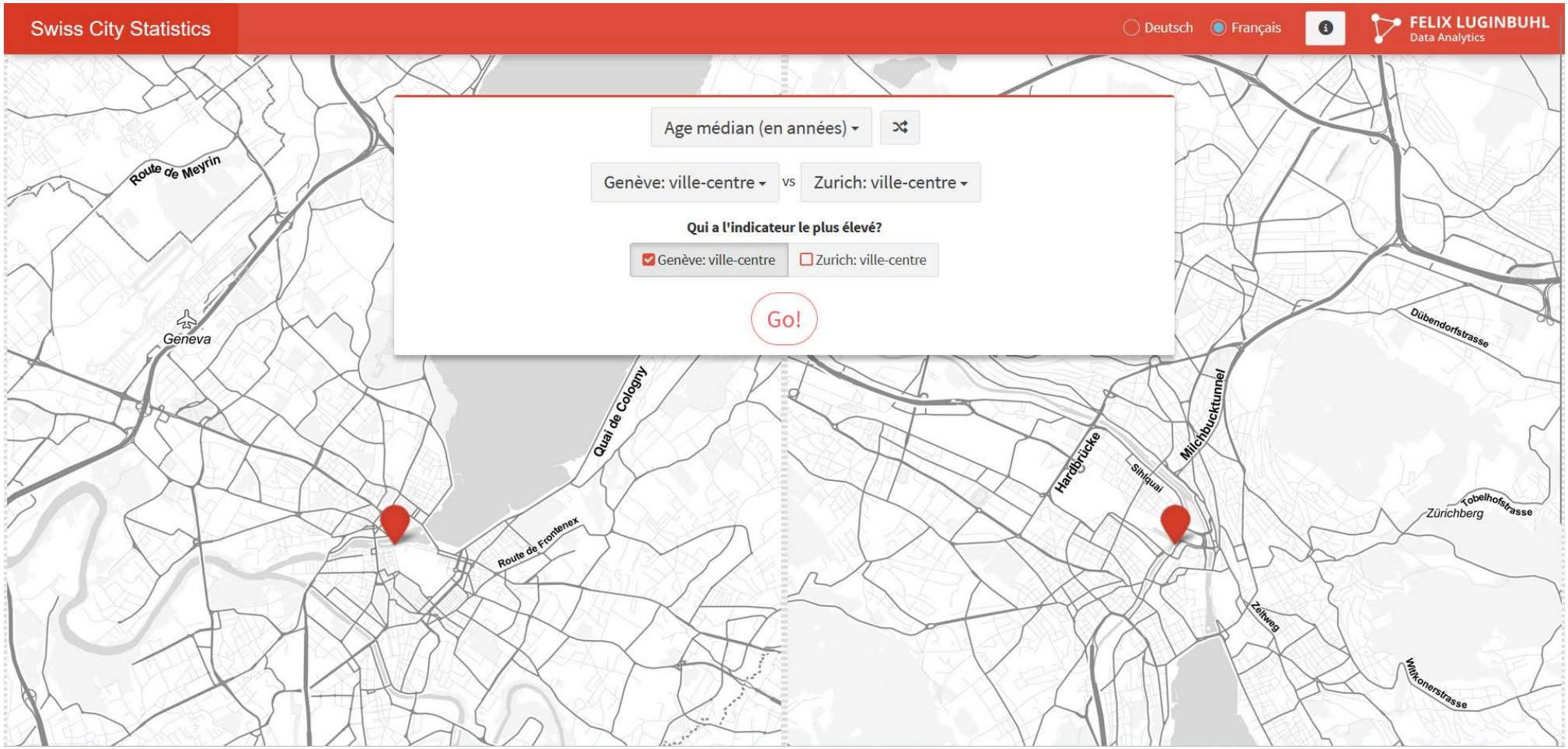
Plan

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)



https://felixluginbuhl.com/BFS

2. Motivations behind the BFS package

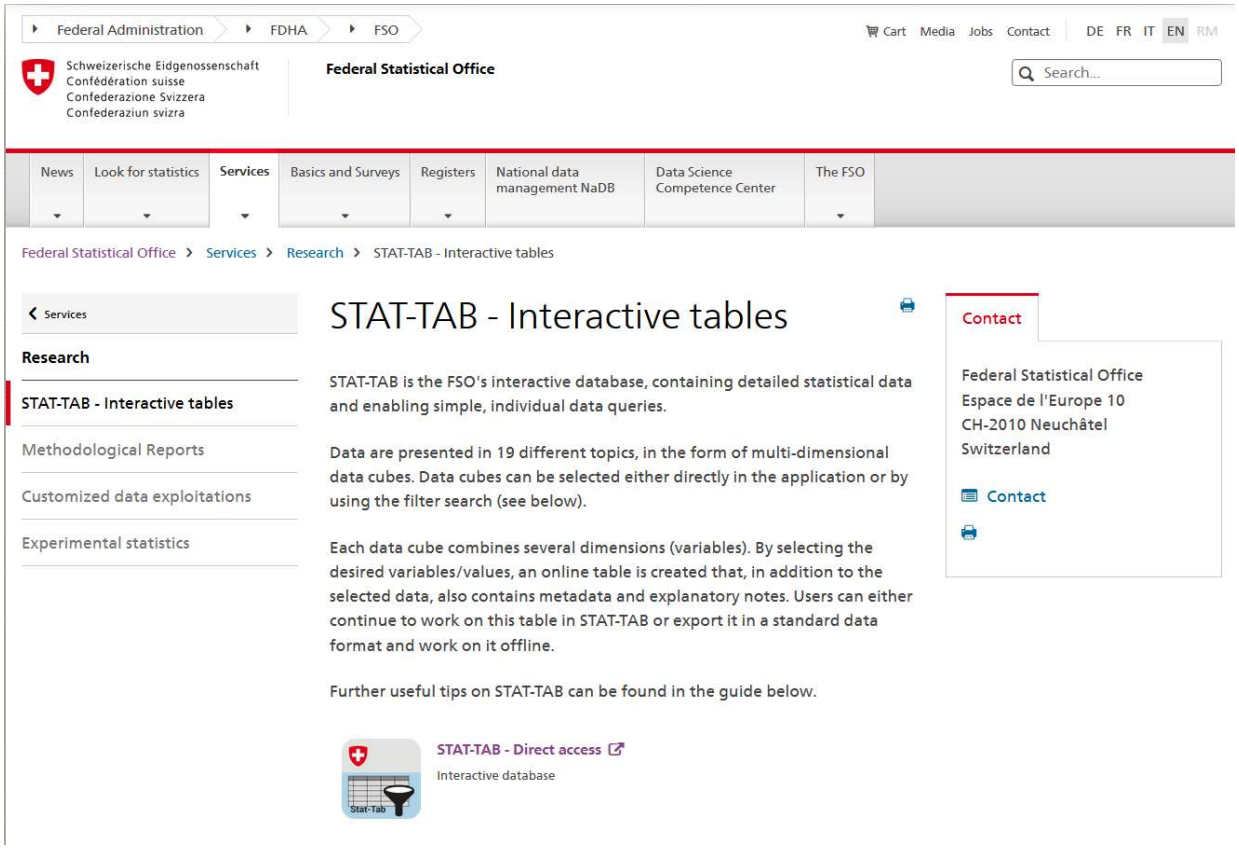
Motivations behind the BFS package

- 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



https://felixluginbuhl.com/BFS

Exploring the data catalog with BFS

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

https://felixluginbuhl.com/BFS

Exploring the data catalog with BFS

- Get available datasets (from official BFS RSS feed)

```r
library(BFS) #install.packages("BFS")

# get the data catalogue
catalog_data_en <- bfs_get_catalog_data(language = "en")

catalog_data_en
```

```
## # A tibble: 179 × 5
##    title                                                langu…¹ publi…² 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…
```

Exploring the data catalog with BFS

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

```
library(BFS) #install.packages("BFS")
library(dplyr) #install.packages("dplyr")

catalog_data_en <- bfs_get_catalog_data(language = "en")

# search for a recent dataset
catalog_data_uni <- catalog_data_en %>%
  filter(title == "University students by year, ISCED field, sex and level

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
```

Exploring the data catalog with BFS

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

```r
library(BFS) #install.packages("BFS")
library(dplyr) #install.packages("dplyr")

catalog_data_en <- bfs_get_catalog_data(language = "en")

catalog_data_uni <- catalog_data_en %>%
  filter(title == "University students by year, ISCED field, sex and level

# get the data
df_uni <- bfs_get_data(url_bfs = catalog_data_uni$url_bfs, language = "en")

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
```

```
##  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

```
library(BFS) #install.packages("BFS")
library(dplyr) #install.packages("dplyr")

catalog_data_en <- bfs_get_catalog_data(language = "en")

catalog_data_uni <- catalog_data_en %>%
  filter(title == "University students by year, ISCED field, sex and level

df_uni <- bfs_get_data(url_bfs = catalog_data_uni$url_bfs, language = "en")

# get data comments
comments <- bfs_get_data_comments(url_bfs = catalog_data_uni$url_bfs, langu

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 [illegible]\r\n- Postgraduate studies
```

4. All you need is

`bfs_get_data()`

All you need is `bfs_get_data()`

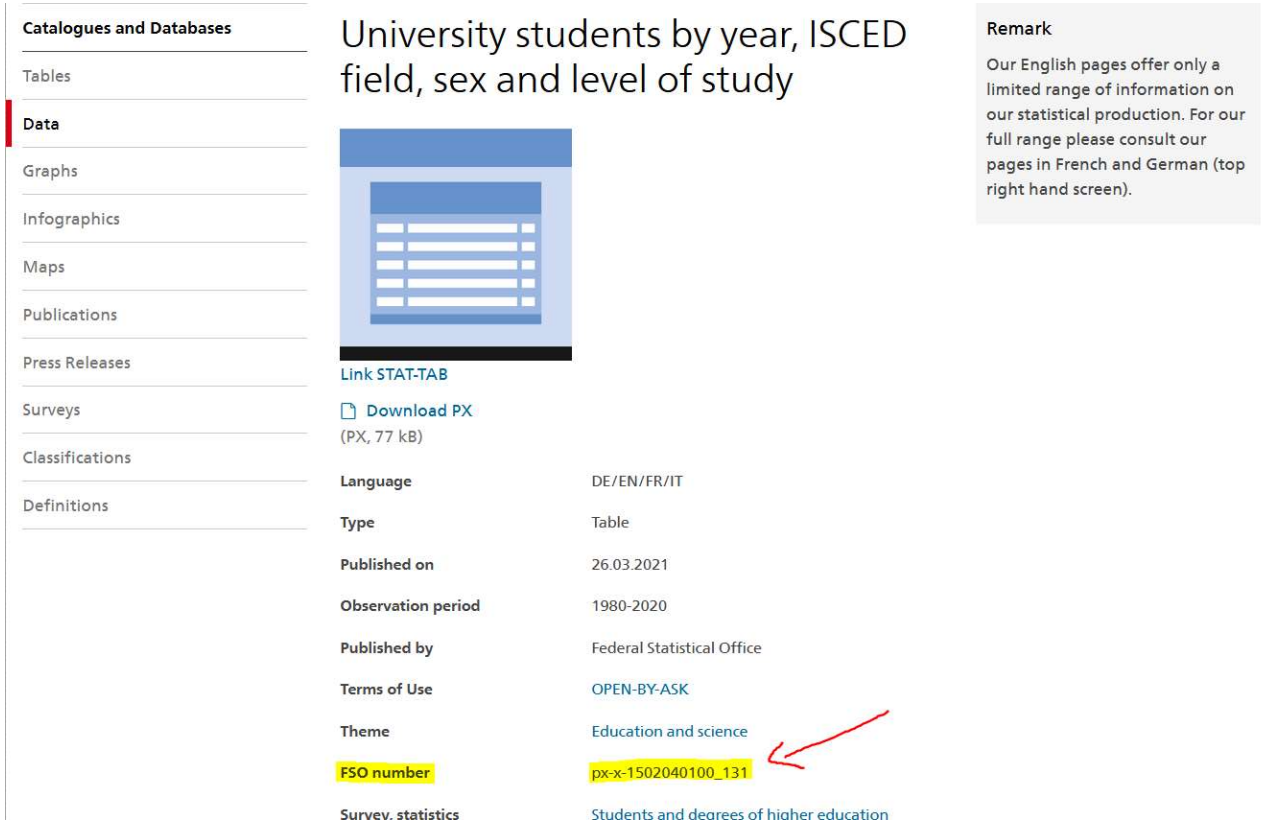
- Using `pxweb` R package functions under the hood to query the Swiss Federal Statistical Office PXWEB API.[1]

1. PXWEB is an API structure developed by Statistics Sweden and other national statistical institutions to disseminate public statistics in a structured way.

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
```



https://felixluginbuhl.com/BFS

All you need is `bfs_get_data()`

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

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

```
# 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`

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

```
# 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)

```
BFS::bfs_get_data(
  number_bfs = "px-x-1502040100_131",
  language = "de",
  clean_names = TRUE
)
```

```
# 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: ¹studierende_an_den_universitaren_hochschulen
```

All you need is `bfs_get_data()`

- Query specific categories with `query`.

```
BFS::bfs_get_data(
  number_bfs = "px-x-1502040100_131",
  language = "de",
  clean_names = TRUE,
  query = "all"
)
```

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.

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

```
# 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
```

https://felixluginbuhl.com/BFS

All you need is `bfs_get_data()`

- Query the `code` variables and value types.

```
BFS::bfs_get_data(
  number_bfs = "px-x-1502040100_131",
  language = "en",
  query = list(
          "Jahr" = c("40", "41"),
          "ISCED Fach" = c("0"),
          "Geschlecht" = c("*"), # Use "*" to select all
          "Studienstufe" = c("2", "3")
          ),
  column_name_type = "code", # "text" by default
  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          2                              143
```

All you need is `bfs_get_data()`

- Documentation: www.felixluginbuhl.com/BFS
- Source code: www.github.com/lgnbhl/BFS

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

5. Questions

Thank you for your attention!

- BFS documentation: felix.luginbuhl.com/BFS
- Swiss City Statistics app: city.felixluginbuhl.com
- LinkedIn: linkedin.com/in/felixluginbuhl/