Nils Sommer
Sahlistrasse 41, 3012 Bern
nils.sommer@unibe.ch

Data Science Project

SatisMap: Mapping Well-Being in Switzerland

Conceptual Design Report

30 October 2021

Abstract

The project SatisMap is an interactive web application mapping and displaying various measures of well-being in Switzerland. With the data from the Swiss Household Panel [1], the application will be able to display average levels of well-being in psychological, sociological, and medical domains per Canton over the years 1999 to 2020. Moreover, SatisMap will be able to calculate predictions of wellbeing in Switzerland. This report gives an overview of the data and data models for this project and shows promising preliminary development results.

Table of Contents

1 Project Objectives

The goal of this project is to create a web application that lets users interactively view a map of well-being of adults in Switzerland over the years 1999 – 2020 (SatisMap). The application will be able to describe and show various forms of well-being from psychological (e.g., self-reported life satisfaction, affect, or self-efficacy), sociological (e.g., education, income, work status), and medical (e.g., overall health status, eating habits, smoking) domains. On top of descriptively showing average values of these well-being measures per year and canton, the application will be able to variably display associations of well-being with possible predictors and / or criteria of well-being (e.g., political orientation or well-being measures from other domains). Aside from its main purpose of displaying various informative maps on the topic of well-being in Switzerland, the application will be able to calculate selected predictions of well-being in cross-sectional and longitudinal analyses.

The data used for all these analyses stems from the Swiss Household Panel (SHP) [1]. The SHP is an ongoing panel study that investigates households and individuals since 1998.

The web application SatisMap will be an important asset to get an overview of different domains of well-being in Switzerland. Governments (federal, cantonal, and municipal) will be able to find starting points to initiate actions to promote well-being. SatisMap will help researchers to explore well-being and to build their hypotheses. And finally, SatisMap will be an interesting application to the general populace interested in well-being in Switzerland.

2 Methods

SatisMap will be created using R software [2] as well as several specific R-packages. To name two of high importance, the packages tmap [3] (used to create maps in R) and shiny [4] (used to create web applications in R) will be used. Infrastructure needs for this project will be limited to personal computers. Webhosting of the final application will be carried by Shinyapps [5], a basic but scalable hosting service for Shiny applications.

The following statistical analyses will be conducted during this project and in the final application. To display overall associations, Pearson-Correlations will be used. To predict well-being, multiple linear regressions (cross-sectional analyses, continuous outcomes), multi-level-regressions (longitudinal analyses, continuous outcomes) and logistic regressions (cross-sectional, binary outcomes) will be employed.

3 Data

The data used for this project stems from the SHP [1], retrieved from SWISSUbase [6]. The SHP is "is a household panel study that follows a random sample of households resident in Switzerland over time" and aims to "observe social change, in particular, the dynamics of changing living conditions and social representations in the population of Switzerland" (SHP-Userguide, Page 5) [7]. In the SHP, representative households and their members are regularly interviewed about various topics concerning their daily lives. This includes several measures of the domains of well-being stated above, but also extensive data on demographics, work, finances, and so forth. Currently, 22 Waves of interviews are provided, conducted in consecutive years from 1999 to 2020. In the SHP, the relevant data is divided into 2 master files (describing households and individuals) and 44 annual files (one household and one individual file per wave) [7].

In figures 1 to 3, exemplary data is displayed from the most recent wave of interviews conducted in the year 2020. Only adults (16 years or older) are included. In Figure 3, one exemplary measure of well-being, life satisfaction, is depicted.
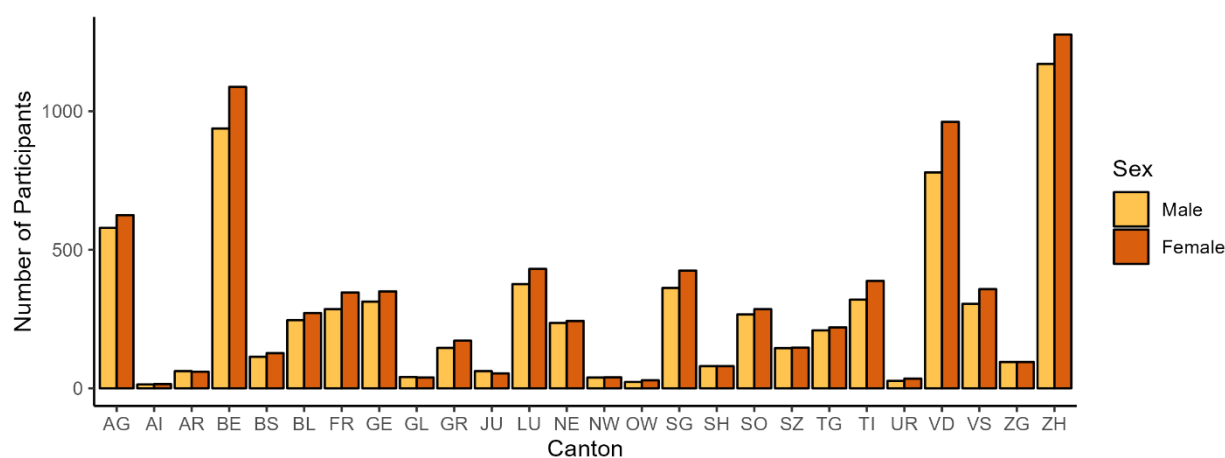

Figure 1. Participants by canton and sex.
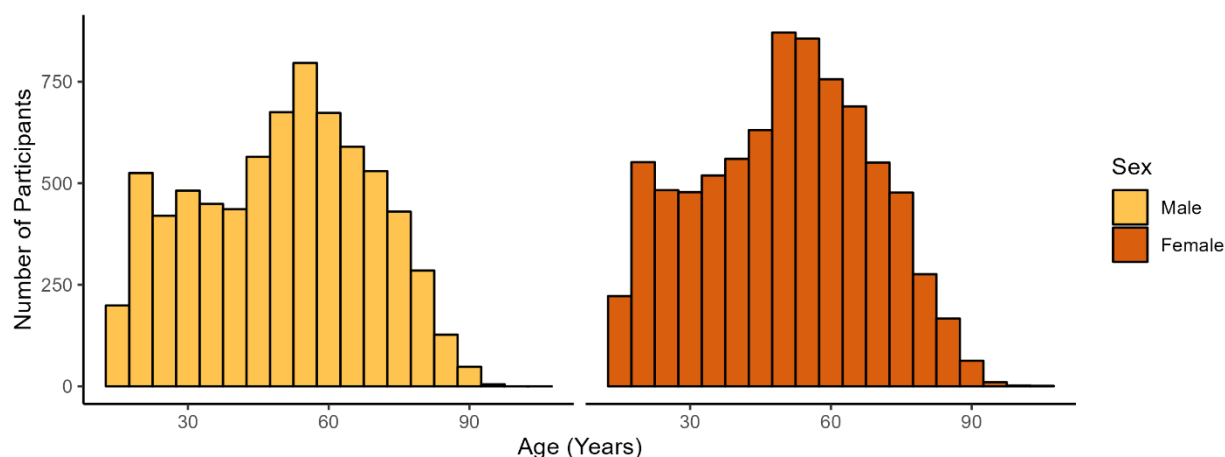SHP wave 22 (2020), n = 15399.

Figure 2. Participants by age and sex.
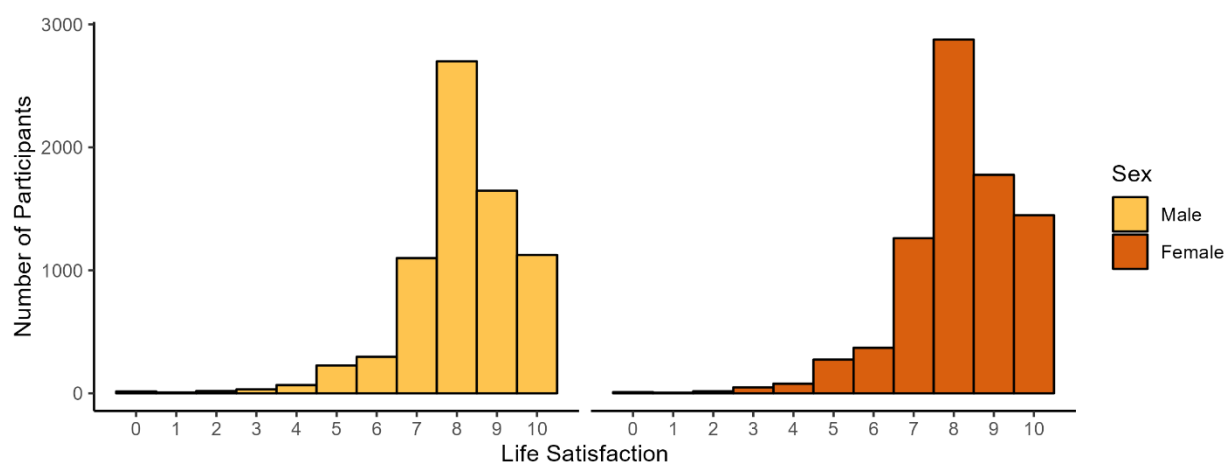SHP wave 22 (2020), n = 15399.

Figure 3. Life satisfaction by sex.
SHP wave 22 (2020), n = 15399. Life satisfaction measured globally with one question: 'In general, how satisfied are you with your life if 0 means "not at all satisfied" and 10 means "completely satisfied"?'

In accordance with the license agreement with the SHP [1], the data is not to be transmitted to third parties and was therefore stored exclusively on a personal, password protected cloud. No personal information of the participants of the SHP are included in the data and therefore no such data will be included in the final application. Aside from this, no security issues arise.

To create maps in R, some basic geospatial data of Switzerland and its cantons is required. This data was retrieved from the Website of the Federal Statistical Office [8]

4 Metadata

Metadata in great detail (such as in the SHP Userguide [7]) concerning the SHP-data is available at SWISSUbase [7] free of charge, however requiring a license agreement to access it.

Metadata concerning calculations and programming of the web application will be stored in a GitHub repository [9] and will be publicly available in the form of R-scripts.

5 Data Quality

To achieve the goals of this project, data needs to be representative for the Swiss Population. Aside from a representative initial Sampling the SHP provides mathematical weights to adjust results for low response rates and high attrition rates [7]. The SHP also conducts primary data cleaning before releasing the data [7].

6 Data Flow

Figure 4 displays the data flow for this project. Data was initially retrieved from SWISSUbase (SHP data) [6] and the Federal Statistical Office (geodata) [8] and was then stored in a personal cloud. These data files will be combined to a data frame in R. From there, the web application will be created, and the data will be analyzed to be displayed as maps and simplified regression outputs. Users can then manipulate the options in the web application and renew the analyses according to their input.
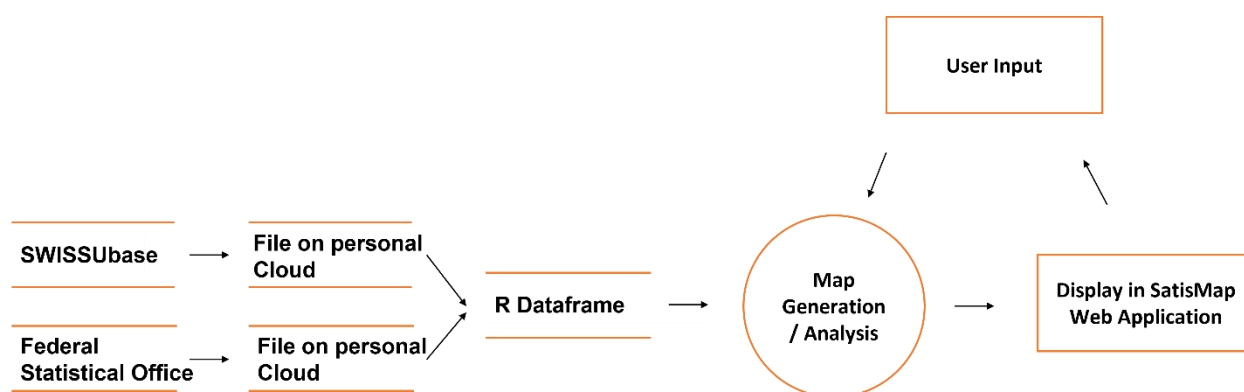

Figure 4. Data flow of the project SatisMap.

7 Data Model

In the following, preliminary designs for the data models at the conceptual, the logical, and the physical level are presented. The conceptual data model is briefly outlined in figure 5. User input in the web application will trigger the R script to gather relevant data from the SHP files and the geodata file in a combined R data frame. R will then calculate the requested analysis and will return it to the web application to be displayed.
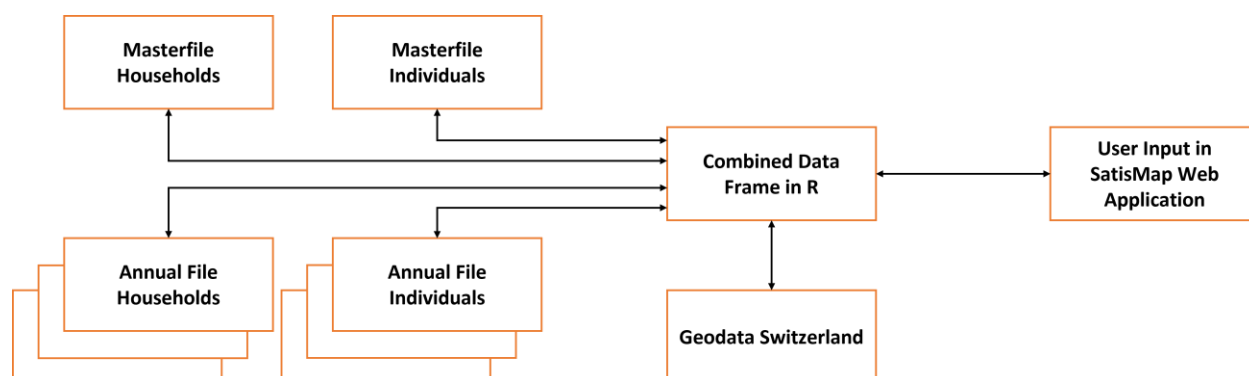

Figure 5. Conceptual data model of the project SatisMap.

Figure 6 contains a brief exemplary visualization of the logical and physical data model (combined). On the logical level, the exemplary request of a correlation of life satisfaction and political orientation by canton (with effects of age and sex partialled out) will trigger the r script to gather the variables sex, age, canton, canton name, canton geodata (for display), life satisfaction, and political orientation and store them into a combined R data frame. It then calculates the correlations per canton and returns the values to the web application.

On the physical level, the same process proceeds as follows: The request made in the web browser of the user is transmitted to the R script on the personal computer of the developer. The R script running on this computer then gathers the relevant information from the personal cloud storage into the R data frame on the personal computer (see key variables in bold in figure 6). The R script calculates the requested analyses and returns them to the web browser of the user.
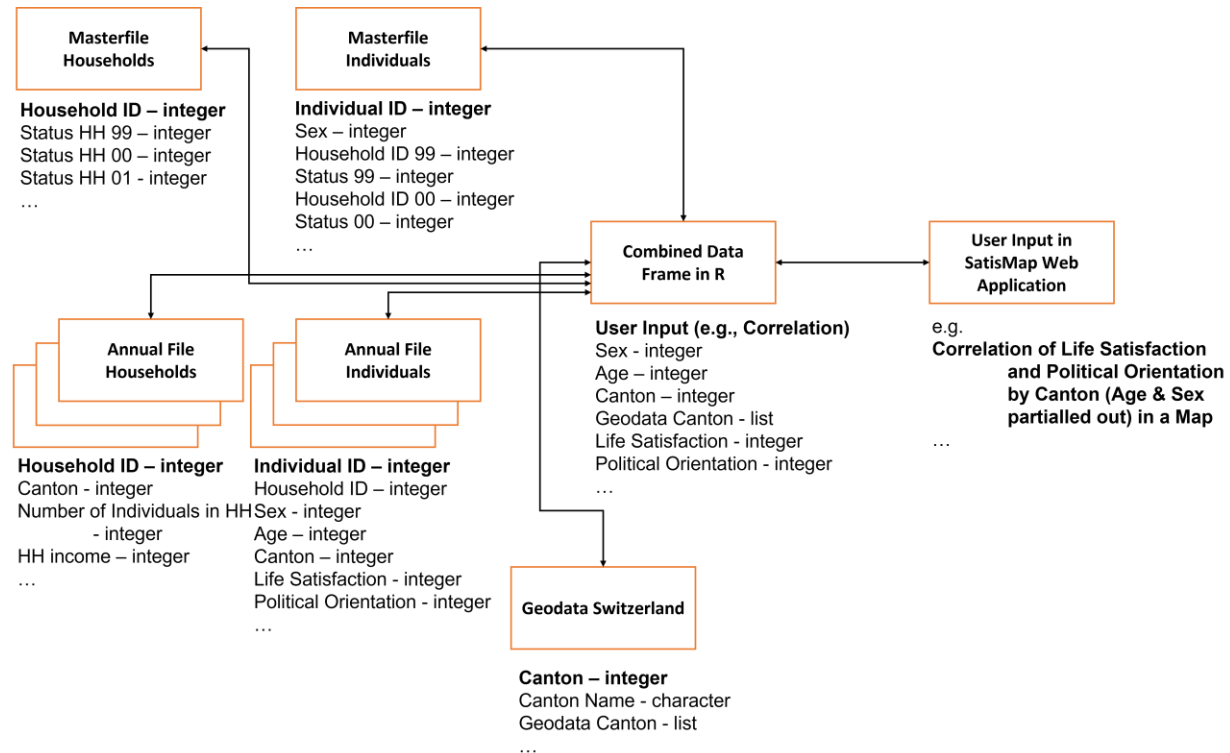
Figure 6. Logical and physical data model of the project SatisMap.

8 Risks

There are no foreseeable risks that could decrease output quality, delay project finalization, or increase project cost.

9 Preliminary Studies

Preliminary examples for maps depicting well-being in Switzerland using the SHP data and the R-package tmap [3] are displayed below. Guidance for this preliminary programming was found in the book "Geocomputation with R" by Robin Lovelace and colleagues [10].

Figures 7 and 8 depict static maps of life satisfaction in the Swiss cantons for men and women, respectively. Figure 9 depicts a screenshot of an interactive map, based on the static map in figure 7 (life satisfaction of women). It is visually evident that satisfaction levels differ to a high amount between cantons and between men and women (see for example cantons Uri and Nidwalden).
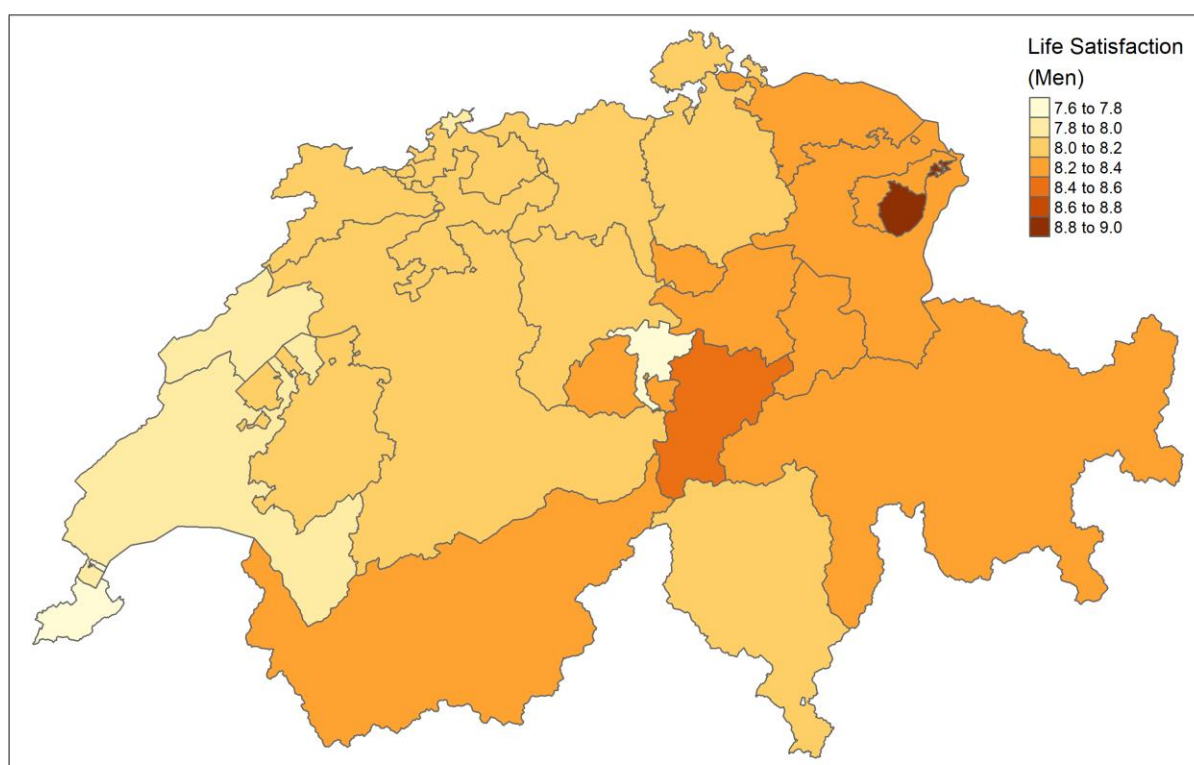

Figure 7. Life satisfaction per canton (men).
SHP wave 22 (2020), n = 7235. Life satisfaction measured globally with one question: 'In general, how satisfied are you with your life if 0 means "not at all satisfied" and 10 means "completely satisfied"?'
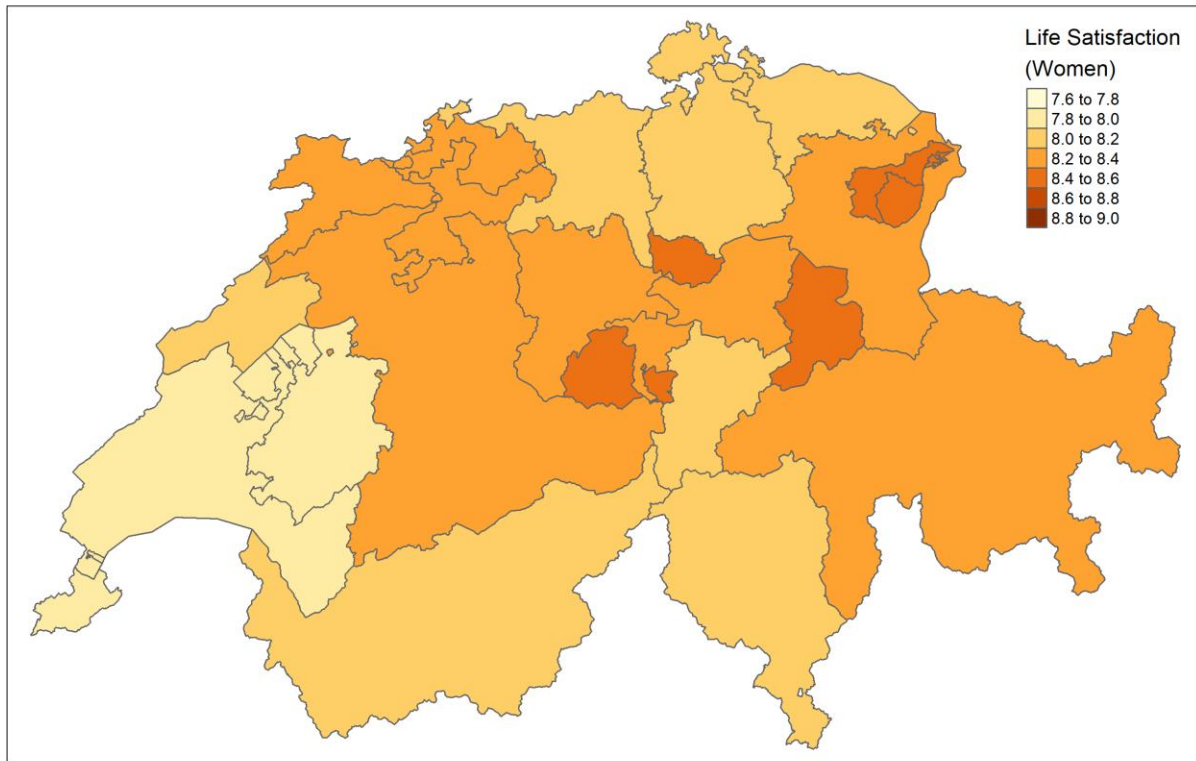
Figure 8. Life satisfaction per canton (women).

SHP wave 22 (2020), n = 8164. Life satisfaction measured globally with one question: 'In general, how satisfied are you with your life if 0 means "not at all satisfied" and 10 means "completely satisfied"?'
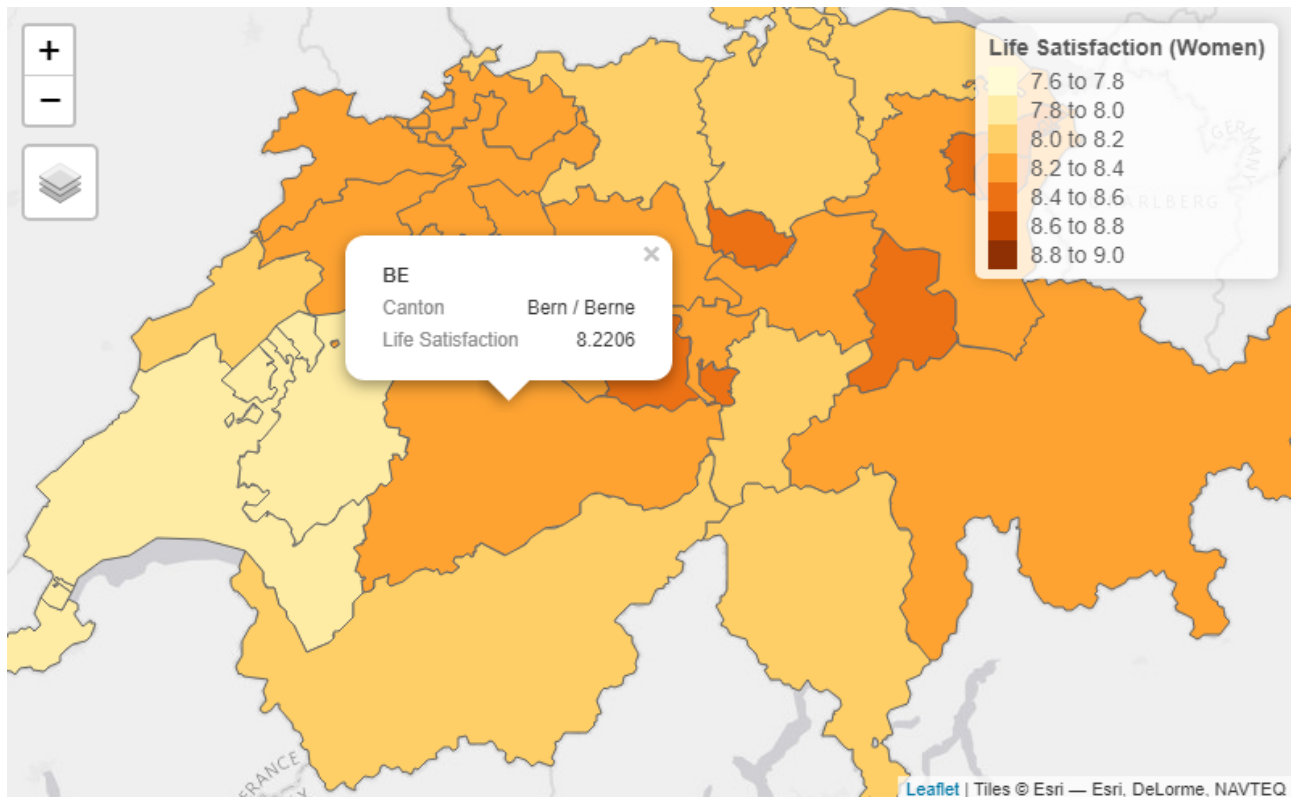
Figure 9. Life satisfaction per canton (women).
Screenshot taken from the interactive version of the map in figure 7. Created with tmap [3].

10 Conclusions

The development of SatisMap is still at an early stage, however, the outlined project plan and the preliminary results and graphs look promising. Satismap will be an important asset to explore well-being in Switzerland for governments, researchers, and the general population.

References and Bibliography

[1] Swiss Household Panel; Lausanne: FORS (https://forscenter.ch/projekte/swiss-household-panel/)

[2] The R Project for Statistical Computing (https://www.r-project.org/)

[3] M. Tennekes; tmap: Thematic maps in R; Journal of Statistical Software, 84(6), 1–39; 2018 (doi:10.18637/jss.v084.i06.; https://r-tmap.github.io/tmap/)

[4] Shiny; Rstudio (https://shiny.rstudio.com/)

[5] Shinyapps; Rstudio (https://www.shinyapps.io/)

[6] SWISSUbase; Lausanne: FORS (https://www.swissubase.ch/en/)

[7] M. Vorpostel et al.; Swiss Household Panel Userguide (1999-2020), Wave 22, January 2022; Lausanne: FORS; 2021

[8] Generalisierte Gemeindegrenzen: Geodaten; Federal Statistical Office, Switzerland (https://www.bfs.admin.ch/bfs/en/home/services/geostat/swiss-federal-statistics-geodata/administrative-boundaries/generalized-boundaries-local-regional-authorities.assetdetail.22484210.html)

[9] GitHub (https://github.com/)

[10] R. Lovelace et al.; Geocomputation with R; 2022 (https://geocompr.robinlovelace.net/references.html)