# Micro location rating
# R User Group Zurich

Jacqueline Schweizer
**Zurich, 16th of May 2018**

wüestpartner

# Location – the magic word in the real estate world

# Location, location, location

Accessibility?

Views?

Surroundings?

Nuisances?

Neighbours?

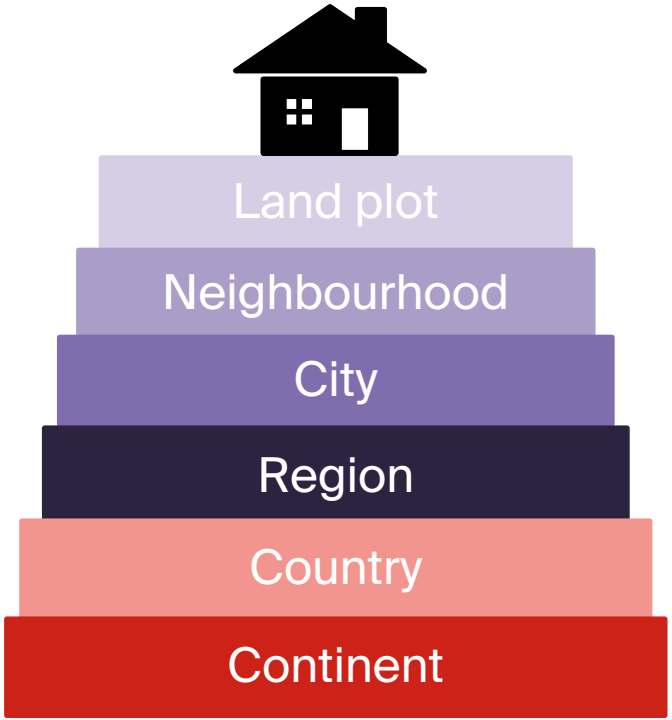
Sun exposure?

«Jiutian International Plaza» in Zhuzhou (central chinese province Huna)

# Location, location, location

«Jiutian International Plaza» in Zhuzhou (central chinese province Huna)

# Location




Noise pollution, slope, exposition, sunlight, nuisances, ...

Day-to-day errands, green space, schools, accessibility, ...

Infrastructure, administrative institutions, ...

Language, topography, tax level, work places, education, ...

Political system, legislative system, part of market spaces, ..

Time zone, climate zone, natural hazards, ...

# Location in real estate valuation

In real estate valuation there is usually two levels of location. Those two levels are relevant in determining the value of a house or an apartment: the micro and the macro location parameter.

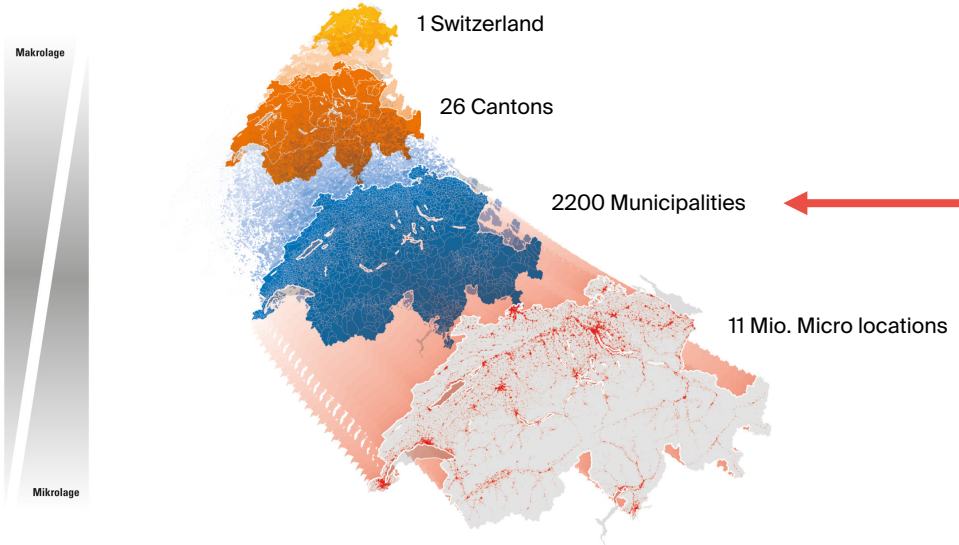
**Micro location**

Noise pollution, slope, exposition, sunlight, nuisances, Day-to-day errands, green space, schools, accessibility, infrastructure ...

**Macro location**

Tax level, accessibility of urban areas, work places, education, accessibility via road, rail, air, ...

# Macro location - Switzerland




In the real estate industry the macro location was established to differentiate on a **municipality level**.

The macro location serves to identify the **rough price level** of a house/ apartment.

# Macro location

A house has a different price regarding its general location (macro location). The exact same house in Meilen (gold coast) is significantly more expensive than it would be in Rorschach by Bodensee even though both municipalities border a big lake.

In real estate valuation, the macro location is determined relatively easy: by knowing in which municipality it is located.

The macro location explains a big chunk of the price of a house or an apartment.

Variance Partition Coefficient (VPC): 23-40% (explained part of total variance)

# Micro location

How beneficial is the location of a house within the macro location? How „good“ or „bad“ is the house located in comparison to all other possible locations within the municipality (macro location)?

→ relative conception of location quality resulting in **relative rating system**

**How did it work in the past?**

- Sight visitation → very time consuming
- Individually done by property valuer → very subjective, relative concept can only really be applied if the valuer knows **all** the other available locations within the municipality (macro location)
- Determination of micro location quality is related to a relatively high amount of effort

**What do we want to achieve?**

- Cost and time efficient estimation
- Objective evaluation of the measurable variables determining location quality
- Use the widely available GIS data
- → Developing an automated GIS based model to establish an objectively derived rating

# What are people searching for?

NZZ and Wüest Partner collaborate on an annual survey regarding relevant and irrelevant criteria when Swiss people are looking for a new apartment. The most recent survey found:

Most relevant criteria:

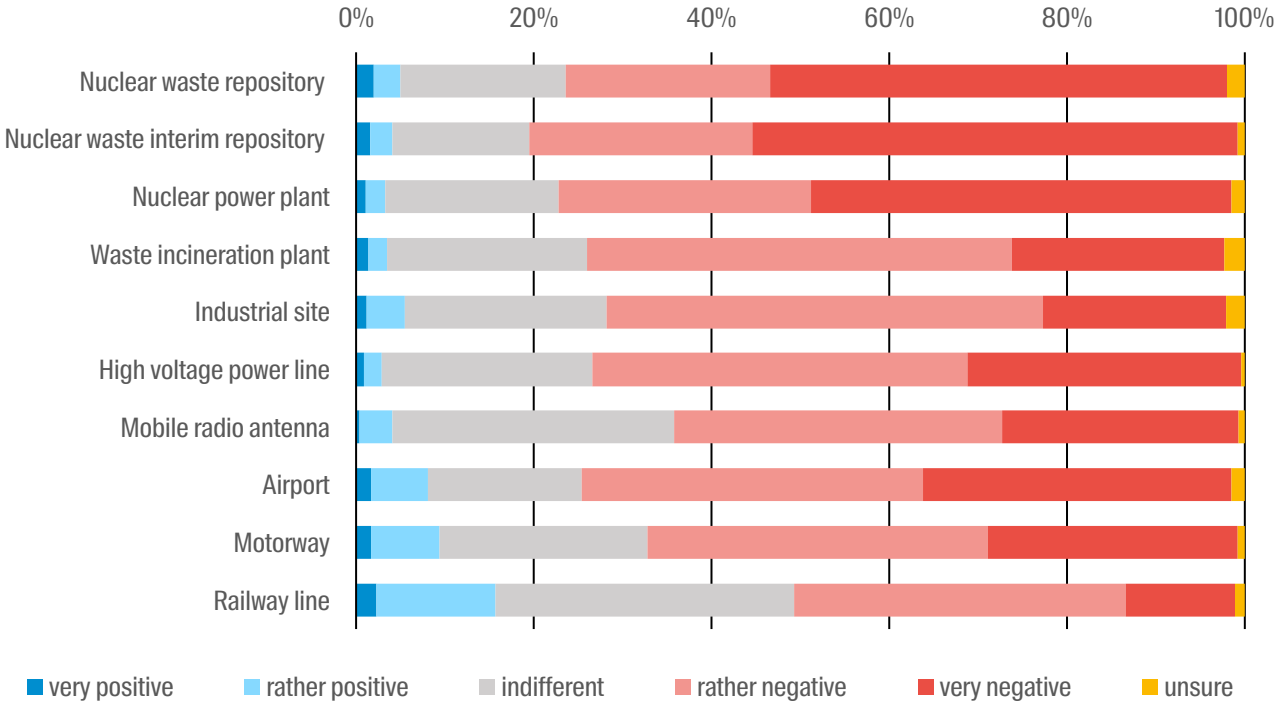
- Access to public transport
- Possibilities for day-to-day shopping
- Commute
- Noise pollution
- Green space

Least relevant criteria:

- Neighbours
- Supply of cultural infrastructure
- Child friendliness

*Immo-Barometer-Studie 2017*

# People‘s reaction to certain infrastructure by their house...




*Immo-Barometer-Studie 2016*

# Method

- Empirically based model
- Hedonic approach: log linear multiple regression
- Empirical data: real estate adverts on platforms like homegate, immoscout, newhome etc.
  - No detailed information about object qualities
  - But highly dense data base across the whole of Switzerland

$$\ln \ Transaction \ \ price_i = \beta_0 + \sum_{k=1}^{K} \beta_{Obj_k} Object_{ki} + \sum_{t=1}^{T} \beta_{Macro_t} Macro \ \ location_{ti} + \sum_{u=1}^{U} \beta_{Micro_u} Micro \ \ location_{ui} + \varepsilon_i$$

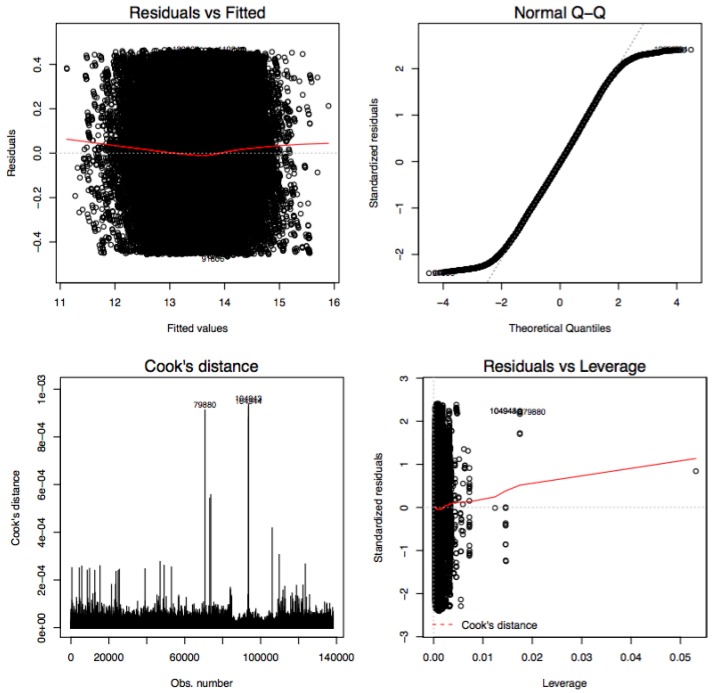
Object qualities e.g.:
- Living space
- Number of rooms
- Year of Construction
- New built
- ...

Macro location e.g.:
- Tax level
- Accessibility
- Infrastructure
- Commodities
- ...

Micro location e.g.:
- Traffic noise
- Railway noise
- Lake view
- Public transport
- Centrality
- ...

# Method

```
reg.lm.efh              <- function(dat) {
                            lm(log(vpreis)                  ~
                              log(makro1)                   +
                              log(makro2)                   +
                              makro3                        +
                              makro4                        +
                              as.factor(makro5)             +
                              as.factor(makro6)             +
                              log(obj1)                     +
                              log(obj2)                     +
                              as.factor(obj3)               +
                              obj4                          +
                              I(obj4^2)                     +
                              as.factor(obj5)               +
                              log(mikro1)                   +
                              log(mikro2)                   +
                              mikro3                        +
                              mikro4                        +
                              mikro5                        +
                              log(mikro6)                   +
                              log(mikro7)                   +
                              log(mikro8)                   +
                              mikro9                        +
                              I(1/(mikro10))                +
                              I(1/(mikro11))                +
                              I(1/(mikro12))                +
                              mikro13                       +
                              I(mikro13^2)                  +
                              as.factor(mikro14)            +
                              as.factor(mikro15)             ,
                              data = dat, na.action=na.exclude,
                              x = T,
                              y = T)
}
```




```
Residual standard error: 0.1939 on 138056 degrees of freedom
Multiple R-squared:  0.8773,   Adjusted R-squared:  0.8772
F-statistic:  9470 on 94 and 138056 DF,  p-value: < 2.2e-16
```

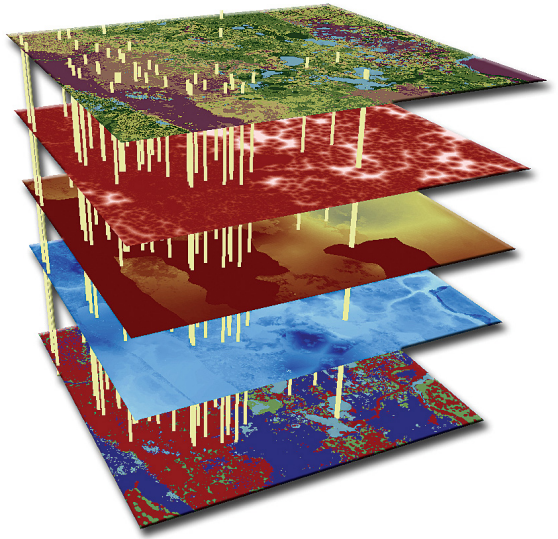
# GIS Data

- All layers are gridded or are being gridded

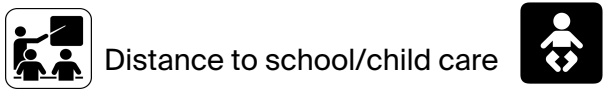
→ 25x25m grid cells
→ 66 Mio. grid cells for Switzerland!
→ Limit the scope to the settlement area plus some additional buildings outside of this area
→ 11 Mio. grid cells are being rated with the micro location rating

- GIS based micro location rating for Switzerland:
  - Each cell contains a value for every variable
  - Floating, classified and binary variables

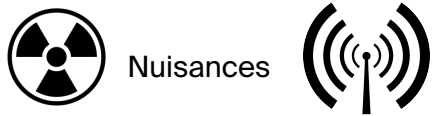
- Price prognosis calculated through the regression model and the designated values per cell



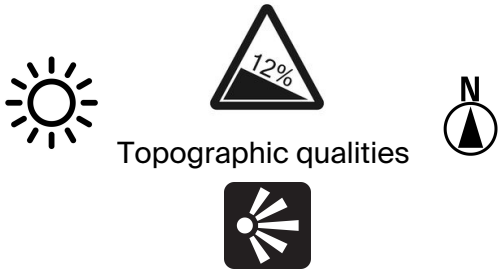
# Variable groups



Distance to school/child care



Nuisances



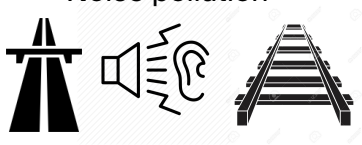

Topographic qualities

Natural features

Distance to centre and other infrastructure

Distance to public transport

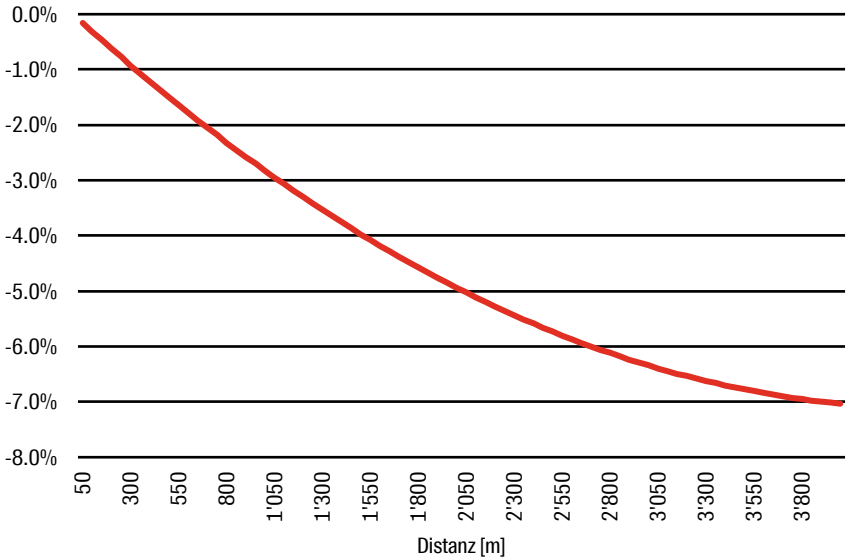
Noise pollution

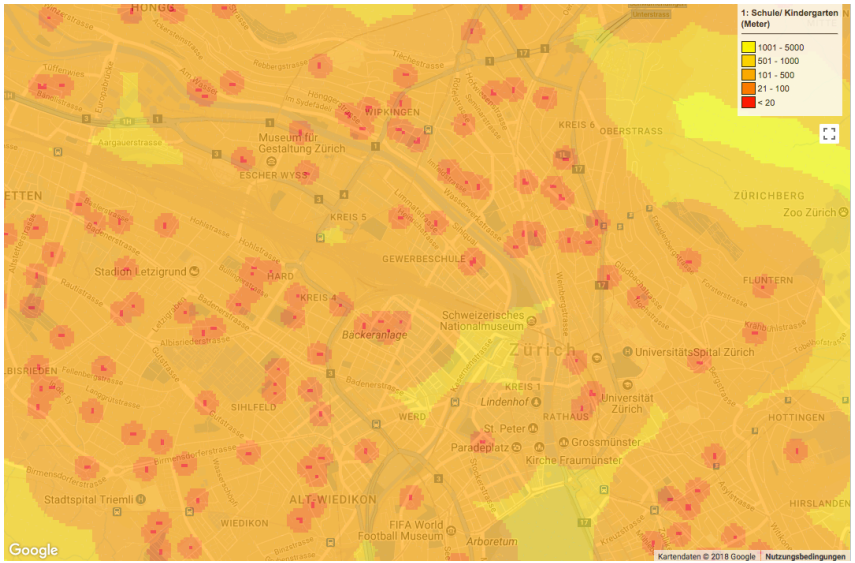


# Distance to primary school

Price effect in the residential rental market:

- 1‘000m approx. -2.8%
- 3‘000m approx. -6.5%

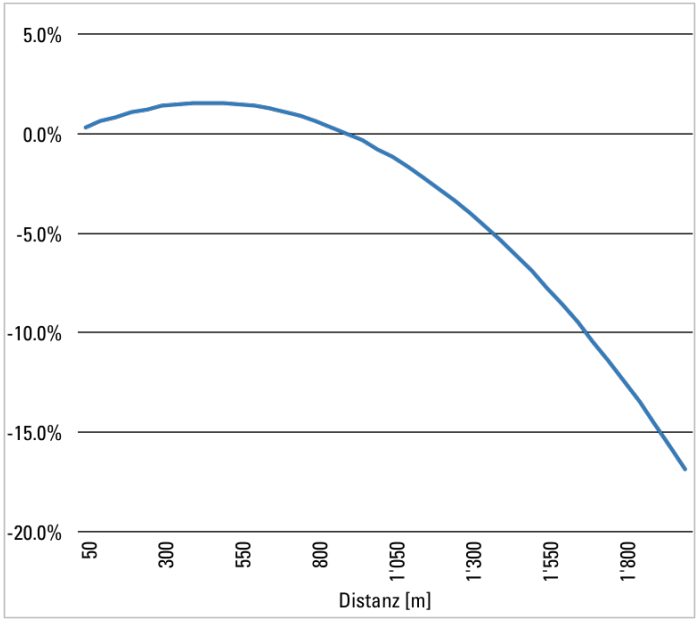


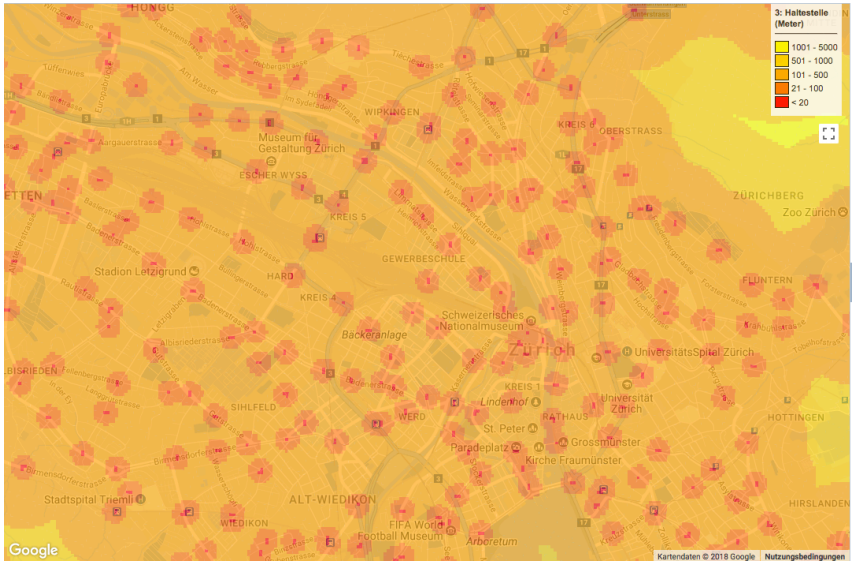



# Distance to public transport

Price effect in the residential sales market:

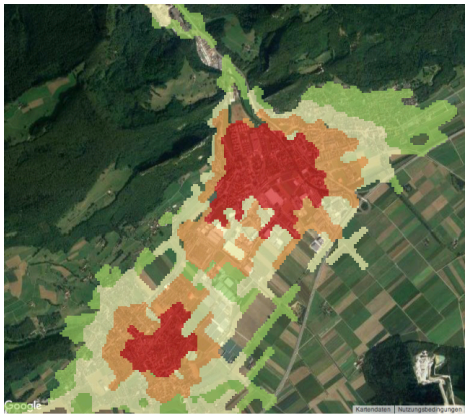
- 400m approx. +1.6%
- 1‘500m approx.  -7%



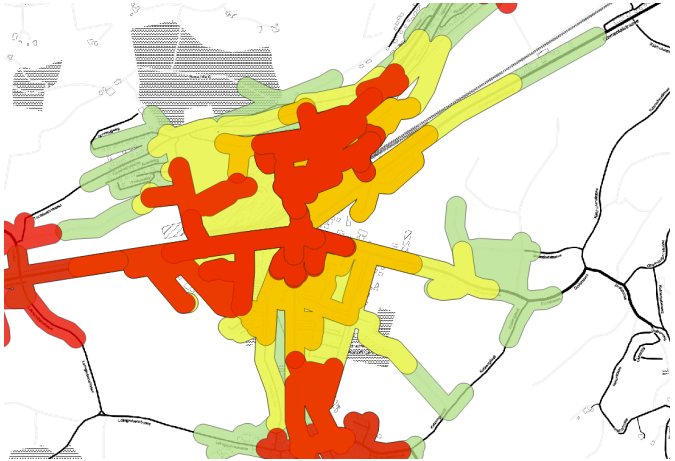
# ÖV-Güteklasse – public transportation quality

- Combination of distance, frequency and mean of transportation → public transportation quality published and calculated by the federal office of spatial planning
- Instead of the Euclidean distance, we calculated the actual walking distance and thereby developed the model further

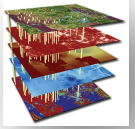
# ÖV-Güteklasse – public transportation quality

```r
get_poly <- function(id, lat, lon, modus, messung, ring) {

    options(warn=-1)
    # Request gemäss Defintion Route360 zusammenstiefeln: benutzte Parameter:
        # Koordinaten (WGS84),
        # modus ("bike", "walk", "transit", "car")
        # speed: wählbar oder per Default: bikeSpeed = 15.0 (bikeUphill = 20.0; bikeDownhill = -10.0), walkSpeed = 5.0 (walkUphill = 10.0; walkDownhill = 0.0)
        # messung: distance or time
        # ring: einzelne Zahl oder vector (100, 300, 400)

    connectRequest          <-    paste(    "https://service.route360.net/switzerland/v1/polygon?cfg={sources:[{lat:",
                                            lat,
                                            ",lng:",
                                            lon,
                                            ",id:source0,tm:{",
                                            modus,
                                            ":{walkSpeed:",
                                            speed,
                                            "}}},elevation:true,reverse:false,edgeWeight:",
                                            messung,
                                            ",polygon:{values:[",
                                            ring,
                                            "],srid:21781,buffer:1,intersectionMode:union,serializer:geojson,pointReduction:true,minPolygonHoleSize:10000000,}}",
                                            sep="")

    request                 <- GET(connectRequest)
        if (request$headers[[2]] == "application/json" ) {
            data_json               <- fromJSON(content(request, "text", encoding="UTF8"))
            options(warn=0)

            # Anzahl Polygone, die zu extrahieren sind
            n.poly      <- length(data_json$data$features)
            list.poly   <- list()

            for ( i in 1:n.poly) {

            t.length            <- length(data_json$data$features[[i]]$geometry$coordinates[[1]][[1]])
                if (t.length == 0) {
                t.xy                <- data.frame(lat84=lat, lon84=lon )
                t.xy$X              <- WGS.to.CH.x(t.xy$lat84, t.xy$lon84)
                t.xy$Y              <- WGS.to.CH.y(t.xy$lat84, t.xy$lon84)
                coordinates(t.xy)   <- ~ X + Y
                d.sp                <- gBuffer(t.xy, width=1)
                d.sp@polygons[[1]]@ID <- as.character(i)
                list.poly[[i]]      <-  d.sp

        } else {

            t.xy                <- matrix(unlist(data_json$data$features[[i]]$geometry$coordinates), nrow=t.length, ncol=2, byrow=TRUE)
            t.xy                <- data.frame(t.xy)
            colnames(t.xy)      <- c("X", "Y")
            coordinates(t.xy)   <- ~ X+Y
            t.poly              <- gBuffer(t.xy, width=1)
            t.poly@polygons[[1]]@ID <- as.character(i)
            list.poly[[i]]      <- t.poly
            }

    t.join.poly             <- SpatialPolygons(lapply(list.poly, function(x){x@polygons[[1]]}))

    return(t.join.poly)

    }
```



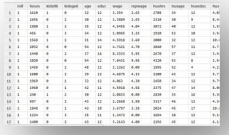
# Approach

- Data preparation: foreign, FNN, raster, maptools, rgeos, SDMTools, jsonlite (using webservices), adehabitatMA

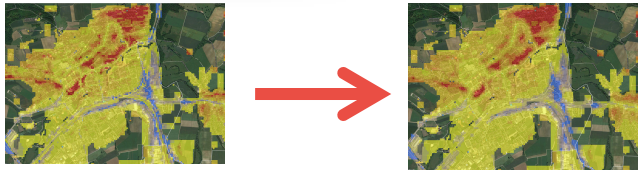


- Calculating regression model: dataframes instead of raster layers, everything stored and exported as tables

- Prognosis and relative rating: dataframes, prognosis and rating exported in raster format (ASCII)



- Smoothing and mapping: ArcGIS



```
# EWG
# Ergänzen der Objekteigenschaften für EWG und MWG
t.data.prog$flae                                <- 120
t.data.prog$zimmer                              <- 4
t.data.prog$alter.cl20                          <- 3

load(paste(c(p.output.mod.ins, today, "_ewg_inserat.rda"), collapse=""))


# Vorhersage mit Logarithmus-Korrektur
t.sum.lm.ewg                    <- summary(reg.ewg)
t.data.prog$PROG.EWG            <- exp( predict(reg.ewg, newdata=t.data.prog  ) + ( t.sum.lm.ewg$sigma ^ 2) / 2 )

# Check der Prognose (NAs: 689 )
summary(t.data.prog$PROG.EWG )
d.prognose.grid.ewg                             <- t.vorlage
colnames( d.prognose.grid.ewg@data )            <- "PROG.EWG"
d.prognose.grid.ewg@data$PROG.EWG               <- NA
d.prognose.grid.ewg@data$PROG.EWG[t.pos.prog]   <- t.data.prog$PROG.EWG

# ASCII rausschreiben für Klassierungsscript
write.asciigrid(d.prognose.grid.ewg, paste(c(p.output.ins, "02_Prognose/", p.prognose, today, "_prognose_EWG.asc"), collapse="
```

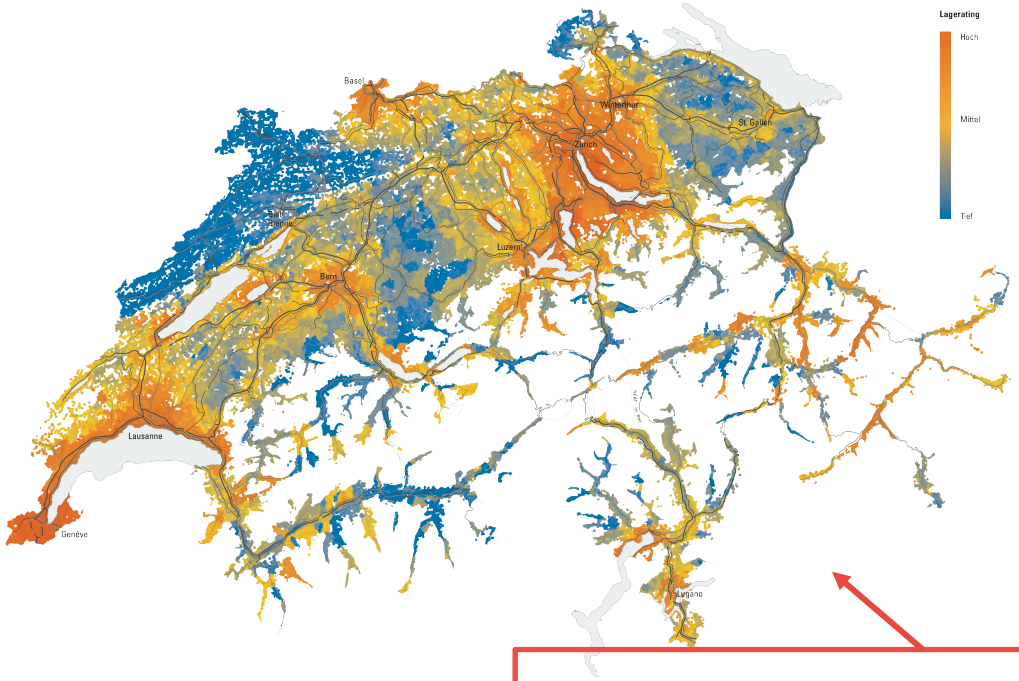
# Benefits and challenges of R in this project

- Combination of statistical methods and spatial data (GIS methods)
- Very heterogeneous data
- Efficient processing once the data is loaded

- Big file sizes to load into main memory
- Long calculation duration at times, looping is a no go

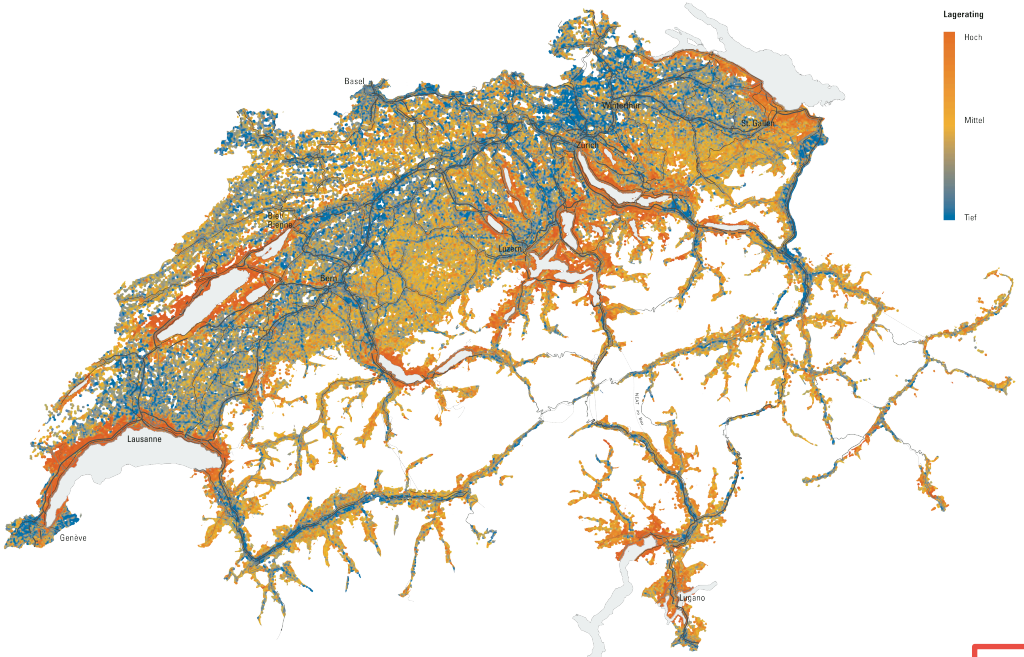
- Calculated on: Mac Pro 2013, 128 RAM, R on OS X El Capitan
- Final rating raster: approx. 700 MB per raster layer
- Prognosis table: 11 GB text files

# Price prognosis – over-all location




$$\ln\ Transaction\ price_i = \beta_0 + \sum_{k=1}^{K} \beta_{Obj_k} Object_{ki} + \sum_{t=1}^{T} \beta_{Macro_t} Macro\ location_{ti} + \sum_{u=1}^{U} \beta_{Micro_u} Micro\ location_{ui} + \varepsilon_i$$

# Price prognosis only regarding mirco location qualities




$$\ln Transaction\ price_i = \beta_0 + \sum_{k=1}^{K} \beta_{Obj_k} Object_{ki} + \sum_{t=1}^{T} \beta_{Macro_t} Macro\ location_{ti} + \sum_{u=1}^{U} \beta_{Micro_u} Micro\ location_{ui} + \varepsilon_i$$

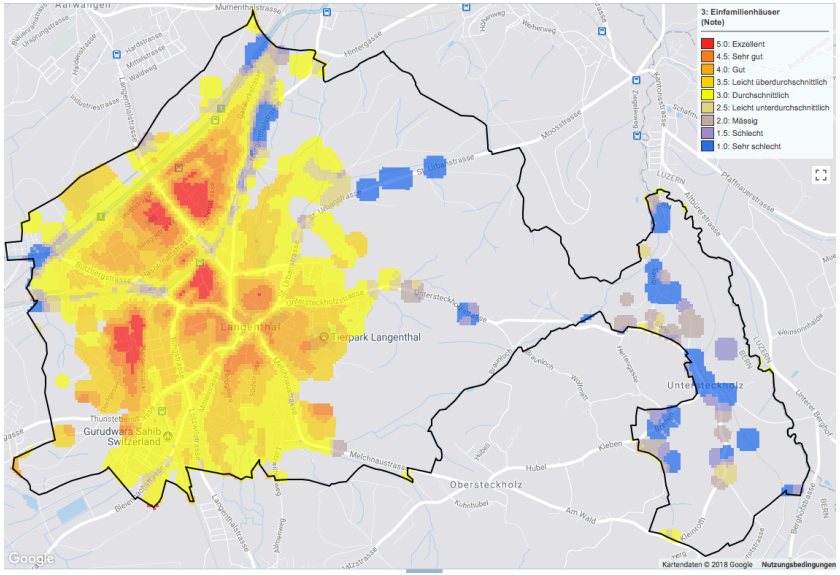
# Relative scoring system

- Price prognosis for every cell in Switzerland
  → absolute score

- The goal is to have a relative rating to rate the small scaled location quality within the municipality (a tranquil location ≠ high value in micro location rating)

- Rating scale from 1.0 (very bad) to 5.0 (excellent)
  → relative score

- Model product: in licensable WEB-GIS-Tool “GeoInfo” and “Wüest Dimensions” available as one of many Raster layers.



# Thank you!

At your disposal for further questions.

**Jacqueline Schweizer**
**T +41 44 289 90 16**
**jacqueline.schweizer@wuestpartner.com**

**Wüest Partner AG**
**Alte Börse**
**Bleicherweg 5**
**8001 Zürich**
**Schweiz**
**wuestpartner.com**

wüestpartner