

Schaffner, Sandra

**Research Report**

FDZ data description: Real-estate data for Germany Campus Files (RWI-GEO-RED city and RWI-GEO-RED cross) - Advertisements on the internet platform ImmobilienScout24 for teaching purposes

RWI Projektberichte

**Provided in Cooperation with:**

RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

*Suggested Citation:* Schaffner, Sandra (2020) : FDZ data description: Real-estate data for Germany Campus Files (RWI-GEO-RED city and RWI-GEO-RED cross) - Advertisements on the internet platform ImmobilienScout24 for teaching purposes, RWI Projektberichte, RWI - Leibniz-Institut für Wirtschaftsforschung, Essen

This Version is available at:

<https://hdl.handle.net/10419/242995>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



Datenbeschreibung

RWI – Leibniz-Institut für Wirtschaftsforschung

**FDZ Data description:  
Real-Estate Data for Germany Campus Files  
(RWI-GEO-RED city and RWI-GEO-RED cross)-  
Advertisements on the Internet Platform  
ImmobilienScout24 for teaching purposes**

**November 2020**

**Sandra Schaffner**

# Impressum

---

## Herausgeber:

RWI – Leibniz-Institut für Wirtschaftsforschung  
Hohenzollernstraße 1-3 | 45128 Essen, Germany

Postanschrift:

Postfach 10 30 54 | 45030 Essen, Germany

Fon: +49 201-81 49-0 | E-Mail: [rwi@rwi-essen.de](mailto:rwi@rwi-essen.de)  
[www.rwi-essen.de](http://www.rwi-essen.de)

## Vorstand

Prof. Dr. Dr. h. c. Christoph M. Schmidt (Präsident)

Prof. Dr. Thomas K. Bauer (Vizepräsident)

Dr. Stefan Rumpf

© RWI 2020

Der Nachdruck, auch auszugsweise, ist nur mit Genehmigung des RWI gestattet.

## RWI Datenbeschreibung

Schriftleitung: Prof. Dr. Dr. h. c. Christoph M. Schmidt

Gestaltung: Daniela Schwindt, Magdalena Franke, Claudia Lohkamp

## FDZ Data description:

Real-Estate Data for Germany Campus Files (RWI-GEO-RED city and RWI-GEO-RED cross)-Advertisements on the Internet Platform ImmobilienScout24 for teaching purposes

November 2020

Sandra Schaffner

# Datenbeschreibung

---

RWI – Leibniz-Institut für Wirtschaftsforschung

**FDZ Data description:  
Real-Estate Data for Germany Campus Files  
(RWI-GEO-RED city and RWI-GEO-RED cross)-  
Advertisements on the Internet Platform  
ImmobilienScout24 for teaching purposes**

**November 2020**

**Sandra Schaffner**

### List of contents

List of tables .....	4
Abstract .....	5
1. Introduction and Short Description .....	6
1.1 Introduction .....	6
1.2 Original data base .....	6
2. The Campus Files .....	7
2.1 City-Campus File .....	7
2.2 Cross-Sectional Campus File .....	8
2.3 Data preparation.....	10
3. Variable Description.....	10
3.1 Identifier .....	10
3.2 Time Period .....	11
3.3 Object Features.....	11
3.4 General Object Information.....	13
3.5 Area Information.....	14
3.6 Energy and Structure Information .....	15
3.7 Price Information .....	16
3.8 Regional Information .....	17
3.9 Meta-Information of Advertisement .....	17
3.10 Generated Technical Variables .....	18

## List of tables

Table 1	(Rough) Number of observations per city and year in the City Campus File.....	7
Table 2	Variables covered in the different campus files .....	8
Table 3	Number of municipalities in the Cross-Sectional Campus file 2019 that can be identified .	10
Table 4	Detailed Variable Descriptions of Identifiers .....	10
Table 5	Detailed Variable Descriptions of Variables Describing Time.....	11
Table 6	Detailed Variable Descriptions of Variables Describing Object Features .....	11
Table 7	Detailed Variable Descriptions of General Object Information .....	13
Table 8	Detailed Variable Descriptions of Area Information.....	14
Table 9	Detailed Variable Descriptions of Energy and Structure Information .....	15
Table 10	Detailed Variable Descriptions of Price Information .....	16
Table 11	Detailed Variable Descriptions of Regional Information .....	17
Table 12	Detailed Variable Descriptions of Variables Containing Meta-information of the Advertisement .....	17
Table 13	Detailed Variable Descriptions of Technical Variables .....	18

### **Abstract**

The FDZ Ruhr provides two campus files on real estate advertisements in Germany: the City Campus File (RWI-GEO-RED city) and a Cross-Sectional Campus file (RWI-GEO-RED cross). The datasets are extractions of the Scientific Use Files of RWI-GEO-RED. The City File covers the 15 biggest cities in Germany over the whole time period whereas the Cross File covers all Germany within one year (2019). The RWI-GEO-RED data are based on information from the internet platform ImmobilienScout24 and cover residential advertisements only. The campus files include apartments for sale and for rent and houses for sale. The data are available for lectures, tutorials, seminar thesis, bachelor and master thesis and scientific research. The provided dataset covers detailed regional information and a rich set of housing characteristics. Both datasets are a sample drawn and not the comprehensive data base like the Scientific Use File. The already implemented data cleaning eases the data work for students. This data report gives a brief overview of the data as well as its limitations and specifics. The data report is addressed to (potential) users of the data as support for their data preparation.

## 1. Introduction and Short Description

### 1.1 Introduction

The campus files of the RWI-GEO-RED data base on the respective Scientific Use Files. Scientific Use Files (SUFs) are available for scientific research on a contractual basis. However, due to increased demand for a campus file usable within teaching etc. the FDZ Ruhr and ImmobilienScout24 decided to release campus files. These files do not cover the whole database regarding variables and observations. Further, the regional information is restricted to municipalities and districts. Due to the process of sampling it is not possible to analyze the number of advertisements over time. However, the files are representative regarding all other characteristics especially prices etc. Further, the size of the data files is much smaller than the SUFs which makes it much easier to work with the data. There exist two types of data: a longitudinal version for the biggest cities in Germany as well as a cross-sectional version for all Germany.

This data description gives a brief overview on the data. Since the data is based on the SUFs we also recommend the data description of the SUFs for more information. This description covers the version RWI-GEO-RED city0719 and RWI-GEO-RED cross2019.

### 1.2 Original data base

The research data center (FDZ) Ruhr at the RWI provides a unique dataset on German real estate prices, obtained by ImmobilienScout24. The dataset entails information of real estate offerings on prices as well as on various observable characteristics that determine the value of a property. The present dataset covers the timespan from January 2007 to June 2020. The data of the campus files are organized into three separate datasets: houses for sale, apartments for sale, and apartments for rent.

ImmobilienScout24 is the largest internet platform on real estate in Germany. It gives real estate owners the opportunity to advertise their objects for a fee. The platform is open to private and commercial entities. Yet, the present dataset only includes residential real estate. It distinguishes between houses and apartments and features objects both for rent and for sale. ImmobilienScout24 has a self-reported market share of about 50% of all real estate objects offered for sale or rent in Germany (Georgi and Barkow 2010). Due to the high market share of ImmobilienScout24, the dataset consists of a substantial number of observations.

To advertise an object, the owner must fill out a questionnaire asking for different characteristics of the property, including the price. This means that the price at which the owner is willing to sell or rent out the object is an offering price. The advertised price is not binding, i.e. the data do not comprise transaction prices. The price information is available for almost all advertisements. In addition to the price, advertisers are free to include further object-specific characteristics. This helps to present an object adequately, and ideally, it increases the chance of selling at a favorable price or renting out an apartment respectively.



## 2. The Campus Files

There exist two different types of campus files: the City Panel File and the Cross-sectional Campus File

### 2.1 City-Campus File

The City Campus File covers the 15 biggest cities in Germany. These cities are Berlin, Hamburg, Munich, Cologne, Frankfurt, Stuttgart, Düsseldorf, Leipzig, Dortmund, Essen, Bremen, Dresden, Hannover, Nürnberg and Duisburg. All of these cities have at least almost 500 thousand residents (Duisburg had about 498 thousand residents in 2019). For each of the cities a large sample is drawn in a way that there is almost the same number of advertisements each year (ending date of the advertisement). The number of observations varies between 1 000 per city and year in the houses for sale sample up to 50 000 per year in Berlin and apartments for rent. Table 1 gives an overview of the number of observations per ending year and city. In most times, due to the randomization process it is not the exact number of observations. The sample drawn includes about 27% apartments for sale and about 37% houses for sale and apartments for rent. However, the samples only cover observations that fulfill some quality standards like less missing values, no outliers etc. A more detailed description of the sample drawn is described in 2.3.

The City Campus File is still in the original spell form. Each advertisement is surveyed as a spell with a beginning and ending date. It is possible that an advertisement is deactivated and reactivated again. Then a new spell starts. Also, changes made by the offeror (more detailed information, different price, ...) lead to a new spell generated by the system. Therefore, the data set allows the identification of price adjustments. Table 2 gives an overview on the variables in each data set.

Table 1

**(Rough) Number of observations per city and year in the City Campus File**

	houses for sale	apartments for sale	apartments for rent
Hamburg	5 000	4 000	15 000
Hannover	1 000	2 000	7 000
Bremen	2 000	2 000	7 000
Düsseldorf	1000	4 000	15 000
Essen	2 000	2 000	7 000
Duisburg	1 000	2 000	7 000
Cologne	2 000	4 000	15 000
Dortmund	2 000	2 000	7 000
Frankfurt	1 000	4 000	15 000
Stuttgart	1000	4 000	7 000
München	1000	10 000	15 000
Nurnberg	1 000	2 000	7 000
Berlin	5 000	10 000	50 000
Dresden	2 000	4 000	15 000
Leipzig	1000	2 000	15 000

*The exact number of observations vary due to the randomization process to a small extent.*

## 2.2 Cross-Sectional Campus File

The Cross-Sectional Campus File is a sample drawn from one recent year (2019) for all Germany. The sample is drawn from all advertisements that were in the database for at least one day within the respective year. For simplification only the last spell of the advertisement is covered. For municipalities with less than 50 observations only the district information is available with at most 100 observations drawn for this district. Afterwards a sample of 50, 200, 1000 or 5000 observations is drawn dependent on the number of observations in the full sample. Table 3 gives an overview of the number of municipalities identified in the data set. While most municipalities can be identified in the houses for sale sample, this is not the case in the apartments for sale since within the majority of municipalities less than 50 apartments (after data cleaning) are advertised. For these observations only the district is available.

Table 2

### Variables covered in the different campus files

	City Campus File			Cross-sectional Campus File		
	houses for sale	apartments for sale	apartments for rent	houses for sale	apartments for sale	apartments for rent
obid	x	x	x	x	x	x
mietekalt			x			x
nebenkosten			x			x
kaufpreis	x	x		x	x	
mieteinnahmenpromonat	x			x		
heizkosten	x			x		
baujahr	x	x	x	x	x	x
letzte_modernisierung	x	x	x	x	x	x
wohnflaeche	x	x	x	x	x	x
grundstuecksfläche	x	x		x	x	
nutzflaeche	x	x	x	x	x	x
etage		x	x		x	x
anzahletagen		x	x		x	x
zimmeranzahl	x	x	x	x	x	x
nebenraeume	x	x	x	x	x	x
schlafzimmer	x	x	x	x	x	x
badezimmer	x	x	x	x	x	x
parkplatzplatzpreis	x	x		x	x	
wohngeld		x			x	
ev_kennwert	x	x	x	x	x	x
laufzeittage	x	x	x	x	x	x
hits	x	x	x	x	x	x
click_schnellkontakte	x	x	x	x	x	x
click_customer	x	x	x	x	x	x
click_weitersagen	x	x	x	x	x	x
click_url	x	x	x	x	x	x
immobilientyp	x	x	x	x	x	x
denkmalobjekt	x	x		x	x	
aufzug		x	x		x	x
balkon		x	x		x	x
einbaukueche	x	x	x	x	x	x
einliegerwohnung	x	x		x	x	

## Campus Files RWI-GEO-RED city and RWI-GEO-RED cross

	City Campus File			Cross-sectional Campus File		
	houses for sale	apartments for sale	apartments for rent	houses for sale	apartments for sale	apartments for rent
ev_wwenthalten	x	x	x	x	x	x
ferienhaus	x			x		
foerderung	x		x			x
gaestewc	x	x	x	x	x	x
garten	x	x	x	x	x	x
kaufvermietet	x			x		
heizkosten			x			x
keller	x	x	x	x	x	x
parkplatz	x	x	x	x	x	x
rollstuhlgerecht	x	x	x	x	x	x
bauphase	x	x	x	x	x	x
ausstattung	x	x	x	x	x	x
energieeffizienzklasse	x	x	x	x	x	x
energieausweistyp	x	x	x	x	x	x
haustier_erlaubt	x	x	x	x	x	x
heizungsart	x	x	x	x	x	x
kategorie_Haus	x	x	x	x	x	x
objektzustand	x	x	x	x	x	x
lieferung			x			x
kid2019	x	x	x	x	x	x
ags2019	x	x	x	x	x	x
bef1	x	x	x	x	x	x
bef2	x	x	x	x	x	x
bef3	x	x	x	x	x	x
bef4	x	x	x	x	x	x
bef5	x	x	x	x	x	x
bef6	x	x	x	x	x	x
bef7	x	x	x	x	x	x
bef8	x	x	x	x	x	x
bef9	x	x	x	x	x	x
bef10	x	x	x	x	x	x
anbieter	x	x	x	x	x	x
duplicateid	x	x	x			
spell	x	x	x			
adat	x	x	x	x	x	x
edat	x	x	x			
price_sqm	x	x		x	x	
rent_sqm			x			x

Table 3

**Number of municipalities in the Cross-Sectional Campus file 2019 that can be identified**

NOBS	houses for sale	apartments for sale	apartments for rent
50	2 287	719	1 025
200	638	215	424
1000	25	31	97
5000	2	4	30
all municipalities	2 952	969	1 576

*All other municipalities are censored due to less than 50 observations after data cleaning. For these observations only district information is available.*

### 2.3 Data preparation

Both samples are based on the Scientific Use File and have therefore gone through some preparation. Further, we implement some additional cleaning before the samples are drawn. Some variables that are not relevant for the respective category or are suffering from high shares of missing variables are deleted from the SUF. Further, the information on the 1km<sup>2</sup> grid cell is not included. See Table 2 for an overview of all variables included. Further, those spells with a missing price/rent, missing information on the lot size are deleted. Further, some outliers are deleted regarding the price/rent, lot size, and price per sqm/ rent per sqm. Those spells below the 1<sup>st</sup> percentile and those above the 99<sup>th</sup> percentile are not included in the sampling. The same is true for objects that are available after 2020 (houses/apartments under construction for more than 1 year). Further, construction years before 1500 are set to missing value.

## 3. Variable Description

In the following, we describe each variable. Variables are ordered by category.

### 3.1 Identifier

Table 4

**Detailed Variable Descriptions of Identifiers**

Feature	Description
Label	Object identifier
Name	obid
Data type	Numeric
Description	Each property is uniquely identified by an artificial ID number. IDs are property-specific and do not change over time even if the object is temporarily withdrawn from the pool of advised real estates and offered again at a later time. Some IDs may be re-used over time when agents re-use previous advertisements. For duplicate spells, we provide a classification of these (see dupID_gen).

## 3.2 Time Period

Table 5

**Detailed Variable Descriptions of Variables Describing Time**

Feature	Description
Label	Beginning of ad, month
Name	adat
Data type	Numeric
Description	This is a numerical variable, which refers to the month during which an object is first advertised. If an object is advertised at least at some point in time during a certain month, this advertisement is included in the respective wave. If an advertisement is updated during a specific month, only the last update is recorded and enters the dataset.
Label	Ending of ad, month
Name	edat
Data type	Numeric
Description	This numeric variable refers to the month of the end of the advertisement. This can be misleading if the advertisement is split into two spells due to the timing of the data extraction from the database (see Data preparation/Data problems for more information). In City Campus File only

## 3.3 Object Features

Table 6

**Detailed Variable Descriptions of Variables Describing Object Features**

Feature	Description
Label	Elevator in object
Name	aufzug
Data type	Numeric
Description	This variable indicates if an object has an elevator.
Label	Facilities of object
Name	ausstattung
Data type	Numeric
Description	This is an artificial category number indicating the facilities of the object.
Label	Number of bathrooms
Name	badezimmer
Data type	Numeric
Description	This indicates the number of bathrooms in the object.
Label	Balcony at object
Name	balkon
Data type	Numeric
Description	This variable indicates the presence of a balcony.
Label	Protected historic building
Name	denkmalobjekt
Data type	Numeric
Description	This is an indicator of whether or not the property is protected.
Label	Kitchenette in object
Name	einbaukueche

Feature	Description
Data type	Numeric
Description	This variable indicates the presence of a fitted kitchen.
Label	Floor on which object is located
Name	etage
Data type	Numeric
Description	Apartment-specific variable indicates the floor the apartment is located in.
Label	Usable as holiday home
Name	ferienhaus
Data type	Numeric
Description	This is a binary indicator for whether the object can be used as a holiday home. It is only filled for apartments.
Label	Guest toilet in object
Name	gaestewc
Data type	Numeric
Description	This binary variable indicates the presence of a guest toilet.
Label	(Shared) garden available
Name	garten
Data type	Numeric
Description	This variable indicates the presence of a garden.
Label	Pets allowed
Name	haustier_erlaubt
Data type	Numeric
Description	This binary indicator shows whether pets are allowed in the object.
Label	House type
Name	kategorie_Haus
Data type	Numeric
Description	The artificial category number indicates which object category a property belongs to. Each property is assigned exactly one category number. This variable is filled for houses only.
Label	Cellar in object
Name	keller
Data type	Numeric
Description	This variable indicates if an object has a cellar or a cellar room.
Label	Garage/parking space available
Name	parkplatz
Data type	Numeric
Description	This is a binary variable indicating whether a parking space is available.
Label	Number of rooms
Name	zimmeranzahl
Data type	Numeric
Description	Number of rooms, excluding kitchen, bath or corridors. In several cases, "zimmeranzahl" is not a natural number, which is not necessarily due to a faulty entry. In Germany there is the concept of half rooms. Following the DIN 283 norm, a half room is defined as a room with a size between 6 and 10 square meters. While this definition is outdated, it is still frequently in use.

## 3.4 General Object Information

Table 7

**Detailed Variable Descriptions of General Object Information**

Feature	Description
Label	Number of floors
Name	anzahletagen
Data type	Numeric
Description	This indicates the number of floors in the building.
Label	Construction phase
Name	bauphase
Data type	Numeric
Description	This variable indicates whether the object is still under construction. Missings likely indicate that the object is not under construction.
Label	Granny flat in object
Name	einliegerwohnung
Data type	Numeric
Description	This variable indicates whether a granny flat is present in the given object.
Label	Public housing
Name	foerderung
Data type	Numeric
Description	This is a binary variable indicating that a certificate of eligibility to public housing is needed to rent the apartment
Label	Type of real estate
Name	immobilientyp
Data type	Numeric
Description	This artificial number indicates the type of a property.
Label	Rented when sold
Name	kaufvermietet
Data type	Numeric
Description	This variable indicates if an object for sale is already rented out.
Label	Rental income per month in EUR
Name	mieteinnahmenpromonat
Data type	Numeric
Description	For objects offered for sale, this indicates the rent income if the object is rented out. This is on a monthly basis and in EUR rounded to two decimal digits.
Label	Number of ancillary rooms
Name	nebenraeume
Data type	Numeric
Description	This is the number of ancillary rooms.
Label	Accessible, no steps
Name	rollstuhlgerecht
Data type	Numeric
Description	This is a binary indicator for step-free access of the object.
Label	Number of bedrooms
Name	schlafzimmer

Feature	Description
Data type	Numeric
Description	This is the number of bedrooms of the object.
Label	Common charge for community association in EUR/month
Name	wohngeld
Data type	Numeric
Description	This variable refers to the amount of the common charge for community association in EUR per month.

### 3.5 Area Information

Table 8

#### Detailed Variable Descriptions of Area Information

Feature	Description
Label	Plot area
Name	grundstuecksflaeche
Data type	Numeric
Description	This variable indicates the plot area of the object in square meters. Numbers are rounded to two decimal digits.
Label	Usable floor space
Name	nutzflaeche
Data type	Numeric
Description	This indicates the usable floor space in square meters. Numberes are rounded to two decimal digits.
Label	Living area
Name	wohnflaeche
Data type	String
Description	Living space in square meters. The precision of entries varies between natural numbers and numbers with two decimal places. If users enter more than two decimal places, numbers are rounded to two decimal places.



## 3.6 Energy and Structure Information

Table 9

**Detailed Variable Descriptions of Energy and Structure Information**

Feature	Description
Label	Year that object was built
Name	baujahr
Data type	Numeric
Description	Year in which the object was built. Observations that lie in the future are not necessarily faulty entries, potentially indicating that an object is still under construction.
Label	Type of Energy Performance Certificates (EPCs)
Name	energieausweistyp
Data type	Numeric
Description	This variable indicates the type of Energy Performance Certificate that the customer has for the object.
Label	Energy Efficiency Rating
Name	energieeffizienzklasse
Data type	Numeric
Description	The Energy Efficiency rating is represented here as an artificially created categorical variable.
Label	Energy consumption per year and square meter
Name	ev_kennwert
Data type	Numeric
Description	This indicated the energy consumption per year and square meter in kWh.
Label	Warm water consumption included in energy consumption
Name	ev_wwenthalten
Data type	Numeric
Description	This is a binary variable indicating whether the warm water consumption was included in the calculation of the energy consumption value.
Label	Heating costs
Name	heizkosten
Data type	Numeric
Description	This indicates the monthly heating costs in EUR and is rounded to two decimal digits.
Label	Type of heating
Name	heizungsart
Data type	Numeric
Description	This is an artificially created category number indicating the type of heating.
Label	Year of last modernisation of object
Name	letzte_modernisierung
Data type	Numeric
Description	Indicator for the year of the last modernisation.
Label	Condition of object
Name	objektzustand
Data type	Numeric

Feature	Description
Description	The artificial condition number indicates the condition of a property. Each property is assigned exactly one out of 11 possible numbers.

### 3.7 Price Information

Table 10

#### Detailed Variable Descriptions of Price Information

Feature	Description
Label	Purchasing price in EUR
Name	kaufpreis
Data type	Numeric
Description	Price at which the owner advertises to sell the object. Prices are expressed in EUR and rounded to two decimal digits. This is only filled for objects offered for sale.
Label	Exclusive rent in EUR
Name	mietekalt
Data type	Numeric
Description	Price at which the owner is willing to rent out the object. The rent covers expenses for the living space only. Amenities as well as expenses for heating or fees such as for garbage disposal are not included. Prices are expressed in EUR and rounded to two decimal digits. This is only filled for objects offered for rent.
Label	Utilities in EUR
Name	nebenkosten
Data type	Numeric
Description	This refers to the extra monthly costs that need to be paid for bills on top of the base rent. This usually does not include electricity, but other bills. Prices are expressed in EUR and rounded to two decimal digits.
Label	Price per sqm in EUR
Name	price_sqm
Data type	Numeric
Description	Calculated price per square meter by price and size of housesize of apartment
Label	Monthly Rent per sqm in EUR
Name	rent_sqm
Data type	Numeric
Description	Calculated rent per square meter by rent and size of apartment
Label	Price of parking space in EUR
Name	parkplatzpreis
Data type	Numeric
Description	This refers to the price of the parking place, expressed in EUR.

## 3.8 Regional Information

Table 11

**Detailed Variable Descriptions of Regional Information**

Feature	Description
Label	Municipality identifier (AGS, 2019)
Name	ags_2019
Data type	Numeric
Description	This is the municipality identifier according to the German Official Municipality Key (Amtlicher Gemeindeschlüssel). It is based on the territorial definition of 2019 (end of year).
Label	District identifier (AGS, 2019)
Name	kid_2019
Data type	Numeric
Description	This is the district identifier according to the German Official Municipality Key (Amtlicher Gemeindeschlüssel). It is based on the territorial definition of 2019 (end of year).

## 3.9 Meta-Information of Advertisement

Table 12

**Detailed Variable Descriptions of Variables Containing Meta-information of the Advertisement**

Feature	Description
Label	Number of clicks on customer profile
Name	click_customer
Data type	Numeric
Description	This variable indicates the number of clicks on the customer profile.
Label	Number of clicks on contact button
Name	click_schnellkontakte
Data type	Numeric
Description	This variable indicates the number of clicks on the contact button.
Label	Number of clicks on customer URL
Name	click_url
Data type	Numeric
Description	This variable indicates the number of clicks on the customer's URL.
Label	Number of clicks on share button
Name	click_weitersagen
Data type	Numeric
Description	This variable indicates the number of clicks on the share button.
Label	Number of hits of ad
Name	hits
Data type	Numeric
Description	This variable shows the number of hits that the advertisement got.
Label	Days of availability of ad
Name	laufzeittage
Data type	Numeric
Description	This variable indicates the number of days the respective advertisement has been online.
Label	Date of data retrieval

Name	lieferung
Data type	Numeric
Description	This variable shows the year and month in which the data was extracted from the database and delivered to RWI. Advertisements which overlap with this cut-off date will hence be split and duplicated. See more under dupID_gen.

### 3.10 Generated Technical Variables

Table 13

#### Detailed Variable Descriptions of Technical Variables

Feature	Description
Label	Firing type
Name	bef1 – bef10
Data type	categorical
Description	Firing type of heating system; if there are several heating types than they are listed in bef2, bef3 etc.
Label	Type of offeror
Name	anbieter
Data type	categorical
Description	Type of offeror like private, real-estate agent etc.
Label	Classification of object identifiers used more than once
Name	duplicateid
Data type	Numeric
Description	This is a generated variable, indicating whether a duplicate spell is likely to be the same object, or a new object. New objects can have the same ID when customers re-use a previous advertisement for another apartment, or when an object is re-rented/resold. There is also the possibility to strategically change certain variables. See more under Data preparation/Data problems. In City Campus File only
Label	Spell counter within object identifier
Name	spell
Data type	Numeric
Description	This is an artificially generated variable indicating the spell within each object ID, should an ID occur more than once. Spells are ordered chronologically. In City Campus File only



Leibniz-Institut für  
Wirtschaftsforschung

Das RWI wird vom Bund und vom Land  
Nordrhein-Westfalen gefördert.

