

Arvato Project Workbook

August 26, 2020

1 Capstone Project: Create a Customer Segmentation Report for Arvato Financial Services

In this project, you will analyze demographics data for customers of a mail-order sales company in Germany, comparing it against demographics information for the general population. You'll use unsupervised learning techniques to perform customer segmentation, identifying the parts of the population that best describe the core customer base of the company. Then, you'll apply what you've learned on a third dataset with demographics information for targets of a marketing campaign for the company, and use a model to predict which individuals are most likely to convert into becoming customers for the company. The data that you will use has been provided by our partners at Bertelsmann Arvato Analytics, and represents a real-life data science task.

If you completed the first term of this program, you will be familiar with the first part of this project, from the unsupervised learning project. The versions of those two datasets used in this project will include many more features and has not been pre-cleaned. You are also free to choose whatever approach you'd like to analyzing the data rather than follow pre-determined steps. In your work on this project, make sure that you carefully document your steps and decisions, since your main deliverable for this project will be a blog post reporting your findings.

```
In [158]: # import libraries here; add more as necessary
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
# from yellowbrick.cluster.elbow import kelbow_visualizer
from sklearn.cluster import KMeans
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.pipeline import Pipeline

pd.set_option('display.max_columns', 500)
```

```
# magic word for producing visualizations in notebook
%matplotlib inline
```

1.1 Part 0: Get to Know the Data

There are four data files associated with this project:

- `Udacity_AZDIAS_052018.csv`: Demographics data for the general population of Germany; 891 211 persons (rows) x 366 features (columns).
- `Udacity_CUSTOMERS_052018.csv`: Demographics data for customers of a mail-order company; 191 652 persons (rows) x 369 features (columns).
- `Udacity_MAILOUT_052018_TRAIN.csv`: Demographics data for individuals who were targets of a marketing campaign; 42 982 persons (rows) x 367 (columns).
- `Udacity_MAILOUT_052018_TEST.csv`: Demographics data for individuals who were targets of a marketing campaign; 42 833 persons (rows) x 366 (columns).

Each row of the demographics files represents a single person, but also includes information outside of individuals, including information about their household, building, and neighborhood. Use the information from the first two files to figure out how customers ("CUSTOMERS") are similar to or differ from the general population at large ("AZDIAS"), then use your analysis to make predictions on the other two files ("MAILOUT"), predicting which recipients are most likely to become a customer for the mail-order company.

The "CUSTOMERS" file contains three extra columns ('CUSTOMER_GROUP', 'ONLINE_PURCHASE', and 'PRODUCT_GROUP'), which provide broad information about the customers depicted in the file. The original "MAILOUT" file included one additional column, "RESPONSE", which indicated whether or not each recipient became a customer of the company. For the "TRAIN" subset, this column has been retained, but in the "TEST" subset it has been removed; it is against that withheld column that your final predictions will be assessed in the Kaggle competition.

Otherwise, all of the remaining columns are the same between the three data files. For more information about the columns depicted in the files, you can refer to two Excel spreadsheets provided in the workspace. [One of them](#) is a top-level list of attributes and descriptions, organized by informational category. [The other](#) is a detailed mapping of data values for each feature in alphabetical order.

In the below cell, we've provided some initial code to load in the first two datasets. Note for all of the .csv data files in this project that they're semicolon (;) delimited, so an additional argument in the `read_csv()` call has been included to read in the data properly. Also, considering the size of the datasets, it may take some time for them to load completely.

You'll notice when the data is loaded in that a warning message will immediately pop up. Before you really start digging into the modeling and analysis, you're going to need to perform some cleaning. Take some time to browse the structure of the data and look over the informational spreadsheets to understand the data values. Make some decisions on which features to keep, which features to drop, and if any revisions need to be made on data formats. It'll be a good idea to create a function with pre-processing steps, since you'll need to clean all of the datasets before you work with them.

Examine azdias dataframe

```
In [2]: # load in the data
        azdias = pd.read_csv('../data/Term2/capstone/arvato_data/Udacity_AZDIAS_052018.csv',

        azdias.head()

/opt/conda/lib/python3.6/site-packages/IPython/core/interactiveshell.py:2785: DtypeWarning: Columns (
interactivity=interactivity, compiler=compiler, result=result)
```

```
Out[2]:
```

| | LNR | AGER_TYP | AKT_DAT_KL | ALTER_HH | ALTER_KIND1 | ALTER_KIND2 | \ |
|---|--------|----------|------------|----------|-------------|-------------|---|
| 0 | 910215 | -1 | NaN | NaN | NaN | NaN | |
| 1 | 910220 | -1 | 9.0 | 0.0 | NaN | NaN | |
| 2 | 910225 | -1 | 9.0 | 17.0 | NaN | NaN | |
| 3 | 910226 | 2 | 1.0 | 13.0 | NaN | NaN | |
| 4 | 910241 | -1 | 1.0 | 20.0 | NaN | NaN | |

| | ALTER_KIND3 | ALTER_KIND4 | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | \ |
|---|-------------|-------------|----------------------|---------------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | NaN | NaN | 21.0 | 11.0 | |
| 2 | NaN | NaN | 17.0 | 10.0 | |
| 3 | NaN | NaN | 13.0 | 1.0 | |
| 4 | NaN | NaN | 14.0 | 3.0 | |

| | ANZ_HH_TITEL | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE | \ |
|---|--------------|------------|--------------|----------------------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 0.0 | 2.0 | 12.0 | |
| 2 | 0.0 | 0.0 | 1.0 | 7.0 | |
| 3 | 0.0 | 0.0 | 0.0 | 2.0 | |
| 4 | 0.0 | 0.0 | 4.0 | 3.0 | |

| | ANZ_TITEL | ARBEIT | BALLRAUM | CAMEO_DEU_2015 | CAMEO_DEUG_2015 | CAMEO_INTL_2015 | \ |
|---|-----------|--------|----------|----------------|-----------------|-----------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 3.0 | 6.0 | 8A | 8 | 51 | |
| 2 | 0.0 | 3.0 | 2.0 | 4C | 4 | 24 | |
| 3 | 0.0 | 2.0 | 4.0 | 2A | 2 | 12 | |
| 4 | 0.0 | 4.0 | 2.0 | 6B | 6 | 43 | |

| | CJT_GESAMTTYP | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 | \ | |
|---|---------------|-------------------|-----------|-----------|-----------|-----|--|
| 0 | 2.0 | | 5.0 | 1.0 | 1.0 | 5.0 | |
| 1 | 5.0 | | 1.0 | 5.0 | 5.0 | 2.0 | |
| 2 | 3.0 | | 2.0 | 4.0 | 4.0 | 1.0 | |
| 3 | 2.0 | | 3.0 | 2.0 | 2.0 | 4.0 | |
| 4 | 5.0 | | 3.0 | 3.0 | 3.0 | 3.0 | |

| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 | D19_BANKEN_ANZ_24 | \ |
|---|-----------|-----------|-----------|-------------------|-------------------|---|
| 0 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 1 | 3.0 | 1.0 | 1.0 | 0 | 0 | |
| 2 | 3.0 | 2.0 | 2.0 | 0 | 0 | |

| | | | | | |
|---|-----|-----|-----|---|---|
| 3 | 4.0 | 5.0 | 3.0 | 0 | 0 |
| 4 | 4.0 | 3.0 | 3.0 | 3 | 5 |

| | D19_BANKEN_DATUM | D19_BANKEN_DIREKT | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | \ |
|---|------------------|-------------------|------------------|------------------|---|
| 0 | 10 | 0 | 0 | 0 | |
| 1 | 10 | 0 | 0 | 0 | |
| 2 | 10 | 0 | 0 | 0 | |
| 3 | 10 | 0 | 0 | 0 | |
| 4 | 5 | 1 | 2 | 0 | |

| | D19_BANKEN_OFFLINE_DATUM | D19_BANKEN_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |
| 2 | 10 | 10 | |
| 3 | 10 | 10 | |
| 4 | 10 | 5 | |

| | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST | D19_BEKLEIDUNG_GEH | \ |
|---|----------------------------|-----------------|--------------------|---|
| 0 | NaN | 0 | 0 | |
| 1 | NaN | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 10.0 | 6 | 6 | |

| | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO | D19_BUCH_CD | \ |
|---|---------------------|-------------|--------------|-------------|---|
| 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 6 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 6 | |
| 4 | 1 | 6 | 0 | 6 | |

| | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE | D19_FREIZEIT | D19_GARTEN | \ |
|---|----------------|---------------------|-------------|--------------|------------|---|
| 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 0 | |
| 4 | 0 | 1 | 5 | 0 | 0 | |

| | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 | D19_GESAMT_DATUM | \ |
|---|-------------------|-------------------|------------------|---|
| 0 | 0 | 0 | 10 | |
| 1 | 0 | 0 | 10 | |
| 2 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 10 | |
| 4 | 6 | 6 | 1 | |

| | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |

| | | |
|---|----|----|
| 2 | 10 | 10 |
| 3 | 10 | 10 |
| 4 | 6 | 1 |

| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO | D19_KINDERARTIKEL | \ |
|---|----------------------------|--------------|---------------|-------------------|---|
| 0 | NaN | 0 | 0 | 0 | |
| 1 | NaN | 0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | 0 | |
| 4 | 10.0 | 0 | 5 | 0 | |

| | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK | D19_LEBENSMITTEL | \ |
|---|---------------|-------------------|--------------|------------------|---|
| 0 | NaN | 9 | 0 | 0 | |
| 1 | NaN | 9 | 0 | 0 | |
| 2 | 9.0 | 8 | 6 | 0 | |
| 3 | 9.0 | 8 | 0 | 0 | |
| 4 | 1.0 | 1 | 0 | 0 | |

| | D19_LETZTER_KAUF_BRANCHE | D19_LOTTO | D19_NAHRUNGSERGAEZUNG | D19_RATGEBER | \ |
|---|--------------------------|-----------|-----------------------|--------------|---|
| 0 | NaN | NaN | 0 | 0 | |
| 1 | NaN | NaN | 0 | 0 | |
| 2 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 3 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 4 | D19_SCHUHE | 0.0 | 0 | 0 | |

| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE | D19_SOZIALES | \ |
|---|------------|-------------------|------------|--------------|--------------|---|
| 0 | 0 | 0 | 0 | 0 | NaN | |
| 1 | 0 | 0 | 0 | 0 | NaN | |
| 2 | 0 | 0 | 0 | 6 | 0.0 | |
| 3 | 6 | 6 | 0 | 6 | 0.0 | |
| 4 | 0 | 0 | 1 | 4 | 0.0 | |

| | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 | D19_TELKO_DATUM | \ |
|---|-------------|------------------|------------------|-----------------|---|
| 0 | 0 | 0 | 0 | 10 | |
| 1 | 0 | 0 | 0 | 10 | |
| 2 | 6 | 0 | 0 | 10 | |
| 3 | 6 | 0 | 0 | 10 | |
| 4 | 5 | 0 | 1 | 6 | |

| | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM | D19_TELKO_ONLINE_DATUM | \ |
|---|------------------|-------------------------|------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 0 | 10 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 6 | 8 | 10 | |

| | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST | D19_TIERARTIKEL | \ |
|---|---------------------------|----------------|-----------------|---|
| 0 | NaN | 0 | 0 | |

| | | | |
|---|-----|---|---|
| 1 | NaN | 0 | 0 |
| 2 | 0.0 | 0 | 0 |
| 3 | 0.0 | 0 | 0 |
| 4 | 0.0 | 5 | 0 |

| | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 | D19_VERSAND_DATUM | \ |
|---|--------------------|--------------------|-------------------|---|
| 0 | 0 | 0 | 10 | |
| 1 | 0 | 0 | 10 | |
| 2 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 10 | |
| 4 | 6 | 6 | 1 | |

| | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM | \ |
|---|---------------------------|--------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |
| 2 | 10 | 10 | |
| 3 | 10 | 10 | |
| 4 | 9 | 1 | |

| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 | \ |
|---|-----------------------------|------------------|------------------|---|
| 0 | NaN | 0 | 0 | |
| 1 | NaN | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 10.0 | 2 | 1 | |

| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM | \ |
|---|------------------|-----------------|-------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 0 | 10 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 3 | 2 | 7 | |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN | \ |
|---|------------------------|---------------------------|--------------------|---|
| 0 | 10 | NaN | 0 | |
| 1 | 10 | NaN | 0 | |
| 2 | 10 | 0.0 | 0 | |
| 3 | 10 | 0.0 | 0 | |
| 4 | 10 | 0.0 | 3 | |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG | EINGEFUEGT_AM | \ |
|---|-------------------|-------------------|----------|---------------------|---|
| 0 | 0 | 0 | NaN | NaN | |
| 1 | 0 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 2 | 7 | 0 | 1.0 | 1992-02-12 00:00:00 | |
| 3 | 0 | 0 | 1.0 | 1997-04-21 00:00:00 | |
| 4 | 0 | 0 | 1.0 | 1992-02-12 00:00:00 | |

| EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER | \ |
|----------------------|----------|-----------|----------------|---|
|----------------------|----------|-----------|----------------|---|

| | | | | |
|---|--------|-----|------|---|
| 0 | NaN | NaN | NaN | 5 |
| 1 | 2004.0 | 3.0 | NaN | 5 |
| 2 | 2000.0 | 4.0 | 14.0 | 2 |
| 3 | 1998.0 | 2.0 | 31.0 | 2 |
| 4 | 2004.0 | 5.0 | NaN | 1 |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER | FINANZ_UNAUFFAELLIGER | \ |
|---|------------------|-------------------|---------------|-----------------------|---|
| 0 | 3 | 3 | 4 | 5 | |
| 1 | 5 | 1 | 5 | 4 | |
| 2 | 5 | 1 | 4 | 3 | |
| 3 | 2 | 4 | 2 | 1 | |
| 4 | 2 | 4 | 3 | 3 | |

| | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE | GEBAEUDETYP | GEBAEUDETYP_RASTER | \ |
|---|------------------|-----------|--------------|-------------|--------------------|---|
| 0 | 3 | 4 | NaN | NaN | NaN | |
| 1 | 2 | 1 | 2.0 | 8.0 | 3.0 | |
| 2 | 1 | 1 | 4.0 | 1.0 | 4.0 | |
| 3 | 5 | 6 | 4.0 | 1.0 | 4.0 | |
| 4 | 4 | 5 | 5.0 | 1.0 | 5.0 | |

| | GEBURTSJAHR | GEMEINDETYP | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP | \ |
|---|-------------|-------------|-----------------|------------------|------------|---|
| 0 | 0 | NaN | 10.0 | 0 | -1 | |
| 1 | 1996 | 22.0 | 10.0 | 0 | 3 | |
| 2 | 1979 | 22.0 | 10.0 | 1 | 3 | |
| 3 | 1957 | 40.0 | 1.0 | 0 | 2 | |
| 4 | 1963 | 21.0 | 5.0 | 0 | 3 | |

| | HH_DELTA_FLAG | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 | \ |
|---|---------------|--------------------|------------|--------------|--------------|---|
| 0 | NaN | 2.0 | NaN | NaN | NaN | |
| 1 | 0.0 | 6.0 | 8.0 | 3.0 | 4.0 | |
| 2 | 0.0 | 4.0 | 4.0 | 2.0 | 3.0 | |
| 3 | NaN | 1.0 | 6.0 | 2.0 | 5.0 | |
| 4 | 0.0 | 5.0 | 1.0 | 0.0 | 4.0 | |

| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 | KBA05_ANTG2 | \ |
|---|--------------|--------------|--------------|-------------|-------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 1.0 | 4.0 | 0.0 | 0.0 | 0.0 | |
| 2 | 3.0 | 3.0 | 0.0 | 1.0 | 3.0 | |
| 3 | 3.0 | 0.0 | 1.0 | 4.0 | 1.0 | |
| 4 | 4.0 | 3.0 | 0.0 | 1.0 | 4.0 | |

| | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT | KBA05_BAUMAX | KBA05_CCM1 | \ |
|---|-------------|-------------|----------------|--------------|------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 2.0 | 1.0 | 5.0 | 1.0 | |
| 2 | 1.0 | 0.0 | 3.0 | 0.0 | 5.0 | |
| 3 | 0.0 | 0.0 | 4.0 | 1.0 | 2.0 | |
| 4 | 1.0 | 0.0 | 3.0 | 0.0 | 4.0 | |

| | KBA05_CCM2 | KBA05_CCM3 | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ \ |
|---|------------|------------|------------|--------------|------------|-------------|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN |
| 1 | 5.0 | 1.0 | 4.0 | 2.0 | 4.0 | 1.0 |
| 2 | 2.0 | 3.0 | 0.0 | 0.0 | 3.0 | 3.0 |
| 3 | 3.0 | 5.0 | 1.0 | 3.0 | 4.0 | 4.0 |
| 4 | 1.0 | 4.0 | 2.0 | 1.0 | 5.0 | 3.0 |

| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 | KBA05_HERST5 \ |
|---|--------------|--------------|--------------|--------------|----------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 5.0 | 5.0 | 2.0 | 2.0 | 0.0 |
| 2 | 2.0 | 2.0 | 3.0 | 2.0 | 5.0 |
| 3 | 4.0 | 3.0 | 3.0 | 2.0 | 3.0 |
| 4 | 2.0 | 2.0 | 4.0 | 1.0 | 4.0 |

| | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 | KBA05_KRSHERST2 \ |
|---|-----------------|----------------|-----------------|-------------------|
| 0 | NaN | NaN | NaN | NaN |
| 1 | 4.0 | 1.0 | 5.0 | 4.0 |
| 2 | 4.0 | 3.0 | 3.0 | 2.0 |
| 3 | 3.0 | 4.0 | 4.0 | 2.0 |
| 4 | 3.0 | 3.0 | 3.0 | 3.0 |

| | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER | KBA05_KRSVAN | KBA05_KRSZUL \ |
|---|-----------------|----------------|---------------|--------------|----------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 2.0 | 1.0 | 2.0 | 1.0 | 2.0 |
| 2 | 3.0 | 3.0 | 2.0 | 2.0 | 3.0 |
| 3 | 3.0 | 1.0 | 2.0 | 2.0 | 3.0 |
| 4 | 3.0 | 2.0 | 2.0 | 3.0 | 2.0 |

| | KBA05_KW1 | KBA05_KW2 | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST \ |
|---|-----------|-----------|-----------|-------------|-------------|------------------|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN |
| 1 | 1.0 | 3.0 | 4.0 | 2.0 | 1.0 | 2.0 |
| 2 | 3.0 | 2.0 | 2.0 | 3.0 | 4.0 | 5.0 |
| 3 | 3.0 | 4.0 | 1.0 | 3.0 | 4.0 | 3.0 |
| 4 | 3.0 | 2.0 | 3.0 | 2.0 | 2.0 | 3.0 |

| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 | KBA05_MOD3 \ |
|---|--------------|---------------|------------|------------|--------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 4.0 | 3.0 | 3.0 | 2.0 | 2.0 |
| 2 | 1.0 | 1.0 | 0.0 | 2.0 | 5.0 |
| 3 | 2.0 | 2.0 | 2.0 | 4.0 | 4.0 |
| 4 | 1.0 | 1.0 | 3.0 | 1.0 | 1.0 |

| | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP | KBA05_MOTOR | KBA05_MOTRAD \ |
|---|------------|------------|---------------|-------------|----------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 0.0 | 0.0 | 1.0 | 3.0 | 0.0 |
| 2 | 1.0 | 1.0 | 4.0 | 1.0 | 1.0 |
| 3 | 2.0 | 1.0 | 3.0 | 3.0 | 3.0 |
| 4 | 4.0 | 2.0 | 3.0 | 4.0 | 1.0 |

| | KBA05_SEG1 | KBA05_SEG10 | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 | \ |
|---|------------|-------------|------------|------------|------------|------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 4.0 | 1.0 | 2.0 | 2.0 | 2.0 | |
| 2 | 2.0 | 1.0 | 5.0 | 3.0 | 3.0 | 1.0 | |
| 3 | 1.0 | 1.0 | 2.0 | 5.0 | 3.0 | 2.0 | |
| 4 | 3.0 | 3.0 | 4.0 | 1.0 | 3.0 | 2.0 | |

| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 | KBA05_VORBO | KBA05_VORB1 | \ |
|---|------------|------------|------------|------------|-------------|-------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 1 | 1.0 | 3.0 | 3.0 | 0.0 | 1.0 | 1.0 | |
| 2 | 0.0 | 0.0 | 0.0 | 1.0 | 4.0 | 2.0 | |
| 3 | 0.0 | 0.0 | 0.0 | 1.0 | 2.0 | 5.0 | |
| 4 | 0.0 | 1.0 | 0.0 | 2.0 | 5.0 | 1.0 | |

| | KBA05_VORB2 | KBA05_ZUL1 | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | \ |
|---|-------------|------------|------------|------------|------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 5.0 | 5.0 | 1.0 | 0.0 | 2.0 | |
| 2 | 3.0 | 2.0 | 3.0 | 4.0 | 4.0 | |
| 3 | 1.0 | 3.0 | 3.0 | 3.0 | 3.0 | |
| 4 | 5.0 | 3.0 | 4.0 | 2.0 | 2.0 | |

| | KBA13_ALTERHALTER_30 | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | \ |
|---|----------------------|----------------------|----------------------|---|
| 0 | NaN | NaN | NaN | |
| 1 | 3.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 2.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 5.0 | |
| 4 | 3.0 | 3.0 | 3.0 | |

| | KBA13_ALTERHALTER_61 | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 | \ |
|---|----------------------|-------------|-------------|-------------|-------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 4.0 | 2.0 | 4.0 | 2.0 | 1.0 | |
| 2 | 3.0 | 2.0 | 3.0 | 1.0 | 0.0 | |
| 3 | 2.0 | 2.0 | 3.0 | 1.0 | 0.0 | |
| 4 | 3.0 | 1.0 | 4.0 | 2.0 | 1.0 | |

| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX | KBA13_BJ_1999 | \ |
|---|------------------|------------|-----------------|--------------|---------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 963.0 | 4.0 | 2.0 | 2.0 | 3.0 | |
| 2 | 712.0 | 3.0 | 3.0 | 1.0 | 2.0 | |
| 3 | 596.0 | 5.0 | 3.0 | 1.0 | 2.0 | |
| 4 | 435.0 | 4.0 | 3.0 | 2.0 | 3.0 | |

| | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 | KBA13_BJ_2008 | KBA13_BJ_2009 | \ |
|---|---------------|---------------|---------------|---------------|---------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 2 | 2.0 | 4.0 | 5.0 | 3.0 | 1.0 | |
| 3 | 2.0 | 3.0 | 3.0 | 4.0 | 3.0 | |

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 4 | 3.0 | 3.0 | 2.0 | 0.0 | 5.0 |
|---|-----|-----|-----|-----|-----|

| | KBA13_BMW | KBA13_CCM_0_1400 | KBA13_CCM_1000 | KBA13_CCM_1200 | \ |
|---|-----------|------------------|----------------|----------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 2.0 | 0.0 | 0.0 | |
| 2 | 4.0 | 1.0 | 1.0 | 2.0 | |
| 3 | 4.0 | 3.0 | 4.0 | 2.0 | |
| 4 | 2.0 | 3.0 | 5.0 | 1.0 | |

| | KBA13_CCM_1400 | KBA13_CCM_1401_2500 | KBA13_CCM_1500 | KBA13_CCM_1600 | \ |
|---|----------------|---------------------|----------------|----------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 4.0 | 3.0 | 1.0 | 2.0 | |
| 2 | 2.0 | 3.0 | 4.0 | 3.0 | |
| 3 | 3.0 | 2.0 | 3.0 | 3.0 | |
| 4 | 2.0 | 1.0 | 4.0 | 1.0 | |

| | KBA13_CCM_1800 | KBA13_CCM_2000 | KBA13_CCM_2500 | KBA13_CCM_2501 | \ |
|---|----------------|----------------|----------------|----------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 2.0 | 5.0 | 3.0 | 3.0 | |
| 2 | 4.0 | 3.0 | 3.0 | 4.0 | |
| 3 | 2.0 | 3.0 | 4.0 | 4.0 | |
| 4 | 3.0 | 3.0 | 3.0 | 5.0 | |

| | KBA13_CCM_3000 | KBA13_CCM_3001 | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE | \ |
|---|----------------|----------------|-----------------|--------------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 5.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 5.0 | 4.0 | 3.0 | |
| 3 | 3.0 | 5.0 | 3.0 | 2.0 | |
| 4 | 5.0 | 5.0 | 3.0 | 2.0 | |

| | KBA13_FIAT | KBA13_FORD | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 | \ |
|---|------------|------------|-----------|-----------------|-----------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 4.0 | 2.0 | 4.0 | 3.0 | 3.0 | |
| 2 | 3.0 | 4.0 | 4.0 | 3.0 | 3.0 | |
| 3 | 3.0 | 3.0 | 4.0 | 2.0 | 2.0 | |
| 4 | 3.0 | 4.0 | 3.0 | 3.0 | 3.0 | |

| | KBA13_HALTER_30 | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 2 | 2.0 | 2.0 | 2.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 4 | 3.0 | 3.0 | 3.0 | 3.0 | |

| | KBA13_HALTER_50 | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 2 | 3.0 | 3.0 | 3.0 | 4.0 | |

| | | | | |
|---|-----|-----|-----|-----|
| 3 | 5.0 | 5.0 | 4.0 | 3.0 |
| 4 | 4.0 | 4.0 | 2.0 | 3.0 |

| | | | | |
|---|-----------------|-------------------|---------------------|---|
| | KBA13_HALTER_66 | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | \ |
| 0 | NaN | NaN | NaN | |
| 1 | 4.0 | 1.0 | 4.0 | |
| 2 | 3.0 | 3.0 | 2.0 | |
| 3 | 2.0 | 3.0 | 4.0 | |
| 4 | 3.0 | 3.0 | 4.0 | |

| | | | | |
|---|----------------------|--------------------|-----------------------|---|
| | KBA13_HERST_BMW_BENZ | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | \ |
| 0 | NaN | NaN | NaN | |
| 1 | 4.0 | 4.0 | 2.0 | |
| 2 | 4.0 | 3.0 | 3.0 | |
| 3 | 3.0 | 2.0 | 2.0 | |
| 4 | 3.0 | 4.0 | 3.0 | |

| | | | | | |
|---|-------------------|-----------|-----------------|---------------|---|
| | KBA13_HERST_SONST | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | \ |
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 5.0 | 3.0 | 1.0 | |
| 2 | 3.0 | 4.0 | 1.0 | 1.0 | |
| 3 | 2.0 | 3.0 | 5.0 | 1.0 | |
| 4 | 2.0 | 3.0 | 5.0 | 1.0 | |

| | | | | | |
|---|---------------|-------------------|---------------|---------------|---|
| | KBA13_KMH_140 | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | \ |
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 3.0 | 2.0 | 4.0 | |
| 2 | 1.0 | 2.0 | 2.0 | 4.0 | |
| 3 | 5.0 | 1.0 | 2.0 | 2.0 | |
| 4 | 5.0 | 1.0 | 1.0 | 3.0 | |

| | | | | | |
|---|---------------|---------------|---------------|----------------|---|
| | KBA13_KMH_211 | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | \ |
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 3.0 | 1.0 | 2.0 | |
| 2 | 4.0 | 4.0 | 1.0 | 3.0 | |
| 3 | 5.0 | 5.0 | 1.0 | 3.0 | |
| 4 | 5.0 | 5.0 | 1.0 | 3.0 | |

| | | | | |
|---|------------------------|-------------------------|--------------------------|---|
| | KBA13_KRSHERST_AUDI_VW | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | \ |
| 0 | NaN | NaN | NaN | |
| 1 | 4.0 | 3.0 | 3.0 | |
| 2 | 3.0 | 3.0 | 2.0 | |
| 3 | 4.0 | 4.0 | 2.0 | |
| 4 | 4.0 | 3.0 | 2.0 | |

| | | | | | |
|---|--------------------|-------------------|------------------|------------------|---|
| | KBA13_KRSSEG_KLEIN | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | \ |
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 2.0 | 2.0 | 2.0 | 1.0 | |

| | | | | |
|---|-----|-----|-----|-----|
| 2 | 2.0 | 3.0 | 2.0 | 1.0 |
| 3 | 2.0 | 2.0 | 2.0 | 2.0 |
| 4 | 2.0 | 2.0 | 1.0 | 3.0 |

| | KBA13_KW_0_60 | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 \ |
|---|---------------|--------------|--------------|--------------|---------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 3.0 | 4.0 | 4.0 | 3.0 | 1.0 |
| 2 | 1.0 | 3.0 | 4.0 | 4.0 | 1.0 |
| 3 | 3.0 | 1.0 | 3.0 | 5.0 | 1.0 |
| 4 | 3.0 | 3.0 | 1.0 | 5.0 | 1.0 |

| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 | KBA13_KW_70 \ |
|---|-------------|-------------|-------------|-----------------|---------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 2.0 | 4.0 | 0.0 | 3.0 | 1.0 |
| 2 | 1.0 | 2.0 | 1.0 | 5.0 | 4.0 |
| 3 | 4.0 | 3.0 | 2.0 | 2.0 | 3.0 |
| 4 | 5.0 | 3.0 | 0.0 | 2.0 | 2.0 |

| | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA | KBA13_MERCEDES | KBA13_MOTOR \ |
|---|-------------|-------------|-------------|----------------|---------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 2.0 | 3.0 | 2.0 | 4.0 | 3.0 |
| 2 | 4.0 | 2.0 | 3.0 | 4.0 | 3.0 |
| 3 | 2.0 | 3.0 | 2.0 | 3.0 | 3.0 |
| 4 | 1.0 | 3.0 | 3.0 | 3.0 | 4.0 |

| | KBA13_NISSAN | KBA13_OPEL | KBA13_PEUGEOT | KBA13_RENAULT \ |
|---|--------------|------------|---------------|-----------------|
| 0 | NaN | NaN | NaN | NaN |
| 1 | 2.0 | 3.0 | 4.0 | 3.0 |
| 2 | 3.0 | 2.0 | 3.0 | 3.0 |
| 3 | 5.0 | 2.0 | 3.0 | 2.0 |
| 4 | 4.0 | 3.0 | 3.0 | 4.0 |

| | KBA13_SEG_GELAENDEWAGEN | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST \ |
|---|-------------------------|-------------------------|---------------------|
| 0 | NaN | NaN | NaN |
| 1 | 2.0 | 3.0 | 2.0 |
| 2 | 5.0 | 3.0 | 3.0 |
| 3 | 3.0 | 4.0 | 3.0 |
| 4 | 3.0 | 3.0 | 3.0 |

| | KBA13_SEG_KLEINWAGEN | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS \ |
|---|----------------------|-------------------------|----------------------|
| 0 | NaN | NaN | NaN |
| 1 | 2.0 | 5.0 | 4.0 |
| 2 | 3.0 | 1.0 | 3.0 |
| 3 | 3.0 | 4.0 | 4.0 |
| 4 | 3.0 | 3.0 | 2.0 |

| | KBA13_SEG_MINIWAGEN | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE \ |
|---|---------------------|------------------------|-------------------------------|
| 0 | NaN | NaN | NaN |

| | | | |
|---|-----|-----|-----|
| 1 | 2.0 | 3.0 | 3.0 |
| 2 | 3.0 | 2.0 | 4.0 |
| 3 | 3.0 | 4.0 | 3.0 |
| 4 | 3.0 | 2.0 | 4.0 |

| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN | \ |
|---|----------------------|--------------------|----------------------|---|
| 0 | NaN | NaN | NaN | |
| 1 | 3.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 2.0 | 4.0 | |
| 3 | 1.0 | 2.0 | 3.0 | |
| 4 | 4.0 | 5.0 | 4.0 | |

| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE | KBA13_SITZE_4 | \ |
|---|---------------------|---------------|----------------------|---------------|---|
| 0 | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 4.0 | 2.0 | 3.0 | |
| 2 | 5.0 | 3.0 | 2.0 | 4.0 | |
| 3 | 2.0 | 4.0 | 2.0 | 3.0 | |
| 4 | 2.0 | 2.0 | 5.0 | 3.0 | |

| | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA | KBA13_VORB_0 | KBA13_VORB_1 | \ |
|---|---------------|---------------|--------------|--------------|--------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 4.0 | 2.0 | 3.0 | 3.0 | |
| 2 | 2.0 | 3.0 | 3.0 | 3.0 | 4.0 | |
| 3 | 3.0 | 3.0 | 3.0 | 4.0 | 3.0 | |
| 4 | 3.0 | 3.0 | 3.0 | 4.0 | 2.0 | |

| | KBA13_VORB_1_2 | KBA13_VORB_2 | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK | \ |
|---|----------------|--------------|--------------|----------|--------------|-----|---|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 1 | 3.0 | 3.0 | 3.0 | 4.0 | NaN | 2.0 | |
| 2 | 4.0 | 2.0 | 2.0 | 2.0 | NaN | 2.0 | |
| 3 | 2.0 | 3.0 | 2.0 | 4.0 | NaN | 0.0 | |
| 4 | 2.0 | 3.0 | 4.0 | 3.0 | 1.0 | 3.0 | |

| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN | LP_FAMILIE_GROB | \ |
|---|------------|-------------|-------------|-----------------|-----------------|---|
| 0 | 9 | NaN | NaN | 2.0 | 2.0 | |
| 1 | 1 | 1.0 | 1.0 | 5.0 | 3.0 | |
| 2 | 2 | 5.0 | 0.0 | 1.0 | 1.0 | |
| 3 | 4 | 4.0 | 0.0 | 0.0 | 0.0 | |
| 4 | 3 | 4.0 | 0.0 | 10.0 | 5.0 | |

| | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB | LP_STATUS_FEIN | LP_STATUS_GROB | \ |
|---|---------------------|---------------------|----------------|----------------|---|
| 0 | 15.0 | 4.0 | 1.0 | 1.0 | |
| 1 | 21.0 | 6.0 | 2.0 | 1.0 | |
| 2 | 3.0 | 1.0 | 3.0 | 2.0 | |
| 3 | 0.0 | 0.0 | 9.0 | 4.0 | |
| 4 | 32.0 | 10.0 | 3.0 | 2.0 | |

| | MIN_GEBAEUDEJAHR | MOBI_RASTER | MOBI_REGIO | NATIONALITAET_KZ | \ |
|--|------------------|-------------|------------|------------------|---|
|--|------------------|-------------|------------|------------------|---|

| | | | | |
|---|--------|-----|-----|---|
| 0 | NaN | NaN | NaN | 0 |
| 1 | 1992.0 | 1.0 | 1.0 | 1 |
| 2 | 1992.0 | 2.0 | 3.0 | 1 |
| 3 | 1997.0 | 4.0 | 4.0 | 1 |
| 4 | 1992.0 | 1.0 | 3.0 | 1 |

| | ONLINE_AFFINITAET | ORTSGR_KLS9 | OST_WEST_KZ | PLZ8_ANTG1 | PLZ8_ANTG2 \ |
|---|-------------------|-------------|-------------|------------|--------------|
| 0 | 1.0 | NaN | NaN | NaN | NaN |
| 1 | 3.0 | 5.0 | W | 2.0 | 3.0 |
| 2 | 2.0 | 5.0 | W | 3.0 | 3.0 |
| 3 | 1.0 | 3.0 | W | 2.0 | 2.0 |
| 4 | 5.0 | 6.0 | W | 2.0 | 4.0 |

| | PLZ8_ANTG3 | PLZ8_ANTG4 | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ \ |
|---|------------|------------|-------------|----------|------------|
| 0 | NaN | NaN | NaN | NaN | NaN |
| 1 | 2.0 | 1.0 | 1.0 | 4.0 | 5.0 |
| 2 | 1.0 | 0.0 | 1.0 | 4.0 | 4.0 |
| 3 | 2.0 | 0.0 | 1.0 | 4.0 | 3.0 |
| 4 | 2.0 | 1.0 | 2.0 | 3.0 | 3.0 |

| | PRAEGENDE_JUGENDJAHRE | REGIOTYP | RELAT_AB | RETOURTYP_BK_S | RT_KEIN_ANREIZ \ |
|---|-----------------------|----------|----------|----------------|------------------|
| 0 | 0 | NaN | NaN | 5.0 | 1.0 |
| 1 | 14 | 3.0 | 4.0 | 1.0 | 5.0 |
| 2 | 15 | 2.0 | 2.0 | 3.0 | 5.0 |
| 3 | 8 | 0.0 | 3.0 | 2.0 | 3.0 |
| 4 | 8 | 5.0 | 5.0 | 5.0 | 3.0 |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL | SEMIO_FAM \ |
|---|-----------------|-----------------|-----------|-----------|-------------|
| 0 | 4.0 | 1.0 | 6 | 3 | 6 |
| 1 | 3.0 | 5.0 | 7 | 2 | 4 |
| 2 | 4.0 | 5.0 | 7 | 6 | 1 |
| 3 | 2.0 | 3.0 | 4 | 7 | 1 |
| 4 | 5.0 | 5.0 | 2 | 4 | 4 |

| | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT \ |
|---|------------|------------|------------|------------|-----------|-----------------|
| 0 | 6 | 7 | 3 | 5 | 5 | 5 |
| 1 | 4 | 4 | 3 | 2 | 3 | 7 |
| 2 | 7 | 7 | 3 | 4 | 3 | 3 |
| 3 | 5 | 4 | 4 | 4 | 1 | 4 |
| 4 | 2 | 3 | 6 | 4 | 2 | 4 |

| | SEMIO_RAT | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT | SHOPPER_TYP \ |
|---|-----------|-----------|-----------|-------------|------------|---------------|
| 0 | 4 | 7 | 2 | 3 | 1 | -1 |
| 1 | 6 | 4 | 5 | 6 | 1 | 3 |
| 2 | 4 | 3 | 4 | 3 | 4 | 2 |
| 3 | 3 | 2 | 5 | 4 | 4 | 1 |
| 4 | 2 | 4 | 6 | 2 | 7 | 2 |

| | SOHO_KZ | STRUKTURTYP | TITEL_KZ | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG | \ |
|---|---------|-------------|----------|------------|-------------|------------------|---|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 1 | 1.0 | 2.0 | 0.0 | 3.0 | 3.0 | 1.0 | |
| 2 | 0.0 | 3.0 | 0.0 | 2.0 | 5.0 | 0.0 | |
| 3 | 0.0 | 1.0 | 0.0 | 4.0 | 5.0 | 0.0 | |
| 4 | 0.0 | 3.0 | 0.0 | 4.0 | 3.0 | 0.0 | |

| | VERDICHTUNGSRAUM | VERS_TYP | VHA | VHN | VK_DHT4A | VK_DISTANZ | VK_ZG11 | \ |
|---|------------------|----------|-----|-----|----------|------------|---------|---|
| 0 | NaN | -1 | NaN | NaN | NaN | NaN | NaN | |
| 1 | 0.0 | 2 | 0.0 | 4.0 | 8.0 | 11.0 | 10.0 | |
| 2 | 1.0 | 1 | 0.0 | 2.0 | 9.0 | 9.0 | 6.0 | |
| 3 | 0.0 | 1 | 1.0 | 0.0 | 7.0 | 10.0 | 11.0 | |
| 4 | 1.0 | 2 | 0.0 | 2.0 | 3.0 | 5.0 | 4.0 | |

| | W_KEIT_KIND_HH | WOHNDAUER_2008 | WOHNLAGE | ZABEOTYP | ANREDE_KZ | \ |
|---|----------------|----------------|----------|----------|-----------|---|
| 0 | NaN | NaN | NaN | 3 | 1 | |
| 1 | 3.0 | 9.0 | 4.0 | 5 | 2 | |
| 2 | 3.0 | 9.0 | 2.0 | 5 | 2 | |
| 3 | NaN | 9.0 | 7.0 | 3 | 2 | |
| 4 | 2.0 | 9.0 | 3.0 | 4 | 1 | |

| | ALTERSKATEGORIE_GROB |
|---|----------------------|
| 0 | 2 |
| 1 | 1 |
| 2 | 3 |
| 3 | 4 |
| 4 | 3 |

In [3]: 'Shape', azdias.shape

Out[3]: ('Shape', (891221, 366))

In [4]: azdias.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891221 entries, 0 to 891220
Columns: 366 entries, LNR to ALTERSKATEGORIE_GROB
dtypes: float64(267), int64(93), object(6)
memory usage: 2.4+ GB
```

In [5]: azdias.describe()

| Out[5]: | LNR | AGER_TYP | AKT_DAT_KL | ALTER_HH | \ |
|---------|--------------|---------------|---------------|---------------|---|
| count | 8.912210e+05 | 891221.000000 | 817722.000000 | 817722.000000 | |
| mean | 6.372630e+05 | -0.358435 | 4.421928 | 10.864126 | |
| std | 2.572735e+05 | 1.198724 | 3.638805 | 7.639683 | |
| min | 1.916530e+05 | -1.000000 | 1.000000 | 0.000000 | |
| 25% | 4.144580e+05 | -1.000000 | 1.000000 | 0.000000 | |

| | | | | |
|-----|--------------|-----------|----------|-----------|
| 50% | 6.372630e+05 | -1.000000 | 3.000000 | 13.000000 |
| 75% | 8.600680e+05 | -1.000000 | 9.000000 | 17.000000 |
| max | 1.082873e+06 | 3.000000 | 9.000000 | 21.000000 |

| | | | | |
|-------|--------------|--------------|-------------|---------------|
| | ALTER_KIND1 | ALTER_KIND2 | ALTER_KIND3 | ALTER_KIND4 \ |
| count | 81058.000000 | 29499.000000 | 6170.000000 | 1205.000000 |
| mean | 11.745392 | 13.402658 | 14.476013 | 15.089627 |
| std | 4.097660 | 3.243300 | 2.712427 | 2.452932 |
| min | 2.000000 | 2.000000 | 4.000000 | 7.000000 |
| 25% | 8.000000 | 11.000000 | 13.000000 | 14.000000 |
| 50% | 12.000000 | 14.000000 | 15.000000 | 15.000000 |
| 75% | 15.000000 | 16.000000 | 17.000000 | 17.000000 |
| max | 18.000000 | 18.000000 | 18.000000 | 18.000000 |

| | | | |
|-------|----------------------|---------------------|----------------|
| | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | ANZ_HH_TITEL \ |
| count | 628274.000000 | 798073.000000 | 794213.000000 |
| mean | 13.700717 | 8.287263 | 0.040647 |
| std | 5.079849 | 15.628087 | 0.324028 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 11.000000 | 1.000000 | 0.000000 |
| 50% | 14.000000 | 4.000000 | 0.000000 |
| 75% | 17.000000 | 9.000000 | 0.000000 |
| max | 25.000000 | 595.000000 | 23.000000 |

| | | | |
|-------|---------------|---------------|------------------------------|
| | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE \ |
| count | 817722.000000 | 817722.000000 | 798073.000000 |
| mean | 0.154018 | 1.727637 | 7.599356 |
| std | 0.502389 | 1.155849 | 14.332201 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 1.000000 | 1.000000 |
| 50% | 0.000000 | 1.000000 | 3.000000 |
| 75% | 0.000000 | 2.000000 | 9.000000 |
| max | 11.000000 | 45.000000 | 449.000000 |

| | | | | |
|-------|---------------|---------------|---------------|-----------------|
| | ANZ_TITEL | ARBEIT | BALLRAUM | CJT_GESAMTTYP \ |
| count | 817722.000000 | 794005.000000 | 797481.000000 | 886367.000000 |
| mean | 0.004162 | 3.167854 | 4.153043 | 3.632838 |
| std | 0.068855 | 1.002376 | 2.183710 | 1.595021 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 3.000000 | 2.000000 | 2.000000 |
| 50% | 0.000000 | 3.000000 | 5.000000 | 4.000000 |
| 75% | 0.000000 | 4.000000 | 6.000000 | 5.000000 |
| max | 6.000000 | 9.000000 | 7.000000 | 6.000000 |

| | | | | |
|-------|-------------------|---------------|---------------|---------------|
| | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 \ |
| count | 886367.000000 | 886367.000000 | 886367.000000 | 886367.000000 |
| mean | 3.335264 | 3.368086 | 3.195014 | 3.351290 |
| std | 1.493633 | 1.368331 | 1.401382 | 1.396508 |

| | | | | |
|-----|----------|----------|----------|----------|
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 4.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 5.000000 | 5.000000 | 5.000000 | 5.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 \ |
|-------|---------------|---------------|---------------|---------------------|
| count | 886367.000000 | 886367.000000 | 886367.000000 | 891221.000000 |
| mean | 3.336151 | 3.360684 | 3.465980 | 0.122336 |
| std | 1.373077 | 1.378992 | 1.328456 | 0.535950 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 0.000000 |
| 50% | 3.000000 | 3.000000 | 4.000000 | 0.000000 |
| 75% | 5.000000 | 5.000000 | 5.000000 | 0.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 6.000000 |

| | D19_BANKEN_ANZ_24 | D19_BANKEN_DATUM | D19_BANKEN_DIREKT \ |
|-------|-------------------|------------------|---------------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.219907 | 9.267420 | 0.892735 |
| std | 0.747903 | 1.735725 | 2.011838 |
| min | 0.000000 | 1.000000 | 0.000000 |
| 25% | 0.000000 | 10.000000 | 0.000000 |
| 50% | 0.000000 | 10.000000 | 0.000000 |
| 75% | 0.000000 | 10.000000 | 0.000000 |
| max | 6.000000 | 10.000000 | 7.000000 |

| | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | D19_BANKEN_OFFLINE_DATUM \ |
|-------|------------------|------------------|----------------------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.568580 | 0.106769 | 9.926794 |
| std | 1.643764 | 0.808179 | 0.605641 |
| min | 0.000000 | 0.000000 | 1.000000 |
| 25% | 0.000000 | 0.000000 | 10.000000 |
| 50% | 0.000000 | 0.000000 | 10.000000 |
| 75% | 0.000000 | 0.000000 | 10.000000 |
| max | 6.000000 | 7.000000 | 10.000000 |

| | D19_BANKEN_ONLINE_DATUM | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST \ |
|-------|-------------------------|----------------------------|-------------------|
| count | 891221.000000 | 634108.000000 | 891221.000000 |
| mean | 9.439073 | 0.705221 | 0.425645 |
| std | 1.547773 | 2.552707 | 1.510782 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | D19_BEKLEIDUNG_GEH | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO \ |
|-------|--------------------|---------------------|---------------|----------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 0.467869 | 1.145516 | 0.485508 | 0.257938 |
| std | 1.542151 | 2.266999 | 1.639340 | 1.252328 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_BUCH_CD | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE \ |
|-------|---------------|----------------|---------------------|---------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 1.585405 | 0.196746 | 0.673292 | 0.346854 |
| std | 2.547925 | 1.036762 | 1.752758 | 1.337269 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 3.000000 | 0.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_FREIZEIT | D19_GARTEN | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 \ |
|-------|---------------|---------------|-------------------|---------------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.590678 | 0.255993 | 0.797683 | 1.240179 |
| std | 1.721602 | 1.213048 | 1.330718 | 1.727867 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 0.000000 | 1.000000 | 2.000000 |
| max | 7.000000 | 7.000000 | 6.000000 | 6.000000 |

| | D19_GESAMT_DATUM | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM \ |
|-------|------------------|--------------------------|---------------------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 7.071743 | 9.034676 | 7.680381 |
| std | 3.212715 | 1.768926 | 3.039867 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 5.000000 | 9.000000 | 5.000000 |
| 50% | 9.000000 | 10.000000 | 10.000000 |
| 75% | 10.000000 | 10.000000 | 10.000000 |
| max | 10.000000 | 10.000000 | 10.000000 |

| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO \ |
|-------|----------------------------|---------------|-----------------|
| count | 634108.000000 | 891221.000000 | 891221.000000 |
| mean | 3.560952 | 0.841994 | 1.017852 |
| std | 4.658538 | 2.119703 | 2.144350 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 7.000000 | 7.000000 |

| | | | | |
|-------|-------------------|---------------|-------------------|----------------|
| | D19_KINDERARTIKEL | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK \ |
| count | 891221.000000 | 634108.000000 | 891221.000000 | 891221.000000 |
| mean | 0.870657 | 5.424540 | 5.849228 | 1.030101 |
| std | 2.077303 | 3.234275 | 3.225762 | 2.347223 |
| min | 0.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 0.000000 | 2.000000 | 2.000000 | 0.000000 |
| 50% | 0.000000 | 5.000000 | 8.000000 | 0.000000 |
| 75% | 0.000000 | 9.000000 | 9.000000 | 0.000000 |
| max | 7.000000 | 9.000000 | 9.000000 | 7.000000 |

| | | | | |
|-------|------------------|---------------|-------------------------|----------------|
| | D19_LEBENSMITTEL | D19_LOTTO | D19_NAHRUNGSERGAEANZUNG | D19_RATGEBER \ |
| count | 891221.000000 | 634108.000000 | 891221.000000 | 891221.000000 |
| mean | 0.319525 | 1.521733 | 0.244524 | 0.502830 |
| std | 1.316090 | 2.832521 | 1.184654 | 1.616871 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | | | | |
|-------|---------------|-------------------|---------------|----------------|
| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE \ |
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 1.048628 | 0.594606 | 0.510818 | 2.365959 |
| std | 2.329347 | 1.798291 | 1.433293 | 2.861577 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 0.000000 | 0.000000 | 6.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | | | | |
|-------|---------------|---------------|------------------|--------------------|
| | D19_SOZIALES | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 \ |
| count | 634108.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.657908 | 1.738563 | 0.049056 | 0.098804 |
| std | 1.457774 | 2.740417 | 0.277552 | 0.393587 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 6.000000 | 0.000000 | 0.000000 |
| max | 5.000000 | 7.000000 | 6.000000 | 6.000000 |

| | | | |
|-------|-----------------|------------------|---------------------------|
| | D19_TELKO_DATUM | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM \ |
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 9.428728 | 1.009937 | 9.828039 |
| std | 1.344289 | 2.187102 | 0.745852 |
| min | 1.000000 | 0.000000 | 1.000000 |
| 25% | 9.000000 | 0.000000 | 10.000000 |
| 50% | 10.000000 | 0.000000 | 10.000000 |
| 75% | 10.000000 | 0.000000 | 10.000000 |

| | | | |
|-----|-----------|----------|-----------|
| max | 10.000000 | 7.000000 | 10.000000 |
|-----|-----------|----------|-----------|

| | | | |
|-------|------------------------|---------------------------|------------------|
| | D19_TELKO_ONLINE_DATUM | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST \ |
| count | 891221.000000 | 634108.000000 | 891221.000000 |
| mean | 9.981780 | 0.012261 | 0.779996 |
| std | 0.241035 | 0.348780 | 1.969724 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | | | |
|-------|-----------------|--------------------|----------------------|
| | D19_TIERARTIKEL | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 \ |
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.243178 | 0.604646 | 0.958868 |
| std | 1.170809 | 1.150455 | 1.529453 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 1.000000 | 2.000000 |
| max | 7.000000 | 6.000000 | 6.000000 |

| | | | |
|-------|-------------------|---------------------------|----------------------------|
| | D19_VERSAND_DATUM | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM \ |
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 7.717380 | 9.326818 | 7.942473 |
| std | 2.989552 | 1.456007 | 2.942631 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 5.000000 | 9.000000 | 6.000000 |
| 50% | 9.000000 | 10.000000 | 10.000000 |
| 75% | 10.000000 | 10.000000 | 10.000000 |
| max | 10.000000 | 10.000000 | 10.000000 |

| | | | |
|-------|-----------------------------|------------------|--------------------|
| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 \ |
| count | 634108.000000 | 891221.000000 | 891221.000000 |
| mean | 3.244432 | 0.829550 | 0.113597 |
| std | 4.586604 | 1.912499 | 0.434877 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 7.000000 | 6.000000 |

| | | | |
|-------|------------------|-----------------|---------------------------|
| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM \ |
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 0.206998 | 9.142563 | 9.922649 |
| std | 0.617016 | 1.911186 | 0.502665 |
| min | 0.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 9.000000 | 10.000000 |

| | | | |
|-----|----------|-----------|-----------|
| 50% | 0.000000 | 10.000000 | 10.000000 |
| 75% | 0.000000 | 10.000000 | 10.000000 |
| max | 6.000000 | 10.000000 | 10.000000 |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN \ |
|-------|------------------------|---------------------------|----------------------|
| count | 891221.000000 | 634108.000000 | 891221.000000 |
| mean | 9.976712 | 0.025228 | 1.259786 |
| std | 0.311191 | 0.498010 | 2.253954 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 2.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG \ |
|-------|-------------------|-------------------|---------------|
| count | 891221.000000 | 891221.000000 | 798073.000000 |
| mean | 1.728497 | 0.377515 | 0.967816 |
| std | 2.605193 | 1.493151 | 0.176488 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 1.000000 |
| 50% | 0.000000 | 0.000000 | 1.000000 |
| 75% | 5.000000 | 0.000000 | 1.000000 |
| max | 7.000000 | 7.000000 | 1.000000 |

| | EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER \ |
|-------|----------------------|---------------|---------------|------------------|
| count | 817722.000000 | 797481.000000 | 237068.000000 | 891221.000000 |
| mean | 2003.729061 | 3.939172 | 33.338392 | 3.033328 |
| std | 7.058204 | 1.718996 | 14.537408 | 1.529603 |
| min | 1900.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 1997.000000 | 2.000000 | 23.000000 | 2.000000 |
| 50% | 2003.000000 | 4.000000 | 34.000000 | 3.000000 |
| 75% | 2010.000000 | 6.000000 | 43.000000 | 5.000000 |
| max | 2018.000000 | 6.000000 | 56.000000 | 5.000000 |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER \ |
|-------|------------------|-------------------|-----------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 3.075121 | 3.074528 | 2.821039 |
| std | 1.353248 | 1.321055 | 1.464749 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | FINANZ_UNAUFFAELLIGER | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE \ |
|-------|-----------------------|------------------|---------------|----------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 798066.000000 |
| mean | 2.874167 | 3.401106 | 3.790586 | 3.456029 |
| std | 1.486731 | 1.322134 | 1.987876 | 1.153415 |

| | | | | |
|-----|----------|----------|----------|----------|
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 4.000000 | 4.000000 |
| 75% | 4.000000 | 5.000000 | 6.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 6.000000 | 5.000000 |

| | GEBAEUDETYP | GEBAEUDETYP_RASTER | GEBURTSJAHR | GEMEINDETYP \ |
|-------|---------------|--------------------|---------------|---------------|
| count | 798073.000000 | 798066.000000 | 891221.000000 | 793947.000000 |
| mean | 2.798641 | 3.738306 | 1101.178533 | 24.186748 |
| std | 2.656713 | 0.923193 | 976.583551 | 12.037852 |
| min | 1.000000 | 1.000000 | 0.000000 | 11.000000 |
| 25% | 1.000000 | 3.000000 | 0.000000 | 12.000000 |
| 50% | 1.000000 | 4.000000 | 1943.000000 | 22.000000 |
| 75% | 3.000000 | 4.000000 | 1970.000000 | 30.000000 |
| max | 8.000000 | 5.000000 | 2017.000000 | 50.000000 |

| | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP | HH_DELTA_FLAG \ |
|-------|-----------------|------------------|---------------|-----------------|
| count | 886367.000000 | 891221.000000 | 891221.000000 | 783619.000000 |
| mean | 7.350304 | 0.196612 | 1.792102 | 0.092745 |
| std | 3.525723 | 0.397437 | 1.269062 | 0.290075 |
| min | 1.000000 | 0.000000 | -1.000000 | 0.000000 |
| 25% | 5.000000 | 0.000000 | 1.000000 | 0.000000 |
| 50% | 8.000000 | 0.000000 | 2.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 3.000000 | 0.000000 |
| max | 12.000000 | 1.000000 | 3.000000 | 1.000000 |

| | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 \ |
|-------|--------------------|---------------|---------------|----------------|
| count | 872873.000000 | 797481.000000 | 757897.000000 | 757897.000000 |
| mean | 4.207243 | 4.549491 | 2.071317 | 3.149416 |
| std | 1.624057 | 2.028919 | 1.532120 | 1.338676 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 3.000000 | 3.000000 | 1.000000 | 2.000000 |
| 50% | 5.000000 | 5.000000 | 2.000000 | 3.000000 |
| 75% | 6.000000 | 6.000000 | 3.000000 | 4.000000 |
| max | 6.000000 | 8.000000 | 9.000000 | 9.000000 |

| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 3.112196 | 2.919489 | 1.106609 | 1.494277 |
| std | 1.349705 | 1.500372 | 1.427606 | 1.403961 |
| min | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 0.000000 | 0.000000 |
| 50% | 3.000000 | 3.000000 | 1.000000 | 1.000000 |
| 75% | 4.000000 | 4.000000 | 1.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 4.000000 |

| | KBA05_ANTG2 | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT \ |
|-------|---------------|---------------|---------------|------------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 1.265584 | 0.624525 | 0.305927 | 3.207994 |
| std | 1.245178 | 1.013443 | 0.638725 | 1.400238 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 2.000000 |
| 50% | 1.000000 | 0.000000 | 0.000000 | 3.000000 |
| 75% | 2.000000 | 1.000000 | 0.000000 | 4.000000 |
| max | 4.000000 | 3.000000 | 2.000000 | 9.000000 |

| | KBA05_BAUMAX | KBA05_CCM1 | KBA05_CCM2 | KBA05_CCM3 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 1.389552 | 3.082453 | 3.115361 | 3.144479 |
| std | 1.779483 | 1.349763 | 1.323141 | 1.351333 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 1.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ \ |
|-------|---------------|---------------|---------------|---------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 1.358658 | 2.147870 | 3.110023 | 3.158580 |
| std | 1.621032 | 1.427566 | 1.349794 | 1.329537 |
| min | 0.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 1.000000 | 2.000000 | 2.000000 |
| 50% | 1.000000 | 2.000000 | 3.000000 | 3.000000 |
| 75% | 2.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 5.000000 |

| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 \ |
|-------|---------------|---------------|---------------|----------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 2.476347 | 3.103159 | 3.035828 | 2.955032 |
| std | 1.635226 | 1.316661 | 1.394468 | 1.496686 |
| min | 0.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 2.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_HERST5 | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 \ |
|-------|---------------|-----------------|----------------|-------------------|
| count | 757897.000000 | 798073.000000 | 757897.000000 | 757897.000000 |
| mean | 2.923587 | 2.836532 | 3.097802 | 3.049426 |
| std | 1.539973 | 1.491578 | 1.396567 | 1.308095 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_KRSHERST2 | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER \ |
|-------|-----------------|-----------------|----------------|-----------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 3.078738 | 3.153802 | 2.127912 | 2.102949 |
| std | 1.345043 | 1.362345 | 1.159441 | 1.145852 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 2.000000 | 2.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 | 2.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_KRSVAN | KBA05_KRSZUL | KBA05_KW1 | KBA05_KW2 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 2.135488 | 2.065225 | 3.093550 | 3.113674 |
| std | 1.126434 | 1.196110 | 1.376706 | 1.333745 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 2.000000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 3.000000 | 3.000000 |
| 75% | 2.000000 | 2.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST \ |
|-------|---------------|---------------|---------------|------------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 1.551242 | 3.386927 | 2.444595 | 2.869010 |
| std | 1.607973 | 1.493682 | 1.496391 | 1.403278 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 2.000000 | 1.000000 | 2.000000 |
| 50% | 1.000000 | 3.000000 | 2.000000 | 3.000000 |
| 75% | 2.000000 | 5.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 2.278277 | 2.218266 | 1.437354 | 3.091425 |
| std | 1.318311 | 1.209227 | 1.643943 | 1.328794 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 2.000000 | 0.000000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 1.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 2.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | KBA05_MOD3 | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP \ |
|-------|---------------|---------------|---------------|-----------------|
| count | 757897.000000 | 757897.000000 | 757897.000000 | 798073.000000 |
| mean | 3.096057 | 2.832567 | 1.379413 | 3.006467 |
| std | 1.350524 | 1.612727 | 1.464159 | 1.255616 |
| min | 1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 0.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 1.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 | 4.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| max | 9.000000 | 9.000000 | 9.000000 | 6.000000 |
|-----|----------|----------|----------|----------|

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_MOTOR | KBA05_MOTRAD | KBA05_SEG1 | KBA05_SEG10 \ |
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 2.649657 | 1.163318 | 1.233541 | 2.005177 |
| std | 1.287217 | 1.359159 | 1.441903 | 1.508828 |
| min | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 0.000000 | 0.000000 | 1.000000 |
| 50% | 3.000000 | 1.000000 | 1.000000 | 2.000000 |
| 75% | 3.000000 | 1.000000 | 2.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 \ |
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 3.098730 | 3.086683 | 3.104275 | 1.597992 |
| std | 1.332421 | 1.343910 | 1.317709 | 1.567611 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 1.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 | 2.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 \ |
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 0.292329 | 0.988600 | 0.901281 | 1.213546 |
| std | 1.269465 | 1.477921 | 1.472815 | 1.445609 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 1.000000 | 0.000000 | 1.000000 |
| 75% | 0.000000 | 2.000000 | 1.000000 | 2.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_VORB0 | KBA05_VORB1 | KBA05_VORB2 | KBA05_ZUL1 \ |
| count | 757897.000000 | 757897.000000 | 757897.000000 | 757897.000000 |
| mean | 2.970043 | 3.111915 | 2.858849 | 3.101210 |
| std | 1.417498 | 1.333840 | 1.616102 | 1.336639 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|------------------------|
| | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | KBA13_ALTERHALTER_30 \ |
| count | 757897.000000 | 757897.000000 | 757897.000000 | 785421.000000 |
| mean | 3.105024 | 2.782602 | 2.270413 | 2.978739 |
| std | 1.341038 | 1.614416 | 1.733541 | 1.061121 |
| min | 1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 | 2.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| 50% | 3.000000 | 3.000000 | 2.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 5.000000 |

| | | | | |
|-------|----------------------|----------------------|----------------------|---|
| | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | KBA13_ALTERHALTER_61 | \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 3.084179 | 2.818780 | 3.119899 | |
| std | 1.114805 | 1.054657 | 1.065411 | |
| min | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 3.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 | \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 2.150965 | 2.777878 | 1.604806 | 0.676414 | |
| std | 0.921052 | 0.932685 | 1.017783 | 0.735776 | |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 | |
| 25% | 1.000000 | 2.000000 | 1.000000 | 0.000000 | |
| 50% | 2.000000 | 3.000000 | 2.000000 | 1.000000 | |
| 75% | 3.000000 | 3.000000 | 2.000000 | 1.000000 | |
| max | 4.000000 | 4.000000 | 3.000000 | 2.000000 | |

| | | | | | |
|-------|------------------|---------------|-----------------|---------------|---|
| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX | \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 619.701439 | 3.005443 | 2.782243 | 2.019615 | |
| std | 340.034318 | 1.003598 | 1.049191 | 1.510167 | |
| min | 0.000000 | 1.000000 | 0.000000 | 1.000000 | |
| 25% | 384.000000 | 2.000000 | 2.000000 | 1.000000 | |
| 50% | 549.000000 | 3.000000 | 3.000000 | 1.000000 | |
| 75% | 778.000000 | 4.000000 | 3.000000 | 3.000000 | |
| max | 2300.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA13_BJ_1999 | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 | \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 2.987050 | 2.990920 | 2.982818 | 2.983200 | |
| std | 0.962333 | 0.997081 | 0.950731 | 0.964752 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|------------------|---|
| | KBA13_BJ_2008 | KBA13_BJ_2009 | KBA13_BMW | KBA13_CCM_0_1400 | \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 2.430025 | 2.453592 | 3.159222 | 2.342429 | |
| std | 1.466413 | 1.448965 | 1.000801 | 1.445697 | |

| | | | | |
|-----|----------|----------|----------|----------|
| min | 0.000000 | 0.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 1.000000 | 3.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|----------------|----------------|----------------|-----------------------|
| | KBA13_CCM_1000 | KBA13_CCM_1200 | KBA13_CCM_1400 | KBA13_CCM_1401_2500 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.426497 | 2.329706 | 3.005170 | 2.910121 |
| std | 1.436357 | 1.459578 | 0.942756 | 0.945072 |
| min | 0.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 1.000000 | 1.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|----------------|----------------|----------------|------------------|
| | KBA13_CCM_1500 | KBA13_CCM_1600 | KBA13_CCM_1800 | KBA13_CCM_2000 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.619311 | 3.029111 | 2.364001 | 3.071686 |
| std | 1.419685 | 0.938871 | 1.436662 | 0.938193 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|----------------|----------------|----------------|------------------|
| | KBA13_CCM_2500 | KBA13_CCM_2501 | KBA13_CCM_3000 | KBA13_CCM_3001 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.482918 | 2.512178 | 2.583847 | 2.625875 |
| std | 1.414943 | 1.449423 | 1.368104 | 1.511502 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|-----------------|--------------------|---------------|---------------|
| | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE | KBA13_FIAT | KBA13_FORD \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.974637 | 2.988924 | 3.128779 | 3.004055 |
| std | 1.017911 | 1.008135 | 1.001766 | 1.040267 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|-----------------|-----------------|-------------------|
| | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 | KBA13_HALTER_30 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 3.419656 | 2.909913 | 2.945092 | 3.056838 |
| std | 1.119775 | 1.001968 | 1.036591 | 1.089204 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 | KBA13_HALTER_50 | \ |
|-------|-----------------|-----------------|-----------------|-----------------|---|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 3.104942 | 3.089957 | 3.027353 | 2.844840 | |
| std | 1.105378 | 1.097062 | 1.078913 | 1.053112 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 3.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | KBA13_HALTER_66 | \ |
|-------|-----------------|-----------------|-----------------|-----------------|---|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 2.840152 | 2.889783 | 3.126904 | 3.103431 | |
| std | 1.062277 | 1.073687 | 1.044822 | 1.075110 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 3.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 3.000000 | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | KBA13_HERST_BMW_BENZ | \ |
|-------|-------------------|---------------------|----------------------|---|
| count | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 3.004371 | 2.924983 | 3.165170 | |
| std | 1.031193 | 1.026814 | 1.022937 | |
| min | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 3.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | |

| | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | KBA13_HERST_SONST | \ |
|-------|--------------------|-----------------------|-------------------|---|
| count | 785421.000000 | 785421.000000 | 785421.000000 | |
| mean | 3.083781 | 2.966992 | 2.988924 | |
| std | 1.013485 | 1.058178 | 1.008135 | |
| min | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | |

| | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | KBA13_KMH_140 \ |
|-------|---------------|-----------------|---------------|-----------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.573068 | 2.329369 | 1.320813 | 2.694810 |
| std | 0.978024 | 1.529004 | 0.676535 | 1.390453 |
| min | 1.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 1.000000 | 1.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 1.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 1.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 3.000000 | 5.000000 |

| | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | KBA13_KMH_211 \ |
|-------|-------------------|---------------|---------------|-----------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.849393 | 2.932968 | 3.037302 | 2.457977 |
| std | 0.971055 | 0.970028 | 0.955909 | 1.491628 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | KBA13_KRSHERST_AUDI_VW \ |
|-------|---------------|---------------|----------------|--------------------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.455954 | 1.268960 | 2.866368 | 2.959944 |
| std | 1.490854 | 0.672798 | 1.062772 | 1.011523 |
| min | 0.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 1.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 1.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 1.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 3.000000 | 5.000000 | 5.000000 |

| | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | KBA13_KRSSEG_KLEIN \ |
|-------|-------------------------|--------------------------|----------------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.083096 | 3.008966 | 1.994527 |
| std | 1.009635 | 1.040945 | 0.292661 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 2.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 |
| max | 5.000000 | 5.000000 | 3.000000 |

| | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | KBA13_KW_0_60 \ |
|-------|-------------------|------------------|------------------|-----------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 1.954656 | 1.945301 | 1.832946 | 2.977747 |
| std | 0.584313 | 0.615499 | 0.781109 | 0.968180 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 75% | 2.000000 | 2.000000 | 2.000000 | 4.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| max | 3.000000 | 3.000000 | 3.000000 | 5.000000 |
|-----|----------|----------|----------|----------|

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.404903 | 2.376123 | 2.506342 | 1.404983 |
| std | 1.429198 | 1.505370 | 1.437424 | 0.681106 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 1.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 1.000000 |
| 75% | 3.000000 | 3.000000 | 3.000000 | 2.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 3.000000 |

| | | | | |
|-------|---------------|---------------|---------------|-------------------|
| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.405066 | 2.340691 | 2.291351 | 2.996188 |
| std | 1.412106 | 1.441927 | 1.411954 | 0.957387 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 2.000000 | 1.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA13_KW_70 | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.336341 | 2.310013 | 2.382884 | 3.085354 |
| std | 1.430272 | 1.410661 | 1.430076 | 0.999931 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|----------------|---------------|---------------|---------------|
| | KBA13_MERCEDES | KBA13_MOTOR | KBA13_NISSAN | KBA13_OPEL \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.144134 | 2.786029 | 3.060305 | 2.971935 |
| std | 1.023847 | 0.768496 | 1.018650 | 1.054774 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 4.000000 | 5.000000 | 5.000000 |

| | | | |
|-------|---------------|---------------|---------------------------|
| | KBA13_PEUGEOT | KBA13_RENAULT | KBA13_SEG_GELAENDEWAGEN \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.071690 | 3.043498 | 2.910411 |
| std | 1.009349 | 1.023270 | 1.001290 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |

| | | | |
|-----|----------|----------|----------|
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | | |
|-------|-------------------------|-------------------|------------------------|
| | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST | KBA13_SEG_KLEINWAGEN \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.076376 | 2.989147 | 2.945591 |
| std | 1.013363 | 1.027775 | 1.016326 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | | |
|-------|-------------------------|--------------------|-----------------------|
| | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS | KBA13_SEG_MINIWAGEN \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.950019 | 3.01906 | 3.117099 |
| std | 1.017614 | 1.01522 | 1.011775 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | |
|-------|------------------------|-------------------------------|
| | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE \ |
| count | 785421.000000 | 785421.000000 |
| mean | 3.058855 | 3.161382 |
| std | 1.026028 | 1.009669 |
| min | 1.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 |

| | | | |
|-------|----------------------|--------------------|------------------------|
| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 2.496483 | 3.071439 | 2.566942 |
| std | 1.479816 | 0.963400 | 1.437284 |
| min | 0.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | | |
|-------|---------------------|---------------|------------------------|
| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE \ |
| count | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.016091 | 3.047092 | 2.538733 |
| std | 0.998701 | 1.011615 | 1.411366 |

| | | | |
|-----|----------|----------|----------|
| min | 1.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_SITZE_4 | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA \ |
|-------|---------------|---------------|---------------|----------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.170897 | 2.842810 | 3.071781 | 3.074966 |
| std | 1.053637 | 1.056045 | 1.041083 | 1.007389 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_VORB_0 | KBA13_VORB_1 | KBA13_VORB_1_2 | KBA13_VORB_2 \ |
|-------|---------------|---------------|----------------|----------------|
| count | 785421.000000 | 785421.000000 | 785421.000000 | 785421.000000 |
| mean | 3.117557 | 2.980660 | 2.916887 | 3.011097 |
| std | 0.975495 | 0.957392 | 0.961899 | 0.943805 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK \ |
|-------|---------------|---------------|---------------|---------------|
| count | 785421.000000 | 785421.000000 | 306609.000000 | 770025.000000 |
| mean | 2.354173 | 2.935193 | 3.410640 | 2.592991 |
| std | 1.469417 | 1.028888 | 1.628844 | 1.119052 |
| min | 0.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 5.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 6.000000 | 4.000000 |

| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN \ |
|-------|---------------|---------------|---------------|-------------------|
| count | 891221.000000 | 817252.000000 | 798066.000000 | 886367.000000 |
| mean | 3.517955 | 3.018452 | 0.236165 | 3.599574 |
| std | 2.110462 | 1.550312 | 0.424725 | 3.926486 |
| min | 1.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 0.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 0.000000 | 1.000000 |
| 75% | 4.000000 | 4.000000 | 0.000000 | 8.000000 |
| max | 9.000000 | 7.000000 | 1.000000 | 11.000000 |

| | LP_FAMILIE_GROB | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB \ |
|-------|-----------------|---------------------|-----------------------|
| count | 886367.000000 | 886367.000000 | 886367.000000 |

| | | | |
|------|----------|-----------|-----------|
| mean | 2.185966 | 14.622637 | 4.453621 |
| std | 1.756537 | 12.616883 | 3.855639 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 1.000000 | 4.000000 | 1.000000 |
| 50% | 1.000000 | 11.000000 | 3.000000 |
| 75% | 4.000000 | 27.000000 | 8.000000 |
| max | 5.000000 | 40.000000 | 12.000000 |

| | LP_STATUS_FEIN | LP_STATUS_GROB | MIN_GEBAEUDEJAHR | MOBI_RASTER \ |
|-------|----------------|----------------|------------------|---------------|
| count | 886367.000000 | 886367.000000 | 798073.000000 | 798073.000000 |
| mean | 4.791151 | 2.432575 | 1993.277011 | 2.378819 |
| std | 3.425305 | 1.474315 | 3.332739 | 1.536927 |
| min | 1.000000 | 1.000000 | 1985.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 1992.000000 | 1.000000 |
| 50% | 4.000000 | 2.000000 | 1992.000000 | 2.000000 |
| 75% | 9.000000 | 4.000000 | 1993.000000 | 4.000000 |
| max | 10.000000 | 5.000000 | 2016.000000 | 6.000000 |

| | MOBI_REGIO | NATIONALITAET_KZ | ONLINE_AFFINITAET | ORTSGR_KLS9 \ |
|-------|---------------|------------------|-------------------|---------------|
| count | 757897.000000 | 891221.000000 | 886367.000000 | 794005.000000 |
| mean | 2.963540 | 1.026827 | 2.698691 | 5.293002 |
| std | 1.428882 | 0.586634 | 1.521524 | 2.303739 |
| min | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 1.000000 | 1.000000 | 4.000000 |
| 50% | 3.000000 | 1.000000 | 3.000000 | 5.000000 |
| 75% | 4.000000 | 1.000000 | 4.000000 | 7.000000 |
| max | 6.000000 | 3.000000 | 5.000000 | 9.000000 |

| | PLZ8_ANTG1 | PLZ8_ANTG2 | PLZ8_ANTG3 | PLZ8_ANTG4 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 774706.000000 | 774706.000000 | 774706.000000 | 774706.000000 |
| mean | 2.253330 | 2.801858 | 1.595426 | 0.699166 |
| std | 0.972008 | 0.920309 | 0.986736 | 0.727137 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 1.000000 | 0.000000 |
| 50% | 2.000000 | 3.000000 | 2.000000 | 1.000000 |
| 75% | 3.000000 | 3.000000 | 2.000000 | 1.000000 |
| max | 4.000000 | 4.000000 | 3.000000 | 2.000000 |

| | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ | PRAEGENDE_JUGENDJAHRE \ |
|-------|---------------|---------------|---------------|-------------------------|
| count | 774706.000000 | 774706.000000 | 774706.000000 | 891221.000000 |
| mean | 1.943913 | 3.381087 | 3.612821 | 8.154346 |
| std | 1.459654 | 1.111598 | 0.973967 | 4.844532 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 3.000000 | 3.000000 | 5.000000 |
| 50% | 1.000000 | 3.000000 | 4.000000 | 8.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 | 14.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 15.000000 |

| | REGIOTYP | RELAT_AB | RETOURTYT_BK_S | RT_KEIN_ANREIZ \ |
|-------|---------------|---------------|----------------|------------------|
| count | 770025.000000 | 794005.000000 | 886367.000000 | 886367.000000 |
| mean | 4.257967 | 3.07222 | 3.419630 | 3.233723 |
| std | 2.030385 | 1.36298 | 1.417741 | 1.388687 |
| min | 0.000000 | 1.00000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.00000 | 2.000000 | 2.000000 |
| 50% | 5.000000 | 3.00000 | 3.000000 | 3.000000 |
| 75% | 6.000000 | 4.00000 | 5.000000 | 4.000000 |
| max | 7.000000 | 9.00000 | 5.000000 | 5.000000 |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL \ |
|-------|-----------------|-----------------|---------------|---------------|
| count | 886367.000000 | 839995.000000 | 891221.000000 | 891221.000000 |
| mean | 3.863797 | 3.023813 | 4.667550 | 4.481405 |
| std | 1.282747 | 1.502725 | 1.795712 | 1.807552 |
| min | 1.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 3.000000 | 3.000000 |
| 50% | 4.000000 | 3.000000 | 5.000000 | 4.000000 |
| 75% | 5.000000 | 4.000000 | 6.000000 | 6.000000 |
| max | 5.000000 | 5.000000 | 7.000000 | 7.000000 |

| | SEMIO_FAM | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 4.272729 | 4.445007 | 4.763223 | 4.025014 |
| std | 1.915885 | 1.852412 | 1.830789 | 1.903816 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 50% | 4.000000 | 5.000000 | 5.000000 | 4.000000 |
| 75% | 6.000000 | 6.000000 | 6.000000 | 5.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT | SEMIO_RAT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 4.359086 | 4.001597 | 4.256076 | 3.910139 |
| std | 2.022829 | 1.857540 | 1.770137 | 1.580306 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 | 3.000000 |
| 50% | 5.000000 | 4.000000 | 4.000000 | 4.000000 |
| 75% | 6.000000 | 5.000000 | 6.000000 | 5.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 891221.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 4.240609 | 3.945860 | 3.661784 | 4.023709 |
| std | 2.007373 | 1.946564 | 1.707637 | 2.077746 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 4.000000 | 4.000000 | 3.000000 | 4.000000 |
| 75% | 6.000000 | 6.000000 | 5.000000 | 6.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |
|-----|----------|----------|----------|----------|

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | SHOPPER_TYP | SOHO_KZ | STRUKTURTYP | TITEL_KZ \ |
| count | 891221.000000 | 817722.000000 | 793947.000000 | 817722.000000 |
| mean | 1.266967 | 0.008423 | 2.539212 | 0.003483 |
| std | 1.287435 | 0.091392 | 0.754926 | 0.084957 |
| min | -1.000000 | 0.000000 | 1.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 2.000000 | 0.000000 |
| 50% | 1.000000 | 0.000000 | 3.000000 | 0.000000 |
| 75% | 2.000000 | 0.000000 | 3.000000 | 0.000000 |
| max | 3.000000 | 1.000000 | 3.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|------------------|--------------------|
| | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG | VERDICHTUNGSRAUM \ |
| count | 793435.000000 | 793435.000000 | 817722.000000 | 793947.000000 |
| mean | 3.223125 | 4.017526 | 0.090067 | 4.58576 |
| std | 1.255785 | 1.118399 | 0.286278 | 8.47152 |
| min | 1.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 3.000000 | 0.000000 | 0.000000 |
| 50% | 3.000000 | 4.000000 | 0.000000 | 1.000000 |
| 75% | 4.000000 | 5.000000 | 0.000000 | 5.000000 |
| max | 5.000000 | 5.000000 | 1.000000 | 45.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | VERS_TYP | VHA | VHN | VK_DHT4A \ |
| count | 891221.000000 | 817722.000000 | 770025.000000 | 815304.000000 |
| mean | 1.197852 | 0.43882 | 2.417322 | 6.001214 |
| std | 0.952532 | 1.14329 | 1.166572 | 2.856091 |
| min | -1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 0.000000 | 2.000000 | 3.000000 |
| 50% | 1.000000 | 0.000000 | 2.000000 | 6.000000 |
| 75% | 2.000000 | 0.000000 | 3.000000 | 9.000000 |
| max | 2.000000 | 5.000000 | 4.000000 | 11.000000 |

| | | | | |
|-------|---------------|---------------|----------------|------------------|
| | VK_DISTANZ | VK_ZG11 | W_KEIT_KIND_HH | WOHNDAUER_2008 \ |
| count | 815304.000000 | 815304.000000 | 783619.000000 | 817722.000000 |
| mean | 7.532130 | 5.945972 | 3.933406 | 7.908791 |
| std | 3.247789 | 2.771464 | 1.964701 | 1.923137 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 5.000000 | 4.000000 | 2.000000 | 8.000000 |
| 50% | 8.000000 | 6.000000 | 4.000000 | 9.000000 |
| 75% | 10.000000 | 8.000000 | 6.000000 | 9.000000 |
| max | 13.000000 | 11.000000 | 6.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|----------------------|
| | WOHNLAGE | ZABEOTYP | ANREDE_KZ | ALTERSKATEGORIE_GROB |
| count | 798073.000000 | 891221.000000 | 891221.000000 | 891221.000000 |
| mean | 4.052836 | 3.362438 | 1.522098 | 2.777398 |
| std | 1.949539 | 1.352704 | 0.499512 | 1.068775 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 3.000000 | 1.000000 | 2.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| 50% | 3.000000 | 3.000000 | 2.000000 | 3.000000 |
| 75% | 5.000000 | 4.000000 | 2.000000 | 4.000000 |
| max | 8.000000 | 6.000000 | 2.000000 | 9.000000 |

Examine customers dataframe

```
In [6]: # load in the data
```

```
customers = pd.read_csv('../data/Term2/capstone/arvato_data/Udacity_CUSTOMERS_052018..
```

```
customers.head()
```

```
/opt/conda/lib/python3.6/site-packages/IPython/core/interactiveshell.py:2785: DtypeWarning: Colu
interactivity=interactivity, compiler=compiler, result=result)
```

```
Out[6]:
```

| | LNR | AGER_TYP | AKT_DAT_KL | ALTER_HH | ALTER_KIND1 | ALTER_KIND2 | \ |
|---|--------|----------|------------|----------|-------------|-------------|---|
| 0 | 9626 | 2 | 1.0 | 10.0 | NaN | NaN | |
| 1 | 9628 | -1 | 9.0 | 11.0 | NaN | NaN | |
| 2 | 143872 | -1 | 1.0 | 6.0 | NaN | NaN | |
| 3 | 143873 | 1 | 1.0 | 8.0 | NaN | NaN | |
| 4 | 143874 | -1 | 1.0 | 20.0 | NaN | NaN | |

| | ALTER_KIND3 | ALTER_KIND4 | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | \ |
|---|-------------|-------------|----------------------|---------------------|---|
| 0 | NaN | NaN | 10.0 | 1.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | NaN | NaN | 0.0 | 1.0 | |
| 3 | NaN | NaN | 8.0 | 0.0 | |
| 4 | NaN | NaN | 14.0 | 7.0 | |

| | ANZ_HH_TITEL | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE | \ |
|---|--------------|------------|--------------|----------------------------|---|
| 0 | 0.0 | 0.0 | 2.0 | 1.0 | |
| 1 | NaN | 0.0 | 3.0 | NaN | |
| 2 | 0.0 | 0.0 | 1.0 | 1.0 | |
| 3 | NaN | 0.0 | 0.0 | 1.0 | |
| 4 | 0.0 | 0.0 | 4.0 | 7.0 | |

| | ANZ_TITEL | ARBEIT | BALLRAUM | CAMEO_DEU_2015 | CAMEO_DEUG_2015 | CAMEO_INTL_2015 | \ |
|---|-----------|--------|----------|----------------|-----------------|-----------------|---|
| 0 | 0.0 | 1.0 | 3.0 | 1A | 1 | 13 | |
| 1 | 0.0 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 0.0 | 3.0 | 7.0 | 5D | 5 | 34 | |
| 3 | 0.0 | 1.0 | 7.0 | 4C | 4 | 24 | |
| 4 | 0.0 | 3.0 | 3.0 | 7B | 7 | 41 | |

| | CJT_GESAMTTYP | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 | \ |
|---|---------------|-------------------|-----------|-----------|-----------|-----|
| 0 | 5.0 | | 4.0 | 1.0 | 1.0 | 5.0 |
| 1 | NaN | | NaN | NaN | NaN | NaN |
| 2 | 2.0 | | 5.0 | 2.0 | 2.0 | 5.0 |
| 3 | 2.0 | | 5.0 | 1.0 | 1.0 | 5.0 |

| | | | | | | |
|---|-----|--|-----|-----|-----|-----|
| 4 | 6.0 | | 4.0 | 3.0 | 3.0 | 3.0 |
|---|-----|--|-----|-----|-----|-----|

| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 | D19_BANKEN_ANZ_24 | \ |
|---|-----------|-----------|-----------|-------------------|-------------------|---|
| 0 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 1 | NaN | NaN | NaN | 0 | 1 | |
| 2 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 3 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 4 | 4.0 | 3.0 | 3.0 | 1 | 2 | |

| | D19_BANKEN_DATUM | D19_BANKEN_DIREKT | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | \ |
|---|------------------|-------------------|------------------|------------------|---|
| 0 | 10 | 0 | 0 | 0 | |
| 1 | 6 | 0 | 5 | 0 | |
| 2 | 10 | 0 | 0 | 0 | |
| 3 | 10 | 0 | 0 | 0 | |
| 4 | 3 | 5 | 0 | 3 | |

| | D19_BANKEN_OFFLINE_DATUM | D19_BANKEN_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |
| 2 | 10 | 10 | |
| 3 | 10 | 10 | |
| 4 | 10 | 7 | |

| | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST | D19_BEKLEIDUNG_GEH | \ |
|---|----------------------------|-----------------|--------------------|---|
| 0 | 0.0 | 0 | 0 | |
| 1 | 0.0 | 6 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 0.0 | 0 | 0 | |

| | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO | D19_BUCH_CD | \ |
|---|---------------------|-------------|--------------|-------------|---|
| 0 | 0 | 0 | 0 | 6 | |
| 1 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 6 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 6 | |
| 4 | 6 | 0 | 0 | 2 | |

| | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE | D19_FREIZEIT | D19_GARTEN | \ |
|---|----------------|---------------------|-------------|--------------|------------|---|
| 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 6 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 0 | |
| 4 | 0 | 4 | 0 | 6 | 0 | |

| | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 | D19_GESAMT_DATUM | \ |
|---|-------------------|-------------------|------------------|---|
| 0 | 0 | 0 | 9 | |
| 1 | 0 | 1 | 6 | |
| 2 | 0 | 0 | 10 | |

| | | | |
|---|---|---|---|
| 3 | 0 | 1 | 6 |
| 4 | 3 | 5 | 1 |

| | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 9 | 10 | |
| 1 | 10 | 9 | |
| 2 | 10 | 10 | |
| 3 | 6 | 10 | |
| 4 | 8 | 1 | |

| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO | D19_KINDERARTIKEL | \ |
|---|----------------------------|--------------|---------------|-------------------|---|
| 0 | 0.0 | 0 | 6 | 0 | |
| 1 | 0.0 | 0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | 0 | |
| 4 | 10.0 | 0 | 6 | 0 | |

| | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK | D19_LEBENSMITTEL | \ |
|---|---------------|-------------------|--------------|------------------|---|
| 0 | 3.0 | 2 | 0 | 6 | |
| 1 | 5.0 | 3 | 0 | 0 | |
| 2 | 3.0 | 2 | 0 | 0 | |
| 3 | 3.0 | 2 | 0 | 0 | |
| 4 | 1.0 | 4 | 0 | 5 | |

| | D19_LETZTER_KAUF_BRANCHE | D19_LOTTO | D19_NAHRUNGSEGAENZUNG | D19_RATGEBER | \ |
|---|--------------------------|-----------|-----------------------|--------------|---|
| 0 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 1 | D19_BANKEN_GROSS | 0.0 | 0 | 0 | |
| 2 | D19_UNBEKANNT | 7.0 | 0 | 0 | |
| 3 | D19_NAHRUNGSEGAENZUNG | 0.0 | 5 | 0 | |
| 4 | D19_SCHUHE | 0.0 | 0 | 6 | |

| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE | D19_SOZIALES | \ |
|---|------------|-------------------|------------|--------------|--------------|---|
| 0 | 0 | 6 | 0 | 6 | 1.0 | |
| 1 | 0 | 0 | 0 | 0 | 0.0 | |
| 2 | 6 | 6 | 0 | 6 | 1.0 | |
| 3 | 0 | 0 | 0 | 0 | 1.0 | |
| 4 | 0 | 6 | 3 | 6 | 1.0 | |

| | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 | D19_TELKO_DATUM | \ |
|---|-------------|------------------|------------------|-----------------|---|
| 0 | 6 | 0 | 0 | 10 | |
| 1 | 0 | 0 | 0 | 10 | |
| 2 | 0 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 0 | 10 | |
| 4 | 6 | 0 | 1 | 7 | |

| | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM | D19_TELKO_ONLINE_DATUM | \ |
|---|------------------|-------------------------|------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 0 | 10 | 10 | |

| | | | |
|---|---|----|----|
| 2 | 0 | 10 | 10 |
| 3 | 0 | 10 | 10 |
| 4 | 5 | 10 | 9 |

| | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST | D19_TIERARTIKEL \ |
|---|---------------------------|----------------|-------------------|
| 0 | 0.0 | 0 | 0 |
| 1 | 0.0 | 0 | 0 |
| 2 | 0.0 | 0 | 0 |
| 3 | 0.0 | 0 | 0 |
| 4 | 0.0 | 6 | 0 |

| | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 | D19_VERSAND_DATUM \ |
|---|--------------------|--------------------|---------------------|
| 0 | 0 | 0 | 9 |
| 1 | 0 | 0 | 9 |
| 2 | 0 | 0 | 10 |
| 3 | 0 | 1 | 6 |
| 4 | 3 | 5 | 1 |

| | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM \ |
|---|---------------------------|----------------------------|
| 0 | 9 | 10 |
| 1 | 10 | 9 |
| 2 | 10 | 10 |
| 3 | 6 | 10 |
| 4 | 8 | 1 |

| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 \ |
|---|-----------------------------|------------------|--------------------|
| 0 | 0.0 | 0 | 0 |
| 1 | 0.0 | 0 | 0 |
| 2 | 0.0 | 0 | 0 |
| 3 | 0.0 | 0 | 0 |
| 4 | 10.0 | 3 | 0 |

| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM \ |
|---|------------------|-----------------|---------------------------|
| 0 | 0 | 10 | 10 |
| 1 | 0 | 10 | 10 |
| 2 | 0 | 10 | 10 |
| 3 | 0 | 9 | 10 |
| 4 | 0 | 10 | 10 |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN \ |
|---|------------------------|---------------------------|----------------------|
| 0 | 10 | 0.0 | 0 |
| 1 | 10 | 0.0 | 0 |
| 2 | 10 | 0.0 | 0 |
| 3 | 10 | 0.0 | 6 |
| 4 | 10 | 0.0 | 0 |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG | EINGEFUEGT_AM \ |
|---|-------------------|-------------------|----------|---------------------|
| 0 | 0 | 0 | 1.0 | 1992-02-12 00:00:00 |

| | | | | |
|---|---|---|-----|---------------------|
| 1 | 6 | 0 | NaN | NaN |
| 2 | 0 | 0 | 1.0 | 1992-02-10 00:00:00 |
| 3 | 6 | 0 | 1.0 | 1992-02-10 00:00:00 |
| 4 | 0 | 0 | 1.0 | 1992-02-12 00:00:00 |

| | EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER | \ |
|---|----------------------|----------|-----------|----------------|---|
| 0 | 1994.0 | 2.0 | 40.0 | 1 | |
| 1 | 2007.0 | NaN | 29.0 | 1 | |
| 2 | 1996.0 | 4.0 | 26.0 | 1 | |
| 3 | 1997.0 | 1.0 | 10.0 | 2 | |
| 4 | 1997.0 | 4.0 | NaN | 4 | |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER | FINANZ_UNAUFFAELLIGER | \ |
|---|------------------|-------------------|---------------|-----------------------|---|
| 0 | 2 | | 5 | 1 | 2 |
| 1 | 2 | | 5 | 1 | 3 |
| 2 | 4 | | 5 | 1 | 4 |
| 3 | 2 | | 5 | 1 | 1 |
| 4 | 2 | | 3 | 1 | 5 |

| | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE | GEBAEUDETYP | GEBAEUDETYP_RASTER | \ |
|---|------------------|-----------|--------------|-------------|--------------------|---|
| 0 | 5 | 2 | 4.0 | 1.0 | 4.0 | |
| 1 | 5 | 2 | NaN | NaN | NaN | |
| 2 | 5 | 2 | 2.0 | 8.0 | 3.0 | |
| 3 | 5 | 6 | 4.0 | 2.0 | 4.0 | |
| 4 | 4 | 2 | 2.0 | 3.0 | 3.0 | |

| | GEBURTSJAHR | GEMEINDETYPE | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP | \ |
|---|-------------|--------------|-----------------|------------------|------------|---|
| 0 | 0 | 50.0 | 4.0 | 1 | 1 | |
| 1 | 0 | NaN | NaN | 0 | 1 | |
| 2 | 0 | 22.0 | 3.0 | 1 | 2 | |
| 3 | 0 | 40.0 | 10.0 | 0 | 2 | |
| 4 | 1960 | 22.0 | 2.0 | 0 | 3 | |

| | HH_DELTA_FLAG | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 | \ |
|---|---------------|--------------------|------------|--------------|--------------|---|
| 0 | 0.0 | 1.0 | 4.0 | 2.0 | 2.0 | |
| 1 | 1.0 | NaN | NaN | NaN | NaN | |
| 2 | 0.0 | 1.0 | 1.0 | 2.0 | 4.0 | |
| 3 | NaN | 4.0 | 7.0 | 2.0 | 3.0 | |
| 4 | 1.0 | 6.0 | 4.0 | 2.0 | 4.0 | |

| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 | KBA05_ANTG2 | \ |
|---|--------------|--------------|--------------|-------------|-------------|---|
| 0 | 4.0 | 4.0 | 1.0 | 2.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 3.0 | 3.0 | 2.0 | 2.0 | |
| 3 | 3.0 | 3.0 | 3.0 | 3.0 | 0.0 | |
| 4 | 4.0 | 1.0 | 0.0 | 0.0 | 3.0 | |

| | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT | KBA05_BAUMAX | KBA05_CCM1 | \ |
|--|-------------|-------------|----------------|--------------|------------|---|
|--|-------------|-------------|----------------|--------------|------------|---|

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 0 | 0.0 | 0.0 | 5.0 | 0.0 | 3.0 |
| 1 | NaN | NaN | NaN | NaN | NaN |
| 2 | 0.0 | 0.0 | 4.0 | 0.0 | 2.0 |
| 3 | 0.0 | 0.0 | 3.0 | 1.0 | 3.0 |
| 4 | 2.0 | 0.0 | 3.0 | 0.0 | 2.0 |

| | KBA05_CCM2 | KBA05_CCM3 | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ | \ |
|---|------------|------------|------------|--------------|------------|-----------|---|
| 0 | 3.0 | 3.0 | 1.0 | 1.0 | 4.0 | 4.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 5.0 | 2.0 | 4.0 | 4.0 | 3.0 | |
| 3 | 4.0 | 2.0 | 0.0 | 3.0 | 2.0 | 4.0 | |
| 4 | 5.0 | 2.0 | 0.0 | 2.0 | 3.0 | 3.0 | |

| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 | KBA05_HERST5 | \ |
|---|--------------|--------------|--------------|--------------|--------------|---|
| 0 | 4.0 | 3.0 | 3.0 | 2.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 4.0 | 4.0 | 1.0 | 2.0 | 3.0 | |
| 3 | 3.0 | 1.0 | 3.0 | 5.0 | 3.0 | |
| 4 | 1.0 | 4.0 | 3.0 | 2.0 | 3.0 | |

| | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 | KBA05_KRSHERST2 | \ |
|---|-----------------|----------------|-----------------|-----------------|---|
| 0 | 2.0 | 4.0 | 4.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 4.0 | 4.0 | 4.0 | |
| 3 | 3.0 | 3.0 | 4.0 | 2.0 | |
| 4 | 1.0 | 2.0 | 2.0 | 4.0 | |

| | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER | KBA05_KRSVAN | KBA05_KRSZUL | \ |
|---|-----------------|----------------|---------------|--------------|--------------|---|
| 0 | 3.0 | 2.0 | 3.0 | 3.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | |
| 3 | 2.0 | 2.0 | 1.0 | 2.0 | 3.0 | |
| 4 | 3.0 | 1.0 | 2.0 | 1.0 | 1.0 | |

| | KBA05_KW1 | KBA05_KW2 | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST | \ |
|---|-----------|-----------|-----------|-------------|-------------|----------------|---|
| 0 | 3.0 | 3.0 | 3.0 | 5.0 | 2.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 3.0 | 3.0 | 3.0 | 2.0 | 2.0 | |
| 3 | 3.0 | 4.0 | 0.0 | 3.0 | 4.0 | 4.0 | |
| 4 | 3.0 | 4.0 | 0.0 | 3.0 | 1.0 | 2.0 | |

| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 | KBA05_MOD3 | \ |
|---|--------------|---------------|------------|------------|------------|---|
| 0 | 3.0 | 2.0 | 2.0 | 4.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 1.0 | 2.0 | 3.0 | 4.0 | |
| 3 | 1.0 | 3.0 | 0.0 | 3.0 | 5.0 | |
| 4 | 2.0 | 3.0 | 1.0 | 3.0 | 3.0 | |

| | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP | KBA05_MOTOR | KBA05_MOTRAD | \ |
|---|------------|------------|---------------|-------------|--------------|---|
| 0 | 4.0 | 2.0 | 2.0 | 4.0 | 1.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 1.0 | 3.0 | 4.0 | 1.0 | |
| 3 | 0.0 | 1.0 | 4.0 | 2.0 | 1.0 | |
| 4 | 1.0 | 1.0 | 4.0 | 2.0 | 0.0 | |

| | KBA05_SEG1 | KBA05_SEG10 | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 | \ |
|---|------------|-------------|------------|------------|------------|------------|---|
| 0 | 1.0 | 2.0 | 3.0 | 2.0 | 3.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 2 | 1.0 | 4.0 | 2.0 | 4.0 | 3.0 | 1.0 | |
| 3 | 3.0 | 2.0 | 4.0 | 3.0 | 4.0 | 0.0 | |
| 4 | 0.0 | 1.0 | 2.0 | 4.0 | 3.0 | 1.0 | |

| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 | KBA05_VORBO | KBA05_VORB1 | \ |
|---|------------|------------|------------|------------|-------------|-------------|---|
| 0 | 1.0 | 0.0 | 1.0 | 2.0 | 4.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 2 | 0.0 | 3.0 | 2.0 | 1.0 | 4.0 | 3.0 | |
| 3 | 0.0 | 0.0 | 2.0 | 0.0 | 3.0 | 3.0 | |
| 4 | 0.0 | 1.0 | 0.0 | 0.0 | 2.0 | 4.0 | |

| | KBA05_VORB2 | KBA05_ZUL1 | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | \ |
|---|-------------|------------|------------|------------|------------|---|
| 0 | 2.0 | 3.0 | 4.0 | 2.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 2.0 | 4.0 | 3.0 | 2.0 | |
| 3 | 2.0 | 2.0 | 3.0 | 4.0 | 4.0 | |
| 4 | 3.0 | 3.0 | 3.0 | 3.0 | 1.0 | |

| | KBA13_ALTERHALTER_30 | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | \ |
|---|----------------------|----------------------|----------------------|---|
| 0 | 1.0 | 2.0 | 5.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 3.0 | 3.0 | 2.0 | |
| 3 | 3.0 | 4.0 | 3.0 | |
| 4 | 1.0 | 3.0 | 3.0 | |

| | KBA13_ALTERHALTER_61 | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 | \ |
|---|----------------------|-------------|-------------|-------------|-------------|---|
| 0 | 3.0 | 3.0 | 3.0 | 1.0 | 0.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 4.0 | 1.0 | 4.0 | 3.0 | 1.0 | |
| 3 | 2.0 | 2.0 | 2.0 | 0.0 | 0.0 | |
| 4 | 4.0 | 2.0 | 4.0 | 2.0 | 1.0 | |

| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX | KBA13_BJ_1999 | \ |
|---|------------------|------------|-----------------|--------------|---------------|---|
| 0 | 1201.0 | 4.0 | 4.0 | 1.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 433.0 | 4.0 | 2.0 | 3.0 | 4.0 | |
| 3 | 755.0 | 2.0 | 4.0 | 1.0 | 3.0 | |
| 4 | 513.0 | 2.0 | 3.0 | 2.0 | 4.0 | |

| | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 | KBA13_BJ_2008 | KBA13_BJ_2009 \ |
|---|---------------|---------------|---------------|---------------|-----------------|
| 0 | 2.0 | 3.0 | 3.0 | 4.0 | 3.0 |
| 1 | NaN | NaN | NaN | NaN | NaN |
| 2 | 3.0 | 1.0 | 2.0 | 0.0 | 5.0 |
| 3 | 2.0 | 3.0 | 4.0 | 3.0 | 2.0 |
| 4 | 3.0 | 3.0 | 3.0 | 5.0 | 1.0 |

| | KBA13_BMW | KBA13_CCM_0_1400 | KBA13_CCM_1000 | KBA13_CCM_1200 \ |
|---|-----------|------------------|----------------|------------------|
| 0 | 5.0 | 2.0 | 3.0 | 2.0 |
| 1 | NaN | NaN | NaN | NaN |
| 2 | 4.0 | 4.0 | 5.0 | 3.0 |
| 3 | 4.0 | 2.0 | 2.0 | 0.0 |
| 4 | 3.0 | 3.0 | 3.0 | 3.0 |

| | KBA13_CCM_1400 | KBA13_CCM_1401_2500 | KBA13_CCM_1500 | KBA13_CCM_1600 \ |
|---|----------------|---------------------|----------------|------------------|
| 0 | 3.0 | 4.0 | 4.0 | 2.0 |
| 1 | NaN | NaN | NaN | NaN |
| 2 | 4.0 | 1.0 | 1.0 | 2.0 |
| 3 | 2.0 | 4.0 | 3.0 | 4.0 |
| 4 | 4.0 | 2.0 | 1.0 | 3.0 |

| | KBA13_CCM_1800 | KBA13_CCM_2000 | KBA13_CCM_2500 | KBA13_CCM_2501 \ |
|---|----------------|----------------|----------------|------------------|
| 0 | 3.0 | 4.0 | 4.0 | 0.0 |
| 1 | NaN | NaN | NaN | NaN |
| 2 | 2.0 | 3.0 | 2.0 | 4.0 |
| 3 | 3.0 | 4.0 | 3.0 | 3.0 |
| 4 | 0.0 | 2.0 | 0.0 | 3.0 |

| | KBA13_CCM_3000 | KBA13_CCM_3001 | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE \ |
|---|----------------|----------------|-----------------|----------------------|
| 0 | 3.0 | 3.0 | 2.0 | 3.0 |
| 1 | NaN | NaN | NaN | NaN |
| 2 | 5.0 | 3.0 | 3.0 | 3.0 |
| 3 | 3.0 | 3.0 | 4.0 | 3.0 |
| 4 | 4.0 | 1.0 | 4.0 | 3.0 |

| | KBA13_FIAT | KBA13_FORD | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 \ |
|---|------------|------------|-----------|-----------------|-------------------|
| 0 | 3.0 | 3.0 | 5.0 | 1.0 | 1.0 |
| 1 | NaN | NaN | NaN | NaN | NaN |
| 2 | 3.0 | 3.0 | 2.0 | 3.0 | 3.0 |
| 3 | 1.0 | 3.0 | 5.0 | 4.0 | 4.0 |
| 4 | 3.0 | 4.0 | 3.0 | 2.0 | 2.0 |

| | KBA13_HALTER_30 | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 \ |
|---|-----------------|-----------------|-----------------|-------------------|
| 0 | 1.0 | 3.0 | 3.0 | 2.0 |
| 1 | NaN | NaN | NaN | NaN |
| 2 | 4.0 | 4.0 | 2.0 | 2.0 |
| 3 | 3.0 | 4.0 | 4.0 | 4.0 |

| | | | | |
|---|-----|-----|-----|-----|
| 4 | 1.0 | 2.0 | 2.0 | 5.0 |
|---|-----|-----|-----|-----|

| | KBA13_HALTER_50 | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | 4.0 | 5.0 | 5.0 | 4.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 3 | 4.0 | 3.0 | 3.0 | 1.0 | |
| 4 | 4.0 | 3.0 | 3.0 | 3.0 | |

| | KBA13_HALTER_66 | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | \ |
|---|-----------------|-------------------|---------------------|---|
| 0 | 3.0 | 2.0 | 4.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 4.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 4.0 | 1.0 | |
| 4 | 4.0 | 3.0 | 2.0 | |

| | KBA13_HERST_BMW_BENZ | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | \ |
|---|----------------------|--------------------|-----------------------|---|
| 0 | 4.0 | 2.0 | 3.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 3.0 | 4.0 | 2.0 | |
| 3 | 4.0 | 5.0 | 2.0 | |
| 4 | 3.0 | 4.0 | 4.0 | |

| | KBA13_HERST_SONST | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | \ |
|---|-------------------|-----------|-----------------|---------------|---|
| 0 | 3.0 | 5.0 | 3.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 3.0 | 4.0 | 1.0 | |
| 3 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 4 | 3.0 | 3.0 | 3.0 | 1.0 | |

| | KBA13_KMH_140 | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | \ |
|---|---------------|-------------------|---------------|---------------|---|
| 0 | 4.0 | 3.0 | 2.0 | 4.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 5.0 | 2.0 | 3.0 | 2.0 | |
| 3 | 3.0 | 3.0 | 3.0 | 4.0 | |
| 4 | 4.0 | 3.0 | 4.0 | 2.0 | |

| | KBA13_KMH_211 | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | \ |
|---|---------------|---------------|---------------|----------------|---|
| 0 | 3.0 | 3.0 | 1.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 3.0 | 1.0 | 3.0 | |
| 3 | 3.0 | 3.0 | 1.0 | 4.0 | |
| 4 | 3.0 | 3.0 | 1.0 | 3.0 | |

| | KBA13_KRSHERST_AUDI_VW | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | \ |
|---|------------------------|-------------------------|--------------------------|---|
| 0 | 4.0 | 3.0 | 3.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 4.0 | 3.0 | 1.0 | |

| | | | | | | |
|---|--|-----|--|-----|--|-----|
| 3 | | 3.0 | | 5.0 | | 1.0 |
| 4 | | 3.0 | | 2.0 | | 3.0 |

| | KBA13_KRSSEG_KLEIN | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | \ |
|---|--------------------|-------------------|------------------|------------------|---|
| 0 | 2.0 | 2.0 | 2.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 1.0 | 1.0 | 3.0 | |
| 3 | 1.0 | 3.0 | 3.0 | 1.0 | |
| 4 | 2.0 | 2.0 | 2.0 | 2.0 | |

| | KBA13_KW_0_60 | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 | \ |
|---|---------------|--------------|--------------|--------------|-------------|---|
| 0 | 3.0 | 4.0 | 3.0 | 3.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 4.0 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 4 | 4.0 | 2.0 | 1.0 | 3.0 | 1.0 | |

| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 | KBA13_KW_70 | \ |
|---|-------------|-------------|-------------|-----------------|-------------|---|
| 0 | 3.0 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 4.0 | 4.0 | 1.0 | 2.0 | |
| 3 | 2.0 | 2.0 | 2.0 | 4.0 | 2.0 | |
| 4 | 3.0 | 4.0 | 4.0 | 1.0 | 3.0 | |

| | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA | KBA13_MERCEDES | KBA13_MOTOR | \ |
|---|-------------|-------------|-------------|----------------|-------------|---|
| 0 | 2.0 | 3.0 | 3.0 | 3.0 | 4.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 1.0 | 0.0 | 3.0 | 3.0 | 3.0 | |
| 3 | 5.0 | 2.0 | 5.0 | 4.0 | 3.0 | |
| 4 | 0.0 | 3.0 | 2.0 | 3.0 | 4.0 | |

| | KBA13_NISSAN | KBA13_OPEL | KBA13_PEUGEOT | KBA13_RENAULT | \ |
|---|--------------|------------|---------------|---------------|---|
| 0 | 3.0 | 3.0 | 2.0 | 2.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 1.0 | 5.0 | 3.0 | |
| 3 | 2.0 | 1.0 | 5.0 | 2.0 | |
| 4 | 4.0 | 4.0 | 3.0 | 4.0 | |

| | KBA13_SEG_GELAENDEWAGEN | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST | \ |
|---|-------------------------|-------------------------|-------------------|---|
| 0 | 3.0 | 5.0 | 2.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 4.0 | 3.0 | 4.0 | |
| 3 | 5.0 | 3.0 | 1.0 | |
| 4 | 3.0 | 4.0 | 4.0 | |

| | KBA13_SEG_KLEINWAGEN | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS | \ |
|---|----------------------|-------------------------|--------------------|---|
| 0 | 2.0 | 3.0 | 3.0 | |
| 1 | NaN | NaN | NaN | |

| | | | |
|---|-----|-----|-----|
| 2 | 3.0 | 3.0 | 2.0 |
| 3 | 2.0 | 3.0 | 4.0 |
| 4 | 5.0 | 3.0 | 4.0 |

| | KBA13_SEG_MINIWAGEN | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE | \ |
|---|---------------------|------------------------|-----------------------------|---|
| 0 | 1.0 | 3.0 | 4.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 5.0 | 2.0 | 2.0 | |
| 3 | 2.0 | 3.0 | 2.0 | |
| 4 | 3.0 | 2.0 | 3.0 | |

| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN | \ |
|---|----------------------|--------------------|----------------------|---|
| 0 | 3.0 | 4.0 | 4.0 | |
| 1 | NaN | NaN | NaN | |
| 2 | 1.0 | 5.0 | 3.0 | |
| 3 | 0.0 | 4.0 | 4.0 | |
| 4 | 1.0 | 3.0 | 3.0 | |

| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE | KBA13_SITZE_4 | \ |
|---|---------------------|---------------|----------------------|---------------|---|
| 0 | 4.0 | 4.0 | 4.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 3.0 | 5.0 | 3.0 | |
| 3 | 3.0 | 4.0 | 2.0 | 3.0 | |
| 4 | 3.0 | 4.0 | 3.0 | 4.0 | |

| | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA | KBA13_VORB_0 | KBA13_VORB_1 | \ |
|---|---------------|---------------|--------------|--------------|--------------|---|
| 0 | 1.0 | 5.0 | 2.0 | 4.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 4.0 | 3.0 | 4.0 | 2.0 | |
| 3 | 3.0 | 3.0 | 5.0 | 4.0 | 3.0 | |
| 4 | 2.0 | 3.0 | 3.0 | 3.0 | 3.0 | |

| | KBA13_VORB_1_2 | KBA13_VORB_2 | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK | \ |
|---|----------------|--------------|--------------|----------|--------------|-----|---|
| 0 | 3.0 | 3.0 | 1.0 | 4.0 | NaN | 1.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | NaN | |
| 2 | 2.0 | 3.0 | 4.0 | 3.0 | NaN | 3.0 | |
| 3 | 2.0 | 2.0 | 3.0 | 1.0 | NaN | 3.0 | |
| 4 | 3.0 | 3.0 | 0.0 | 2.0 | 2.0 | 4.0 | |

| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN | LP_FAMILIE_GROB | \ |
|---|------------|-------------|-------------|-----------------|-----------------|---|
| 0 | 4 | 5.0 | 0.0 | 2.0 | 2.0 | |
| 1 | 4 | 5.0 | NaN | NaN | NaN | |
| 2 | 4 | 1.0 | 1.0 | 1.0 | 1.0 | |
| 3 | 4 | 2.0 | 0.0 | 0.0 | 0.0 | |
| 4 | 3 | 1.0 | 1.0 | 10.0 | 5.0 | |

| | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB | LP_STATUS_FEIN | LP_STATUS_GROB | \ |
|---|---------------------|---------------------|----------------|----------------|---|
| 0 | 20.0 | 5.0 | 10.0 | 5.0 | |

| | | | | |
|---|------|------|------|-----|
| 1 | NaN | NaN | NaN | NaN |
| 2 | 13.0 | 3.0 | 10.0 | 5.0 |
| 3 | 0.0 | 0.0 | 9.0 | 4.0 |
| 4 | 31.0 | 10.0 | 1.0 | 1.0 |

| | MIN_GEBAEUDEJAHR | MOBI_RASTER | MOBI_REGIO | NATIONALITAET_KZ | \ |
|---|------------------|-------------|------------|------------------|---|
| 0 | 1992.0 | 3.0 | 4.0 | 1 | |
| 1 | NaN | NaN | NaN | 1 | |
| 2 | 1992.0 | 1.0 | 3.0 | 1 | |
| 3 | 1992.0 | 3.0 | 4.0 | 1 | |
| 4 | 1992.0 | 1.0 | 3.0 | 1 | |

| | ONLINE_AFFINITAET | ORTSGR_KLS9 | OST_WEST_KZ | PLZ8_ANTG1 | PLZ8_ANTG2 | \ |
|---|-------------------|-------------|-------------|------------|------------|---|
| 0 | 3.0 | 2.0 | W | 3.0 | 3.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 1.0 | 5.0 | W | 2.0 | 3.0 | |
| 3 | 2.0 | 3.0 | W | 3.0 | 2.0 | |
| 4 | 5.0 | 5.0 | W | 2.0 | 4.0 | |

| | PLZ8_ANTG3 | PLZ8_ANTG4 | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ | \ |
|---|------------|------------|-------------|----------|----------|---|
| 0 | 1.0 | 0.0 | 1.0 | 5.0 | 5.0 | |
| 1 | NaN | NaN | NaN | NaN | NaN | |
| 2 | 3.0 | 1.0 | 3.0 | 2.0 | 3.0 | |
| 3 | 1.0 | 0.0 | 1.0 | 4.0 | 3.0 | |
| 4 | 2.0 | 1.0 | 2.0 | 3.0 | 3.0 | |

| | PRAEGENDE_JUGENDJAHRE | REGIOTYP | RELAT_AB | RETOURTYP_BK_S | RT_KEIN_ANREIZ | \ |
|---|-----------------------|----------|----------|----------------|----------------|---|
| 0 | 4 | 1.0 | 1.0 | 5.0 | 1.0 | |
| 1 | 0 | NaN | NaN | NaN | NaN | |
| 2 | 4 | 7.0 | 3.0 | 5.0 | 1.0 | |
| 3 | 1 | 6.0 | 1.0 | 3.0 | 1.0 | |
| 4 | 8 | 7.0 | 1.0 | 5.0 | 4.0 | |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL | SEMIO_FAM | \ |
|---|-----------------|-----------------|-----------|-----------|-----------|---|
| 0 | 5.0 | 3.0 | 1 | 3 | 5 | |
| 1 | NaN | NaN | 3 | 3 | 6 | |
| 2 | 5.0 | 1.0 | 5 | 7 | 2 | |
| 3 | 5.0 | 2.0 | 3 | 3 | 5 | |
| 4 | 3.0 | 5.0 | 5 | 4 | 5 | |

| | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT | \ |
|---|------------|------------|------------|------------|-----------|---------------|---|
| 0 | 1 | 3 | 4 | 7 | 6 | 2 | |
| 1 | 2 | 3 | 4 | 5 | 6 | 4 | |
| 2 | 6 | 7 | 1 | 7 | 3 | 4 | |
| 3 | 3 | 3 | 4 | 5 | 4 | 3 | |
| 4 | 2 | 3 | 5 | 6 | 6 | 5 | |

| | SEMIO_RAT | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT | SHOPPER_TYP | \ |
|--|-----------|-----------|-----------|-------------|------------|-------------|---|
|--|-----------|-----------|-----------|-------------|------------|-------------|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 6 | 1 | 6 | 3 |
| 1 | 1 | 2 | 3 | 1 | 7 | 3 |
| 2 | 2 | 1 | 2 | 1 | 3 | 1 |
| 3 | 3 | 3 | 6 | 4 | 7 | 0 |
| 4 | 5 | 4 | 4 | 4 | 5 | 1 |

| | SOHO_KZ | STRUKTURTYP | TITEL_KZ | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG \ |
|---|---------|-------------|----------|------------|-------------|--------------------|
| 0 | 0.0 | 3.0 | 0.0 | 4.0 | 4.0 | 0.0 |
| 1 | 0.0 | NaN | 0.0 | NaN | NaN | 0.0 |
| 2 | 0.0 | 3.0 | 0.0 | 1.0 | 5.0 | 0.0 |
| 3 | 0.0 | 1.0 | 0.0 | 3.0 | 4.0 | 0.0 |
| 4 | 0.0 | 3.0 | 0.0 | 2.0 | 4.0 | 0.0 |

| | VERDICHTUNGSRAUM | VERS_TYP | VHA | VHN | VK_DHT4A | VK_DISTANZ | VK_ZG11 \ |
|---|------------------|----------|-----|-----|----------|------------|-----------|
| 0 | 8.0 | 1 | 0.0 | 3.0 | 5.0 | 3.0 | 2.0 |
| 1 | NaN | 1 | 0.0 | NaN | 6.0 | 6.0 | 3.0 |
| 2 | 0.0 | 2 | 0.0 | 4.0 | 10.0 | 13.0 | 11.0 |
| 3 | 0.0 | 1 | 0.0 | 2.0 | 6.0 | 4.0 | 2.0 |
| 4 | 1.0 | 2 | 0.0 | 4.0 | 3.0 | 5.0 | 4.0 |

| | W_KEIT_KIND_HH | WOHNDAUER_2008 | WOHNLAGE | ZABEOTYP | PRODUCT_GROUP \ |
|---|----------------|----------------|----------|----------|-------------------|
| 0 | 6.0 | 9.0 | 7.0 | 3 | COSMETIC_AND_FOOD |
| 1 | 0.0 | 9.0 | NaN | 3 | FOOD |
| 2 | 6.0 | 9.0 | 2.0 | 3 | COSMETIC_AND_FOOD |
| 3 | NaN | 9.0 | 7.0 | 1 | COSMETIC |
| 4 | 2.0 | 9.0 | 3.0 | 1 | FOOD |

| | CUSTOMER_GROUP | ONLINE_PURCHASE | ANREDE_KZ | ALTERSKATEGORIE_GROB |
|---|----------------|-----------------|-----------|----------------------|
| 0 | MULTI_BUYER | 0 | 1 | 4 |
| 1 | SINGLE_BUYER | 0 | 1 | 4 |
| 2 | MULTI_BUYER | 0 | 2 | 4 |
| 3 | MULTI_BUYER | 0 | 1 | 4 |
| 4 | MULTI_BUYER | 0 | 1 | 3 |

```
In [7]: 'Shape', customers.shape
```

```
Out[7]: ('Shape', (191652, 369))
```

```
In [8]: customers.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 191652 entries, 0 to 191651
Columns: 369 entries, LNR to ALTERSKATEGORIE_GROB
dtypes: float64(267), int64(94), object(8)
memory usage: 539.5+ MB
```

```
In [9]: customers.describe()
```


Out[9]:

| | LNR | AGER_TYP | AKT_DAT_KL | ALTER_HH \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 191652.000000 | 145056.000000 | 145056.000000 |
| mean | 95826.500000 | 0.344359 | 1.747525 | 11.352009 |
| std | 55325.311233 | 1.391672 | 1.966334 | 6.275026 |
| min | 1.000000 | -1.000000 | 1.000000 | 0.000000 |
| 25% | 47913.750000 | -1.000000 | 1.000000 | 8.000000 |
| 50% | 95826.500000 | 0.000000 | 1.000000 | 11.000000 |
| 75% | 143739.250000 | 2.000000 | 1.000000 | 16.000000 |
| max | 191652.000000 | 3.000000 | 9.000000 | 21.000000 |

| | ALTER_KIND1 | ALTER_KIND2 | ALTER_KIND3 | ALTER_KIND4 \ |
|-------|--------------|-------------|-------------|---------------|
| count | 11766.000000 | 5100.000000 | 1275.000000 | 236.000000 |
| mean | 12.337243 | 13.672353 | 14.647059 | 15.377119 |
| std | 4.006050 | 3.243335 | 2.753787 | 2.307653 |
| min | 2.000000 | 2.000000 | 5.000000 | 8.000000 |
| 25% | 9.000000 | 11.000000 | 13.000000 | 14.000000 |
| 50% | 13.000000 | 14.000000 | 15.000000 | 16.000000 |
| 75% | 16.000000 | 16.000000 | 17.000000 | 17.000000 |
| max | 18.000000 | 18.000000 | 18.000000 | 18.000000 |

| | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | ANZ_HH_TITEL \ |
|-------|----------------------|---------------------|----------------|
| count | 139810.000000 | 141725.000000 | 139542.000000 |
| mean | 10.331579 | 4.965863 | 0.067413 |
| std | 4.134828 | 14.309694 | 0.545576 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 9.000000 | 1.000000 | 0.000000 |
| 50% | 10.000000 | 1.000000 | 0.000000 |
| 75% | 13.000000 | 4.000000 | 0.000000 |
| max | 25.000000 | 523.000000 | 23.000000 |

| | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE \ |
|-------|---------------|---------------|------------------------------|
| count | 145056.000000 | 145056.000000 | 141725.000000 |
| mean | 0.136402 | 2.267828 | 4.701288 |
| std | 0.493249 | 1.390620 | 14.184081 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 1.000000 | 1.000000 |
| 50% | 0.000000 | 2.000000 | 1.000000 |
| 75% | 0.000000 | 3.000000 | 3.000000 |
| max | 8.000000 | 21.000000 | 375.000000 |

| | ANZ_TITEL | ARBEIT | BALLRAUM | CJT_GESAMTTYP \ |
|-------|---------------|---------------|---------------|-----------------|
| count | 145056.000000 | 141176.000000 | 141693.000000 | 188439.000000 |
| mean | 0.020392 | 2.824850 | 4.301758 | 3.677928 |
| std | 0.152234 | 1.012415 | 2.114614 | 1.813975 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 0.000000 | 3.000000 | 5.000000 | 4.000000 |
| 75% | 0.000000 | 4.000000 | 6.000000 | 6.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| max | 5.000000 | 9.000000 | 7.000000 | 6.000000 |
|-----|----------|----------|----------|----------|

| | | | | |
|-------|-------------------|---------------|---------------|---------------|
| | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 \ |
| count | 188439.00000 | 188439.000000 | 188439.000000 | 188439.000000 |
| mean | 4.00994 | 2.665441 | 2.548490 | 4.516369 |
| std | 1.37894 | 1.545714 | 1.557262 | 0.874722 |
| min | 1.00000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.00000 | 1.000000 | 1.000000 | 4.000000 |
| 50% | 5.00000 | 2.000000 | 2.000000 | 5.000000 |
| 75% | 5.00000 | 4.000000 | 4.000000 | 5.000000 |
| max | 5.00000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------------|
| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 \ |
| count | 188439.000000 | 188439.000000 | 188439.000000 | 191652.000000 |
| mean | 4.415317 | 4.519956 | 4.538838 | 0.091040 |
| std | 1.025281 | 0.895371 | 0.886091 | 0.416684 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 4.000000 | 4.000000 | 4.000000 | 0.000000 |
| 50% | 5.000000 | 5.000000 | 5.000000 | 0.000000 |
| 75% | 5.000000 | 5.000000 | 5.000000 | 0.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 6.000000 |

| | | | |
|-------|-------------------|------------------|---------------------|
| | D19_BANKEN_ANZ_24 | D19_BANKEN_DATUM | D19_BANKEN_DIREKT \ |
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.159524 | 9.367599 | 0.646166 |
| std | 0.589824 | 1.643262 | 1.771525 |
| min | 0.000000 | 1.000000 | 0.000000 |
| 25% | 0.000000 | 10.000000 | 0.000000 |
| 50% | 0.000000 | 10.000000 | 0.000000 |
| 75% | 0.000000 | 10.000000 | 0.000000 |
| max | 6.000000 | 10.000000 | 7.000000 |

| | | | |
|-------|------------------|------------------|----------------------------|
| | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | D19_BANKEN_OFFLINE_DATUM \ |
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.424478 | 0.132970 | 9.866472 |
| std | 1.443739 | 0.907925 | 0.772253 |
| min | 0.000000 | 0.000000 | 1.000000 |
| 25% | 0.000000 | 0.000000 | 10.000000 |
| 50% | 0.000000 | 0.000000 | 10.000000 |
| 75% | 0.000000 | 0.000000 | 10.000000 |
| max | 6.000000 | 7.000000 | 10.000000 |

| | | | |
|-------|-------------------------|----------------------------|-------------------|
| | D19_BANKEN_ONLINE_DATUM | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST \ |
| count | 191652.000000 | 143955.000000 | 191652.000000 |
| mean | 9.600646 | 0.462019 | 0.444477 |
| std | 1.379280 | 2.087402 | 1.546226 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |

| | | | |
|-----|-----------|-----------|----------|
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | D19_BEKLEIDUNG_GEH | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO \ |
|-------|--------------------|---------------------|---------------|----------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.978727 | 1.527780 | 0.998565 | 0.538492 |
| std | 2.097215 | 2.562741 | 2.215235 | 1.746448 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 3.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_BUCH_CD | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE \ |
|-------|---------------|----------------|---------------------|---------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 2.443987 | 0.218265 | 0.757999 | 0.474266 |
| std | 2.822639 | 1.084785 | 1.857884 | 1.522716 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 6.000000 | 0.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_FREIZEIT | D19_GARTEN | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 \ |
|-------|---------------|---------------|-------------------|---------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.688414 | 0.344525 | 0.937679 | 1.438581 |
| std | 1.838987 | 1.383708 | 1.376459 | 1.742774 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 75% | 0.000000 | 0.000000 | 2.000000 | 3.000000 |
| max | 7.000000 | 7.000000 | 6.000000 | 6.000000 |

| | D19_GESAMT_DATUM | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM \ |
|-------|------------------|--------------------------|---------------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 6.529496 | 8.412435 | 7.445714 |
| std | 3.246552 | 2.185122 | 3.117772 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 4.000000 | 8.000000 | 5.000000 |
| 50% | 7.000000 | 9.000000 | 9.000000 |
| 75% | 10.000000 | 10.000000 | 10.000000 |
| max | 10.000000 | 10.000000 | 10.000000 |

| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO \ |
|-------|----------------------------|---------------|-----------------|
| count | 143955.000000 | 191652.000000 | 191652.000000 |
| mean | 3.522879 | 1.535257 | 1.537031 |
| std | 4.561253 | 2.668880 | 2.459250 |

| | | | |
|-----|-----------|----------|----------|
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 3.000000 | 3.000000 |
| max | 10.000000 | 7.000000 | 7.000000 |

| | D19_KINDERARTIKEL | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK \ |
|-------|-------------------|---------------|-------------------|----------------|
| count | 191652.000000 | 143955.000000 | 191652.000000 | 191652.000000 |
| mean | 1.083516 | 3.027654 | 4.224469 | 1.756110 |
| std | 2.277333 | 2.206507 | 3.198298 | 2.883393 |
| min | 0.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 0.000000 | 2.000000 | 2.000000 | 0.000000 |
| 50% | 0.000000 | 3.000000 | 2.000000 | 0.000000 |
| 75% | 0.000000 | 3.000000 | 8.000000 | 6.000000 |
| max | 7.000000 | 9.000000 | 9.000000 | 7.000000 |

| | D19_LEBENSMITTEL | D19_LOTTO | D19_NAHRUNGSERGAENZUNG | D19_RATGEBER \ |
|-------|------------------|---------------|------------------------|----------------|
| count | 191652.000000 | 143955.000000 | 191652.000000 | 191652.000000 |
| mean | 0.577636 | 2.633733 | 0.505917 | 0.799757 |
| std | 1.723675 | 3.332828 | 1.651736 | 1.959001 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 7.000000 | 0.000000 | 0.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE \ |
|-------|---------------|-------------------|---------------|----------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 1.775557 | 1.432179 | 0.598788 | 3.269436 |
| std | 2.803230 | 2.551136 | 1.584454 | 2.880428 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 | 3.000000 |
| 75% | 6.000000 | 0.000000 | 0.000000 | 6.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | D19_SOZIALES | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 \ |
|-------|---------------|---------------|------------------|--------------------|
| count | 143955.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 1.243590 | 2.324234 | 0.045228 | 0.086871 |
| std | 1.158867 | 2.973775 | 0.246678 | 0.346445 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 75% | 1.000000 | 6.000000 | 0.000000 | 0.000000 |
| max | 5.000000 | 7.000000 | 6.000000 | 6.000000 |

| | D19_TELKO_DATUM | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM \ |
|-------|-----------------|------------------|---------------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |

| | | | |
|------|-----------|----------|-----------|
| mean | 9.482014 | 0.928694 | 9.799339 |
| std | 1.288103 | 2.124832 | 0.832611 |
| min | 1.000000 | 0.000000 | 1.000000 |
| 25% | 10.000000 | 0.000000 | 10.000000 |
| 50% | 10.000000 | 0.000000 | 10.000000 |
| 75% | 10.000000 | 0.000000 | 10.000000 |
| max | 10.000000 | 7.000000 | 10.000000 |

| | D19_TELKO_ONLINE_DATUM | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST \ |
|-------|------------------------|---------------------------|------------------|
| count | 191652.000000 | 143955.000000 | 191652.000000 |
| mean | 9.978002 | 0.013601 | 0.651379 |
| std | 0.260237 | 0.367646 | 1.811244 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | D19_TIERARTIKEL | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 \ |
|-------|-----------------|--------------------|----------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.234247 | 0.771283 | 1.206285 |
| std | 1.167001 | 1.254807 | 1.622334 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 0.000000 | 1.000000 | 2.000000 |
| max | 7.000000 | 6.000000 | 6.000000 |

| | D19_VERSAND_DATUM | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM \ |
|-------|-------------------|---------------------------|----------------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 7.164167 | 8.691237 | 7.699784 |
| std | 3.094218 | 1.987110 | 3.033627 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 5.000000 | 8.000000 | 5.000000 |
| 50% | 9.000000 | 10.000000 | 10.000000 |
| 75% | 10.000000 | 10.000000 | 10.000000 |
| max | 10.000000 | 10.000000 | 10.000000 |

| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 \ |
|-------|-----------------------------|------------------|--------------------|
| count | 143955.000000 | 191652.000000 | 191652.000000 |
| mean | 3.216088 | 0.776767 | 0.101747 |
| std | 4.486796 | 1.884116 | 0.393303 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 |
| 50% | 0.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 7.000000 | 6.000000 |

| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM \ |
|-------|------------------|-----------------|---------------------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 0.175887 | 9.209171 | 9.917298 |
| std | 0.539539 | 1.856680 | 0.563425 |
| min | 0.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 10.000000 | 10.000000 |
| 50% | 0.000000 | 10.000000 | 10.000000 |
| 75% | 0.000000 | 10.000000 | 10.000000 |
| max | 6.000000 | 10.000000 | 10.000000 |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN \ |
|-------|------------------------|---------------------------|----------------------|
| count | 191652.000000 | 143955.000000 | 191652.000000 |
| mean | 9.983162 | 0.017499 | 1.168613 |
| std | 0.261332 | 0.415305 | 2.199257 |
| min | 1.000000 | 0.000000 | 0.000000 |
| 25% | 10.000000 | 0.000000 | 0.000000 |
| 50% | 10.000000 | 0.000000 | 0.000000 |
| 75% | 10.000000 | 0.000000 | 0.000000 |
| max | 10.000000 | 10.000000 | 7.000000 |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG \ |
|-------|-------------------|-------------------|---------------|
| count | 191652.000000 | 191652.000000 | 141725.000000 |
| mean | 2.365892 | 0.788126 | 0.977202 |
| std | 2.841914 | 2.065434 | 0.149258 |
| min | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.000000 | 0.000000 | 1.000000 |
| 50% | 0.000000 | 0.000000 | 1.000000 |
| 75% | 6.000000 | 0.000000 | 1.000000 |
| max | 7.000000 | 7.000000 | 1.000000 |

| | EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER \ |
|-------|----------------------|---------------|---------------|------------------|
| count | 145056.000000 | 141693.000000 | 106369.000000 | 191652.000000 |
| mean | 1999.185053 | 3.881702 | 38.418599 | 2.439808 |
| std | 6.178099 | 1.607621 | 13.689466 | 1.697932 |
| min | 1986.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 1994.000000 | 2.000000 | 29.000000 | 1.000000 |
| 50% | 1997.000000 | 4.000000 | 36.000000 | 2.000000 |
| 75% | 2004.000000 | 5.000000 | 53.000000 | 5.000000 |
| max | 2018.000000 | 6.000000 | 56.000000 | 5.000000 |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER \ |
|-------|------------------|-------------------|-----------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 2.781176 | 3.963773 | 2.057051 |
| std | 1.147353 | 1.036230 | 1.319422 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 | 1.000000 |
| 50% | 3.000000 | 4.000000 | 1.000000 |
| 75% | 3.000000 | 5.000000 | 4.000000 |

| | | | | |
|-------|-----------------------|------------------|---------------|----------------|
| max | 5.000000 | 5.000000 | 5.000000 | |
| | FINANZ_UNAUFFAELLIGER | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE \ |
| count | 191652.000000 | 191652.000000 | 191652.000000 | 141725.000000 |
| mean | 2.646176 | 4.183317 | 4.137958 | 3.580018 |
| std | 1.586506 | 0.995365 | 1.503946 | 1.051072 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 1.000000 | 3.000000 | 2.000000 | 3.000000 |
| 50% | 2.000000 | 5.000000 | 4.000000 | 4.000000 |
| 75% | 5.000000 | 5.000000 | 5.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 6.000000 | 5.000000 |

| | | | | |
|-------|---------------|--------------------|---------------|---------------|
| | GEBAEUDETYP | GEBAEUDETYP_RASTER | GEBURTSJAHR | GEMEINDETYP \ |
| count | 141725.000000 | 141725.000000 | 191652.000000 | 141176.000000 |
| mean | 2.369942 | 3.852524 | 1003.392733 | 24.776683 |
| std | 2.434227 | 0.830285 | 974.531081 | 11.758510 |
| min | 1.000000 | 1.000000 | 0.000000 | 11.000000 |
| 25% | 1.000000 | 3.000000 | 0.000000 | 12.000000 |
| 50% | 1.000000 | 4.000000 | 1926.000000 | 22.000000 |
| 75% | 3.000000 | 4.000000 | 1949.000000 | 30.000000 |
| max | 8.000000 | 5.000000 | 2017.000000 | 50.000000 |

| | | | | |
|-------|-----------------|------------------|---------------|-----------------|
| | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP | HH_DELTA_FLAG \ |
| count | 188439.000000 | 191652.000000 | 191652.000000 | 137910.000000 |
| mean | 6.302268 | 0.367171 | 1.198396 | 0.149714 |
| std | 2.877181 | 0.482035 | 1.450937 | 0.356792 |
| min | 1.000000 | 0.000000 | -1.000000 | 0.000000 |
| 25% | 5.000000 | 0.000000 | -1.000000 | 0.000000 |
| 50% | 5.000000 | 0.000000 | 2.000000 | 0.000000 |
| 75% | 9.000000 | 1.000000 | 2.000000 | 0.000000 |
| max | 12.000000 | 1.000000 | 3.000000 | 1.000000 |

| | | | | |
|-------|--------------------|---------------|---------------|----------------|
| | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 \ |
| count | 188684.000000 | 141693.000000 | 135672.000000 | 135672.000000 |
| mean | 2.942481 | 4.784577 | 1.592075 | 2.797548 |
| std | 1.533347 | 1.961473 | 1.201312 | 1.138182 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 | 1.000000 | 2.000000 |
| 50% | 2.000000 | 5.000000 | 2.000000 | 3.000000 |
| 75% | 4.000000 | 6.000000 | 2.000000 | 3.000000 |
| max | 6.000000 | 8.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 3.170507 | 3.315607 | 1.164795 | 2.207058 |
| std | 1.144635 | 1.197248 | 1.138407 | 1.383509 |
| min | 1.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 3.000000 | 0.000000 | 1.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| 50% | 3.000000 | 3.000000 | 1.000000 | 2.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 4.000000 |

| | | | | |
|-------|---------------|---------------|---------------|------------------|
| | KBA05_ANTG2 | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 1.209004 | 0.309968 | 0.153385 | 3.585139 |
| std | 1.086202 | 0.765292 | 0.476992 | 1.081234 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 0.000000 | 0.000000 | 0.000000 | 3.000000 |
| 50% | 1.000000 | 0.000000 | 0.000000 | 4.000000 |
| 75% | 2.000000 | 0.000000 | 0.000000 | 4.000000 |
| max | 4.000000 | 3.000000 | 2.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_BAUMAX | KBA05_CCM1 | KBA05_CCM2 | KBA05_CCM3 \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 1.043826 | 2.812924 | 2.940784 | 3.276630 |
| std | 1.384973 | 1.098850 | 1.086483 | 1.081889 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 0.000000 | 2.000000 | 2.000000 | 3.000000 |
| 50% | 1.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 1.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 1.544423 | 2.229421 | 3.097640 | 3.630403 |
| std | 1.374065 | 1.109139 | 1.100046 | 1.168496 |
| min | 0.000000 | 0.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 | 3.000000 |
| 50% | 1.000000 | 2.000000 | 3.000000 | 4.000000 |
| 75% | 2.000000 | 3.000000 | 4.000000 | 5.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|----------------|
| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 2.791873 | 3.186457 | 2.832508 | 2.781775 |
| std | 1.382877 | 1.081792 | 1.183915 | 1.251338 |
| min | 0.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 3.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|-----------------|----------------|-------------------|
| | KBA05_HERST5 | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 \ |
| count | 135672.000000 | 141725.000000 | 135672.000000 | 135672.000000 |
| mean | 2.637287 | 2.635901 | 3.432882 | 3.153569 |
| std | 1.281040 | 1.345303 | 1.146148 | 1.102288 |

| | | | | |
|-----|----------|----------|----------|----------|
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 | 2.000000 |
| 50% | 3.000000 | 2.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|-----------------|-----------------|----------------|-----------------|
| | KBA05_KRSHERST2 | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 3.131597 | 2.941042 | 2.013024 | 2.104215 |
| std | 1.128022 | 1.164365 | 0.835436 | 0.829555 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 2.000000 | 2.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 | 2.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_KRSVAN | KBA05_KRSZUL | KBA05_KW1 | KBA05_KW2 \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 2.060749 | 2.095628 | 2.764638 | 3.117489 |
| std | 0.815105 | 0.874546 | 1.145709 | 1.084035 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 50% | 2.000000 | 2.000000 | 3.000000 | 3.000000 |
| 75% | 2.000000 | 2.000000 | 3.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|------------------|
| | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 1.820693 | 3.801086 | 2.600669 | 2.538998 |
| std | 1.373208 | 1.269736 | 1.305333 | 1.175724 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 1.000000 | 3.000000 | 1.000000 | 2.000000 |
| 50% | 2.000000 | 4.000000 | 3.000000 | 2.000000 |
| 75% | 3.000000 | 5.000000 | 4.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 |
| mean | 2.278495 | 1.893906 | 1.681246 | 3.044173 |
| std | 1.112226 | 0.916158 | 1.387116 | 1.078154 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 1.000000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 75% | 3.000000 | 2.000000 | 3.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 |

| | | | | |
|-------|---------------|---------------|---------------|-----------------|
| | KBA05_MOD3 | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 141725.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 3.076766 | 2.759272 | 1.446024 | 2.912831 |
| std | 1.117153 | 1.360024 | 1.133161 | 1.245163 |
| min | 1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 1.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 2.000000 | 4.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 6.000000 |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA05_MOTOR | KBA05_MOTRAD | KBA05_SEG1 | KBA05_SEG10 | \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 | |
| mean | 2.814354 | 1.247280 | 1.158470 | 2.162701 | |
| std | 1.045462 | 1.098361 | 1.087014 | 1.193769 | |
| min | 1.000000 | 0.000000 | 0.000000 | 0.000000 | |
| 25% | 2.000000 | 1.000000 | 0.000000 | 1.000000 | |
| 50% | 3.000000 | 1.000000 | 1.000000 | 2.000000 | |
| 75% | 3.000000 | 2.000000 | 2.000000 | 3.000000 | |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 | \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 | |
| mean | 2.921185 | 2.890265 | 3.049848 | 1.784230 | |
| std | 1.090810 | 1.092005 | 1.068622 | 1.295017 | |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | 1.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 2.000000 | |
| 75% | 3.000000 | 3.000000 | 4.000000 | 3.000000 | |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 | \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 | |
| mean | 0.236792 | 1.001776 | 1.000170 | 1.237050 | |
| std | 0.822108 | 1.142617 | 1.154433 | 1.106325 | |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 | |
| 25% | 0.000000 | 0.000000 | 0.000000 | 0.000000 | |
| 50% | 0.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 75% | 0.000000 | 2.000000 | 2.000000 | 2.000000 | |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 | |

| | | | | | |
|-------|---------------|---------------|---------------|---------------|---|
| | KBA05_VORBO | KBA05_VORB1 | KBA05_VORB2 | KBA05_ZUL1 | \ |
| count | 135672.000000 | 135672.000000 | 135672.000000 | 135672.000000 | |
| mean | 3.305074 | 3.025930 | 2.306968 | 2.785217 | |
| std | 1.156245 | 1.101034 | 1.345094 | 1.076783 | |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 | |
| 25% | 3.000000 | 2.000000 | 1.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 2.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 3.000000 | 3.000000 | |
| max | 9.000000 | 9.000000 | 9.000000 | 9.000000 | |

| | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | KBA13_ALTERHALTER_30 \ |
|-------|---------------|---------------|---------------|------------------------|
| count | 135672.000000 | 135672.000000 | 135672.000000 | 140371.000000 |
| mean | 3.088161 | 2.969124 | 2.453063 | 2.625621 |
| std | 1.086681 | 1.306667 | 1.470195 | 0.967255 |
| min | 1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 2.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 3.000000 |
| max | 9.000000 | 9.000000 | 9.000000 | 5.000000 |

| | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | KBA13_ALTERHALTER_61 \ |
|-------|----------------------|----------------------|------------------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.004203 | 2.926082 | 3.271623 |
| std | 1.048777 | 1.021501 | 1.029104 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 \ |
|-------|---------------|---------------|---------------|---------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.405882 | 2.753076 | 1.406843 | 0.519951 |
| std | 0.858058 | 0.852022 | 0.922689 | 0.644013 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 1.000000 | 0.000000 |
| 50% | 2.000000 | 3.000000 | 1.000000 | 0.000000 |
| 75% | 3.000000 | 3.000000 | 2.000000 | 1.000000 |
| max | 4.000000 | 4.000000 | 3.000000 | 2.000000 |

| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX \ |
|-------|------------------|---------------|-----------------|----------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 667.231216 | 3.218934 | 2.990148 | 1.630358 |
| std | 340.481722 | 0.970351 | 0.961144 | 1.249188 |
| min | 5.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 430.000000 | 3.000000 | 2.000000 | 1.000000 |
| 50% | 593.000000 | 3.000000 | 3.000000 | 1.000000 |
| 75% | 828.000000 | 4.000000 | 4.000000 | 2.000000 |
| max | 2300.000000 | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_BJ_1999 | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 \ |
|-------|---------------|---------------|---------------|-----------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.789608 | 2.715938 | 3.024535 | 3.120260 |
| std | 0.937110 | 0.950800 | 0.919405 | 0.911582 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 4.000000 |

| | | | | |
|-----|----------|----------|----------|----------|
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |
|-----|----------|----------|----------|----------|

| | | | | | |
|-------|---------------|---------------|---------------|------------------|---|
| | KBA13_BJ_2008 | KBA13_BJ_2009 | KBA13_BMW | KBA13_CCM_0_1400 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.698805 | 2.694046 | 3.465652 | 2.125681 | |
| std | 1.474305 | 1.469413 | 0.948646 | 1.392846 | |
| min | 0.000000 | 0.000000 | 1.000000 | 0.000000 | |
| 25% | 2.000000 | 2.000000 | 3.000000 | 1.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 2.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 3.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|----------------|----------------|----------------|---------------------|---|
| | KBA13_CCM_1000 | KBA13_CCM_1200 | KBA13_CCM_1400 | KBA13_CCM_1401_2500 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.248627 | 2.147701 | 2.871405 | 2.938641 | |
| std | 1.389221 | 1.429985 | 0.933020 | 0.919215 | |
| min | 0.000000 | 0.000000 | 1.000000 | 1.000000 | |
| 25% | 1.000000 | 1.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 2.000000 | 3.000000 | 3.000000 | |
| 75% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|----------------|----------------|----------------|----------------|---|
| | KBA13_CCM_1500 | KBA13_CCM_1600 | KBA13_CCM_1800 | KBA13_CCM_2000 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.578645 | 2.942759 | 2.364591 | 3.229193 | |
| std | 1.397120 | 0.915023 | 1.428590 | 0.917430 | |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 | |
| 25% | 1.000000 | 2.000000 | 2.000000 | 3.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 3.000000 | 3.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|----------------|----------------|----------------|----------------|---|
| | KBA13_CCM_2500 | KBA13_CCM_2501 | KBA13_CCM_3000 | KBA13_CCM_3001 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.693177 | 2.778608 | 2.820590 | 2.871946 | |
| std | 1.444370 | 1.454251 | 1.370009 | 1.501274 | |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | 1.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|-----------------|--------------------|---------------|---------------|---|
| | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE | KBA13_FIAT | KBA13_FORD | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.743344 | 2.803314 | 3.242671 | 2.856922 | |
| std | 0.947306 | 0.941188 | 0.957729 | 1.072144 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 3.000000 | 2.000000 | |

| | | | | |
|-----|----------|----------|----------|----------|
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | | |
|-------|---------------|-----------------|-----------------|-----------------|---|
| | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 | KBA13_HALTER_30 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 3.663477 | 2.770537 | 2.615540 | 2.749728 | |
| std | 1.014806 | 0.923539 | 0.931909 | 1.010695 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 4.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 5.000000 | 3.000000 | 3.000000 | 3.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|-----------------|-----------------|-----------------|-----------------|---|
| | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 | KBA13_HALTER_50 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.918886 | 3.027698 | 3.090047 | 2.902131 | |
| std | 1.069563 | 1.024388 | 1.010756 | 1.023940 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | | |
|-------|-----------------|-----------------|-----------------|-----------------|---|
| | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | KBA13_HALTER_66 | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.983593 | 2.985203 | 3.336145 | 3.225709 | |
| std | 1.030085 | 1.021826 | 0.961205 | 1.049287 | |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 2.000000 | 3.000000 | 3.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 4.000000 | 4.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 | |

| | | | | |
|-------|-------------------|---------------------|----------------------|---|
| | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | KBA13_HERST_BMW_BENZ | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 2.760029 | 3.055539 | 3.495637 | |
| std | 0.953098 | 0.974195 | 0.938108 | |
| min | 1.000000 | 1.000000 | 1.000000 | |
| 25% | 2.000000 | 3.000000 | 3.000000 | |
| 50% | 3.000000 | 3.000000 | 3.000000 | |
| 75% | 3.000000 | 4.000000 | 4.000000 | |
| max | 5.000000 | 5.000000 | 5.000000 | |

| | | | | |
|-------|--------------------|-----------------------|-------------------|---|
| | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | KBA13_HERST_SONST | \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | |
| mean | 3.049526 | 2.812960 | 2.803314 | |
| std | 1.028904 | 1.095075 | 0.941188 | |

| | | | |
|-----|----------|----------|----------|
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|-----------------|---------------|-----------------|
| | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | KBA13_KMH_140 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.595429 | 2.188700 | 1.282879 | 2.638479 |
| std | 0.935729 | 1.524135 | 0.636897 | 1.351953 |
| min | 1.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 1.000000 | 1.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 1.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 1.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 3.000000 | 5.000000 |

| | | | | |
|-------|-------------------|---------------|---------------|-----------------|
| | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | KBA13_KMH_211 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.678837 | 2.730386 | 3.138056 | 2.745482 |
| std | 0.981039 | 0.971710 | 0.926822 | 1.510260 |
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|----------------|--------------------------|
| | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | KBA13_KRSHERST_AUDI_VW \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.73912 | 1.340818 | 2.996431 | 3.087148 |
| std | 1.50962 | 0.741228 | 0.995513 | 0.983040 |
| min | 0.00000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.00000 | 1.000000 | 2.000000 | 3.000000 |
| 50% | 3.00000 | 1.000000 | 3.000000 | 3.000000 |
| 75% | 4.00000 | 1.000000 | 4.000000 | 4.000000 |
| max | 5.00000 | 3.000000 | 5.000000 | 5.000000 |

| | | | |
|-------|-------------------------|--------------------------|----------------------|
| | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | KBA13_KRSSEG_KLEIN \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.268268 | 2.826674 | 1.976712 |
| std | 0.949651 | 1.058525 | 0.278902 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 2.000000 |
| 75% | 4.000000 | 3.000000 | 2.000000 |
| max | 5.000000 | 5.000000 | 3.000000 |

| | | | | |
|-------|-------------------|------------------|------------------|-----------------|
| | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | KBA13_KW_0_60 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 2.025198 | 1.980815 | 1.999964 | 2.767651 |
| std | 0.538547 | 0.591234 | 0.739867 | 0.954970 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 75% | 2.000000 | 2.000000 | 3.000000 | 3.000000 |
| max | 3.000000 | 3.000000 | 3.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.607077 | 2.508488 | 2.775751 | 1.360979 |
| std | 1.425966 | 1.554176 | 1.474828 | 0.643797 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 2.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 1.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 | 2.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 3.000000 |

| | | | | |
|-------|---------------|---------------|---------------|-------------------|
| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.190944 | 2.170370 | 2.165063 | 3.084868 |
| std | 1.360301 | 1.412521 | 1.380792 | 0.923064 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 1.000000 | 1.000000 | 3.000000 |
| 50% | 2.000000 | 2.000000 | 2.000000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|---------------|
| | KBA13_KW_70 | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.255231 | 2.269942 | 2.42335 | 2.951564 |
| std | 1.415183 | 1.389212 | 1.42951 | 0.978634 |
| min | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | 1.000000 | 1.000000 | 2.00000 | 2.000000 |
| 50% | 2.000000 | 2.000000 | 3.00000 | 3.000000 |
| 75% | 3.000000 | 3.000000 | 3.00000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.00000 | 5.000000 |

| | | | | |
|-------|----------------|---------------|---------------|---------------|
| | KBA13_MERCEDES | KBA13_MOTOR | KBA13_NISSAN | KBA13_OPEL \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.413917 | 2.774889 | 2.817306 | 2.858226 |
| std | 0.933214 | 0.781071 | 0.983231 | 1.094689 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 4.000000 | 5.000000 | 5.000000 |

| | KBA13_PEUGEOT | KBA13_RENAULT | KBA13_SEG_GELAENDEWAGEN \ |
|-------|---------------|---------------|---------------------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.129842 | 2.884221 | 3.144937 |
| std | 1.010289 | 1.002776 | 0.937455 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST | KBA13_SEG_KLEINWAGEN \ |
|-------|-------------------------|-------------------|------------------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.308247 | 2.816992 | 2.760848 |
| std | 0.962788 | 1.010607 | 0.970130 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS | KBA13_SEG_MINIWAGEN \ |
|-------|-------------------------|--------------------|-----------------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.743964 | 3.081491 | 3.114311 |
| std | 0.920321 | 0.976794 | 0.947762 |
| min | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE \ |
|-------|------------------------|-------------------------------|
| count | 140371.000000 | 140371.000000 |
| mean | 2.981143 | 3.355237 |
| std | 0.973991 | 0.935366 |
| min | 1.000000 | 1.000000 |
| 25% | 2.000000 | 3.000000 |
| 50% | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 |

| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN \ |
|-------|----------------------|--------------------|------------------------|
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 2.728291 | 2.962093 | 2.956465 |
| std | 1.474939 | 0.905255 | 1.445584 |
| min | 0.000000 | 1.000000 | 0.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 |

| | | | |
|-----|----------|----------|----------|
| max | 5.000000 | 5.000000 | 5.000000 |
|-----|----------|----------|----------|

| | | | |
|-------|---------------------|---------------|------------------------|
| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.080038 | 3.214296 | 2.669405 |
| std | 0.940034 | 0.971683 | 1.404804 |
| min | 1.000000 | 1.000000 | 0.000000 |
| 25% | 3.000000 | 3.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|---------------|----------------|
| | KBA13_SITZE_4 | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.454531 | 2.514159 | 3.331058 | 3.057426 |
| std | 0.939839 | 0.951726 | 0.977886 | 0.983471 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 2.000000 | 3.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 3.000000 | 4.000000 | 4.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|----------------|----------------|
| | KBA13_VORB_0 | KBA13_VORB_1 | KBA13_VORB_1_2 | KBA13_VORB_2 \ |
| count | 140371.000000 | 140371.000000 | 140371.000000 | 140371.000000 |
| mean | 3.375626 | 3.047132 | 2.852861 | 2.854308 |
| std | 0.956641 | 0.925362 | 0.961431 | 0.921046 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 3.000000 | 2.000000 | 2.000000 |
| 50% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 75% | 4.000000 | 4.000000 | 3.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 5.000000 |

| | | | | |
|-------|---------------|---------------|--------------|---------------|
| | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK \ |
| count | 140371.000000 | 140371.000000 | 79715.000000 | 137392.000000 |
| mean | 2.011911 | 3.002358 | 3.421803 | 2.321569 |
| std | 1.375629 | 0.972860 | 1.623890 | 1.100512 |
| min | 0.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 2.000000 | 2.000000 | 1.000000 |
| 50% | 2.000000 | 3.000000 | 3.000000 | 2.000000 |
| 75% | 3.000000 | 4.000000 | 5.000000 | 3.000000 |
| max | 5.000000 | 5.000000 | 6.000000 | 4.000000 |

| | | | | |
|-------|---------------|---------------|---------------|-------------------|
| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN \ |
| count | 191652.000000 | 145001.000000 | 141725.000000 | 188439.000000 |
| mean | 4.99784 | 3.129978 | 0.177146 | 4.254448 |
| std | 2.34032 | 1.439740 | 0.381793 | 4.492807 |
| min | 1.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 4.000000 | 2.000000 | 0.000000 | 0.000000 |

| | | | | |
|-----|---------|----------|----------|-----------|
| 50% | 4.00000 | 3.000000 | 0.000000 | 2.000000 |
| 75% | 4.00000 | 4.000000 | 0.000000 | 10.000000 |
| max | 9.00000 | 7.000000 | 1.000000 | 11.000000 |

| | LP_FAMILIE_GROB | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB | \ |
|-------|-----------------|---------------------|---------------------|---|
| count | 188439.000000 | 188439.000000 | 188439.000000 | |
| mean | 2.355043 | 18.181571 | 5.422694 | |
| std | 2.052142 | 15.009985 | 4.717907 | |
| min | 0.000000 | 0.000000 | 0.000000 | |
| 25% | 0.000000 | 0.000000 | 0.000000 | |
| 50% | 2.000000 | 16.000000 | 4.000000 | |
| 75% | 5.000000 | 36.000000 | 12.000000 | |
| max | 5.000000 | 40.000000 | 12.000000 | |

| | LP_STATUS_FEIN | LP_STATUS_GROB | MIN_GEBAEUDEJAHR | MOBI_RASTER | \ |
|-------|----------------|----------------|------------------|---------------|---|
| count | 188439.000000 | 188439.000000 | 141725.000000 | 141725.000000 | |
| mean | 6.687910 | 3.209251 | 1993.056659 | 2.900363 | |
| std | 3.090573 | 1.436958 | 3.080241 | 1.527411 | |
| min | 1.000000 | 1.000000 | 1985.000000 | 1.000000 | |
| 25% | 5.000000 | 2.000000 | 1992.000000 | 1.000000 | |
| 50% | 7.000000 | 3.000000 | 1992.000000 | 3.000000 | |
| 75% | 10.000000 | 5.000000 | 1992.000000 | 4.000000 | |
| max | 10.000000 | 5.000000 | 2016.000000 | 6.000000 | |

| | MOBI_REGIO | NATIONALITAET_KZ | ONLINE_AFFINITAET | ORTSGR_KLS9 | \ |
|-------|---------------|------------------|-------------------|---------------|---|
| count | 135672.000000 | 191652.000000 | 188439.000000 | 141176.000000 | |
| mean | 3.627425 | 0.773433 | 2.764327 | 5.119517 | |
| std | 1.282444 | 0.495983 | 1.266050 | 2.159184 | |
| min | 1.000000 | 0.000000 | 0.000000 | 1.000000 | |
| 25% | 3.000000 | 0.000000 | 2.000000 | 4.000000 | |
| 50% | 4.000000 | 1.000000 | 2.000000 | 5.000000 | |
| 75% | 5.000000 | 1.000000 | 4.000000 | 7.000000 | |
| max | 6.000000 | 3.000000 | 5.000000 | 9.000000 | |

| | PLZ8_ANTG1 | PLZ8_ANTG2 | PLZ8_ANTG3 | PLZ8_ANTG4 | \ |
|-------|---------------|---------------|---------------|---------------|---|
| count | 138888.000000 | 138888.000000 | 138888.000000 | 138888.000000 | |
| mean | 2.537404 | 2.731510 | 1.388817 | 0.539939 | |
| std | 0.904927 | 0.840922 | 0.885038 | 0.637048 | |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 | |
| 25% | 2.000000 | 2.000000 | 1.000000 | 0.000000 | |
| 50% | 3.000000 | 3.000000 | 1.000000 | 0.000000 | |
| 75% | 3.000000 | 3.000000 | 2.000000 | 1.000000 | |
| max | 4.000000 | 4.000000 | 3.000000 | 2.000000 | |

| | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ | PRAEGENDE_JUGENDJAHRE | \ |
|-------|---------------|---------------|---------------|-----------------------|---|
| count | 138888.000000 | 138888.000000 | 138888.000000 | 191652.000000 | |
| mean | 1.556607 | 3.622192 | 3.634893 | 4.248273 | |
| std | 1.185736 | 1.008472 | 0.933660 | 3.807671 | |

| | | | | |
|-----|----------|----------|----------|-----------|
| min | 1.000000 | 1.000000 | 1.000000 | 0.000000 |
| 25% | 1.000000 | 3.000000 | 3.000000 | 0.000000 |
| 50% | 1.000000 | 4.000000 | 3.000000 | 4.000000 |
| 75% | 1.000000 | 4.000000 | 4.000000 | 6.000000 |
| max | 5.000000 | 5.000000 | 5.000000 | 15.000000 |

| | REGIOTYP | RELAT_AB | RETOURTYT_BK_S | RT_KEIN_ANREIZ \ |
|-------|---------------|---------------|----------------|------------------|
| count | 137392.000000 | 141176.000000 | 188439.000000 | 188439.000000 |
| mean | 3.814341 | 2.898510 | 3.716311 | 2.462484 |
| std | 2.075155 | 1.422683 | 1.113933 | 1.315405 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 3.000000 | 1.000000 |
| 50% | 4.000000 | 3.000000 | 3.000000 | 2.000000 |
| 75% | 6.000000 | 4.000000 | 5.000000 | 4.000000 |
| max | 7.000000 | 9.000000 | 5.000000 | 5.000000 |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL \ |
|-------|-----------------|-----------------|---------------|---------------|
| count | 188439.000000 | 147460.000000 | 191652.000000 | 191652.000000 |
| mean | 4.528516 | 2.519653 | 4.483835 | 4.408021 |
| std | 0.939708 | 1.356235 | 1.631941 | 1.720909 |
| min | 1.000000 | 0.000000 | 1.000000 | 1.000000 |
| 25% | 4.000000 | 1.000000 | 3.000000 | 3.000000 |
| 50% | 5.000000 | 2.000000 | 5.000000 | 4.000000 |
| 75% | 5.000000 | 3.000000 | 6.000000 | 6.000000 |
| max | 5.000000 | 5.000000 | 7.000000 | 7.000000 |

| | SEMIO_FAM | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 4.414026 | 4.187246 | 4.674535 | 3.682497 |
| std | 1.733128 | 1.872047 | 2.041059 | 1.573090 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 3.000000 | 3.000000 | 3.000000 | 3.000000 |
| 50% | 5.000000 | 4.000000 | 5.000000 | 4.000000 |
| 75% | 6.000000 | 6.000000 | 7.000000 | 5.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT | SEMIO_RAT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 5.366477 | 3.883163 | 3.528254 | 3.165879 |
| std | 1.443103 | 1.840131 | 1.493916 | 1.316221 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 5.000000 | 2.000000 | 2.000000 | 2.000000 |
| 50% | 5.000000 | 4.000000 | 4.000000 | 3.000000 |
| 75% | 7.000000 | 5.000000 | 5.000000 | 4.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 191652.000000 | 191652.000000 | 191652.000000 |

| | | | | |
|------|----------|----------|----------|----------|
| mean | 4.112788 | 3.742137 | 2.919161 | 4.185279 |
| std | 2.070958 | 1.718039 | 1.219224 | 2.367407 |
| min | 1.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 2.000000 | 2.000000 | 2.000000 | 1.000000 |
| 50% | 4.000000 | 3.000000 | 3.000000 | 5.000000 |
| 75% | 7.000000 | 6.000000 | 4.000000 | 6.000000 |
| max | 7.000000 | 7.000000 | 7.000000 | 7.000000 |

| | SHOPPER_TYP | SOHO_KZ | STRUKTURTYP | TITEL_KZ \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 145056.000000 | 141176.000000 | 145056.000000 |
| mean | 0.987754 | 0.009865 | 2.567759 | 0.021681 |
| std | 1.538679 | 0.098833 | 0.724788 | 0.210424 |
| min | -1.000000 | 0.000000 | 1.000000 | 0.000000 |
| 25% | -1.000000 | 0.000000 | 2.000000 | 0.000000 |
| 50% | 1.000000 | 0.000000 | 3.000000 | 0.000000 |
| 75% | 3.000000 | 0.000000 | 3.000000 | 0.000000 |
| max | 3.000000 | 1.000000 | 3.000000 | 5.000000 |

| | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG | VERDICHTUNGSRAUM \ |
|-------|---------------|---------------|------------------|--------------------|
| count | 141204.000000 | 141204.000000 | 145056.000000 | 141176.000000 |
| mean | 2.968011 | 4.335770 | 0.084747 | 5.055087 |
| std | 1.312128 | 0.979701 | 0.278505 | 9.250235 |
| min | 1.000000 | 1.000000 | 0.000000 | 0.000000 |
| 25% | 2.000000 | 4.000000 | 0.000000 | 0.000000 |
| 50% | 3.000000 | 5.000000 | 0.000000 | 1.000000 |
| 75% | 4.000000 | 5.000000 | 0.000000 | 6.000000 |
| max | 5.000000 | 5.000000 | 1.000000 | 45.000000 |

| | VERS_TYP | VHA | VHN | VK_DHT4A \ |
|-------|---------------|---------------|---------------|---------------|
| count | 191652.000000 | 145056.000000 | 137392.000000 | 143781.000000 |
| mean | 0.849008 | 0.868534 | 2.429508 | 4.374417 |
| std | 1.166162 | 1.320530 | 1.148821 | 2.924355 |
| min | -1.000000 | 0.000000 | 0.000000 | 1.000000 |
| 25% | -1.000000 | 0.000000 | 2.000000 | 2.000000 |
| 50% | 1.000000 | 0.000000 | 2.000000 | 4.000000 |
| 75% | 2.000000 | 1.000000 | 3.000000 | 7.000000 |
| max | 2.000000 | 5.000000 | 4.000000 | 11.000000 |

| | VK_DISTANZ | VK_ZG11 | W_KEIT_KIND_HH | WOHNDAUER_2008 \ |
|-------|---------------|---------------|----------------|------------------|
| count | 143781.000000 | 143781.000000 | 137910.000000 | 145056.000000 |
| mean | 4.564769 | 3.168868 | 4.152716 | 8.646371 |
| std | 2.887035 | 2.233516 | 1.974375 | 1.154001 |
| min | 1.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 2.000000 | 9.000000 |
| 50% | 4.000000 | 3.000000 | 5.000000 | 9.000000 |
| 75% | 7.000000 | 4.000000 | 6.000000 | 9.000000 |
| max | 13.000000 | 11.000000 | 6.000000 | 9.000000 |

| | WOHNLAG | ZABEOTYP | ONLINE_PURCHASE | ANREDE_KZ \ |
|-------|---------------|---------------|-----------------|---------------|
| count | 141725.000000 | 191652.000000 | 191652.000000 | 191652.000000 |
| mean | 3.723133 | 2.576806 | 0.090247 | 1.376432 |
| std | 2.095540 | 1.168486 | 0.286536 | 0.484492 |
| min | 0.000000 | 1.000000 | 0.000000 | 1.000000 |
| 25% | 2.000000 | 1.000000 | 0.000000 | 1.000000 |
| 50% | 3.000000 | 3.000000 | 0.000000 | 1.000000 |
| 75% | 5.000000 | 3.000000 | 0.000000 | 2.000000 |
| max | 8.000000 | 6.000000 | 1.000000 | 2.000000 |

| | ALTERSKATEGORIE_GROB |
|-------|----------------------|
| count | 191652.000000 |
| mean | 3.060907 |
| std | 1.086254 |
| min | 1.000000 |
| 25% | 3.000000 |
| 50% | 3.000000 |
| 75% | 4.000000 |
| max | 9.000000 |

Further consideration of datasets, categorical variables

In [10]: # Look at extra columns

```

extra = ['CUSTOMER_GROUP', 'ONLINE_PURCHASE', 'PRODUCT_GROUP']
customers['ONLINE_PURCHASE'] = customers['ONLINE_PURCHASE'].astype(str)
customers['ONLINE_PURCHASE'].replace({'0': 'No', '1': 'Yes'}, inplace=True)

print(customers[extra].dtypes)

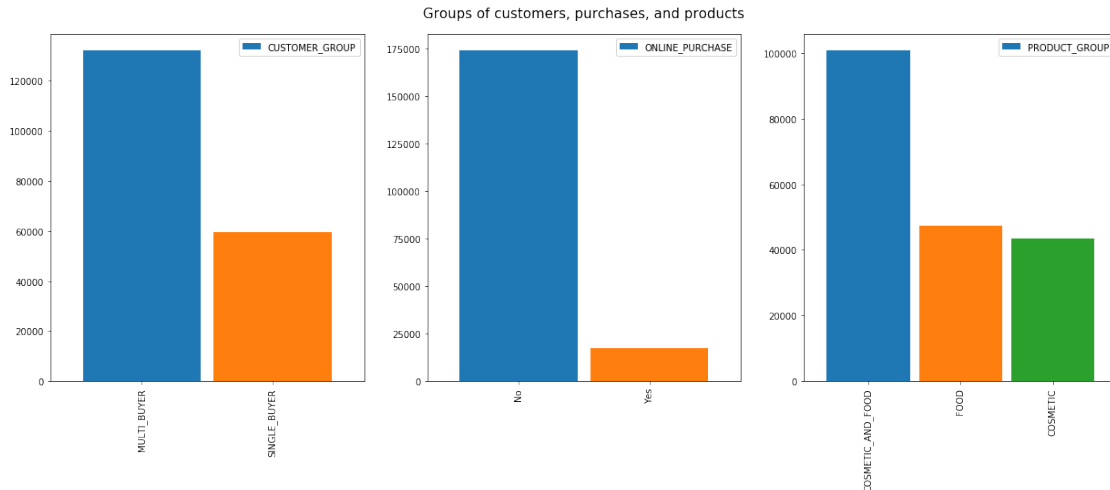
fig, ax = plt.subplots(nrows=1, ncols=3, figsize=(21,7))
for i in range(3):
    customers[extra[i]].value_counts().plot(kind='bar', ax=ax[i], width=.9)
    ax[i].legend(loc='best')
ax[1].set_title('Groups of customers, purchases, and products', fontsize=15, y=1.03)
plt.show()

```

```

CUSTOMER_GROUP    object
ONLINE_PURCHASE    object
PRODUCT_GROUP      object
dtype: object

```



In [11]: *# Save this additional info in another customers_extra DF*

```
extra = ['CUSTOMER_GROUP', 'ONLINE_PURCHASE', 'PRODUCT_GROUP']
customers_extra = customers[extra]
customers.drop(extra, axis=1, inplace=True)
```

In [12]: `cat_cols_azdias = azdias.select_dtypes(include='object').columns`
`cat_cols_customers = customers.select_dtypes(include='object').columns`
`set(cat_cols_customers).difference(set(cat_cols_azdias))`

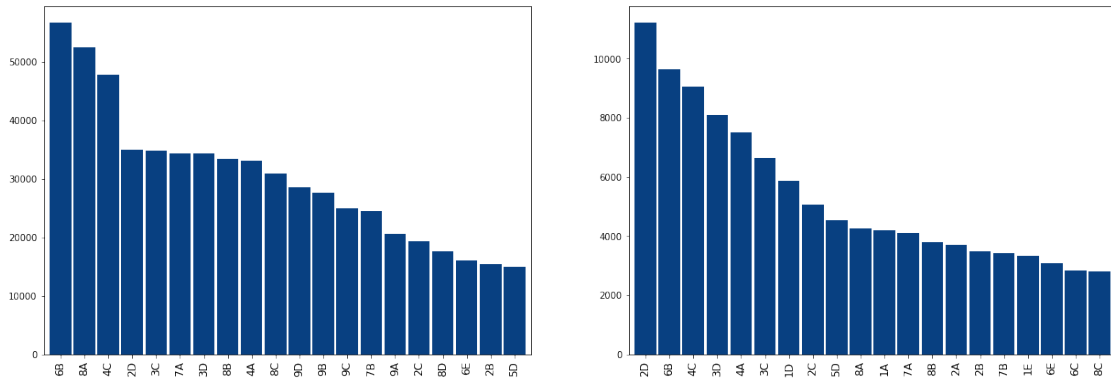
Out[12]: `set()`

In [13]: *# Plot up to top-20*

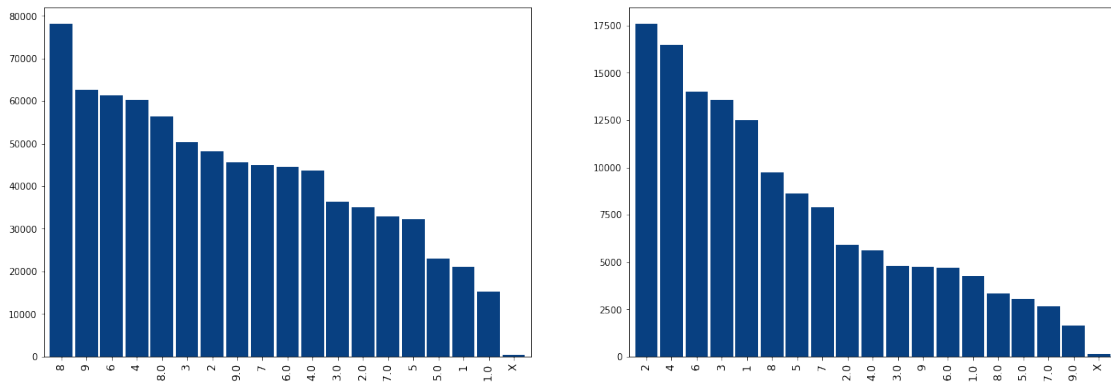
```
plt.rc('xtick', labels=12)

for column in cat_cols_azdias:
    if column != 'EINGEFUEGT_AM':
        fig, ax = plt.subplots(nrows=1, ncols=2, figsize=(21,7))
        azdias[column].value_counts().iloc[:20].plot(kind='bar', ax=ax[0], width=0.9, c
        customers[column].value_counts().iloc[:20].plot(kind='bar', ax=ax[1], width=0.9
        plt.suptitle(column, fontsize=15)
        plt.show()
```

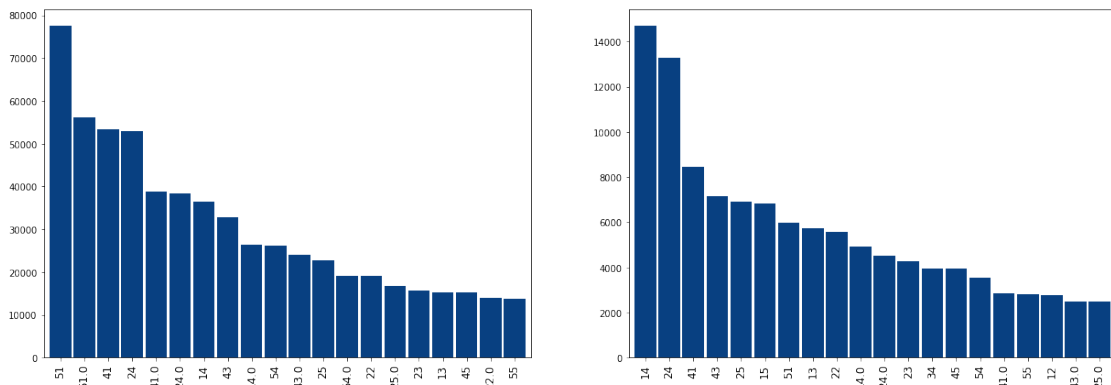
CAMEO_DEU_2015



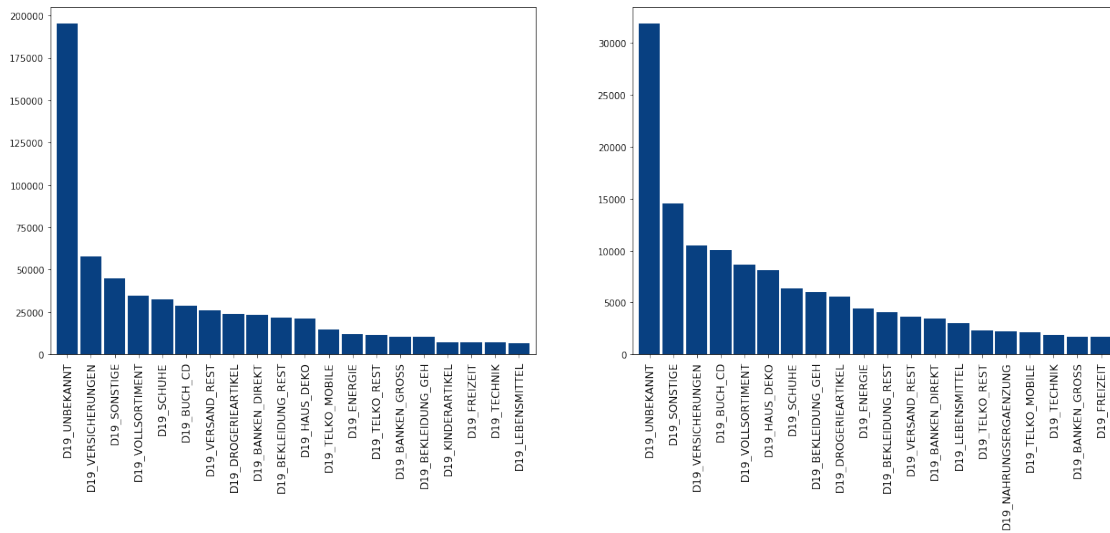
CAMEO_DEUG_2015



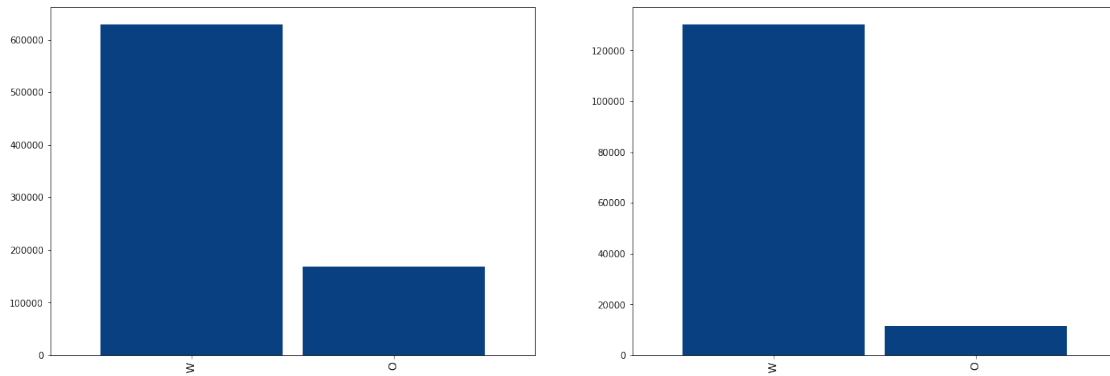
CAMEO_INTL_2015



D19_LETZTER_KAUF_BRANCHE



OST_WEST_KZ



```
In [14]: print(azdias['EINGEFUEGT_AM'].dtype)
          print(azdias['EINGEFUEGT_AM'].unique())
          azdias['EINGEFUEGT_AM'].value_counts().describe()
```

object

```
[nan '1992-02-10 00:00:00' '1992-02-12 00:00:00' ..., '2010-12-02 00:00:00'
 '2005-03-19 00:00:00' '2011-11-18 00:00:00']
```

```
Out[14]: count      5162.000000
          mean        154.605386
          std         5977.496782
          min           1.000000
```



```

25%          2.000000
50%          6.000000
75%         19.000000
max        383738.000000
Name: EINGEFUEGT_AM, dtype: float64

```

```

In [15]: print(customers['EINGEFUEGT_AM'].dtype)
print(customers['EINGEFUEGT_AM'].unique())
customers['EINGEFUEGT_AM'].value_counts().describe()

```

```

object
['1992-02-12 00:00:00' nan '1992-02-10 00:00:00' ..., '2008-04-25 00:00:00'
 '2005-03-30 00:00:00' '2008-07-14 00:00:00']

```

```

Out[15]: count      3034.000000
mean         46.712261
std         1417.970544
min           1.000000
25%           1.000000
50%           3.000000
75%           8.000000
max        64744.000000
Name: EINGEFUEGT_AM, dtype: float64

```

Work with DIAS Information Levels and Values / 2017

```

In [16]: # load in a top-level list of attributes and descriptions, organized by informational c

```

```

levels_df = pd.read_excel('DIAS Information Levels - Attributes 2017.xlsx', skiprows=1)
levels_df = levels_df[['Information level', 'Attribute', 'Description']]
levels_df.loc[0, 'Information level'] = 'Person'
levels_df.loc[88:96, 'Information level'] = 'Microcell (RR3_ID)'
levels_df['Information level'] = levels_df['Information level'].fillna(method='ffill')
levels_df.head()

```

```

Out[16]:  Information level      Attribute \
0          Person      AGER_TYP
1          Person  ALTERSKATEGORIE_GROB
2          Person      ANREDE_KZ
3          Person      CJT_GESAMTTYP
4          Person  FINANZ_MINIMALIST

      Description
0      best-ager typology
1      age through prename analysis
2          gender
3  Customer-Journey-Typology relating to the pref...
4      financial typology: low financial interest

```

```

In [17]: levels_df['Information level'].unique()

Out[17]: array(['Person', 'Household', 'Building', 'Microcell (RR4_ID)',
                'Microcell (RR3_ID)', '125m x 125m Grid', 'Postcode ', 'RR1_ID',
                'PLZ8', 'Community'], dtype=object)

In [18]: # load in a detailed mapping of data values for each feature in alphabetical order

values_df = pd.read_excel('DIAS Attributes - Values 2017.xlsx', skiprows=1)
values_df.drop('Unnamed: 0', axis=1, inplace=True)
values_df = values_df.fillna(method='ffill')
values_df.head()

Out[18]:
  Attribute      Description Value      Meaning
0  AGER_TYP  best-ager typology   -1      unknown
1  AGER_TYP  best-ager typology    0  no classification possible
2  AGER_TYP  best-ager typology    1      passive elderly
3  AGER_TYP  best-ager typology    2      cultural elderly
4  AGER_TYP  best-ager typology    3  experience-driven elderly

In [19]: details = values_df.merge(levels_df, on=['Attribute', 'Description'], how='left')
details.head()

Out[19]:
  Attribute      Description Value      Meaning \
0  AGER_TYP  best-ager typology   -1      unknown
1  AGER_TYP  best-ager typology    0  no classification possible
2  AGER_TYP  best-ager typology    1      passive elderly
3  AGER_TYP  best-ager typology    2      cultural elderly
4  AGER_TYP  best-ager typology    3  experience-driven elderly

  Information level
0      Person
1      Person
2      Person
3      Person
4      Person

In [20]: details.isnull().sum()

Out[20]:
Attribute      0
Description    0
Value          0
Meaning        0
Information level    662
dtype: int64

In [21]: details.head(10)

Out[21]:
  Attribute      Description Value \
0  AGER_TYP  best-ager typology   -1

```

| | | | |
|---|----------------------|---|-------|
| 1 | AGER_TYP | best-ager typology | 0 |
| 2 | AGER_TYP | best-ager typology | 1 |
| 3 | AGER_TYP | best-ager typology | 2 |
| 4 | AGER_TYP | best-ager typology | 3 |
| 5 | ALTERSKATEGORIE_GROB | age classification through prename analysis | -1, 0 |
| 6 | ALTERSKATEGORIE_GROB | age classification through prename analysis | 1 |
| 7 | ALTERSKATEGORIE_GROB | age classification through prename analysis | 2 |
| 8 | ALTERSKATEGORIE_GROB | age classification through prename analysis | 3 |
| 9 | ALTERSKATEGORIE_GROB | age classification through prename analysis | 4 |

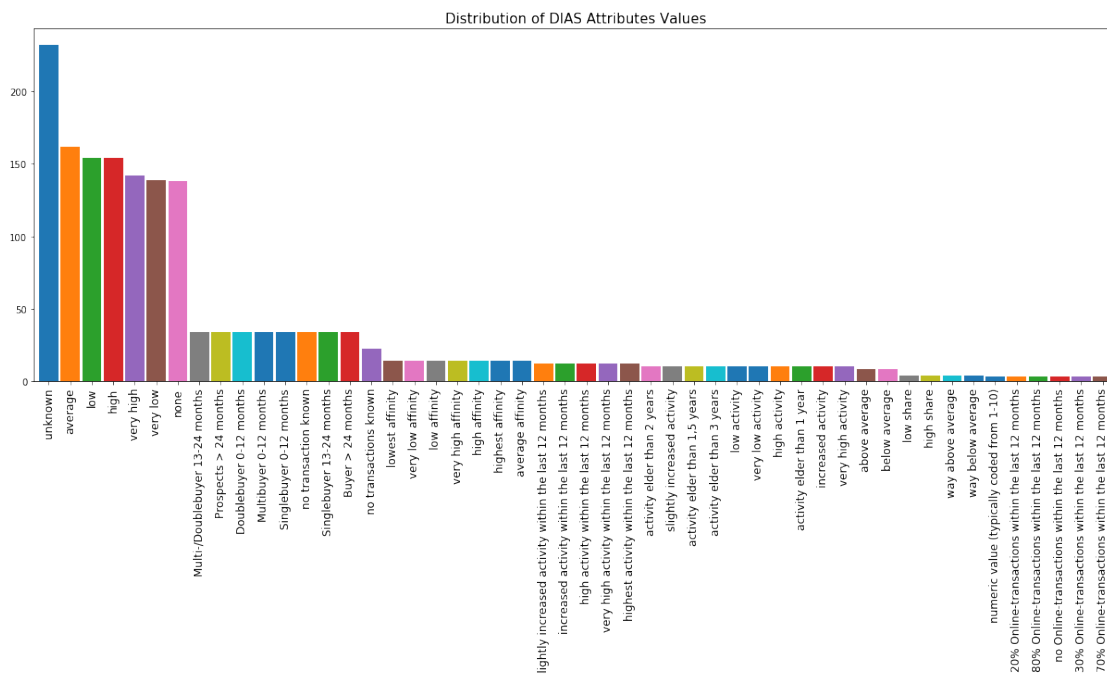
| | Meaning | Information level |
|---|----------------------------|-------------------|
| 0 | unknown | Person |
| 1 | no classification possible | Person |
| 2 | passive elderly | Person |
| 3 | cultural elderly | Person |
| 4 | experience-driven elderly | Person |
| 5 | unknown | NaN |
| 6 | < 30 years | NaN |
| 7 | 30 - 45 years | NaN |
| 8 | 46 - 60 years | NaN |
| 9 | > 60 years | NaN |

In [22]: # Number of all values and unique Attributes in DIAS data

details.shape, details.Attribute.unique().shape

Out[22]: ((2258, 5), (314,))

In [23]: details.Meaning.value_counts().iloc[:50].plot.bar(figsize=(21,7), width=.9)
plt.title('Distribution of DIAS Attributes Values', fontsize=15)
plt.show()



In the following part such values as "unknown" or "unknown / no main age detectable", etc. will be replaced with NaNs.

Replace unknown values in azdias and customers with NaN using details DF

```
In [24]: # Get list of Attribute - Value where "unknown" or "no transaction(s) known" is present
```

```
details['meaning'] = details.Meaning.str.lower().astype(str)
unknown = details[(details.meaning.str.contains('unknown')) |
                  (details.meaning.isin(['no transactions known', 'no transaction known'])
unknown.set_index('Attribute', inplace=True)
unknown.head()
```

```
Out[24]:
```

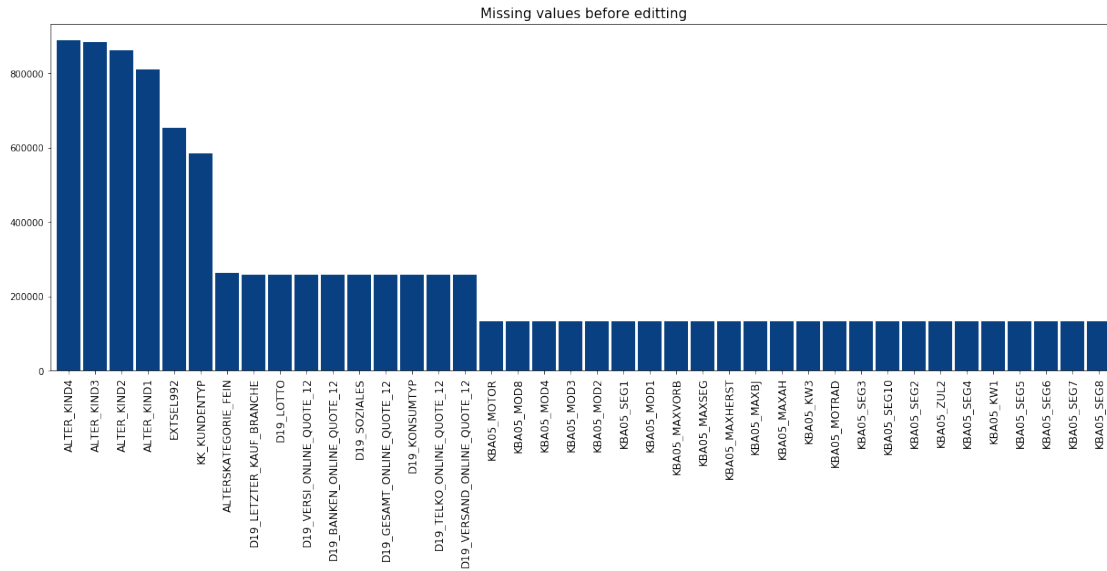
| | Value |
|----------------------|-------|
| Attribute | |
| AGER_TYP | -1 |
| ALTERSKATEGORIE_GROB | -1, 0 |
| ALTER_HH | 0 |
| ANREDE_KZ | -1, 0 |
| BALLRAUM | -1 |

```
In [25]: def replace_unknown_with_nan(df, unknown):
```

```
    cols_df = df.columns
    for column in unknown.index:
        if column in cols_df:
            col_values = df[column].unique().tolist()
            unknown_vals = unknown.loc[column]['Value']
            for val in col_values:
                if isinstance(unknown_vals, int):
                    if val == unknown_vals:
                        df[column] = df[column].replace(val, np.nan)
                else:
                    if str(val) in unknown_vals.split():
                        df[column] = df[column].replace(val, np.nan)
    return df
```

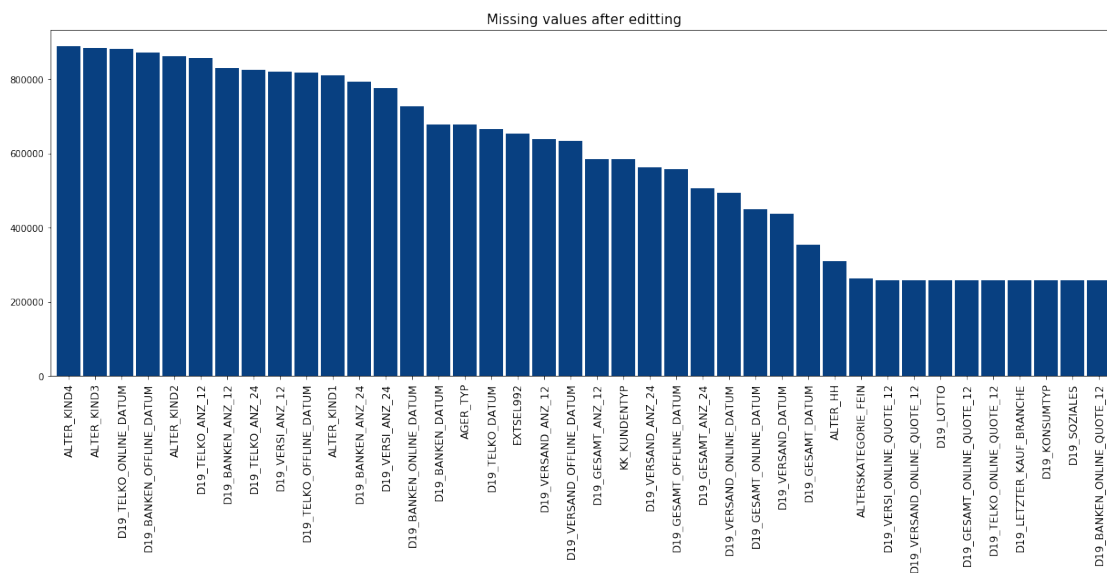
Apply function replace_unknown_with_nan to azdias dataframe

```
In [26]: azdias.isnull().sum().sort_values(ascending=False).iloc[:40].plot.bar(figsize=(21,7), w
plt.title('Missing values before editing', fontsize=15)
plt.show()
```



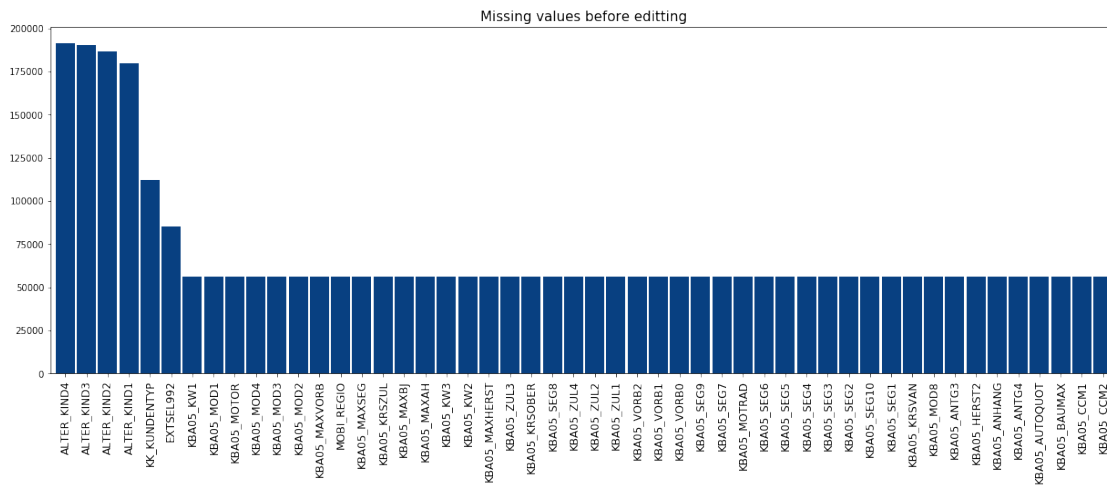
```
In [27]: azdias = replace_unknown_with_nan(azdias, unknown)
```

```
In [28]: azdias.isnull().sum().sort_values(ascending=False).iloc[:40].plot.bar(figsize=(21,7), w
plt.title('Missing values after editing', fontsize=15)
plt.show()
```



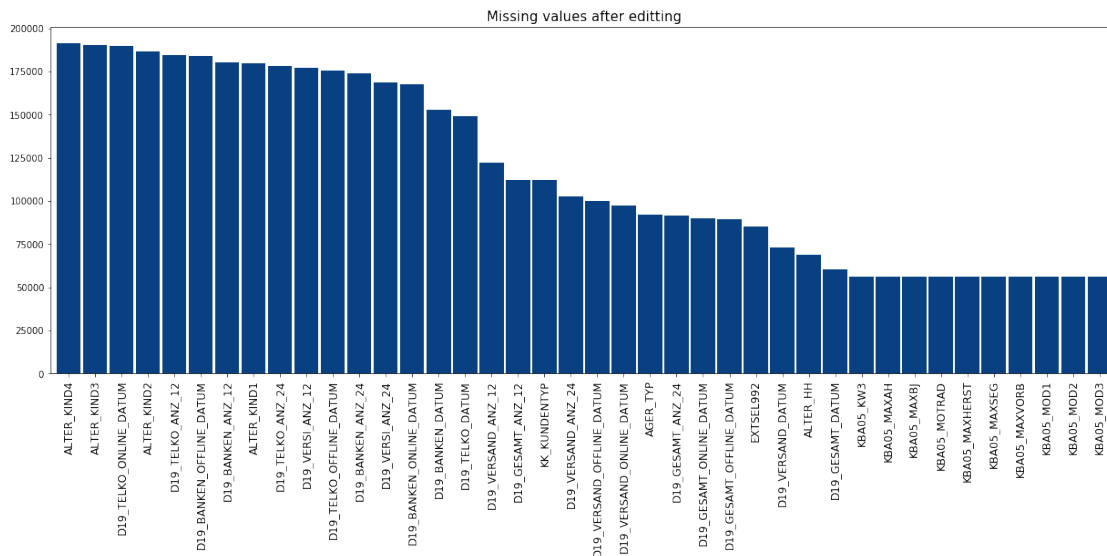
Apply function `replace_unknown_with_nan` to customers dataframe

```
In [29]: customers.isnull().sum().sort_values(ascending=False).iloc[:50].plot.bar(figsize=(21,7))
plt.title('Missing values before editing', fontsize=15)
plt.show()
```



```
In [30]: customers = replace_unknown_with_nan(customers, unknown)
```

```
In [31]: customers.isnull().sum().sort_values(ascending=False).iloc[:40].plot.bar(figsize=(21,7))
plt.title('Missing values after editing', fontsize=15)
plt.show()
```



Drop columns from azdias and customers where there's no corresponding value in Attribute in details DF

```
In [32]: # Differences and similarities in azdias and details DF
```

```
common_cols = set(azdias.columns.tolist()).intersection(set(details.Attribute.unique()))
different_cols = set(azdias.columns.tolist()).symmetric_difference(set(details.Attribute.unique()))
different_azdias_cols = set(azdias.columns.tolist()).difference(set(details.Attribute.unique()))
different_details_cols = set(details.Attribute.tolist()).difference(set(azdias.columns.tolist()))

print('Same for both azdias and details', len(common_cols))
print('Different for both azdias and details', len(different_cols))
print('Different for azdias', len(different_azdias_cols))
print('Different for details', len(different_details_cols))
```

Same for both azdias and details 272

Different for both azdias and details 136

Different for azdias 94

Different for details 42

```
In [33]: # Differences and similarities in customers and details DF
```

```
common_cols = set(customers.columns.tolist()).intersection(set(details.Attribute.unique()))
different_cols = set(customers.columns.tolist()).symmetric_difference(set(details.Attribute.unique()))
different_customers_cols = set(customers.columns.tolist()).difference(set(details.Attribute.unique()))
different_details_cols = set(details.Attribute.tolist()).difference(set(customers.columns.tolist()))

print('Same for both customers and details', len(common_cols))
print('Different for both customers and details', len(different_cols))
print('Different for customers', len(different_customers_cols))
print('Different for details', len(different_details_cols))
```

Same for both customers and details 272

Different for both customers and details 136

Different for customers 94

Different for details 42

It seems that D19 columns with "RZ" in the end of column names in DIAS Attributes DF ("RZ" stands for "Herzogtum Lauenburg" region in Germany) means the same as the corresponding columns of total population. So, the unknown or missing values could also be replaced with NaN. The same could be applied to columns CAMEO_INTL_2015 (azdias DF) and CAMEO_DEU_INTL_2015 () are the same (and probably "DEU" means "Deutsche Eislauf-Union", but it's not exact).

```
In [34]: details[details.Attribute == 'D19_BANKEN_GROSS_RZ'].head()
```

```
Out[34]:
```

| | Attribute | Description \ |
|-----|---------------------|---|
| 170 | D19_BANKEN_GROSS_RZ | transactional activity based on the product gr... |

```

171 D19_BANKEN_GROSS_RZ - on grid level -
172 D19_BANKEN_GROSS_RZ - on grid level -
173 D19_BANKEN_GROSS_RZ - on grid level -
174 D19_BANKEN_GROSS_RZ - on grid level -

```

| | Value | Meaning | Information level \ |
|-----|-------|---------------------------------|---------------------|
| 170 | 0 | no transaction known | NaN |
| 171 | 1 | Multibuyer 0-12 months | NaN |
| 172 | 2 | Doublebuyer 0-12 months | NaN |
| 173 | 3 | Singlebuyer 0-12 months | NaN |
| 174 | 4 | Multi-/Doublebuyer 13-24 months | NaN |

| | meaning |
|-----|---------------------------------|
| 170 | no transaction known |
| 171 | multibuyer 0-12 months |
| 172 | doublebuyer 0-12 months |
| 173 | singlebuyer 0-12 months |
| 174 | multi-/doublebuyer 13-24 months |

```

In [35]: details.Attribute = details.Attribute.str.replace('_RZ', '')
details[details.Attribute == 'D19_BANKEN_GROSS'].head()

```

```

Out[35]:
Attribute Description \
170 D19_BANKEN_GROSS transactional activity based on the product gr...
171 D19_BANKEN_GROSS - on grid level -
172 D19_BANKEN_GROSS - on grid level -
173 D19_BANKEN_GROSS - on grid level -
174 D19_BANKEN_GROSS - on grid level -

```

| | Value | Meaning | Information level \ |
|-----|-------|---------------------------------|---------------------|
| 170 | 0 | no transaction known | NaN |
| 171 | 1 | Multibuyer 0-12 months | NaN |
| 172 | 2 | Doublebuyer 0-12 months | NaN |
| 173 | 3 | Singlebuyer 0-12 months | NaN |
| 174 | 4 | Multi-/Doublebuyer 13-24 months | NaN |

| | meaning |
|-----|---------------------------------|
| 170 | no transaction known |
| 171 | multibuyer 0-12 months |
| 172 | doublebuyer 0-12 months |
| 173 | singlebuyer 0-12 months |
| 174 | multi-/doublebuyer 13-24 months |

```

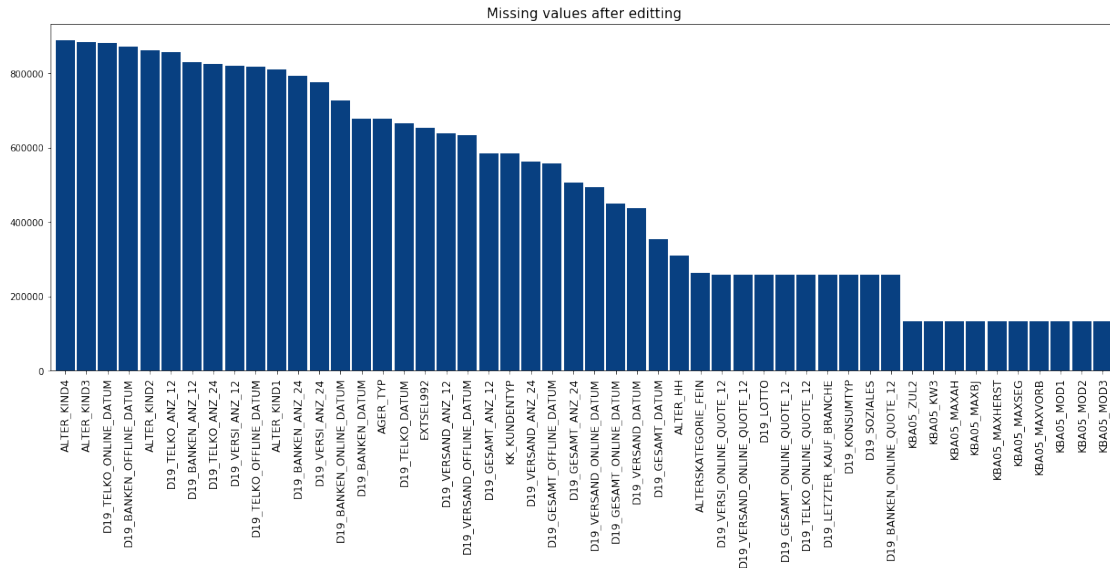
In [36]: azdias = replace_unknown_with_nan(azdias, unknown)

```

```

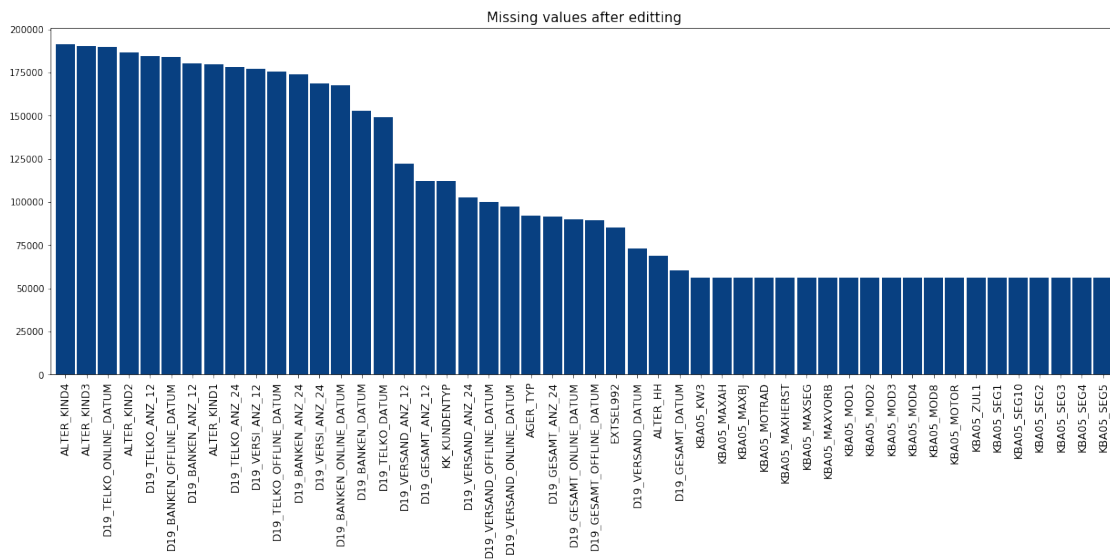
azdias.isnull().sum().sort_values(ascending=False).iloc[:50].plot.bar(figsize=(21,7), w
plt.title('Missing values after editing', fontsize=15)
plt.show()

```

```
In [37]: customers = replace_unknown_with_nan(customers, unknown)
```

```
customers.isnull().sum().sort_values(ascending=False).iloc[:50].plot.bar(figsize=(21,7))
plt.title('Missing values after editing', fontsize=15)
plt.show()
```



Convert categorical columns to numeric where possible

```
In [38]: def object_toNumeric(df):
```

```
    cat_cols = df.select_dtypes(include='object').columns
    for column in cat_cols:
        print(column)
        unique_vals = df[column].unique()
        print('Unique values in', column, unique_vals, '\n')
        if ('X' in unique_vals) or ('XX' in unique_vals):
            df[column] = df[column].replace({'X': np.nan, 'XX': np.nan})
            df[column] = pd.to_numeric(df[column], errors='coerse')

    return df
```

```
In [39]: azdias = object_toNumeric(azdias)
        azdias.select_dtypes(include='object').head()
```

CAMEO_DEU_2015

```
Unique values in CAMEO_DEU_2015 [nan '8A' '4C' '2A' '6B' '8C' '4A' '2D' '1A' '1E' '9D' '5C' '8B'
'9E' '9B' '1B' '3D' '4E' '4B' '3C' '5A' '7B' '9A' '6D' '6E' '2C' '7C' '9C'
'7D' '5E' '1D' '8D' '6C' '6A' '5B' '4D' '3A' '2B' '7E' '3B' '6F' '5F' '1C'
'XX']
```

CAMEO_DEUG_2015

```
Unique values in CAMEO_DEUG_2015 [nan 8.0 4.0 2.0 6.0 1.0 9.0 5.0 7.0 3.0 '4' '3' '7' '2' '8' '9'
'1' 'X']
```

CAMEO_INTL_2015

```
Unique values in CAMEO_INTL_2015 [nan 51.0 24.0 12.0 43.0 54.0 22.0 14.0 13.0 15.0 33.0 41.0 34.
23.0 31.0 52.0 35.0 45.0 44.0 32.0 '22' '24' '41' '12' '54' '51' '44' '35'
'23' '25' '14' '34' '52' '55' '31' '32' '15' '13' '43' '33' '45' 'XX']
```

D19_LETZTER_KAUF_BRANCHE

```
Unique values in D19_LETZTER_KAUF_BRANCHE [nan 'D19_UNBEKANNT' 'D19_SCHUHE' 'D19_ENERGIE' 'D19_K
'D19_VOLLSORTIMENT' 'D19_SONSTIGE' 'D19_BANKEN_GROSS'
'D19_DROGERIEARTIKEL' 'D19_HANDWERK' 'D19_BUCH_CD' 'D19_VERSICHERUNGEN'
'D19_VERSAND_REST' 'D19_TELKO_REST' 'D19_BANKEN_DIREKT' 'D19_BANKEN_REST'
'D19_FREIZEIT' 'D19_LEBENSMITTEL' 'D19_HAUS_DEKO' 'D19_BEKLEIDUNG_REST'
'D19_SAMMELARTIKEL' 'D19_TELKO_MOBILE' 'D19_REISEN' 'D19_BEKLEIDUNG_GEH'
'D19_TECHNIK' 'D19_NAHRUNGSEGAENZUNG' 'D19_DIGIT_SERV' 'D19_LOTTO'
'D19_RATGEBER' 'D19_TIERARTIKEL' 'D19_KINDERARTIKEL' 'D19_BIO_OEKO'
'D19_WEIN_FEINKOST' 'D19_GARTEN' 'D19_BILDUNG' 'D19_BANKEN_LOKAL']
```

EINGEFUEGT_AM

```
Unique values in EINGEFUEGT_AM [nan '1992-02-10 00:00:00' '1992-02-12 00:00:00' ..., '2010-12-02
'2005-03-19 00:00:00' '2011-11-18 00:00:00']
```

OST_WEST_KZ

```
Unique values in OST_WEST_KZ [nan 'W' 'O']
```

```
Out[39]:
```

| | D19_LETZTER_KAUF_BRANCHE | EINGEFUEGT_AM | OST_WEST_KZ |
|---|--------------------------|---------------------|-------------|
| 0 | NaN | NaN | NaN |
| 1 | NaN | 1992-02-10 00:00:00 | W |
| 2 | D19_UNBEKANNT | 1992-02-12 00:00:00 | W |
| 3 | D19_UNBEKANNT | 1997-04-21 00:00:00 | W |
| 4 | D19_SCHUHE | 1992-02-12 00:00:00 | W |

```
In [40]: customers = object_toNumeric(customers)
customers.select_dtypes(include='object').head()
```

CAMEO_DEU_2015

```
Unique values in CAMEO_DEU_2015 ['1A' nan '5D' '4C' '7B' '3B' '1D' '9E' '2D' '4A' '6B' '9D' '8B'
'4E' '6C' '8C' '8A' '5B' '9B' '3D' '2A' '3C' '5F' '7A' '1E' '2C' '7C' '5A'
'2B' '6D' '7E' '5E' '6E' '3A' '9A' '4B' '1C' '1B' '6A' '8D' '7D' '6F' '4D'
'XX']
```

CAMEO_DEUG_2015

```
Unique values in CAMEO_DEUG_2015 [1.0 nan 5.0 4.0 7.0 3.0 9.0 2.0 6.0 8.0 '6' '3' '8' '9' '2' '4'
'5' 'X']
```

CAMEO_INTL_2015

```
Unique values in CAMEO_INTL_2015 [13.0 nan 34.0 24.0 41.0 23.0 15.0 55.0 14.0 22.0 43.0 51.0 33.
54.0 32.0 12.0 35.0 31.0 45.0 52.0 '45' '25' '55' '51' '14' '54' '43' '22'
'15' '24' '35' '23' '12' '44' '41' '52' '31' '13' '34' '32' '33' 'XX']
```

D19_LETZTER_KAUF_BRANCHE

```
Unique values in D19_LETZTER_KAUF_BRANCHE ['D19_UNBEKANNT' 'D19_BANKEN_GROSS' 'D19_NAHRUNGSERGAE
'D19_BUCH_CD' 'D19_DROGERIEARTIKEL' 'D19_SONSTIGE' 'D19_TECHNIK'
'D19_VERSICHERUNGEN' 'D19_TELKO_MOBILE' 'D19_VOLLSORTIMENT' nan
'D19_HAUS_DEKO' 'D19_ENERGIE' 'D19_REISEN' 'D19_BANKEN_LOKAL'
'D19_VERSAND_REST' 'D19_BEKLEIDUNG_REST' 'D19_FREIZEIT'
'D19_BEKLEIDUNG_GEH' 'D19_TELKO_REST' 'D19_SAMMELARTIKEL'
'D19_BANKEN_DIREKT' 'D19_KINDERARTIKEL' 'D19_BANKEN_REST'
'D19_LEBENSMITTEL' 'D19_GARTEN' 'D19_HANDWERK' 'D19_RATGEBER'
'D19_DIGIT_SERV' 'D19_BIO_OEKO' 'D19_BILDUNG' 'D19_WEIN_FEINKOST'
'D19_TIERARTIKEL' 'D19_LOTTO' 'D19_KOSMETIK']
```

EINGEFUEGT_AM

```
Unique values in EINGEFUEGT_AM ['1992-02-12 00:00:00' nan '1992-02-10 00:00:00' ..., '2008-04-25
'2005-03-30 00:00:00' '2008-07-14 00:00:00']
```

OST_WEST_KZ

```
Unique values in OST_WEST_KZ ['W' nan '0']
```

```
Out[40]:
```

| | D19_LETZTER_KAUF_BRANCHE | EINGEFUEGT_AM | OST_WEST_KZ |
|---|--------------------------|---------------------|-------------|
| 0 | D19_UNBEKANNT | 1992-02-12 00:00:00 | W |
| 1 | D19_BANKEN_GROSS | NaN | NaN |
| 2 | D19_UNBEKANNT | 1992-02-10 00:00:00 | W |
| 3 | D19_NAHRUNGSErgAENZUNG | 1992-02-10 00:00:00 | W |
| 4 | D19_SCHUHE | 1992-02-12 00:00:00 | W |

Convert categorical columns to datetime where possible

```
In [41]: def object_toDatetime(df, column):
```

```
    df[column] = pd.to_datetime(df[column])
    df['YEAR'] = df[column].dt.year
    df['MONTH'] = df[column].dt.month
    df.drop(column, axis=1, inplace=True)
```

```
    return df
```

```
In [42]: azdias = object_toDatetime(azdias, 'EINGEFUEGT_AM')
        azdias[['YEAR', 'MONTH']].head()
```

```
Out[42]:
```

| | YEAR | MONTH |
|---|--------|-------|
| 0 | NaN | NaN |
| 1 | 1992.0 | 2.0 |
| 2 | 1992.0 | 2.0 |
| 3 | 1997.0 | 4.0 |
| 4 | 1992.0 | 2.0 |

```
In [43]: customers = object_toDatetime(customers, 'EINGEFUEGT_AM')
        customers[['YEAR', 'MONTH']].head()
```

```
Out[43]:
```

| | YEAR | MONTH |
|---|--------|-------|
| 0 | 1992.0 | 2.0 |
| 1 | NaN | NaN |
| 2 | 1992.0 | 2.0 |
| 3 | 1992.0 | 2.0 |
| 4 | 1992.0 | 2.0 |

Missing values

```
In [44]: nan_azdias = azdias.isnull().sum()
        nan_azdias = nan_azdias.sort_values(ascending=False)
        nan_azdias_cols = nan_azdias[nan_azdias > 0]
        print('Number of columns with NaNs is', len(nan_azdias_cols))
        proportion_azdias = nan_azdias_cols / azdias.shape[0]
        print(proportion_azdias[proportion_azdias > .9])
```

```
Number of columns with NaNs is 302
CAMEO_DEU_2015          1.000000
```

| | |
|--------------------------|----------|
| ALTER_KIND4 | 0.998648 |
| ALTER_KIND3 | 0.993077 |
| D19_TELKO_ONLINE_DATUM | 0.990796 |
| D19_BANKEN_OFFLINE_DATUM | 0.977911 |
| ALTER_KIND2 | 0.966900 |
| D19_TELKO_ANZ_12 | 0.962713 |
| D19_BANKEN_ANZ_12 | 0.933252 |
| D19_TELKO_ANZ_24 | 0.927052 |
| D19_VERSI_ANZ_12 | 0.921532 |
| D19_TELKO_OFFLINE_DATUM | 0.919092 |
| ALTER_KIND1 | 0.909048 |

dtype: float64

```
In [45]: nan_customers = customers.isnull().sum()
        nan_customers = nan_customers.sort_values(ascending=False)
        nan_customers_cols = nan_customers[nan_customers > 0]
        print('Number of columns with NaNs is', len(nan_customers_cols))
        proportion_customers = nan_customers_cols / customers.shape[0]
        print(proportion_customers[proportion_customers > .9])
```

Number of columns with NaNs is 302

| | |
|--------------------------|----------|
| CAMEO_DEU_2015 | 1.000000 |
| ALTER_KIND4 | 0.998769 |
| ALTER_KIND3 | 0.993347 |
| D19_TELKO_ONLINE_DATUM | 0.988855 |
| ALTER_KIND2 | 0.973389 |
| D19_TELKO_ANZ_12 | 0.962510 |
| D19_BANKEN_OFFLINE_DATUM | 0.961127 |
| D19_BANKEN_ANZ_12 | 0.939985 |
| ALTER_KIND1 | 0.938607 |
| D19_TELKO_ANZ_24 | 0.930911 |
| D19_VERSI_ANZ_12 | 0.924780 |
| D19_TELKO_OFFLINE_DATUM | 0.916635 |
| D19_BANKEN_ANZ_24 | 0.906335 |

dtype: float64

```
In [46]: # Delete columns with more than 90% missing values
```

```
        cols_to_drop = nan_azdias_cols[nan_azdias_cols / azdias.shape[0] > .9].index.tolist()
        azdias.drop(cols_to_drop, axis=1, inplace=True)
```

```
In [47]: cols_to_drop = nan_customers_cols[nan_customers_cols / customers.shape[0] > .9].index.t
        customers.drop(cols_to_drop, axis=1, inplace=True)
```

```
In [48]: # Column which has a little bit less missing values in azdias than in customers DF
```

```
        difference = (set(proportion_customers[proportion_customers > .9].index)
```

```

        .difference(set(proportion_azdias[proportion_azdias > .9].index)))
print(difference)
print(proportion_azdias[difference])

azdias.drop(difference, axis=1, inplace=True)

{'D19_BANKEN_ANZ_24'}
D19_BANKEN_ANZ_24    0.891025
dtype: float64

```

In [49]: # Visualize distribution of missing values in top-30 columns in azdias

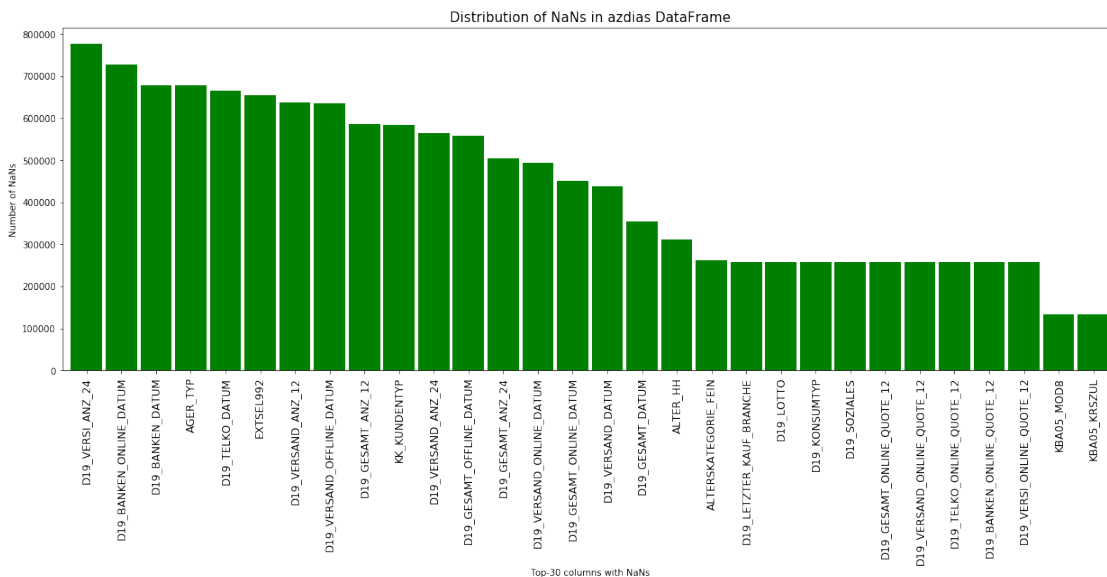
```

nan_azdias = azdias.isnull().sum()
nan_azdias = nan_azdias.sort_values(ascending=False)
nan_azdias_cols = nan_azdias[nan_azdias > 0]

nan_azdias_cols.iloc[:30].plot(kind='bar', figsize=(21,7), width=.9, color='green')
plt.xticks(rotation=90)
plt.title('Distribution of NaNs in azdias DataFrame', fontsize=15)
plt.xlabel('Top-30 columns with NaNs')
plt.ylabel('Number of NaNs')

plt.show()

```



In [50]: # Visualize distribution of missing values in top-30 columns in customers

```

nan_customers = customers.isnull().sum()

```

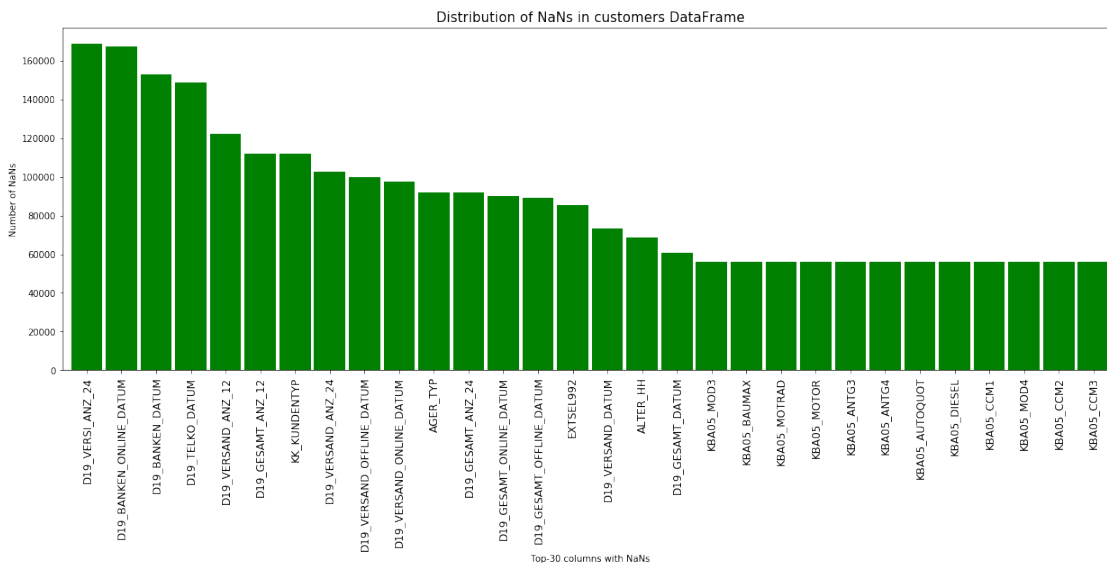
```

nan_customers = nan_customers.sort_values(ascending=False)
nan_customers_cols = nan_customers[nan_customers > 0]

nan_customers_cols.iloc[:30].plot(kind='bar', figsize=(21,7), width=.9, color='green')
plt.xticks(rotation=90)
plt.title('Distribution of NaNs in customers DataFrame', fontsize=15)
plt.xlabel('Top-30 columns with NaNs')
plt.ylabel('Number of NaNs')

plt.show()

```



```

In [51]: print('Shape after deleting columns with more than 90% missing values')
          print('Azdias', azdias.shape)
          print('Customers', customers.shape)

```

```

Shape after deleting columns with more than 90% missing values
Azdias (891221, 354)
Customers (191652, 354)

```

Drop columns and create new ones

```

In [52]: # Create column which represent sum of NaN in each row

```

```

def create_num_missing_column(df):
    df['Num_missing'] = df.isnull().sum(axis=1)

    return df

```

```

In [53]: azdias = create_num_missing_column(azdias)

In [54]: customers = create_num_missing_column(customers)

In [55]: # For columns with NaN more than 10% and less 90% replace missing values with 0s, and n

def tranform_cols_with_missing(df):
    nan_df = df.isnull().sum() / df.shape[0]
    nan_cols = nan_df[(nan_df > .1) & (nan_df < .9)]
    for column in nan_cols.index:
        print(column)
        df[column] = np.where(df[column].isnull(), 0, 1)

    return df

In [56]: azdias = tranform_cols_with_missing(azdias)

AGER_TYP
ALTER_HH
ALTERSKATEGORIE_FEIN
ANZ_HAUSHALTE_AKTIV
ANZ_HH_TITEL
ANZ_STATISTISCHE_HAUSHALTE
ARBEIT
BALLRAUM
CAMEO_DEUG_2015
CAMEO_INTL_2015
D19_BANKEN_DATUM
D19_BANKEN_ONLINE_DATUM
D19_BANKEN_ONLINE_QUOTE_12
D19_GESAMT_ANZ_12
D19_GESAMT_ANZ_24
D19_GESAMT_DATUM
D19_GESAMT_OFFLINE_DATUM
D19_GESAMT_ONLINE_DATUM
D19_GESAMT_ONLINE_QUOTE_12
D19_KONSUMTYP
D19_LETZTER_KAUF_BRANCHE
D19_LOTTO
D19_SOZIALES
D19_TELKO_DATUM
D19_TELKO_ONLINE_QUOTE_12
D19_VERSAND_ANZ_12
D19_VERSAND_ANZ_24
D19_VERSAND_DATUM
D19_VERSAND_OFFLINE_DATUM
D19_VERSAND_ONLINE_DATUM
D19_VERSAND_ONLINE_QUOTE_12
D19_VERSI_ANZ_24

```


D19_VERSI_ONLINE_QUOTE_12
DSL_FLAG
EWDICHTE
EXTSEL992
FIRMENDICHTE
GEBAEUDETYP
GEBAEUDETYP_RASTER
GEMEINDETYP
HEALTH_TYP
HH_DELTA_FLAG
INNENSTADT
KBA05_ALTER1
KBA05_ALTER2
KBA05_ALTER3
KBA05_ALTER4
KBA05_ANHANG
KBA05_ANTG1
KBA05_ANTG2
KBA05_ANTG3
KBA05_ANTG4
KBA05_AUTOQUOT
KBA05_BAUMAX
KBA05_CCM1
KBA05_CCM2
KBA05_CCM3
KBA05_CCM4
KBA05_DIESEL
KBA05_FRAU
KBA05_GBZ
KBA05_HERST1
KBA05_HERST2
KBA05_HERST3
KBA05_HERST4
KBA05_HERST5
KBA05_HERSTTEMP
KBA05_KRSAQUOT
KBA05_KRSHERST1
KBA05_KRSHERST2
KBA05_KRSHERST3
KBA05_KRSKLEIN
KBA05_KRSOBER
KBA05_KRSVAN
KBA05_KRSZUL
KBA05_KW1
KBA05_KW2
KBA05_KW3
KBA05_MAXAH
KBA05_MAXBJ

KBA05_MAXHERST
KBA05_MAXSEG
KBA05_MAXVORB
KBA05_MOD1
KBA05_MOD2
KBA05_MOD3
KBA05_MOD4
KBA05_MOD8
KBA05_MODTEMP
KBA05_MOTOR
KBA05_MOTRAD
KBA05_SEG1
KBA05_SEG10
KBA05_SEG2
KBA05_SEG3
KBA05_SEG4
KBA05_SEG5
KBA05_SEG6
KBA05_SEG7
KBA05_SEG8
KBA05_SEG9
KBA05_VORB0
KBA05_VORB1
KBA05_VORB2
KBA05_ZUL1
KBA05_ZUL2
KBA05_ZUL3
KBA05_ZUL4
KBA13_ALTERHALTER_30
KBA13_ALTERHALTER_45
KBA13_ALTERHALTER_60
KBA13_ALTERHALTER_61
KBA13_ANTG1
KBA13_ANTG2
KBA13_ANTG3
KBA13_ANTG4
KBA13_ANZAHL_PKW
KBA13_AUDI
KBA13_AUTOQUOTE
KBA13_BAUMAX
KBA13_BJ_1999
KBA13_BJ_2000
KBA13_BJ_2004
KBA13_BJ_2006
KBA13_BJ_2008
KBA13_BJ_2009
KBA13_BMW
KBA13_CCM_0_1400

KBA13_CCM_1000
KBA13_CCM_1200
KBA13_CCM_1400
KBA13_CCM_1401_2500
KBA13_CCM_1500
KBA13_CCM_1600
KBA13_CCM_1800
KBA13_CCM_2000
KBA13_CCM_2500
KBA13_CCM_2501
KBA13_CCM_3000
KBA13_CCM_3001
KBA13_FAB_ASIEN
KBA13_FAB_SONSTIGE
KBA13_FIAT
KBA13_FORD
KBA13_GBZ
KBA13_HALTER_20
KBA13_HALTER_25
KBA13_HALTER_30
KBA13_HALTER_35
KBA13_HALTER_40
KBA13_HALTER_45
KBA13_HALTER_50
KBA13_HALTER_55
KBA13_HALTER_60
KBA13_HALTER_65
KBA13_HALTER_66
KBA13_HERST_ASIEN
KBA13_HERST_AUDI_VW
KBA13_HERST_BMW_BENZ
KBA13_HERST_EUROPA
KBA13_HERST_FORD_OPEL
KBA13_HERST_SONST
KBA13_HHZ
KBA13_KMH_0_140
KBA13_KMH_110
KBA13_KMH_140
KBA13_KMH_140_210
KBA13_KMH_180
KBA13_KMH_210
KBA13_KMH_211
KBA13_KMH_250
KBA13_KMH_251
KBA13_KRSAQUOT
KBA13_KRSHERST_AUDI_VW
KBA13_KRSHERST_BMW_BENZ
KBA13_KRSHERST_FORD_OPEL

KBA13_KRSSEG_KLEIN
KBA13_KRSSEG_OBER
KBA13_KRSSEG_VAN
KBA13_KRSZUL_NEU
KBA13_KW_0_60
KBA13_KW_110
KBA13_KW_120
KBA13_KW_121
KBA13_KW_30
KBA13_KW_40
KBA13_KW_50
KBA13_KW_60
KBA13_KW_61_120
KBA13_KW_70
KBA13_KW_80
KBA13_KW_90
KBA13_MAZDA
KBA13_MERCEDES
KBA13_MOTOR
KBA13_NISSAN
KBA13_OPEL
KBA13_PEUGEOT
KBA13_RENAULT
KBA13_SEG_GELAENDEWAGEN
KBA13_SEG_GROSSRAUMVANS
KBA13_SEG_KLEINST
KBA13_SEG_KLEINWAGEN
KBA13_SEG_KOMPAKTKLASSE
KBA13_SEG_MINIVANS
KBA13_SEG_MINIWAGEN
KBA13_SEG_MITTELKLASSE
KBA13_SEG_OBEREMITTELKLASSE
KBA13_SEG_OBERKLASSE
KBA13_SEG_SONSTIGE
KBA13_SEG_SPORTWAGEN
KBA13_SEG_UTILITIES
KBA13_SEG_VAN
KBA13_SEG_WOHNMOBILE
KBA13_SITZE_4
KBA13_SITZE_5
KBA13_SITZE_6
KBA13_TOYOTA
KBA13_VORB_0
KBA13_VORB_1
KBA13_VORB_1_2
KBA13_VORB_2
KBA13_VORB_3
KBA13_VW

KK_KUNDENTYP
KKK
KONSUMZELLE
MIN_GEBAEUDEJAHR
MOBI_RASTER
MOBI_REGIO
NATIONALITAET_KZ
ORTSGR_KLS9
OST_WEST_KZ
PLZ8_ANTG1
PLZ8_ANTG2
PLZ8_ANTG3
PLZ8_ANTG4
PLZ8_BAUMAX
PLZ8_GBZ
PLZ8_HHZ
PRAEGENDE_JUGENDJAHRE
REGIOTYP
RELAT_AB
SHOPPER_TYP
STRUKTURTYP
UMFELD_ALT
UMFELD_JUNG
VERDICHTUNGSRaum
VERS_TYP
VHN
W_KEIT_KIND_HH
WOHNLAGe
YEAR
MONTH

```
In [57]: customers = tranform_cols_with_missing(customers)
```

AGER_TYP
AKT_DAT_KL
ALTER_HH
ALTERSKATEGORIE_FEIN
ANZ_HAUSHALTE_AKTIV
ANZ_HH_TITEL
ANZ_KINDER
ANZ_PERSONEN
ANZ_STATISTISCHE_HAUSHALTE
ANZ_TITEL
ARBEIT
BALLRAUM
CAMEO_DEUG_2015
CAMEO_INTL_2015

D19_BANKEN_DATUM
D19_BANKEN_ONLINE_DATUM
D19_BANKEN_ONLINE_QUOTE_12
D19_GESAMT_ANZ_12
D19_GESAMT_ANZ_24
D19_GESAMT_DATUM
D19_GESAMT_OFFLINE_DATUM
D19_GESAMT_ONLINE_DATUM
D19_GESAMT_ONLINE_QUOTE_12
D19_KONSUMTYP
D19_LETZTER_KAUF_BRANCHE
D19_LOTTO
D19_SOZIALES
D19_TELKO_DATUM
D19_TELKO_ONLINE_QUOTE_12
D19_VERSAND_ANZ_12
D19_VERSAND_ANZ_24
D19_VERSAND_DATUM
D19_VERSAND_OFFLINE_DATUM
D19_VERSAND_ONLINE_DATUM
D19_VERSAND_ONLINE_QUOTE_12
D19_VERSI_ANZ_24
D19_VERSI_ONLINE_QUOTE_12
DSL_FLAG
EINGEZOGENAM_HH_JAHR
EWDICHTE
EXTSEL992
FIRMENDICHTE
GEBAEUDETYP
GEBAEUDETYP_RASTER
GEMEINDETYP
HEALTH_TYP
HH_DELTA_FLAG
INNENSTADT
KBA05_ALTER1
KBA05_ALTER2
KBA05_ALTER3
KBA05_ALTER4
KBA05_ANHANG
KBA05_ANTG1
KBA05_ANTG2
KBA05_ANTG3
KBA05_ANTG4
KBA05_AUTOQUOT
KBA05_BAUMAX
KBA05_CCM1
KBA05_CCM2
KBA05_CCM3

KBA05_CCM4
KBA05_DIESEL
KBA05_FRAU
KBA05_GBZ
KBA05_HERST1
KBA05_HERST2
KBA05_HERST3
KBA05_HERST4
KBA05_HERST5
KBA05_HERSTTEMP
KBA05_KRSAQUOT
KBA05_KRSHERST1
KBA05_KRSHERST2
KBA05_KRSHERST3
KBA05_KRSKLEIN
KBA05_KRSOBER
KBA05_KRSVAN
KBA05_KRSZUL
KBA05_KW1
KBA05_KW2
KBA05_KW3
KBA05_MAXAH
KBA05_MAXBJ
KBA05_MAXHERST
KBA05_MAXSEG
KBA05_MAXVORB
KBA05_MOD1
KBA05_MOD2
KBA05_MOD3
KBA05_MOD4
KBA05_MOD8
KBA05_MODTEMP
KBA05_MOTOR
KBA05_MOTRAD
KBA05_SEG1
KBA05_SEG10
KBA05_SEG2
KBA05_SEG3
KBA05_SEG4
KBA05_SEG5
KBA05_SEG6
KBA05_SEG7
KBA05_SEG8
KBA05_SEG9
KBA05_VORB0
KBA05_VORB1
KBA05_VORB2
KBA05_ZUL1

KBA05_ZUL2
KBA05_ZUL3
KBA05_ZUL4
KBA13_ALTERHALTER_30
KBA13_ALTERHALTER_45
KBA13_ALTERHALTER_60
KBA13_ALTERHALTER_61
KBA13_ANTG1
KBA13_ANTG2
KBA13_ANTG3
KBA13_ANTG4
KBA13_ANZAHL_PKW
KBA13_AUDI
KBA13_AUTOQUOTE
KBA13_BAUMAX
KBA13_BJ_1999
KBA13_BJ_2000
KBA13_BJ_2004
KBA13_BJ_2006
KBA13_BJ_2008
KBA13_BJ_2009
KBA13_BMW
KBA13_CCM_0_1400
KBA13_CCM_1000
KBA13_CCM_1200
KBA13_CCM_1400
KBA13_CCM_1401_2500
KBA13_CCM_1500
KBA13_CCM_1600
KBA13_CCM_1800
KBA13_CCM_2000
KBA13_CCM_2500
KBA13_CCM_2501
KBA13_CCM_3000
KBA13_CCM_3001
KBA13_FAB_ASIEN
KBA13_FAB_SONSTIGE
KBA13_FIAT
KBA13_FORD
KBA13_GBZ
KBA13_HALTER_20
KBA13_HALTER_25
KBA13_HALTER_30
KBA13_HALTER_35
KBA13_HALTER_40
KBA13_HALTER_45
KBA13_HALTER_50
KBA13_HALTER_55

KBA13_HALTER_60
KBA13_HALTER_65
KBA13_HALTER_66
KBA13_HERST_ASIEN
KBA13_HERST_AUDI_VW
KBA13_HERST_BMW_BENZ
KBA13_HERST_EUROPA
KBA13_HERST_FORD_OPEL
KBA13_HERST_SONST
KBA13_HHZ
KBA13_KMH_O_140
KBA13_KMH_110
KBA13_KMH_140
KBA13_KMH_140_210
KBA13_KMH_180
KBA13_KMH_210
KBA13_KMH_211
KBA13_KMH_250
KBA13_KMH_251
KBA13_KRSAQUOT
KBA13_KRSHERST_AUDI_VW
KBA13_KRSHERST_BMW_BENZ
KBA13_KRSHERST_FORD_OPEL
KBA13_KRSSEG_KLEIN
KBA13_KRSSEG_OBER
KBA13_KRSSEG_VAN
KBA13_KRSZUL_NEU
KBA13_KW_O_60
KBA13_KW_110
KBA13_KW_120
KBA13_KW_121
KBA13_KW_30
KBA13_KW_40
KBA13_KW_50
KBA13_KW_60
KBA13_KW_61_120
KBA13_KW_70
KBA13_KW_80
KBA13_KW_90
KBA13_MAZDA
KBA13_MERCEDES
KBA13_MOTOR
KBA13_NISSAN
KBA13_OPEL
KBA13_PEUGEOT
KBA13_RENAULT
KBA13_SEG_GELAEDEWAGEN
KBA13_SEG_GROSSRAUMVANS

KBA13_SEG_KLEINST
KBA13_SEG_KLEINWAGEN
KBA13_SEG_KOMPAKTKLASSE
KBA13_SEG_MINIVANS
KBA13_SEG_MINIWAGEN
KBA13_SEG_MITTELKLASSE
KBA13_SEG_OBEREMITTELKLASSE
KBA13_SEG_OBERKLASSE
KBA13_SEG_SONSTIGE
KBA13_SEG_SPORTWAGEN
KBA13_SEG_UTILITIES
KBA13_SEG_VAN
KBA13_SEG_WOHNMOBILE
KBA13_SITZE_4
KBA13_SITZE_5
KBA13_SITZE_6
KBA13_TOYOTA
KBA13_VORB_0
KBA13_VORB_1
KBA13_VORB_1_2
KBA13_VORB_2
KBA13_VORB_3
KBA13_VW
KK_KUNDENTYP
KKK
KONSUMNAEHE
KONSUMZELLE
MIN_GEBAEUDEJAHR
MOBI_RASTER
MOBI_REGIO
NATIONALITAET_KZ
ORTSGR_KLS9
OST_WEST_KZ
PLZ8_ANTG1
PLZ8_ANTG2
PLZ8_ANTG3
PLZ8_ANTG4
PLZ8_BAUMAX
PLZ8_GBZ
PLZ8_HHZ
PRAEGENDE_JUGENDJAHRE
REGIOTYP
RELAT_AB
RT_UEBERGROESSE
SHOPPER_TYP
SOHO_KZ
STRUKTURTYP
TITEL_KZ

```

UMFELD_ALT
UMFELD_JUNG
UNGLEICHENN_FLAG
VERDICHTUNGSRAUM
VERS_TYP
VHA
VHN
VK_DHT4A
VK_DISTANZ
VK_ZG11
W_KEIT_KIND_HH
WOHNDAUER_2008
WOHNLAGE
YEAR
MONTH

```

```

In [58]: print('# of missing in azdias', azdias.isnull().sum().sum())
         print('# of missing in customers', customers.isnull().sum().sum())

```

```

# of missing in azdias 1198510
# of missing in customers 64015

```

```

In [59]: # Work with azdias DF where # of missing values < 10%

```

```

nan_azdias = azdias.isnull().sum() / azdias.shape[0]
nan_cols = nan_azdias[(nan_azdias > 0) & (nan_azdias < .1)]
print(nan_cols)

```

```

AKT_DAT_KL          0.082470
ANZ_KINDER          0.082470
ANZ_PERSONEN        0.082470
ANZ_TITEL           0.082470
CJT_GESAMTTYP       0.005446
CJT_KATALOGNUTZER   0.005446
CJT_TYP_1           0.005446
CJT_TYP_2           0.005446
CJT_TYP_3           0.005446
CJT_TYP_4           0.005446
CJT_TYP_5           0.005446
CJT_TYP_6           0.005446
EINGEZOGENAM_HH_JAHR 0.082470
GFK_URLAUBERTYP     0.005446
HH_EINKOMMEN_SCORE   0.020587
KONSUMNAEHE         0.082997
LP_FAMILIE_FEIN     0.005446
LP_FAMILIE_GROB     0.005446
LP_LEBENSPHASE_FEIN 0.005446

```

| | |
|---------------------|----------|
| LP_LEBENSPHASE_GROB | 0.005446 |
| LP_STATUS_FEIN | 0.005446 |
| LP_STATUS_GROB | 0.005446 |
| ONLINE_AFFINITAET | 0.005446 |
| RETOURTYP_BK_S | 0.005446 |
| RT_KEIN_ANREIZ | 0.005446 |
| RT_SCHNAEPPCHEN | 0.005446 |
| RT_UEBERGROESSE | 0.057478 |
| SOHO_KZ | 0.082470 |
| TITEL_KZ | 0.082470 |
| UNGLEICHENN_FLAG | 0.082470 |
| VHA | 0.082470 |
| VK_DHT4A | 0.085183 |
| VK_DISTANZ | 0.085183 |
| VK_ZG11 | 0.085183 |
| WOHNDAUER_2008 | 0.082470 |

dtype: float64

Replace missing values with modes of corresponding columns.

```
In [60]: for column in azdias[nan_cols.index]:
          col_mode = azdias[column].mode().values[0]
          azdias[column] = azdias[column].replace(np.nan, col_mode)

          azdias.isnull().sum().sum()
```

Out[60]: 0

```
In [61]: # Work with customers DF where # of missing values < 10%

          nan_customers = customers.isnull().sum() / customers.shape[0]
          nan_cols = nan_customers[(nan_customers > 0) & (nan_customers < .1)]
          print(nan_cols)
```

| | |
|---------------------|----------|
| CJT_GESAMTTYP | 0.016765 |
| CJT_KATALOGNUTZER | 0.016765 |
| CJT_TYP_1 | 0.016765 |
| CJT_TYP_2 | 0.016765 |
| CJT_TYP_3 | 0.016765 |
| CJT_TYP_4 | 0.016765 |
| CJT_TYP_5 | 0.016765 |
| CJT_TYP_6 | 0.016765 |
| GFK_URLAUBERTYP | 0.016765 |
| HH_EINKOMMEN_SCORE | 0.015486 |
| LP_FAMILIE_FEIN | 0.016765 |
| LP_FAMILIE_GROB | 0.016765 |
| LP_LEBENSPHASE_FEIN | 0.016765 |
| LP_LEBENSPHASE_GROB | 0.016765 |

```

LP_STATUS_FEIN          0.016765
LP_STATUS_GROB          0.016765
ONLINE_AFFINITAET       0.016765
RETOURTYP_BK_S          0.016765
RT_KEIN_ANREIZ          0.016765
RT_SCHNAEPPCHEN         0.016765
dtype: float64

```

```
In [62]: # Check dtypes of columns with missing values
```

```

    for column in customers[nan_cols.index]:
        print(column)
        print(customers[column].value_counts())

```

```
CJT_GESAMTTYP
```

```

6.0    51907
2.0    42841
4.0    26912
3.0    24343
1.0    24229
5.0    18207

```

```
Name: CJT_GESAMTTYP, dtype: int64
```

```
CJT_KATALOGNUTZER
```

```

5.0    107544
4.0    27271
3.0    22087
1.0    20510
2.0    11027

```

```
Name: CJT_KATALOGNUTZER, dtype: int64
```

```
CJT_TYP_1
```

```

1.0    55916
2.0    53362
5.0    46069
3.0    23080
4.0    10012

```

```
Name: CJT_TYP_1, dtype: int64
```

```
CJT_TYP_2
```

```

1.0    64716
2.0    51794
5.0    44294
3.0    20079
4.0     7556

```

```
Name: CJT_TYP_2, dtype: int64
```

```
CJT_TYP_3
```

```

5.0    133124
4.0    30494
3.0    15564

```

```

2.0      7515
1.0      1742
Name: CJT_TYP_3, dtype: int64
CJT_TYP_4
5.0     127878
4.0      32339
3.0      11982
2.0      11086
1.0       5154
Name: CJT_TYP_4, dtype: int64
CJT_TYP_5
5.0     136309
4.0      24855
3.0      18855
2.0       5786
1.0       2634
Name: CJT_TYP_5, dtype: int64
CJT_TYP_6
5.0     137643
4.0      27282
3.0      12791
2.0       8855
1.0       1868
Name: CJT_TYP_6, dtype: int64
GFK_URLAUBERTYP
5.0      58113
10.0     27291
8.0      17631
4.0      16361
3.0      14684
7.0      13826
1.0       8782
11.0      7486
12.0      7303
6.0       6373
9.0       5516
2.0       5073
Name: GFK_URLAUBERTYP, dtype: int64
HH_EINKOMMEN_SCORE
2.0      70160
1.0      29936
4.0      27674
5.0      23923
3.0      22438
6.0      14553
Name: HH_EINKOMMEN_SCORE, dtype: int64
LP_FAMILIE_FEIN
0.0      47369

```

| | |
|------|-------|
| 1.0 | 40769 |
| 10.0 | 36568 |
| 2.0 | 28937 |
| 11.0 | 22289 |
| 8.0 | 4686 |
| 7.0 | 2960 |
| 9.0 | 2428 |
| 5.0 | 903 |
| 6.0 | 831 |
| 4.0 | 544 |
| 3.0 | 155 |

Name: LP_FAMILIE_FEIN, dtype: int64

LP_FAMILIE_GROB

| | |
|-----|-------|
| 5.0 | 61285 |
| 0.0 | 47369 |
| 1.0 | 40769 |
| 2.0 | 28937 |
| 4.0 | 8477 |
| 3.0 | 1602 |

Name: LP_FAMILIE_GROB, dtype: int64

LP_LEBENSPHASE_FEIN

| | |
|------|-------|
| 0.0 | 47840 |
| 40.0 | 18299 |
| 20.0 | 10851 |
| 13.0 | 9972 |
| 36.0 | 8821 |
| 38.0 | 8648 |
| 39.0 | 6785 |
| 6.0 | 6646 |
| 8.0 | 6297 |
| 19.0 | 6143 |
| 12.0 | 5951 |
| 32.0 | 5415 |
| 37.0 | 4592 |
| 31.0 | 3983 |
| 16.0 | 3962 |
| 28.0 | 3709 |
| 17.0 | 3475 |
| 15.0 | 3252 |
| 9.0 | 2939 |
| 27.0 | 2556 |
| 5.0 | 2261 |
| 11.0 | 1985 |
| 35.0 | 1784 |
| 7.0 | 1700 |
| 23.0 | 1113 |
| 26.0 | 1015 |
| 34.0 | 919 |

| | |
|------|-----|
| 10.0 | 852 |
| 30.0 | 826 |
| 25.0 | 669 |
| 2.0 | 663 |
| 33.0 | 623 |
| 18.0 | 600 |
| 14.0 | 565 |
| 1.0 | 553 |
| 24.0 | 528 |
| 29.0 | 466 |
| 4.0 | 434 |
| 3.0 | 258 |
| 22.0 | 253 |
| 21.0 | 236 |

Name: LP_LEBENSPHASE_FEIN, dtype: int64

LP_LEBENSPHASE_GROB

| | |
|------|-------|
| 0.0 | 47728 |
| 12.0 | 47145 |
| 3.0 | 21739 |
| 5.0 | 21069 |
| 2.0 | 16904 |
| 10.0 | 9398 |
| 4.0 | 7851 |
| 8.0 | 7280 |
| 11.0 | 3326 |
| 1.0 | 1908 |
| 6.0 | 1602 |
| 9.0 | 1292 |
| 7.0 | 1197 |

Name: LP_LEBENSPHASE_GROB, dtype: int64

LP_STATUS_FEIN

| | |
|------|-------|
| 10.0 | 54653 |
| 5.0 | 41912 |
| 9.0 | 32916 |
| 1.0 | 19271 |
| 3.0 | 15364 |
| 7.0 | 10574 |
| 6.0 | 6502 |
| 4.0 | 5004 |
| 2.0 | 1404 |
| 8.0 | 839 |

Name: LP_STATUS_FEIN, dtype: int64

LP_STATUS_GROB

| | |
|-----|-------|
| 2.0 | 62280 |
| 5.0 | 54653 |
| 4.0 | 33755 |
| 1.0 | 20675 |
| 3.0 | 17076 |


```

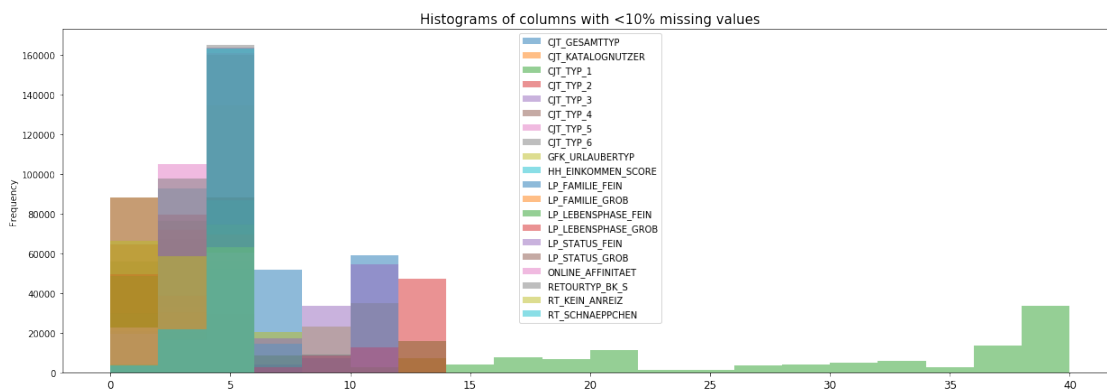
Name: LP_STATUS_GROB, dtype: int64
ONLINE_AFFINITAET
2.0    76440
4.0    39104
3.0    28791
5.0    21311
1.0    18683
0.0     4110
Name: ONLINE_AFFINITAET, dtype: int64
RETOURTYP_BK_S
3.0    83297
5.0    70985
4.0    15653
2.0    14366
1.0     4138
Name: RETOURTYP_BK_S, dtype: int64
RT_KEIN_ANREIZ
1.0    66533
4.0    58452
2.0    36421
3.0    22160
5.0     4873
Name: RT_KEIN_ANREIZ, dtype: int64
RT_SCHNAEPPCHEN
5.0    140281
4.0    23168
3.0    12628
2.0     9026
1.0     3336
Name: RT_SCHNAEPPCHEN, dtype: int64

```

```

In [63]: customers[nan_cols.index].plot.hist(figsize=(21,7), alpha=.5, bins=20)
plt.legend(loc='best')
plt.title('Histograms of columns with <10% missing values', fontsize=15)
plt.show()

```



In this case less than 2% of answers are missing. Since, values in this columns are discrete, and distributions are more or less normal (except "CJT_TYP_1" column), missing values will be replaced with modes of corresponding columns.

```
In [64]: for column in customers[nan_cols.index]:
          col_mode = customers[column].mode().values[0]
          customers[column] = customers[column].replace(np.nan, col_mode)

          customers.isnull().sum().sum()
```

Out[64]: 0

```
In [65]: print('Whether all values in LNR column are unique in azdias DF -', len(azdias.LNR.unique))
          print('Whether all values in LNR column are unique in customers DF - ',
                len(customers.LNR.unique()) == customers.shape[0])

          print('# of common values in LNR column', len(set(customers.LNR).intersection(set(azdias.LNR))))
          print('# of different values in LNR column', len(set(customers.LNR).difference(set(azdias.LNR))))

          # So, we can drop LNR column in both datasets
          azdias.drop('LNR', axis=1, inplace=True)
          customers.drop('LNR', axis=1, inplace=True)
```

```
Whether all values in LNR column are unique in azdias DF - True
Whether all values in LNR column are unique in customers DF - True
# of common values in LNR column 0
# of different values in LNR column 191652
```

Create dummies from categorical variables

```
In [66]: def create_dummies(df):

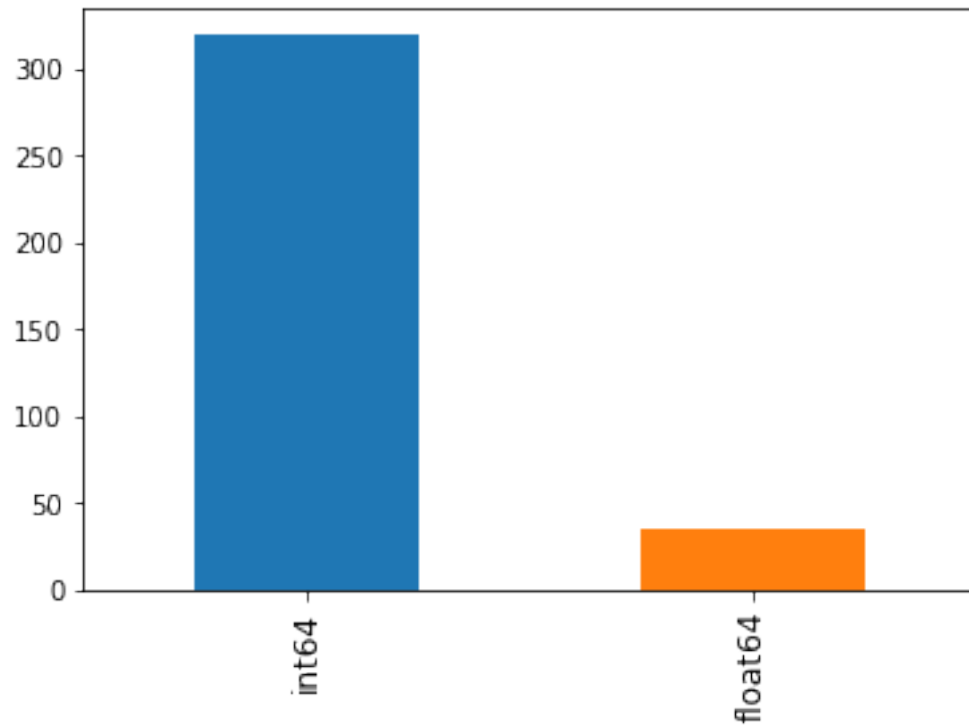
          cat_cols = df.select_dtypes(include='object').columns
          for column in cat_cols:
              num_unique = len(df[column].value_counts())
              if (num_unique == 2) or (num_unique == 3 and np.nan in df[column].value_counts()):
                  values = df[column].value_counts()
                  df[column] = df[column].replace({values.index[0]: 0, values.index[1]: 1})
              else:
                  df = pd.concat([df.drop(column, axis=1), pd.get_dummies(df[column], drop_first=True)], axis=1)

          return df

In [67]: azdias = create_dummies(azdias)
          azdias.dtypes.value_counts()
```

```
Out[67]: int64      319  
float64      35  
dtype: int64
```

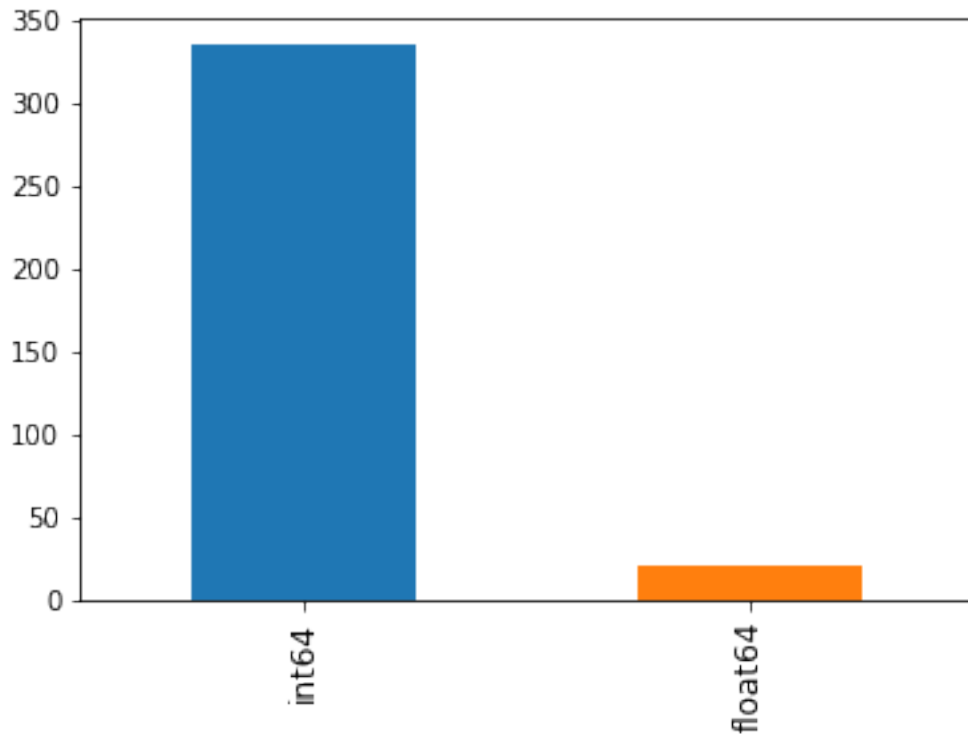
```
In [68]: azdias.dtypes.value_counts().plot.bar()  
plt.show()
```



```
In [69]: customers = create_dummies(customers)  
customers.dtypes.value_counts()
```

```
Out[69]: int64      334  
float64      20  
dtype: int64
```

```
In [70]: customers.dtypes.value_counts().plot.bar()  
plt.show()
```



```
In [71]: azdias.shape, customers.shape, customers_extra.shape
```

```
Out[71]: ((891221, 354), (191652, 354), (191652, 3))
```

Function with pre-processing steps to clean all of the datasets before work with them.

```
In [133]: def clean_demographic_data(df, unknown, cat_column='EINGEFUEGT_AM'):
```

```
#####
# Replace unknown values with NaNs

cols_df = df.columns
for column in unknown.index:
    if column in cols_df:
        col_values = df[column].unique().tolist()
        unknown_vals = unknown.loc[column]['Value']
        for val in col_values:
            if isinstance(unknown_vals, int):
                if val == unknown_vals:
                    df[column] = df[column].replace(val, np.nan)
            else:
                if str(val) in unknown_vals.split():
                    df[column] = df[column].replace(val, np.nan)
```

```
#####
# Convert categorical columns to numeric where possible

cat_cols = df.select_dtypes(include='object').columns
for column in cat_cols:
    unique_vals = df[column].unique()
    if ('X' in unique_vals) or ('XX' in unique_vals):
        df[column] = df[column].replace({'X': np.nan, 'XX': np.nan})
        df[column] = pd.to_numeric(df[column], errors='coerse')

#####
# Convert categorical columns to datetime where possible

df[cat_column] = pd.to_datetime(df[cat_column])
df['YEAR'] = df[cat_column].dt.year
df['MONTH'] = df[cat_column].dt.month
df.drop(cat_column, axis=1, inplace=True)

#####
# Missing values

nan_df = df.isnull().sum()
nan_df = nan_df.sort_values(ascending=False)
nan_df_cols = nan_df[nan_df > 0]
proportion_df = nan_df_cols / df.shape[0]

# Delete columns with more than 90% missing values
cols_to_drop = nan_df_cols[nan_df_cols / df.shape[0] > .9].index.tolist()
df.drop(cols_to_drop, axis=1, inplace=True)

#####
# Create column which represent sum of NaN in each row
df['Num_missing'] = df.isnull().sum(axis=1)

#####
# Create column which represent sum of NaN in each row

nan_df = df.isnull().sum() / df.shape[0]
nan_cols = nan_df[(nan_df > .1) & (nan_df < .9)]
for column in nan_cols.index:
    df[column] = np.where(df[column].isnull(), 0, 1)

#####
# LNR column
print('Whether all values in LNR column are unique in azdias DF -', len(df.LNR.unique()))
df.drop('LNR', axis=1, inplace=True)
```

```
#####
# Work with customers DF where # of missing values < 10%

nan_df = df.isnull().sum() / df.shape[0]
nan_cols = nan_df[(nan_df > 0) & (nan_df < .1)]

for column in df[nan_cols.index]:
    col_mode = df[column].mode().values[0]
    df[column] = df[column].replace(np.nan, col_mode)

#####
# Create dummies from categorical variables

cat_cols = df.select_dtypes(include='object').columns
for column in cat_cols:
    num_unique = len(df[column].value_counts())
    if (num_unique == 2) or (num_unique == 3 and np.nan in df[column].value_counts):
        values = df[column].value_counts()
        df[column] = df[column].replace({values.index[0]: 0, values.index[1]: 1})
    else:
        df = pd.concat([df.drop(column, axis=1), pd.get_dummies(df[column], drop_f

print('Final check of missing values', df.isnull().sum().sum())

return df
```

In []:

In []:

1.2 Part 1: Customer Segmentation Report

The main bulk of your analysis will come in this part of the project. Here, you should use unsupervised learning techniques to describe the relationship between the demographics of the company's existing customers and the general population of Germany. By the end of this part, you should be able to describe parts of the general population that are more likely to be part of the mail-order company's main customer base, and which parts of the general population are less so.

Standardize values

```
In [73]: azdias_cols = azdias.columns
         customers_cols = customers.columns

scaler = StandardScaler()
azdias_scaled = scaler.fit_transform(azdias)
print('azdias done!')
customers_scaled = scaler.transform(customers)
```

azdias done!

```
In [74]: pd.DataFrame(azdias_scaled).head()
```

```
Out[74]:
```

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -0.561649 | -0.869634 | -1.368369 | -1.545754 | -2.927080 | -2.861309 | -0.292527 | |
| 1 | -0.561649 | 1.346190 | -1.368369 | 0.646933 | 0.341637 | 0.349490 | -0.292527 | |
| 2 | -0.561649 | 1.346190 | 0.730797 | 0.646933 | 0.341637 | 0.349490 | -0.292527 | |
| 3 | 1.780472 | -0.869634 | 0.730797 | 0.646933 | 0.341637 | 0.349490 | -0.292527 | |
| 4 | -0.561649 | -0.869634 | 0.730797 | 0.646933 | 0.341637 | 0.349490 | -0.292527 | |

| | 7 | 8 | 9 | 10 | 11 | 12 | 13 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -0.593390 | -2.927080 | -0.057885 | -2.857872 | -2.916740 | -2.823179 | -2.823179 | |
| 1 | 0.295412 | 0.341637 | -0.057885 | 0.349911 | 0.342849 | 0.354211 | 0.354211 | |
| 2 | -0.593390 | 0.341637 | -0.057885 | 0.349911 | 0.342849 | 0.354211 | 0.354211 | |
| 3 | -1.482193 | 0.341637 | -0.057885 | 0.349911 | 0.342849 | 0.354211 | 0.354211 | |
| 4 | 2.073018 | 0.341637 | -0.057885 | 0.349911 | 0.342849 | 0.354211 | 0.354211 | |

| | 14 | 15 | 16 | 17 | 18 | 19 | 20 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -1.027618 | 1.107775 | -1.735176 | -1.570560 | 1.172932 | 1.203659 | 1.181001 | |
| 1 | 0.858107 | -1.568545 | 1.184799 | 1.278729 | -0.973028 | -0.251101 | -1.716472 | |
| 2 | -0.399043 | -0.899465 | 0.454805 | 0.566407 | -1.688349 | -0.251101 | -0.992104 | |
| 3 | -1.027618 | -0.230385 | -1.005182 | -0.858238 | 0.457612 | 0.476279 | 1.181001 | |
| 4 | 0.858107 | -0.230385 | -0.275189 | -0.145915 | -0.257708 | 0.476279 | -0.267735 | |

| | 21 | 22 | 23 | 24 | 25 | 26 | 27 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | 1.147432 | -0.560218 | -0.443741 | -0.345901 | -0.132111 | -0.475309 | -1.570434 | |
| 1 | -1.860914 | -0.560218 | -0.443741 | -0.345901 | -0.132111 | -0.475309 | -1.570434 | |
| 2 | -1.108827 | -0.560218 | -0.443741 | -0.345901 | -0.132111 | -0.475309 | 0.636767 | |
| 3 | -0.356741 | -0.560218 | -0.443741 | -0.345901 | -0.132111 | -0.475309 | 0.636767 | |
| 4 | -0.356741 | 1.785020 | 0.053317 | 0.870819 | -0.132111 | 2.103893 | 0.636767 | |

| | 28 | 29 | 30 | 31 | 32 | 33 | 34 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|----------|---|
| 0 | -0.281739 | -0.303388 | -0.505301 | -0.296161 | -0.205967 | -0.622234 | -0.18977 | |
| 1 | -0.281739 | -0.303388 | -0.505301 | -0.296161 | -0.205967 | -0.622234 | -0.18977 | |
| 2 | -0.281739 | -0.303388 | -0.505301 | 3.363851 | -0.205967 | -0.622234 | -0.18977 | |
| 3 | -0.281739 | -0.303388 | -0.505301 | -0.296161 | -0.205967 | 1.732624 | -0.18977 | |
| 4 | 3.689717 | 3.587285 | -0.064189 | 3.363851 | -0.205967 | 1.732624 | -0.18977 | |

| | 35 | 36 | 37 | 38 | 39 | 40 | 41 | \ |
|---|-----------|-----------|-----------|-----------|-----------|----------|-----------|---|
| 0 | -0.384133 | -0.259375 | -0.343098 | -0.211033 | -0.723867 | -0.87392 | -1.231408 | |
| 1 | -0.384133 | -0.259375 | -0.343098 | -0.211033 | -0.723867 | -0.87392 | -1.231408 | |
| 2 | -0.384133 | -0.259375 | -0.343098 | -0.211033 | -0.723867 | -0.87392 | -1.231408 | |
| 3 | -0.384133 | -0.259375 | -0.343098 | -0.211033 | -0.723867 | -0.87392 | -1.231408 | |
| 4 | 0.186397 | 3.479590 | -0.343098 | -0.211033 | 1.381469 | 1.14427 | 0.812079 | |

| | 42 | 43 | 44 | 45 | 46 | 47 | 48 | \ |
|--|----|----|----|----|----|----|----|---|
|--|----|----|----|----|----|----|----|---|

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0 | -0.771735 | -0.987989 | -1.570434 | -0.397223 | -0.474667 | -0.419129 | -1.570434 |
| 1 | -0.771735 | -0.987989 | -1.570434 | -0.397223 | -0.474667 | -0.419129 | -1.570434 |
| 2 | -0.771735 | -0.987989 | 0.636767 | -0.397223 | -0.474667 | -0.419129 | 0.636767 |
| 3 | -0.771735 | -0.987989 | 0.636767 | -0.397223 | -0.474667 | -0.419129 | 0.636767 |
| 4 | 1.295782 | 1.012157 | 0.636767 | -0.397223 | 1.857043 | -0.419129 | 0.636767 |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|----------|----------|---|
| | 49 | 50 | 51 | 52 | 53 | 54 | 55 | \ |
| 0 | 0.976753 | -0.438860 | -0.242783 | -1.570434 | -1.570434 | -0.20641 | -0.31099 | |
| 1 | 0.976753 | -0.438860 | -0.242783 | -1.570434 | -1.570434 | -0.20641 | -0.31099 | |
| 2 | 0.666749 | 2.117354 | -0.242783 | 0.636767 | 0.636767 | -0.20641 | -0.31099 | |
| 3 | 0.666749 | -0.438860 | -0.242783 | 0.636767 | 0.636767 | -0.20641 | -0.31099 | |
| 4 | -1.503282 | -0.438860 | -0.242783 | 0.636767 | 0.636767 | -0.20641 | -0.31099 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 56 | 57 | 58 | 59 | 60 | 61 | 62 | \ |
| 0 | -0.450181 | -0.330651 | -0.356395 | -0.826803 | -1.570434 | -0.634416 | -0.581873 | |
| 1 | -0.450181 | -0.330651 | -0.356395 | -0.826803 | -1.570434 | -0.634416 | -0.581873 | |
| 2 | -0.450181 | -0.330651 | -0.356395 | 1.269944 | 0.636767 | 1.555033 | -0.581873 | |
| 3 | 2.125649 | 3.005852 | -0.356395 | 1.269944 | 0.636767 | 1.555033 | -0.581873 | |
| 4 | -0.450181 | -0.330651 | 0.341299 | 0.571029 | 0.636767 | 1.190125 | 1.718589 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 63 | 64 | 65 | 66 | 67 | 68 | 69 | \ |
| 0 | -0.461770 | -1.570434 | -0.395993 | -0.207701 | -0.630047 | -0.762030 | -1.017488 | |
| 1 | -0.461770 | -1.570434 | -0.395993 | -0.207701 | -0.630047 | -0.762030 | -1.017488 | |
| 2 | -0.461770 | 0.636767 | -0.395993 | -0.207701 | -0.630047 | -0.762030 | -1.017488 | |
| 3 | -0.461770 | 0.636767 | -0.395993 | -0.207701 | -0.630047 | -0.762030 | -1.017488 | |
| 4 | 2.281588 | 0.636767 | 2.142436 | -0.207701 | 1.587183 | 1.312285 | 0.982813 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 70 | 71 | 72 | 73 | 74 | 75 | 76 | \ |
| 0 | -0.636549 | -0.895767 | -1.570434 | -0.433752 | -0.383338 | 0.448642 | 0.153882 | |
| 1 | -0.636549 | -0.895767 | -1.570434 | -0.433752 | -0.383338 | 0.448642 | 0.153882 | |
| 2 | -0.636549 | -0.895767 | 0.636767 | -0.433752 | -0.383338 | 0.448642 | 0.153882 | |
| 3 | -0.636549 | -0.895767 | 0.636767 | -0.433752 | -0.383338 | 0.448642 | 0.153882 | |
| 4 | 1.570970 | 1.116362 | 0.636767 | 0.612001 | 2.608664 | -3.737244 | -5.814308 | |

| | | | | | | | | |
|---|----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 77 | 78 | 79 | 80 | 81 | 82 | 83 | \ |
| 0 | 0.074836 | -1.570434 | -0.558923 | -0.663482 | -0.252831 | -2.927080 | -1.227660 | |
| 1 | 0.074836 | -1.570434 | -0.558923 | -0.663482 | -0.252831 | 0.341637 | 0.147607 | |
| 2 | 0.074836 | 0.636767 | -0.558923 | 2.023461 | -0.252831 | 0.341637 | -0.402500 | |
| 3 | 0.074836 | 0.636767 | -0.558923 | -0.663482 | -0.252831 | 0.341637 | -0.677553 | |
| 4 | 0.074836 | 0.636767 | 0.772072 | -0.663482 | -0.252831 | 0.341637 | 0.147607 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 84 | 85 | 86 | 87 | 88 | 89 | 90 | \ |
| 0 | -2.916740 | -0.602000 | 1.285741 | -0.055511 | -0.056416 | 0.804890 | 1.429871 | |
| 1 | 0.342849 | -0.602000 | 1.285741 | 1.422415 | -1.570358 | 1.487601 | 0.757254 | |
| 2 | 0.342849 | 1.661128 | -0.675554 | 1.422415 | -1.570358 | 0.804890 | 0.084637 | |
| 3 | 0.342849 | 1.661128 | -0.675554 | -0.794475 | 0.700556 | -0.560532 | -1.260597 | |
| 4 | 0.342849 | -0.602000 | -1.329319 | -0.794475 | 0.700556 | 0.122179 | 0.084637 | |

| | 91 | 92 | 93 | 94 | 95 | 96 | 97 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -0.303378 | 0.105346 | -2.926957 | -2.927080 | -2.926957 | -1.127583 | -2.856915 | |
| 1 | -1.059731 | -1.403804 | 0.341652 | 0.341637 | 0.341652 | 0.916278 | 0.350028 | |
| 2 | -1.816084 | -1.403804 | 0.341652 | 0.341637 | 0.341652 | 0.898870 | 0.350028 | |
| 3 | 1.209329 | 1.111445 | 0.341652 | 0.341637 | 0.341652 | 0.876343 | 0.350028 | |
| 4 | 0.452976 | 0.608395 | 0.341652 | 0.341637 | 0.341652 | 0.882487 | 0.350028 | |

| | 98 | 99 | 100 | 101 | 102 | 103 | 104 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | 0.742875 | -0.494701 | -2.648559 | -2.698624 | -1.379075 | -2.916740 | -2.384245 | |
| 1 | 0.742875 | -0.494701 | 0.377564 | 0.370559 | 1.079004 | 0.342849 | 0.419420 | |
| 2 | 0.742875 | 2.021423 | 0.377564 | 0.370559 | -0.150036 | 0.342849 | 0.419420 | |
| 3 | -1.804736 | -0.494701 | 0.377564 | -2.698624 | -1.993595 | 0.342849 | 0.419420 | |
| 4 | -0.672464 | -0.494701 | 0.377564 | 0.370559 | 0.464484 | 0.342849 | 0.419420 | |

| | 105 | 106 | 107 | 108 | 109 | 110 | 111 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | 112 | 113 | 114 | 115 | 116 | 117 | 118 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | 119 | 120 | 121 | 122 | 123 | 124 | 125 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | 126 | 127 | 128 | 129 | 130 | 131 | 132 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -2.384245 | -2.927080 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | 133 | 134 | 135 | 136 | 137 | 138 | 139 | \ |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 140 | 141 | 142 | 143 | 144 | 145 | 146 | \ |
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 147 | 148 | 149 | 150 | 151 | 152 | 153 | \ |
| 0 | -2.384245 | -2.384245 | -2.927080 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.341637 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 154 | 155 | 156 | 157 | 158 | 159 | 160 | \ |
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 161 | 162 | 163 | 164 | 165 | 166 | 167 | \ |
| 0 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | -2.384245 | |
| 1 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 2 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 3 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |
| 4 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | 0.419420 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 168 | 169 | 170 | 171 | 172 | 173 | 174 | \ |
| 0 | -2.384245 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | |
| 1 | 0.419420 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 2 | 0.419420 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 3 | 0.419420 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 4 | 0.419420 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 175 | 176 | 177 | 178 | 179 | 180 | 181 | \ |
| 0 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | |
| 1 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 2 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 3 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 4 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---|
| | 182 | 183 | 184 | 185 | 186 | 187 | 188 | \ |
| 0 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | -2.724636 | |
| 1 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 2 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |
| 3 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | |

[illegible]

| | | | | | | | |
|---|----------|----------|----------|----------|----------|-----------|----------|
| 2 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | -0.724200 | 0.396727 |
| 3 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | -0.724200 | 0.396727 |
| 4 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 0.367022 | 1.380834 | 0.396727 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 287 | 288 | 289 | 290 | 291 | 292 | 293 \ |
| 0 | 2.597558 | -1.167350 | -2.926957 | -0.404397 | -0.102346 | 0.036189 | -0.114372 |
| 1 | -1.193083 | -1.167350 | 0.341652 | 0.360822 | 0.467806 | 0.511308 | 0.405195 |
| 2 | -0.719253 | 1.355387 | 0.341652 | -0.659470 | -0.672497 | -0.914050 | -0.893722 |
| 3 | 0.228408 | 0.724703 | 0.341652 | -0.914542 | -1.242649 | -1.151609 | -1.153505 |
| 4 | -0.245422 | 0.724703 | 0.341652 | 1.636186 | 1.608109 | 1.382359 | 1.444329 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 294 | 295 | 296 | 297 | 298 | 299 | 300 \ |
| 0 | -1.100125 | -0.966557 | -2.927080 | -2.927080 | -2.384245 | -2.688503 | -1.116346 |
| 1 | -0.808354 | -0.966557 | 0.341637 | 0.341637 | 0.419420 | 0.371954 | 0.200965 |
| 2 | -0.516582 | -0.288163 | 0.341637 | 0.341637 | 0.419420 | 0.371954 | -0.457690 |
| 3 | 1.234047 | 1.068626 | 0.341637 | 0.341637 | 0.419420 | 0.371954 | -1.116346 |
| 4 | -0.516582 | -0.288163 | 0.341637 | 0.341637 | 0.419420 | 0.371954 | 1.518276 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 301 | 302 | 303 | 304 | 305 | 306 | 307 \ |
| 0 | -2.857872 | -2.927080 | -2.578562 | -2.578562 | -2.578562 | -2.578562 | -2.578562 |
| 1 | 0.349911 | 0.341637 | 0.387813 | 0.387813 | 0.387813 | 0.387813 | 0.387813 |
| 2 | 0.349911 | 0.341637 | 0.387813 | 0.387813 | 0.387813 | 0.387813 | 0.387813 |
| 3 | 0.349911 | 0.341637 | 0.387813 | 0.387813 | 0.387813 | 0.387813 | 0.387813 |
| 4 | 0.349911 | 0.341637 | 0.387813 | 0.387813 | 0.387813 | 0.387813 | 0.387813 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 308 | 309 | 310 | 311 | 312 | 313 | 314 \ |
| 0 | -2.578562 | -2.578562 | -2.690638 | -2.520625 | -2.857872 | 1.107928 | -1.612770 |
| 1 | 0.387813 | 0.387813 | 0.371659 | 0.396727 | 0.349911 | -1.711653 | 1.262885 |
| 2 | 0.387813 | 0.387813 | 0.371659 | 0.396727 | 0.349911 | -0.301862 | 1.262885 |
| 3 | 0.387813 | 0.387813 | 0.371659 | 0.396727 | 0.349911 | -1.006758 | -0.174942 |
| 4 | 0.387813 | 0.387813 | 0.371659 | 0.396727 | 0.349911 | 1.107928 | -0.174942 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 315 | 316 | 317 | 318 | 319 | 320 | 321 \ |
| 0 | 0.101417 | -1.397277 | 0.742018 | -0.819565 | 0.901553 | 0.839442 | 1.221756 |
| 1 | -0.678627 | 1.217631 | 1.298900 | -1.372799 | -0.142352 | -0.240231 | -0.416882 |
| 2 | 0.101417 | 1.217631 | 1.298900 | 0.840140 | -1.708208 | 1.379279 | 1.221756 |
| 3 | -1.458671 | -0.089823 | -0.371747 | 1.393374 | -1.708208 | 0.299605 | -0.416882 |
| 4 | 0.881461 | 1.217631 | -1.485512 | -0.266330 | -0.142352 | -1.319905 | -0.963095 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 322 | 323 | 324 | 325 | 326 | 327 | 328 \ |
| 0 | -0.538400 | 0.316841 | 0.537487 | 0.420264 | 0.056863 | 1.374629 | -0.999639 |
| 1 | -0.538400 | -1.166231 | -0.539206 | 1.550121 | 1.322441 | -0.119863 | 0.541539 |
| 2 | -0.538400 | -0.177517 | -0.539206 | -0.709593 | 0.056863 | -0.618027 | 0.027813 |
| 3 | -0.013139 | -0.177517 | -1.615900 | -0.144664 | -0.575926 | -1.116191 | 0.541539 |
| 4 | 1.037383 | -0.177517 | -1.077553 | -0.144664 | -1.208715 | -0.119863 | 1.055265 |

| | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | 329 | 330 | 331 | 332 | 333 | 334 | 335 \ |
| 0 | -0.387544 | -1.455284 | -2.648559 | -0.088255 | -2.856915 | -0.039266 | -2.848507 |

| | | | | | | | |
|---|-----------|-----------|----------|-----------|----------|-----------|----------|
| 1 | 1.369271 | -1.455284 | 0.377564 | 11.330820 | 0.350028 | -0.039266 | 0.351061 |
| 2 | -0.387544 | -0.011411 | 0.377564 | -0.088255 | 0.350028 | -0.039266 | 0.351061 |
| 3 | 0.198061 | -0.011411 | 0.377564 | -0.088255 | 0.350028 | -0.039266 | 0.351061 |
| 4 | -0.973149 | 1.432462 | 0.377564 | -0.088255 | 0.350028 | -0.039266 | 0.351061 |

| | | | | | | | | |
|---|-----------|----------|-----------|-----------|-----------|-----------|-----------|---|
| | 336 | 337 | 338 | 339 | 340 | 341 | 342 | \ |
| 0 | -2.848507 | -0.30014 | -2.856915 | -2.648559 | -0.365442 | -2.520625 | 1.239628 | |
| 1 | 0.351061 | 3.33178 | 0.350028 | 0.377564 | -0.365442 | 0.396727 | 0.561895 | |
| 2 | 0.351061 | -0.30014 | 0.350028 | 0.377564 | -0.365442 | 0.396727 | 0.900762 | |
| 3 | 0.351061 | -0.30014 | 0.350028 | 0.377564 | 0.542193 | 0.396727 | 0.223028 | |
| 4 | 0.351061 | -0.30014 | 0.350028 | 0.377564 | -0.365442 | 0.396727 | -1.132440 | |

| | | | | | | | | |
|---|-----------|-----------|-----------|----------|-----------|-----------|-----------|---|
| | 343 | 344 | 345 | 346 | 347 | 348 | 349 | \ |
| 0 | 0.709538 | 1.286729 | -2.698624 | 0.536436 | -2.927080 | -0.267936 | -1.045218 | |
| 1 | 1.023821 | 1.286729 | 0.370559 | 0.536436 | 0.341637 | 1.210585 | 0.956738 | |
| 2 | 0.395256 | -0.101069 | 0.370559 | 0.536436 | 0.341637 | 1.210585 | 0.956738 | |
| 3 | 0.709538 | 1.633678 | -2.698624 | 0.536436 | 0.341637 | -0.267936 | 0.956738 | |
| 4 | -0.861872 | -0.794968 | 0.370559 | 0.536436 | 0.341637 | 0.471325 | -1.045218 | |

| | | | | |
|---|-----------|-----------|-----------|-----------|
| | 350 | 351 | 352 | 353 |
| 0 | -0.727373 | -2.927080 | -2.927080 | 2.867684 |
| 1 | -1.663024 | 0.341637 | 0.341637 | -0.223239 |
| 2 | 0.208278 | 0.341637 | 0.341637 | -0.364319 |
| 3 | 1.143930 | 0.341637 | 0.341637 | -0.351493 |
| 4 | 0.208278 | 0.341637 | 0.341637 | -0.543874 |

Implement PCA

```
In [75]: pca = PCA(.9)
          azdias_reduced = pca.fit_transform(azdias_scaled)
          print('azdias done!')
          customers_reduced = pca.transform(customers_scaled)
          print('Explained variance ratio of azdias DF', pca.explained_variance_ratio_)

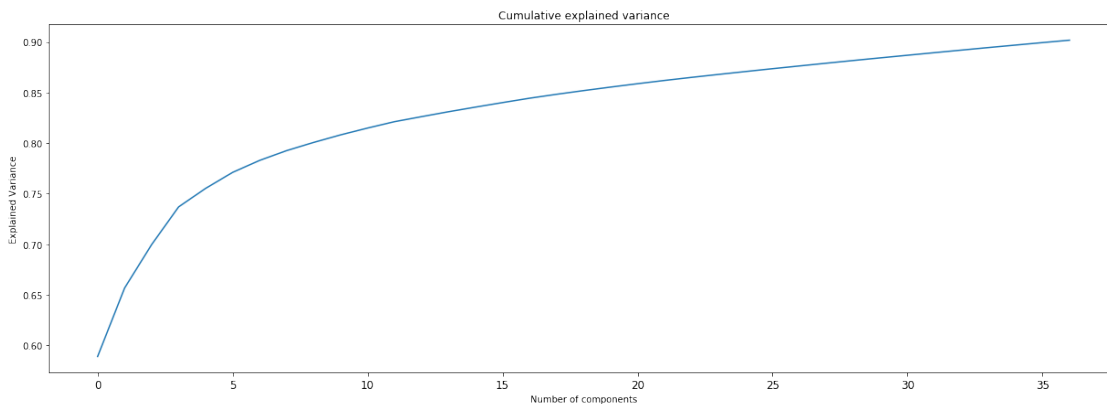
azdias done!
Explained variance ratio of azdias DF [ 0.5889759  0.06774619  0.04287888  0.03731728  0.018290
 0.01184706  0.00967404  0.00810277  0.00746895  0.00675233  0.00632234
 0.00495845  0.00486202  0.00458882  0.00444364  0.0043382  0.00386995
 0.00363423  0.00349306  0.00334006  0.003217  0.00305584  0.00291218
 0.00284511  0.00282003  0.00273326  0.0027142  0.00266649  0.00261064
 0.00257671  0.00255237  0.00252196  0.00252053  0.00246275  0.0024434
 0.00237766]
```



```
In [76]: print('New shape of azdias', azdias_reduced.shape)
          print('New shape of customers', customers_reduced.shape)

New shape of azdias (891221, 37)
New shape of customers (191652, 37)
```

```
In [77]: plt.subplots(figsize=(21,7))
plt.plot(pca.explained_variance_ratio_.cumsum())
plt.title('Cumulative explained variance')
plt.xlabel('Number of components')
plt.ylabel('Explained Variance')
plt.show()
```

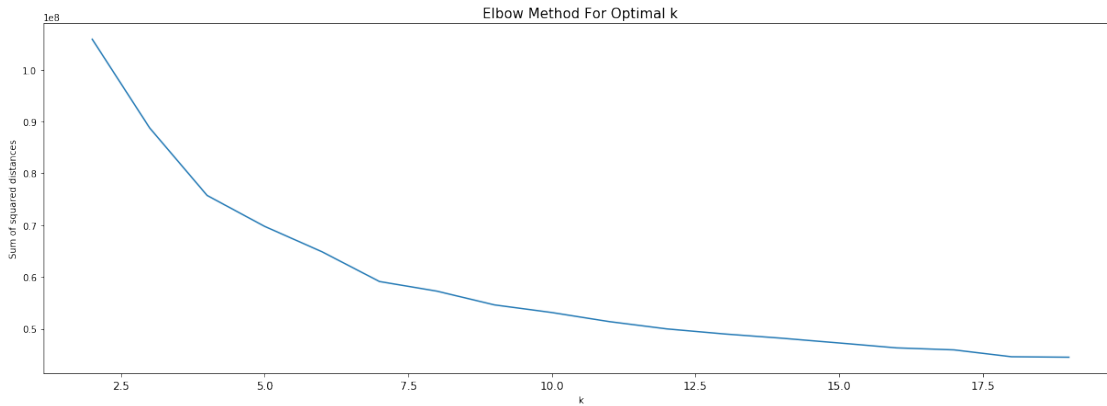


KMeans

```
In [78]: # Choose an optimal # of clusters
# With help of "How to determine the optimal number of clusters for k-means clustering"
# https://blog.cambridgespark.com/how-to-determine-the-optimal-number-of-clusters-for-k
```

```
sum_of_squared_distances = []
for k in range(2,20):
    kmeans = KMeans(n_clusters=k)
    kmeans = kmeans.fit(azdias_reduced)
    sum_of_squared_distances.append(kmeans.inertia_)
```

```
In [79]: plt.subplots(figsize=(21,7))
plt.plot(range(2,20), sum_of_squared_distances)
plt.xlabel('k')
plt.ylabel('Sum of squared distances')
plt.title('Elbow Method For Optimal k', fontsize=15)
plt.show()
```



It's very difficult to determine the optimal # of clusters from this figure. So, in the further analysis of 2-10 clusters will be performed.

```
In [80]: def kmeans(df, n_clusters):
    df = pd.DataFrame(df)

    kmeans = KMeans(n_clusters=n_clusters).fit(df)

    cluster_map = pd.DataFrame()
    cluster_map['cluster'] = kmeans.labels_

    return kmeans, cluster_map
```

k=10

```
In [81]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 10)
```

```
In [82]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
    preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

    preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].
                                                                axis=1, sort=False)
```

preds

```
Out[82]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 34090 | 1.0 |
| 1 | 103054 | NaN |
| 2 | 74893 | 50753.0 |
| 3 | 171095 | NaN |
| 4 | 145983 | NaN |
| 5 | 78740 | 91928.0 |
| 6 | 6566 | 528.0 |
| 7 | 130548 | 48442.0 |


```

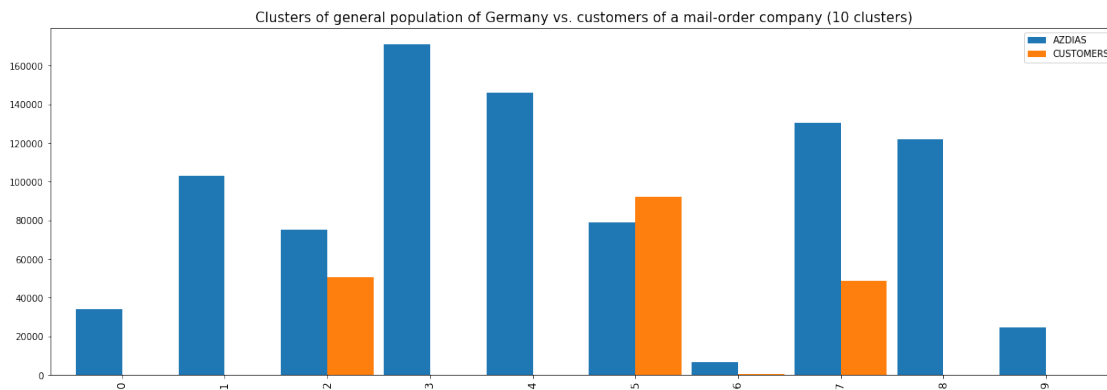
8  121911      NaN
9   24341      NaN

```

```

In [83]: preds.plot(kind='bar', figsize=(21,7), width=.9)
plt.title('Clusters of general population of Germany vs. customers of a mail-order comp
plt.legend()
plt.show()

```



k=9

```

In [84]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 9)

```

```

In [85]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                  axis=1, sort=False)

```

preds

```

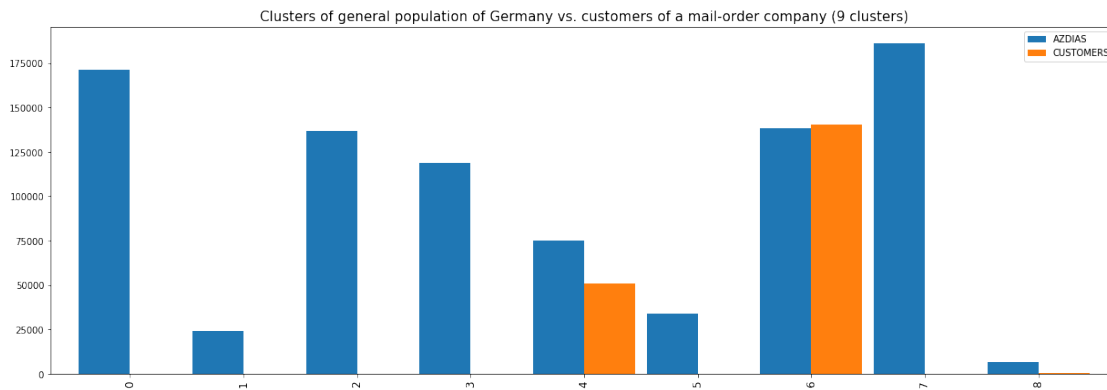
Out[85]:
   AZDIAS  CUSTOMERS
0  171094         NaN
1   24341         NaN
2  136907         NaN
3  118779         NaN
4   74893    50753.0
5   34090         NaN
6  138280   140371.0
7  186271         NaN
8    6566     528.0

```

```

In [86]: preds.plot(kind='bar', figsize=(21,7), width=.9)
plt.title('Clusters of general population of Germany vs. customers of a mail-order comp
plt.legend()
plt.show()

```



k=8

```
In [87]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 8)
```

```
In [88]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
         preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

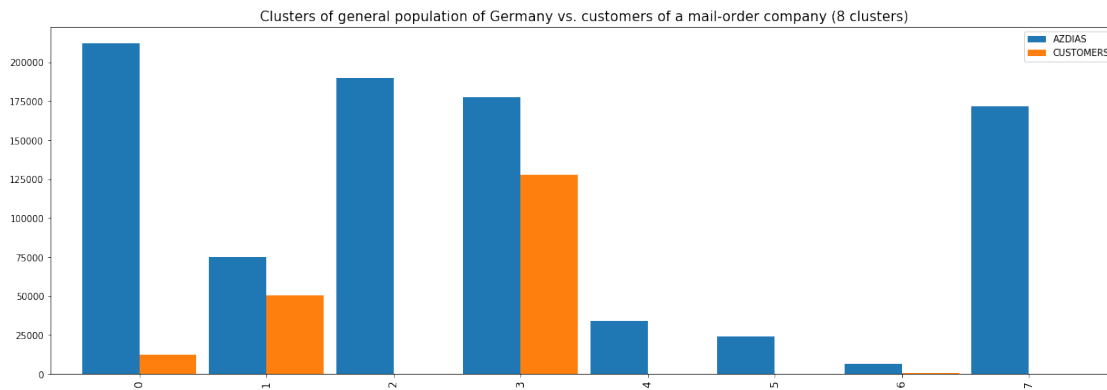
         preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                           axis=1, sort=False)
```

preds

```
Out[88]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 212283 | 12605.0 |
| 1 | 74893 | 50753.0 |
| 2 | 189846 | NaN |
| 3 | 177307 | 127692.0 |
| 4 | 34090 | 74.0 |
| 5 | 24341 | NaN |
| 6 | 6566 | 528.0 |
| 7 | 171895 | NaN |

```
In [89]: preds.plot(kind='bar', figsize=(21,7), width=.9)
         plt.title('Clusters of general population of Germany vs. customers of a mail-order company')
         plt.legend()
         plt.show()
```



k=7

```
In [90]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 7)
```

```
In [91]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
         preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

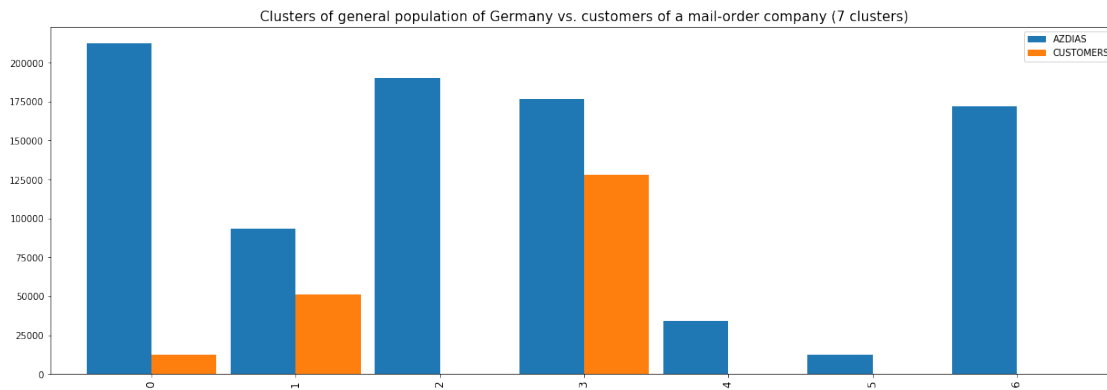
         preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                           axis=1, sort=False)
```

preds

```
Out[91]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 212598 | 12560.0 |
| 1 | 93153 | 51281.0 |
| 2 | 189985 | NaN |
| 3 | 176845 | 127738.0 |
| 4 | 34090 | 73.0 |
| 5 | 12647 | NaN |
| 6 | 171903 | NaN |

```
In [92]: preds.plot(kind='bar', figsize=(21,7), width=.9)
         plt.title('Clusters of general population of Germany vs. customers of a mail-order company')
         plt.legend()
         plt.show()
```



k=6

```
In [93]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 6)
```

```
In [94]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
         preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

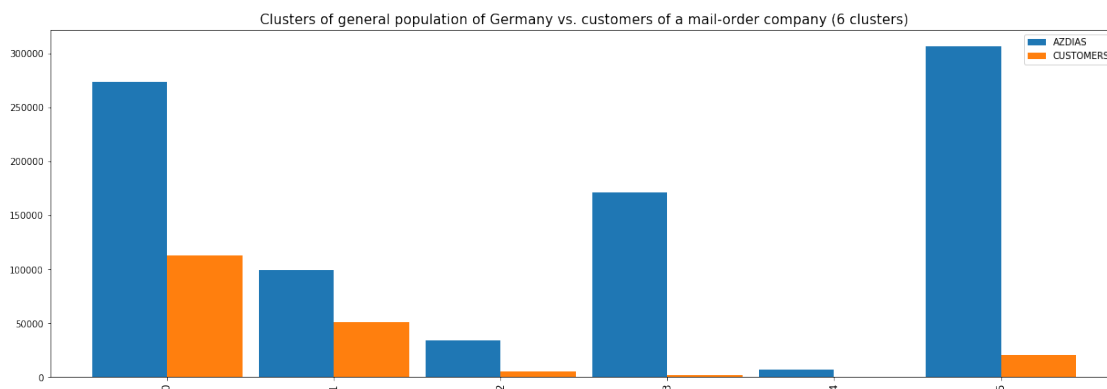
         preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                           axis=1, sort=False)
```

preds

```
Out[94]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 273671 | 112508 |
| 1 | 99234 | 50753 |
| 2 | 34090 | 5227 |
| 3 | 171105 | 1918 |
| 4 | 6566 | 528 |
| 5 | 306555 | 20718 |

```
In [95]: preds.plot(kind='bar', figsize=(21,7), width=.9)
         plt.title('Clusters of general population of Germany vs. customers of a mail-order company (6 clusters)')
         plt.legend()
         plt.show()
```



k=5

```
In [96]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 5)
```

```
In [97]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
         preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

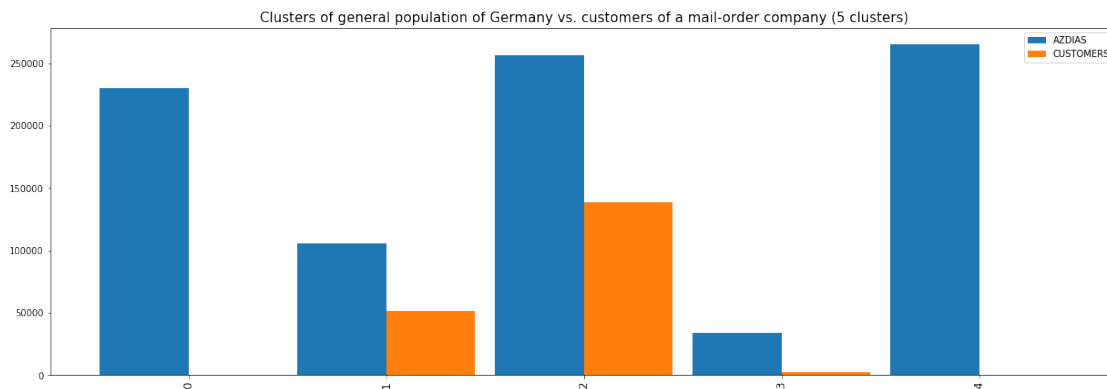
         preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                           axis=1, sort=False)
```

preds

```
Out[97]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 230007 | NaN |
| 1 | 105800 | 51281.0 |
| 2 | 256113 | 138184.0 |
| 3 | 34090 | 2187.0 |
| 4 | 265211 | NaN |

```
In [98]: preds.plot(kind='bar', figsize=(21,7), width=.9)
         plt.title('Clusters of general population of Germany vs. customers of a mail-order company (5 clusters)')
         plt.legend()
         plt.show()
```



k=4

```
In [99]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 4)
```

```
In [100]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
          preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])
```

```

preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                  axis=1, sort=False)

```

preds

```

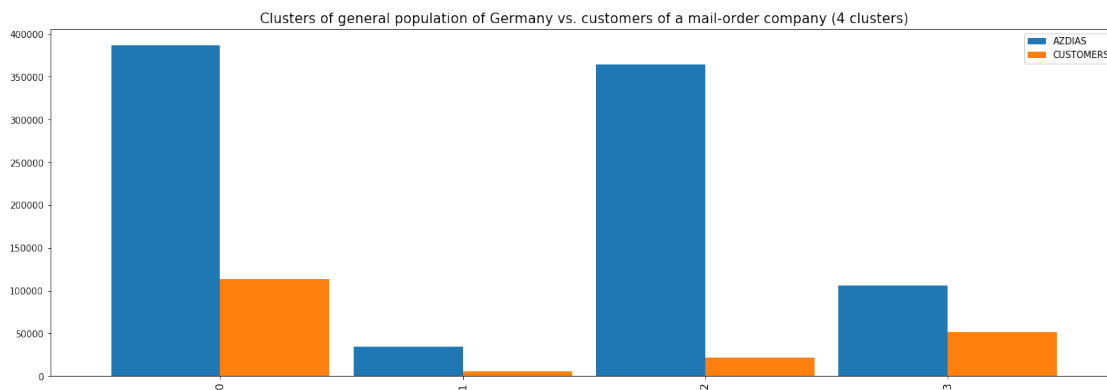
Out[100]:
   AZDIAS  CUSTOMERS
0  386994     113128
1   34090       5227
2  364337     22016
3  105800     51281

```

```

In [101]: preds.plot(kind='bar', figsize=(21,7), width=.9)
plt.title('Clusters of general population of Germany vs. customers of a mail-order company')
plt.legend()
plt.show()

```



k=3

```

In [102]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 3)

```

```

In [103]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                  axis=1, sort=False)

```

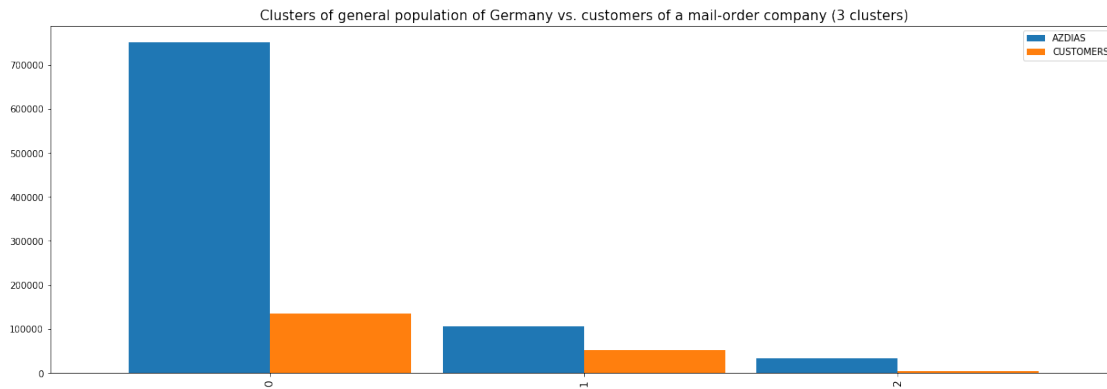
preds

```

Out[103]:
   AZDIAS  CUSTOMERS
0  751331     135144
1  105800      51281
2   34090       5227

```

```
In [104]: preds.plot(kind='bar', figsize=(21,7), width=.9)
plt.title('Clusters of general population of Germany vs. customers of a mail-order com
plt.legend()
plt.show()
```



k=2

```
In [105]: kmeans_azdias, cluster_map = kmeans(azdias_reduced, 2)

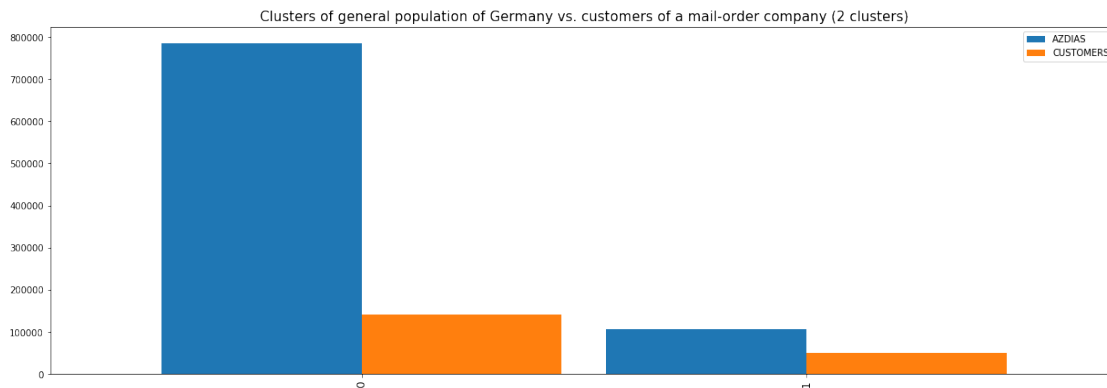
In [106]: preds_azdias = pd.DataFrame(kmeans_azdias.predict(azdias_reduced), columns=['AZDIAS'])
preds_customers = pd.DataFrame(kmeans_azdias.predict(customers_reduced), columns=['CUSTOMERS'])

preds = pd.concat([preds_azdias['AZDIAS'].value_counts(), preds_customers['CUSTOMERS'].value_counts()],
                  axis=1, sort=False)

preds
```

```
Out[106]:   AZDIAS  CUSTOMERS
0   785421     140371
1   105800      51281
```

```
In [107]: preds.plot(kind='bar', figsize=(21,7), width=.9)
plt.title('Clusters of general population of Germany vs. customers of a mail-order com
plt.legend()
plt.show()
```



Discuss results For the prepared demographic data of the general population of Germany and customers, the use of more than 4 clusters leads to the fact that a disproportion appears in the results, and some of the people from the customers simply do not fall into the corresponding clusters of the entire population.

On the one hand, we are losing some of the information, reducing the number of clusters, on the other hand, hoping that customers will be exactly like the general population of Germany is only one of the possible options for segmentation of demographic data.

For further analysis, it is necessary to choose the most acceptable 1-2 types of clustering, and on their basis try to implement certain marketing strategies. Thus, significantly reducing costs and not covering all customers. And of course, further research needs to look at the results of hierarchical clustering in order to compare several clustering methods. This analysis will take more resources, including time, than KMeans clustering

```
In [108]: kmeans_azdias_4, cluster_map_4 = kmeans(azdias_reduced, 4)
```

```
preds_azdias_4 = pd.DataFrame(kmeans_azdias_4.predict(azdias_reduced), columns=['AZDIAS'])
preds_customers_4 = pd.DataFrame(kmeans_azdias_4.predict(customers_reduced), columns=['CUSTOMERS'])
```

```
preds_4 = pd.concat([preds_azdias_4['AZDIAS'].value_counts(), preds_customers_4['CUSTOMERS'].value_counts()],
                    axis=1, sort=False)
```

```
preds_4
```

```
Out[108]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 362017 | 22227 |
| 1 | 105800 | 51281 |
| 2 | 34090 | 5227 |
| 3 | 389314 | 112917 |

```
In [109]: preds_4['AZDIAS'] = round(preds_4['AZDIAS'] / preds_4['AZDIAS'].sum() * 100, 1)
preds_4['CUSTOMERS'] = round(preds_4['CUSTOMERS'] / preds_4['CUSTOMERS'].sum() * 100, 1)
preds_4
```



```
Out[109]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 0 | 40.6 | 11.6 |
| 1 | 11.9 | 26.8 |
| 2 | 3.8 | 2.7 |
| 3 | 43.7 | 58.9 |

```
In [110]: preds_azdias_4.shape, preds_customers_4.shape
```

```
Out[110]: ((891221, 1), (191652, 1))
```

```
In [111]: kmeans_azdias_3, cluster_map_3 = kmeans(azdias_reduced, 3)
```

```
preds_azdias_3 = pd.DataFrame(kmeans_azdias_3.predict(azdias_reduced), columns=['AZDIAS'])
preds_customers_3 = pd.DataFrame(kmeans_azdias_3.predict(customers_reduced), columns=['CUSTOMERS'])

preds_3 = pd.concat([preds_azdias_3['AZDIAS'].value_counts(), preds_customers_3['CUSTOMERS'].value_counts()],
                    axis=1, sort=False)
```

```
preds_3
```

```
Out[111]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 1 | 751331 | 135144 |
| 0 | 105800 | 51281 |
| 2 | 34090 | 5227 |

```
In [112]: preds_3['AZDIAS'] = round(preds_3['AZDIAS'] / preds_3['AZDIAS'].sum() * 100, 1)
preds_3['CUSTOMERS'] = round(preds_3['CUSTOMERS'] / preds_3['CUSTOMERS'].sum() * 100, 1)
preds_3
```

```
Out[112]:
```

| | AZDIAS | CUSTOMERS |
|---|--------|-----------|
| 1 | 84.3 | 70.5 |
| 0 | 11.9 | 26.8 |
| 2 | 3.8 | 2.7 |

```
In [113]: preds_azdias_3.shape, preds_customers_3.shape
```

```
Out[113]: ((891221, 1), (191652, 1))
```

```
In [114]: 'Shape extra customers DF' + str(customers_extra.shape)
```

```
Out[114]: 'Shape extra customers DF(191652, 3)'
```

Analyze 4 clusters scenario

```
In [223]: preds_customers_4 = pd.concat([preds_customers_4, customers_extra], axis=1, sort=False)
print('Shape (4 clusters)', preds_customers_4.shape)
preds_customers_4.head()
```

```
Shape (4 clusters) (191652, 4)
```

```
Out[223]:
```

| | CUSTOMERS | CUSTOMER_GROUP | ONLINE_PURCHASE | PRODUCT_GROUP |
|---|-----------|----------------|-----------------|-------------------|
| 0 | 2 | MULTI_BUYER | No | COSMETIC_AND_FOOD |
| 1 | 1 | SINGLE_BUYER | No | FOOD |
| 2 | 0 | MULTI_BUYER | No | COSMETIC_AND_FOOD |
| 3 | 2 | MULTI_BUYER | No | COSMETIC |
| 4 | 2 | MULTI_BUYER | No | FOOD |

```
In [224]: round(pd.crosstab(preds_customers_4['CUSTOMERS'], preds_customers_4['CUSTOMER_GROUP'])
               / preds_customers_4.shape[0] * 100, 1)
```

```
Out[224]:
```

| | CUSTOMER_GROUP | MULTI_BUYER | SINGLE_BUYER |
|-----------|----------------|-------------|--------------|
| CUSTOMERS | | | |
| 0 | | 9.5 | 2.8 |
| 1 | | 17.6 | 9.2 |
| 2 | | 40.1 | 18.0 |
| 3 | | 1.8 | 1.0 |

```
In [225]: round(pd.crosstab(preds_customers_4['CUSTOMERS'], preds_customers_4['ONLINE_PURCHASE'])
               / preds_customers_4.shape[0] * 100, 1)
```

```
Out[225]:
```

| | ONLINE_PURCHASE | No | Yes |
|-----------|-----------------|------|-----|
| CUSTOMERS | | | |
| 0 | | 12.0 | 0.4 |
| 1 | | 24.1 | 2.7 |
| 2 | | 52.7 | 5.5 |
| 3 | | 2.2 | 0.5 |

```
In [226]: round(pd.crosstab(preds_customers_4['CUSTOMERS'], preds_customers_4['PRODUCT_GROUP'])
               / preds_customers_4.shape[0] * 100, 1)
```

```
Out[226]:
```

| | PRODUCT_GROUP | COSMETIC | COSMETIC_AND_FOOD | FOOD |
|-----------|---------------|----------|-------------------|------|
| CUSTOMERS | | | | |
| 0 | | 2.3 | 7.6 | 2.5 |
| 1 | | 6.4 | 13.2 | 7.1 |
| 2 | | 13.3 | 30.5 | 14.4 |
| 3 | | 0.6 | 1.3 | 0.8 |

```
In [227]: preds_customers_4['ONLINE_PURCHASE'] = preds_customers_4['ONLINE_PURCHASE'].replace({'
round(pd.pivot_table(data=preds_customers_4, index='CUSTOMERS', columns='CUSTOMER_GROU
aggfunc='mean') * 100, 1)
```

```
Out[227]:
```

| | CUSTOMER_GROUP | MULTI_BUYER | SINGLE_BUYER |
|-----------|----------------|-------------|--------------|
| CUSTOMERS | | | |
| 0 | | 2.5 | 4.6 |
| 1 | | 8.9 | 12.3 |
| 2 | | 9.2 | 9.8 |
| 3 | | 17.4 | 19.3 |

```
In [229]: round(pd.pivot_table(data=preds_customers_4, index='CUSTOMERS', columns='PRODUCT_GROUP
aggfunc='mean') * 100, 1)
```

```
Out[229]:
```

| | PRODUCT_GROUP | COSMETIC | COSMETIC_AND_FOOD | FOOD |
|-----------|---------------|----------|-------------------|------|
| CUSTOMERS | | | | |
| 0 | | 5.3 | 2.3 | 3.1 |
| 1 | | 16.2 | 7.9 | 8.7 |
| 2 | | 10.9 | 8.9 | 9.2 |
| 3 | | 23.9 | 16.6 | 15.9 |

Analyze 3 clusters scenario

```
In [230]: preds_customers_3 = pd.concat([preds_customers_3, customers_extra], axis=1, sort=False)
print('Shape (3 clusters)', preds_customers_3.shape)
preds_customers_3.head()
```

Shape (3 clusters) (191652, 4)

```
Out[230]:
```

| | CUSTOMERS | CUSTOMER_GROUP | ONLINE_PURCHASE | PRODUCT_GROUP |
|---|-----------|----------------|-----------------|-------------------|
| 0 | 0 | MULTI_BUYER | No | COSMETIC_AND_FOOD |
| 1 | 1 | SINGLE_BUYER | No | FOOD |
| 2 | 0 | MULTI_BUYER | No | COSMETIC_AND_FOOD |
| 3 | 0 | MULTI_BUYER | No | COSMETIC |
| 4 | 0 | MULTI_BUYER | No | FOOD |

```
In [231]: round(pd.crosstab(preds_customers_3['CUSTOMERS'], preds_customers_3['CUSTOMER_GROUP'])
/ preds_customers_3.shape[0] * 100, 1)
```

```
Out[231]:
```

| | CUSTOMER_GROUP | MULTI_BUYER | SINGLE_BUYER |
|-----------|----------------|-------------|--------------|
| CUSTOMERS | | | |
| 0 | | 49.7 | 20.8 |
| 1 | | 17.6 | 9.2 |
| 2 | | 1.8 | 1.0 |

```
In [232]: round(pd.crosstab(preds_customers_3['CUSTOMERS'], preds_customers_3['ONLINE_PURCHASE'])
/ preds_customers_3.shape[0] * 100, 1)
```

```
Out[232]:
```

| | ONLINE_PURCHASE | No | Yes |
|-----------|-----------------|------|-----|
| CUSTOMERS | | | |
| 0 | | 64.7 | 5.8 |
| 1 | | 24.1 | 2.7 |
| 2 | | 2.2 | 0.5 |

```
In [233]: round(pd.crosstab(preds_customers_3['CUSTOMERS'], preds_customers_3['PRODUCT_GROUP'])
/ preds_customers_3.shape[0] * 100, 1)
```

```
Out[233]:
```

| | PRODUCT_GROUP | COSMETIC | COSMETIC_AND_FOOD | FOOD |
|-----------|---------------|----------|-------------------|------|
| CUSTOMERS | | | | |
| 0 | | 15.6 | 38.1 | 16.8 |
| 1 | | 6.4 | 13.2 | 7.1 |
| 2 | | 0.6 | 1.3 | 0.8 |

```
In [234]: preds_customers_3['ONLINE_PURCHASE'] = preds_customers_3['ONLINE_PURCHASE'].replace({'
round(pd.pivot_table(data=preds_customers_3, index='CUSTOMERS', columns='CUSTOMER_GROU
aggfunc='mean') * 100, 1)
```

```
Out[234]: CUSTOMER_GROUP  MULTI_BUYER  SINGLE_BUYER
CUSTOMERS
0                7.9          9.1
1                8.9         12.3
2               17.4         19.3
```

```
In [235]: round(pd.pivot_table(data=preds_customers_3, index='CUSTOMERS', columns='PRODUCT_GROUP
aggfunc='mean') * 100, 1)
```

```
Out[235]: PRODUCT_GROUP  COSMETIC  COSMETIC_AND_FOOD  FOOD
CUSTOMERS
0                10.0                7.5    8.3
1                16.2                7.9    8.7
2                23.9               16.6   15.9
```

```
In [ ]:
```

```
In [ ]:
```

1.3 Part 2: Supervised Learning Model

Now that you've found which parts of the population are more likely to be customers of the mail-order company, it's time to build a prediction model. Each of the rows in the "MAILOUT" data files represents an individual that was targeted for a mailout campaign. Ideally, we should be able to use the demographic information from each individual to decide whether or not it will be worth it to include that person in the campaign.

The "MAILOUT" data has been split into two approximately equal parts, each with almost 43 000 data rows. In this part, you can verify your model with the "TRAIN" partition, which includes a column, "RESPONSE", that states whether or not a person became a customer of the company following the campaign. In the next part, you'll need to create predictions on the "TEST" partition, where the "RESPONSE" column has been withheld.

```
In [116]: mailout_train = pd.read_csv('../data/Term2/capstone/arvato_data/Udacity_MAILOUT_052
mailout_train.head()
```

```
/opt/conda/lib/python3.6/site-packages/IPython/core/interactiveshell.py:2785: DtypeWarning: Colu
interactivity=interactivity, compiler=compiler, result=result)
```

```
Out[116]:   LNR  AGER_TYP  AKT_DAT_KL  ALTER_HH  ALTER_KIND1  ALTER_KIND2  \
0  1763         2         1.0        8.0           NaN           NaN
1  1771         1         4.0       13.0           NaN           NaN
2  1776         1         1.0        9.0           NaN           NaN
3  1460         2         1.0        6.0           NaN           NaN
4  1783         2         1.0        9.0           NaN           NaN
```

| | ALTER_KIND3 | ALTER_KIND4 | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | \ |
|---|-------------|-------------|----------------------|---------------------|---|
| 0 | NaN | NaN | 8.0 | 15.0 | |
| 1 | NaN | NaN | 13.0 | 1.0 | |
| 2 | NaN | NaN | 7.0 | 0.0 | |
| 3 | NaN | NaN | 6.0 | 4.0 | |
| 4 | NaN | NaN | 9.0 | 53.0 | |

| | ANZ_HH_TITEL | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE | \ |
|---|--------------|------------|--------------|----------------------------|---|
| 0 | 0.0 | 0.0 | 1.0 | 13.0 | |
| 1 | 0.0 | 0.0 | 2.0 | 1.0 | |
| 2 | NaN | 0.0 | 0.0 | 1.0 | |
| 3 | 0.0 | 0.0 | 2.0 | 4.0 | |
| 4 | 0.0 | 0.0 | 1.0 | 44.0 | |

| | ANZ_TITEL | ARBEIT | BALLRAUM | CAMEO_DEU_2015 | CAMEO_DEUG_2015 | CAMEO_INTL_2015 | \ |
|---|-----------|--------|----------|----------------|-----------------|-----------------|---|
| 0 | 0.0 | 3.0 | 5.0 | 5D | 5 | 34 | |
| 1 | 0.0 | 2.0 | 5.0 | 5B | 5 | 32 | |
| 2 | 0.0 | 4.0 | 1.0 | 2D | 2 | 14 | |
| 3 | 0.0 | 4.0 | 2.0 | 2D | 2 | 14 | |
| 4 | 0.0 | 3.0 | 4.0 | 7B | 7 | 41 | |

| | CJT_GESAMTTYP | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 | \ |
|---|---------------|-------------------|-----------|-----------|-----------|-----|
| 0 | 2.0 | | 5.0 | 2.0 | 2.0 | 5.0 |
| 1 | 2.0 | | 2.0 | 2.0 | 2.0 | 4.0 |
| 2 | 4.0 | | 5.0 | 1.0 | 1.0 | 5.0 |
| 3 | 2.0 | | 5.0 | 2.0 | 2.0 | 5.0 |
| 4 | 6.0 | | 5.0 | 1.0 | 2.0 | 5.0 |

| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 | D19_BANKEN_ANZ_24 | \ |
|---|-----------|-----------|-----------|-------------------|-------------------|---|
| 0 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 1 | 3.0 | 5.0 | 4.0 | 1 | 1 | |
| 2 | 5.0 | 5.0 | 5.0 | 0 | 1 | |
| 3 | 5.0 | 5.0 | 4.0 | 0 | 0 | |
| 4 | 5.0 | 5.0 | 5.0 | 0 | 0 | |

| | D19_BANKEN_DATUM | D19_BANKEN_DIREKT | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | \ |
|---|------------------|-------------------|------------------|------------------|---|
| 0 | 10 | 0 | 0 | 0 | |
| 1 | 5 | 0 | 0 | 3 | |
| 2 | 7 | 0 | 0 | 0 | |
| 3 | 10 | 0 | 0 | 0 | |
| 4 | 10 | 0 | 0 | 0 | |

| | D19_BANKEN_OFFLINE_DATUM | D19_BANKEN_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |
| 2 | 10 | 7 | |
| 3 | 10 | 10 | |

| | | | | | |
|---|--|----|--|----|--|
| 4 | | 10 | | 10 | |
|---|--|----|--|----|--|

| | | | | |
|---|----------------------------|-----------------|--------------------|---|
| | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST | D19_BEKLEIDUNG_GEH | \ |
| 0 | 0.0 | 0 | 0 | |
| 1 | 0.0 | 0 | 0 | |
| 2 | 0.0 | 5 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 0.0 | 0 | 5 | |

| | | | | | |
|---|---------------------|-------------|--------------|-------------|---|
| | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO | D19_BUCH_CD | \ |
| 0 | 0 | 0 | 0 | 0 | |
| 1 | 6 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | |
| 3 | 6 | 0 | 0 | 6 | |
| 4 | 0 | 7 | 0 | 0 | |

| | | | | | | |
|---|----------------|---------------------|-------------|--------------|------------|---|
| | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE | D19_FREIZEIT | D19_GARTEN | \ |
| 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 6 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 0 | |
| 4 | 0 | 0 | 0 | 0 | 0 | |

| | | | | |
|---|-------------------|-------------------|------------------|---|
| | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 | D19_GESAMT_DATUM | \ |
| 0 | 0 | 0 | 9 | |
| 1 | 2 | 2 | 1 | |
| 2 | 0 | 2 | 6 | |
| 3 | 0 | 0 | 10 | |
| 4 | 0 | 1 | 7 | |

| | | | |
|---|--------------------------|-------------------------|---|
| | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM | \ |
| 0 | 9 | 9 | |
| 1 | 1 | 8 | |
| 2 | 6 | 7 | |
| 3 | 10 | 10 | |
| 4 | 7 | 10 | |

| | | | | | |
|---|----------------------------|--------------|---------------|-------------------|---|
| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO | D19_KINDERARTIKEL | \ |
| 0 | 0.0 | 0 | 0 | 0 | |
| 1 | 0.0 | 6 | 6 | 0 | |
| 2 | 0.0 | 0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | 0 | |
| 4 | 0.0 | 0 | 0 | 0 | |

| | | | | | |
|---|---------------|-------------------|--------------|------------------|---|
| | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK | D19_LEBENSMITTEL | \ |
| 0 | 3.0 | 2 | 7 | 0 | |
| 1 | 1.0 | 4 | 6 | 0 | |
| 2 | 2.0 | 2 | 0 | 5 | |

| | | | | |
|---|-----|---|---|---|
| 3 | 3.0 | 2 | 7 | 0 |
| 4 | 2.0 | 2 | 0 | 0 |

| | D19_LETZTER_KAUF_BRANCHE | D19_LOTTO | D19_NAHRUNGSEGAENZUNG | D19_RATGEBER | \ |
|---|--------------------------|-----------|-----------------------|--------------|---|
| 0 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 1 | D19_TELKO_MOBILE | 0.0 | 0 | 0 | |
| 2 | D19_LEBENSMITTEL | 0.0 | 0 | 0 | |
| 3 | D19_UNBEKANNT | 7.0 | 0 | 0 | |
| 4 | D19_BEKLEIDUNG_GEH | 0.0 | 0 | 2 | |

| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE | D19_SOZIALES | \ |
|---|------------|-------------------|------------|--------------|--------------|---|
| 0 | 0 | 0 | 0 | 0 | 1.0 | |
| 1 | 0 | 0 | 0 | 6 | 5.0 | |
| 2 | 0 | 0 | 0 | 6 | 2.0 | |
| 3 | 0 | 0 | 0 | 6 | 1.0 | |
| 4 | 0 | 0 | 0 | 7 | 1.0 | |

| | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 | D19_TELKO_DATUM | \ |
|---|-------------|------------------|------------------|-----------------|---|
| 0 | 0 | 0 | 0 | 10 | |
| 1 | 0 | 1 | 1 | 1 | |
| 2 | 6 | 0 | 0 | 10 | |
| 3 | 7 | 0 | 0 | 10 | |
| 4 | 0 | 0 | 0 | 10 | |

| | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM | D19_TELKO_ONLINE_DATUM | \ |
|---|------------------|-------------------------|------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 3 | 1 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 0 | 10 | 10 | |

| | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST | D19_TIERARTIKEL | \ |
|---|---------------------------|----------------|-----------------|---|
| 0 | 0.0 | 0 | 0 | |
| 1 | 0.0 | 0 | 6 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 0.0 | 0 | 0 | |

| | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 | D19_VERSAND_DATUM | \ |
|---|--------------------|--------------------|-------------------|---|
| 0 | 0 | 0 | 9 | |
| 1 | 1 | 1 | 5 | |
| 2 | 0 | 1 | 6 | |
| 3 | 0 | 0 | 10 | |
| 4 | 0 | 1 | 7 | |

| | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM | \ |
|---|---------------------------|--------------------------|---|
| 0 | 9 | 9 | |
| 1 | 5 | 8 | |

| | | |
|---|----|----|
| 2 | 6 | 10 |
| 3 | 10 | 10 |
| 4 | 7 | 10 |

| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 | \ |
|---|-----------------------------|------------------|------------------|---|
| 0 | 0.0 | 0 | 0 | |
| 1 | 0.0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 0.0 | 0 | 0 | |

| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM | \ |
|---|------------------|-----------------|-------------------------|---|
| 0 | 0 | 9 | 10 | |
| 1 | 0 | 10 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 0 | 10 | 10 | |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN | \ |
|---|------------------------|---------------------------|--------------------|---|
| 0 | 10 | 0.0 | 6 | |
| 1 | 10 | 0.0 | 0 | |
| 2 | 10 | 0.0 | 0 | |
| 3 | 10 | 0.0 | 0 | |
| 4 | 10 | 0.0 | 0 | |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG | EINGEFUEGT_AM | \ |
|---|-------------------|-------------------|----------|---------------------|---|
| 0 | 6 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 1 | 6 | 3 | 1.0 | 1997-05-14 00:00:00 | |
| 2 | 6 | 0 | 1.0 | 1995-05-24 00:00:00 | |
| 3 | 6 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 4 | 7 | 0 | 1.0 | 1992-02-10 00:00:00 | |

| | EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER | \ |
|---|----------------------|----------|-----------|----------------|---|
| 0 | 2004.0 | 5.0 | 47.0 | 2 | |
| 1 | 1994.0 | 1.0 | 56.0 | 2 | |
| 2 | 1997.0 | 6.0 | 36.0 | 1 | |
| 3 | 1994.0 | 6.0 | 41.0 | 1 | |
| 4 | 1994.0 | 5.0 | 55.0 | 1 | |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER | FINANZ_UNAUFFAELLIGER | \ |
|---|------------------|-------------------|---------------|-----------------------|---|
| 0 | 5 | 3 | 1 | 1 | |
| 1 | 3 | 5 | 1 | 1 | |
| 2 | 2 | 5 | 1 | 1 | |
| 3 | 5 | 4 | 1 | 2 | |
| 4 | 5 | 3 | 1 | 1 | |

| | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE | GEBAEUDETYP | GEBAEUDETYP_RASTER | \ |
|---|------------------|-----------|--------------|-------------|--------------------|---|
| 0 | 5 | 6 | 2.0 | 3.0 | 3.0 | |

| | | | | | |
|---|---|---|-----|-----|-----|
| 1 | 5 | 6 | 4.0 | 1.0 | 4.0 |
| 2 | 4 | 5 | 4.0 | 2.0 | 4.0 |
| 3 | 5 | 5 | 3.0 | 1.0 | 3.0 |
| 4 | 5 | 5 | 3.0 | 3.0 | 4.0 |

| | GEBURTSJAHR | GEMEINDETYPE | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP \ |
|---|-------------|--------------|-----------------|------------------|--------------|
| 0 | 0 | 12.0 | 8.0 | 0 | 2 |
| 1 | 1957 | 40.0 | 8.0 | 0 | 2 |
| 2 | 1929 | 11.0 | 3.0 | 1 | 1 |
| 3 | 1924 | 11.0 | 5.0 | 1 | 2 |
| 4 | 1936 | 12.0 | 4.0 | 0 | 3 |

| | HH_DELTA_FLAG | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 \ |
|---|---------------|--------------------|------------|--------------|----------------|
| 0 | 0.0 | 6.0 | 2.0 | 1.0 | 3.0 |
| 1 | 0.0 | 1.0 | 7.0 | 3.0 | 1.0 |
| 2 | NaN | 1.0 | 3.0 | 1.0 | 3.0 |
| 3 | 0.0 | 4.0 | 4.0 | 2.0 | 3.0 |
| 4 | 0.0 | 4.0 | 2.0 | 1.0 | 1.0 |

| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 | KBA05_ANTG2 \ |
|---|--------------|--------------|--------------|-------------|---------------|
| 0 | 3.0 | 4.0 | 0.0 | 0.0 | 0.0 |
| 1 | 5.0 | 3.0 | 3.0 | 3.0 | 0.0 |
| 2 | 1.0 | 5.0 | 1.0 | 3.0 | 0.0 |
| 3 | 1.0 | 4.0 | 0.0 | 2.0 | 3.0 |
| 4 | 3.0 | 5.0 | 1.0 | 1.0 | 0.0 |

| | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT | KBA05_BAUMAX | KBA05_CCM1 \ |
|---|-------------|-------------|----------------|--------------|--------------|
| 0 | 0.0 | 2.0 | 3.0 | 4.0 | 1.0 |
| 1 | 0.0 | 0.0 | 3.0 | 1.0 | 3.0 |
| 2 | 0.0 | 0.0 | 4.0 | 1.0 | 3.0 |
| 3 | 0.0 | 0.0 | 3.0 | 0.0 | 3.0 |
| 4 | 0.0 | 1.0 | 3.0 | 0.0 | 2.0 |

| | KBA05_CCM2 | KBA05_CCM3 | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ \ |
|---|------------|------------|------------|--------------|------------|-------------|
| 0 | 5.0 | 2.0 | 3.0 | 3.0 | 3.0 | 1.0 |
| 1 | 3.0 | 3.0 | 1.0 | 3.0 | 3.0 | 5.0 |
| 2 | 3.0 | 2.0 | 1.0 | 2.0 | 3.0 | 5.0 |
| 3 | 2.0 | 2.0 | 4.0 | 1.0 | 4.0 | 3.0 |
| 4 | 3.0 | 4.0 | 0.0 | 2.0 | 3.0 | 1.0 |

| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 | KBA05_HERST5 \ |
|---|--------------|--------------|--------------|--------------|----------------|
| 0 | 4.0 | 2.0 | 4.0 | 1.0 | 3.0 |
| 1 | 1.0 | 4.0 | 4.0 | 1.0 | 2.0 |
| 2 | 2.0 | 3.0 | 3.0 | 4.0 | 2.0 |
| 3 | 5.0 | 3.0 | 1.0 | 2.0 | 3.0 |
| 4 | 3.0 | 4.0 | 2.0 | 3.0 | 1.0 |

| | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 | KBA05_KRSHERST2 \ |
|--|-----------------|----------------|-----------------|-------------------|
|--|-----------------|----------------|-----------------|-------------------|

| | | | | |
|---|-----|-----|-----|-----|
| 0 | 4.0 | 3.0 | 4.0 | 2.0 |
| 1 | 2.0 | 2.0 | 2.0 | 4.0 |
| 2 | 3.0 | 5.0 | 4.0 | 3.0 |
| 3 | 2.0 | 4.0 | 5.0 | 3.0 |
| 4 | 2.0 | 3.0 | 3.0 | 4.0 |

| | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER | KBA05_KRSVAN | KBA05_KRSZUL | \ |
|---|-----------------|----------------|---------------|--------------|--------------|---|
| 0 | 5.0 | 1.0 | 2.0 | 2.0 | 1.0 | |
| 1 | 4.0 | 2.0 | 1.0 | 3.0 | 2.0 | |
| 2 | 3.0 | 3.0 | 2.0 | 2.0 | 2.0 | |
| 3 | 2.0 | 2.0 | 2.0 | 2.0 | 3.0 | |
| 4 | 3.0 | 1.0 | 2.0 | 2.0 | 3.0 | |

| | KBA05_KW1 | KBA05_KW2 | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST | \ |
|---|-----------|-----------|-----------|-------------|-------------|----------------|---|
| 0 | 1.0 | 4.0 | 2.0 | 5.0 | 3.0 | 3.0 | |
| 1 | 3.0 | 4.0 | 0.0 | 3.0 | 1.0 | 2.0 | |
| 2 | 3.0 | 3.0 | 1.0 | 5.0 | 4.0 | 4.0 | |
| 3 | 2.0 | 1.0 | 4.0 | 5.0 | 4.0 | 1.0 | |
| 4 | 2.0 | 4.0 | 2.0 | 5.0 | 4.0 | 2.0 | |

| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 | KBA05_MOD3 | \ |
|---|--------------|---------------|------------|------------|------------|---|
| 0 | 2.0 | 3.0 | 3.0 | 3.0 | 5.0 | |
| 1 | 3.0 | 3.0 | 0.0 | 5.0 | 1.0 | |
| 2 | 1.0 | 2.0 | 1.0 | 2.0 | 4.0 | |
| 3 | 4.0 | 1.0 | 3.0 | 3.0 | 3.0 | |
| 4 | 2.0 | 1.0 | 2.0 | 3.0 | 4.0 | |

| | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP | KBA05_MOTOR | KBA05_MOTRAD | \ |
|---|------------|------------|---------------|-------------|--------------|---|
| 0 | 1.0 | 0.0 | 4.0 | 2.0 | 1.0 | |
| 1 | 3.0 | 2.0 | 2.0 | 3.0 | 1.0 | |
| 2 | 4.0 | 0.0 | 4.0 | 2.0 | 0.0 | |
| 3 | 3.0 | 0.0 | 1.0 | 4.0 | 0.0 | |
| 4 | 2.0 | 1.0 | 3.0 | 3.0 | 1.0 | |

| | KBA05_SEG1 | KBA05_SEG10 | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 | \ |
|---|------------|-------------|------------|------------|------------|------------|---|
| 0 | 0.0 | 2.0 | 2.0 | 3.0 | 3.0 | 2.0 | |
| 1 | 2.0 | 2.0 | 3.0 | 1.0 | 5.0 | 0.0 | |
| 2 | 1.0 | 1.0 | 4.0 | 3.0 | 2.0 | 2.0 | |
| 3 | 0.0 | 4.0 | 3.0 | 3.0 | 1.0 | 1.0 | |
| 4 | 0.0 | 1.0 | 2.0 | 4.0 | 3.0 | 2.0 | |

| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 | KBA05_VORBO | KBA05_VORB1 | \ |
|---|------------|------------|------------|------------|-------------|-------------|---|
| 0 | 0.0 | 0.0 | 2.0 | 0.0 | 3.0 | 3.0 | |
| 1 | 0.0 | 0.0 | 0.0 | 2.0 | 2.0 | 5.0 | |
| 2 | 0.0 | 0.0 | 0.0 | 1.0 | 3.0 | 4.0 | |
| 3 | 0.0 | 1.0 | 3.0 | 0.0 | 5.0 | 2.0 | |
| 4 | 0.0 | 0.0 | 1.0 | 1.0 | 4.0 | 3.0 | |

| | KBA05_VORB2 | KBA05_ZUL1 | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | \ |
|---|-------------|------------|------------|------------|------------|---|
| 0 | 3.0 | 3.0 | 4.0 | 3.0 | 1.0 | |
| 1 | 2.0 | 3.0 | 3.0 | 4.0 | 1.0 | |
| 2 | 1.0 | 3.0 | 2.0 | 4.0 | 2.0 | |
| 3 | 3.0 | 2.0 | 1.0 | 5.0 | 4.0 | |
| 4 | 0.0 | 3.0 | 2.0 | 4.0 | 3.0 | |

| | KBA13_ALTERHALTER_30 | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | \ |
|---|----------------------|----------------------|----------------------|---|
| 0 | 4.0 | 4.0 | 1.0 | |
| 1 | 2.0 | 3.0 | 4.0 | |
| 2 | 5.0 | 4.0 | 1.0 | |
| 3 | 1.0 | 1.0 | 1.0 | |
| 4 | 3.0 | 4.0 | 2.0 | |

| | KBA13_ALTERHALTER_61 | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 | \ |
|---|----------------------|-------------|-------------|-------------|-------------|---|
| 0 | 4.0 | 1.0 | 4.0 | 3.0 | 1.0 | |
| 1 | 3.0 | 3.0 | 2.0 | 1.0 | 0.0 | |
| 2 | 3.0 | 4.0 | 1.0 | 0.0 | 0.0 | |
| 3 | 5.0 | 4.0 | 2.0 | 1.0 | 0.0 | |
| 4 | 4.0 | 2.0 | 2.0 | 1.0 | 1.0 | |

| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX | KBA13_BJ_1999 | \ |
|---|------------------|------------|-----------------|--------------|---------------|---|
| 0 | 390.0 | 5.0 | 2.0 | 5.0 | 3.0 | |
| 1 | 586.0 | 3.0 | 4.0 | 1.0 | 4.0 | |
| 2 | 297.0 | 4.0 | 2.0 | 1.0 | 3.0 | |
| 3 | 373.0 | 4.0 | 3.0 | 1.0 | 2.0 | |
| 4 | 285.0 | 4.0 | 3.0 | 1.0 | 1.0 | |

| | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 | KBA13_BJ_2008 | KBA13_BJ_2009 | \ |
|---|---------------|---------------|---------------|---------------|---------------|---|
| 0 | 3.0 | 4.0 | 3.0 | 2.0 | 3.0 | |
| 1 | 3.0 | 3.0 | 3.0 | 2.0 | 3.0 | |
| 2 | 2.0 | 3.0 | 3.0 | 5.0 | 2.0 | |
| 3 | 2.0 | 3.0 | 4.0 | 5.0 | 1.0 | |
| 4 | 1.0 | 1.0 | 4.0 | 5.0 | 3.0 | |

| | KBA13_BMW | KBA13_CCM_0_1400 | KBA13_CCM_1000 | KBA13_CCM_1200 | \ |
|---|-----------|------------------|----------------|----------------|---|
| 0 | 4.0 | 2.0 | 3.0 | 2.0 | |
| 1 | 1.0 | 0.0 | 3.0 | 0.0 | |
| 2 | 3.0 | 2.0 | 2.0 | 0.0 | |
| 3 | 5.0 | 0.0 | 4.0 | 1.0 | |
| 4 | 4.0 | 0.0 | 1.0 | 0.0 | |

| | KBA13_CCM_1400 | KBA13_CCM_1401_2500 | KBA13_CCM_1500 | KBA13_CCM_1600 | \ |
|---|----------------|---------------------|----------------|----------------|---|
| 0 | 3.0 | 3.0 | 1.0 | 4.0 | |
| 1 | 3.0 | 4.0 | 1.0 | 2.0 | |
| 2 | 4.0 | 4.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 1.0 | 4.0 | 3.0 | |
| 4 | 2.0 | 4.0 | 1.0 | 2.0 | |

| | KBA13_CCM_1800 | KBA13_CCM_2000 | KBA13_CCM_2500 | KBA13_CCM_2501 | \ |
|---|----------------|----------------|----------------|----------------|---|
| 0 | 3.0 | 3.0 | 3.0 | 4.0 | |
| 1 | 4.0 | 4.0 | 3.0 | 2.0 | |
| 2 | 0.0 | 5.0 | 2.0 | 2.0 | |
| 3 | 0.0 | 2.0 | 4.0 | 5.0 | |
| 4 | 5.0 | 3.0 | 5.0 | 3.0 | |

| | KBA13_CCM_3000 | KBA13_CCM_3001 | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE | \ |
|---|----------------|----------------|-----------------|--------------------|---|
| 0 | 3.0 | 5.0 | 2.0 | 2.0 | |
| 1 | 2.0 | 1.0 | 3.0 | 2.0 | |
| 2 | 2.0 | 1.0 | 2.0 | 4.0 | |
| 3 | 5.0 | 5.0 | 1.0 | 4.0 | |
| 4 | 4.0 | 1.0 | 2.0 | 3.0 | |

| | KBA13_FIAT | KBA13_FORD | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 | \ |
|---|------------|------------|-----------|-----------------|-----------------|---|
| 0 | 5.0 | 3.0 | 2.0 | 3.0 | 4.0 | |
| 1 | 2.0 | 4.0 | 4.0 | 3.0 | 2.0 | |
| 2 | 1.0 | 3.0 | 4.0 | 3.0 | 5.0 | |
| 3 | 3.0 | 1.0 | 4.0 | 1.0 | 1.0 | |
| 4 | 3.0 | 2.0 | 2.0 | 2.0 | 3.0 | |

| | KBA13_HALTER_30 | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | 4.0 | 5.0 | 4.0 | 2.0 | |
| 1 | 2.0 | 1.0 | 3.0 | 4.0 | |
| 2 | 5.0 | 5.0 | 4.0 | 3.0 | |
| 3 | 1.0 | 1.0 | 1.0 | 2.0 | |
| 4 | 4.0 | 4.0 | 4.0 | 3.0 | |

| | KBA13_HALTER_50 | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | 2.0 | 1.0 | 1.0 | 4.0 | |
| 1 | 4.0 | 4.0 | 3.0 | 3.0 | |
| 2 | 2.0 | 1.0 | 1.0 | 2.0 | |
| 3 | 1.0 | 1.0 | 2.0 | 5.0 | |
| 4 | 2.0 | 3.0 | 2.0 | 3.0 | |

| | KBA13_HALTER_66 | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | \ |
|---|-----------------|-------------------|---------------------|---|
| 0 | 4.0 | 2.0 | 3.0 | |
| 1 | 3.0 | 3.0 | 5.0 | |
| 2 | 3.0 | 3.0 | 3.0 | |
| 3 | 5.0 | 1.0 | 3.0 | |
| 4 | 4.0 | 2.0 | 4.0 | |

| | KBA13_HERST_BMW_BENZ | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | \ |
|---|----------------------|--------------------|-----------------------|---|
| 0 | 4.0 | 5.0 | 2.0 | |
| 1 | 2.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 4.0 | 3.0 | |
| 3 | 5.0 | 2.0 | 1.0 | |

| | | | |
|---|-----|-----|-----|
| 4 | 4.0 | 3.0 | 2.0 |
|---|-----|-----|-----|

| | KBA13_HERST_SONST | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | \ |
|---|-------------------|-----------|-----------------|---------------|---|
| 0 | 2.0 | 3.0 | 3.0 | 1.0 | |
| 1 | 2.0 | 3.0 | 4.0 | 3.0 | |
| 2 | 4.0 | 3.0 | 1.0 | 1.0 | |
| 3 | 4.0 | 3.0 | 5.0 | 1.0 | |
| 4 | 3.0 | 2.0 | 1.0 | 1.0 | |

| | KBA13_KMH_140 | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | \ |
|---|---------------|-------------------|---------------|---------------|---|
| 0 | 4.0 | 2.0 | 3.0 | 2.0 | |
| 1 | 4.0 | 3.0 | 3.0 | 3.0 | |
| 2 | 1.0 | 4.0 | 2.0 | 5.0 | |
| 3 | 5.0 | 1.0 | 1.0 | 2.0 | |
| 4 | 1.0 | 2.0 | 1.0 | 4.0 | |

| | KBA13_KMH_211 | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | \ |
|---|---------------|---------------|---------------|----------------|---|
| 0 | 3.0 | 3.0 | 1.0 | 2.0 | |
| 1 | 2.0 | 2.0 | 1.0 | 4.0 | |
| 2 | 2.0 | 2.0 | 1.0 | 3.0 | |
| 3 | 5.0 | 5.0 | 1.0 | 4.0 | |
| 4 | 5.0 | 5.0 | 1.0 | 3.0 | |

| | KBA13_KRSHERST_AUDI_VW | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | \ |
|---|------------------------|-------------------------|--------------------------|---|
| 0 | 3.0 | 3.0 | 3.0 | |
| 1 | 4.0 | 2.0 | 3.0 | |
| 2 | 4.0 | 4.0 | 2.0 | |
| 3 | 3.0 | 5.0 | 1.0 | |
| 4 | 3.0 | 3.0 | 2.0 | |

| | KBA13_KRSSEG_KLEIN | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | \ |
|---|--------------------|-------------------|------------------|------------------|---|
| 0 | 2.0 | 2.0 | 1.0 | 2.0 | |
| 1 | 2.0 | 3.0 | 2.0 | 2.0 | |
| 2 | 2.0 | 2.0 | 2.0 | 2.0 | |
| 3 | 1.0 | 3.0 | 2.0 | 2.0 | |
| 4 | 2.0 | 2.0 | 2.0 | 1.0 | |

| | KBA13_KW_0_60 | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 | \ |
|---|---------------|--------------|--------------|--------------|-------------|---|
| 0 | 2.0 | 4.0 | 1.0 | 3.0 | 3.0 | |
| 1 | 3.0 | 3.0 | 3.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 4.0 | 3.0 | 2.0 | 1.0 | |
| 3 | 2.0 | 2.0 | 3.0 | 5.0 | 1.0 | |
| 4 | 1.0 | 2.0 | 5.0 | 3.0 | 1.0 | |

| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 | KBA13_KW_70 | \ |
|---|-------------|-------------|-------------|-----------------|-------------|---|
| 0 | 3.0 | 2.0 | 2.0 | 4.0 | 4.0 | |
| 1 | 3.0 | 2.0 | 0.0 | 4.0 | 3.0 | |
| 2 | 2.0 | 3.0 | 3.0 | 4.0 | 1.0 | |

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 3 | 3.0 | 0.0 | 1.0 | 2.0 | 0.0 |
| 4 | 1.0 | 0.0 | 2.0 | 5.0 | 0.0 |

| | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA | KBA13_MERCEDES | KBA13_MOTOR \ |
|---|-------------|-------------|-------------|----------------|---------------|
| 0 | 3.0 | 2.0 | 2.0 | 3.0 | 3.0 |
| 1 | 4.0 | 3.0 | 2.0 | 3.0 | 4.0 |
| 2 | 3.0 | 4.0 | 4.0 | 3.0 | 2.0 |
| 3 | 2.0 | 3.0 | 2.0 | 5.0 | 2.0 |
| 4 | 5.0 | 4.0 | 3.0 | 4.0 | 2.0 |

| | KBA13_NISSAN | KBA13_OPEL | KBA13_PEUGEOT | KBA13_RENAULT \ |
|---|--------------|------------|---------------|-----------------|
| 0 | 3.0 | 2.0 | 4.0 | 3.0 |
| 1 | 3.0 | 3.0 | 3.0 | 3.0 |
| 2 | 3.0 | 3.0 | 5.0 | 4.0 |
| 3 | 1.0 | 1.0 | 3.0 | 2.0 |
| 4 | 2.0 | 2.0 | 4.0 | 3.0 |

| | KBA13_SEG_GELAENDEWAGEN | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST \ |
|---|-------------------------|-------------------------|---------------------|
| 0 | 3.0 | 1.0 | 2.0 |
| 1 | 3.0 | 4.0 | 3.0 |
| 2 | 2.0 | 2.0 | 3.0 |
| 3 | 4.0 | 3.0 | 1.0 |
| 4 | 3.0 | 4.0 | 2.0 |

| | KBA13_SEG_KLEINWAGEN | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS \ |
|---|----------------------|-------------------------|----------------------|
| 0 | 2.0 | 3.0 | 1.0 |
| 1 | 3.0 | 2.0 | 3.0 |
| 2 | 3.0 | 4.0 | 2.0 |
| 3 | 1.0 | 1.0 | 2.0 |
| 4 | 2.0 | 4.0 | 3.0 |

| | KBA13_SEG_MINIWAGEN | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE \ |
|---|---------------------|------------------------|-------------------------------|
| 0 | 2.0 | 5.0 | 4.0 |
| 1 | 4.0 | 4.0 | 3.0 |
| 2 | 2.0 | 5.0 | 2.0 |
| 3 | 3.0 | 3.0 | 5.0 |
| 4 | 3.0 | 4.0 | 4.0 |

| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN \ |
|---|----------------------|--------------------|------------------------|
| 0 | 0.0 | 3.0 | 5.0 |
| 1 | 1.0 | 3.0 | 2.0 |
| 2 | 0.0 | 5.0 | 2.0 |
| 3 | 5.0 | 4.0 | 5.0 |
| 4 | 3.0 | 2.0 | 4.0 |

| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE | KBA13_SITZE_4 \ |
|---|---------------------|---------------|----------------------|-----------------|
| 0 | 4.0 | 1.0 | 0.0 | 3.0 |
| 1 | 5.0 | 3.0 | 0.0 | 1.0 |

| | | | | |
|---|-----|-----|-----|-----|
| 2 | 2.0 | 2.0 | 2.0 | 2.0 |
| 3 | 4.0 | 3.0 | 3.0 | 5.0 |
| 4 | 2.0 | 4.0 | 3.0 | 3.0 |

| | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA | KBA13_VORB_0 | KBA13_VORB_1 \ |
|---|---------------|---------------|--------------|--------------|----------------|
| 0 | 3.0 | 3.0 | 3.0 | 3.0 | 4.0 |
| 1 | 4.0 | 5.0 | 4.0 | 3.0 | 3.0 |
| 2 | 4.0 | 2.0 | 3.0 | 3.0 | 3.0 |
| 3 | 1.0 | 2.0 | 1.0 | 4.0 | 3.0 |
| 4 | 3.0 | 3.0 | 3.0 | 5.0 | 1.0 |

| | KBA13_VORB_1_2 | KBA13_VORB_2 | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK \ |
|---|----------------|--------------|--------------|----------|--------------|-------|
| 0 | 4.0 | 4.0 | 2.0 | 3.0 | NaN | 3.0 |
| 1 | 4.0 | 3.0 | 3.0 | 5.0 | 2.0 | 2.0 |
| 2 | 3.0 | 3.0 | 2.0 | 2.0 | NaN | 1.0 |
| 3 | 2.0 | 2.0 | 3.0 | 3.0 | NaN | 1.0 |
| 4 | 1.0 | 2.0 | 0.0 | 4.0 | NaN | 3.0 |

| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN | LP_FAMILIE_GROB \ |
|---|------------|-------------|-------------|-----------------|-------------------|
| 0 | 4 | 1.0 | 1.0 | 1.0 | 1.0 |
| 1 | 4 | 4.0 | 0.0 | 2.0 | 2.0 |
| 2 | 4 | 2.0 | 0.0 | 0.0 | 0.0 |
| 3 | 4 | 3.0 | 0.0 | 2.0 | 2.0 |
| 4 | 4 | 3.0 | 0.0 | 1.0 | 1.0 |

| | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB | LP_STATUS_FEIN | LP_STATUS_GROB \ |
|---|---------------------|---------------------|----------------|------------------|
| 0 | 8.0 | 2.0 | 3.0 | 2.0 |
| 1 | 19.0 | 5.0 | 9.0 | 4.0 |
| 2 | 0.0 | 0.0 | 10.0 | 5.0 |
| 3 | 16.0 | 4.0 | 3.0 | 2.0 |
| 4 | 9.0 | 3.0 | 6.0 | 3.0 |

| | MIN_GEBAEUDEJAHR | MOBI_RASTER | MOBI_REGIO | NATIONALITAET_KZ \ |
|---|------------------|-------------|------------|--------------------|
| 0 | 1992.0 | 1.0 | 1.0 | 1 |
| 1 | 1994.0 | 4.0 | 5.0 | 2 |
| 2 | 1995.0 | 5.0 | 5.0 | 1 |
| 3 | 1992.0 | 1.0 | 3.0 | 1 |
| 4 | 1992.0 | 1.0 | 1.0 | 1 |

| | ONLINE_AFFINITAET | ORTSGR_KLS9 | OST_WEST_KZ | PLZ8_ANTG1 | PLZ8_ANTG2 \ |
|---|-------------------|-------------|-------------|------------|--------------|
| 0 | 1.0 | 7.0 | W | 1.0 | 4.0 |
| 1 | 2.0 | 2.0 | W | 3.0 | 2.0 |
| 2 | 3.0 | 8.0 | O | 4.0 | 0.0 |
| 3 | 1.0 | 9.0 | W | 4.0 | 2.0 |
| 4 | 2.0 | 7.0 | W | 2.0 | 3.0 |

| | PLZ8_ANTG3 | PLZ8_ANTG4 | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ \ |
|---|------------|------------|-------------|----------|------------|
| 0 | 2.0 | 1.0 | 5.0 | 2.0 | 3.0 |

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 1 | 1.0 | 0.0 | 1.0 | 4.0 | 3.0 |
| 2 | 0.0 | 1.0 | 1.0 | 4.0 | 3.0 |
| 3 | 1.0 | 0.0 | 1.0 | 4.0 | 3.0 |
| 4 | 0.0 | 2.0 | 4.0 | 1.0 | 2.0 |

| | PRAEGENDE_JUGENDJAHRE | REGIOTYP | RELAT_AB | RETOURTYP_BK_S | RT_KEIN_ANREIZ | \ |
|---|-----------------------|----------|----------|----------------|----------------|-----|
| 0 | | 1 | 5.0 | 5.0 | 5.0 | 2.0 |
| 1 | | 8 | 3.0 | 1.0 | 2.0 | 3.0 |
| 2 | | 2 | 1.0 | 4.0 | 3.0 | 1.0 |
| 3 | | 2 | 1.0 | 3.0 | 5.0 | 2.0 |
| 4 | | 3 | 7.0 | 3.0 | 5.0 | 1.0 |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL | SEMIO_FAM | \ |
|---|-----------------|-----------------|-----------|-----------|-----------|---|
| 0 | 5.0 | 1.0 | 6 | 7 | 2 | |
| 1 | 1.0 | 3.0 | 6 | 6 | 1 | |
| 2 | 5.0 | 2.0 | 3 | 3 | 4 | |
| 3 | 5.0 | 1.0 | 5 | 7 | 1 | |
| 4 | 5.0 | 1.0 | 5 | 4 | 6 | |

| | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT | \ |
|---|------------|------------|------------|------------|-----------|---------------|---|
| 0 | 6 | 6 | 1 | 7 | 1 | 3 | |
| 1 | 5 | 6 | 3 | 6 | 5 | 5 | |
| 2 | 1 | 3 | 4 | 7 | 6 | 1 | |
| 3 | 6 | 7 | 1 | 7 | 1 | 4 | |
| 4 | 2 | 5 | 5 | 6 | 6 | 4 | |

| | SEMIO_RAT | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT | SHOPPER_TYP | \ |
|---|-----------|-----------|-----------|-------------|------------|-------------|---|
| 0 | 2 | 1 | 5 | 1 | 3 | 3 | |
| 1 | 4 | 3 | 1 | 3 | 4 | 2 | |
| 2 | 1 | 5 | 3 | 1 | 7 | 3 | |
| 3 | 2 | 2 | 5 | 3 | 2 | 1 | |
| 4 | 5 | 4 | 6 | 2 | 7 | 1 | |

| | SOHO_KZ | STRUKTURTYP | TITEL_KZ | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG | \ |
|---|---------|-------------|----------|------------|-------------|------------------|---|
| 0 | 0.0 | 3.0 | 0.0 | 2.0 | 4.0 | 0.0 | |
| 1 | 0.0 | 1.0 | 0.0 | 2.0 | 5.0 | 0.0 | |
| 2 | 0.0 | 3.0 | 0.0 | 1.0 | 5.0 | 0.0 | |
| 3 | 0.0 | 3.0 | 0.0 | 3.0 | 5.0 | 0.0 | |
| 4 | 0.0 | 3.0 | 0.0 | 2.0 | 5.0 | 0.0 | |

| | VERDICHTUNGSRAUM | VERS_TYP | VHA | VHN | VK_DHT4A | VK_DISTANZ | VK_ZG11 | \ |
|---|------------------|----------|-----|-----|----------|------------|---------|---|
| 0 | 4.0 | 2 | 1.0 | 2.0 | 5.0 | 2.0 | 1.0 | |
| 1 | 0.0 | 1 | 1.0 | 3.0 | 1.0 | 2.0 | 1.0 | |
| 2 | 10.0 | 1 | 4.0 | 1.0 | 6.0 | 4.0 | 2.0 | |
| 3 | 5.0 | 2 | 1.0 | 4.0 | 8.0 | 11.0 | 11.0 | |
| 4 | 4.0 | 1 | 0.0 | 4.0 | 2.0 | 2.0 | 1.0 | |

| W_KEIT_KIND_HH | WOHNDAUER_2008 | WOHNLAGE | ZABEOTYP | RESPONSE | ANREDE_KZ | \ |
|----------------|----------------|----------|----------|----------|-----------|---|
|----------------|----------------|----------|----------|----------|-----------|---|

| | | | | | | |
|---|-----|-----|-----|---|---|---|
| 0 | 6.0 | 9.0 | 3.0 | 3 | 0 | 2 |
| 1 | 4.0 | 9.0 | 7.0 | 1 | 0 | 2 |
| 2 | NaN | 9.0 | 2.0 | 3 | 0 | 1 |
| 3 | 6.0 | 9.0 | 1.0 | 3 | 0 | 2 |
| 4 | 6.0 | 9.0 | 3.0 | 3 | 0 | 1 |

| | ALTERSKATEGORIE_GROB |
|---|----------------------|
| 0 | 4 |
| 1 | 3 |
| 2 | 4 |
| 3 | 4 |
| 4 | 3 |

```
In [117]: 'Shape', mailout_train.shape
```

```
Out[117]: ('Shape', (42962, 367))
```

```
In [134]: y = mailout_train['RESPONSE']
          X = mailout_train.drop('RESPONSE', axis=1)
          X = clean_demographic_data(X, unknown)
```

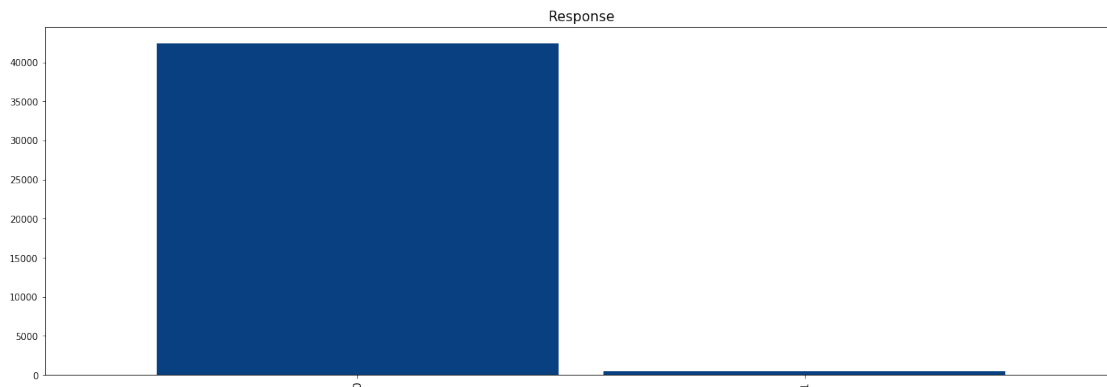
Whether all values in LNR column are unique in azdias DF - True
 Final check of missing values 0

```
In [135]: 'Shape X', X.shape
```

```
Out[135]: ('Shape X', (42962, 354))
```

```
In [136]: plt.rc('xtick', labels=12)
```

```
plt.subplots(figsize=(21,7))
y.value_counts().iloc[:20].plot(kind='bar', width=0.9, cmap='GnBu_r')
plt.title('Response', fontsize=15)
plt.show()
```



```

In [137]: # Split data into train and test sets
          X_train, X_test, y_train, y_test = train_test_split(X, y)

In [147]: X_train.shape, X_test.shape, y_train.shape, y_test.shape

Out[147]: ((32221, 354), (10741, 354), (32221,), (10741,))

In [138]: pipeline = Pipeline([
            ('scale', scaler),
            ('clf', RandomForestClassifier(class_weight='balanced'))
        ])

In [139]: # Train pipeline
          pipeline.fit(X_train, y_train)

Out[139]: Pipeline(memory=None,
                  steps=[('scale', StandardScaler(copy=True, with_mean=True, with_std=True)), ('clf',
                  criterion='gini', max_depth=None, max_features='auto',
                  max_leaf_nodes=None, min_impurity_decrease=0.0,
                  min_impurity_...imators=10, n_jobs=1, oob_score=False, random_state=None,
                  verbose=0, warm_start=False))])

In [160]: pred = pipeline.predict(X_test)
          print(pred.shape, y_test.shape)

          precision_ = precision_score(y_test, pred)
          recall_ = recall_score(y_test, pred)
          f1_ = f1_score(y_test, pred)
          accuracy_ = accuracy_score(y_test, pred)

          precision_, recall_, f1_, accuracy_

(10741,) (10741,)

Out[160]: (0.0096618357487922701, 0.03125, 0.014760147601476014, 0.95028395866306681)

In [161]: y_train.mean(), y_test.mean(), pred.mean()

Out[161]: (0.01253840662921697, 0.011916953728703101, 0.038543897216274089)

In [162]: # Use grid search to find better parameters

          parameters = {
              'clf__n_estimators': [10, 20]
          }

          cv = GridSearchCV(pipeline, param_grid=parameters)

In [163]: cv.fit(X_train, y_train)

```

```
Out[163]: GridSearchCV(cv=None, error_score='raise',
                      estimator=Pipeline(memory=None,
                      steps=[('scale', StandardScaler(copy=True, with_mean=True, with_std=True)), ('clf',
                      criterion='gini', max_depth=None, max_features='auto',
                      max_leaf_nodes=None, min_impurity_decrease=0.0,
                      min_impurity_...imators=10, n_jobs=1, oob_score=False, random_state=None,
                      verbose=0, warm_start=False))]),
                      fit_params=None, iid=True, n_jobs=1,
                      param_grid={'clf__n_estimators': [10, 20]}, pre_dispatch='2*n_jobs',
                      refit=True, return_train_score='warn', scoring=None, verbose=0)
```

```
In [164]: pred_2 = cv.predict(X_test)
          print(pred_2.shape, y_test.shape)

          precision_ = precision_score(y_test, pred_2)
          recall_ = recall_score(y_test, pred_2)
          f1_ = f1_score(y_test, pred_2)
          accuracy_ = accuracy_score(y_test, pred_2)

          precision_, recall_, f1_, accuracy_

(10741,) (10741,)
```

```
Out[164]: (0.0096618357487922701, 0.03125, 0.014760147601476014, 0.95028395866306681)
```

```
In [ ]:
```

```
In [ ]:
```

1.4 Part 3: Kaggle Competition

Now that you've created a model to predict which individuals are most likely to respond to a mailout campaign, it's time to test that model in competition through Kaggle. If you click on the link [here](#), you'll be taken to the competition page where, if you have a Kaggle account, you can enter. If you're one of the top performers, you may have the chance to be contacted by a hiring manager from Arvato or Bertelsmann for an interview!

Your entry to the competition should be a CSV file with two columns. The first column should be a copy of "LNR", which acts as an ID number for each individual in the "TEST" partition. The second column, "RESPONSE", should be some measure of how likely each individual became a customer – this might not be a straightforward probability. As you should have found in Part 2, there is a large output class imbalance, where most individuals did not respond to the mailout. Thus, predicting individual classes and using accuracy does not seem to be an appropriate performance evaluation method. Instead, the competition will be using AUC to evaluate performance. The exact values of the "RESPONSE" column do not matter as much: only that the higher values try to capture as many of the actual customers as possible, early in the ROC curve sweep.

```
In [165]: mailout_test = pd.read_csv('../data/Term2/capstone/arvato_data/Udacity_MAILOUT_0520
          mailout_test.head()
```

```
/opt/conda/lib/python3.6/site-packages/IPython/core/interactiveshell.py:2785: DtypeWarning: Column
interactivity=interactivity, compiler=compiler, result=result)
```

```
Out[165]:
```

| | LNR | AGER_TYP | AKT_DAT_KL | ALTER_HH | ALTER_KIND1 | ALTER_KIND2 | \ |
|---|------|----------|------------|----------|-------------|-------------|---|
| 0 | 1754 | 2 | 1.0 | 7.0 | NaN | NaN | |
| 1 | 1770 | -1 | 1.0 | 0.0 | NaN | NaN | |
| 2 | 1465 | 2 | 9.0 | 16.0 | NaN | NaN | |
| 3 | 1470 | -1 | 7.0 | 0.0 | NaN | NaN | |
| 4 | 1478 | 1 | 1.0 | 21.0 | NaN | NaN | |

| | ALTER_KIND3 | ALTER_KIND4 | ALTERSKATEGORIE_FEIN | ANZ_HAUSHALTE_AKTIV | \ |
|---|-------------|-------------|----------------------|---------------------|---|
| 0 | NaN | NaN | 6.0 | 2.0 | |
| 1 | NaN | NaN | 0.0 | 20.0 | |
| 2 | NaN | NaN | 11.0 | 2.0 | |
| 3 | NaN | NaN | 0.0 | 1.0 | |
| 4 | NaN | NaN | 13.0 | 1.0 | |

| | ANZ_HH_TITEL | ANZ_KINDER | ANZ_PERSONEN | ANZ_STATISTISCHE_HAUSHALTE | \ |
|---|--------------|------------|--------------|----------------------------|---|
| 0 | 0.0 | 0.0 | 2.0 | 2.0 | |
| 1 | 0.0 | 0.0 | 1.0 | 21.0 | |
| 2 | 0.0 | 0.0 | 4.0 | 2.0 | |
| 3 | 0.0 | 0.0 | 0.0 | 1.0 | |
| 4 | 0.0 | 0.0 | 4.0 | 1.0 | |

| | ANZ_TITEL | ARBEIT | BALLRAUM | CAMEO_DEU_2015 | CAMEO_DEUG_2015 | CAMEO_INTL_2015 | \ |
|---|-----------|--------|----------|----------------|-----------------|-----------------|---|
| 0 | 0.0 | 3.0 | 6.0 | 2B | 2 | 13 | |
| 1 | 0.0 | 4.0 | 7.0 | 5A | 5 | 31 | |
| 2 | 0.0 | 4.0 | 1.0 | 7A | 7 | 41 | |
| 3 | 0.0 | 4.0 | 1.0 | 2B | 2 | 13 | |
| 4 | 0.0 | 3.0 | 6.0 | 5A | 5 | 31 | |

| | CJT_GESAMTTYP | CJT_KATALOGNUTZER | CJT_TYP_1 | CJT_TYP_2 | CJT_TYP_3 | \ |
|---|---------------|-------------------|-----------|-----------|-----------|-----|
| 0 | 5.0 | | 5.0 | 1.0 | 2.0 | 5.0 |
| 1 | 1.0 | | 5.0 | 2.0 | 2.0 | 5.0 |
| 2 | 2.0 | | 5.0 | 2.0 | 2.0 | 5.0 |
| 3 | 4.0 | | 5.0 | 2.0 | 1.0 | 5.0 |
| 4 | 6.0 | | 1.0 | 3.0 | 2.0 | 4.0 |

| | CJT_TYP_4 | CJT_TYP_5 | CJT_TYP_6 | D19_BANKEN_ANZ_12 | D19_BANKEN_ANZ_24 | \ |
|---|-----------|-----------|-----------|-------------------|-------------------|---|
| 0 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 1 | 5.0 | 4.0 | 5.0 | 0 | 0 | |
| 2 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 3 | 5.0 | 5.0 | 5.0 | 0 | 0 | |
| 4 | 4.0 | 4.0 | 3.0 | 3 | 4 | |

| | D19_BANKEN_DATUM | D19_BANKEN_DIREKT | D19_BANKEN_GROSS | D19_BANKEN_LOKAL | \ |
|---|------------------|-------------------|------------------|------------------|---|
| 0 | 10 | 0 | 0 | 0 | |

| | | | | |
|---|----|---|---|---|
| 1 | 10 | 0 | 0 | 0 |
| 2 | 10 | 0 | 0 | 0 |
| 3 | 10 | 0 | 0 | 0 |
| 4 | 2 | 2 | 2 | 0 |

| | D19_BANKEN_OFFLINE_DATUM | D19_BANKEN_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 10 | 10 | |
| 1 | 10 | 10 | |
| 2 | 10 | 10 | |
| 3 | 10 | 10 | |
| 4 | 8 | 2 | |

| | D19_BANKEN_ONLINE_QUOTE_12 | D19_BANKEN_REST | D19_BEKLEIDUNG_GEH | \ |
|---|----------------------------|-----------------|--------------------|---|
| 0 | 0.0 | 0 | 3 | |
| 1 | 0.0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 10.0 | 6 | 6 | |

| | D19_BEKLEIDUNG_REST | D19_BILDUNG | D19_BIO_OEKO | D19_BUCH_CD | \ |
|---|---------------------|-------------|--------------|-------------|---|
| 0 | 6 | 0 | 0 | 0 | |
| 1 | 6 | 0 | 0 | 7 | |
| 2 | 0 | 0 | 0 | 6 | |
| 3 | 0 | 0 | 0 | 0 | |
| 4 | 6 | 5 | 0 | 3 | |

| | D19_DIGIT_SERV | D19_DROGERIEARTIKEL | D19_ENERGIE | D19_FREIZEIT | D19_GARTEN | \ |
|---|----------------|---------------------|-------------|--------------|------------|---|
| 0 | 0 | 6 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 6 | |
| 4 | 0 | 2 | 0 | 5 | 5 | |

| | D19_GESAMT_ANZ_12 | D19_GESAMT_ANZ_24 | D19_GESAMT_DATUM | \ |
|---|-------------------|-------------------|------------------|---|
| 0 | 1 | 2 | 4 | |
| 1 | 0 | 0 | 9 | |
| 2 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 8 | |
| 4 | 4 | 6 | 2 | |

| | D19_GESAMT_OFFLINE_DATUM | D19_GESAMT_ONLINE_DATUM | \ |
|---|--------------------------|-------------------------|---|
| 0 | 8 | 4 | |
| 1 | 9 | 10 | |
| 2 | 10 | 10 | |
| 3 | 8 | 8 | |
| 4 | 3 | 2 | |

| | D19_GESAMT_ONLINE_QUOTE_12 | D19_HANDWERK | D19_HAUS_DEKO | D19_KINDERARTIKEL | \ |
|--|----------------------------|--------------|---------------|-------------------|---|
|--|----------------------------|--------------|---------------|-------------------|---|

| | | | | |
|---|------|---|---|---|
| 0 | 10.0 | 0 | 6 | 0 |
| 1 | 0.0 | 0 | 6 | 0 |
| 2 | 0.0 | 7 | 0 | 0 |
| 3 | 0.0 | 0 | 0 | 0 |
| 4 | 7.0 | 7 | 2 | 5 |

| | D19_KONSUMTYP | D19_KONSUMTYP_MAX | D19_KOSMETIK | D19_LEBENSMITTEL | \ |
|---|---------------|-------------------|--------------|------------------|---|
| 0 | 3.0 | 2 | 6 | 0 | |
| 1 | 3.0 | 2 | 6 | 0 | |
| 2 | 9.0 | 8 | 0 | 0 | |
| 3 | 9.0 | 8 | 6 | 6 | |
| 4 | 1.0 | 1 | 7 | 0 | |

| | D19_LETZTER_KAUF_BRANCHE | D19_LOTTO | D19_NAHRUNGSEGAENZUNG | D19_RATGEBER | \ |
|---|--------------------------|-----------|-----------------------|--------------|---|
| 0 | D19_BEKLEIDUNG_GEH | 0.0 | 0 | 0 | |
| 1 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 2 | D19_UNBEKANNT | 0.0 | 0 | 0 | |
| 3 | D19_LEBENSMITTEL | 0.0 | 6 | 0 | |
| 4 | D19_BANKEN_GROSS | 7.0 | 0 | 6 | |

| | D19_REISEN | D19_SAMMELARTIKEL | D19_SCHUHE | D19_SONSTIGE | D19_SOZIALES | \ |
|---|------------|-------------------|------------|--------------|--------------|---|
| 0 | 0 | 0 | 0 | 6 | 1.0 | |
| 1 | 0 | 0 | 0 | 0 | 1.0 | |
| 2 | 6 | 6 | 0 | 6 | 0.0 | |
| 3 | 0 | 0 | 0 | 6 | 0.0 | |
| 4 | 6 | 6 | 0 | 6 | 0.0 | |

| | D19_TECHNIK | D19_TELKO_ANZ_12 | D19_TELKO_ANZ_24 | D19_TELKO_DATUM | \ |
|---|-------------|------------------|------------------|-----------------|---|
| 0 | 0 | 0 | 0 | 10 | |
| 1 | 0 | 0 | 0 | 10 | |
| 2 | 6 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 0 | 10 | |
| 4 | 3 | 0 | 1 | 7 | |

| | D19_TELKO_MOBILE | D19_TELKO_OFFLINE_DATUM | D19_TELKO_ONLINE_DATUM | \ |
|---|------------------|-------------------------|------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 0 | 10 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 5 | 8 | 9 | |

| | D19_TELKO_ONLINE_QUOTE_12 | D19_TELKO_REST | D19_TIERARTIKEL | \ |
|---|---------------------------|----------------|-----------------|---|
| 0 | 0.0 | 0 | 0 | |
| 1 | 0.0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 0.0 | 6 | 7 | |

| | D19_VERSAND_ANZ_12 | D19_VERSAND_ANZ_24 | D19_VERSAND_DATUM | \ |
|---|--------------------|--------------------|-------------------|---|
| 0 | 1 | 2 | 4 | |
| 1 | 0 | 0 | 9 | |
| 2 | 0 | 0 | 10 | |
| 3 | 0 | 0 | 8 | |
| 4 | 4 | 5 | 3 | |

| | D19_VERSAND_OFFLINE_DATUM | D19_VERSAND_ONLINE_DATUM | \ |
|---|---------------------------|--------------------------|---|
| 0 | 8 | 4 | |
| 1 | 9 | 10 | |
| 2 | 10 | 10 | |
| 3 | 8 | 8 | |
| 4 | 3 | 4 | |

| | D19_VERSAND_ONLINE_QUOTE_12 | D19_VERSAND_REST | D19_VERSI_ANZ_12 | \ |
|---|-----------------------------|------------------|------------------|---|
| 0 | 10.0 | 0 | 0 | |
| 1 | 0.0 | 0 | 0 | |
| 2 | 0.0 | 0 | 0 | |
| 3 | 0.0 | 0 | 0 | |
| 4 | 5.0 | 6 | 0 | |

| | D19_VERSI_ANZ_24 | D19_VERSI_DATUM | D19_VERSI_OFFLINE_DATUM | \ |
|---|------------------|-----------------|-------------------------|---|
| 0 | 0 | 10 | 10 | |
| 1 | 0 | 10 | 10 | |
| 2 | 0 | 10 | 10 | |
| 3 | 0 | 10 | 10 | |
| 4 | 1 | 6 | 10 | |

| | D19_VERSI_ONLINE_DATUM | D19_VERSI_ONLINE_QUOTE_12 | D19_VERSICHERUNGEN | \ |
|---|------------------------|---------------------------|--------------------|---|
| 0 | 10 | 0.0 | 0 | |
| 1 | 10 | 0.0 | 0 | |
| 2 | 10 | 0.0 | 0 | |
| 3 | 10 | 0.0 | 0 | |
| 4 | 9 | 0.0 | 5 | |

| | D19_VOLLSORTIMENT | D19_WEIN_FEINKOST | DSL_FLAG | EINGEFUEGT_AM | \ |
|---|-------------------|-------------------|----------|---------------------|---|
| 0 | 5 | 0 | 1.0 | 1993-11-03 00:00:00 | |
| 1 | 6 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 2 | 0 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 3 | 0 | 0 | 1.0 | 1992-02-10 00:00:00 | |
| 4 | 5 | 6 | 1.0 | 1992-02-10 00:00:00 | |

| | EINGEZOGENAM_HH_JAHR | EWDICHTE | EXTSEL992 | FINANZ_ANLEGER | \ |
|---|----------------------|----------|-----------|----------------|---|
| 0 | 1994.0 | 5.0 | 34.0 | 1 | |
| 1 | 1998.0 | 4.0 | 47.0 | 1 | |
| 2 | 1994.0 | 4.0 | 35.0 | 1 | |
| 3 | 1994.0 | 4.0 | 31.0 | 2 | |
| 4 | 2000.0 | 1.0 | 36.0 | 3 | |

| | FINANZ_HAUSBAUER | FINANZ_MINIMALIST | FINANZ_SPARER | FINANZ_UNAUFFAELLIGER | \ |
|---|------------------|-------------------|---------------|-----------------------|---|
| 0 | 3 | 5 | 1 | 2 | |
| 1 | 5 | 3 | 1 | 1 | |
| 2 | 4 | 4 | 1 | 2 | |
| 3 | 4 | 4 | 1 | 1 | |
| 4 | 2 | 5 | 1 | 1 | |

| | FINANZ_VORSORGER | FINANZTYP | FIRMENDICHTE | GEBAEUDETYP | GEBAEUDETYP_RASTER | \ |
|---|------------------|-----------|--------------|-------------|--------------------|---|
| 0 | 5 | 5 | 4.0 | 1.0 | 3.0 | |
| 1 | 5 | 6 | 4.0 | 8.0 | 4.0 | |
| 2 | 5 | 2 | 3.0 | 1.0 | 4.0 | |
| 3 | 4 | 6 | 4.0 | 1.0 | 4.0 | |
| 4 | 5 | 6 | 5.0 | 1.0 | 5.0 | |

| | GEBURTSJAHR | GEMEINDETYP | GFK_URLAUBERTYP | GREEN_AVANTGARDE | HEALTH_TYP | \ |
|---|-------------|-------------|-----------------|------------------|------------|---|
| 0 | 0 | 30.0 | 10.0 | 1 | 1 | |
| 1 | 0 | 22.0 | 10.0 | 0 | 2 | |
| 2 | 1946 | 21.0 | 5.0 | 1 | 2 | |
| 3 | 0 | 11.0 | 6.0 | 0 | 2 | |
| 4 | 1956 | 40.0 | 10.0 | 0 | 2 | |

| | HH_DELTA_FLAG | HH_EINKOMMEN_SCORE | INNENSTADT | KBA05_ALTER1 | KBA05_ALTER2 | \ |
|---|---------------|--------------------|------------|--------------|--------------|---|
| 0 | 0.0 | 4.0 | 4.0 | 1.0 | 1.0 | |
| 1 | 0.0 | 5.0 | 8.0 | 1.0 | 5.0 | |
| 2 | 1.0 | 5.0 | 3.0 | 0.0 | 4.0 | |
| 3 | NaN | 3.0 | 2.0 | 3.0 | 2.0 | |
| 4 | 1.0 | 4.0 | 8.0 | 1.0 | 3.0 | |

| | KBA05_ALTER3 | KBA05_ALTER4 | KBA05_ANHANG | KBA05_ANTG1 | KBA05_ANTG2 | \ |
|---|--------------|--------------|--------------|-------------|-------------|---|
| 0 | 3.0 | 5.0 | 1.0 | 3.0 | 1.0 | |
| 1 | 3.0 | 1.0 | 0.0 | 0.0 | 2.0 | |
| 2 | 2.0 | 4.0 | 1.0 | 1.0 | 3.0 | |
| 3 | 3.0 | 4.0 | 3.0 | 4.0 | 1.0 | |
| 4 | 2.0 | 4.0 | 3.0 | 3.0 | 0.0 | |

| | KBA05_ANTG3 | KBA05_ANTG4 | KBA05_AUTOQUOT | KBA05_BAUMAX | KBA05_CCM1 | \ |
|---|-------------|-------------|----------------|--------------|------------|---|
| 0 | 0.0 | 0.0 | 4.0 | 1.0 | 2.0 | |
| 1 | 1.0 | 1.0 | 3.0 | 0.0 | 3.0 | |
| 2 | 1.0 | 0.0 | 2.0 | 5.0 | 2.0 | |
| 3 | 0.0 | 0.0 | 3.0 | 1.0 | 5.0 | |
| 4 | 0.0 | 0.0 | 5.0 | 1.0 | 2.0 | |

| | KBA05_CCM2 | KBA05_CCM3 | KBA05_CCM4 | KBA05_DIESEL | KBA05_FRAU | KBA05_GBZ | \ |
|---|------------|------------|------------|--------------|------------|-----------|---|
| 0 | 2.0 | 3.0 | 3.0 | 2.0 | 5.0 | 5.0 | |
| 1 | 1.0 | 5.0 | 2.0 | 2.0 | 2.0 | 3.0 | |
| 2 | 3.0 | 4.0 | 3.0 | 3.0 | 2.0 | 2.0 | |
| 3 | 2.0 | 3.0 | 0.0 | 1.0 | 4.0 | 4.0 | |

| | | | | | | |
|---|-----|-----|-----|-----|-----|-----|
| 4 | 2.0 | 5.0 | 1.0 | 4.0 | 3.0 | 5.0 |
|---|-----|-----|-----|-----|-----|-----|

| | KBA05_HERST1 | KBA05_HERST2 | KBA05_HERST3 | KBA05_HERST4 | KBA05_HERST5 | \ |
|---|--------------|--------------|--------------|--------------|--------------|---|
| 0 | 4.0 | 3.0 | 1.0 | 3.0 | 2.0 | |
| 1 | 2.0 | 5.0 | 2.0 | 1.0 | 3.0 | |
| 2 | 3.0 | 3.0 | 5.0 | 2.0 | 0.0 | |
| 3 | 2.0 | 1.0 | 2.0 | 5.0 | 4.0 | |
| 4 | 4.0 | 4.0 | 2.0 | 2.0 | 2.0 | |

| | KBA05_HERSTTEMP | KBA05_KRSAQUOT | KBA05_KRSHERST1 | KBA05_KRSHERST2 | \ |
|---|-----------------|----------------|-----------------|-----------------|---|
| 0 | 2.0 | 3.0 | 4.0 | 3.0 | |
| 1 | 2.0 | 2.0 | 3.0 | 5.0 | |
| 2 | 4.0 | 2.0 | 4.0 | 3.0 | |
| 3 | 3.0 | 4.0 | 3.0 | 2.0 | |
| 4 | 2.0 | 5.0 | 4.0 | 4.0 | |

| | KBA05_KRSHERST3 | KBA05_KRSKLEIN | KBA05_KRSOBER | KBA05_KRSVAN | KBA05_KRSZUL | \ |
|---|-----------------|----------------|---------------|--------------|--------------|---|
| 0 | 1.0 | 1.0 | 3.0 | 2.0 | 2.0 | |
| 1 | 1.0 | 2.0 | 2.0 | 1.0 | 1.0 | |
| 2 | 5.0 | 1.0 | 3.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 2.0 | 2.0 | 2.0 | |
| 4 | 2.0 | 2.0 | 3.0 | 2.0 | 2.0 | |

| | KBA05_KW1 | KBA05_KW2 | KBA05_KW3 | KBA05_MAXAH | KBA05_MAXBJ | KBA05_MAXHERST | \ |
|---|-----------|-----------|-----------|-------------|-------------|----------------|---|
| 0 | 2.0 | 3.0 | 4.0 | 5.0 | 1.0 | 2.0 | |
| 1 | 2.0 | 4.0 | 1.0 | 3.0 | 2.0 | 2.0 | |
| 2 | 1.0 | 4.0 | 2.0 | 2.0 | 4.0 | 3.0 | |
| 3 | 4.0 | 2.0 | 1.0 | 5.0 | 1.0 | 4.0 | |
| 4 | 4.0 | 2.0 | 1.0 | 2.0 | 1.0 | 2.0 | |

| | KBA05_MAXSEG | KBA05_MAXVORB | KBA05_MOD1 | KBA05_MOD2 | KBA05_MOD3 | \ |
|---|--------------|---------------|------------|------------|------------|---|
| 0 | 4.0 | 1.0 | 4.0 | 3.0 | 2.0 | |
| 1 | 3.0 | 3.0 | 1.0 | 4.0 | 4.0 | |
| 2 | 4.0 | 1.0 | 4.0 | 2.0 | 1.0 | |
| 3 | 1.0 | 2.0 | 0.0 | 3.0 | 2.0 | |
| 4 | 2.0 | 3.0 | 3.0 | 3.0 | 2.0 | |

| | KBA05_MOD4 | KBA05_MOD8 | KBA05_MODTEMP | KBA05_MOTOR | KBA05_MOTRAD | \ |
|---|------------|------------|---------------|-------------|--------------|---|
| 0 | 1.0 | 1.0 | 1.0 | 4.0 | 1.0 | |
| 1 | 2.0 | 0.0 | 2.0 | 3.0 | 1.0 | |
| 2 | 2.0 | 3.0 | 1.0 | 3.0 | 0.0 | |
| 3 | 1.0 | 2.0 | 4.0 | 1.0 | 3.0 | |
| 4 | 3.0 | 2.0 | 4.0 | 3.0 | 3.0 | |

| | KBA05_SEG1 | KBA05_SEG10 | KBA05_SEG2 | KBA05_SEG3 | KBA05_SEG4 | KBA05_SEG5 | \ |
|---|------------|-------------|------------|------------|------------|------------|---|
| 0 | 1.0 | 3.0 | 2.0 | 2.0 | 2.0 | 4.0 | |
| 1 | 1.0 | 1.0 | 3.0 | 3.0 | 5.0 | 1.0 | |
| 2 | 2.0 | 4.0 | 2.0 | 1.0 | 2.0 | 4.0 | |

| | | | | | | |
|---|-----|-----|-----|-----|-----|-----|
| 3 | 2.0 | 1.0 | 5.0 | 2.0 | 3.0 | 1.0 |
| 4 | 1.0 | 2.0 | 2.0 | 3.0 | 3.0 | 4.0 |

| | KBA05_SEG6 | KBA05_SEG7 | KBA05_SEG8 | KBA05_SEG9 | KBA05_VORB0 | KBA05_VORB1 | \ |
|---|------------|------------|------------|------------|-------------|-------------|---|
| 0 | 1.0 | 1.0 | 3.0 | 1.0 | 4.0 | 3.0 | |
| 1 | 0.0 | 0.0 | 1.0 | 0.0 | 2.0 | 4.0 | |
| 2 | 0.0 | 3.0 | 0.0 | 3.0 | 5.0 | 2.0 | |
| 3 | 0.0 | 1.0 | 0.0 | 0.0 | 2.0 | 3.0 | |
| 4 | 1.0 | 1.0 | 1.0 | 1.0 | 3.0 | 3.0 | |

| | KBA05_VORB2 | KBA05_ZUL1 | KBA05_ZUL2 | KBA05_ZUL3 | KBA05_ZUL4 | \ |
|---|-------------|------------|------------|------------|------------|---|
| 0 | 2.0 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 1 | 3.0 | 3.0 | 4.0 | 4.0 | 0.0 | |
| 2 | 0.0 | 1.0 | 3.0 | 5.0 | 5.0 | |
| 3 | 3.0 | 3.0 | 2.0 | 3.0 | 3.0 | |
| 4 | 4.0 | 4.0 | 2.0 | 2.0 | 1.0 | |

| | KBA13_ALTERHALTER_30 | KBA13_ALTERHALTER_45 | KBA13_ALTERHALTER_60 | \ |
|---|----------------------|----------------------|----------------------|---|
| 0 | 3.0 | 4.0 | 3.0 | |
| 1 | 4.0 | 3.0 | 3.0 | |
| 2 | 2.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 1.0 | 3.0 | |
| 4 | 3.0 | 5.0 | 4.0 | |

| | KBA13_ALTERHALTER_61 | KBA13_ANTG1 | KBA13_ANTG2 | KBA13_ANTG3 | KBA13_ANTG4 | \ |
|---|----------------------|-------------|-------------|-------------|-------------|---|
| 0 | 3.0 | 2.0 | 4.0 | 2.0 | 1.0 | |
| 1 | 3.0 | 2.0 | 3.0 | 3.0 | 1.0 | |
| 2 | 4.0 | 4.0 | 2.0 | 1.0 | 0.0 | |
| 3 | 5.0 | 3.0 | 1.0 | 0.0 | 0.0 | |
| 4 | 2.0 | 4.0 | 2.0 | 0.0 | 0.0 | |

| | KBA13_ANZAHL_PKW | KBA13_AUDI | KBA13_AUTOQUOTE | KBA13_BAUMAX | KBA13_BJ_1999 | \ |
|---|------------------|------------|-----------------|--------------|---------------|---|
| 0 | 412.0 | 3.0 | 2.0 | 2.0 | 2.0 | |
| 1 | 935.0 | 5.0 | 2.0 | 3.0 | 5.0 | |
| 2 | 693.0 | 2.0 | 3.0 | 1.0 | 1.0 | |
| 3 | 326.0 | 2.0 | 3.0 | 1.0 | 4.0 | |
| 4 | 134.0 | 3.0 | 4.0 | 1.0 | 5.0 | |

| | KBA13_BJ_2000 | KBA13_BJ_2004 | KBA13_BJ_2006 | KBA13_BJ_2008 | KBA13_BJ_2009 | \ |
|---|---------------|---------------|---------------|---------------|---------------|---|
| 0 | 2.0 | 4.0 | 4.0 | 3.0 | 4.0 | |
| 1 | 4.0 | 3.0 | 2.0 | 2.0 | 2.0 | |
| 2 | 1.0 | 3.0 | 4.0 | 5.0 | 4.0 | |
| 3 | 4.0 | 2.0 | 2.0 | 0.0 | 2.0 | |
| 4 | 4.0 | 4.0 | 3.0 | 0.0 | 1.0 | |

| | KBA13_BMW | KBA13_CCM_0_1400 | KBA13_CCM_1000 | KBA13_CCM_1200 | \ |
|---|-----------|------------------|----------------|----------------|---|
| 0 | 5.0 | 0.0 | 2.0 | 1.0 | |
| 1 | 4.0 | 0.0 | 3.0 | 0.0 | |

| | | | | |
|---|-----|-----|-----|-----|
| 2 | 3.0 | 3.0 | 0.0 | 4.0 |
| 3 | 3.0 | 3.0 | 2.0 | 4.0 |
| 4 | 1.0 | 0.0 | 1.0 | 1.0 |

| | KBA13_CCM_1400 | KBA13_CCM_1401_2500 | KBA13_CCM_1500 | KBA13_CCM_1600 | \ |
|---|----------------|---------------------|----------------|----------------|---|
| 0 | 2.0 | 3.0 | 1.0 | 2.0 | |
| 1 | 2.0 | 3.0 | 4.0 | 3.0 | |
| 2 | 4.0 | 2.0 | 4.0 | 3.0 | |
| 3 | 4.0 | 3.0 | 4.0 | 5.0 | |
| 4 | 3.0 | 5.0 | 5.0 | 1.0 | |

| | KBA13_CCM_1800 | KBA13_CCM_2000 | KBA13_CCM_2500 | KBA13_CCM_2501 | \ |
|---|----------------|----------------|----------------|----------------|---|
| 0 | 4.0 | 4.0 | 3.0 | 5.0 | |
| 1 | 0.0 | 5.0 | 3.0 | 3.0 | |
| 2 | 2.0 | 3.0 | 2.0 | 3.0 | |
| 3 | 1.0 | 2.0 | 2.0 | 2.0 | |
| 4 | 0.0 | 5.0 | 2.0 | 1.0 | |

| | KBA13_CCM_3000 | KBA13_CCM_3001 | KBA13_FAB_ASIEN | KBA13_FAB_SONSTIGE | \ |
|---|----------------|----------------|-----------------|--------------------|---|
| 0 | 4.0 | 5.0 | 2.0 | 4.0 | |
| 1 | 3.0 | 1.0 | 3.0 | 2.0 | |
| 2 | 3.0 | 4.0 | 2.0 | 3.0 | |
| 3 | 2.0 | 1.0 | 4.0 | 4.0 | |
| 4 | 1.0 | 1.0 | 4.0 | 2.0 | |

| | KBA13_FIAT | KBA13_FORD | KBA13_GBZ | KBA13_HALTER_20 | KBA13_HALTER_25 | \ |
|---|------------|------------|-----------|-----------------|-----------------|---|
| 0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | |
| 1 | 3.0 | 1.0 | 3.0 | 4.0 | 4.0 | |
| 2 | 3.0 | 3.0 | 5.0 | 2.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 3.0 | 3.0 | 3.0 | |
| 4 | 2.0 | 1.0 | 2.0 | 4.0 | 3.0 | |

| | KBA13_HALTER_30 | KBA13_HALTER_35 | KBA13_HALTER_40 | KBA13_HALTER_45 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | 3.0 | 3.0 | 4.0 | 4.0 | |
| 1 | 4.0 | 3.0 | 3.0 | 2.0 | |
| 2 | 2.0 | 3.0 | 3.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 2.0 | 1.0 | |
| 4 | 3.0 | 4.0 | 5.0 | 4.0 | |

| | KBA13_HALTER_50 | KBA13_HALTER_55 | KBA13_HALTER_60 | KBA13_HALTER_65 | \ |
|---|-----------------|-----------------|-----------------|-----------------|---|
| 0 | 3.0 | 3.0 | 3.0 | 3.0 | |
| 1 | 3.0 | 3.0 | 3.0 | 3.0 | |
| 2 | 3.0 | 3.0 | 3.0 | 4.0 | |
| 3 | 1.0 | 2.0 | 5.0 | 5.0 | |
| 4 | 4.0 | 4.0 | 3.0 | 2.0 | |

| | KBA13_HALTER_66 | KBA13_HERST_ASIEN | KBA13_HERST_AUDI_VW | \ |
|---|-----------------|-------------------|---------------------|---|
| 0 | 2.0 | 1.0 | 3.0 | |

| | | | |
|---|-----|-----|-----|
| 1 | 3.0 | 3.0 | 3.0 |
| 2 | 4.0 | 3.0 | 3.0 |
| 3 | 5.0 | 3.0 | 3.0 |
| 4 | 2.0 | 3.0 | 5.0 |

| | KBA13_HERST_BMW_BENZ | KBA13_HERST_EUROPA | KBA13_HERST_FORD_OPEL | \ |
|---|----------------------|--------------------|-----------------------|---|
| 0 | 4.0 | 3.0 | 3.0 | |
| 1 | 3.0 | 5.0 | 1.0 | |
| 2 | 4.0 | 3.0 | 3.0 | |
| 3 | 3.0 | 3.0 | 3.0 | |
| 4 | 4.0 | 1.0 | 1.0 | |

| | KBA13_HERST_SONST | KBA13_HHZ | KBA13_KMH_0_140 | KBA13_KMH_110 | \ |
|---|-------------------|-----------|-----------------|---------------|---|
| 0 | 4.0 | 3.0 | 1.0 | 1.0 | |
| 1 | 2.0 | 5.0 | 4.0 | 1.0 | |
| 2 | 3.0 | 4.0 | 3.0 | 1.0 | |
| 3 | 4.0 | 2.0 | 3.0 | 2.0 | |
| 4 | 2.0 | 1.0 | 1.0 | 1.0 | |

| | KBA13_KMH_140 | KBA13_KMH_140_210 | KBA13_KMH_180 | KBA13_KMH_210 | \ |
|---|---------------|-------------------|---------------|---------------|---|
| 0 | 1.0 | 1.0 | 1.0 | 4.0 | |
| 1 | 4.0 | 3.0 | 2.0 | 4.0 | |
| 2 | 4.0 | 3.0 | 4.0 | 2.0 | |
| 3 | 1.0 | 4.0 | 5.0 | 2.0 | |
| 4 | 1.0 | 4.0 | 3.0 | 4.0 | |

| | KBA13_KMH_211 | KBA13_KMH_250 | KBA13_KMH_251 | KBA13_KRSAQUOT | \ |
|---|---------------|---------------|---------------|----------------|---|
| 0 | 5.0 | 5.0 | 1.0 | 2.0 | |
| 1 | 3.0 | 3.0 | 1.0 | 1.0 | |
| 2 | 0.0 | 0.0 | 1.0 | 4.0 | |
| 3 | 2.0 | 2.0 | 1.0 | 5.0 | |
| 4 | 1.0 | 1.0 | 1.0 | 3.0 | |

| | KBA13_KRSHERST_AUDI_VW | KBA13_KRSHERST_BMW_BENZ | KBA13_KRSHERST_FORD_OPEL | \ |
|---|------------------------|-------------------------|--------------------------|---|
| 0 | 4.0 | 5.0 | 2.0 | |
| 1 | 3.0 | 4.0 | 1.0 | |
| 2 | 3.0 | 4.0 | 3.0 | |
| 3 | 5.0 | 4.0 | 3.0 | |
| 4 | 5.0 | 4.0 | 1.0 | |

| | KBA13_KRSSEG_KLEIN | KBA13_KRSSEG_OBER | KBA13_KRSSEG_VAN | KBA13_KRSZUL_NEU | \ |
|---|--------------------|-------------------|------------------|------------------|---|
| 0 | 2.0 | 2.0 | 1.0 | 3.0 | |
| 1 | 2.0 | 2.0 | 2.0 | 1.0 | |
| 2 | 2.0 | 2.0 | 2.0 | 3.0 | |
| 3 | 2.0 | 3.0 | 2.0 | 2.0 | |
| 4 | 2.0 | 2.0 | 2.0 | 0.0 | |

| | KBA13_KW_0_60 | KBA13_KW_110 | KBA13_KW_120 | KBA13_KW_121 | KBA13_KW_30 | \ |
|--|---------------|--------------|--------------|--------------|-------------|---|
|--|---------------|--------------|--------------|--------------|-------------|---|

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 0 | 1.0 | 5.0 | 4.0 | 5.0 | 1.0 |
| 1 | 3.0 | 4.0 | 3.0 | 3.0 | 1.0 |
| 2 | 4.0 | 3.0 | 0.0 | 3.0 | 1.0 |
| 3 | 3.0 | 2.0 | 3.0 | 0.0 | 2.0 |
| 4 | 4.0 | 0.0 | 1.0 | 2.0 | 1.0 |

| | KBA13_KW_40 | KBA13_KW_50 | KBA13_KW_60 | KBA13_KW_61_120 | KBA13_KW_70 \ |
|---|-------------|-------------|-------------|-----------------|---------------|
| 0 | 1.0 | 2.0 | 1.0 | 4.0 | 3.0 |
| 1 | 3.0 | 2.0 | 3.0 | 4.0 | 0.0 |
| 2 | 3.0 | 3.0 | 3.0 | 2.0 | 4.0 |
| 3 | 2.0 | 3.0 | 3.0 | 3.0 | 5.0 |
| 4 | 2.0 | 5.0 | 4.0 | 2.0 | 1.0 |

| | KBA13_KW_80 | KBA13_KW_90 | KBA13_MAZDA | KBA13_MERCEDES | KBA13_MOTOR \ |
|---|-------------|-------------|-------------|----------------|---------------|
| 0 | 2.0 | 0.0 | 3.0 | 3.0 | 2.0 |
| 1 | 2.0 | 3.0 | 3.0 | 3.0 | 3.0 |
| 2 | 1.0 | 2.0 | 3.0 | 4.0 | 4.0 |
| 3 | 2.0 | 1.0 | 3.0 | 2.0 | 3.0 |
| 4 | 0.0 | 5.0 | 3.0 | 5.0 | 2.0 |

| | KBA13_NISSAN | KBA13_OPEL | KBA13_PEUGEOT | KBA13_RENAULT \ |
|---|--------------|------------|---------------|-----------------|
| 0 | 3.0 | 3.0 | 3.0 | 3.0 |
| 1 | 3.0 | 2.0 | 3.0 | 5.0 |
| 2 | 3.0 | 3.0 | 3.0 | 3.0 |
| 3 | 3.0 | 3.0 | 5.0 | 3.0 |
| 4 | 2.0 | 2.0 | 1.0 | 1.0 |

| | KBA13_SEG_GELAENDEWAGEN | KBA13_SEG_GROSSRAUMVANS | KBA13_SEG_KLEINST \ |
|---|-------------------------|-------------------------|---------------------|
| 0 | 5.0 | 2.0 | 3.0 |
| 1 | 2.0 | 4.0 | 2.0 |
| 2 | 2.0 | 3.0 | 4.0 |
| 3 | 1.0 | 2.0 | 4.0 |
| 4 | 4.0 | 4.0 | 1.0 |

| | KBA13_SEG_KLEINWAGEN | KBA13_SEG_KOMPAKTKLASSE | KBA13_SEG_MINIVANS \ |
|---|----------------------|-------------------------|----------------------|
| 0 | 3.0 | 3.0 | 2.0 |
| 1 | 1.0 | 3.0 | 3.0 |
| 2 | 2.0 | 3.0 | 3.0 |
| 3 | 4.0 | 4.0 | 2.0 |
| 4 | 2.0 | 2.0 | 3.0 |

| | KBA13_SEG_MINIWAGEN | KBA13_SEG_MITTELKLASSE | KBA13_SEG_OBEREMITTELKLASSE \ |
|---|---------------------|------------------------|-------------------------------|
| 0 | 3.0 | 2.0 | 4.0 |
| 1 | 3.0 | 5.0 | 3.0 |
| 2 | 5.0 | 3.0 | 4.0 |
| 3 | 3.0 | 4.0 | 2.0 |
| 4 | 1.0 | 3.0 | 5.0 |

| | KBA13_SEG_OBERKLASSE | KBA13_SEG_SONSTIGE | KBA13_SEG_SPORTWAGEN | \ |
|---|----------------------|--------------------|----------------------|---|
| 0 | 4.0 | 3.0 | 5.0 | |
| 1 | 3.0 | 3.0 | 2.0 | |
| 2 | 3.0 | 3.0 | 1.0 | |
| 3 | 1.0 | 4.0 | 0.0 | |
| 4 | 5.0 | 1.0 | 2.0 | |

| | KBA13_SEG_UTILITIES | KBA13_SEG_VAN | KBA13_SEG_WOHNMOBILE | KBA13_SITZE_4 | \ |
|---|---------------------|---------------|----------------------|---------------|---|
| 0 | 2.0 | 2.0 | 0.0 | 4.0 | |
| 1 | 4.0 | 3.0 | 3.0 | 3.0 | |
| 2 | 3.0 | 3.0 | 0.0 | 3.0 | |
| 3 | 4.0 | 2.0 | 3.0 | 2.0 | |
| 4 | 3.0 | 4.0 | 3.0 | 1.0 | |

| | KBA13_SITZE_5 | KBA13_SITZE_6 | KBA13_TOYOTA | KBA13_VORB_0 | KBA13_VORB_1 | \ |
|---|---------------|---------------|--------------|--------------|--------------|---|
| 0 | 3.0 | 2.0 | 2.0 | 3.0 | 4.0 | |
| 1 | 2.0 | 4.0 | 4.0 | 2.0 | 2.0 | |
| 2 | 3.0 | 3.0 | 4.0 | 5.0 | 3.0 | |
| 3 | 4.0 | 3.0 | 3.0 | 3.0 | 2.0 | |
| 4 | 4.0 | 4.0 | 1.0 | 1.0 | 4.0 | |

| | KBA13_VORB_1_2 | KBA13_VORB_2 | KBA13_VORB_3 | KBA13_VW | KK_KUNDENTYP | KKK | \ |
|---|----------------|--------------|--------------|----------|--------------|-----|---|
| 0 | 3.0 | 2.0 | 0.0 | 3.0 | 4.0 | 2.0 | |
| 1 | 3.0 | 4.0 | 5.0 | 3.0 | NaN | 2.0 | |
| 2 | 2.0 | 2.0 | 0.0 | 3.0 | NaN | 3.0 | |
| 3 | 2.0 | 3.0 | 5.0 | 3.0 | NaN | 3.0 | |
| 4 | 4.0 | 4.0 | 4.0 | 5.0 | 1.0 | 3.0 | |

| | KOMBIALTER | KONSUMNAEHE | KONSUMZELLE | LP_FAMILIE_FEIN | LP_FAMILIE_GROB | \ |
|---|------------|-------------|-------------|-----------------|-----------------|---|
| 0 | 4 | 4.0 | 0.0 | 2.0 | 2.0 | |
| 1 | 4 | 2.0 | 0.0 | 1.0 | 1.0 | |
| 2 | 4 | 1.0 | 1.0 | 10.0 | 5.0 | |
| 3 | 4 | 4.0 | 0.0 | 0.0 | 0.0 | |
| 4 | 4 | 6.0 | 0.0 | 11.0 | 5.0 | |

| | LP_LEBENSPHASE_FEIN | LP_LEBENSPHASE_GROB | LP_STATUS_FEIN | LP_STATUS_GROB | \ |
|---|---------------------|---------------------|----------------|----------------|---|
| 0 | 20.0 | 5.0 | 10.0 | 5.0 | |
| 1 | 6.0 | 2.0 | 1.0 | 1.0 | |
| 2 | 40.0 | 12.0 | 10.0 | 5.0 | |
| 3 | 0.0 | 0.0 | 3.0 | 2.0 | |
| 4 | 37.0 | 12.0 | 9.0 | 4.0 | |

| | MIN_GEBAEUDEJAHR | MOBI_RASTER | MOBI_REGIO | NATIONALITAET_KZ | \ |
|---|------------------|-------------|------------|------------------|---|
| 0 | 1993.0 | 5.0 | 4.0 | 1 | |
| 1 | 1992.0 | 1.0 | 2.0 | 1 | |
| 2 | 1992.0 | 2.0 | 3.0 | 1 | |
| 3 | 1992.0 | 5.0 | 5.0 | 1 | |
| 4 | 1992.0 | 5.0 | 5.0 | 1 | |

| | ONLINE_AFFINITAET | ORTSGR_KLS9 | OST_WEST_KZ | PLZ8_ANTG1 | PLZ8_ANTG2 | \ |
|---|-------------------|-------------|-------------|------------|------------|---|
| 0 | 4.0 | 4.0 | W | 2.0 | 3.0 | |
| 1 | 1.0 | 5.0 | W | 2.0 | 3.0 | |
| 2 | 0.0 | 6.0 | W | 3.0 | 2.0 | |
| 3 | 3.0 | 8.0 | 0 | 4.0 | 1.0 | |
| 4 | 5.0 | 3.0 | W | 4.0 | 2.0 | |

| | PLZ8_ANTG3 | PLZ8_ANTG4 | PLZ8_BAUMAX | PLZ8_GBZ | PLZ8_HHZ | \ |
|---|------------|------------|-------------|----------|----------|---|
| 0 | 2.0 | 1.0 | 1.0 | 3.0 | 3.0 | |
| 1 | 2.0 | 1.0 | 1.0 | 3.0 | 5.0 | |
| 2 | 1.0 | 0.0 | 1.0 | 5.0 | 4.0 | |
| 3 | 0.0 | 0.0 | 1.0 | 3.0 | 3.0 | |
| 4 | 0.0 | 0.0 | 1.0 | 2.0 | 1.0 | |

| | PRAEGENDE_JUGENDJAHRE | REGIOTYP | RELAT_AB | RETOURTYP_BK_S | RT_KEIN_ANREIZ | \ |
|---|-----------------------|----------|----------|----------------|----------------|---|
| 0 | 2 | 3.0 | 4.0 | 5.0 | 2.0 | |
| 1 | 5 | 4.0 | 5.0 | 5.0 | 1.0 | |
| 2 | 6 | 6.0 | 5.0 | 5.0 | 2.0 | |
| 3 | 5 | 5.0 | 4.0 | 5.0 | 1.0 | |
| 4 | 8 | 4.0 | 2.0 | 3.0 | 3.0 | |

| | RT_SCHNAEPPCHEN | RT_UEBERGROESSE | SEMIO_DOM | SEMIO_ERL | SEMIO_FAM | \ |
|---|-----------------|-----------------|-----------|-----------|-----------|---|
| 0 | 5.0 | 1.0 | 3 | 3 | 4 | |
| 1 | 5.0 | 1.0 | 3 | 7 | 4 | |
| 2 | 5.0 | 1.0 | 5 | 7 | 1 | |
| 3 | 3.0 | 1.0 | 5 | 7 | 1 | |
| 4 | 2.0 | 3.0 | 6 | 7 | 2 | |

| | SEMIO_KAEM | SEMIO_KRIT | SEMIO_KULT | SEMIO_LUST | SEMIO_MAT | SEMIO_PFLICHT | \ |
|---|------------|------------|------------|------------|-----------|---------------|---|
| 0 | 2 | 3 | 6 | 7 | 1 | 1 | |
| 1 | 3 | 3 | 4 | 5 | 4 | 3 | |
| 2 | 6 | 4 | 2 | 7 | 1 | 4 | |
| 3 | 6 | 7 | 2 | 7 | 2 | 2 | |
| 4 | 5 | 4 | 2 | 5 | 2 | 3 | |

| | SEMIO_RAT | SEMIO_REL | SEMIO_SOZ | SEMIO_TRADV | SEMIO_VERT | SHOPPER_TYP | \ |
|---|-----------|-----------|-----------|-------------|------------|-------------|---|
| 0 | 1 | 2 | 6 | 1 | 7 | 3 | |
| 1 | 3 | 3 | 6 | 1 | 6 | 2 | |
| 2 | 3 | 1 | 4 | 3 | 3 | 3 | |
| 3 | 3 | 2 | 3 | 3 | 2 | 3 | |
| 4 | 3 | 2 | 5 | 3 | 3 | 3 | |

| | SOHO_KZ | STRUKTURTYP | TITEL_KZ | UMFELD_ALT | UMFELD_JUNG | UNGLEICHENN_FLAG | \ |
|---|---------|-------------|----------|------------|-------------|------------------|---|
| 0 | 0.0 | 3.0 | 0.0 | 2.0 | 5.0 | 0.0 | |
| 1 | 0.0 | 3.0 | 0.0 | 4.0 | 3.0 | 0.0 | |
| 2 | 0.0 | 3.0 | 0.0 | 1.0 | 5.0 | 1.0 | |
| 3 | 0.0 | 3.0 | 0.0 | 2.0 | 5.0 | 0.0 | |

| | | | | | | |
|---|-----|-----|-----|-----|-----|-----|
| 4 | 0.0 | 2.0 | 0.0 | 5.0 | 5.0 | 0.0 |
|---|-----|-----|-----|-----|-----|-----|

| | VERDICHTUNGSRAUM | VERS_TYP | VHA | VHN | VK_DHT4A | VK_DISTANZ | VK_ZG11 \ |
|---|------------------|----------|-----|-----|----------|------------|-----------|
| 0 | 23.0 | 1 | 1.0 | 4.0 | 5.0 | 6.0 | 3.0 |
| 1 | 0.0 | 1 | 1.0 | 1.0 | 5.0 | 2.0 | 1.0 |
| 2 | 15.0 | 1 | 1.0 | 3.0 | 9.0 | 6.0 | 3.0 |
| 3 | 10.0 | 2 | 1.0 | 2.0 | 6.0 | 6.0 | 3.0 |
| 4 | 0.0 | 1 | 1.0 | 1.0 | 2.0 | 4.0 | 3.0 |

| | W_KEIT_KIND_HH | WOHNDAUER_2008 | WOHNLAGE | ZABEOTYP | ANREDE_KZ \ |
|---|----------------|----------------|----------|----------|-------------|
| 0 | 6.0 | 9.0 | 3.0 | 3 | 1 |
| 1 | 6.0 | 9.0 | 5.0 | 3 | 1 |
| 2 | 2.0 | 9.0 | 4.0 | 3 | 2 |
| 3 | NaN | 9.0 | 2.0 | 3 | 2 |
| 4 | 3.0 | 9.0 | 7.0 | 4 | 2 |

| | ALTERSKATEGORIE_GROB |
|---|----------------------|
| 0 | 4 |
| 1 | 4 |
| 2 | 4 |
| 3 | 4 |
| 4 | 4 |

In [166]: mailout_test.shape

Out[166]: (42833, 366)

```
In [167]: lnr = mailout_test['LNR']
          mailout_test_cleaned = clean_demographic_data(mailout_test, unknown)
```

Whether all values in LNR column are unique in azdias DF - True

Final check of missing values 0

```
In [168]: pred_3 = pipeline.predict(mailout_test_cleaned)
          pred_4 = cv.predict(mailout_test_cleaned)
```

```
In [171]: arvato_capstone_submission_3 = pd.DataFrame({'LNR': lnr, 'RESPONSE': pred_3})
          arvato_capstone_submission_4 = pd.DataFrame({'LNR': lnr, 'RESPONSE': pred_4})
```

```
In [172]: arvato_capstone_submission_3.to_csv('Arvato_Capstone_Submission_3.csv', index=False)
          arvato_capstone_submission_4.to_csv('Arvato_Capstone_Submission_4.csv', index=False)
```

In []:

In []:

In []: