

# project

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Setting up our relevante Dataset:</b>	<b>2</b>
<b>3</b>	<b>Descriptice Analysis</b>	<b>5</b>
3.1	Numeric Variables . . . . .	5
3.1.1	dseitz: working in the current job since. . . (in months) . . . . .	6

```

library(tidyverse)
library(ggplot2)
library(reshape2)
library(dplyr)
library(tidyr)
library(MASS) # For Box-Cox
library(moments) # For skewness/kurtosis
library(ggfortify)
library(viridis) # For a professional color palette
library(ggpubr) # For boxplots # For normality tests and transformations
library(ggthemes) # For better themes
library(kableExtra)
library(GGally)

```

## 1 Introduction

fgfgfgfg

## 2 Setting up our relevante Dataset:

```

data_loc = "mc.csv"
data <- read.csv(data_loc)
#-3 values represent null values so we assing NA
data[data == -3] <- NA
#subset = styria -> NUTS2 = AT22
data <- data |> filter(xnuts2 == 22)
#filtering only for the relevant Predictors
data <- data |> dplyr::select(werr, dseitz, dstd, kjahr, xanzkind, xminalt,
                           balt5, bsex, bfst, xbstaat, xbgeblan, xhatlevel, xeinw, xlfi, xpatch)
head(data)

```

	werr	dseitz	dstd	kjahr	xanzkind	xminalt	balt5	bsex	bfst	xbstaat	xbgeblan
1	3	17	30	26	NA	NA	7	1	4	1	1
2	6	NA	NA	50	0	25	11	1	2	1	1
3	6	NA	NA	47	0	25	10	2	2	1	1
4	8	NA	NA	69	1	24	14	2	3	1	1
5	8	NA	NA	45	1	24	10	2	1	1	1
6	4	NA	NA	49	NA	NA	14	1	2	1	1

	xhatlevel	xeinw	xlfi	xpatch
1	32	4	1	NA
2	32	1	3	NA
3	32	1	3	NA
4	21	2	3	NA
5	21	2	3	NA
6	51	4	3	NA

```
data <- data %>%
  mutate(
    werr = factor(werr, levels = 1:8,
                  labels = c("before 1919", "1919-1944", "1945-1960",
                             "1961-1970", "1971-1980", "1981-1990",
                             "1991-2000", "after 2000")),

    balt5 = factor(balt5, levels = 0:15,
                   labels = c("0-14", "15-19", "20-24", "25-29", "30-34", "35-39",
                              "40-44", "45-49", "50-54", "55-59", "60-64", "65-69",
                              "70-74", "75-79", "80-84", "85+")),

    bsex = factor(bsex, levels = c(1, 2), labels = c("Male", "Female")),

    bfst = factor(bfst, levels = 1:4,
                  labels = c("Single", "Married", "Widowed", "Divorced")),

    xbstaat = factor(xbstaat, levels = 1:7,
                     labels = c("Austria", "EU15 without Austria", "EU15 10 new members",
                                "Former Yugoslavia", "Turkey", "Other countries", "Bulgaria/1

    xbgeblan = factor(xbgeblan, levels = 1:7,
                      labels = c("Austria", "EU15 without Austria", "EU15 10 new members",
                                 "Former Yugoslavia", "Turkey", "Other countries", "Bulgaria,

    xhatlevel = factor(xhatlevel, levels = c(0, 11, 21, 22, 30, 31, 32, 41, 42, 43, 51, 52, 6
                  labels = c("ISCED 0/1", "ISCED 1", "ISCED 2", "ISCED 3c <2 years",
                              "ISCED 3", "ISCED 3c 2+ years", "ISCED 3a, b", "ISCED 4a, b",
                              "ISCED 4c", "ISCED 4", "ISCED 5b", "ISCED 5a", "ISCED 6", "

    xeinw = factor(xeinw, levels = 1:4,
                   labels = c("up to 2000", "2001-10000", "10001-100000", "100001+")),

    xlfi = factor(xlfi, levels = 1:3,
```

```

        labels = c("Employed", "Unemployed", "Not in labor force")),

    xpatch = factor(xpatch, levels = c(1, 2), labels = c("Yes", "No"))
)
data <- na.omit(data)

head(data)

```

	werr	dseitz	dstd	kjahr	xanzkind	xminalt	balt5	bsex	bfst	xbstaat
10	after 2000	4	30	30	2	13	50-54	Male	Married	Austria
11	after 2000	156	70	15	2	13	45-49	Female	Married	Austria
14	1991-2000	74	38	31	2	14	45-49	Male	Single	Austria
15	1991-2000	216	34	24	2	14	40-44	Female	Single	Austria
16	1991-2000	11	40	3	2	14	20-24	Female	Single	Austria
21	1981-1990	12	39	3	1	20	20-24	Male	Single	Austria
	xbgeblan	xhatlevel		xeinw	xlfi	xpatch				
10	Austria	ISCED 3a, b		100001+	Employed	No				
11	Austria	ISCED 5a		100001+	Employed	No				
14	Austria	ISCED 3a, b up to 2000			Employed	No				
15	Austria	ISCED 3a, b up to 2000			Employed	No				
16	Austria	ISCED 3a, b up to 2000			Employed	No				
21	Austria	ISCED 3a, b 2001-10000			Employed	No				

```

data.numeric <-c("dseitz","dstd","kjahr","xanzkind")
data.polytomous <- c("balt5","bfst","xbstaat","xbgeblan","xhatlevel","xeinw","xlfi")
data.categorical <- c("balt5","bsex","bfst","xbstaat","xbgeblan","xhatlevel","xeinw","xlfi",

```

## 3 Descriptive Analysis

### 3.1 Numeric Variables

```
# Load required libraries
library(ggplot2)
library(ggpubr)

# Define the function
plot_numeric_variable <- function(data, column_name, target_variable, plot_title) {

  # Histogram with density line
  hist_plot <- ggplot(data, aes_string(x = column_name)) +
    geom_histogram(aes(y = ..density..), bins = 30) +
    geom_density() +
    labs(x = "Data", y = "Density") +
    ggtitle("Histogram") +
    theme_grey() +
    scale_colour_grey()

  # Boxplot
  boxplot <- ggplot(data, aes_string(y = column_name)) +
    geom_boxplot() +
    xlim(-1, 1) +
    labs(y = "Data") +
    ggtitle("Boxplot") +
    theme_grey() +
    scale_colour_grey() +
    scale_fill_grey()

  # Relationship between the numeric variable and the categorical target variable
  relationship_plot <- ggplot(data, aes_string(x = target_variable, y = column_name)) +
    geom_boxplot() +
    labs(x = target_variable, y = column_name) +
    ggtitle("relationship with target werr") +
    theme_grey() +
    scale_colour_grey() +
    scale_fill_grey()

  # Arrange all three plots in one row
  plot <- ggarrange(hist_plot, boxplot, relationship_plot, ncol = 3, nrow = 1, widths = c(0.33, 0.33, 0.33))
}
```

```

# Add title to the combined plot
plot_with_title <- annotate_figure(plot,
                                   top = text_grob(paste(plot_title, target_variable),
                                                    face = "bold", size = 14))

# Return the plot with title
return(plot_with_title)
}

```

### 3.1.1 dseitz: working in the current job since. . . (in months)

```
summary(data$dseitz)
```

```

      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      0.0    29.5    95.5   127.6   187.5   469.0

```

```
sd(data$dseitz)
```

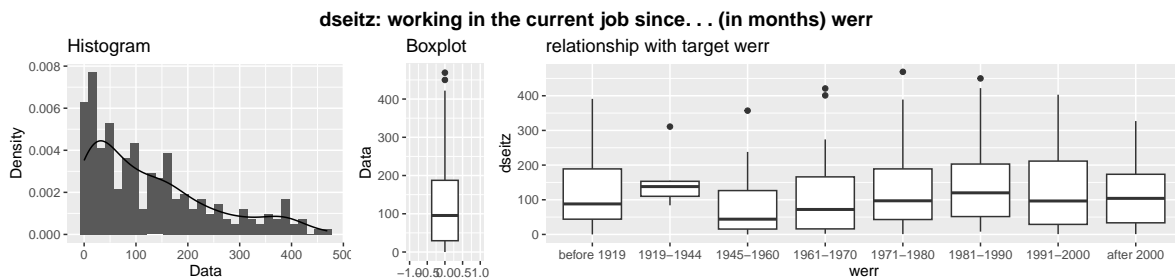
```
[1] 117.1715
```

```

#|warning: false
#|

```

```
plot_numeric_variable(data, "dseitz", "werr", "dseitz: working in the current job since. . . (in months) werr")
```



```

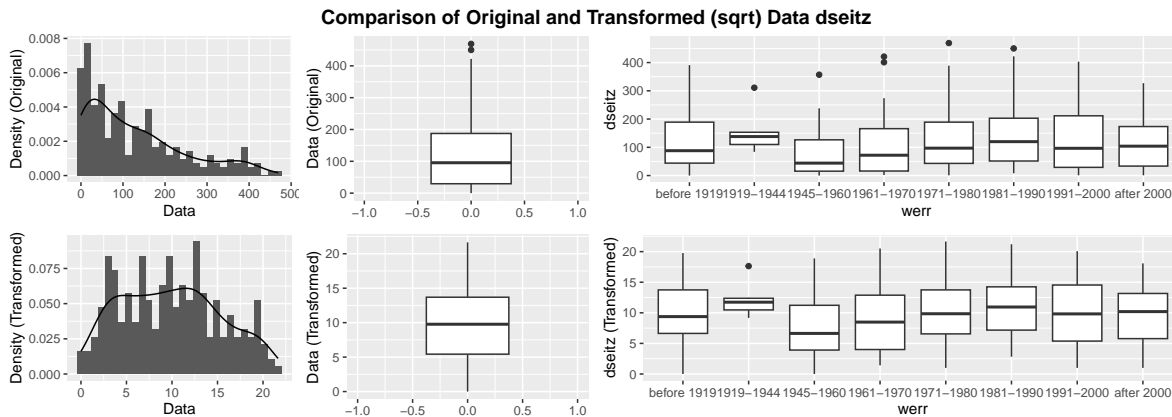
#plot_numeric_variable(data, "dstd", "werr", "dstd: normal weekly working hours")
#plot_numeric_variable(data, "kjahr", "werr", "kjahr: years since completing the highest level")
#plot_numeric_variable(data, "xanzkind", "werr", "xanzkind: number of children under 18 in the family")
#plot_numeric_variable(data, "xminalt", "werr", "age of youngest child in the family (in years)")

```

The histogram shows that the variable follows a right-skewed distribution and has a high spread. Through the boxplot we see that most values are between 0 and 200 with some outliers above 400. The box plot categorized by the buildings-Year shows that the distribution of dseitz differs across the categories. Additionally we see some outliers, but none of them seem to strongly influence the mean of the category, except 1919-1944 which shows a mean skewed towards the outlier.

The skewness and the different distributions throughout the categories might indicate that a transformation would help to conform more to a normal-distribution.

```
plot_numeric_variable_with_transformation(data,"dseitz","werr")
```



```
# Perform Shapiro-Wilk test before transformation
shapiro_before <- shapiro.test(data$dseitz)
data_sqrt <- sqrt(data$dseitz)
shapiro_after <- shapiro.test(data_sqrt)

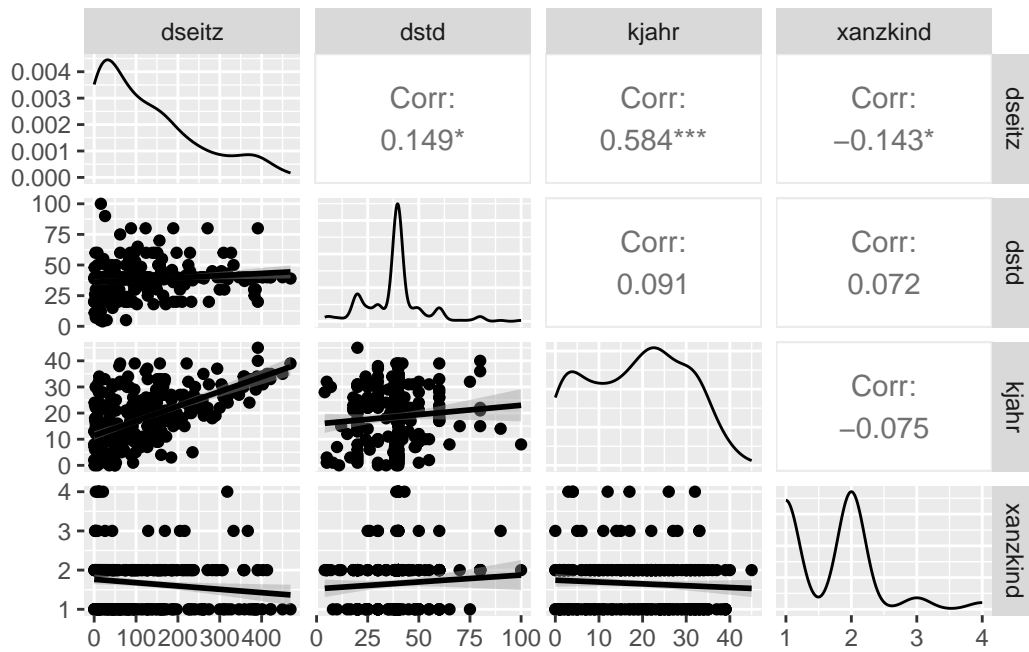
# Create a formatted table of test results
shapiro_results <- data.frame(
  Test = c("Original Data", "Square Root Transformed"),
  W_Statistic = c(shapiro_before$statistic, shapiro_after$statistic),
  P_Value = c(shapiro_before$p.value, shapiro_after$p.value)
)
kable(shapiro_results, caption = "Shapiro-Wilk Normality Test Results", digits = 5)
```

Table 1: Shapiro-Wilk Normality Test Results

Test	W_Statistic	P_Value
Original Data	0.88551	0e+00
Square Root Transformed	0.97086	4e-05

After applying the

```
pair_data <- data[,data.numeric]
ggpairs(pair_data,
  lower = list(continuous = "smooth"), # Smoothed scatterplots on the lower panel
  diag = list(continuous = "densityDiag"), # Density plots on the diagonal
  upper = list(continuous = "cor")) # Add correlation coefficients on the upper panel
```

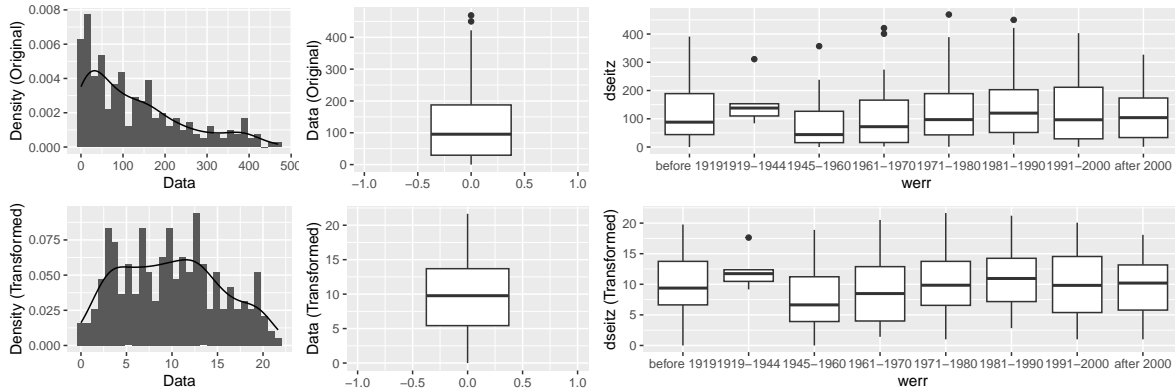


### 3.1.1.1 Examine Outliers

```
plot_numeric_variable_with_transformation(data, "dseitz", "werr")
```

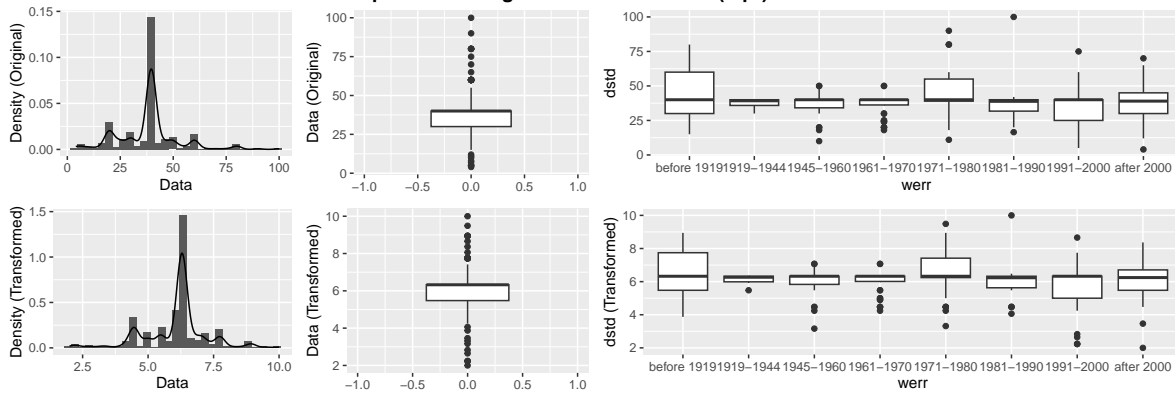


Comparison of Original and Transformed (sqrt) Data dseitz



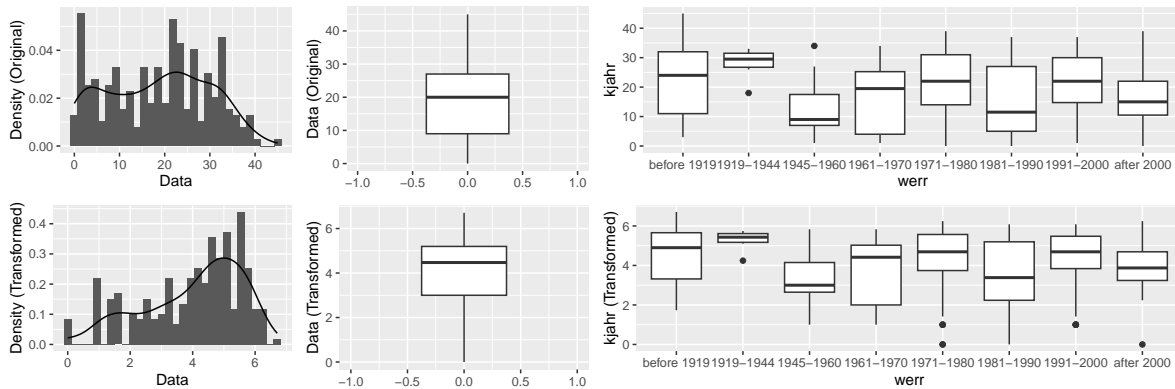
```
plot_numeric_variable_with_transformation(data,"dstd","werr")
```

Comparison of Original and Transformed (sqrt) Data dstd



```
plot_numeric_variable_with_transformation(data,"kjahr","werr")
```

Comparison of Original and Transformed (sqrt) Data kjahr



```
plot_numeric_variable_with_transformation(data, "xanzkind", "werr")
```



### 3.1.1.2 Consider Transformations for Normality

```
# Function to remove outliers and compare distributions
remove_outliers_IQR <- function(data, column_name) {

  # Step 1: Identify Outliers using IQR
  Q1 <- quantile(data[[column_name]], 0.25)
  Q3 <- quantile(data[[column_name]], 0.75)
  IQR <- Q3 - Q1
  lower_bound <- Q1 - 1.5 * IQR
  upper_bound <- Q3 + 1.5 * IQR

  # Step 2: Create a new column 'is_outlier' to indicate outliers
  data$is_outlier <- ifelse(data[[column_name]] < lower_bound | data[[column_name]] > upper_bound, TRUE, FALSE)

  # Step 3: Remove the outliers by filtering the rows where 'is_outlier' is FALSE
  data_no_outliers <- data[data$is_outlier == FALSE, ]

  # Step 4: Return the dataset without outliers (with the 'is_outlier' column still present)
  return(data_no_outliers)
}

temp_data <- remove_outliers_IQR(data, "dstd")
temp_data
```

```
werr dseitz dstd kjahr xanzkind xminalt balt5 bsex bfst
```

10	after 2000	4	30.0	30	2	13	50-54	Male	Married
14	1991-2000	74	38.0	31	2	14	45-49	Male	Single
15	1991-2000	216	34.0	24	2	14	40-44	Female	Single
16	1991-2000	11	40.0	3	2	14	20-24	Female	Single
21	1981-1990	12	39.0	3	1	20	20-24	Male	Single
22	1961-1970	52	50.0	21	1	20	50-54	Male	Divorced
39	1981-1990	13	16.5	3	1	20	20-24	Female	Single
40	1981-1990	127	30.0	37	1	20	55-59	Female	Married
41	1981-1990	450	40.0	35	1	20	50-54	Male	Married
51	1945-1960	151	40.0	13	2	9	35-39	Male	Single
58	1961-1970	138	38.5	26	1	12	40-44	Male	Married
59	1961-1970	258	40.0	21	1	12	40-44	Female	Married
68	1991-2000	7	39.0	1	1	16	15-19	Male	Single
77	1945-1960	357	40.0	34	2	20	50-54	Male	Married
80	1945-1960	44	40.0	2	2	20	20-24	Male	Single
90	after 2000	23	25.0	14	2	9	35-39	Female	Married
94	1945-1960	41	50.0	9	2	0	25-29	Male	Married
99	after 2000	60	38.5	31	1	21	45-49	Female	Married
100	after 2000	319	38.5	24	1	21	40-44	Male	Married
101	after 2000	127	38.5	7	1	21	25-29	Male	Single
102	1981-1990	413	40.0	35	2	19	50-54	Female	Married
103	1981-1990	167	38.5	27	2	19	50-54	Male	Widowed
105	1981-1990	17	40.0	2	2	19	15-19	Female	Single
109	before 1919	161	38.5	27	1	11	45-49	Male	Single
110	before 1919	12	25.0	24	1	11	40-44	Female	Divorced
117	after 2000	127	40.0	15	1	0	30-34	Female	Single
118	after 2000	189	36.0	20	1	0	35-39	Male	Single
126	1991-2000	127	40.0	22	2	9	40-44	Male	Married
127	1991-2000	210	20.0	18	2	9	35-39	Female	Married
139	1961-1970	229	40.0	25	1	16	55-59	Male	Married
140	1961-1970	274	20.0	31	1	16	50-54	Female	Married
147	1971-1980	28	50.0	1	1	20	20-24	Male	Single
148	1971-1980	162	55.0	31	1	20	45-49	Male	Divorced
150	1991-2000	187	40.0	21	2	9	35-39	Male	Married
155	1961-1970	128	25.0	22	2	5	35-39	Female	Single
163	1961-1970	172	40.0	11	1	9	25-29	Male	Single
166	1961-1970	74	24.0	10	1	9	25-29	Female	Single
169	1945-1960	112	48.5	7	1	20	20-24	Male	Single
170	1961-1970	11	40.0	32	4	20	45-49	Female	Divorced
172	1961-1970	21	40.0	3	4	20	20-24	Female	Single
173	1961-1970	4	40.0	12	4	20	25-29	Male	Divorced
174	1961-1970	10	40.0	4	4	20	20-24	Male	Single
175	1961-1970	13	40.0	4	4	20	20-24	Male	Single

179	1991-2000	388	40.0	33	1	20	50-54	Male	Married
185	1971-1980	56	20.0	16	2	10	30-34	Female	Married
201	1971-1980	389	40.0	34	1	19	50-54	Male	Married
212	1961-1970	16	30.0	10	2	4	25-29	Female	Married
252	1945-1960	30	30.0	10	2	3	25-29	Female	Married
253	1945-1960	153	50.0	9	2	3	25-29	Male	Married
259	1991-2000	115	40.0	25	2	6	40-44	Male	Married
260	1991-2000	32	20.0	21	2	6	40-44	Female	Married
263	1991-2000	80	20.0	14	1	19	45-49	Female	Single
278	1971-1980	155	40.0	17	2	0	30-34	Male	Married
283	1991-2000	107	38.5	23	2	11	40-44	Female	Married
288	1991-2000	241	40.0	30	1	20	45-49	Female	Married
289	1991-2000	45	20.0	4	1	20	20-24	Male	Single
298	1981-1990	89	42.0	10	2	1	30-34	Male	Married
301	before 1919	333	50.0	28	3	13	50-54	Male	Married
302	before 1919	169	40.0	33	3	13	50-54	Female	Married
308	after 2000	235	37.5	5	1	21	45-49	Female	Married
309	after 2000	1	48.0	7	1	21	25-29	Female	Single
325	1991-2000	157	50.0	21	1	2	35-39	Male	Married
326	1991-2000	93	20.0	1	1	2	30-34	Female	Married
331	1991-2000	10	20.0	6	2	2	25-29	Female	Single
335	1961-1970	21	25.0	23	1	15	40-44	Female	Single
338	1961-1970	8	40.0	1	1	15	15-19	Male	Single
342	1991-2000	95	42.0	20	1	18	50-54	Male	Married
343	1991-2000	79	20.0	30	1	18	45-49	Female	Married
344	1991-2000	31	40.0	4	1	18	15-19	Female	Single
347	1991-2000	8	40.0	1	2	17	15-19	Female	Single
350	1991-2000	11	40.0	28	1	19	45-49	Female	Married
363	after 2000	96	30.0	9	1	12	40-44	Female	Single
368	1991-2000	8	50.0	24	1	15	35-39	Male	Married
369	1991-2000	262	40.0	22	1	15	35-39	Female	Married
383	1991-2000	1	40.0	24	2	6	40-44	Male	Single
390	after 2000	98	53.0	10	2	1	35-39	Male	Single
391	after 2000	45	21.0	15	2	1	35-39	Female	Single
401	1991-2000	168	38.5	32	1	20	50-54	Male	Married
403	1991-2000	14	38.5	5	1	20	20-24	Female	Single
404	1945-1960	52	20.0	25	2	18	35-39	Female	Married
405	1945-1960	238	40.0	27	2	18	40-44	Male	Married
406	1945-1960	9	40.0	1	2	18	20-24	Female	Single
407	1945-1960	34	40.0	3	2	18	15-19	Male	Single
415	1991-2000	307	40.0	30	2	19	50-54	Male	Married
429	after 2000	4	25.0	22	3	5	35-39	Female	Single
437	1991-2000	261	40.0	35	1	16	50-54	Male	Single

438	1991-2000	40	25.0	26	1	16	40-44	Female	Single
441	1991-2000	44	47.0	31	2	18	45-49	Female	Divorced
445	1991-2000	80	39.0	20	1	10	35-39	Male	Married
448	after 2000	157	40.0	39	1	18	55-59	Female	Married
450	1991-2000	403	40.0	31	2	13	45-49	Male	Married
459	1971-1980	253	40.0	25	1	11	40-44	Female	Married
480	1991-2000	5	20.0	2	2	16	20-24	Female	Single
481	1991-2000	25	40.0	33	2	16	50-54	Male	Single
482	1991-2000	357	40.0	30	2	16	45-49	Female	Single
486	1961-1970	4	40.0	16	2	0	30-34	Male	Married
487	1961-1970	71	40.0	6	2	0	25-29	Female	Married
490	1991-2000	24	40.0	11	2	6	40-44	Female	Married
495	before 1919	27	40.0	3	2	15	15-19	Male	Single
498	before 1919	382	30.0	32	1	21	55-59	Female	Widowed
499	before 1919	0	20.0	8	1	21	25-29	Female	Single
507	after 2000	216	40.0	15	3	1	30-34	Male	Single
539	1961-1970	142	18.0	26	2	16	40-44	Female	Married
540	1961-1970	224	38.5	24	2	16	50-54	Male	Married
541	1961-1970	46	38.3	1	2	16	15-19	Female	Single
542	1961-1970	10	38.3	1	2	16	15-19	Male	Single
557	1971-1980	189	20.0	16	1	6	30-34	Female	Single
561	1961-1970	73	20.0	31	1	15	45-49	Female	Married
562	1961-1970	93	40.0	26	1	15	40-44	Male	Married
568	1981-1990	113	38.0	22	2	9	40-44	Male	Single
569	1981-1990	197	20.0	20	2	9	35-39	Female	Divorced
577	1991-2000	359	40.0	32	2	20	50-54	Female	Married
588	1971-1980	383	25.0	32	1	13	50-54	Female	Married
589	1971-1980	377	39.0	33	1	13	50-54	Male	Married
595	1981-1990	160	40.0	4	2	0	35-39	Male	Married
596	1981-1990	170	20.0	17	2	0	35-39	Female	Married
599	after 2000	161	45.0	13	2	7	30-34	Male	Married
603	1961-1970	34	40.0	15	1	13	35-39	Female	Married
604	1961-1970	164	38.5	19	1	13	35-39	Male	Married
608	after 2000	31	38.0	10	2	0	25-29	Male	Divorced
609	after 2000	54	25.0	9	2	0	20-24	Female	Divorced
616	after 2000	94	34.0	18	2	3	35-39	Female	Single
625	after 2000	186	46.0	22	2	7	40-44	Female	Married
626	after 2000	160	46.0	17	2	7	40-44	Male	Married
633	1919-1944	104	40.0	18	2	3	30-34	Male	Single
639	after 2000	37	42.0	18	2	1	35-39	Male	Married
640	after 2000	92	20.0	15	2	1	30-34	Female	Married
643	1981-1990	371	38.5	24	1	14	45-49	Male	Married
656	1961-1970	46	40.0	24	1	15	40-44	Male	Married

657	1961-1970	132	40.0	17	1	15	30-34	Female	Married
661	1971-1980	304	45.0	31	2	10	45-49	Male	Married
673	1991-2000	150	30.0	24	2	14	40-44	Female	Married
674	1991-2000	128	40.0	19	2	14	45-49	Male	Married
679	1981-1990	67	38.7	9	1	11	25-29	Female	Single
693	after 2000	117	35.0	11	1	6	30-34	Female	Divorced
694	after 2000	262	39.0	19	1	6	35-39	Male	Divorced
700	1971-1980	97	40.0	39	1	20	55-59	Male	Married
702	1971-1980	10	38.5	1	1	20	20-24	Female	Single
704	after 2000	229	45.0	22	1	3	45-49	Male	Single
705	after 2000	300	30.0	22	1	3	40-44	Female	Single
707	1971-1980	49	45.0	21	2	10	40-44	Male	Married
708	1971-1980	44	18.0	22	2	10	40-44	Female	Married
717	1945-1960	102	18.0	23	2	10	40-44	Female	Married
718	1945-1960	141	39.0	22	2	10	40-44	Male	Married
719	1945-1960	0	39.0	2	2	10	15-19	Female	Single
723	1981-1990	359	37.0	34	1	21	50-54	Female	Single
724	1981-1990	65	40.0	5	1	21	25-29	Female	Single
725	1971-1980	184	40.0	27	2	13	45-49	Male	Married
726	1971-1980	185	40.0	25	2	13	40-44	Female	Married
731	after 2000	36	40.0	0	1	20	20-24	Female	Single
742	1991-2000	346	40.0	27	1	20	50-54	Male	Married
743	1991-2000	165	25.0	30	1	20	45-49	Female	Married
744	1991-2000	99	38.5	5	1	20	20-24	Male	Single
747	after 2000	282	45.0	18	2	9	50-54	Male	Married
748	after 2000	13	40.0	12	2	9	35-39	Female	Married
752	1991-2000	171	18.0	21	2	8	35-39	Female	Divorced
761	after 2000	2	40.0	18	2	6	30-34	Female	Single
764	1991-2000	11	39.0	17	4	3	40-44	Female	Married
765	1991-2000	318	43.0	26	4	3	50-54	Male	Married
778	1981-1990	422	40.0	35	1	20	55-59	Male	Single
779	1981-1990	50	39.0	0	1	20	20-24	Male	Single
781	1991-2000	294	40.0	37	2	3	55-59	Male	Married
802	1961-1970	225	50.0	32	1	20	50-54	Male	Married
803	1961-1970	81	24.0	34	1	20	50-54	Female	Married
804	1961-1970	16	30.0	1	1	20	20-24	Female	Single
805	1971-1980	76	50.0	21	2	20	45-49	Male	Married
806	1971-1980	182	41.0	26	2	20	45-49	Female	Married
812	1961-1970	421	40.0	33	1	20	50-54	Male	Married
813	1961-1970	33	40.0	3	1	20	20-24	Male	Single
823	1971-1980	22	55.0	2	2	17	25-29	Male	Single
826	before 1919	44	15.0	12	1	5	40-44	Male	Divorced
827	before 1919	44	48.0	11	1	5	25-29	Female	Divorced

831	1981-1990	56	20.0	13	1	7	25-29	Female	Single
833	after 2000	129	40.0	11	3	4	40-44	Male	Married
834	after 2000	2	26.0	6	3	4	35-39	Female	Married
852	1919-1944	84	38.5	30	1	18	45-49	Male	Married
853	1919-1944	311	30.0	26	1	18	45-49	Female	Married
861	1945-1960	103	40.0	8	2	1	25-29	Female	Single
862	1945-1960	19	38.3	7	2	1	25-29	Male	Single
865	1971-1980	469	39.0	39	1	18	55-59	Male	Married
866	1971-1980	60	30.0	38	1	18	55-59	Female	Married
869	1991-2000	59	38.5	8	2	9	40-44	Male	Married
876	1945-1960	6	40.0	13	1	7	30-34	Male	Single
877	1945-1960	0	20.0	13	1	7	30-34	Female	Single
880	1971-1980	221	42.0	39	2	19	55-59	Male	Married
881	1971-1980	12	30.0	34	2	19	50-54	Female	Married
882	1971-1980	98	50.0	12	2	19	30-34	Male	Single
883	1971-1980	49	34.0	0	2	19	15-19	Female	Single
885	1971-1980	213	40.0	27	2	20	45-49	Male	Married
888	1971-1980	4	39.0	2	2	20	20-24	Male	Single
898	1991-2000	235	40.0	23	2	13	40-44	Male	Married
899	1991-2000	98	30.0	23	2	13	40-44	Female	Married
900	1991-2000	26	39.0	2	2	13	15-19	Male	Single
902	1961-1970	11	20.0	24	1	20	40-44	Female	Married
903	1961-1970	184	38.5	3	1	20	40-44	Male	Married
904	1961-1970	2	38.5	2	1	20	20-24	Male	Single
907	1991-2000	290	42.5	19	2	3	35-39	Male	Married
912	1971-1980	2	40.0	23	2	14	40-44	Female	Married
923	1991-2000	44	38.5	30	1	20	45-49	Male	Married
924	1991-2000	386	30.0	29	1	20	45-49	Female	Married
929	1961-1970	126	40.0	20	1	17	35-39	Female	Married
930	1961-1970	25	40.0	2	1	17	15-19	Female	Single
938	1961-1970	259	40.0	25	1	18	45-49	Male	Single
939	1961-1970	401	40.0	34	1	18	50-54	Female	Single
946	1919-1944	148	35.0	29	1	10	45-49	Female	Divorced
950	1919-1944	128	40.0	33	2	17	50-54	Male	Married
951	1919-1944	155	40.0	32	2	17	45-49	Female	Married
958	1981-1990	367	40.0	33	3	20	50-54	Male	Married
959	1981-1990	205	30.0	27	3	20	50-54	Female	Married
960	1981-1990	8	30.0	0	3	20	25-29	Male	Single
961	1981-1990	43	39.0	5	3	20	20-24	Male	Single
974	1991-2000	198	40.0	16	2	13	40-44	Male	Married
975	1991-2000	232	20.0	20	2	13	35-39	Female	Married
978	before 1919	391	20.0	45	2	20	60-64	Male	Married
981	before 1919	88	40.0	4	2	20	20-24	Female	Single

993	before 1919	75 40.0	11	2	20 25-29	Male	Single
998	1981-1990	149 38.5	9	1	1 25-29	Male	Married
999	1981-1990	61 40.0	6	1	1 25-29	Female	Married
1001	1991-2000	4 20.0	15	2	10 40-44	Female	Divorced
1004	1971-1980	84 55.0	10	1	21 50-54	Male	Married
1005	1971-1980	43 20.0	3	1	21 45-49	Female	Married
1009	1971-1980	52 39.0	1	1	20 20-24	Male	Married
1010	1971-1980	159 40.0	33	1	20 50-54	Male	Married
	xbstaat			xbgeblan		xhatlevel	xeinw
10	Austria			Austria		ISCED 3a, b	100001+
14	Austria			Austria		ISCED 3a, b	up to 2000
15	Austria			Austria		ISCED 3a, b	up to 2000
16	Austria			Austria		ISCED 3a, b	up to 2000
21	Austria			Austria		ISCED 3a, b	2001-10000
22	Austria			Austria		ISCED 6	10001-100000
39	Austria			Austria		ISCED 4a, b	2001-10000
40	Austria			Austria		ISCED 3a, b	2001-10000
41	Austria			Austria		ISCED 3a, b	2001-10000
51	Austria			Austria		ISCED 5a	100001+
58	Austria			Austria		ISCED 3a, b	up to 2000
59	Austria			Austria		ISCED 3a, b	up to 2000
68	Austria			Austria		ISCED 2	up to 2000
77	Austria			Austria		ISCED 3a, b	up to 2000
80	Austria			Austria		ISCED 3a, b	up to 2000
90	Austria		Other countries			ISCED 5a	100001+
94	Austria			Austria		ISCED 3a, b	2001-10000
99	Austria			Austria		ISCED 3a, b	2001-10000
100	Austria			Austria		ISCED 3a, b	2001-10000
101	Austria			Austria		ISCED 3a, b	2001-10000
102	Austria			Austria		ISCED 3a, b	2001-10000
103	Austria			Austria		ISCED 3a, b	2001-10000
105	Austria			Austria		ISCED 3a, b	2001-10000
109	Austria			Austria		ISCED 3a, b	2001-10000
110	Austria			Austria		ISCED 4a, b	2001-10000
117	Austria			Austria		ISCED 3a, b	up to 2000
118	Austria			Austria		ISCED 3a, b	up to 2000
126	Austria			Austria		ISCED 3a, b	up to 2000
127	Austria			Austria		ISCED 3a, b	up to 2000
139	Austria			Austria		ISCED 6	2001-10000
140	Austria			Austria		ISCED 5b	2001-10000
147	Austria			Austria		ISCED 3a, b	up to 2000
148	Austria			Austria		ISCED 3a, b	up to 2000
150	Austria			Austria		ISCED 3a, b	up to 2000



155	Austria	Austria	ISCED 4a, b	up to 2000
163	Austria	Austria	ISCED 3a, b	up to 2000
166	Austria	Austria	ISCED 3a, b	up to 2000
169	Austria	Austria	ISCED 3a, b	2001-10000
170	Austria	Austria	ISCED 2	up to 2000
172	Austria	Austria	ISCED 3a, b	up to 2000
173	Austria	Austria	ISCED 2	up to 2000
174	Austria	Austria	ISCED 3a, b	up to 2000
175	Austria	Austria	ISCED 3a, b	up to 2000
179	Austria	Austria	ISCED 5b	up to 2000
185	Austria	Austria	ISCED 3a, b	up to 2000
201	Austria	Austria	ISCED 3a, b	up to 2000
212	Austria	Former Yugoslavia	ISCED 3a, b	100001+
252	Austria	Austria	ISCED 3a, b	2001-10000
253	Austria	Austria	ISCED 3a, b	2001-10000
259	Austria	Austria	ISCED 3a, b	up to 2000
260	Austria	Austria	ISCED 3a, b	up to 2000
263	Austria	Austria	ISCED 3a, b	2001-10000
278	Austria	Austria	ISCED 3a, b	2001-10000
283	Austria	Austria	ISCED 3a, b	up to 2000
288	Austria	Austria	ISCED 3a, b	2001-10000
289	Austria	Austria	ISCED 3a, b	2001-10000
298	Austria	Austria	ISCED 3a, b	2001-10000
301	Other countries	Other countries	ISCED 5b	100001+
302	Austria	Austria	ISCED 3a, b	100001+
308	Austria	Austria	ISCED 3a, b	up to 2000
309	Austria	Austria	ISCED 3a, b	up to 2000
325	EU15 without Austria	EU15 without Austria	ISCED 3a, b	2001-10000
326	Austria	Austria	ISCED 4a, b	2001-10000
331	Austria	Austria	ISCED 3a, b	up to 2000
335	Austria	Austria	ISCED 3a, b	2001-10000
338	Austria	Austria	ISCED 2	2001-10000
342	Austria	Austria	ISCED 4a, b	up to 2000
343	Austria	Austria	ISCED 3a, b	up to 2000
344	Austria	Austria	ISCED 2	up to 2000
347	Austria	Austria	ISCED 4a, b	up to 2000
350	Austria	Austria	ISCED 3a, b	2001-10000
363	Austria	Austria	ISCED 3a, b	2001-10000
368	Austria	Austria	ISCED 2	2001-10000
369	Austria	Austria	ISCED 3a, b	2001-10000
383	Austria	Austria	ISCED 3a, b	up to 2000
390	Austria	Austria	ISCED 3c <2 years	2001-10000
391	Austria	Austria	ISCED 5b	2001-10000

401	Austria	Austria	ISCED 3c <2 years	up to 2000
403	Austria	Austria	ISCED 3a, b	up to 2000
404	Austria	Austria	ISCED 2	up to 2000
405	Austria	Austria	ISCED 3a, b	up to 2000
406	Austria	Austria	ISCED 4a, b	up to 2000
407	Austria	Austria	ISCED 2	up to 2000
415	Austria	Austria	ISCED 3a, b	up to 2000
429	Austria	Austria	ISCED 2	up to 2000
437	Austria	Austria	ISCED 3a, b	2001-10000
438	Austria	Austria	ISCED 2	2001-10000
441	Austria	Austria	ISCED 4a, b	2001-10000
445	Austria	Austria	ISCED 3a, b	2001-10000
448	Austria	Austria	ISCED 3a, b	up to 2000
450	Austria	Austria	ISCED 3a, b	2001-10000
459	Austria	Austria	ISCED 2	2001-10000
480	Austria	Austria	ISCED 4a, b	2001-10000
481	Austria	Austria	ISCED 3a, b	2001-10000
482	Austria	Austria	ISCED 3a, b	2001-10000
486	Former Yugoslavia	Former Yugoslavia	ISCED 3a, b	10001-100000
487	Austria	Former Yugoslavia	ISCED 3a, b	10001-100000
490	Austria	Austria	ISCED 6	2001-10000
495	Austria	Austria	ISCED 2	2001-10000
498	Austria	Austria	ISCED 5b	100001+
499	Austria	Austria	ISCED 3a, b	100001+
507	Austria	Austria	ISCED 3a, b	up to 2000
539	Austria	Austria	ISCED 3a, b	2001-10000
540	Austria	Austria	ISCED 3a, b	2001-10000
541	Austria	Austria	ISCED 3a, b	2001-10000
542	Austria	Austria	ISCED 2	2001-10000
557	Austria	Austria	ISCED 2	2001-10000
561	Austria	Austria	ISCED 3a, b	2001-10000
562	Austria	Austria	ISCED 3a, b	2001-10000
568	Austria	Austria	ISCED 3a, b	2001-10000
569	Austria	Austria	ISCED 3a, b	2001-10000
577	Austria	Austria	ISCED 5b	up to 2000
588	Austria	Austria	ISCED 3a, b	up to 2000
589	Austria	Austria	ISCED 3a, b	up to 2000
595	Austria	Austria	ISCED 5a	2001-10000
596	Austria	Austria	ISCED 3a, b	2001-10000
599	Austria	Austria	ISCED 3a, b	up to 2000
603	Austria	Former Yugoslavia	ISCED 3a, b	2001-10000
604	Austria	Former Yugoslavia	ISCED 5b	2001-10000
608	Austria	Austria	ISCED 3a, b	up to 2000

609		Austria		Austria	ISCED 2	up to 2000
616		Austria		Austria	ISCED 3a, b	up to 2000
625	EU15 without	Austria	EU15 without	Austria	ISCED 3a, b	2001-10000
626		Austria		Austria	ISCED 5a	2001-10000
633		Austria		Austria	ISCED 2	100001+
639		Austria		Austria	ISCED 3a, b	up to 2000
640		Austria		Austria	ISCED 3a, b	up to 2000
643		Austria		Austria	ISCED 4a, b	2001-10000
656		Austria		Austria	ISCED 3a, b	up to 2000
657		Austria		Austria	ISCED 3a, b	up to 2000
661		Austria		Austria	ISCED 2	up to 2000
673		Austria		Austria	ISCED 4a, b	2001-10000
674		Austria		Austria	ISCED 5a	2001-10000
679		Austria		Austria	ISCED 3a, b	2001-10000
693		Austria		Austria	ISCED 5a	2001-10000
694		Austria		Austria	ISCED 3a, b	2001-10000
700		Austria		Austria	ISCED 3a, b	2001-10000
702		Austria		Austria	ISCED 4a, b	2001-10000
704		Austria		Austria	ISCED 5a	2001-10000
705		Austria		Austria	ISCED 3a, b	2001-10000
707		Austria		Austria	ISCED 5b	up to 2000
708		Austria		Austria	ISCED 3a, b	up to 2000
717		Austria		Austria	ISCED 3a, b	up to 2000
718		Austria		Austria	ISCED 3a, b	up to 2000
719		Austria		Austria	ISCED 2	up to 2000
723		Austria		Austria	ISCED 3a, b	2001-10000
724		Austria		Austria	ISCED 5b	2001-10000
725		Austria	Other countries		ISCED 4a, b	100001+
726		Austria	Other countries		ISCED 4a, b	100001+
731		Austria		Austria	ISCED 4a, b	2001-10000
742		Austria		Austria	ISCED 5b	up to 2000
743		Austria		Austria	ISCED 3a, b	up to 2000
744		Austria		Austria	ISCED 3a, b	up to 2000
747		Austria		Austria	ISCED 6	2001-10000
748		Austria		Austria	ISCED 5a	2001-10000
752		Austria		Austria	ISCED 3a, b	2001-10000
761		Austria		Austria	ISCED 2	2001-10000
764		Austria		Austria	ISCED 5a	2001-10000
765		Austria		Austria	ISCED 5b	2001-10000
778		Austria		Austria	ISCED 5b	up to 2000
779		Austria		Austria	ISCED 3a, b	up to 2000
781		Austria		Austria	ISCED 4a, b	2001-10000
802		Austria		Austria	ISCED 3a, b	2001-10000

803		Austria		Austria	ISCED 4a, b	2001-10000
804		Austria		Austria	ISCED 3a, b	2001-10000
805		Austria		Austria	ISCED 5a	100001+
806		Austria		Austria	ISCED 5b	100001+
812		Austria		Austria	ISCED 3a, b	up to 2000
813		Austria		Austria	ISCED 3a, b	up to 2000
823		Austria		Austria	ISCED 5a	up to 2000
826		Austria	Other countries		ISCED 5a	100001+
827	Bulgaria/Romania		Bulgaria/Romania		ISCED 3a, b	100001+
831		Austria		Austria	ISCED 2	2001-10000
833		Austria		Austria	ISCED 5a	up to 2000
834		Austria		Austria	ISCED 5a	up to 2000
852		Austria		Austria	ISCED 3a, b	10001-100000
853		Austria		Austria	ISCED 3a, b	10001-100000
861		Austria		Austria	ISCED 3a, b	up to 2000
862		Austria		Austria	ISCED 3a, b	up to 2000
865		Austria		Austria	ISCED 3a, b	up to 2000
866		Austria		Austria	ISCED 3a, b	up to 2000
869		Austria		Austria	ISCED 3c <2 years	up to 2000
876	EU15 without	Austria	EU15 without	Austria	ISCED 3a, b	2001-10000
877	EU15 10 new	members	EU15 10 new	members	ISCED 3a, b	2001-10000
880		Austria		Austria	ISCED 3c <2 years	up to 2000
881		Austria		Austria	ISCED 3a, b	up to 2000
882		Austria		Austria	ISCED 3a, b	up to 2000
883		Austria		Austria	ISCED 3a, b	up to 2000
885		Austria		Austria	ISCED 3a, b	up to 2000
888		Austria		Austria	ISCED 3a, b	up to 2000
898		Austria		Austria	ISCED 3a, b	up to 2000
899		Austria		Austria	ISCED 3a, b	up to 2000
900		Austria		Austria	ISCED 2	up to 2000
902		Austria		Austria	ISCED 3a, b	2001-10000
903		Austria		Austria	ISCED 5b	2001-10000
904		Austria		Austria	ISCED 3a, b	2001-10000
907		Austria		Austria	ISCED 3a, b	2001-10000
912		Austria	Bulgaria/Romania		ISCED 2	2001-10000
923		Austria		Austria	ISCED 3a, b	2001-10000
924		Austria		Austria	ISCED 3a, b	2001-10000
929		Austria		Austria	ISCED 3a, b	up to 2000
930		Austria		Austria	ISCED 2	up to 2000
938		Austria		Austria	ISCED 4a, b	up to 2000
939		Austria		Austria	ISCED 3a, b	up to 2000
946		Austria		Austria	ISCED 3a, b	100001+
950		Austria		Austria	ISCED 3a, b	2001-10000

951	Austria	Former Yugoslavia	ISCED 3a, b	2001-10000
958	Austria	Austria	ISCED 3a, b	2001-10000
959	Austria	Austria	ISCED 4a, b	2001-10000
960	Austria	Austria	ISCED 5a	2001-10000
961	Austria	Austria	ISCED 3a, b	2001-10000
974	Austria	Austria	ISCED 4a, b	2001-10000
975	Austria	Austria	ISCED 4a, b	2001-10000
978	Austria	Austria	ISCED 2	up to 2000
981	Austria	Austria	ISCED 3a, b	up to 2000
993	Austria	Austria	ISCED 3a, b	2001-10000
998	Austria	Austria	ISCED 3a, b	up to 2000
999	Austria	Austria	ISCED 4a, b	up to 2000
1001	EU15 10 new members	EU15 10 new members	ISCED 6	2001-10000
1004	Austria	Austria	ISCED 5a	up to 2000
1005	Austria	Other countries	ISCED 5a	up to 2000
1009	Former Yugoslavia	Former Yugoslavia	ISCED 3a, b	up to 2000
1010	Former Yugoslavia	Former Yugoslavia	ISCED 3a, b	up to 2000

	xlfi	xpatch	is_outlier
10	Employed	No	FALSE
14	Employed	No	FALSE
15	Employed	No	FALSE
16	Employed	No	FALSE
21	Employed	No	FALSE
22	Employed	No	FALSE
39	Employed	No	FALSE
40	Employed	No	FALSE
41	Employed	No	FALSE
51	Employed	No	FALSE
58	Employed	No	FALSE
59	Employed	No	FALSE
68	Employed	No	FALSE
77	Employed	No	FALSE
80	Employed	No	FALSE
90	Employed	No	FALSE
94	Employed	No	FALSE
99	Employed	Yes	FALSE
100	Employed	Yes	FALSE
101	Employed	Yes	FALSE
102	Employed	No	FALSE
103	Employed	No	FALSE
105	Employed	No	FALSE
109	Employed	Yes	FALSE
110	Employed	Yes	FALSE

117	Employed	No	FALSE
118	Employed	No	FALSE
126	Employed	No	FALSE
127	Employed	No	FALSE
139	Employed	No	FALSE
140	Employed	No	FALSE
147	Employed	No	FALSE
148	Employed	No	FALSE
150	Employed	No	FALSE
155	Employed	No	FALSE
163	Employed	No	FALSE
166	Employed	No	FALSE
169	Employed	No	FALSE
170	Employed	Yes	FALSE
172	Employed	Yes	FALSE
173	Employed	Yes	FALSE
174	Employed	Yes	FALSE
175	Employed	Yes	FALSE
179	Employed	No	FALSE
185	Employed	No	FALSE
201	Employed	No	FALSE
212	Employed	No	FALSE
252	Employed	No	FALSE
253	Employed	No	FALSE
259	Employed	No	FALSE
260	Employed	No	FALSE
263	Employed	No	FALSE
278	Employed	Yes	FALSE
283	Employed	No	FALSE
288	Employed	No	FALSE
289	Employed	No	FALSE
298	Employed	Yes	FALSE
301	Employed	No	FALSE
302	Employed	No	FALSE
308	Employed	Yes	FALSE
309	Employed	Yes	FALSE
325	Employed	No	FALSE
326	Employed	No	FALSE
331	Employed	No	FALSE
335	Employed	No	FALSE
338	Employed	No	FALSE
342	Employed	No	FALSE
343	Employed	No	FALSE

344	Employed	No	FALSE
347	Employed	No	FALSE
350	Employed	No	FALSE
363	Employed	No	FALSE
368	Employed	Yes	FALSE
369	Employed	Yes	FALSE
383	Employed	No	FALSE
390	Employed	No	FALSE
391	Employed	No	FALSE
401	Employed	No	FALSE
403	Employed	No	FALSE
404	Employed	No	FALSE
405	Employed	No	FALSE
406	Employed	No	FALSE
407	Employed	No	FALSE
415	Employed	No	FALSE
429	Employed	No	FALSE
437	Employed	No	FALSE
438	Employed	No	FALSE
441	Employed	No	FALSE
445	Employed	No	FALSE
448	Employed	No	FALSE
450	Employed	No	FALSE
459	Employed	No	FALSE
480	Employed	No	FALSE
481	Employed	No	FALSE
482	Employed	No	FALSE
486	Employed	No	FALSE
487	Employed	No	FALSE
490	Employed	No	FALSE
495	Employed	No	FALSE
498	Employed	No	FALSE
499	Employed	No	FALSE
507	Employed	Yes	FALSE
539	Employed	No	FALSE
540	Employed	No	FALSE
541	Employed	No	FALSE
542	Employed	No	FALSE
557	Employed	No	FALSE
561	Employed	No	FALSE
562	Employed	No	FALSE
568	Employed	Yes	FALSE
569	Employed	Yes	FALSE

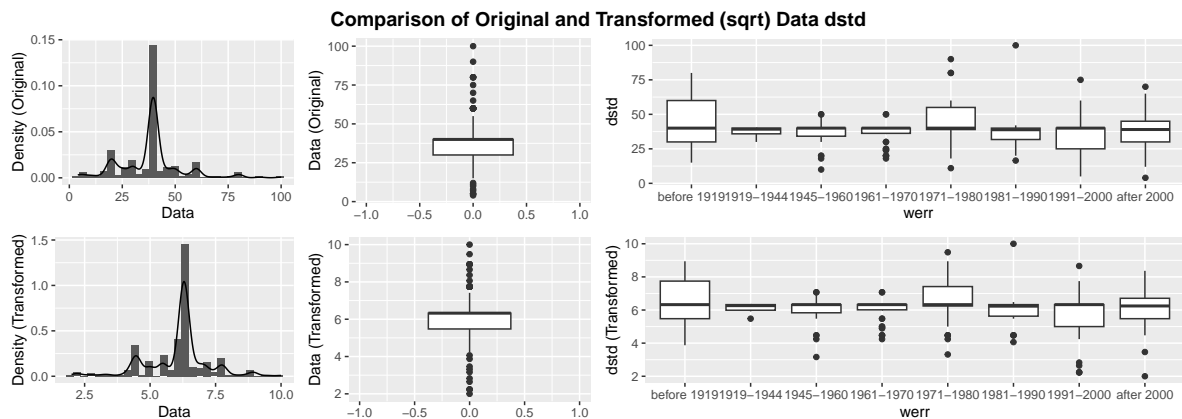
577	Employed	No	FALSE
588	Employed	No	FALSE
589	Employed	No	FALSE
595	Employed	No	FALSE
596	Employed	No	FALSE
599	Employed	No	FALSE
603	Employed	No	FALSE
604	Employed	No	FALSE
608	Employed	No	FALSE
609	Employed	No	FALSE
616	Employed	No	FALSE
625	Employed	Yes	FALSE
626	Employed	Yes	FALSE
633	Employed	No	FALSE
639	Employed	No	FALSE
640	Employed	No	FALSE
643	Employed	No	FALSE
656	Employed	No	FALSE
657	Employed	No	FALSE
661	Employed	No	FALSE
673	Employed	No	FALSE
674	Employed	No	FALSE
679	Employed	No	FALSE
693	Employed	No	FALSE
694	Employed	No	FALSE
700	Employed	No	FALSE
702	Employed	No	FALSE
704	Employed	No	FALSE
705	Employed	No	FALSE
707	Employed	No	FALSE
708	Employed	No	FALSE
717	Employed	Yes	FALSE
718	Employed	Yes	FALSE
719	Employed	Yes	FALSE
723	Employed	No	FALSE
724	Employed	No	FALSE
725	Employed	No	FALSE
726	Employed	No	FALSE
731	Employed	No	FALSE
742	Employed	No	FALSE
743	Employed	No	FALSE
744	Employed	No	FALSE
747	Employed	No	FALSE



748	Employed	No	FALSE
752	Employed	No	FALSE
761	Employed	No	FALSE
764	Employed	No	FALSE
765	Employed	No	FALSE
778	Employed	No	FALSE
779	Employed	No	FALSE
781	Employed	No	FALSE
802	Employed	No	FALSE
803	Employed	No	FALSE
804	Employed	No	FALSE
805	Employed	No	FALSE
806	Employed	No	FALSE
812	Employed	No	FALSE
813	Employed	No	FALSE
823	Employed	No	FALSE
826	Employed	No	FALSE
827	Employed	No	FALSE
831	Employed	No	FALSE
833	Employed	No	FALSE
834	Employed	No	FALSE
852	Employed	No	FALSE
853	Employed	No	FALSE
861	Employed	No	FALSE
862	Employed	No	FALSE
865	Employed	No	FALSE
866	Employed	No	FALSE
869	Employed	No	FALSE
876	Employed	Yes	FALSE
877	Employed	Yes	FALSE
880	Employed	No	FALSE
881	Employed	No	FALSE
882	Employed	No	FALSE
883	Employed	No	FALSE
885	Employed	No	FALSE
888	Employed	No	FALSE
898	Employed	No	FALSE
899	Employed	No	FALSE
900	Employed	No	FALSE
902	Employed	No	FALSE
903	Employed	No	FALSE
904	Employed	No	FALSE
907	Employed	No	FALSE

912	Employed	No	FALSE
923	Employed	No	FALSE
924	Employed	No	FALSE
929	Employed	No	FALSE
930	Employed	No	FALSE
938	Employed	No	FALSE
939	Employed	No	FALSE
946	Employed	No	FALSE
950	Employed	No	FALSE
951	Employed	No	FALSE
958	Employed	No	FALSE
959	Employed	No	FALSE
960	Employed	No	FALSE
961	Employed	No	FALSE
974	Employed	No	FALSE
975	Employed	No	FALSE
978	Employed	No	FALSE
981	Employed	No	FALSE
993	Employed	No	FALSE
998	Employed	No	FALSE
999	Employed	No	FALSE
1001	Employed	No	FALSE
1004	Employed	Yes	FALSE
1005	Employed	Yes	FALSE
1009	Employed	No	FALSE
1010	Employed	No	FALSE

```
plot_numeric_variable_with_transformation(data,"dstd","werr")
```



```
plot_numeric_variable_with_transformation(temp_data,"dstd","werr")
```



### 2.1.1 Introduction [20]

- Formulate research questions (see chapter 2.4.1 of the Meyer/Wurzer script) and research hypotheses, based on the predictors you selected
- Motivate the selection of the predictors - what was the reason for choosing them?
- Describe starting point and objectives of your analysis
- State the regression method you will use for your analysis (binary logistic, ordinal logistic, Poisson, . . . ) and justify your decision

2.1.2 Data collection [10]

- Type of survey; facts concerning the execution of the survey (period etc.)
- Description of the data set/operationalization (type of sample, sample size, variables, scale levels, missing values etc.)
- Data preparation (missing value treatment, transformations, . . . )

2.1.3 Descriptive analysis of the sample [70]

- Descriptive analysis of the analyzed variable(s) – Diagrams, numerical measures, tables, . . . – All statistics have to be commented, in particular diagrams! – Are there any distinctive features? (e.g., group differences, trends, outliers, . . . )
- In detail, the following plots have to be produced: – Univariate visualizations of all variables – Bivariate relationships between predictors and response to show the influence of the former on the latter – Joint influences of all possible pairs of predictors on the response to show potential interactions (exception: the interaction between the two metric variables doesn't have to be visualized)
- Summary of the descriptive analysis. Based on these descriptive findings, segue to the analysis of the questions about the population