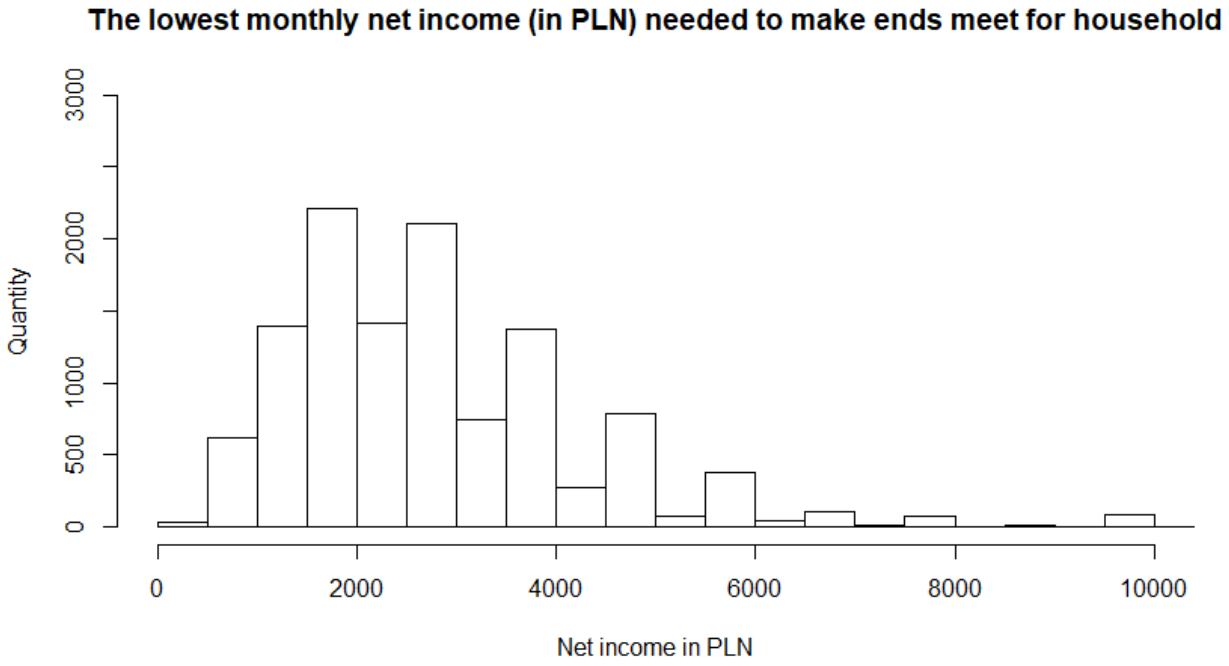
Bartek Podgórski

Grupa 2

175852

Ex.1

In the first task I used the variable representing the lowest monthly net income (in PLN) needed to make ends meet for household from year 2015.



Mean of variable is 3697.574 PLN.
Standard deviation is 10272.07 PLN(sample version).
Standard deviation is 8325.792 PLN(population version).

The table shows the quantiles.

0%	25%	50%	75%	100%
300	2000	2900	4000	99999

It's 5th percentile is 1000 PLN.

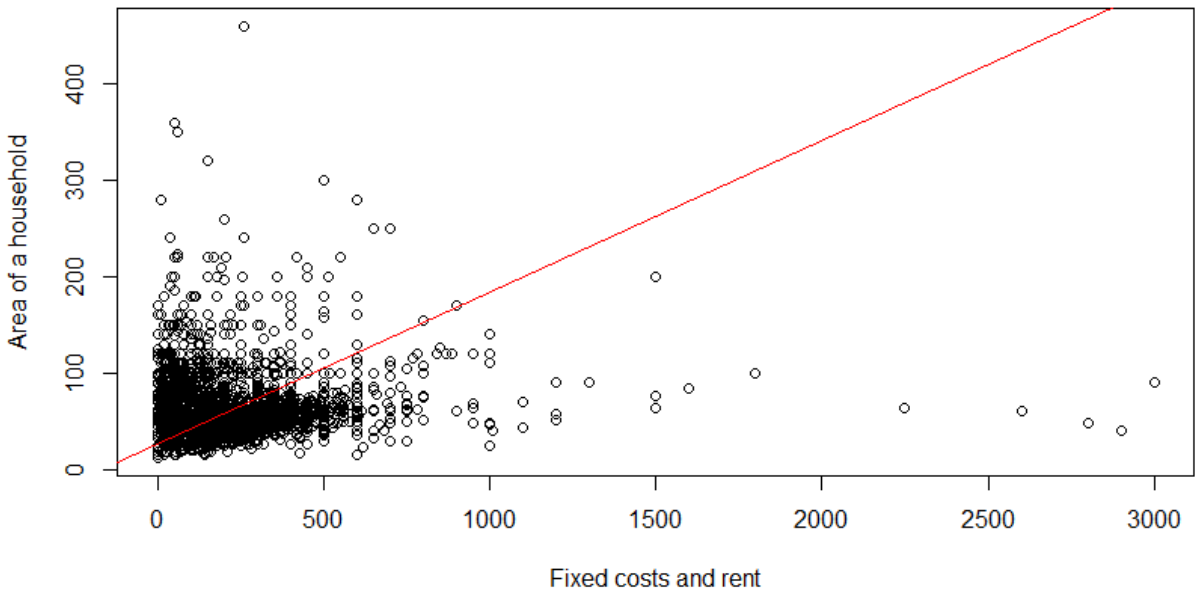
It’s 95th percentile is 6000 PLN.

It’s skewness is 12.78363. It is positive which indicates that the distribution is skewed to the right, it’s right tail is longer and most of the distribution is at the left . It means that it’s mean is greater then it’s mode.

It’s kurtosis is 144.8594 ,which tells us that the histogram in comparison with a histogram with normal distribution has longer and fatter tails and it’s central peak is higher and sharper.

Ex.2

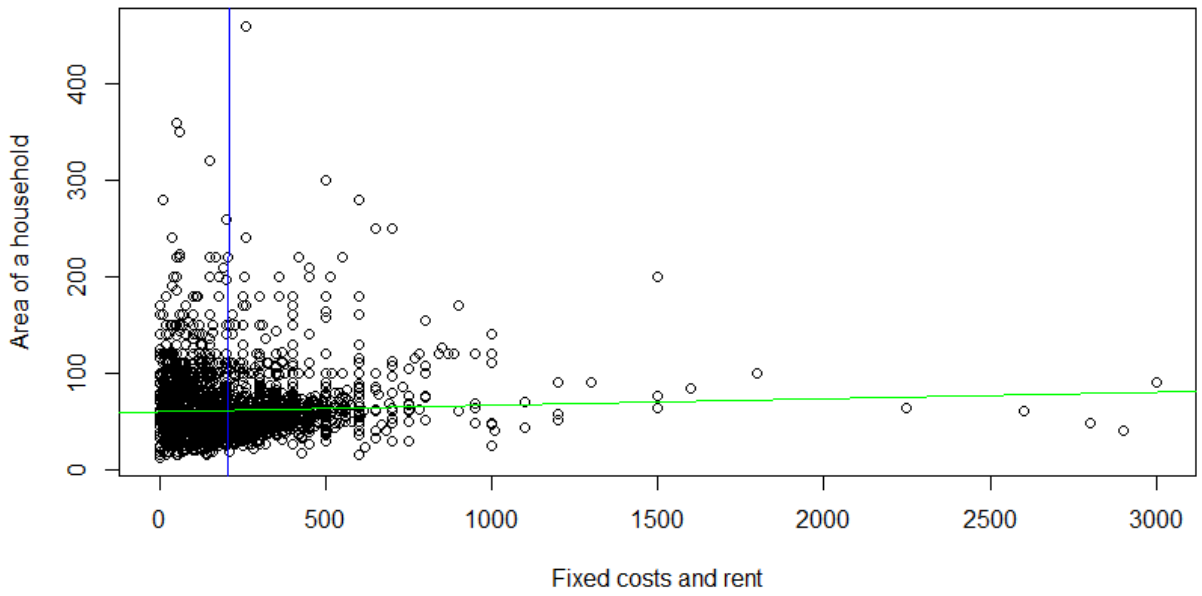
For this task I used two variables , it’s Area of a household(m^2) and fixed costs and rent(PLN) in the year 2000.



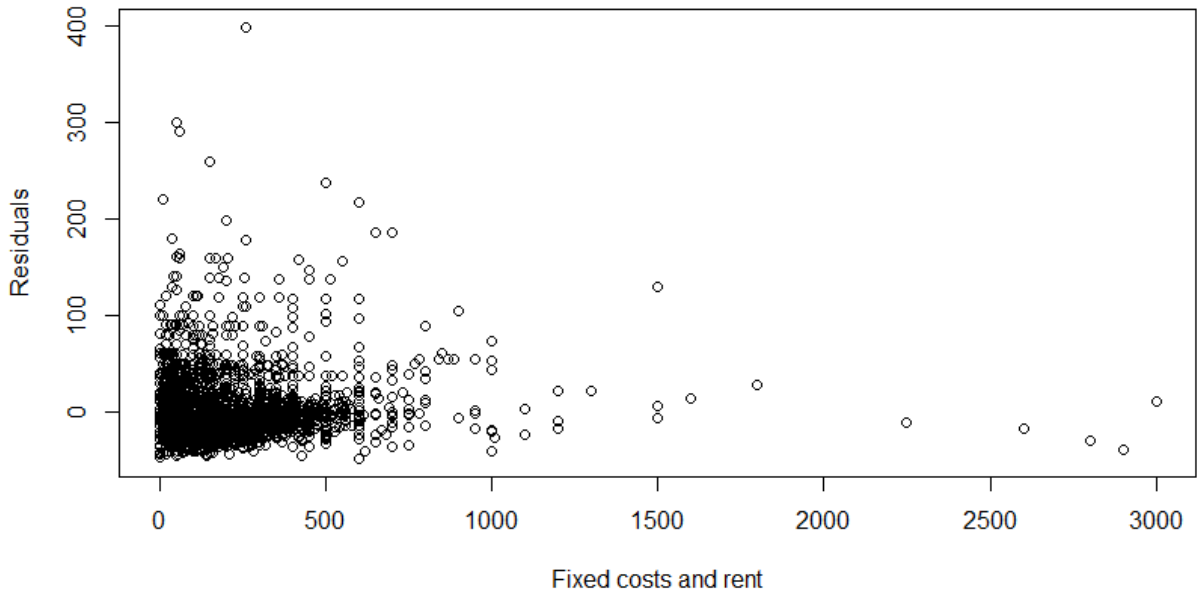
Correlation is 0.04262775, so we can tell that we have very small correlation between area of a household and costs. In sum , when we have bigger household we have to pay more but it is logical.

Ex.3

Scatter plot from task 2 with both regression lines.



Plot of residuals



Call:
lm(formula = ah8 ~ ah15_2, weights = waga_gd_2000)

Weighted Residuals:
Powierzchnia użytkowa całego mieszkania
   Min   1Q Median   3Q   Max
-101.09 -15.38   -3.79   11.87 341.89

Coefficients:
             Estimate Std. Error t value
(Intercept) 59.960658   0.895291   66.97
ah15_2   0.006699   0.002938   2.28
            Pr(>|t|)
(Intercept)   <2e-16 ***
ah15_2   0.0227 *

Signif. codes:
 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
 ‘ ’ 1

Residual standard error: 32.55 on 2855 degrees of freedom
Multiple R-squared: 0.001817, Adjusted R-squared: 0.001467
F-statistic: 5.197 on 1 and 2855 DF, p-value: 0.02269

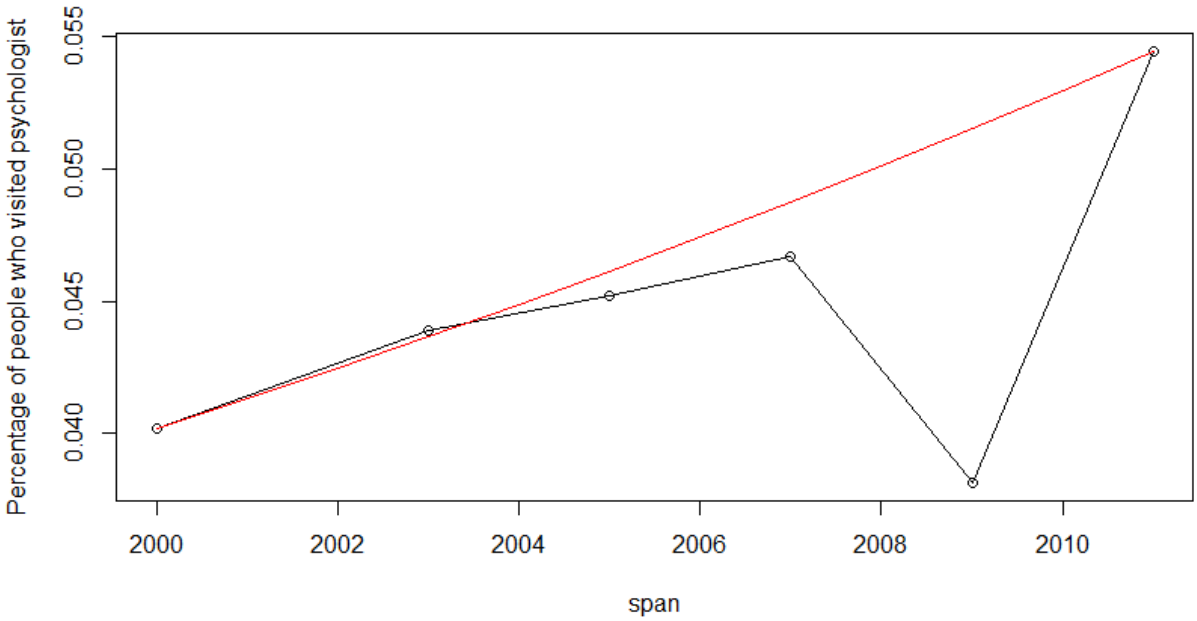
The regression slope of y on x equals 0.006699 , which means that for every 1PLN increase in area of household is 0.006699m^2.

The R 2 equals : 0.001817 which means that about 0,18% of the variance found in the variable Area (whole usable area of the flat) can be explained by the variable Cost(Fixed Cost).

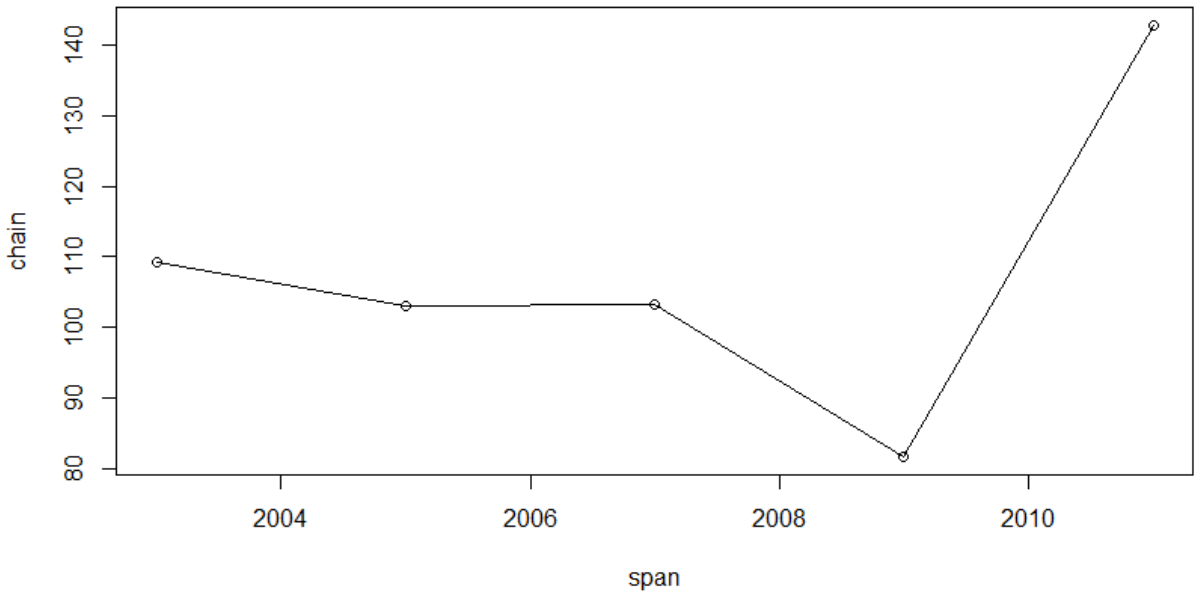
Ex.4

In this task I took question Did you lose relevant?

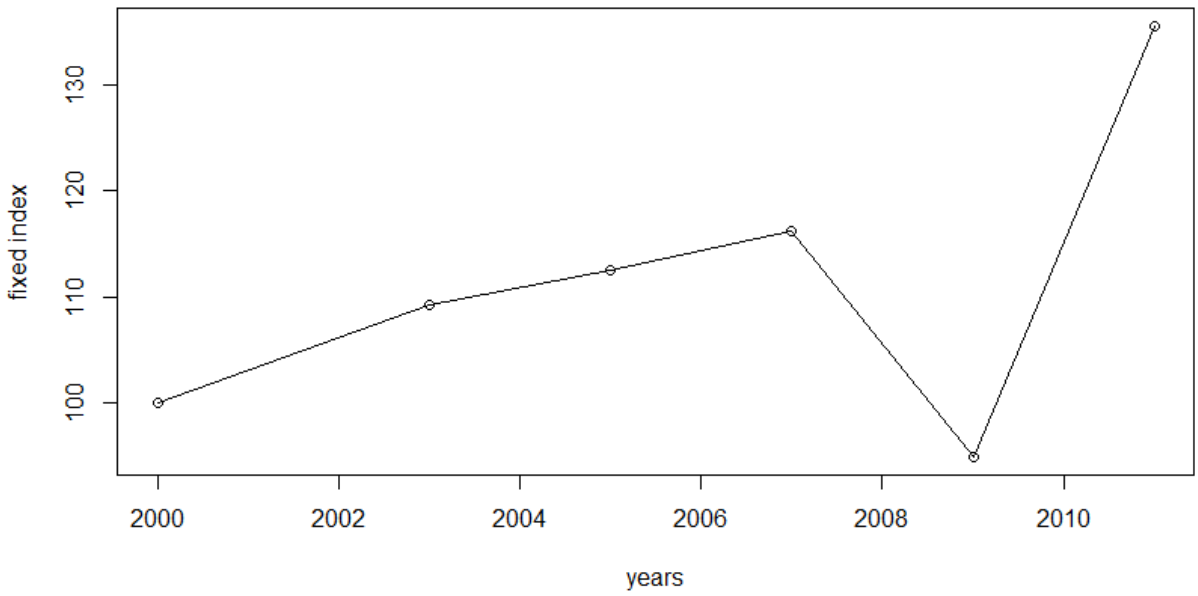
CAGR



Chain Index



Fixed Index



Number of people who participated in the therapy grew by 30% compared to number od people in 2000. It also shows that only small part of society needed such thing, but number of these people is growing, not by any constant value, but it is growing nontheless. Biggest grow occured from 2009 to 2011 where ammount of people who said yes grew by around 50%, right after the biggest drop which occured from 2007 to 2009 and was around 20%.

Value of CAGR: 3%

Ammount of people who participate in therapy grows on yearly basis by 3% on average.