

---

Inteligencja Obliczeniowa  
Projekt 3 – Time Series Analysis

---

Streszczenie projektu:

**Celem projektu jest analiza danych statystycznych dotyczących średniej długości życia Polaków w latach 1950-2017r. oraz przedstawienie prognozy długości życia w kolejnych 50 latach przy wykorzystaniu programu Rstudio.**

**Głównym celem jest oszacowanie na podstawie danych statystycznych, kiedy średnia długość życia ma szansę osiągnąć 150 lat i zostanie zrealizowane marzenie Ś.P. Michaela Jacksona.**

W projekcie wykorzystano metody analizy danych typu TimeSeries na podstawie:

1. Time Series and Forecasting

<https://www.statmethods.net/advstats/timeseries.html>

2. A Complete Tutorial on Time Series Modeling in R

<https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/>

3. Welcome to a Little Book of R for Time Series!

<https://a-little-book-of-r-for-timeseries.readthedocs.io/en/latest/>

Źródło danych do analizy:

Portal Głównego Urzędu Statystycznego:

<http://stat.gov.pl/obszary-tematyczne/ludnosc/trwanie-zycia/trwanie-zycia-tablice,1,1.html>

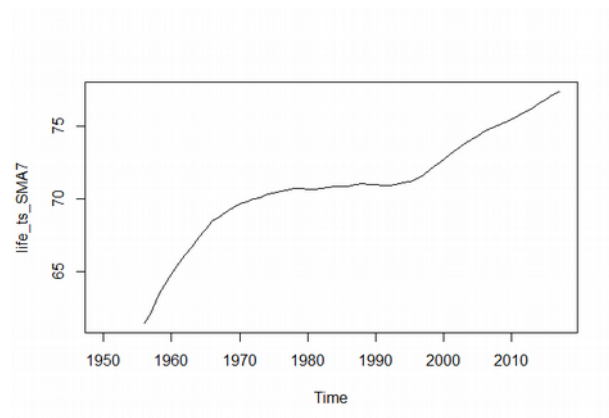
Link do wykorzystanej bazy danych:

<http://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5470/1/1/7/ex1950-2017.xls>

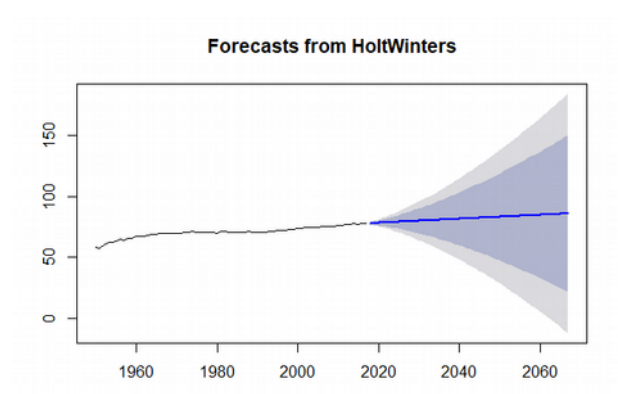
Repozytorium projektu:

[https://github.com/sslowik/Mgr\\_Inf/tree/master/Rstudio/Proj3](https://github.com/sslowik/Mgr_Inf/tree/master/Rstudio/Proj3)

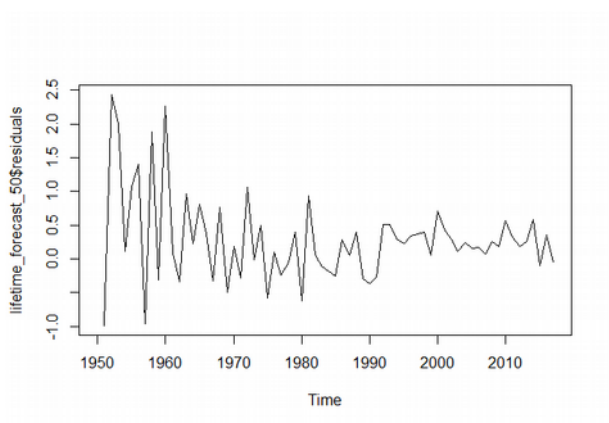
## I. Przykładowe wykresy:



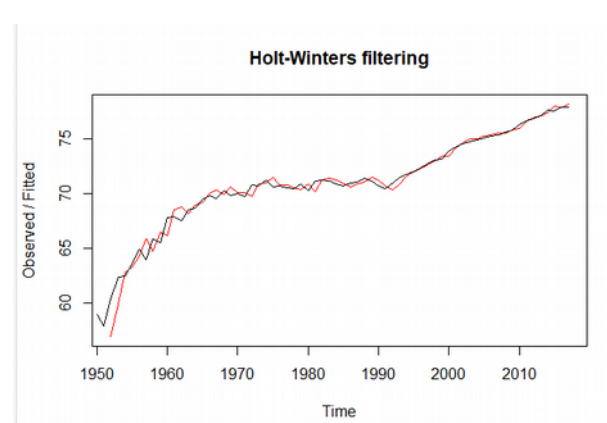
Rys. 1. Linia trendu dla danych z lat 1950-2017



Rys. 2. Linia trendu dla danych z lat 1950-2017



Rys. 3. Rozkład błędów prognozy



Rys. 4. Wartości prognozowane a rzeczywiste

## II. Wnioski z analizy:

1. Średnia długość życia w Polsce systematycznie rośnie. Tempo przyrostu spadło w latach 1970 – 1990, a od roku 1995 przyspieszyło i utrzymuje stałe tempo przyrostu na poziomie niespełna 3 lat na dekadę.
2. Marzenie Ś.P. Michaela Jacksona zostanie zrealizowane w 2067r. - zgodnie z prognozą maksymalna długość życia osiągnie wtedy 150 lat.

## III. Dane tabelaryczne:

Dane te zawierają dużo nieistotnych w gruncie rzeczy danych, więc po namyśle postanowiłem nie nabijać sztucznie ilości stron i odpuścić sobie ich umieszczenie w sprawozdaniu.

#### IV. Wykorzystany skrypt R:

```
getwd()
setwd("d:/Mgr_Inf/Rstudio/Proj3")

lifetime <- read.csv("lifetime_pl.csv", dec = ",", stringsAsFactors=FALSE)
average_lifetime <- as.vector(lifetime$Srednia)
average_lifetime

life_ts <- ts(average_lifetime, start=c(1950))
plot.ts(life_ts)

log_life_ts <- log(life_ts)

plot.ts(log_life_ts)

# smoothing to find trend

install.packages("TTR")
library("TTR")

life_ts_SMA3 <- SMA(life_ts, n=3)
plot.ts(life_ts_SMA3)

life_ts_SMA7 <- SMA(life_ts, n=7)
plot.ts(life_ts_SMA7)

# Forecasts using Exponential Smoothing

life_ts_forecasts <- HoltWinters(life_ts, beta=FALSE, gamma=FALSE)
life_ts_forecasts$SSE
life_ts_forecasts$fitted
plot(life_ts_forecasts)

#Installing library with forecast methods

install.packages("forecast")
library("forecast")

lifetime_forecast_50 <- forecast.HoltWinters(life_ts_forecasts, h=50)
lifetime_forecast_50

plot.forecast(lifetime_forecast_50)

Box.test(lifetime_forecast_50$residuals, lag=20, type="Ljung-Box")
plot.ts(lifetime_forecast_50$residuals)

# Holt's Exponential Smoothing

life_ts_forecasts2 <- HoltWinters(life_ts, gamma=FALSE)
life_ts_forecasts2
life_ts_forecasts2$SSE

plot(life_ts_forecasts2)

life_ts_forecasts2_50 <- forecast.HoltWinters(life_ts_forecasts2, h=50)
plot.forecast(life_ts_forecasts2_50)
```