

Akademia Górniczo-Hutnicza w Krakowie
Wydział Geologii, Geofizyki i Ochrony Środowiska



Sprawozdanie z Projektu 1
w ramach przedmiotu Przetwarzanie Danych Środowiskowych

autorzy: Maciej Bąk, Filip Giermek gr.1

Kraków/Olszyny, 08.04.2020

Przedmiot analizy

Przedmiotem analizy niniejszego projektu jest tabela danych *dynow_t* pochodzących ze stacji meteorologicznej w Dynowie, składająca się z wartości średniej temperatury dla poszczególnych miesięcy na przestrzeni lat 1951-2016.

Projekt 1 - cz.1

1.1. Wczytanie danych i bibliotek

Wczytujemy odpowiednie biblioteki, które będą wykorzystywane w analizie. Do wczytania, wykorzystujemy komendę *read.delim()*.

```
library(dplyr)

library(ggplot2)
library(tidyr)

dynow<-read.delim("dynow_t.txt")
View(dynow)
```

1.2. Brakujące wartości

W pakiecie danych *dynow_t* znajdują się brakujące wartości - NA. Zastępujemy je średnią arytmetyczną kolumn w których dane braki występowały. Realizujemy to z wykorzystaniem funkcji *which()* oraz *is.na()*. Ich odpowiednia kombinacja wyszukuje wartości NA w zadanych kolumnach. Następnie podmieniamy je średnią arytmetyczną obliczoną funkcją *mean()*. Proces powtarzany jest dla każdej kolumny (z wyłączeniem kolumny *rok*).

```
dynow$I[which(is.na(dynow$I))]<-mean(dynow$I, na.rm=TRUE)
dynow$II[which(is.na(dynow$II))]<-mean(dynow$II, na.rm=TRUE)
dynow$III[which(is.na(dynow$III))]<-mean(dynow$III, na.rm=TRUE)
dynow$IV[which(is.na(dynow$IV))]<-mean(dynow$IV, na.rm=TRUE)
dynow$V[which(is.na(dynow$V))]<-mean(dynow$V, na.rm=TRUE)
dynow$VI[which(is.na(dynow$VI))]<-mean(dynow$VI, na.rm=TRUE)
dynow$VII[which(is.na(dynow$VII))]<-mean(dynow$VII, na.rm=TRUE)
dynow$VIII[which(is.na(dynow$VIII))]<-mean(dynow$VIII, na.rm=TRUE)
dynow$IX[which(is.na(dynow$IX))]<-mean(dynow$IX, na.rm=TRUE)
dynow$X[which(is.na(dynow$X))]<-mean(dynow$X, na.rm=TRUE)
dynow$XI[which(is.na(dynow$XI))]<-mean(dynow$XI, na.rm=TRUE)
dynow$XII[which(is.na(dynow$XII))]<-mean(dynow$XII, na.rm=TRUE)
```

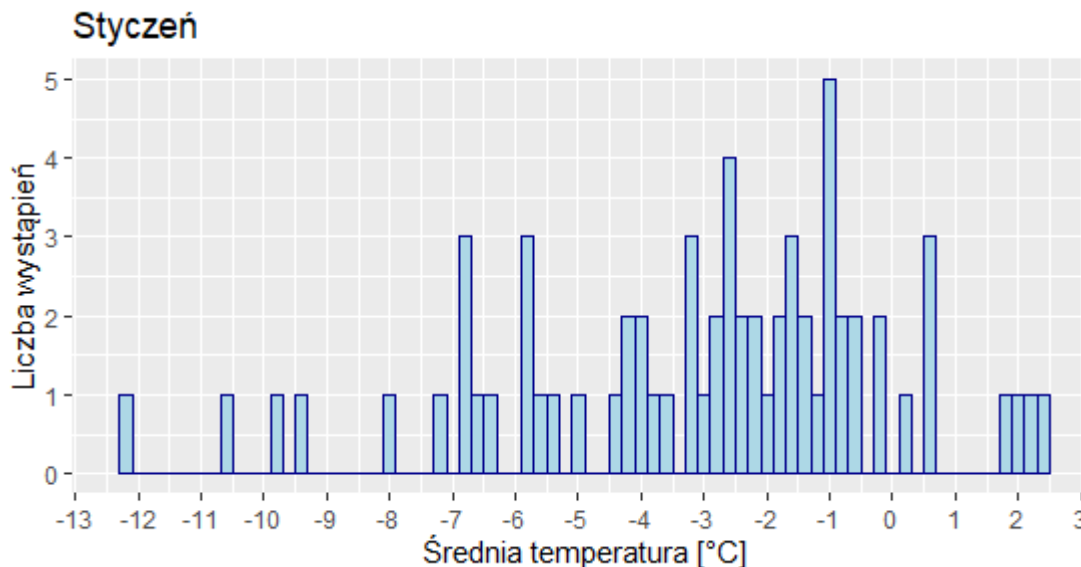
1.3. Rozkład średniej temperatury

W celu zobrazowania rozkładu temperatury w poszczególnych miesiącach na przestrzeni lat zastosujemy histogramy. Do rysowania wykorzystamy pakiet *ggplot2* oraz zawartą w nim funkcję *ggplot()* (gdzie wskazujemy zakres danych do wykresu) w połączeniu z *geom_histogram()*. Zmieniamy nazwy osi oraz dodajemy tytuł aby zachować czytelność histogramu. Ustawiamy też odpowiedni zakres dla osi - *scale_x/y_continuous()*. Funkcja *ggsave()* pozwala zapisać ostatni wygenerowany wykres do pliku.

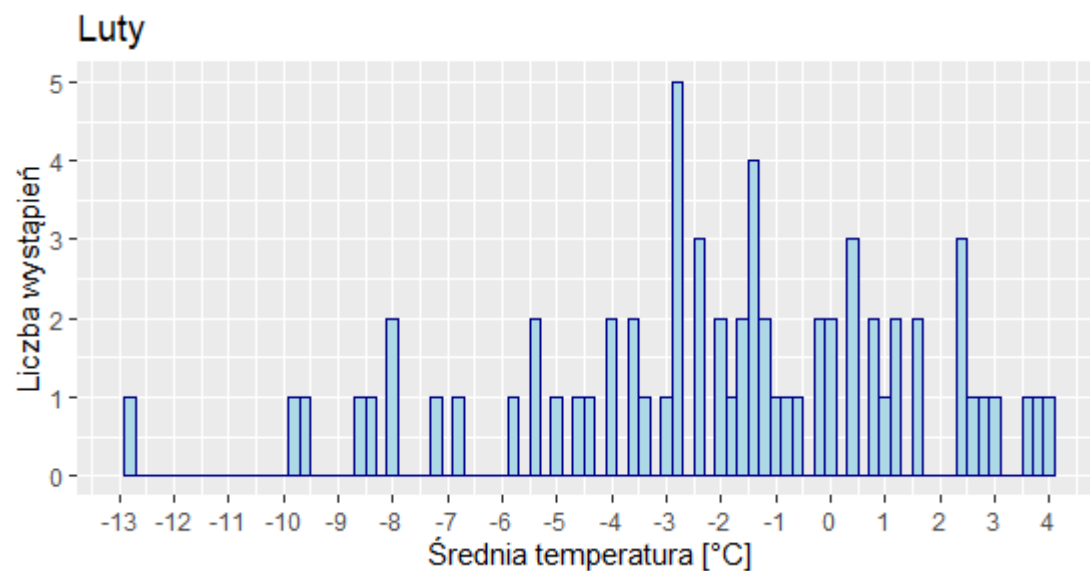
***Dla każdego wykresu w komendach zmienia się:**

aes('numer miesiąca'), ggtitle('nazwa miesiąca') oraz ggsave('nazwa pliku')

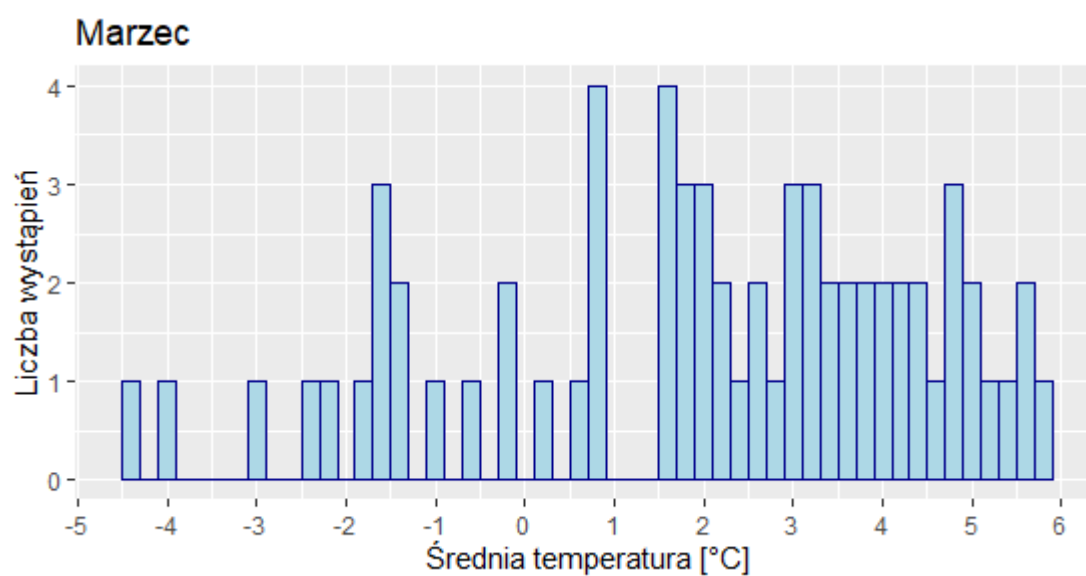
```
ggplot(dynow, aes(I))+  
  geom_histogram(color="darkblue", fill="lightblue", binwidth=0.2)+  
  xlab("Średnia temperatura [°C]")+  
  ylab("Liczba wystąpień")+  
  scale_x_continuous(breaks=seq(-13,20,by=1))+  
  scale_y_continuous(breaks=seq(0,30,by=1))+  
  ggtitle("Styczeń (wykr. 1.3.1)")  
  
ggsave("trend_I.png")
```



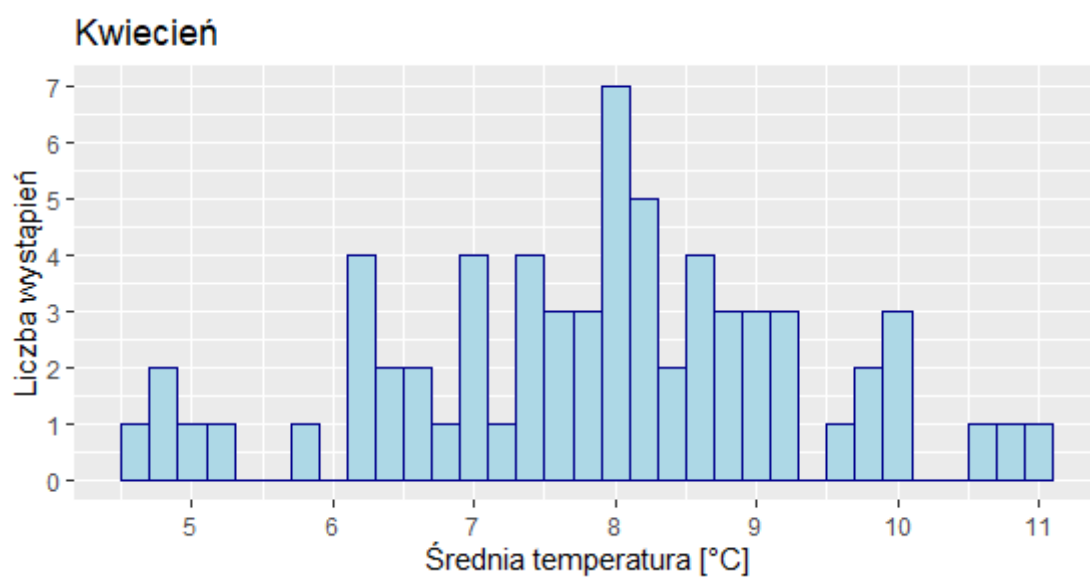
- wykr. 1.3.1



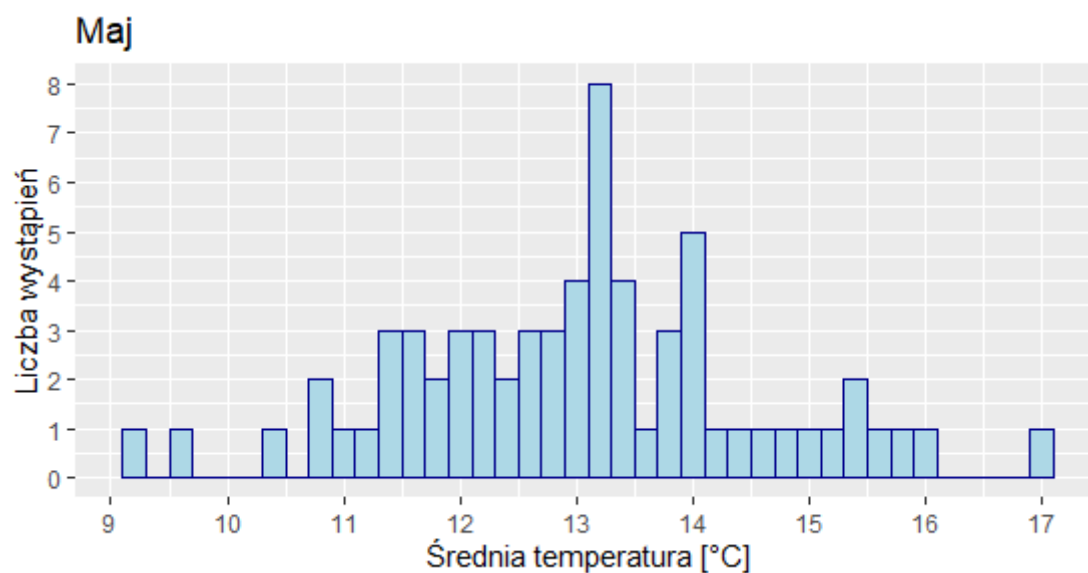
- wykr. 1.3.2



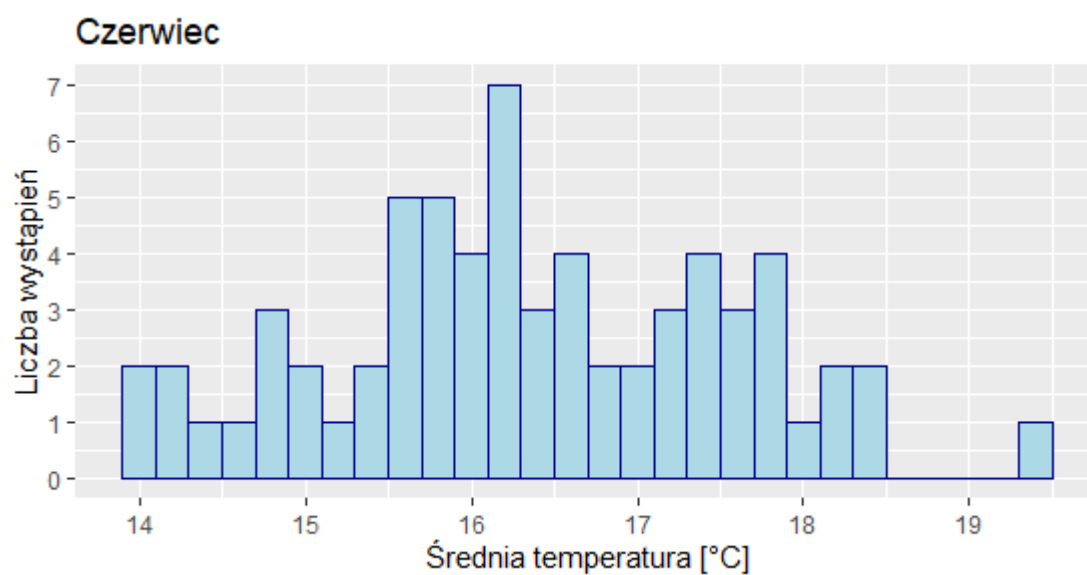
- wykr. 1.3.3



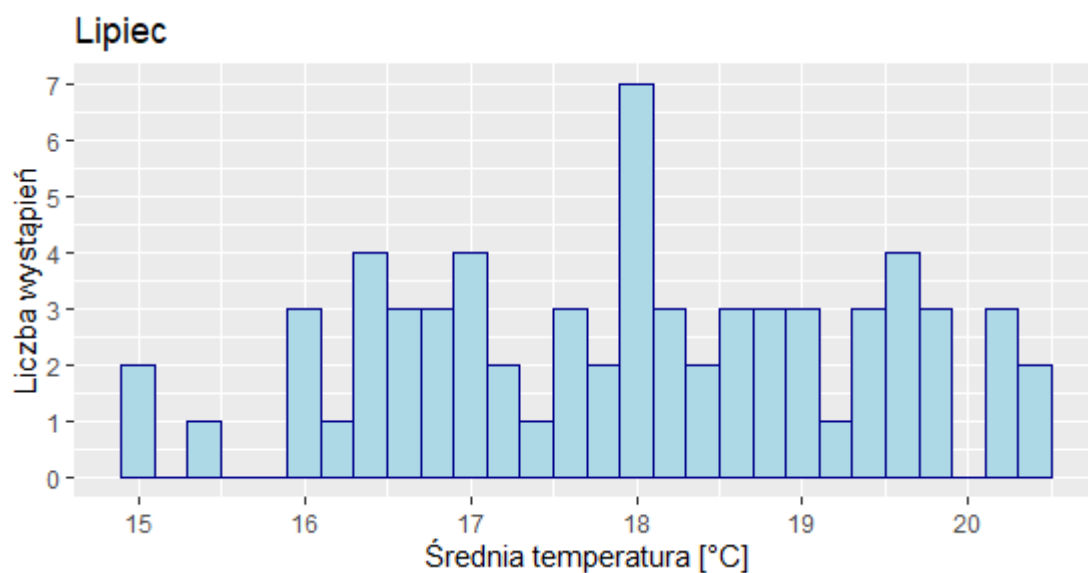
- wykr. 1.3.4



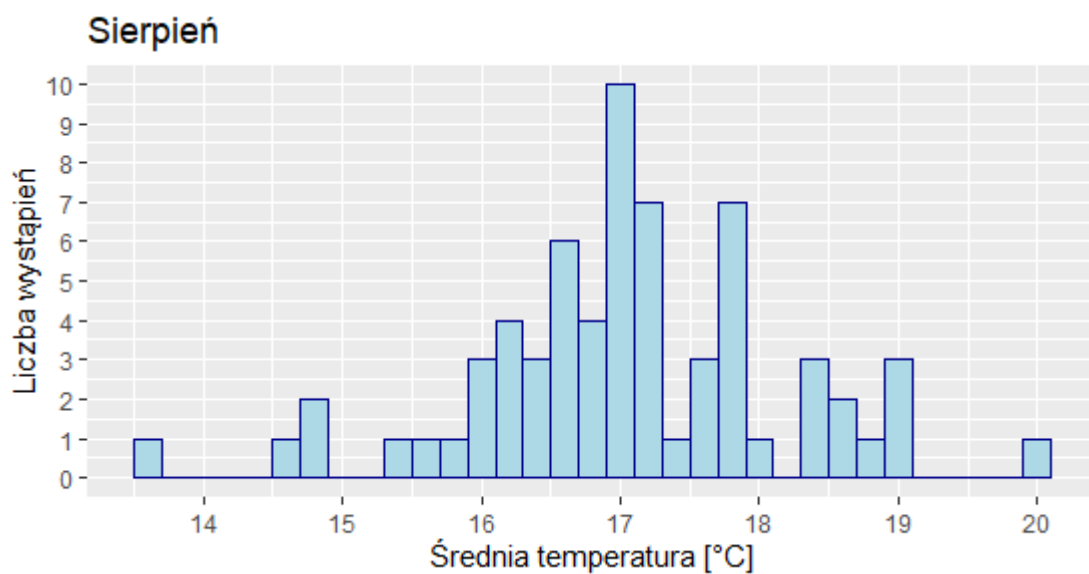
- wykr. 1.3.5



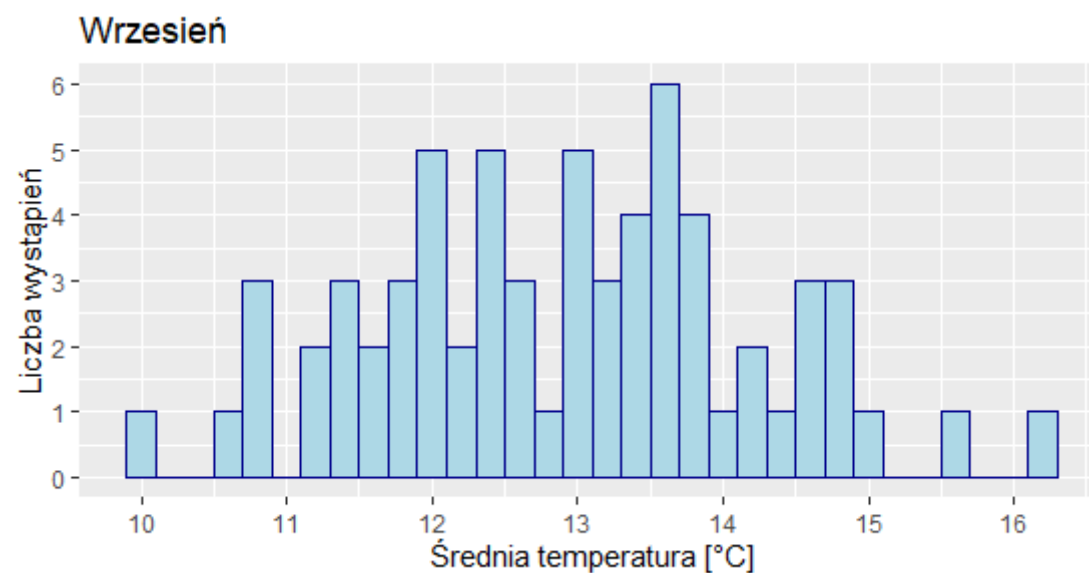
- wykr. 1.3.6



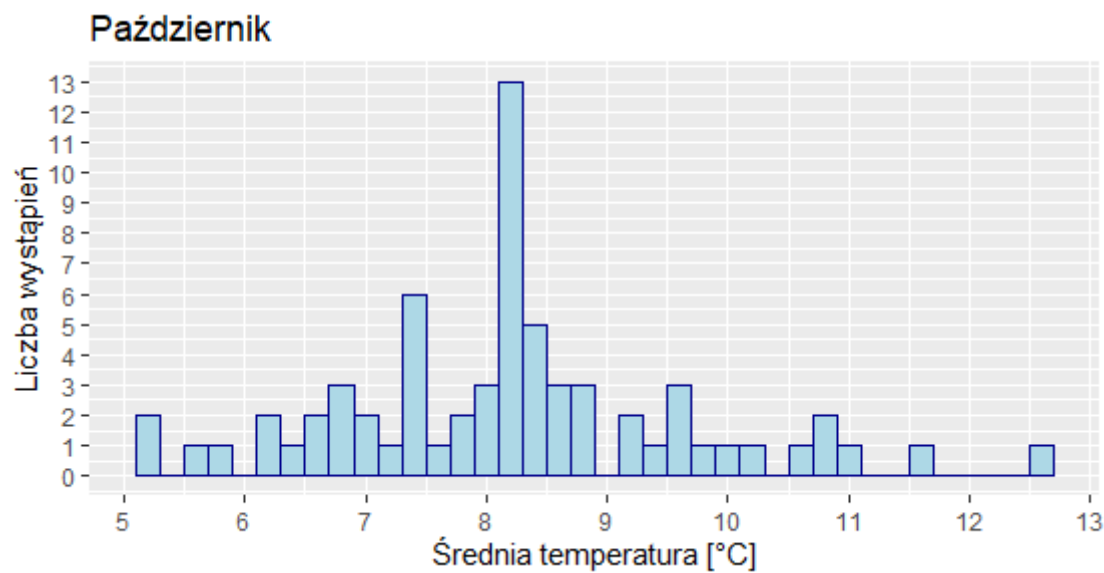
- wykr. 1.3.7



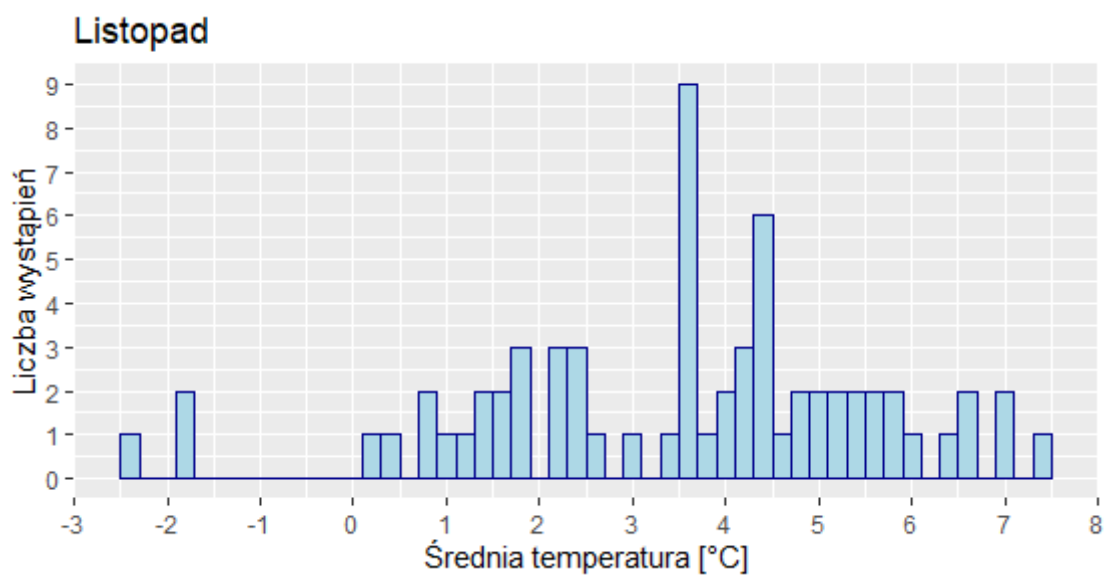
- wykr. 1.3.8



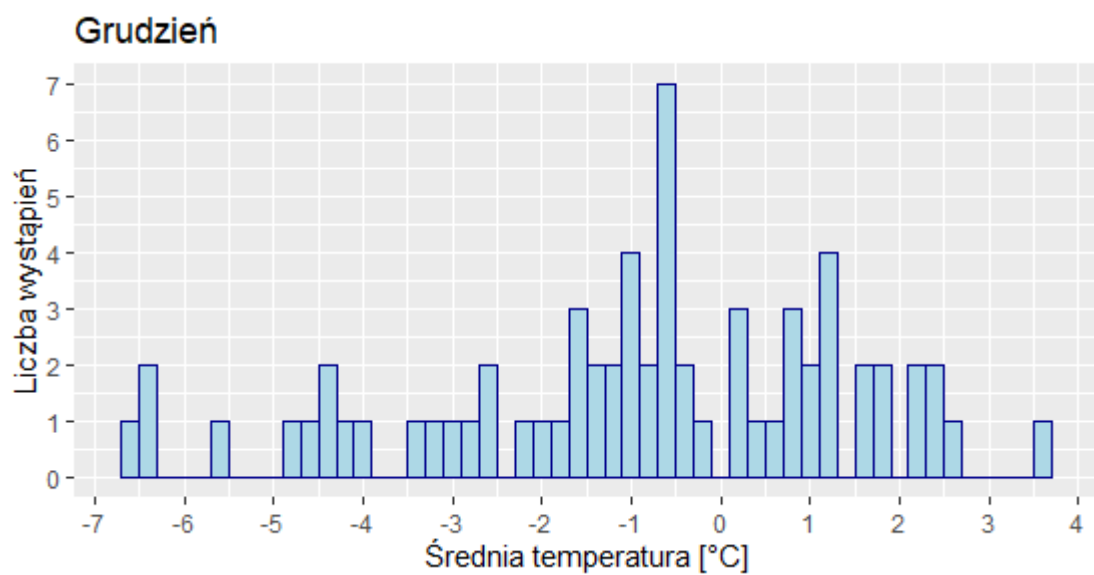
- wykr. 1.3.9



- wykr. 1.3.10



- wyk. 1.3.11



- wyk. 1.3.12

1.4. Statystyka opisowa

Aby poprawnie wykonać podstawową statystykę opisową dla średnich wartości temperatur dla poszczególnych miesięcy kolejnych lat stworzymy nowy *data.frame* - *dynow_stat*. Wcześniej jednak tworzymy wektory użyte do stworzenia *dynow_stat*. Liczymy kolejno wartość minimalną, maksymalną, średnią arytmetyczną, medianę, odchylenie standardowe oraz rozstęp międzykwartylowy.

```
miesiąc<-c("I","II","III","IV","V","VI","VII","VIII","IX","X","XI","XII")

min<-c(min(dynow$I, na.rm=TRUE),min(dynow$II, na.rm=TRUE),min(dynow$III, na.r
m=TRUE),min(dynow$IV, na.rm=TRUE),min(dynow$V, na.rm=TRUE),min(dynow$VI, na.r
m=TRUE),min(dynow$VII, na.rm=TRUE),min(dynow$VIII, na.rm=TRUE),min(dynow$IX,
na.rm=TRUE),min(dynow$X, na.rm=TRUE),min(dynow$XI, na.rm=TRUE),min(dynow$XII,
na.rm=TRUE))

max<-c(max(dynow$I, na.rm=TRUE),max(dynow$II, na.rm=TRUE),max(dynow$III, na.r
m=TRUE),max(dynow$IV, na.rm=TRUE),max(dynow$V, na.rm=TRUE),max(dynow$VI, na.r
m=TRUE),max(dynow$VII, na.rm=TRUE),max(dynow$VIII, na.rm=TRUE),max(dynow$IX,
na.rm=TRUE),max(dynow$X, na.rm=TRUE),max(dynow$XI, na.rm=TRUE),max(dynow$XII,
na.rm=TRUE))

średnia<-c(mean(dynow$I, na.rm=TRUE),mean(dynow$II, na.rm=TRUE),mean(dynow$III
I, na.rm=TRUE),mean(dynow$IV, na.rm=TRUE),mean(dynow$V, na.rm=TRUE),mean(dyno
w$VI, na.rm=TRUE),mean(dynow$VII, na.rm=TRUE),mean(dynow$VIII, na.rm=TRUE),me
an(dynow$IX, na.rm=TRUE),mean(dynow$X, na.rm=TRUE),mean(dynow$XI, na.rm=TRUE)
,mean(dynow$XII, na.rm=TRUE))

mediana<-c(median(dynow$I, na.rm=TRUE),median(dynow$II, na.rm=TRUE),median(dy
now$III, na.rm=TRUE),median(dynow$IV, na.rm=TRUE),median(dynow$V, na.rm=TRUE)
,median(dynow$VI, na.rm=TRUE),median(dynow$VII, na.rm=TRUE),median(dynow$VIII
, na.rm=TRUE),median(dynow$IX, na.rm=TRUE),median(dynow$X, na.rm=TRUE),median
(dynow$XI, na.rm=TRUE),median(dynow$XII, na.rm=TRUE))

odchylenie_standardowe<-c(sd(dynow$I, na.rm=TRUE),sd(dynow$II, na.rm=TRUE),sd
(dynow$III, na.rm=TRUE),sd(dynow$IV, na.rm=TRUE),sd(dynow$V, na.rm=TRUE),sd(d
ynow$VI, na.rm=TRUE),sd(dynow$VII, na.rm=TRUE),sd(dynow$VIII, na.rm=TRUE),sd(
dynow$IX, na.rm=TRUE),sd(dynow$X, na.rm=TRUE),sd(dynow$XI, na.rm=TRUE),sd(dyn
ow$XII, na.rm=TRUE))

rozstęp_międzykwartylowy<-c(IQR(dynow$I, na.rm=TRUE),IQR(dynow$II, na.rm=TRUE
),IQR(dynow$III, na.rm=TRUE),IQR(dynow$IV, na.rm=TRUE),IQR(dynow$V, na.rm=TRU
E),IQR(dynow$VI, na.rm=TRUE),IQR(dynow$VII, na.rm=TRUE),IQR(dynow$VIII, na.rm
=TRUE),IQR(dynow$IX, na.rm=TRUE),IQR(dynow$X, na.rm=TRUE),IQR(dynow$XI, na.rm
=TRUE),IQR(dynow$XII, na.rm=TRUE))

dynow_stat<-data.frame(miesiąc,min,max,średnia,mediana,odchylenie_standardowe
,rozstęp_międzykwartylowy)
```


1.5. Eksport danych

Stworzoną ramkę eksportujemy do tabeli txt za pomocą funkcji `write.table()` i następnie importujemy do arkusza Excel.

```
write.table(dynow_stat, file="dynow_stat.txt")
```

Powstała tabela statystyk opisowych:

miesiąc	min	max	zakres	średnia	mediana	odchylenie standardowe	rozstęp międzykwartyłowy
I	-12.2	2.4	-12.2 : 2.4	-3.02	-2.55	3.1	3.75
II	-12.9	4.1	-12.9 : 4.1	-2.01	-1.55	3.64	4.325
III	-4.5	5.8	-4.5 : 5.8	2.04	2.4	2.54	3.325
IV	4.6	11	4.6 : 11	7.92	8	1.5	1.9
V	9.2	17	9.2 : 17	13.06	13.15	1.52	1.875
VI	13.9	19.5	13.9 : 20.5	16.34	16.3	1.21	1.675
VII	14.9	20.5	14.9 : 20.5	18.04	18.05	1.41	2.175
VIII	13.6	20	13.6 : 20	17.01	17.1	1.13	1.275
IX	9.9	16.2	9.9 : 16.2	12.96	13.05	1.31	1.775
X	5.2	12.6	5.2 : 12.6	8.26	8.26	1.5	1.4
XI	-2.4	7.5	-2.4 : 7.5	3.54	3.7	2.13	2.75
XII	-6.7	3.7	-6.7 : 3.7	-0.91	-0.6	2.36	2.95

1.6. Utworzenie wykresów liniowych z trendem dla średniej temperatury w miesiącu na przestrzeni lat

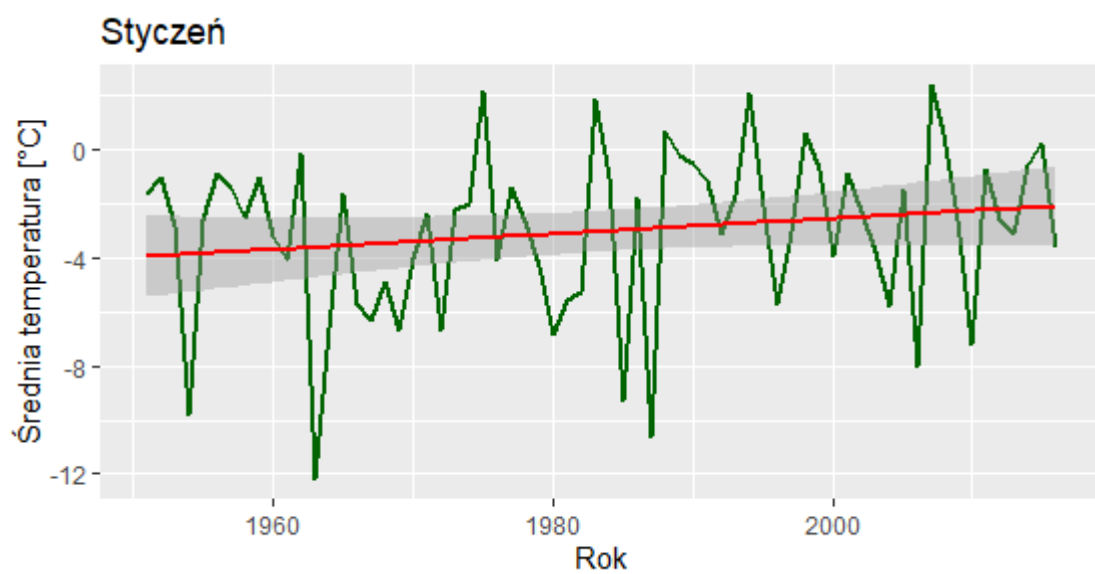
Do stworzenia wykresów liniowych również wykorzystujemy pakiet `ggplot2` oraz zawartą w nim funkcję `ggplot()` oraz `geom_line` - wykres liniowy. Aby dodać trend liniowy stosujemy funkcję `geom_smooth()`. Dodatkowo uzupełniamy wykres pozostałymi informacjami.

***Dla każdego wykresu w komendach zmieniać się:**

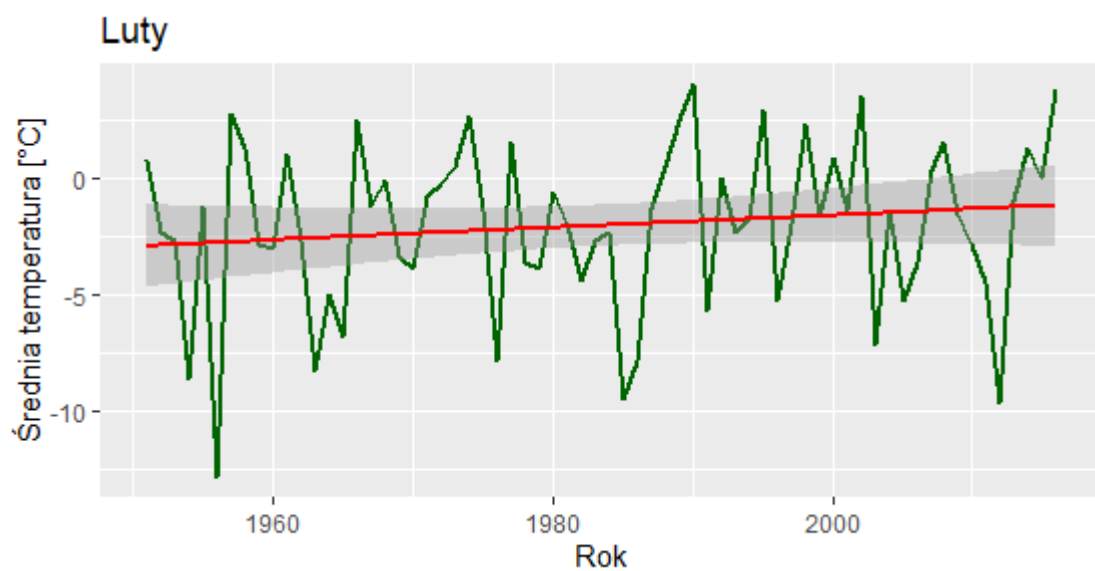
`aes('numer miesiąca')`, `ggtitle('nazwa miesiąca')` oraz `ggsave('nazwa pliku')`

```
library(ggplot2)
ggplot(dynow, aes(rok, I))+
  geom_line(color="darkgreen", size=1)+
  geom_smooth(method="lm", color="red")+
  xlab("Rok")+
  ylab("Średnia temperatura [°C]")+
  ggtitle("Styczeń")

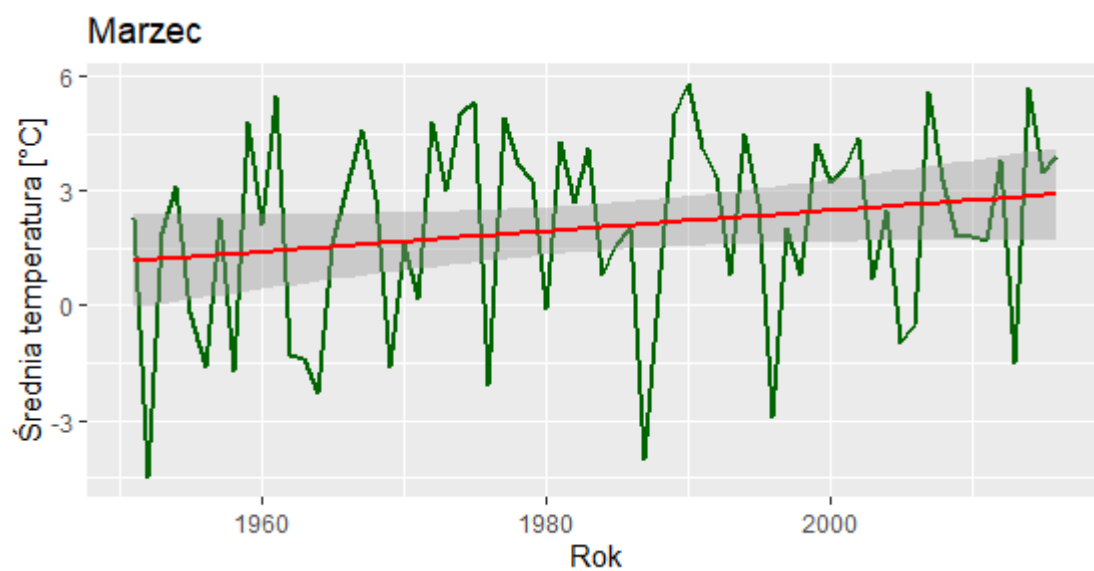
ggsave("trend_I.png")
```



- wykr. 1.6.1

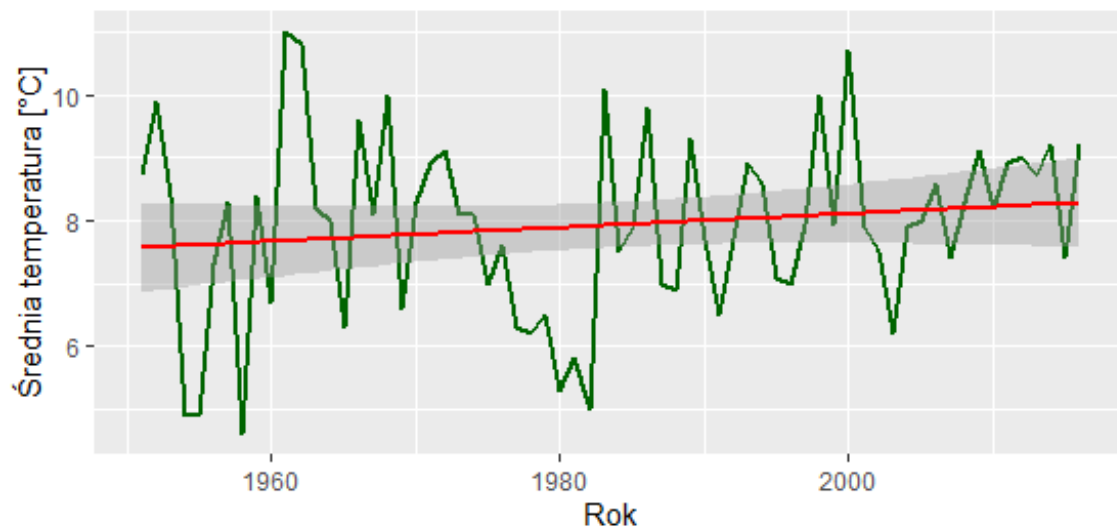


- wykr. 1.6.2



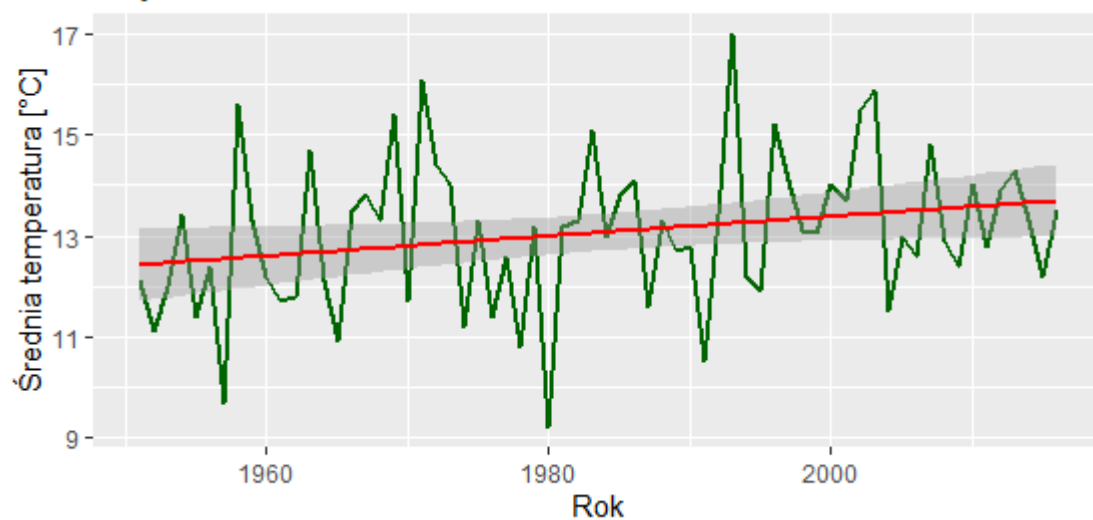
- wykr. 1.6.3

Kwiecień



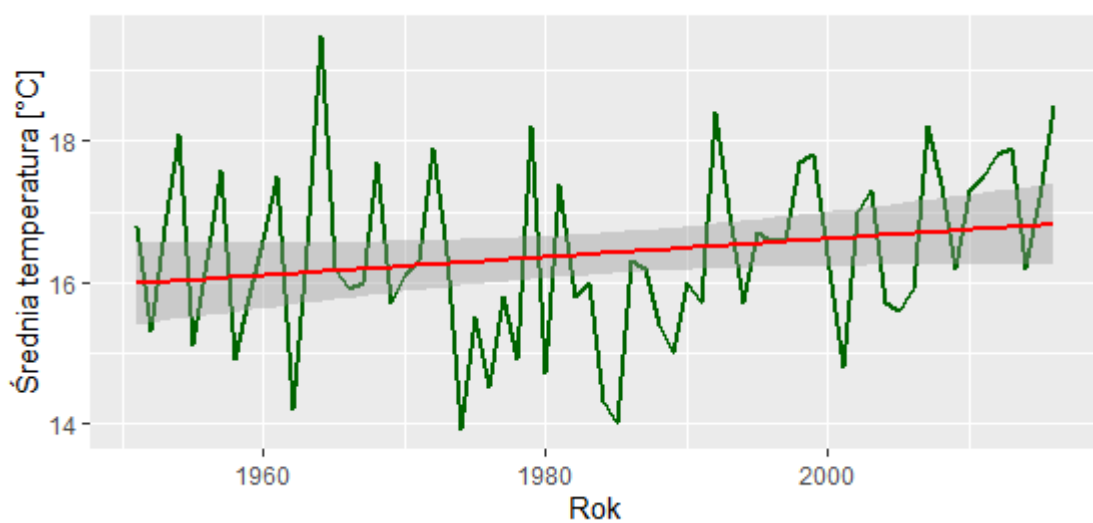
- wykr. 1.6.4

Maj

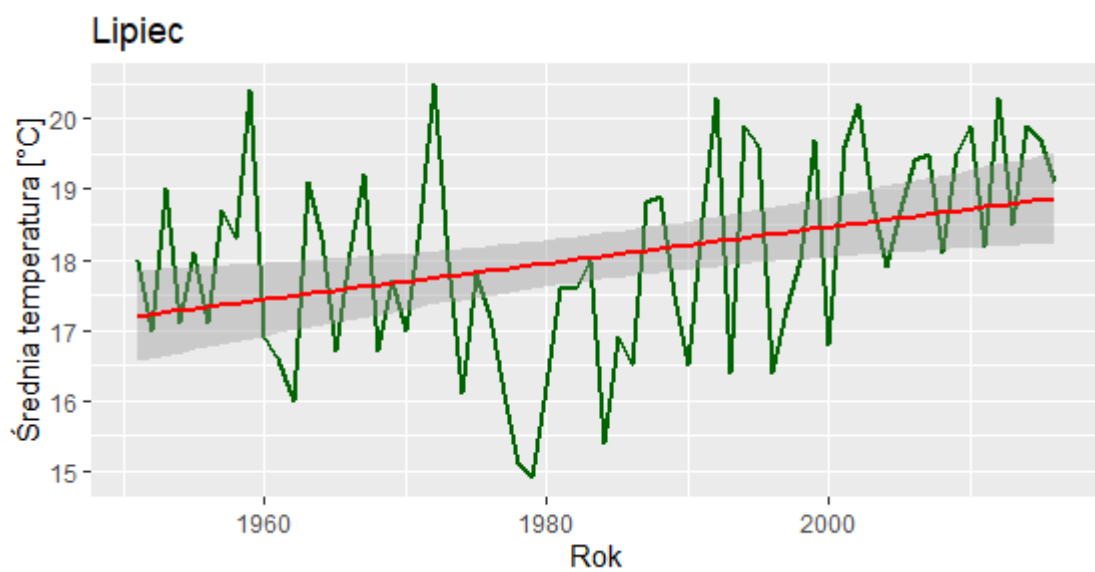


- wykr. 1.6.5

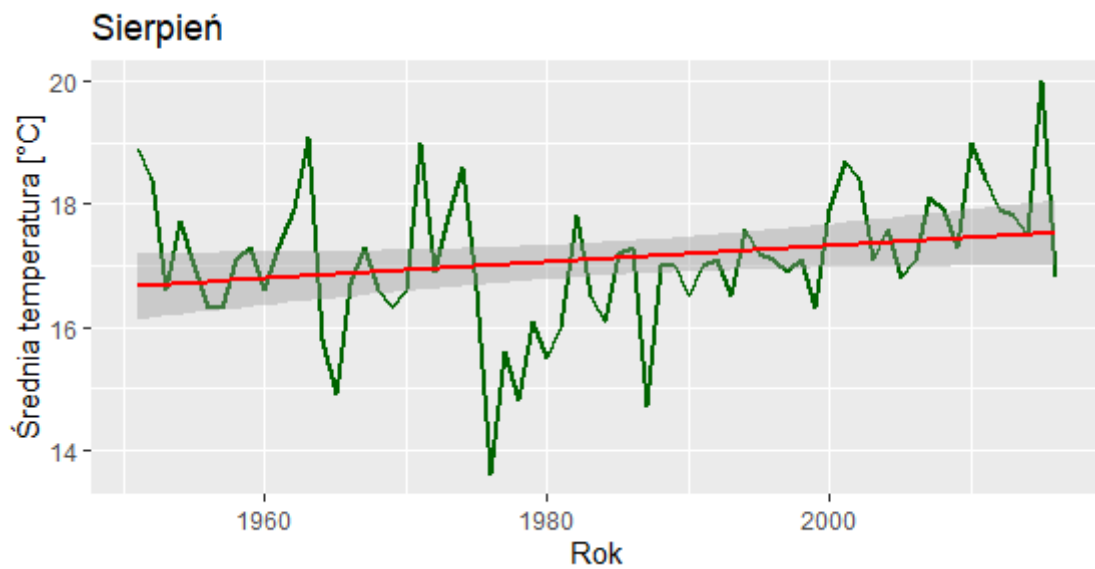
Czerwiec



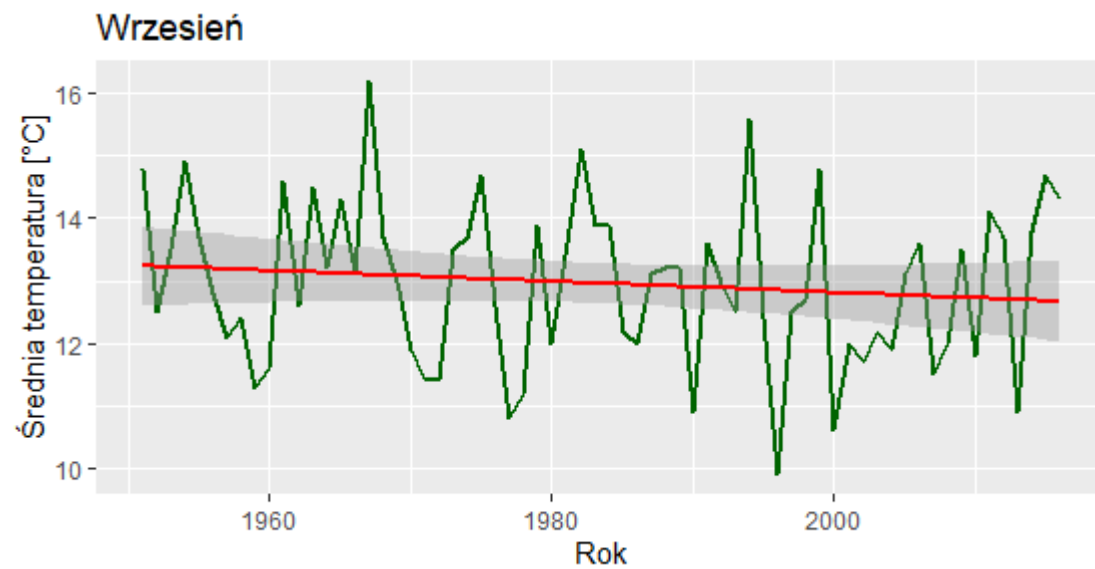
- wykr. 1.6.6



- wyk. 1.6.7

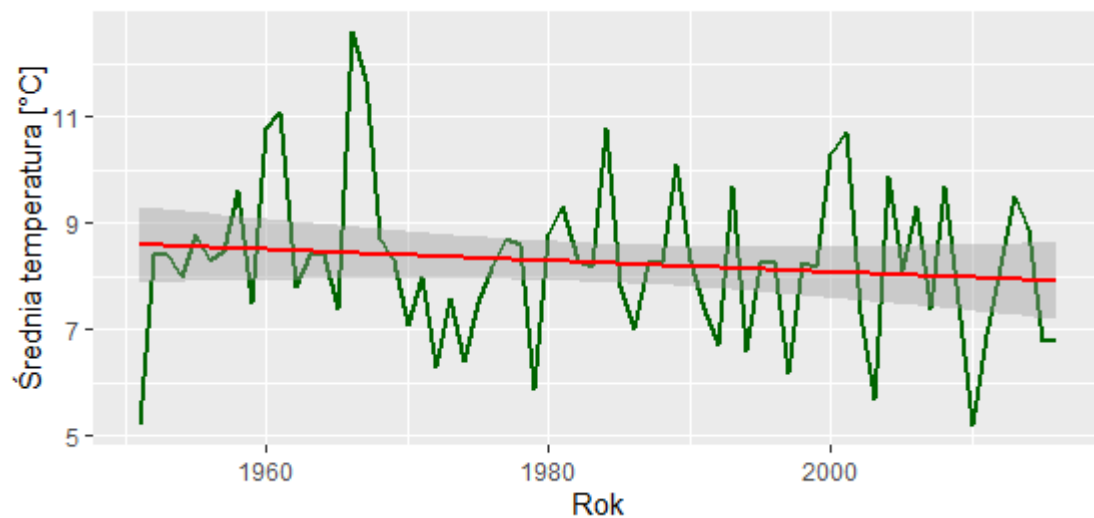


- wyk. 1.6.8



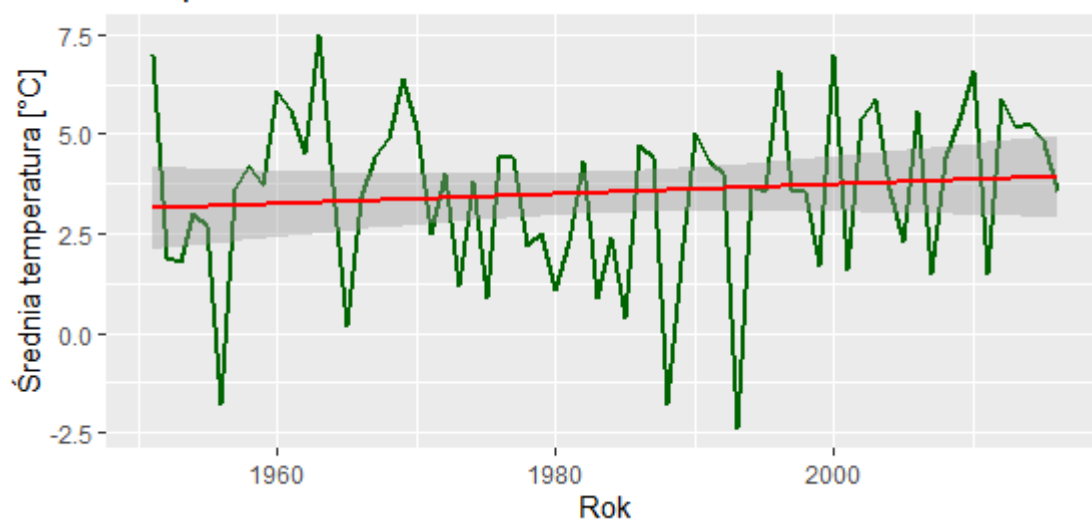
- wyk. 1.6.9

Październik



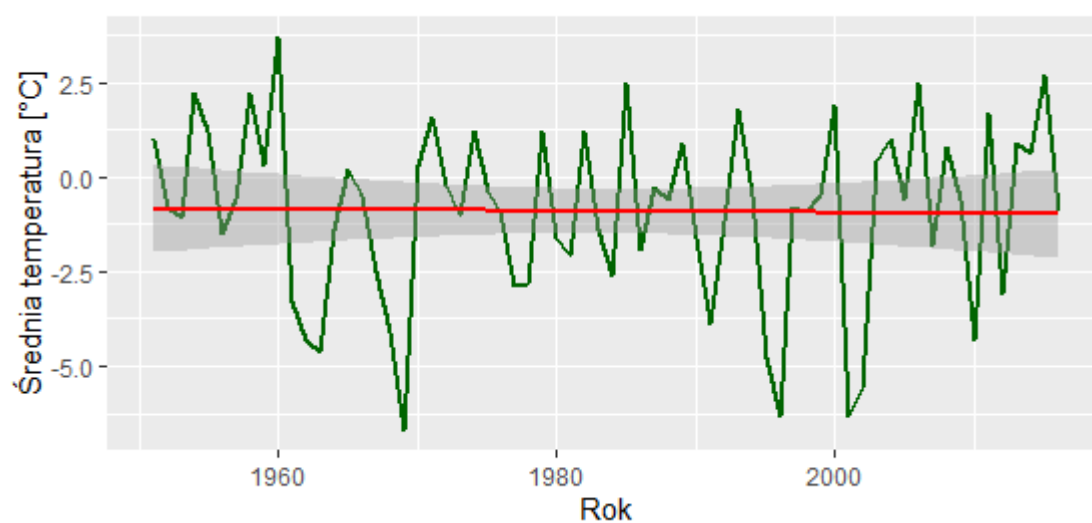
- wykr. 1.6.10

Listopad



- wykr. 1.6.11

Grudzień



- wykr. 1.6.12

Projekt 1 - cz.2

2.1. Wczytanie danych i bibliotek

Podobnie jak w punkcie 1.1 wczytujemy potrzebne biblioteki i pakiet danych.

```
library(dplyr)
library(ggplot2)
library(tidyr)
dynow_t <- read.delim("dynow_t.txt")
```

2.2. Przygotowanie danych

Dzielimy pakiet danych *dynow_t* na mniejsze *data.frame* obejmujące poszczególne pory roku oraz okres wegetacyjny (V-IX) z wykorzystaniem komendy *select()*. Przypisujemy je do nowych zmiennych.

```
wiosna <- select(dynow_t, rok, III, IV, V)
lato <- select(dynow_t, rok, VI, VII, VIII)
jesien <- select(dynow_t, rok, IX, X, XI)
zima <- select(dynow_t, rok, XII, I, II)
ow <- select(dynow_t, rok, V, VI, VII, VIII, IX)
```

2.3. Brakujące wartości

Podobnie jak w punkcie 1.2 zastępujemy nieznane wartości średnią z całej kolumny.

```
wiosna$III[which(is.na(wiosna$III))] <- mean(wiosna$III, na.rm=T)
wiosna$IV[which(is.na(wiosna$IV))] <- mean(wiosna$IV, na.rm=T)
wiosna$V[which(is.na(wiosna$V))] <- mean(wiosna$V, na.rm=T)
lato$VI[which(is.na(lato$VI))] <- mean(lato$VI, na.rm=T)
lato$VII[which(is.na(lato$VII))] <- mean(lato$VII, na.rm=T)
lato$VIII[which(is.na(lato$VIII))] <- mean(lato$VIII, na.rm=T)
jesien$IX[which(is.na(jesien$IX))] <- mean(jesien$IX, na.rm=T)
jesien$X[which(is.na(jesien$X))] <- mean(jesien$X, na.rm=T)
jesien$XI[which(is.na(jesien$XI))] <- mean(jesien$XI, na.rm=T)
zima$XII[which(is.na(zima$XII))] <- mean(zima$XII, na.rm=T)
zima$I[which(is.na(zima$I))] <- mean(zima$I, na.rm=T)
zima$II[which(is.na(zima$II))] <- mean(zima$II, na.rm=T)
ow$V[which(is.na(ow$V))] <- mean(ow$V, na.rm=T)
ow$VI[which(is.na(ow$VI))] <- mean(ow$VI, na.rm=T)
ow$VII[which(is.na(ow$VII))] <- mean(ow$VII, na.rm=T)
ow$VIII[which(is.na(ow$VIII))] <- mean(ow$VIII, na.rm=T)
ow$IX[which(is.na(ow$IX))] <- mean(ow$IX, na.rm=T)
```

2.4. Średnia temperatura w porach roku

Dodajemy do wcześniej utworzonych zmiennych (punkt 2.2) nowe kolumny zawierające średnią temperaturę w poszczególnych porach roku. Odbywa się to z wykorzystaniem funkcji `mutate()`. Następnie zaokrąglamy obliczone wartości do 2 miejsc po przecinku za pomocą funkcji `round()`.

```
wiosna <- wiosna %>% as_tibble() %>%  
  mutate(wiosna_avg=(III+IV+V)/3)  
wiosna <- round(wiosna, digits = 2)  
  
lato <- lato %>% as_tibble() %>%  
  mutate(lato_avg=(VI+VII+VIII)/3)  
lato <- round(lato, digits = 2)  
  
jesien <- jesien %>% as_tibble() %>%  
  mutate(jesien_avg=(IX+X+XI)/3)  
jesien <- round(jesien, digits = 2)  
  
zima <- zima %>% as_tibble() %>%  
  mutate(zima_avg=(XII+I+II)/3)  
zima <- round(zima, digits = 2)  
  
ow <- ow %>% as_tibble() %>%  
  mutate(ow_avg=(V+VI+VII+VIII+IX)/5)  
ow <- round(ow, digits = 2)
```

2.5. Rozkład średniej temperatury

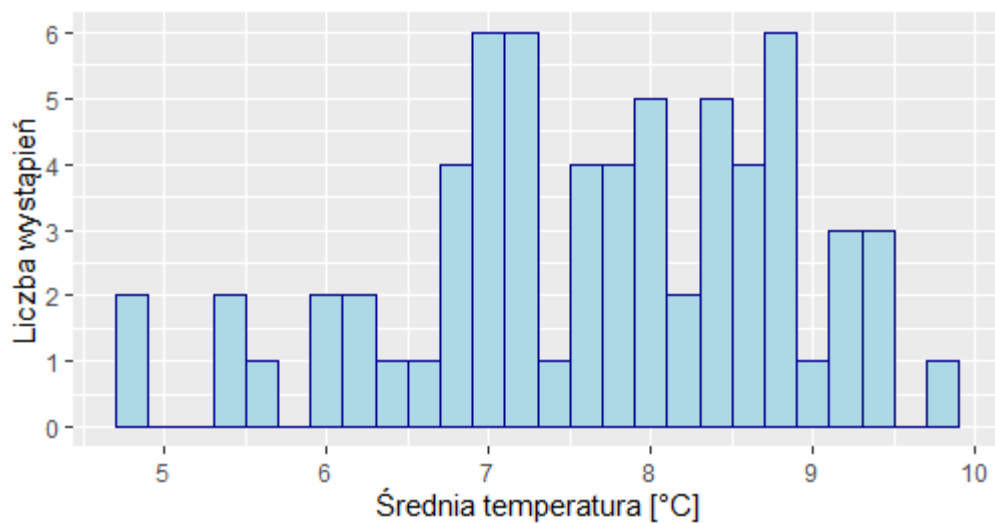
Tak jak w punkcie 1.3 do zobrazowania rozkładu temperatur w poszczególnych porach roku na przestrzeni lat wykorzystujemy histogramy.

***Dla każdego wykresu w komendach zmienia się:**

aes('zmienna'), ggtitle('nazwa miesiąca') oraz ggsave('nazwa pliku')

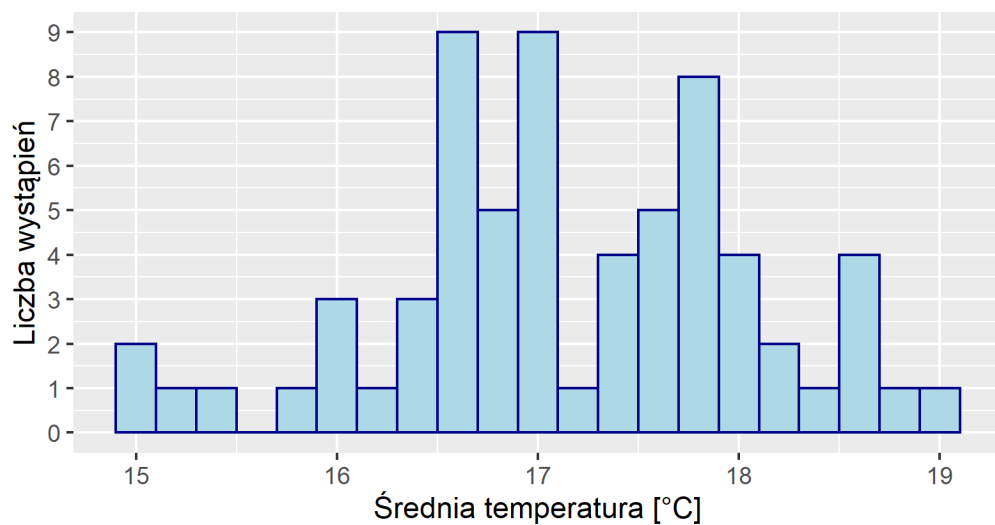
```
ggplot(wiosna, aes(wiosna_avg)) +  
  geom_histogram(color="darkblue", fill="lightblue", binwidth = 0.2)+  
  xlab("Średnia temperatura [°C]")+  
  ylab("Liczba wystąpień")+  
  scale_x_continuous(breaks = seq(4,10, by=1))+  
  scale_y_continuous(breaks = seq(0,18, by=1))+  
  ggtitle("Wiosna")ggsave("wiosna_histogram.png")  
  
ggsave("wiosna_histogram.png")
```

Wiosna



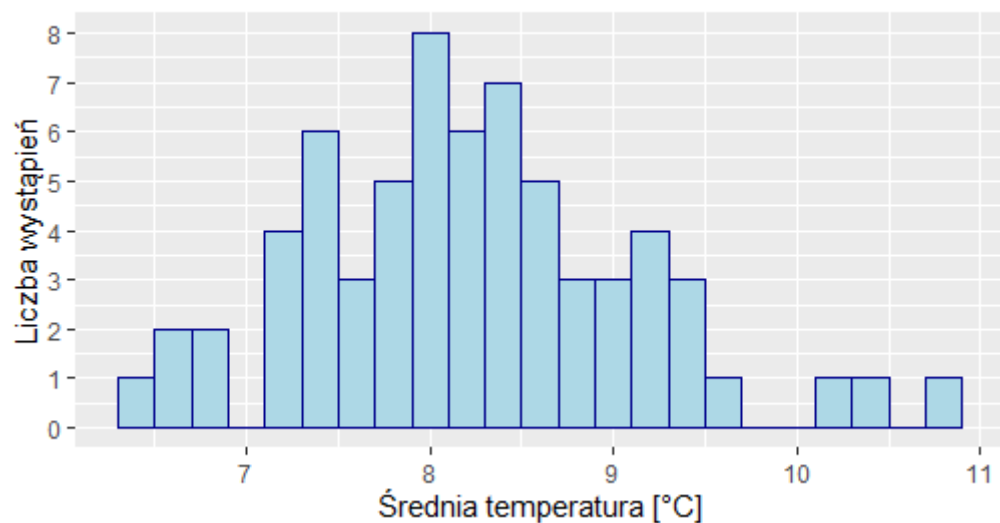
- wykr. 2.5.1

Lato

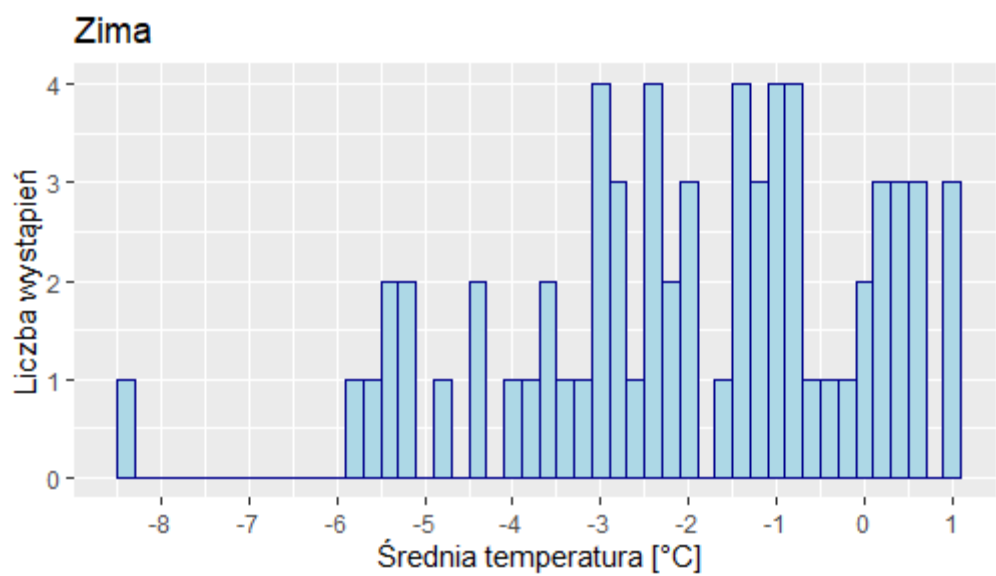


- wykr. 2.5.2

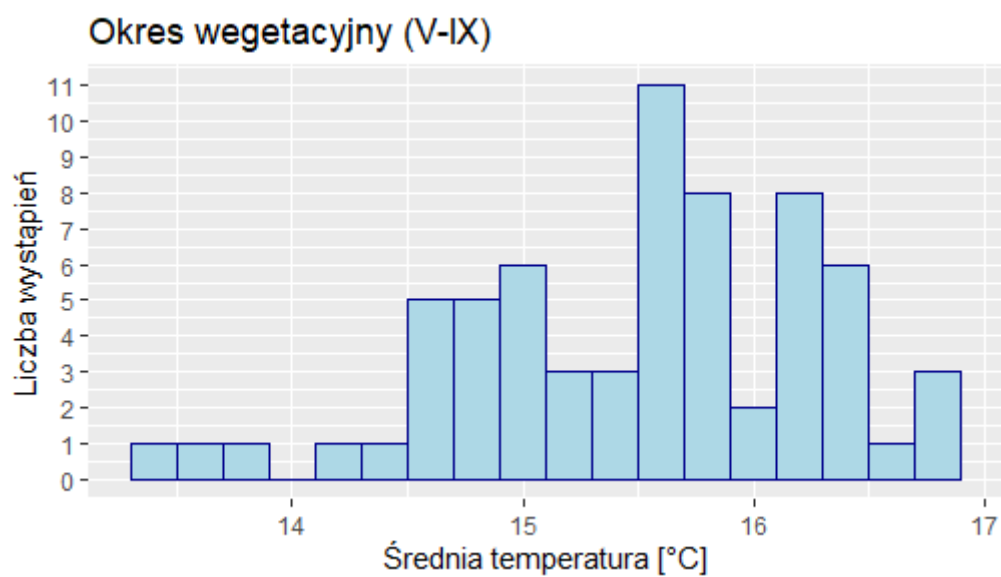
Jesień



- wykr. 2.5.3



- wykr. 2.5.4



- wykr. 2.5.5

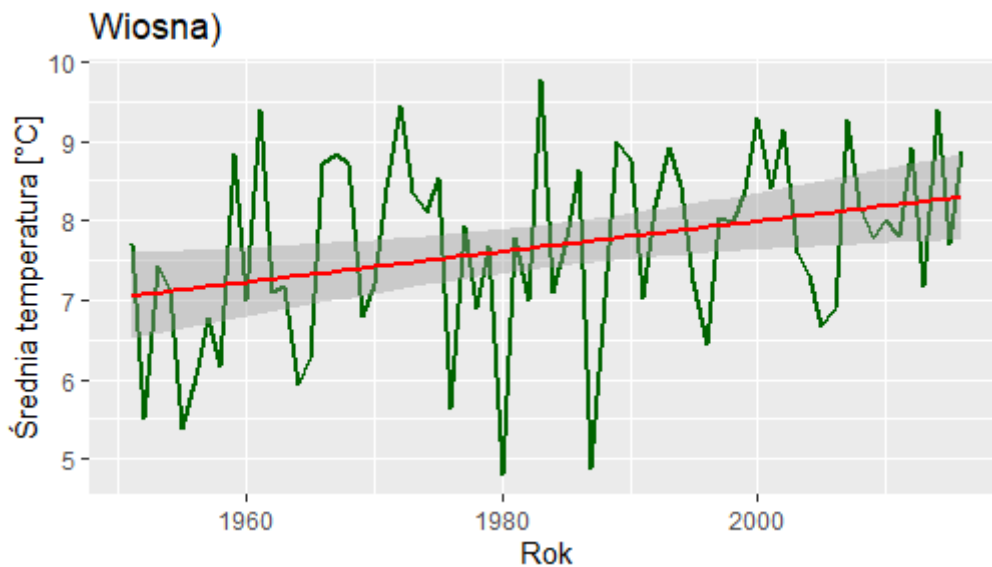
2.6. Zmienność wartości średniej temperatury

Podobnie jak w punkcie 1.6. tworzymy wykres liniowy z trendem liniowym.

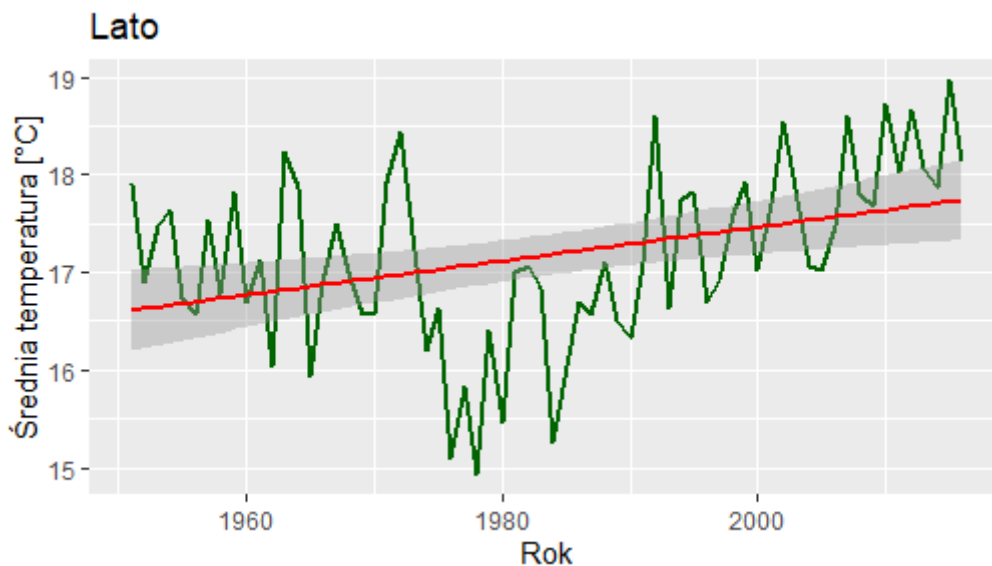
***Dla każdego wykresu w komendach zmienia się:**

aes('zmienna'), ggtitle('nazwa miesiąca') oraz ggsave('nazwa pliku')

```
ggplot(wiosna, aes(rok, wiosna_avg)) +  
  geom_line(color="darkgreen", size=1) +  
  geom_smooth(method = "lm", color="red") +  
  xlab("Rok")+  
  ylab("Średnia temperatura [°C]")+  
  ggtitle("Wiosna")  
  
ggsave("wiosna_zmienosc_srtemp.png")
```

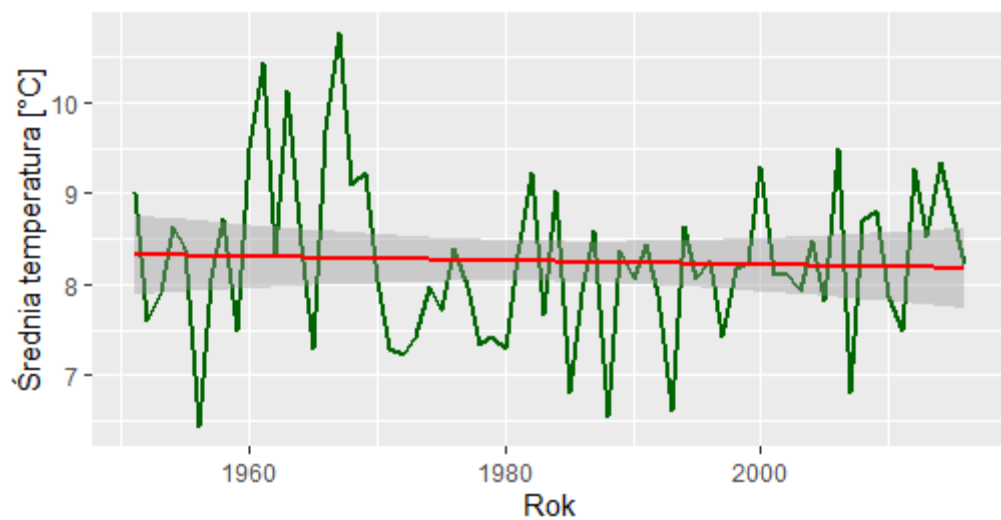


- wykr. 2.6.1



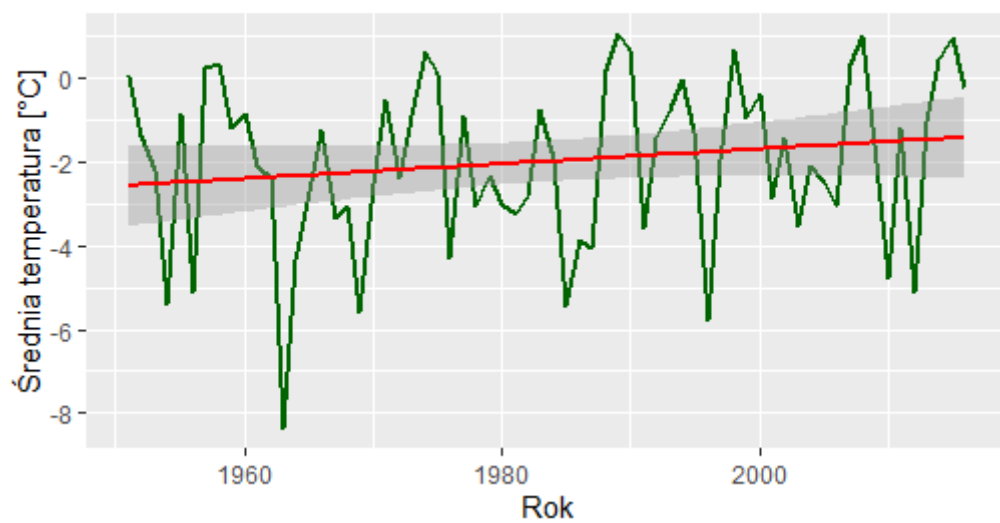
- wykr. 2.6.2

Jesień



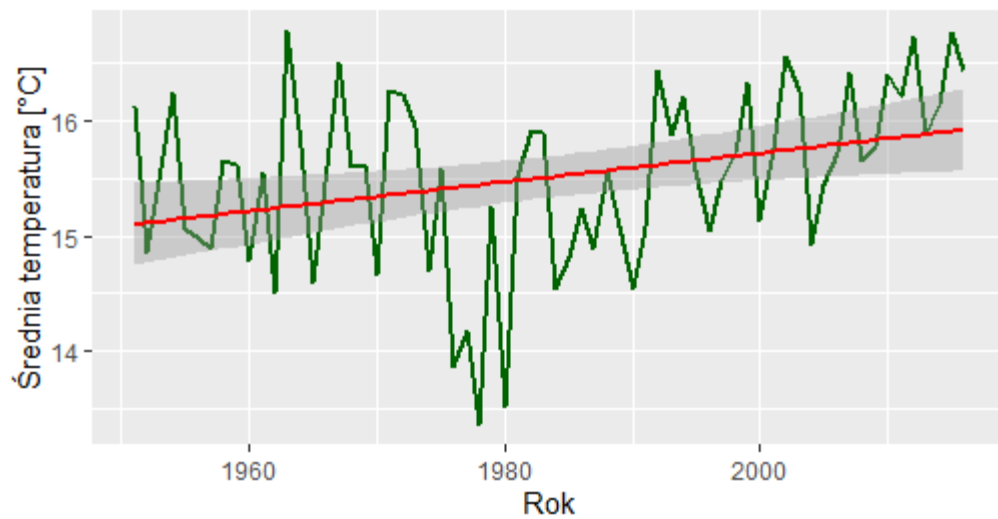
- wykr. 2.6.3

Zima



- wykr. 2.6.4

Okres wegetacyjny (V-IX)



- wykr. 2.6.5

2.7. Statystyka opisowa

Tak jak w punkcie 1.4. tworzymy *data.frame*. Nazywamy go *statopis* i tworzymy z wektorów wartości minimalnej, maksymalnej, średniej arytmetycznej, mediany, odchylenia standardowego oraz rozstępu międzykwartylowego.

```
pora <- c("wiosna", "lato", "jesien", "zima", "ow")

min <- c(min(wiosna$wiosna_avg), min(lato$lato_avg), min(jesien$jesien_avg),
min(zima$zima_avg), min(ow$ow_avg))

max <- c(max(wiosna$wiosna_avg), max(lato$lato_avg), max(jesien$jesien_avg),
max(zima$zima_avg), max(ow$ow_avg))

mean <- c(mean(wiosna$wiosna_avg), mean(lato$lato_avg), mean(jesien$jesien_avg),
mean(zima$zima_avg), mean(ow$ow_avg))

median <- c(median(wiosna$wiosna_avg), median(lato$lato_avg), median(jesien$jesien_avg),
median(zima$zima_avg), median(ow$ow_avg))

sd <- c(sd(wiosna$wiosna_avg), sd(lato$lato_avg), sd(jesien$jesien_avg), sd(zima$zima_avg),
sd(ow$ow_avg))

IQR <- c(IQR(wiosna$wiosna_avg), IQR(lato$lato_avg), IQR(jesien$jesien_avg),
IQR(zima$zima_avg), IQR(ow$ow_avg))

#nowa ramka danych
statopis <- data.frame(okres = pora, min_t = min, max_t = max, mean_t = mean,
median_t = median, sd_t = sd, iqr_t = IQR)
```

2.8. Eksport danych

Stworzoną wcześniej ramkę danych *statopis* eksportujemy do arkusza Excel. Wykorzystujemy bibliotekę "xlsx" w której dostępna jest funkcja *write.xlsx()*. Jako argumenty tej funkcji podajemy źródło danych do eksportu, nazwę pliku oraz nazwę arkusza.

```
#install.packages("xlsx")
library("xlsx")
write.xlsx(statopis, file = "analiza_dynowche.xlsx", sheetName = "Analiza_temperatury", append = F)
```

Powstała tabela statystyk opisowych:

pora roku	min	max	zakres	średnia	mediana	odchylenie standardowe	rozstęp międzykwartyłowy
wiosna	4,8	9,77	4,8 : 9,77	7,67	7,77	1,16	1,60
lato	14,93	18,97	14,93 : 18,97	17,18	17,09	0,90	1,20
jesien	6,43	10,77	6,43 : 10,77	8,25	8,19	0,89	1,05
zima	-8,37	1,07	-8,37 : 1,07	-1,98	-1,76	2,01	2,45
okres wegetacyjny (V-IX)	13,36	16,78	13,36 : 16,78	15,51	15,59	0,76	1,14

Projekt 1 - cz.3

3.1. Wczytanie danych i bibliotek

Wczytujemy odpowiednie biblioteki oraz pakiet danych w formie pliku .txt

```
library(dplyr)

library(ggplot2)
library(tidyr)
dynow_t <- read.delim("dynow_t.txt")
```

3.2. Przygotowanie danych

Tworzymy *data.frame* dla poszczególnych dekad razem ze średnią temperaturą dla pór roku i okresu wegetacyjnego oraz zaokrąglamy wartości do dwóch miejsc po przecinku.

Uwaga! Numery dekad odpowiadają poszczególnym okresom:

1. 1951 - 1960; 2. 1961 - 1970; 3. 1971 - 1980; 4. 1981 - 1990; 5. 1991 - 2000; 6. 2001 - 2010

```
d1 <- filter(dynow_t, rok >= 1951 & rok <= 1960) %>%
  mutate(wiosna_d1 = (III+IV+V)/3) %>%
  mutate(lato_d1 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d1=(IX+X+XI)/3) %>%
  mutate(zima_d1=(XII+I+II)/3) %>%
  mutate(ow_d1=(V+VI+VII+VIII+IX)/5)
d1 <- round(d1, digits = 2)

d2 <- filter(dynow_t, rok >= 1961 & rok <= 1970) %>%
  mutate(wiosna_d2 = (III+IV+V)/3) %>%
  mutate(lato_d2 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d2=(IX+X+XI)/3) %>%
  mutate(zima_d2=(XII+I+II)/3) %>%
  mutate(ow_d2=(V+VI+VII+VIII+IX)/5)
d2 <- round(d2, digits = 2)

d3 <- filter(dynow_t, rok >= 1971 & rok <= 1980) %>%
  mutate(wiosna_d3 = (III+IV+V)/3) %>%
  mutate(lato_d3 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d3 = (IX+X+XI)/3) %>%
  mutate(zima_d3 = (XII+I+II)/3) %>%
  mutate(ow_d3 = (V+VI+VII+VIII+IX)/5)
d3 <- round(d3, digits = 2)
```

```

d4 <- filter(dynow_t, rok >= 1981 & rok <= 1990) %>%
  mutate(wiosna_d4 = (III+IV+V)/3) %>%
  mutate(lato_d4 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d4 = (IX+X+XI)/3) %>%
  mutate(zima_d4 = (XII+I+II)/3) %>%
  mutate(ow_d4 = (V+VI+VII+VIII+IX)/5)
d4 <- round(d4, digits = 2)

d5 <- filter(dynow_t, rok >= 1991 & rok <= 2000) %>%
  mutate(wiosna_d5 = (III+IV+V)/3) %>%
  mutate(lato_d5 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d5 = (IX+X+XI)/3) %>%
  mutate(zima_d5 = (XII+I+II)/3) %>%
  mutate(ow_d5 = (V+VI+VII+VIII+IX)/5)
d5 <- round(d5, digits = 2)

d6 <- filter(dynow_t, rok >= 2001 & rok <= 2010) %>%
  mutate(wiosna_d6 = (III+IV+V)/3) %>%
  mutate(lato_d6 = (VI+VII+VIII)/3) %>%
  mutate(jesien_d6 = (IX+X+XI)/3) %>%
  mutate(zima_d6 = (XII+I+II)/3) %>%
  mutate(ow_d6 = (V+VI+VII+VIII+IX)/5)
d6 <- round(d6, digits = 2)

```

3.3 Przygotowanie danych - cd.

Tworzymy *data.frame srednia_dekady* - zawiera informacje o średnich temperaturach w kolejnych dekadach dla poszczególnych miesięcy, pór roku i okresu wegetacyjnego. Zostawiamy wartości NA. *

```

#wektor do numerowania dekad. Numeracja zgodna z opisem w punkcie 3.2
lp <- c(1, 2, 3, 4, 5, 6)

#wektory ze średnią temperaturą w poszczególnych dekadach i porach roku + okres wegetacyjny
styczen_avg <- c(mean(d1$I), mean(d2$I), mean(d3$I), mean(d4$I), mean(d5$I),
mean(d6$I))

luty_avg <- c(mean(d1$II), mean(d2$II), mean(d3$II), mean(d4$II), mean(d5$II),
, mean(d6$II))

marzec_avg <- c(mean(d1$III), mean(d2$III), mean(d3$III), mean(d4$III), mean(
d5$III), mean(d6$III))

kwiecień_avg <- c(mean(d1$IV), mean(d2$IV), mean(d3$IV), mean(d4$IV), mean(d5
$IV), mean(d6$IV))

```

```

maj_avg <- c(mean(d1$V), mean(d2$V), mean(d3$V), mean(d4$V), mean(d5$V), mean(
d6$V))

czerwiec_avg <- c(mean(d1$VI), mean(d2$VI), mean(d3$VI), mean(d4$VI), mean(d5
$VI), mean(d6$VI))

lipiec_avg <- c(mean(d1$VII), mean(d2$VII), mean(d3$VII), mean(d4$VII), mean(
d5$VII), mean(d6$VII))

sierpien_avg <- c(mean(d1$VIII), mean(d2$VIII), mean(d3$VIII), mean(d4$VIII),
mean(d5$VIII), mean(d6$VIII))

wrzesien_avg <- c(mean(d1$IX), mean(d2$IX), mean(d3$IX), mean(d4$IX), mean(d5
$IX), mean(d6$IX))

pazdziernik_avg <- c(mean(d1$X), mean(d2$X), mean(d3$X), mean(d4$X), mean(d5$
X), mean(d6$X))

listopad_avg <- c(mean(d1$XI), mean(d2$XI), mean(d3$XI), mean(d4$XI), mean(d5
$XI), mean(d6$XI))

grudzien_avg <- c(mean(d1$XII), mean(d2$XII), mean(d3$XII), mean(d4$XII), mea
n(d5$XII), mean(d6$XII))

wiosna_avg <- c(mean(d1$wiosna_d1), mean(d2$wiosna_d2), mean(d3$wiosna_d3), m
ean(d4$wiosna_d4), mean(d5$wiosna_d5), mean(d6$wiosna_d6))

lato_avg <- c(mean(d1$lato_d1), mean(d2$lato_d2), mean(d3$lato_d3), mean(d4$l
ato_d4), mean(d5$lato_d5), mean(d6$lato_d6))

jesien_avg <- c(mean(d1$jesien_d1), mean(d2$jesien_d2), mean(d3$jesien_d3), m
ean(d4$jesien_d4), mean(d5$jesien_d5), mean(d6$jesien_d6))

zima_avg <- c(mean(d1$zima_d1), mean(d2$zima_d2), mean(d3$zima_d3), mean(d4$z
ima_d4), mean(d5$zima_d5), mean(d6$zima_d6))

ow_avg <- c(mean(d1$ow_d1), mean(d2$ow_d2), mean(d3$ow_d3), mean(d4$ow_d4), m
ean(d5$ow_d5), mean(d6$ow_d6))

#nowy data.frame
srednia_dekady <- data.frame(dekada = lp, styczen_avgt = styczen_avg, luty_av
gt = luty_avg, marzec_avgt = marzec_avg, kwiecień_avgt = kwiecień_avg, maj_av
gt = maj_avg, czerwiec_avgt = czerwiec_avg, lipiec_avgt = lipiec_avg, sierpie
n_avgt = sierpien_avg, wrzesien_avgt = wrzesien_avg, pazdziernik_avgt = pazdzi
ernik_avg, listopad_avgt = listopad_avg, grudzien_avgt = grudzien_avg, wiosna_
avgt = wiosna_avg, lato_avgt = lato_avg, jesien_avgt = jesien_avg, zima_avgt
= zima_avg, ow_avgt = ow_avg)

```

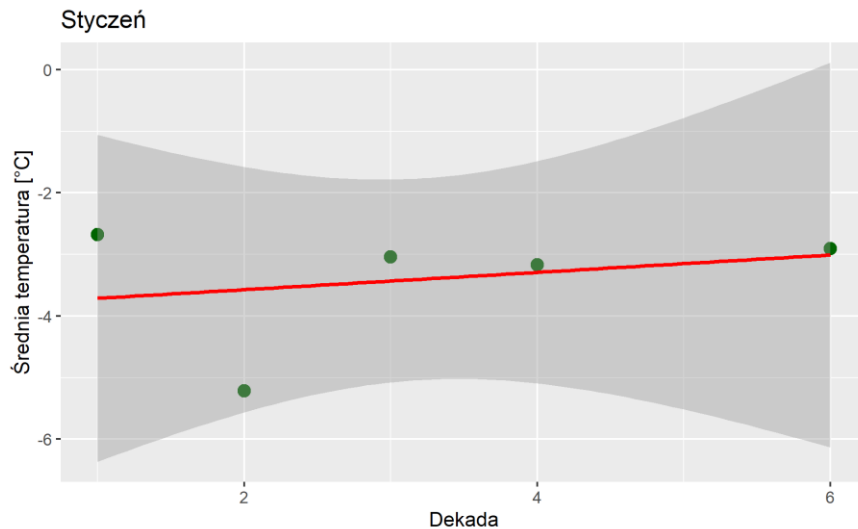

3.4 Wykresy rozrzutu średniej temperatury

Tworzymy wykresy rozrzutu średniej temperatury powietrza dla poszczególnych miesięcy, pór roku i okresu wegetacyjnego z trendem liniowym.

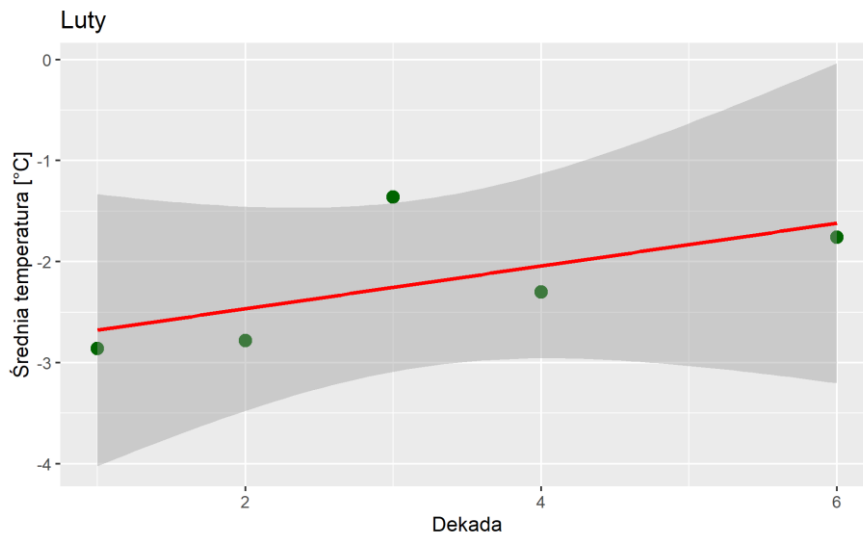
***Dla każdego wykresu w komendach zmienia się:**

aes(-, 'zmienna'), ggtitle('nazwa miesiąca/pora roku') oraz ggsave('nazwa pliku')

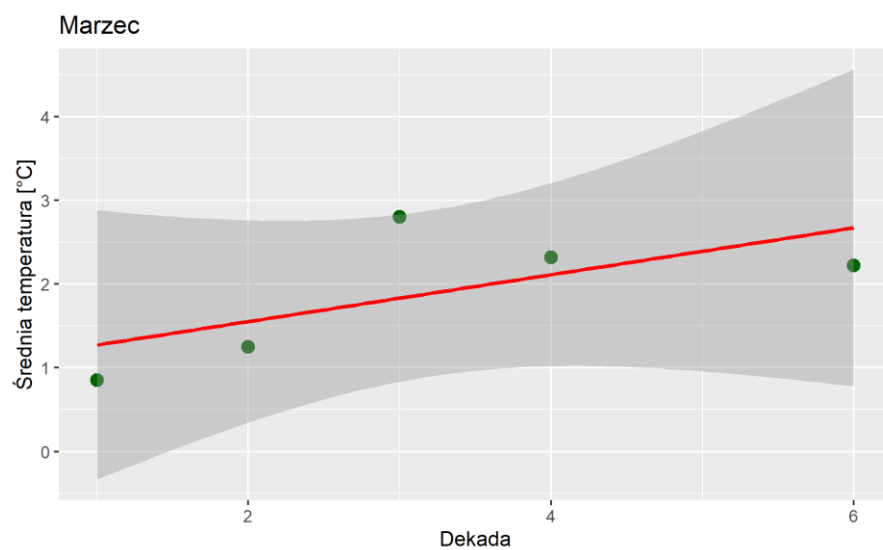
```
ggplot(srednia_dekady, aes(dekada, styczen_avgt)) +  
  geom_point(color = "darkgreen", size = 3) +  
  geom_smooth(method = "lm", color = "red") +  
  xlab("Dekada") +  
  ylab("Średnia temperatura [°C]") +  
  ggtitle("Styczeń")  
  
ggsave("dekady_styczen.png")
```



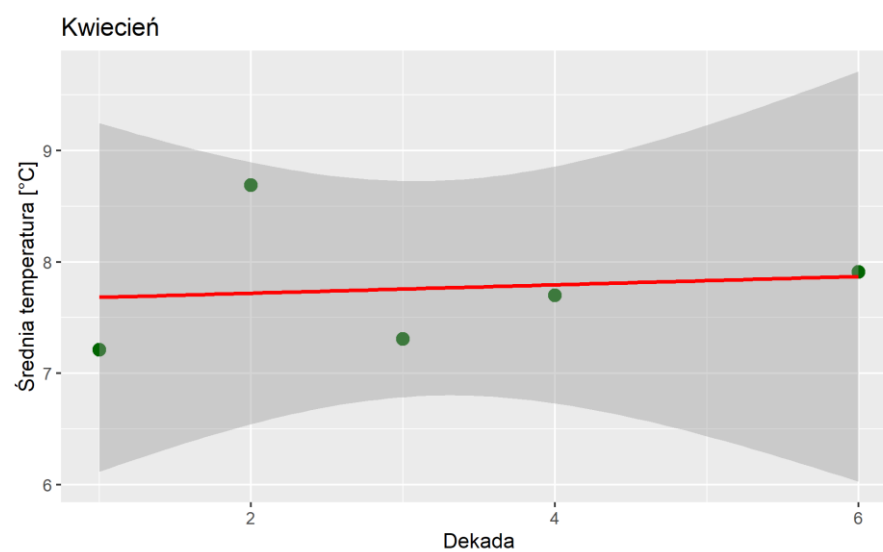
- wyk. 3.4.1



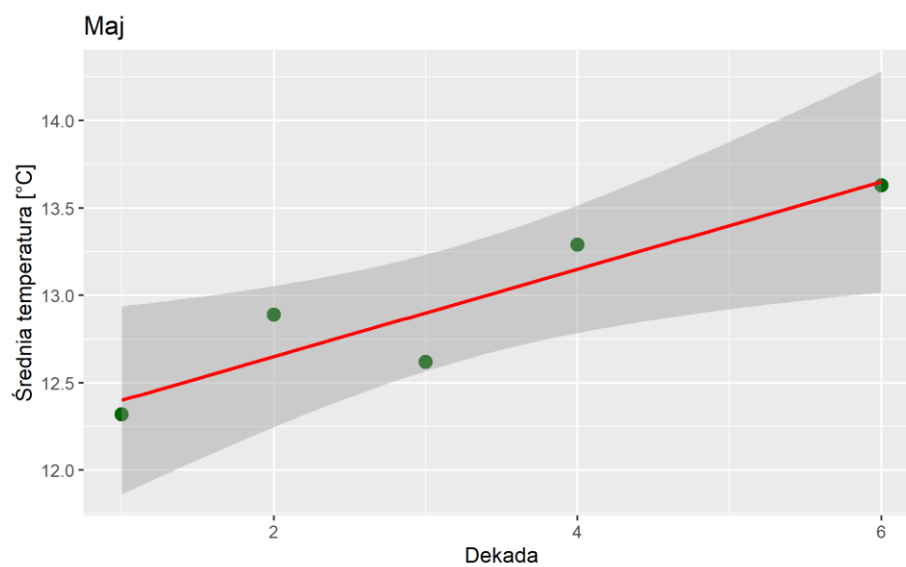
- wyk. 3.4.2



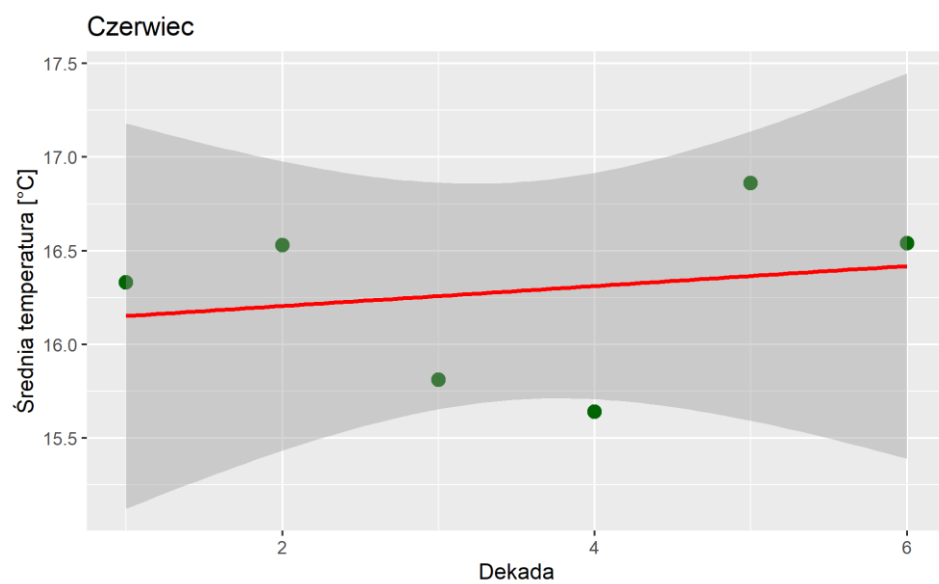
- wykr. 3.4.3



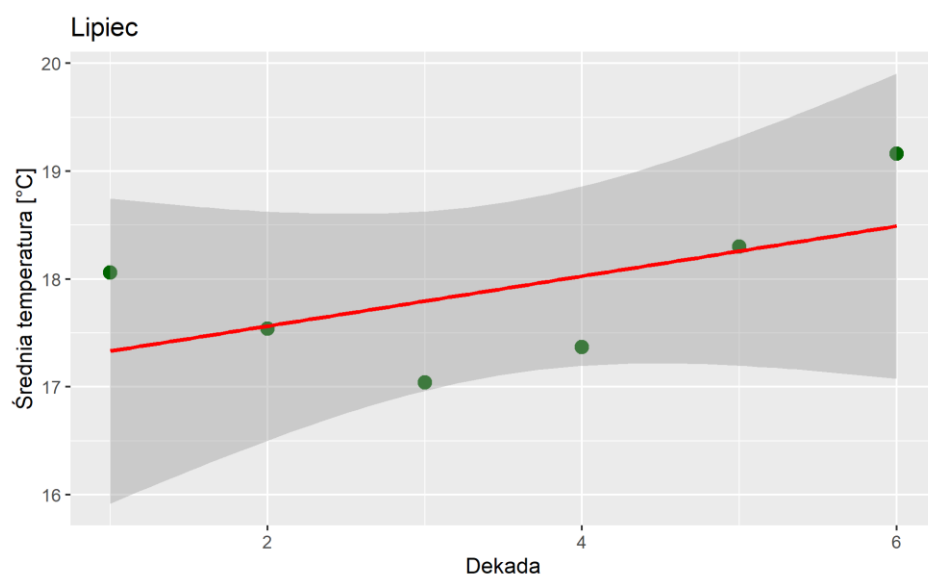
- wykr. 3.4.4



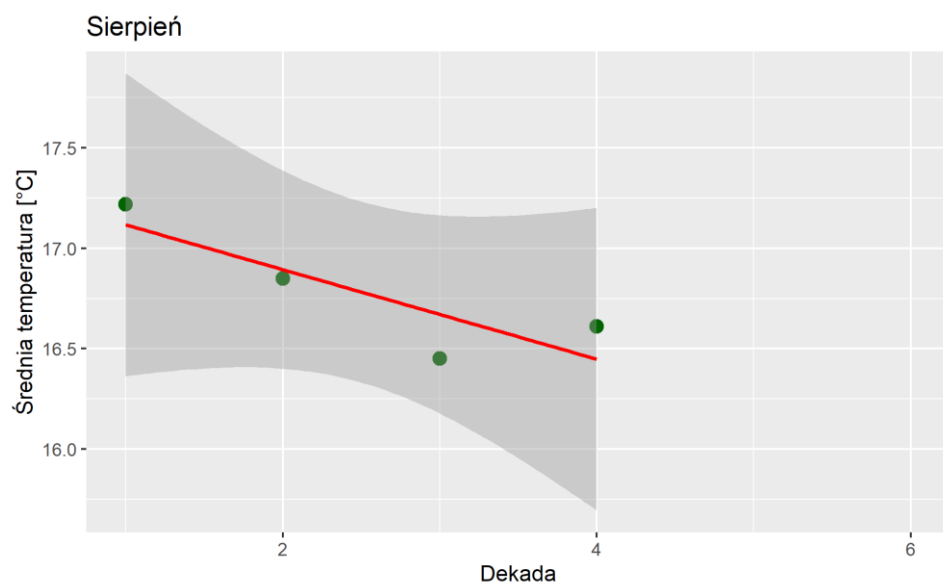
- wykr. 3.4.5



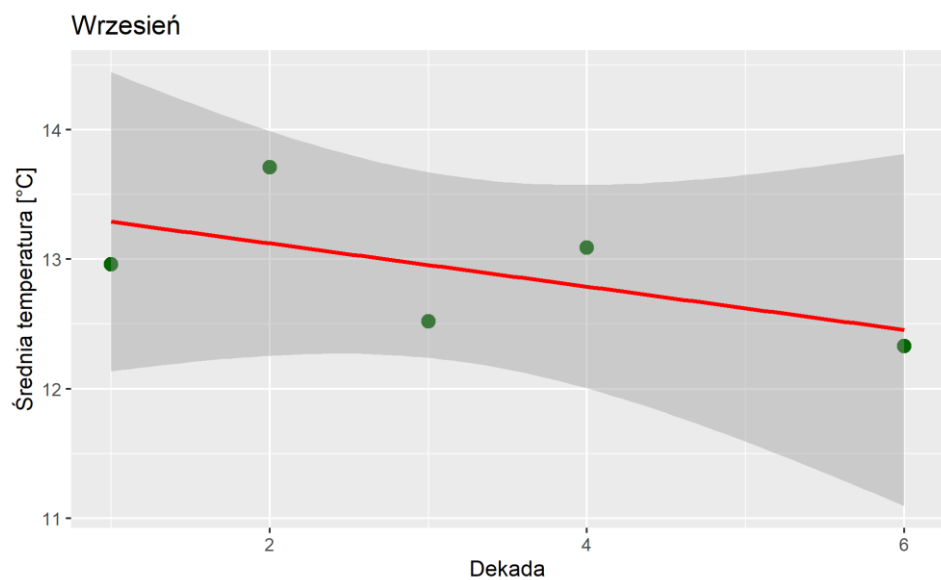
- wykr. 3.4.6



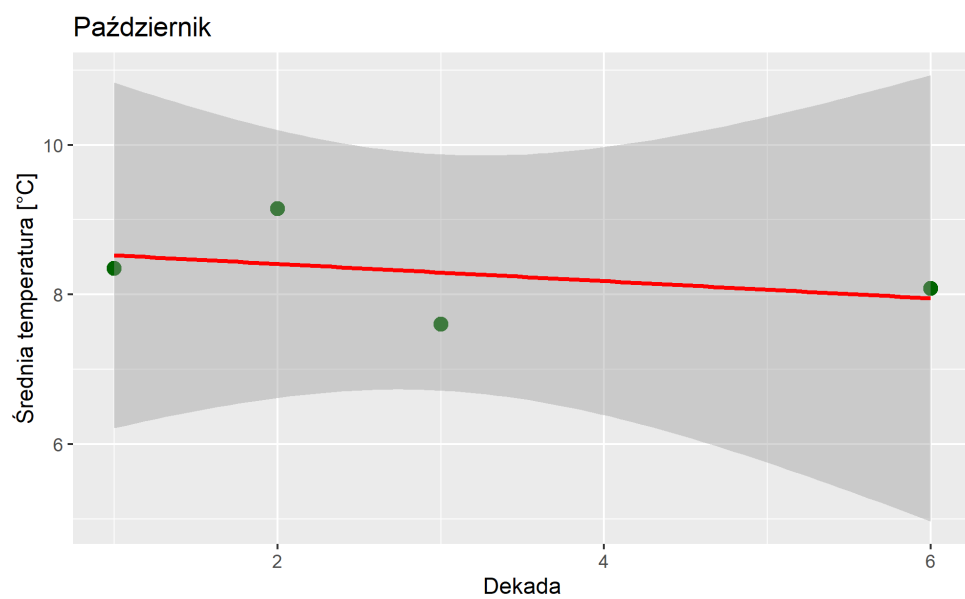
- wykr. 3.4.7



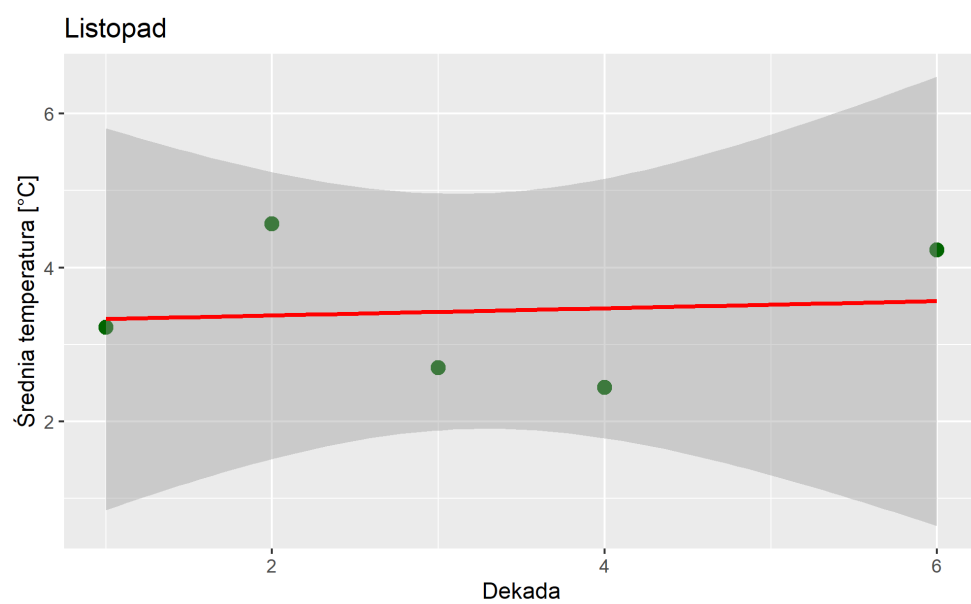
- wykr. 3.4.8



- wyk. 3.4.9

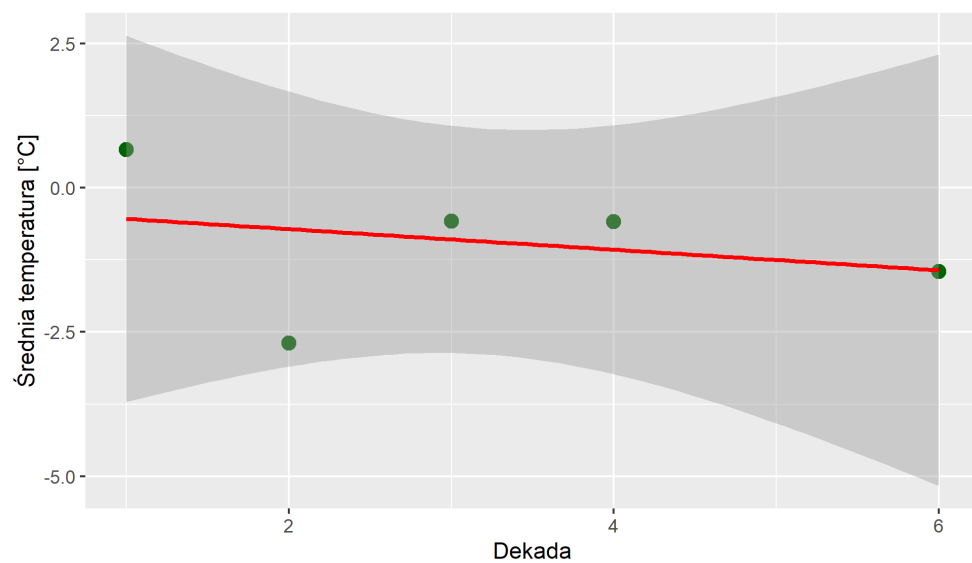


- wyk. 3.4.10



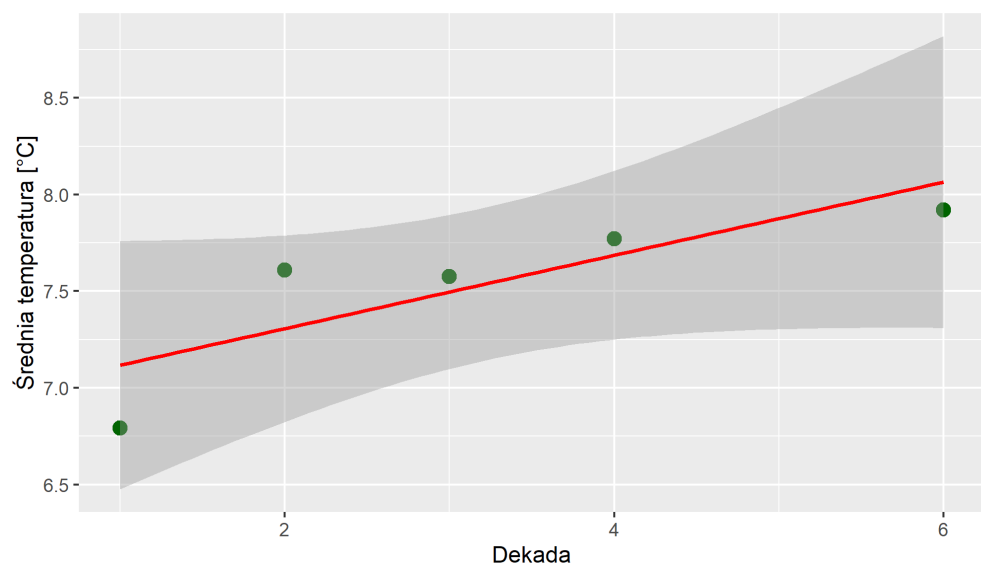
- wyk. 3.4.11

Grudzień



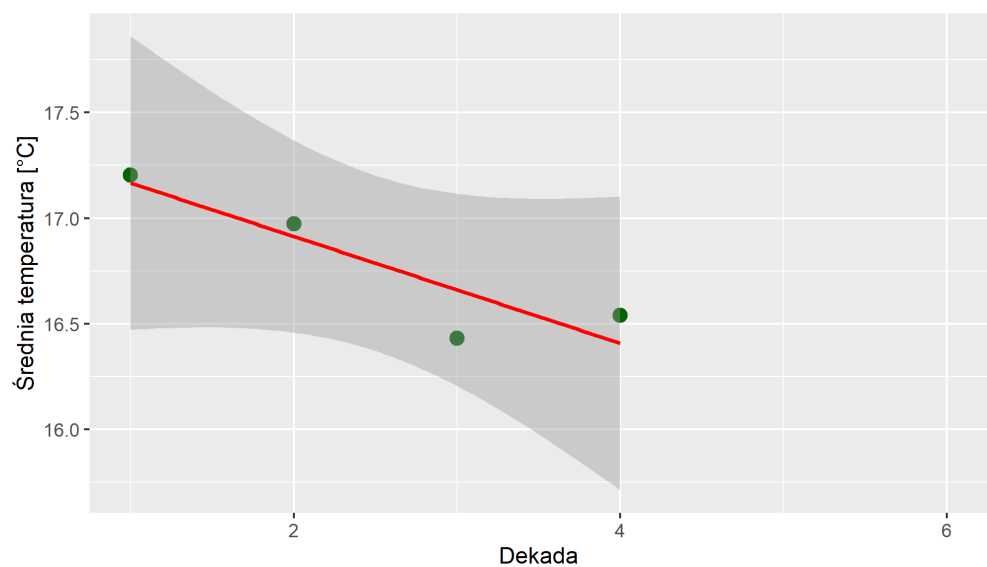
- wykr. 3.4.12

Wiosna



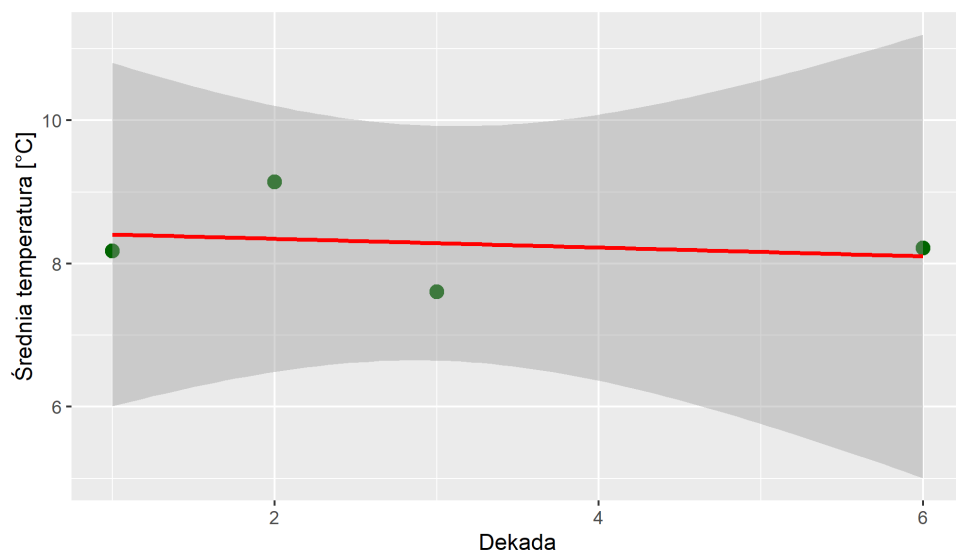
- wykr. 3.4.13

Lato



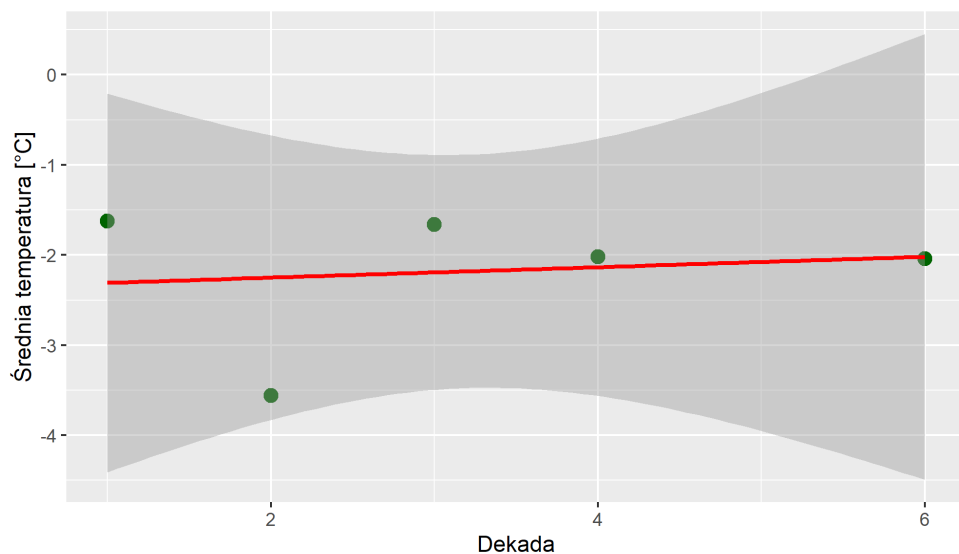
- wykr. 3.4.14

Jesień



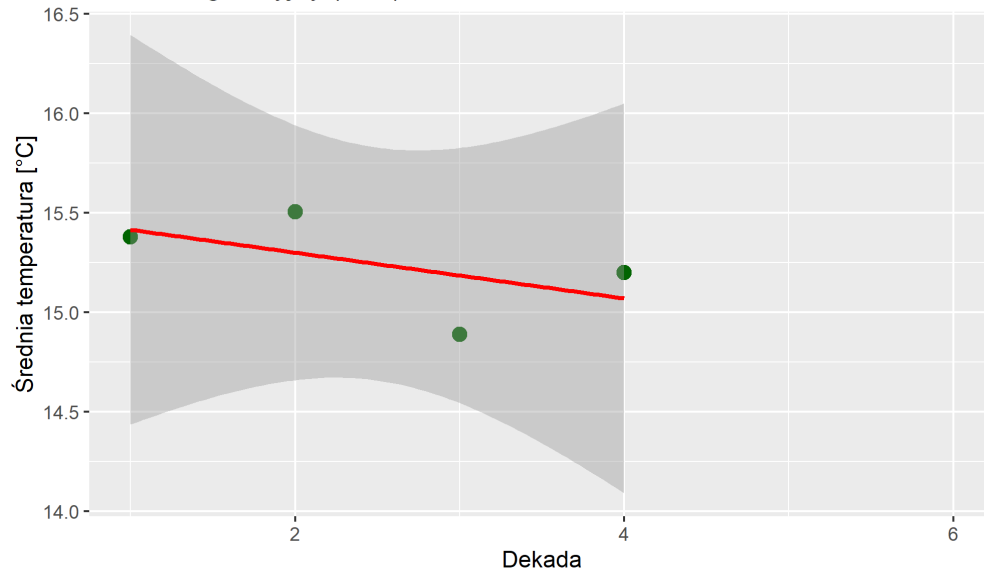
- wyk. 3.4.15

Zima



- wyk. 3.4.16

Okres wegetacyjny (V-IX)



- wyk. 3.4.17

4. Obserwacje:

Dokonana analiza średnich temperatur zarejestrowanych na stacji meteorologicznej w Dynowie pozwala na wyłuskanie wielu interesujących informacji. Przetworzenie nieczytelного pliku tekstowego zawierającego zanotowane pomiary do postaci graficznej – czytelnych wykresów, zdecydowanie ułatwia analizę i interpretację zapisanych informacji.

Utworzone wykresy liniowe z trendem w punktach 1.6 oraz 2.6 pozwalają zauważyć tendencję zmian wartości temperatury dla poszczególnych miesięcy, pór roku i okresu wegetacyjnego na przestrzeni ponad 60 lat. Pozwalają również wnioskować o temperaturze w latach nadchodzących.

Spośród wszystkich miesięcy cztery nie wykazują tendencji wzrostowej temperatury. Co ważniejsze, są to miesiące głównie jednej pory roku - jesieni: wrzesień, październik i listopad. Dodatkowo również grudzień (wykresy 1.6.9-1.6.12).

Największy wzrost średniej temperatury zachodzi dla lipca (wykres 1.6.7), największy spadek dla września (wykres 1.6.9). Wśród pór roku największy wzrost zachodzi dla lata (wykres 2.6.2).

Tabele statystyk opisowych (punkty 1.5 i 2.8) ukazują kilka znaczących informacji. Największy przedział zanotowanej temperatury zachodzi dla lutego (od -12,9 do 4,1), a najmniejszy dla lipca (od 14,9 do 20,5). Jednocześnie luty posiada największe odchylenie standardowe (3,64). Wśród pór roku największy zakres osiągnęła zima (od -8,37 do 1,07).

Kwestia skrajnych wartości również zamyka się dla tych samych dwóch miesięcy. Najniższa wystąpiła w lutym (-12,9), a najwyższa w lipcu (20,5).

Największy rozstęp międzykwartylowy wystąpił dla lutego (4,325), najmniejszy dla października (1,4). Koresponduje z tym histogram października (wykres 1.3.10), zgodnie z którym aż 13 razy wystąpiła ta sama wartość, a reszta wartości wystąpiła znacznie mniejszą i jednocześnie podobną ilość razy (1-2).

Okres wegetacyjny (V-IX) zanotował najwyższą średnią temperaturę (16,78), a także najmniejsze odchylenie standardowe (0,76)

Na przestrzeni dekad największy wzrost odnotowano dla maja, wśród pór roku dla wiosny.

5. Wnioski:

Analiza pokazuje generalny wzrost średniej temperatury w okresie 1951-2016, w niektórych przypadkach znaczący. Pomiarom odbiegającym od oczekiwanych wyników jest tendencja spadkowa dla września. Przystępując do analizy oczekiwaliśmy, że wszystkie miesiące będą wykazywać tendencję wzrostową. Zaskakująca jest również jesień, gdzie tendencja wzrostowa nie jest widoczna. W ujęciu ogólnym wartości te są prawie jednakowe, można zaobserwować nawet delikatny spadek. Jednak nie należy się spodziewać, że w najbliższej przyszłości ta temperatura będzie utrzymywać się na jednakowym poziomie. Prawdopodobnie będzie rosła. Obserwując wykres dla lata (wykr. 2.6.2) wyraźnie widać stopniowy wzrost. Ciekawy jest jednak okres obejmujący lata 1970 – 1980 kiedy nastąpił spadek temperatury w tej porze roku. Gdy spojrzymy na ten sam okres w innych porach roku nie zaobserwujemy tak gwałtownego spadku. Widać go w okresie wegetacyjnym (wykr. 2.6.5) jednak jest on znacznie mniej wyraźny, ze względu na to, że okres wegetacyjny obejmuje również część wiosny.

*W przypadku analizy średniej temperatury dla poszczególnych dekad (punkt 3.4) zdecydowaliśmy się nie zastępować wartości NA średnią arytmetyczną. Po obliczeniu średnich temperatur dla poszczególnych pór roku i okresu wegetacyjnego otrzymaliśmy wiele brakujących wartości. Uznaliśmy, że zastąpienie ich średnią arytmetyczną zaburzyłoby oczekiwane wyniki. Okres objęty analizą podzieliliśmy na 6 dekad. Zastąpienie brakujących danych średnią z pięciu (lub nawet czterech) wartości nie byłoby miarodajne.

6. Bibliografia

- prezentacje z przedmiotu PDŚ autorstwa Pani dr inż. Małgorzaty Danek
- <https://rstudio.com/resources/cheatsheets/>
- <https://bookdown.org/yihui/rmarkdown/notebook.html>
- http://uc-r.github.io/r_notebook
- <https://ggplot2.tidyverse.org/reference/>