

Metody probabilistyczne i statystyka - wykład 8

Materiały uzupełniające.

Jarosław Kotowicz^{1,*}

wersja z roku ak.2020/21

Spis treści

1	Statystyki graficzne i tabelaryczne	1
1.1	Dane z przykładu	2
1.2	Szereg statystyczne	2
1.3	Tabela kontyngencji (tabela krzyżowa)	3
1.4	Statystyki graficzne	3
1.4.1	Wykresy słupkowe	3
1.4.2	Wykresy punktowe	5
1.4.3	Wykres kołowy	7
1.4.4	Wykresy <i>łodyga-liść</i>	8

Spis rysunków

1	Wykres słupkowy 1	4
2	Wykres słupkowy 2	5
3	Wykres punktowy 1	6
4	Wykres punktowy 2	7
5	Wykres kołowy	8

¹ Zakład Bioinformatyki, Instytut Informatyki, Uniwersytet w Białymstoku

* Correspondence: [Jarosław Kotowicz <j.kotowicz@uwb.edu.pl>](mailto:Jaroslaw.Kotowicz@uwb.edu.pl)

1 Statystyki graficzne i tabelaryczne

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.1.2      v dplyr  1.0.7
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(DT)
```

1.1 Dane z przykładu

```
Example_01 %>% datatable()
```

Show entries

Search:

	Liczba pasażerów	Prędkość	Płeć
1	2	64	K
2	0	77	M
3	2	51	M
4	1	70	M
5	2	69	K
6	1	50	M
7	0	72	M
8	3	47	K
9	1	93	M
10	3	52	K

Showing 1 to 10 of 25 entries

Previous 2 3 Next

1.2 Szereg statystyczne

```
library(sjmisc )
```

```
Example_01 %>% frq
```

```
##
## Liczba pasażerów <numeric>
## # total N=25 valid N=25 mean=1.80 sd=1.15
##
## Value | N | Raw % | Valid % | Cum. %
## -----
##      0 | 3 | 12.00 | 12.00 | 12
##      1 | 8 | 32.00 | 32.00 | 44
##      2 | 7 | 28.00 | 28.00 | 72
##      3 | 5 | 20.00 | 20.00 | 92
##      4 | 2 |  8.00 |  8.00 | 100
## <NA> | 0 |  0.00 | <NA> | <NA>
##
##
## Prędkość <numeric>
## # total N=25 valid N=25 mean=64.24 sd=10.22
##
## Value | N | Raw % | Valid % | Cum. %
## -----
##     47 | 1 |  4.00 |  4.00 |  4.00
##     50 | 1 |  4.00 |  4.00 |  8.00
##     51 | 1 |  4.00 |  4.00 | 12.00
##     52 | 1 |  4.00 |  4.00 | 16.00
```

```

## 56 | 1 | 4 | 4 | 20.00
## 58 | 1 | 4 | 4 | 24.00
## 59 | 1 | 4 | 4 | 28.00
## 60 | 2 | 8 | 8 | 36.00
## 61 | 1 | 4 | 4 | 40.00
## 62 | 1 | 4 | 4 | 44.00
## 63 | 2 | 8 | 8 | 52.00
## 64 | 1 | 4 | 4 | 56.00
## 65 | 1 | 4 | 4 | 60.00
## 66 | 1 | 4 | 4 | 64.00
## 67 | 1 | 4 | 4 | 68.00
## 68 | 1 | 4 | 4 | 72.00
## 69 | 1 | 4 | 4 | 76.00
## 70 | 1 | 4 | 4 | 80.00
## 71 | 1 | 4 | 4 | 84.00
## 72 | 1 | 4 | 4 | 88.00
## 77 | 1 | 4 | 4 | 92.00
## 82 | 1 | 4 | 4 | 96.00
## 93 | 1 | 4 | 4 | 100.00
## <NA> | 0 | 0 | <NA> | <NA>
##
##
## Płeć <ategorical>
## # total N=25 valid N=25 mean=1.56 sd=0.51
##
## Value | N | Raw % | Valid % | Cum. %
## -----
## K | 11 | 44.00 | 44.00 | 44
## M | 14 | 56.00 | 56.00 | 100
## <NA> | 0 | 0.00 | <NA> | <NA>

```

1.3 Tabela kontyngencji (tabela krzyżowa)

```
sjmisc::flat_table(Example_01[,c(1,3)])
```

```

##          Płeć K M
## Liczba pasażerów
## 0          0 3
## 1          4 4
## 2          3 4
## 3          3 2
## 4          1 1

```

1.4 Statystyki graficzne

```
library(ggplot2)
```

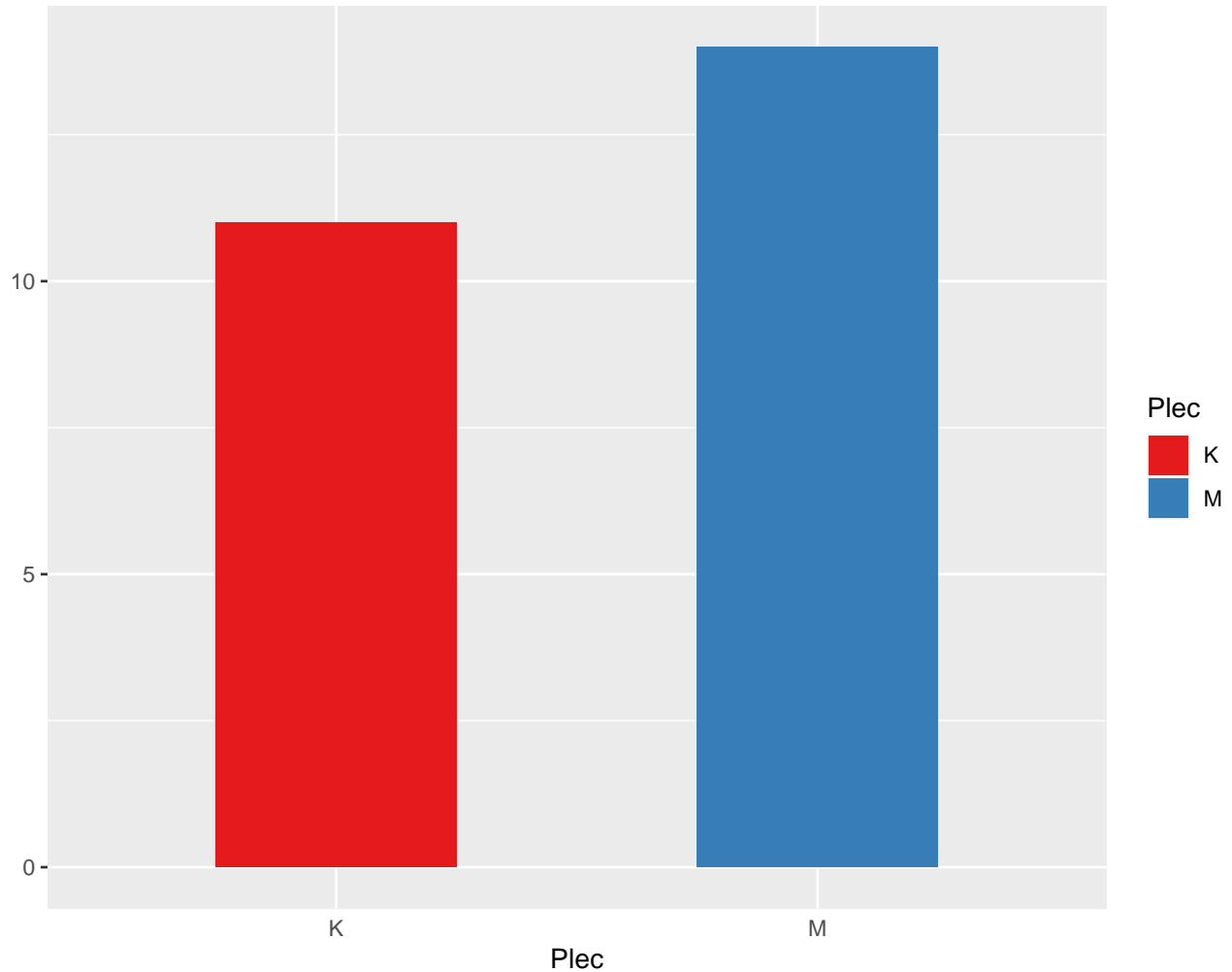
```
rysunek01 <- ggplot(Example_01, aes(x = Płeć, fill = `Płeć`))
```

1.4.1 Wykresy słupkowe

```
rysunek01 +
  geom_bar(width = 0.5) +
```

```
scale_fill_brewer(palette = "Set1") +
theme(axis.ticks.x = element_blank()) +
labs(title = "Wykres słupkowy", y = NULL, caption = "Źródło: Opracowanie własne.")
```

Wykres słupkowy



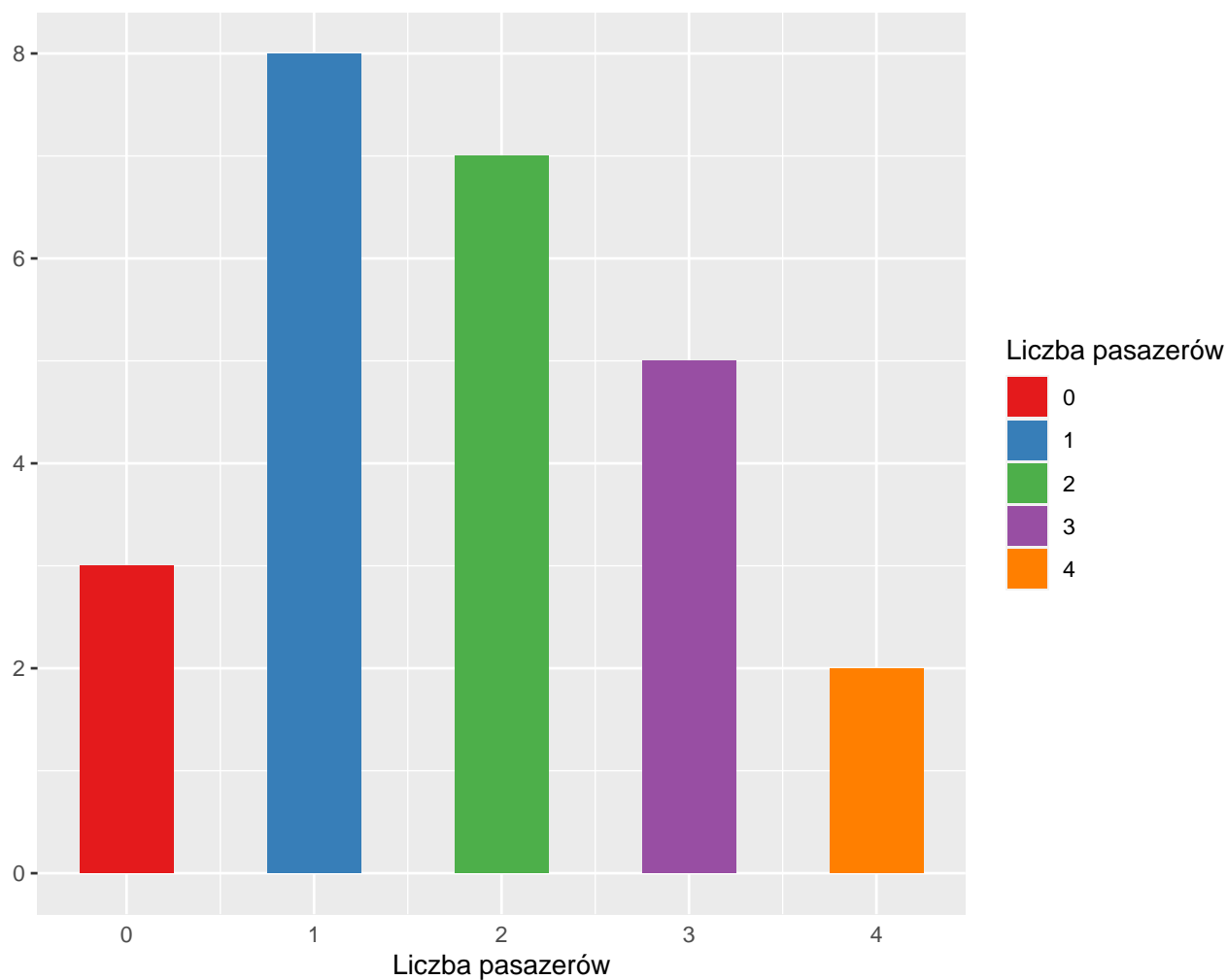
Źródło: Opracowanie własne.

Rysunek 1: Wykres słupkowy 1

```
rysunek02 <- ggplot(Example_01, aes(x=`Liczba pasażerów`, fill = factor(`Liczba pasażerów`)))
```

```
rysunek02 + geom_bar(width = 0.5) +
scale_fill_brewer(palette = "Set1") +
theme(
  axis.ticks.x = element_blank(),
  plot.subtitle = element_text(vjust = 1),
  plot.caption = element_text(vjust = 1)) +
labs(
  title = "Wykres słupkowy",
  y = NULL,
  caption = "Źródło: Opracowanie własne.",
  fill = "Liczba pasażerów")
```

Wykres słupkowy



Zródło: Opracowanie własne.

Rysunek 2: Wykres słupkowy 2

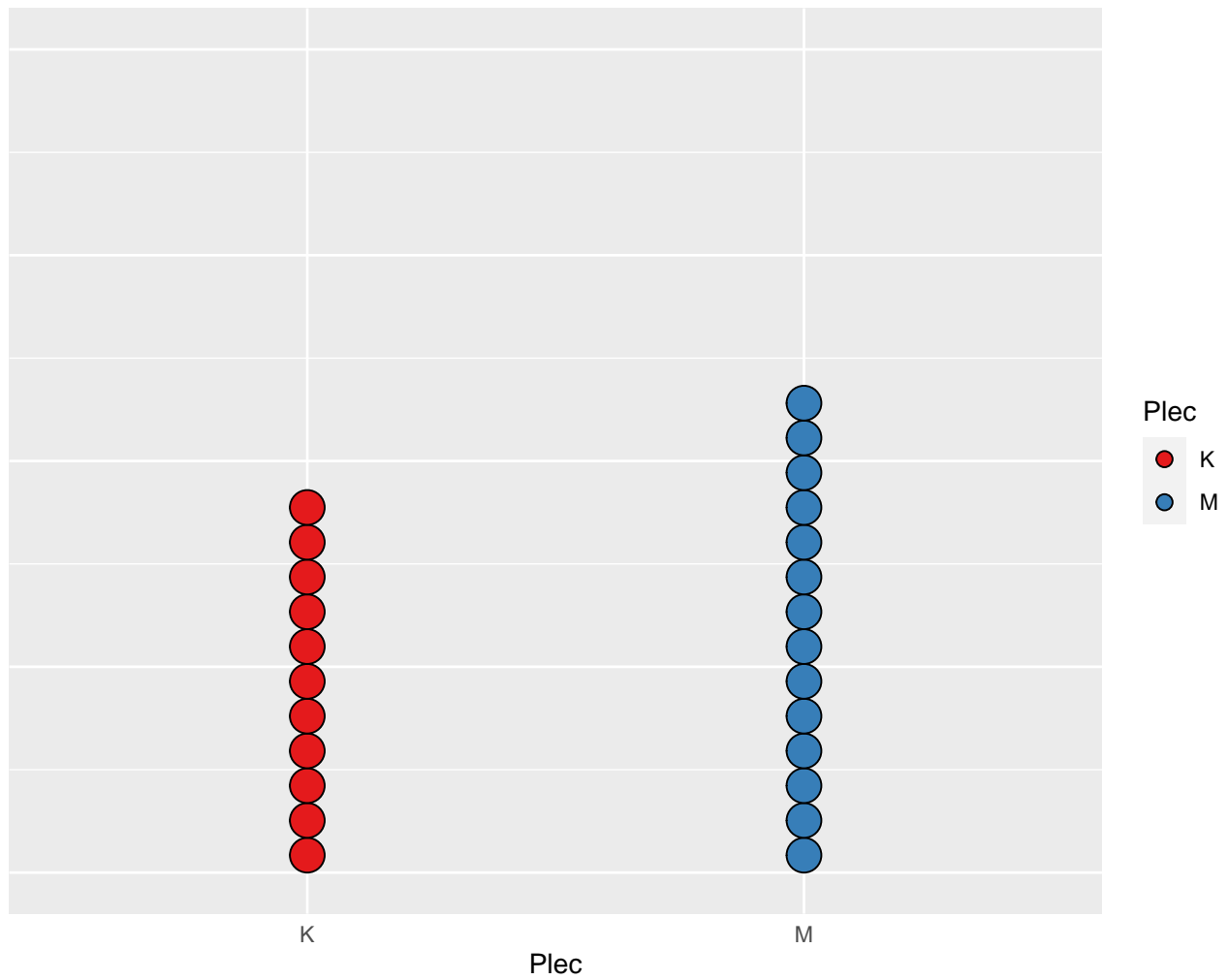
1.4.2 Wykresy punktowe

```
rysunek01 +  
  geom_dotplot(binwidth = .07) +  
  scale_fill_brewer(palette = "Set1") +  
  theme(  
    axis.text.y = element_blank(),  
    axis.ticks = element_blank()) +  
  labs(title = "Wykres punktowy", y = NULL, caption = "Źródło: Opracowanie własne.")
```

```
rysunek03 <- ggplot(Example_01, aes(x=`Prędkość`, fill = `Płeć`))
```

```
rysunek03 + geom_dotplot(binwidth = 1.2) +  
  scale_fill_brewer(palette = "Set1") +  
  theme(  
    axis.text.y = element_blank(),  
    axis.ticks = element_blank())
```

Wykres punktowy

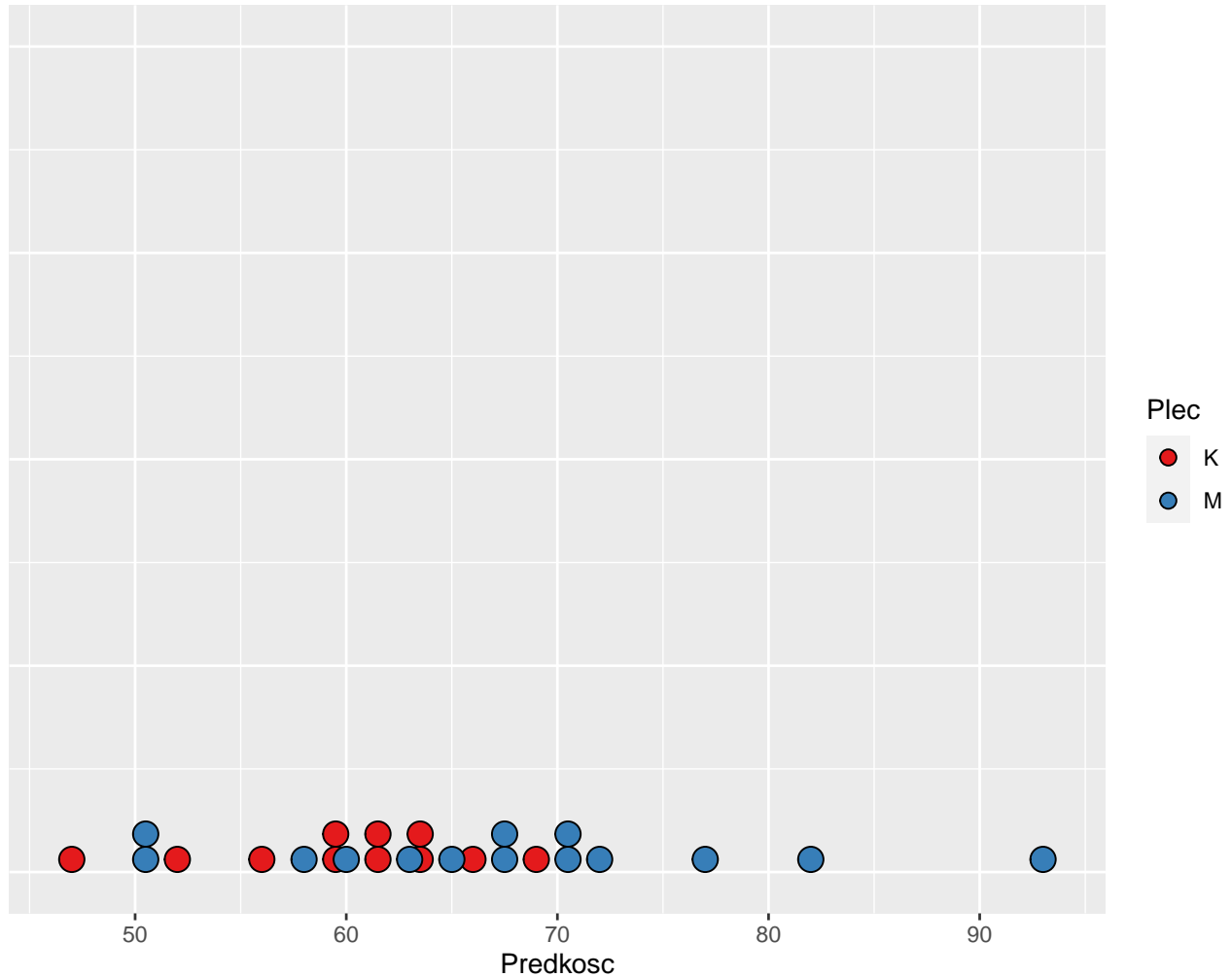


Zródło: Opracowanie własne.

Rysunek 3: Wykres punktowy 1

```
axis.ticks.y = element_blank(),
axis.text.y = element_blank()) +
labs(title = "Wykres punktowy", y = NULL, caption = "Źródło: Opracowanie własne.")
```

Wykres punktowy



Zródło: Opracowanie własne.

Rysunek 4: Wykres punktowy 2

1.4.3 Wykres kołowy

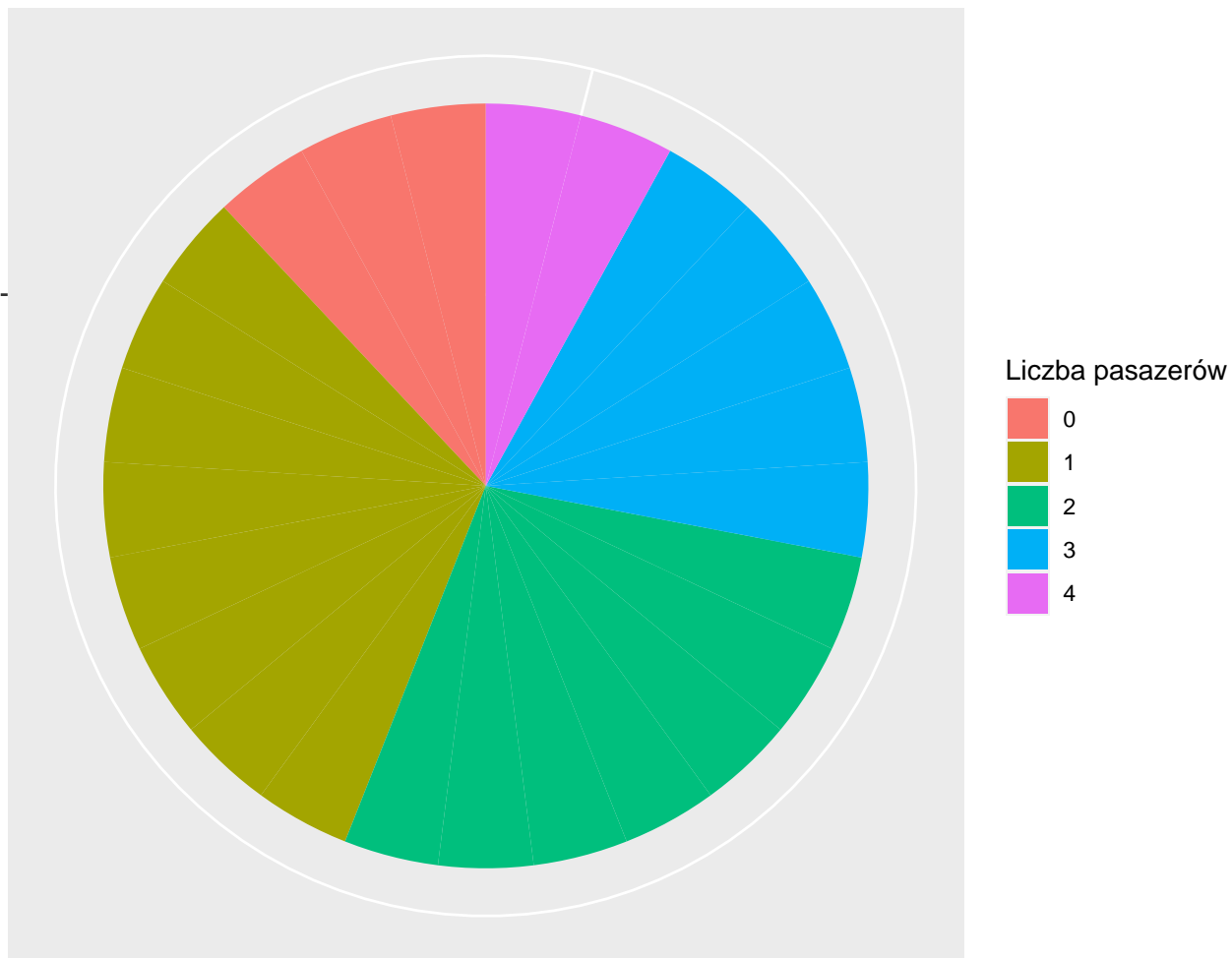
```
rysunek02B <- ggplot(Example_01, aes(x="", y="", fill=factor(`Liczba pasażerów`)))
rysunek02B +
  geom_bar(width = 1, stat = "identity") +
  theme(plot.subtitle = element_text(vjust = 1),
        plot.caption = element_text(vjust = 1),
        plot.title = element_text(hjust = 0.25)) +
  labs(
    title = "Wykres kołowy",
    x = NULL,
```

```

y = NULL,
caption = "Źródło: Opracowanie własne.",
fill = "Liczba pasażerów") +
coord_polar("y", start=0)

```

Wykres kołowy



Zródło: Opracowanie własne.

Rysunek 5: Wykres kołowy

1.4.4 Wykresy *łodyga-liść*

```
stem(Example_01$`Prędkość`)
```

```

##
## The decimal point is 1 digit(s) to the right of the |
##
## 4 | 7
## 5 | 012689
## 6 | 001233456789
## 7 | 0127

```



```
## 8 | 2
## 9 | 3
```

```
stem(Example_01$`Liczba pasażerów`)
```

```
##
## The decimal point is at the |
##
## 0 | 000
## 1 | 0000000
## 2 | 0000000
## 3 | 00000
## 4 | 00
```