

# PWO

Adam Stuller

## PWO - Past weather observation

Jedna sa o kategórickú zakódovanú informáciu o pozorovanom počasí. Z dookumentácie su hodnoty nasledovne:

Domain Value ID: Domain Value Definition Text

0 = Cloud covering 1/2 or less of the sky throughout the appropriate period

1 = Cloud covering more than 1/2 of the sky during part of the appropriate period and covering 1/2 or less during part of the period

2 = Cloud covering more than 1/2 of the sky throughout the appropriate period

3 = Sandstorm, duststorm or blowing snow

4 = Fog or ice fog or thick haze

5 = Drizzle

6 = Rain

7 = Snow, or rain and snow mixed

8 = Shower(s)

9 = Thunderstorm(s) with or without precipitation

### Centralna poloha dat

Vidno, že najčastejšiou formou počasia na slovensku je typ 2, teda zamračené s viac ako polovicou oblohy zakrytou oblakmi.

Relativne pocetnost jednotlivych tried:

```
0 - 0.1720426528 %
1 - 0.1068872541 %
2 - 0.2627742086 %
3 - 0.0002995906 %
4 - 0.1205130766 %
5 - 0.0141029482 %
6 - 0.1845144969 %
7 - 0.0661762258 %
8 - 0.0390577322 %
9 - 0.0336318143 %
```

```
all_data <- read.csv(file= "../data/all.csv")
```

```
unique(all_data$PWO)
```

```
## [1] NA 2 7 6 1 8 0 4 9 5 3
```

```
counts <- table(all_data$PW0)
counts / sum(counts)
```

```
##
##           0           1           2           3           4           5
## 0.1720426528 0.1068872541 0.2627742086 0.0002995906 0.1205130766 0.0141029482
##           6           7           8           9
## 0.1845144969 0.0661762258 0.0390577322 0.0336318143
```

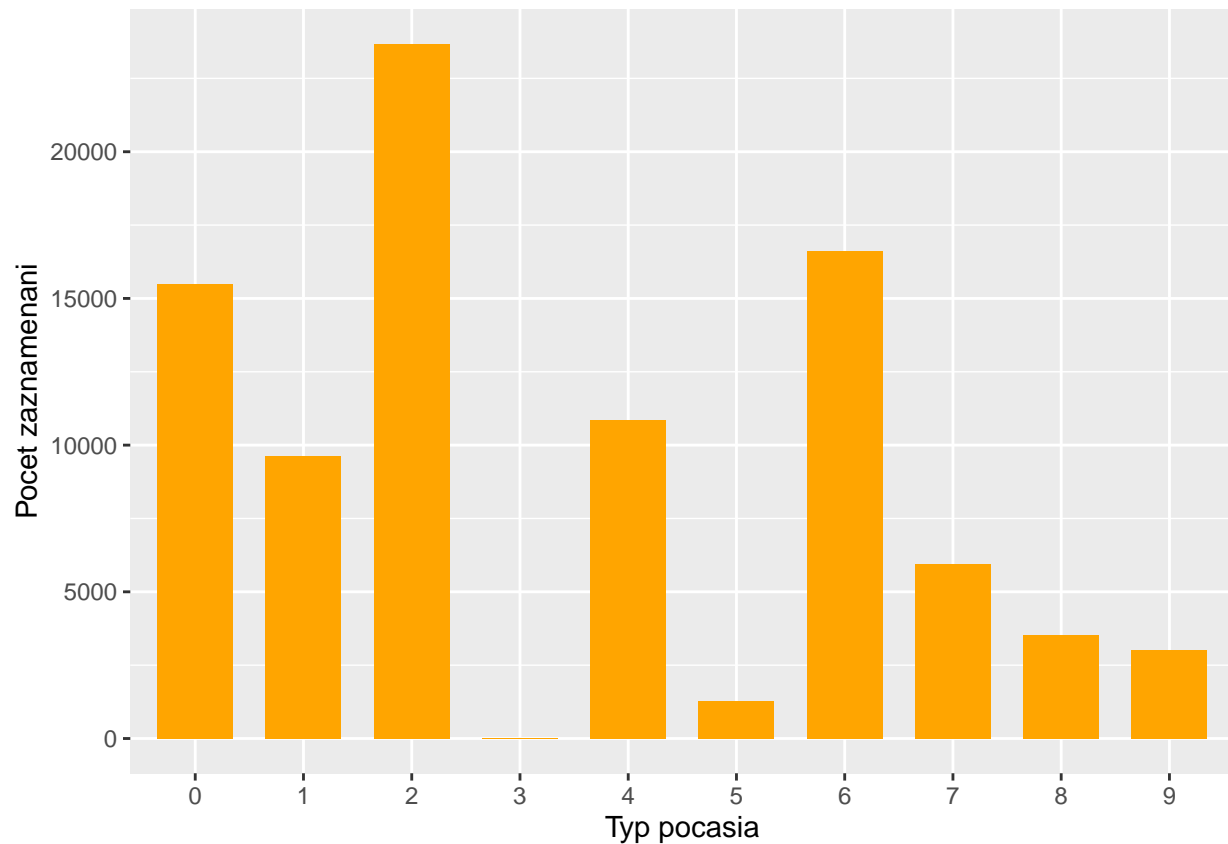
```
getmode(na.omit(all_data$PW0)) %>%
  print(cat("Modus: " ))
```

```
## Modus: [1] 2
```

## Barplot

```
df <- all_data %>%
  dplyr::mutate(
    PW0 = map_chr(.$PW0, as.character)
  ) %>%
  dplyr::filter(!is.na(PW0))

ggplot(df, aes(x = PW0)) +
  geom_bar(fill = "orange", width = 0.7) +
  xlab("Typ pocasia") +
  ylab("Pocet zaznamenani")
```



Jedna sa o kategoricku veličinu, s veľkým množstvom NA hodnot. NA sa tam objavujú iba preto, že sa tato veličina merala v konkrétnych intervaloch a medzi nimi nie.