

Statistika

Patrik Landeka

2025-08-07

Učitavanje podataka

```
path <- "./data.csv"

data1 <- read.csv2(path)
data1[data1 == "/"] <- NA
View(data1)
```

Prepoznavanje okoline i predmeta

Ispisuje koliko puta je objekt prepoznat ako je primljen slikovni ulaz.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
data1 %>%
  filter(vrsta.ulaza != "tekst") %>%
  group_by(naziv.modela, prepoznat.objekt) %>%
  summarise(broj = n())
```

```
## 'summarise()' has grouped output by 'naziv.modela'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 6 x 3
## # Groups:   naziv.modela [3]
##   naziv.modela prepoznat.objekt broj
##   <chr>         <chr>         <int>
```

```
## 1 claude      DA      35
## 2 claude      NE      1
## 3 claude      <NA>     1
## 4 gpt         DA     37
## 5 qwen        DA     36
## 6 qwen        NE      1
```

Ispisuje koliko puta je okolina prepoznata.

```
data1 %>%
  group_by(naziv.modela, prepoznato.okruženje) %>%
  summarise(broj = n())
```

```
## 'summarise()' has grouped output by 'naziv.modela'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 5 x 3
## # Groups:   naziv.modela [3]
##   naziv.modela prepoznato.okruženje broj
##   <chr>        <chr>                <int>
## 1 claude      DA                    73
## 2 claude      <NA>                     1
## 3 gpt         DA                    74
## 4 qwen        DA                    72
## 5 qwen        <NA>                     2
```

```
data1 <- data1 %>%
  filter(!is.na(prepoznato.okruženje))
```

Potruga za predmetom kad se on nalazi u slici

Prikazuje koliko puta je model pronašao predmet ako je on stvarno bio u slici ili nije bio za slikovni unos.

```
library(ggplot2)

data1 %>%
  filter(vrsta.ulaza == "slika") %>%
  group_by(naziv.modela, E..predmet.u.slici) %>%
  summarise(broj = n())
```

```
## 'summarise()' has grouped output by 'naziv.modela'. You can override using the
## '.groups' argument.
```

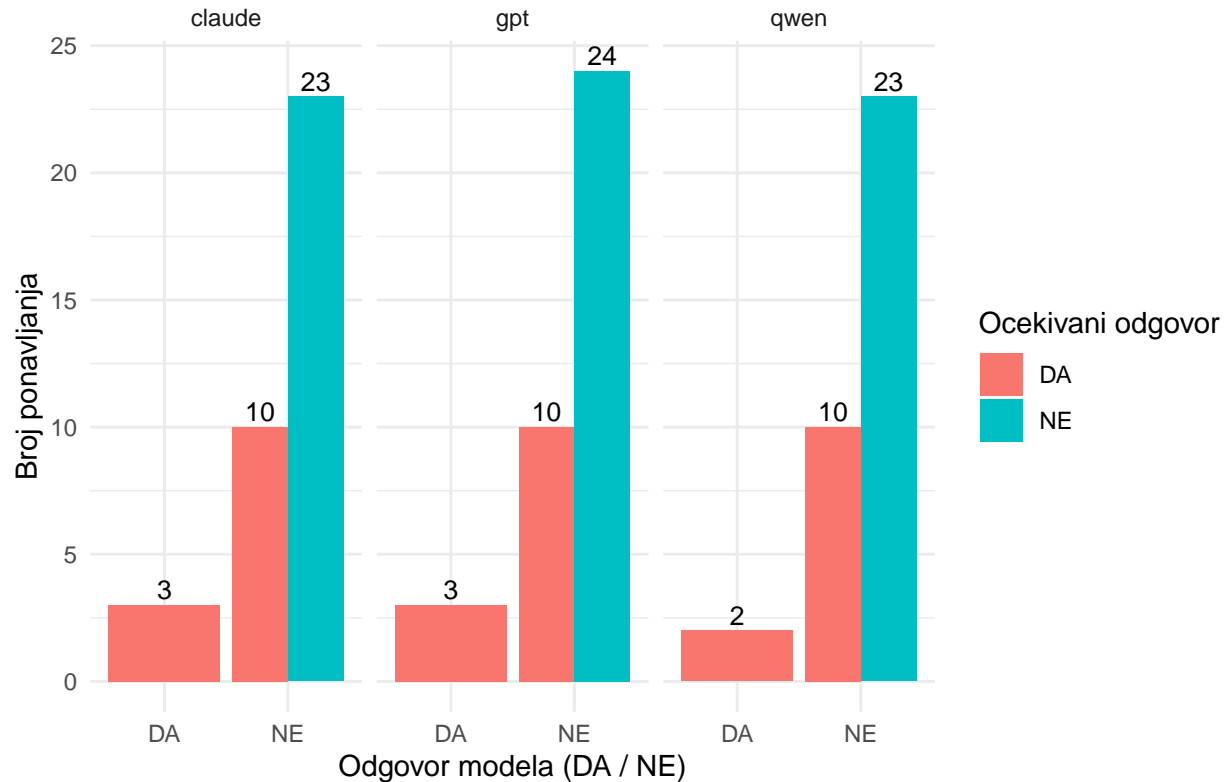
```
## # A tibble: 6 x 3
## # Groups:   naziv.modela [3]
##   naziv.modela E..predmet.u.slici broj
##   <chr>        <chr>                <int>
## 1 claude      DA                    13
## 2 claude      NE                    23
## 3 gpt         DA                    13
## 4 gpt         NE                    24
## 5 qwen        DA                    12
## 6 qwen        NE                    24
```

```
data1 %>%
  filter(vrsta.ulaza == "slika") %>%
  group_by( naziv.modela, E..predmet.u.slici, predmet.u.slici) %>%
  summarise(broj = n(), .groups = "drop") %>%
  filter(predmet.u.slici %in% c("DA", "NE"))
```

```
## # A tibble: 9 x 4
##   naziv.modela E..predmet.u.slici predmet.u.slici broj
##   <chr>      <chr>                <chr>        <int>
## 1 claude     DA                      DA             3
## 2 claude     DA                      NE            10
## 3 claude     NE                      NE            23
## 4 gpt        DA                      DA             3
## 5 gpt        DA                      NE            10
## 6 gpt        NE                      NE            24
## 7 qwen       DA                      DA             2
## 8 qwen       DA                      NE            10
## 9 qwen       NE                      NE            23
```

```
data1 %>%
  filter(vrsta.ulaza == "slika") %>%
  filter(predmet.u.slici %in% c("DA", "NE")) %>%
  group_by(naziv.modela, E..predmet.u.slici, predmet.u.slici) %>%
  summarise(broj = n(), .groups = "drop") %>%
  ggplot(aes(x = predmet.u.slici, y = broj, fill = E..predmet.u.slici)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_text(aes(label = broj), position = position_dodge(width = 0.9), vjust = -0.3, size = 3.5) +
  facet_wrap(~naziv.modela) +
  labs(
    title = "Usporedba modela po stvarnom i dobivenom odgovoru za slikovni unos",
    x = "Odgovor modela (DA / NE)",
    y = "Broj ponavljanja",
    fill = "Očekivani odgovor"
  ) +
  theme_minimal()
```

Usporedba modela po stvarnom i dobivenom odgovoru za slikovni unos



Prikazuje koliko puta je model pronašao predmet ako je on stvarno bio u slici ili nije bio za tekstualni unos.

```
data1 %>%
  filter(vrsta.ulaza == "tekst") %>%
  group_by( naziv.modela, E..predmet.u.slici) %>%
  summarise(broj = n())
```

'summarise()' has grouped output by 'naziv.modela'. You can override using the
'.groups' argument.

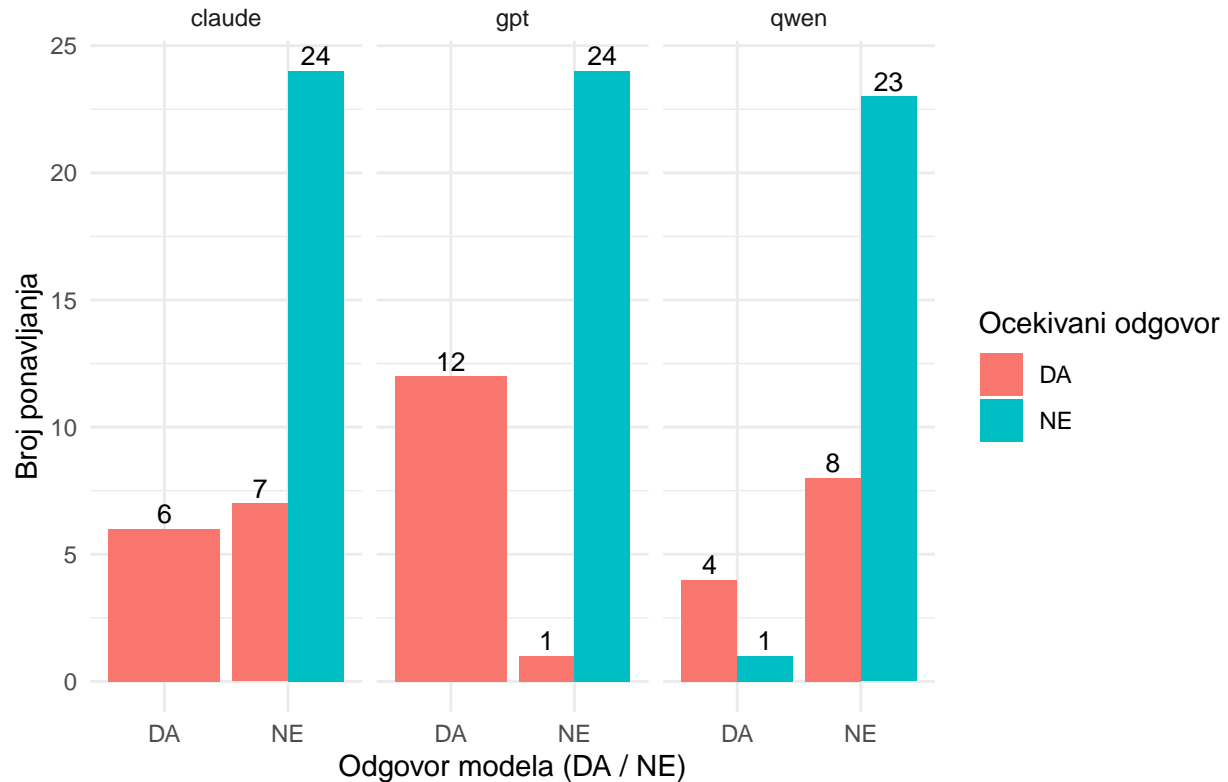
```
## # A tibble: 6 x 3
## # Groups:   naziv.modela [3]
##   naziv.modela E..predmet.u.slici broj
##   <chr>        <chr>             <int>
## 1 claude      DA                 13
## 2 claude      NE                 24
## 3 gpt         DA                 13
## 4 gpt         NE                 24
## 5 qwen       DA                 12
## 6 qwen       NE                 24
```

```
data1 %>%
  filter(vrsta.ulaza == "tekst") %>%
  group_by( naziv.modela, E..predmet.u.slici, predmet.u.slici) %>%
  summarise(broj = n(), .groups = "drop") %>%
  filter(predmet.u.slici %in% c("DA", "NE"))
```

```
## # A tibble: 10 x 4
##   naziv.modela E..predmet.u.slici predmet.u.slici broj
##   <chr>        <chr>                <chr>      <int>
## 1 claude      DA                DA          6
## 2 claude      DA                NE          7
## 3 claude      NE                NE         24
## 4 gpt         DA                DA         12
## 5 gpt         DA                NE          1
## 6 gpt         NE                NE         24
## 7 qwen        DA                DA          4
## 8 qwen        DA                NE          8
## 9 qwen        NE                DA          1
## 10 qwen       NE                NE         23
```

```
data1 %>%
  filter(vrsta.ulaza == "tekst") %>%
  filter(predmet.u.slici %in% c("DA", "NE")) %>%
  group_by(naziv.modela, E..predmet.u.slici, predmet.u.slici) %>%
  summarise(broj = n(), .groups = "drop") %>%
  ggplot(aes(x = predmet.u.slici, y = broj, fill = E..predmet.u.slici)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_text(aes(label = broj), position = position_dodge(width = 0.9), vjust = -0.3, size = 3.5) +
  facet_wrap(~naziv.modela) +
  labs(
    title = "Usporedba modela po stvarnom i dobivenom odgovoru za tekstualni unos",
    x = "Odgovor modela (DA / NE)",
    y = "Broj ponavljanja",
    fill = "Očekivani odgovor"
  ) +
  theme_minimal()
```

Usporedba modela po stvarnom i dobivenom odgovoru za tekstualni unos



Potruga za predmetom kad se ne nalazi u slici

Visoka vjerojatnost

Slikovni unos

```
library(readr)

data1 <- data1 %>%
  mutate(
    ocjena_koordinata = as.numeric(ocjena.koordinata..1.5.),
    ocjena_uputa = as.numeric(ocjena.uputa..1.5.),
    ocjena_vjerojatnosti = as.numeric(ocjena.vjerojatnosti..1.5.),
    vrijeme_izvođenja = as.numeric(vrijeme.izvođenja)
  )

# Filtriraj testove s rbr od 14 do 22
filtered_data <- data1 %>%
  filter(rbr >= 14 & rbr <= 22)

# Izračunaj prosjeke i standardne devijacije za filtrirane podatke
statistika_po_modelu <- filtered_data %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = round(mean(ocjena_koordinata, na.rm = TRUE), 2),
```

```

sd_koordinata = round(sd(ocjena_koordinata, na.rm = TRUE), 2),
prosjek_uputa = round(mean(ocjena_uputa, na.rm = TRUE), 2),
sd_uputa = round(sd(ocjena_uputa, na.rm = TRUE), 2),
prosjek_vjerojatnosti = round(mean(ocjena_vjerojatnosti, na.rm = TRUE), 2),
sd_vjerojatnosti = round(sd(ocjena_vjerojatnosti, na.rm = TRUE), 2)
)

# Prikaz rezultata
print(statistika_po_modelu)

```

```

## # A tibble: 3 x 7
##   naziv.modela prosjek_koordinata sd_koordinata prosjek_uputa sd_uputa
##   <chr>          <dbl>          <dbl>          <dbl>    <dbl>
## 1 claude          3.67            0.87            4.56     1.01
## 2 gpt              3.33            1              4.33     0.87
## 3 qwen             1.89            1.54            3.78     1.09
## # i 2 more variables: prosjek_vjerojatnosti <dbl>, sd_vjerojatnosti <dbl>

```

Tekstualni unos

```

filtered_data <- data1 %>%
  filter(rbr >= 51 & rbr <= 59)

# Izračunaj prosjeke i standardne devijacije za filtrirane podatke
statistika_po_modelu <- filtered_data %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = round(mean(ocjena_koordinata, na.rm = TRUE), 2),
    sd_koordinata = round(sd(ocjena_koordinata, na.rm = TRUE), 2),
    prosjek_uputa = round(mean(ocjena_uputa, na.rm = TRUE), 2),
    sd_uputa = round(sd(ocjena_uputa, na.rm = TRUE), 2),
    prosjek_vjerojatnosti = round(mean(ocjena_vjerojatnosti, na.rm = TRUE), 2),
    sd_vjerojatnosti = round(sd(ocjena_vjerojatnosti, na.rm = TRUE), 2)
  )

# Prikaz rezultata
print(statistika_po_modelu)

```

```

## # A tibble: 3 x 7
##   naziv.modela prosjek_koordinata sd_koordinata prosjek_uputa sd_uputa
##   <chr>          <dbl>          <dbl>          <dbl>    <dbl>
## 1 claude          4.11            0.6            4.89     0.33
## 2 gpt              4.22            1.3            4.67     0.71
## 3 qwen              2              1.32           3.44     1.01
## # i 2 more variables: prosjek_vjerojatnosti <dbl>, sd_vjerojatnosti <dbl>

```

Srednje visoka vjerojatnost

Slikovni unos

```

filtered_data <- data1 %>%
  filter(rbr >= 23 & rbr <= 32)

# Izračunaj prosjeke i standardne devijacije za filtrirane podatke
statistika_po_modelu <- filtered_data %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = round(mean(ocjena_koordinata, na.rm = TRUE), 2),
    sd_koordinata = round(sd(ocjena_koordinata, na.rm = TRUE), 2),
    prosjek_uputa = round(mean(ocjena_uputa, na.rm = TRUE), 2),
    sd_uputa = round(sd(ocjena_uputa, na.rm = TRUE), 2),
    prosjek_vjerojatnosti = round(mean(ocjena_vjerojatnosti, na.rm = TRUE), 2),
    sd_vjerojatnosti = round(sd(ocjena_vjerojatnosti, na.rm = TRUE), 2)
  )

# Prikaz rezultata
print(statistika_po_modelu)

```

```

## # A tibble: 3 x 7
##   naziv.modela prosjek_koordinata sd_koordinata prosjek_uputa sd_uputa
##   <chr>          <dbl>          <dbl>          <dbl>    <dbl>
## 1 claude          3.33            1.32            4.33     1.32
## 2 gpt              2.7            1.57            4.1      0.74
## 3 qwen             1.4            1.26             3      1.49
## # i 2 more variables: prosjek_vjerojatnosti <dbl>, sd_vjerojatnosti <dbl>

```

Tekstualni unos

```

filtered_data <- data1 %>%
  filter(rbr >= 60 & rbr <= 69)

# Izračunaj prosjeke i standardne devijacije za filtrirane podatke
statistika_po_modelu <- filtered_data %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = round(mean(ocjena_koordinata, na.rm = TRUE), 2),
    sd_koordinata = round(sd(ocjena_koordinata, na.rm = TRUE), 2),
    prosjek_uputa = round(mean(ocjena_uputa, na.rm = TRUE), 2),
    sd_uputa = round(sd(ocjena_uputa, na.rm = TRUE), 2),
    prosjek_vjerojatnosti = round(mean(ocjena_vjerojatnosti, na.rm = TRUE), 2),
    sd_vjerojatnosti = round(sd(ocjena_vjerojatnosti, na.rm = TRUE), 2)
  )

# Prikaz rezultata
print(statistika_po_modelu)

```

```

## # A tibble: 3 x 7
##   naziv.modela prosjek_koordinata sd_koordinata prosjek_uputa sd_uputa
##   <chr>          <dbl>          <dbl>          <dbl>    <dbl>
## 1 claude          3.4            1.17            4.4     0.97
## 2 gpt              3.3            1.34            4.3     0.95
## 3 qwen             1.75           1.39            2.78     1.2
## # i 2 more variables: prosjek_vjerojatnosti <dbl>, sd_vjerojatnosti <dbl>

```


Slikovni i tekstualni unos

```
filtered_data <- data1 %>%
  filter( (rbr >= 23 & rbr <= 32) | (rbr >= 60 & rbr <= 69))

# Izračunaj prosjeke i standardne devijacije za filtrirane podatke
statistika_po_modelu <- filtered_data %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = round(mean(ocjena_koordinata, na.rm = TRUE), 2),
    sd_koordinata = round(sd(ocjena_koordinata, na.rm = TRUE), 2),
    prosjek_uputa = round(mean(ocjena_uputa, na.rm = TRUE), 2),
    sd_uputa = round(sd(ocjena_uputa, na.rm = TRUE), 2),
    prosjek_vjerojatnosti = round(mean(ocjena_vjerojatnosti, na.rm = TRUE), 2),
    sd_vjerojatnosti = round(sd(ocjena_vjerojatnosti, na.rm = TRUE), 2)
  )

# Prikaz rezultata
print(statistika_po_modelu)
```

```
## # A tibble: 3 x 7
##   naziv.modela prosjek_koordinata sd_koordinata prosjek_uputa sd_uputa
##   <chr>          <dbl>          <dbl>          <dbl>    <dbl>
## 1 claudé         3.37            1.21            4.37     1.12
## 2 gpt             3              1.45            4.2       0.83
## 3 qwen           1.56            1.29            2.89     1.33
## # i 2 more variables: prosjek_vjerojatnosti <dbl>, sd_vjerojatnosti <dbl>
```

Vizualno

```
library(tidyr)

# Prvi raspon: rbr 23-32 i 60-69
filtered_data1 <- data1 %>%
  filter((rbr >= 23 & rbr <= 32) | (rbr >= 60 & rbr <= 69)) %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = mean(ocjena_koordinata, na.rm = TRUE),
    sd_koordinata = sd(ocjena_koordinata, na.rm = TRUE),
    prosjek_uputa = mean(ocjena_uputa, na.rm = TRUE),
    sd_uputa = sd(ocjena_uputa, na.rm = TRUE),
    prosjek_vjerojatnosti = mean(ocjena_vjerojatnosti, na.rm = TRUE),
    sd_vjerojatnosti = sd(ocjena_vjerojatnosti, na.rm = TRUE)
  ) %>%
  pivot_longer(cols = starts_with("prosjek"), names_to = "kategorija", values_to = "vrijednost") %>%
  mutate(sd = case_when(
    kategorija == "prosjek_koordinata" ~ sd_koordinata,
    kategorija == "prosjek_uputa" ~ sd_uputa,
    kategorija == "prosjek_vjerojatnosti" ~ sd_vjerojatnosti
  ),
  raspon = "Srednja vjerojatnost")

# Drugi raspon: rbr 51-59 i 14-22
```

```

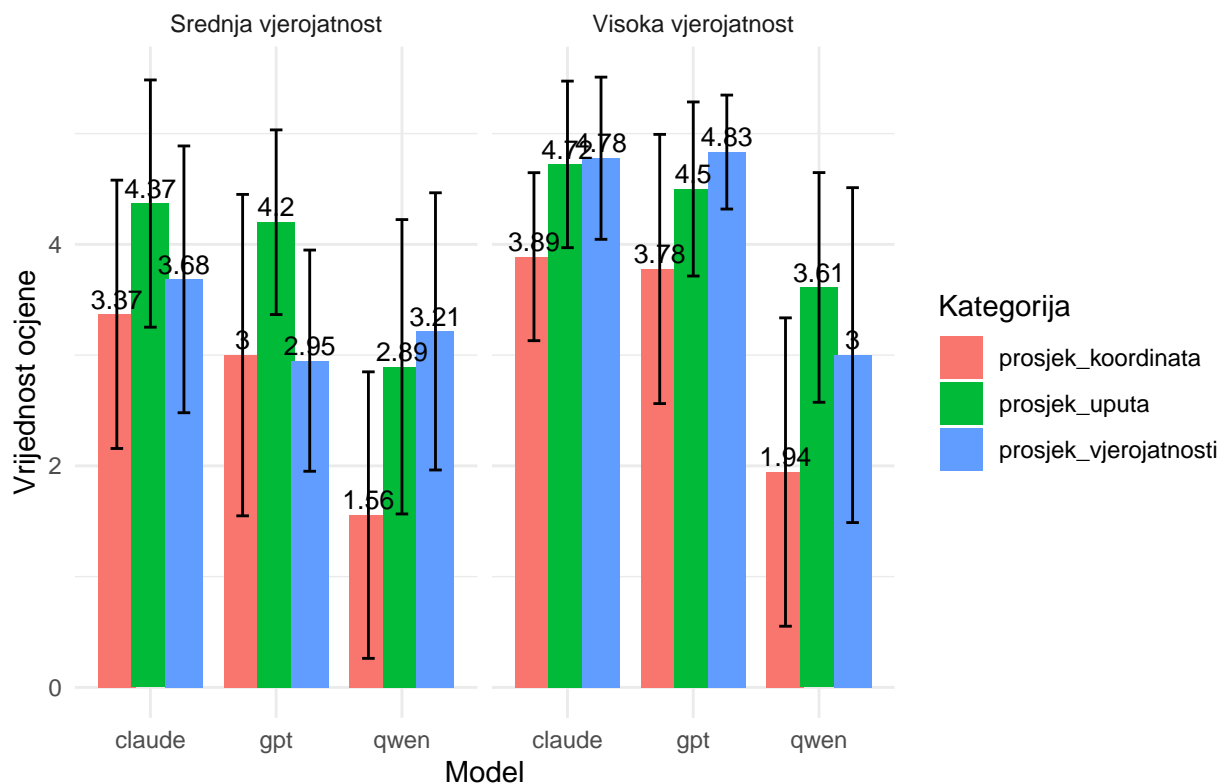
filtered_data2 <- data1 %>%
  filter((rbr >= 51 & rbr <= 59) | (rbr >= 14 & rbr <= 22)) %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = mean(ocjena_koordinata, na.rm = TRUE),
    sd_koordinata = sd(ocjena_koordinata, na.rm = TRUE),
    prosjek_uputa = mean(ocjena_uputa, na.rm = TRUE),
    sd_uputa = sd(ocjena_uputa, na.rm = TRUE),
    prosjek_vjerojatnosti = mean(ocjena_vjerojatnosti, na.rm = TRUE),
    sd_vjerojatnosti = sd(ocjena_vjerojatnosti, na.rm = TRUE)
  ) %>%
  pivot_longer(cols = starts_with("prosjek"), names_to = "kategorija", values_to = "vrijednost") %>%
  mutate(sd = case_when(
    kategorija == "prosjek_koordinata" ~ sd_koordinata,
    kategorija == "prosjek_uputa" ~ sd_uputa,
    kategorija == "prosjek_vjerojatnosti" ~ sd_vjerojatnosti
  ),
  raspon = "Visoka vjerojatnost")

# Kombiniraj oba seta podataka
viz_data <- bind_rows(filtered_data1, filtered_data2)

# Vizualizacija
ggplot(viz_data, aes(x = naziv.modela, y = vrijednost, fill = kategorija)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8)) +
  geom_errorbar(aes(ymin = vrijednost - sd, ymax = vrijednost + sd),
    position = position_dodge(width = 0.8), width = 0.3) +
  geom_text(aes(label = round(vrijednost, 2)),
    position = position_dodge(width = 0.8),
    vjust = -0.3,
    size = 3.5) +
  facet_wrap(~raspon) +
  labs(title = "Prosjeci i SD po modelima za različite skupine primjera",
    x = "Model",
    y = "Vrijednost ocjene",
    fill = "Kategorija") +
  theme_minimal()

```

Prosjeci i SD po modelima za razlicite skupine primjera



Niska vjerojatnost

```
# Filtriranje samo jednog raspona
filtered_data <- data1 %>%
  filter((rbr >= 33 & rbr <= 37) | (rbr >= 70 & rbr <= 74)) %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjeck_koordinata = mean(ocjena_koordinata, na.rm = TRUE),
    sd_koordinata = sd(ocjena_koordinata, na.rm = TRUE),
    prosjeck_uputa = mean(ocjena_uputa, na.rm = TRUE),
    sd_uputa = sd(ocjena_uputa, na.rm = TRUE),
    prosjeck_vjerojatnosti = mean(ocjena_vjerojatnosti, na.rm = TRUE),
    sd_vjerojatnosti = sd(ocjena_vjerojatnosti, na.rm = TRUE)
  ) %>%
  pivot_longer(
    cols = starts_with("prosjeck"),
    names_to = "kategorija",
    values_to = "vrijednost"
  ) %>%
  mutate(
    sd = case_when(
      kategorija == "prosjeck_koordinata" ~ sd_koordinata,
      kategorija == "prosjeck_uputa" ~ sd_uputa,
      kategorija == "prosjeck_vjerojatnosti" ~ sd_vjerojatnosti
    )
  )
```

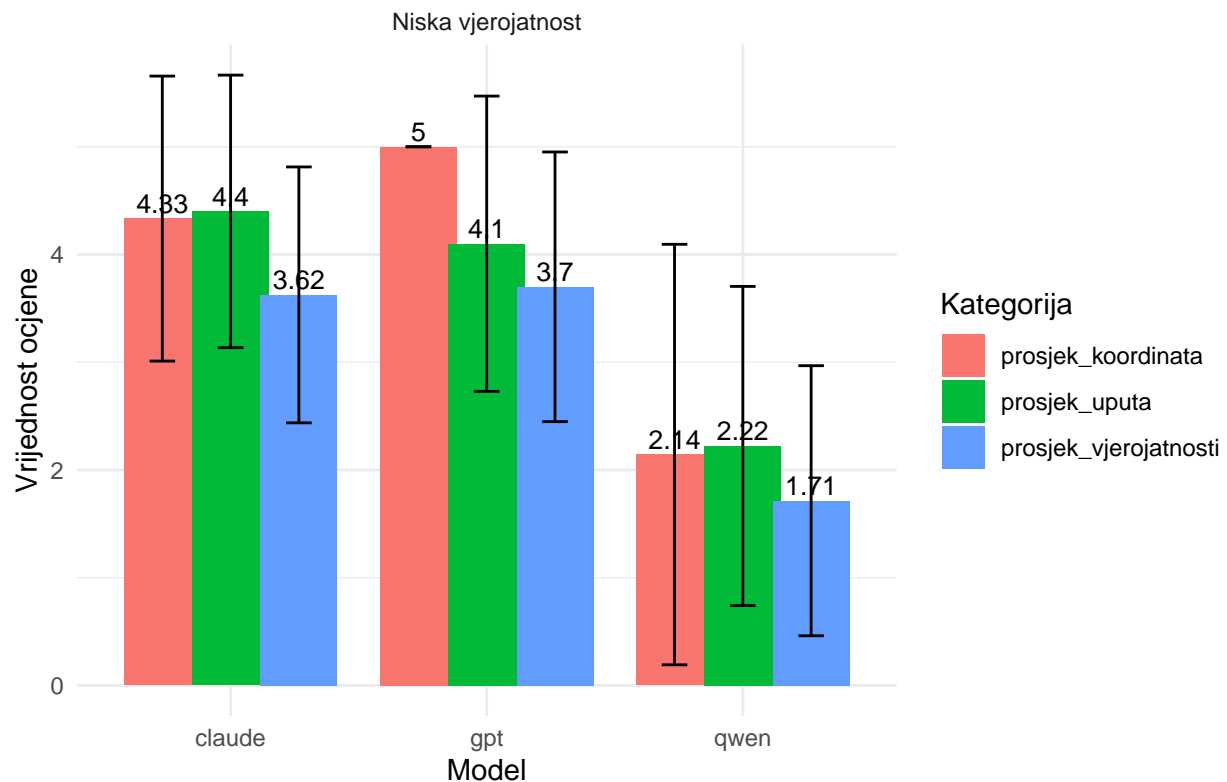
```

    ),
    raspon = "Niska vjerojatnost"
)

# Vizualizacija s jednim rasponom
ggplot(filtered_data, aes(x = naziv.modela, y = vrijednost, fill = kategorija)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8)) +
  geom_errorbar(
    aes(ymin = vrijednost - sd, ymax = vrijednost + sd),
    position = position_dodge(width = 0.8),
    width = 0.3
  ) +
  geom_text(
    aes(label = round(vrijednost, 2)),
    position = position_dodge(width = 0.8),
    vjust = -0.3,
    size = 3.5
  ) +
  facet_wrap(~raspon) +
  labs(
    title = "Prosjeci i SD po modelima - niska vjerojatnost",
    x = "Model",
    y = "Vrijednost ocjene",
    fill = "Kategorija"
  ) +
  theme_minimal()

```

Prosjeci i SD po modelima – niska vjerojatnost



Brzina izvođenja

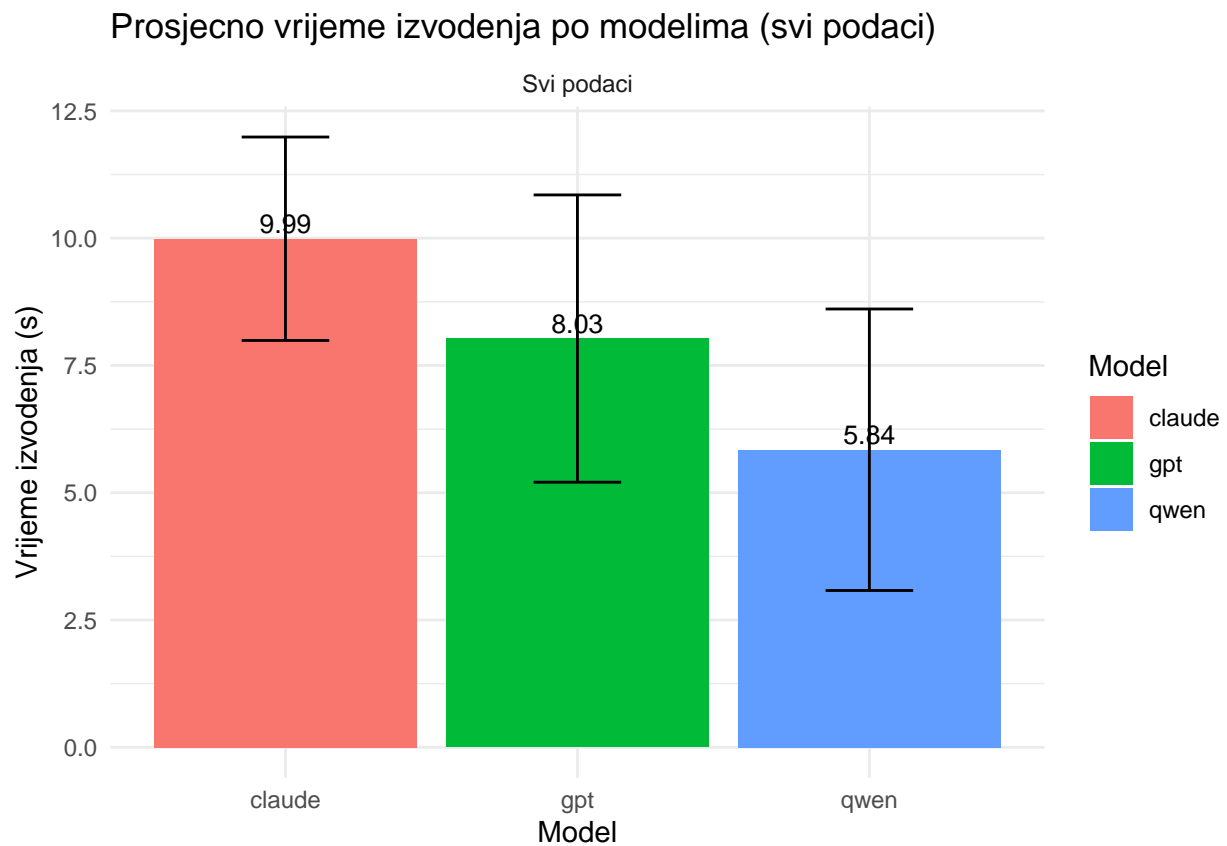
```
# Agregacija podataka bez filtriranja
filtered_data <- data1 %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_vremena = mean(vrijeme_izvođenja, na.rm = TRUE),
    sd_vremena = sd(vrijeme_izvođenja, na.rm = TRUE)
  ) %>%
  filter(!is.na(prosjek_vremena), !is.na(sd_vremena)) %>%
  mutate(raspon = "Svi podaci")

# Vizualizacija
ggplot(filtered_data, aes(x = naziv.modela, y = prosjek_vremena, fill = naziv.modela)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8)) +
  geom_errorbar(
    aes(ymin = prosjek_vremena - sd_vremena, ymax = prosjek_vremena + sd_vremena),
    width = 0.3,
    position = position_dodge(width = 0.8)
  ) +
  geom_text(
    aes(label = round(prosjek_vremena, 2)),
    vjust = -0.3,
```

```

position = position_dodge(width = 0.8),
size = 3.5
) +
facet_wrap(~raspon) +
labs(
  title = "Prosječno vrijeme izvođenja po modelima (svi podaci)",
  x = "Model",
  y = "Vrijeme izvođenja (s)",
  fill = "Model"
) +
theme_minimal()

```



Usporedba slikovnog i tekstualnog upita

```

#Pretvori ocjene u numerički format
data1 <- data1 %>%
  mutate(
    grupa = ifelse(rbr <= 37, "slikovni", "tekstualni")
  )

# Grupiraj po modelu i grupi, izračunaj prosjeke i SD
statistika <- data1 %>%
  group_by(grupa) %>%

```

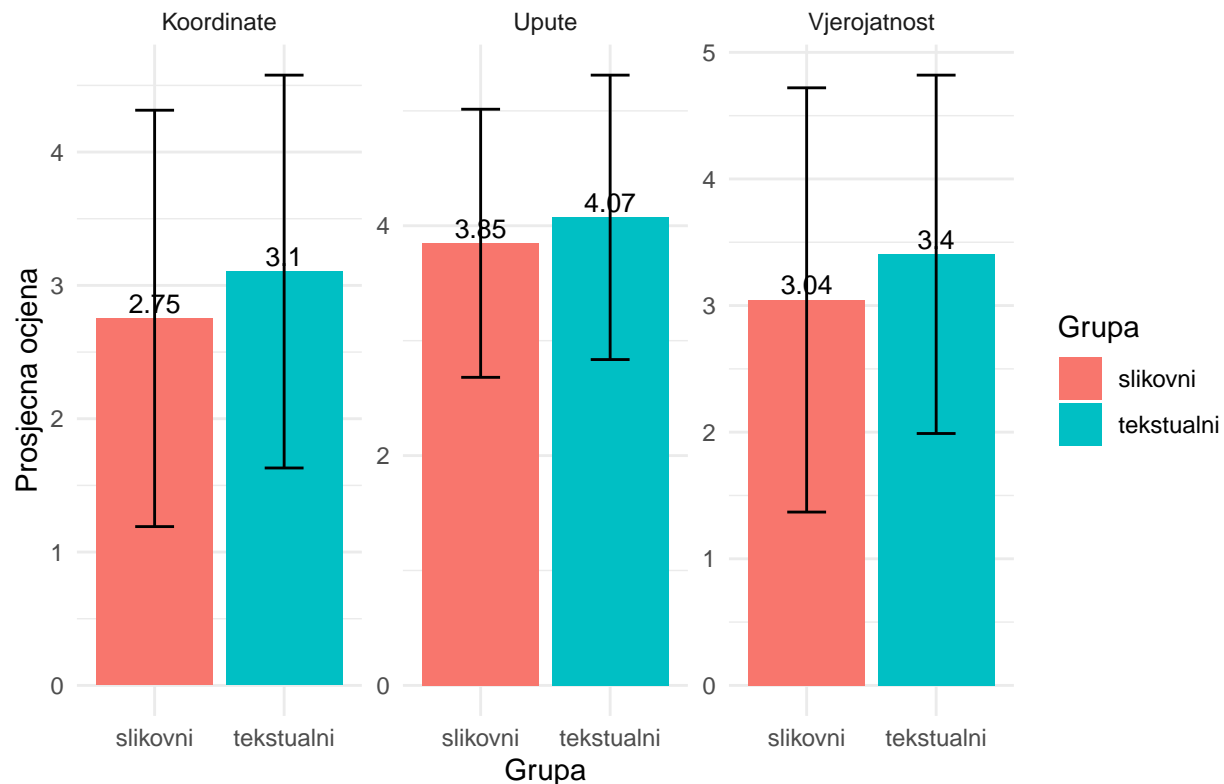
```

summarise(
  prosjek_koordinata = mean(ocjena_koordinata, na.rm = TRUE),
  sd_koordinata = sd(ocjena_koordinata, na.rm = TRUE),
  prosjek_uputa = mean(ocjena_uputa, na.rm = TRUE),
  sd_uputa = sd(ocjena_uputa, na.rm = TRUE),
  prosjek_vjerojatnosti = mean(ocjena_vjerojatnosti, na.rm = TRUE),
  sd_vjerojatnosti = sd(ocjena_vjerojatnosti, na.rm = TRUE),
  .groups = "drop"
) %>%
pivot_longer(
  cols = starts_with("prosjek"),
  names_to = "kategorija",
  values_to = "vrijednost"
) %>%
mutate(
  sd = case_when(
    kategorija == "prosjek_koordinata" ~ sd_koordinata,
    kategorija == "prosjek_uputa" ~ sd_uputa,
    kategorija == "prosjek_vjerojatnosti" ~ sd_vjerojatnosti
  ),
  kategorija = case_when(
    kategorija == "prosjek_koordinata" ~ "Koordinate",
    kategorija == "prosjek_uputa" ~ "Upute",
    kategorija == "prosjek_vjerojatnosti" ~ "Vjerojatnost"
  )
)

# Vizualizacija
ggplot(statistika, aes(x = grupa, y = vrijednost, fill = grupa)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8)) +
  geom_errorbar(
    aes(ymin = vrijednost - sd, ymax = vrijednost + sd),
    position = position_dodge(width = 0.8),
    width = 0.3
  ) +
  geom_text(
    aes(label = round(vrijednost, 2)),
    position = position_dodge(width = 0.8),
    vjust = -0.3,
    size = 3.5
  ) +
  facet_wrap(~kategorija, scales = "free_y") +
  labs(
    title = "Usporedba prosječnih ocjena za slikovni i tekstualni upit",
    x = "Grupa",
    y = "Prosječna ocjena",
    fill = "Grupa"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 0, hjust = 0.5))

```

Usporedba prosječnih ocjena za slikovni i tekstualni upit



Usporedba različitih jezičnih modela

```
# Grupiraj po modelu i izračunaj prosjeke i SD
statistika <- data1 %>%
  group_by(naziv.modela) %>%
  summarise(
    prosjek_koordinata = mean(ocjena_koordinata, na.rm = TRUE),
    sd_koordinata = sd(ocjena_koordinata, na.rm = TRUE),
    prosjek_uputa = mean(ocjena_uputa, na.rm = TRUE),
    sd_uputa = sd(ocjena_uputa, na.rm = TRUE),
    prosjek_vjerojatnosti = mean(ocjena_vjerojatnosti, na.rm = TRUE),
    sd_vjerojatnosti = sd(ocjena_vjerojatnosti, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  pivot_longer(
    cols = starts_with("prosjek"),
    names_to = "kategorija",
    values_to = "vrijednost"
  ) %>%
  mutate(
    sd = case_when(
      kategorija == "prosjek_koordinata" ~ sd_koordinata,
      kategorija == "prosjek_uputa" ~ sd_uputa,
      kategorija == "prosjek_vjerojatnosti" ~ sd_vjerojatnosti
    )
  )
```



```

    ),
    kategorija = case_when(
      kategorija == "prosjek_koordinata" ~ "Koordinate",
      kategorija == "prosjek_uputa" ~ "Upute",
      kategorija == "prosjek_vjerojatnosti" ~ "Vjerojatnost"
    )
  )

# Vizualizacija
ggplot(statistika, aes(x = naziv.modela, y = vrijednost, fill = naziv.modela)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8)) +
  geom_errorbar(
    aes(ymin = vrijednost - sd, ymax = vrijednost + sd),
    position = position_dodge(width = 0.8),
    width = 0.3
  ) +
  geom_text(
    aes(label = round(vrijednost, 2)),
    position = position_dodge(width = 0.8),
    vjust = -0.3,
    size = 3.5
  ) +
  facet_wrap(~kategorija, scales = "free_y") +
  labs(
    title = "Prosjeci ocjena i SD po modelima (svi podaci)",
    x = "Model",
    y = "Prosječna ocjena",
    fill = "Model"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

```

Prosjeci ocjena i SD po modelima (svi podaci)

