Laryngectomy Analysis

AUTHOR
Your Name

Introduction

This document explores the relationship between hospital stay in days and the number of complications among patients in the `analysis` dataset from the `Open.Visualization.Academy` package.

Methods

The data for this analysis come from the `analysis` dataset in the `Open.Visualization.Academy` R package.

Results

Scatterplot with linear trend (no confidence bounds)

```
ggplot(analysis, aes(x = complications_count, y = hospital_stay_days)) +
  geom_point() +
  geom_smooth(method = 'lm', se = FALSE) +
  labs(
    x = 'Complication count',
    y = 'Hospital stay (days)'
  )
```

```
`geom_smooth()` using formula = 'y ~ x'
```

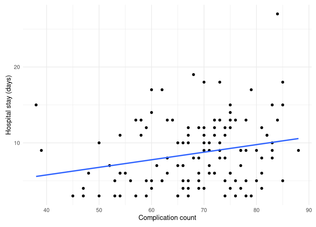

Figure 1: Scatterplot of hospital stay (days) versus complication count with linear trend.

Scatterplot with transparency and linear trend

```
ggplot(analysis, aes(x = complications_count, y = hospital_stay_days)) +
  geom_point(alpha = 0.3) +
  geom_smooth(method = 'lm', se = FALSE) +
  labs(
    x = 'Complication count',
    y = 'Hospital stay (days)'
  )
```

```
`geom_smooth()` using formula = 'y ~ x'
```

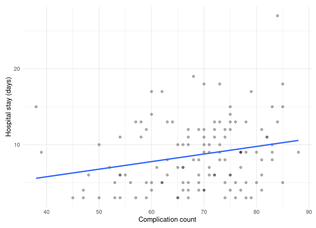

Figure 2: Scatterplot of hospital stay versus complication count using point transparency to reduce overplotting.

Sunflower-like plot for overlapping points

```
library(dplyr)
library(tidyr)

# Calculate counts at each (x, y) position
counted_data <- analysis |>
  group_by(complications_count, hospital_stay_days) |>
  summarise(n = n(), .groups = 'drop')

# Create sunflower-like plot with radiating lines
create_sunflower <- function(data) {
  data |>
    filter(n > 1) |>
    rowwise() |>
    mutate(
      angles = list(seq(0, 2 * pi, length.out = n + 1)[1:n]),
      x_end  = list(complications_count + 0.03 * cos(angles)),
      y_end  = list(hospital_stay_days + 0.03 * sin(angles))
    ) |>
    unnest(cols = c(angles, x_end, y_end))
}

sunflower_segments <- create_sunflower(counted_data)

ggplot() +
  geom_segment(
    data = sunflower_segments,
    aes(x = complications_count, y = hospital_stay_days,
        xend = x_end, yend = y_end),
    alpha = 0.7
  ) +
  geom_point(
    data = counted_data,
    aes(x = complications_count, y = hospital_stay_days),
    size = 1
  ) +
  labs(
    x = 'Complication count',
    y = 'Hospital stay (days)'
  )
```

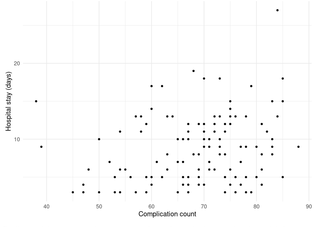

Figure 3: Sunflower plot showing overlapping combinations of complication count and hospital stay.

Hospital stay vs. complication count as a categorical variable

Side-by-side boxplot

```
ggplot(analysis,
       aes(x = factor(complications_count), y = hospital_stay_days)) +
  geom_boxplot() +
  labs(
    x = 'Complication count (categorical)',
    y = 'Hospital stay (days)'
  )
```

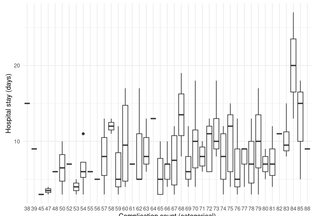

Figure 4: Hospital stay (days) by complication count: side-by-side boxplots.

Violin plot

```
ggplot(analysis,
       aes(x = factor(complications_count), y = hospital_stay_days)) +
  geom_violin(trim = FALSE) +
  labs(
    x = 'Complication count (categorical)',
    y = 'Hospital stay (days)'
  )
```

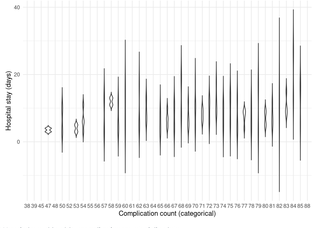

Figure 5: Hospital stay (days) by complication count: violin plots.

Discussion

These visualizations show how hospital stay in days changes with the number of complications. The scatterplots (Figures Figure 1 and Figure 2) and sunflower plot (Figure Figure 3) provide a detailed view of individual patient-level relationships. The boxplot and violin plot (Figures Figure 4 and Figure 5) summarize the distribution of hospital stays for each complication count.