DailyCheck5

2025-04-14

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```r
google <- c(46,49,51,42)
waze <- c(44,47,47,43)
gut <- c(50,51,45,43)

library(ggplot2)
library(gridExtra)
theme_update(text = element_text(size = 20))

g_boxplot <- ggplot() +
  geom_boxplot(aes(x = "Google Maps", y = google)) +
  geom_boxplot(aes(x = "Waze", y = waze)) +
  geom_boxplot(aes(x = "Follow their gut", y = gut)) +
  scale_x_discrete(limits = c("Google Maps", "Waze", "Follow their gut")) +
  ggtitle("Comparison of Commute Times by Method") +
  ylab("Commute Time (minutes)") +
  xlab("Method") +
  theme(axis.text.x = element_text(size = 10),
        axis.title.x = element_text(size = 12),
        axis.text.y = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        plot.title = element_text(size = 11))

g_point <- ggplot() +
  geom_point(aes(x = "Google Maps", y = google)) +
  geom_point(aes(x = "Waze", y = waze)) +
  geom_point(aes(x = "Follow their gut", y = gut)) +
  scale_x_discrete(limits = c("Google Maps", "Waze", "Follow their gut")) +
  ggtitle("Comparison of Commute Times by Method") +
  ylab("Commute Time (minutes)") +
  xlab("Method") +
  theme(axis.text.x = element_text(size = 10),
        axis.title.x = element_text(size = 12),
        axis.text.y = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        plot.title = element_text(size = 11))

grid.arrange(g_boxplot, g_point, ncol = 2)
```

Comparison of Commute Times by Meth... | Comparison of Commute Times by Meth...
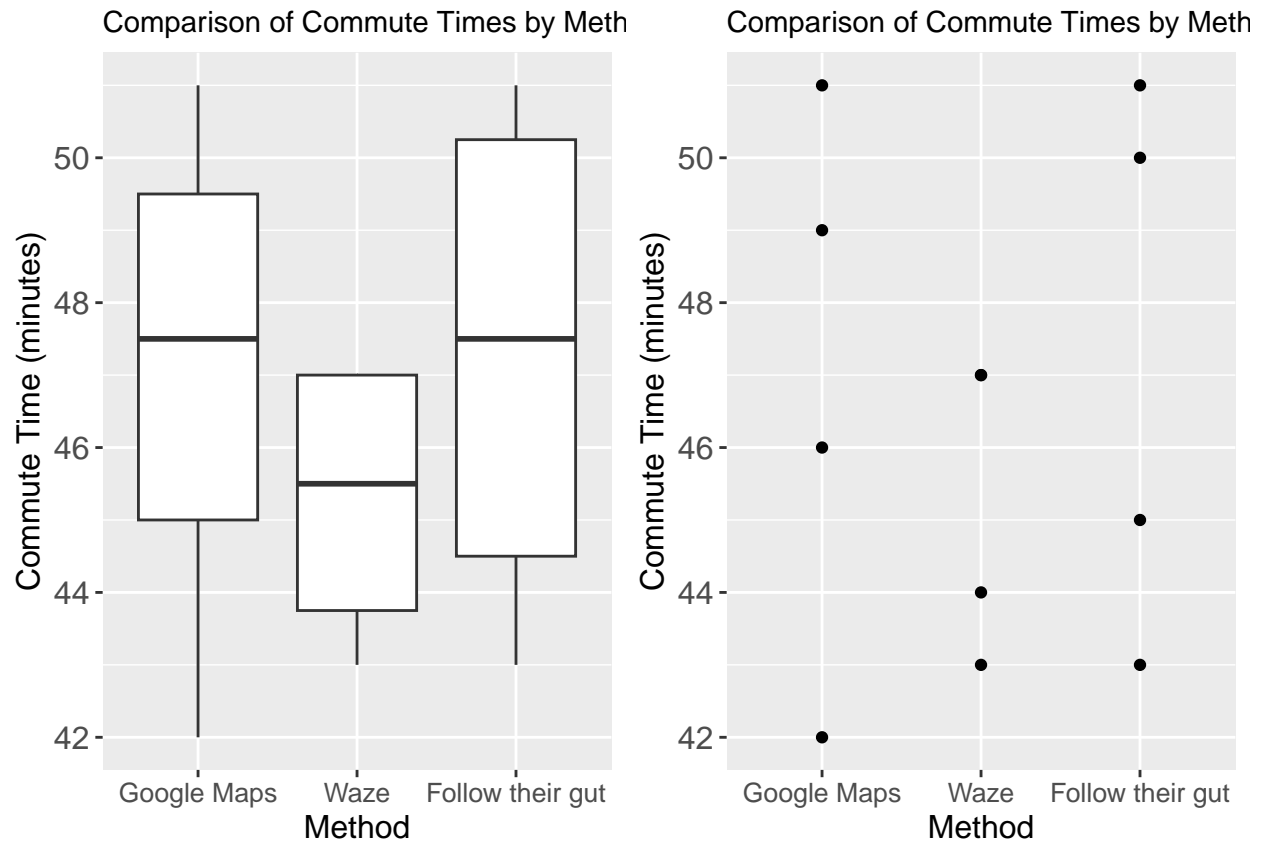
Comments on Trends ## Google maps: Seems to be quite reliable, with most commute times clustering around a ## similar range. It is a solid choice if you're looking for predictability in your daily travel ## Waze: Waze offers similar commute times to google maps but with slightly less variation. ## It might be a bit more consistent, which is great for those who like to avoid surprises on the ## road. ## Gut: This leads to more varied commute times. This method seems to be a bit of a gamble. It can ## either lead to a quick trip or take a bit longer than expected.

```r
method_df <- data.frame(
  Commute_Time_Minutes = c(google, waze, gut),
  Method = c(rep("google", length(google)),
             rep("waze", length(waze)),
             rep("gut", length(gut)))
)
method_df
```

```
##    Commute_Time_Minutes Method
## 1                    46 google
## 2                    49 google
## 3                    51 google
## 4                    42 google
## 5                    44   waze
## 6                    47   waze
## 7                    47   waze
## 8                    43   waze
## 9                    50    gut
## 10                   51    gut
```

```
## 11                    45      gut
## 12                    43      gut
```

```
ggplot(data = method_df, mapping = aes(x = Method, y = Commute_Time_Minutes)) + geom_boxplot() +
  theme(axis.title.x = element_text(size = 10), axis.text.x = element_text(size = 12),
        axis.title.y = element_text(size = 12), axis.text.y = element_text(size = 10))
```
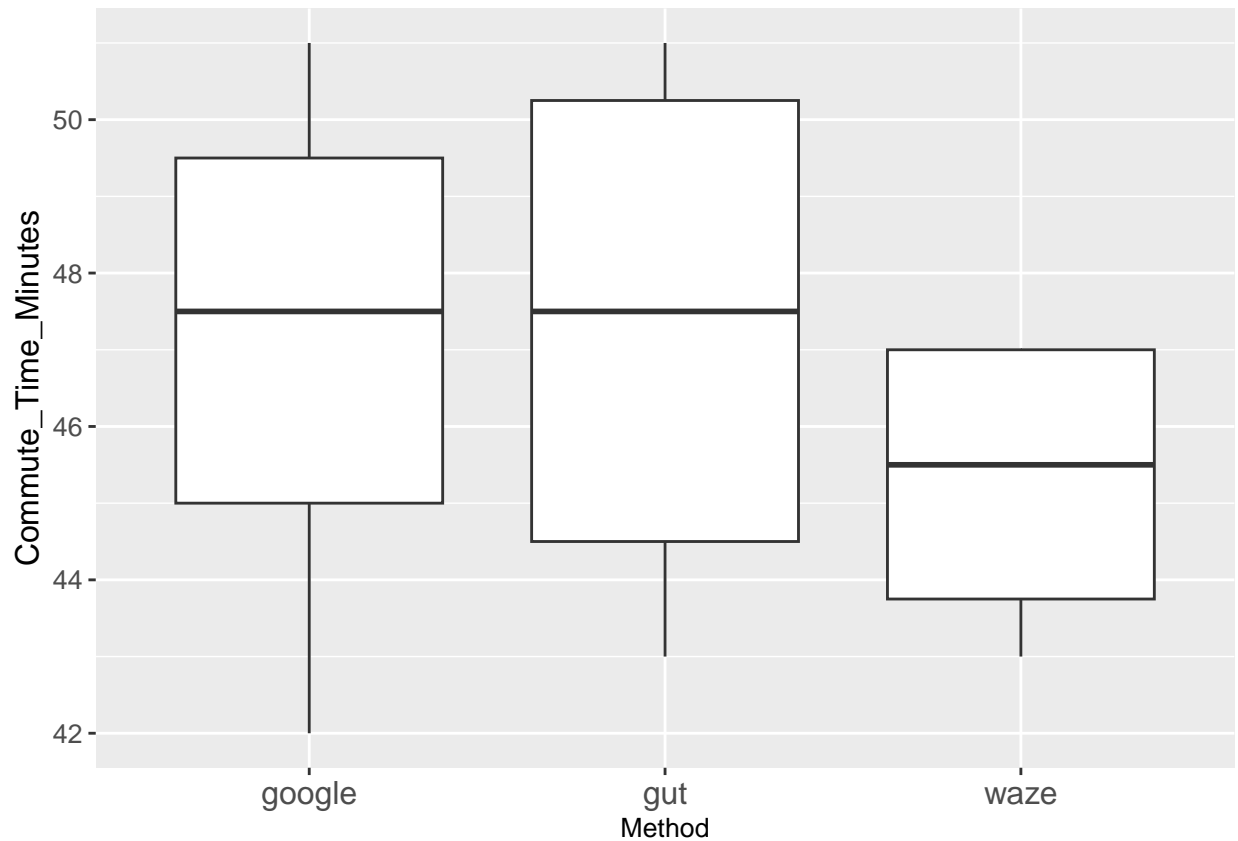


```
ggplot(data = method_df, mapping = aes(x = Method, y = Commute_Time_Minutes)) + geom_point() +
  theme(axis.title.x = element_text(size = 12), axis.text.x = element_text(size = 10),
        axis.title.y = element_text(size = 12), axis.text.y = element_text(size = 10))
```
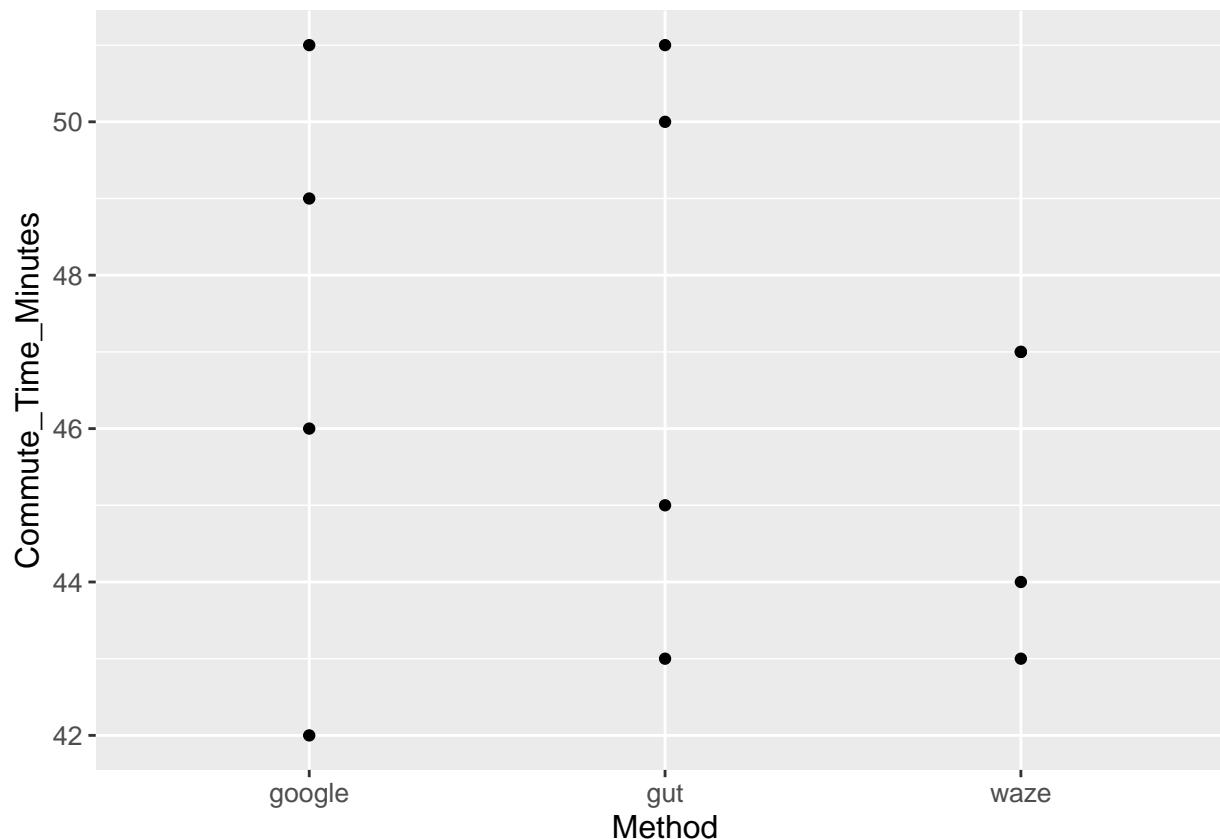
```r
method_anova <- aov(Commute_Time_Minutes ~ Method, data = method_df)
anova_summary <- summary(method_anova)
print(anova_summary)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Method       2    9.5    4.75   0.413  0.674
## Residuals    9  103.5   11.50
```

Pr(>F) the p-value is 0.674. This high p-value suggests that the observed data, or anything more ## extreme, are quite likely under the assumption that the null hypothesis H_0 is true. Thus, ## at a ## significance level of $\alpha = 0.05$, we would not reject H_0. This aligns with ## the with the visual data, showing no significant differences in commute times across methods.

##As discussed in class, what could be better about the graph on Slide 8? ## For the title add a descriptive title and axis labels would help clarify what the graph is about and what the axes represent. ## For the color and style: Using color to differentiate between the different categories can make the graph more visually appealing and easier to read. For the data points adding data points on top of the boxplots could provide more insight into the distribution of the data.

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.