

R_Activity_4

Dahee Ahn

2024-09-19

```
rm(list=ls())
```

```
library(datasets)
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4    ✓ readr      2.1.5
## ✓ forcats    1.0.0    ✓ stringr    1.5.1
## ✓ ggplot2    3.5.1    ✓ tibble     3.2.1
## ✓ lubridate  1.9.3    ✓ tidyr      1.3.1
## ✓ purrr      1.0.2
## — Conflicts — tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
```

1. Download the pig_teeth data set and load it into R. Name the dataset pigs in R. Convert the dose column to a factor using `pigs$dose <- as.factor(pigs$dose)` and provide a `summary()` of the resulting data set.

```
pigs <- read.table("C:/Users/chemk/OneDrive/Desktop/Classes/ENT6707_DataAnalysis/week5/pig_teeth.csv", header = TRUE, sep = ",")
head(pigs)
```

```
##      len dose
## 1   4.2  Low
## 2  11.5  Low
## 3   7.3  Low
## 4   5.8  Low
## 5   6.4  Low
## 6  10.0  Low
```

```
tail(pigs)
```

```
##      len dose
## 55 24.8 High
## 56 30.9 High
## 57 26.4 High
## 58 27.3 High
## 59 29.4 High
## 60 23.0 High
```

```
summary(pigs)
```

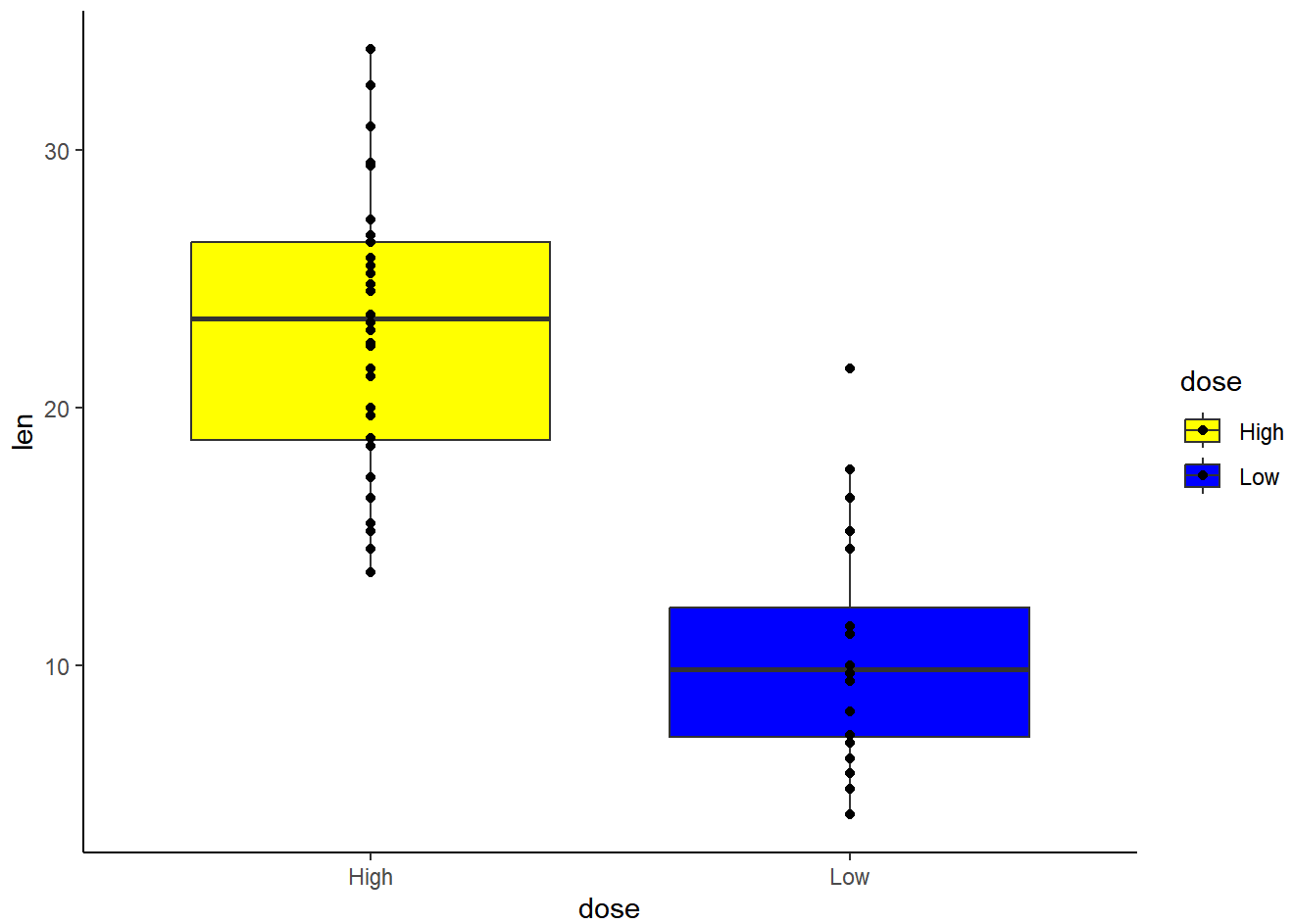
```
##      len      dose
## Min.   : 4.20   Length:60
## 1st Qu.:13.07   Class :character
## Median :19.25   Mode  :character
## Mean   :18.81
## 3rd Qu.:25.27
## Max.   :33.90
```

```
pigs$dose <- as.factor(pigs$dose)
summary(pigs$dose)
```

```
## High Low
##   40  20
```

2. We are going to compare len (length of the cells) by dose levels. First, we need to graph the data (if possible, always always always graph your data before analyzing it!). Make a boxplot with ggplot2 with Length of cells as the y-axis and Dose on the x-axis. Use the theme `theme_classic()`.

```
ggplot(pigs, mapping=aes(x=dose, y=len, fill=dose)) + geom_boxplot() + geom_point() + theme_classic() + scale_fill_manual(values =c("yellow", "blue"))
```



3. Look at the graph you just created. Do the variances look equal? If not, what might you do to fix the issue? Please write a response - we are not asking for an example using R code for this question.

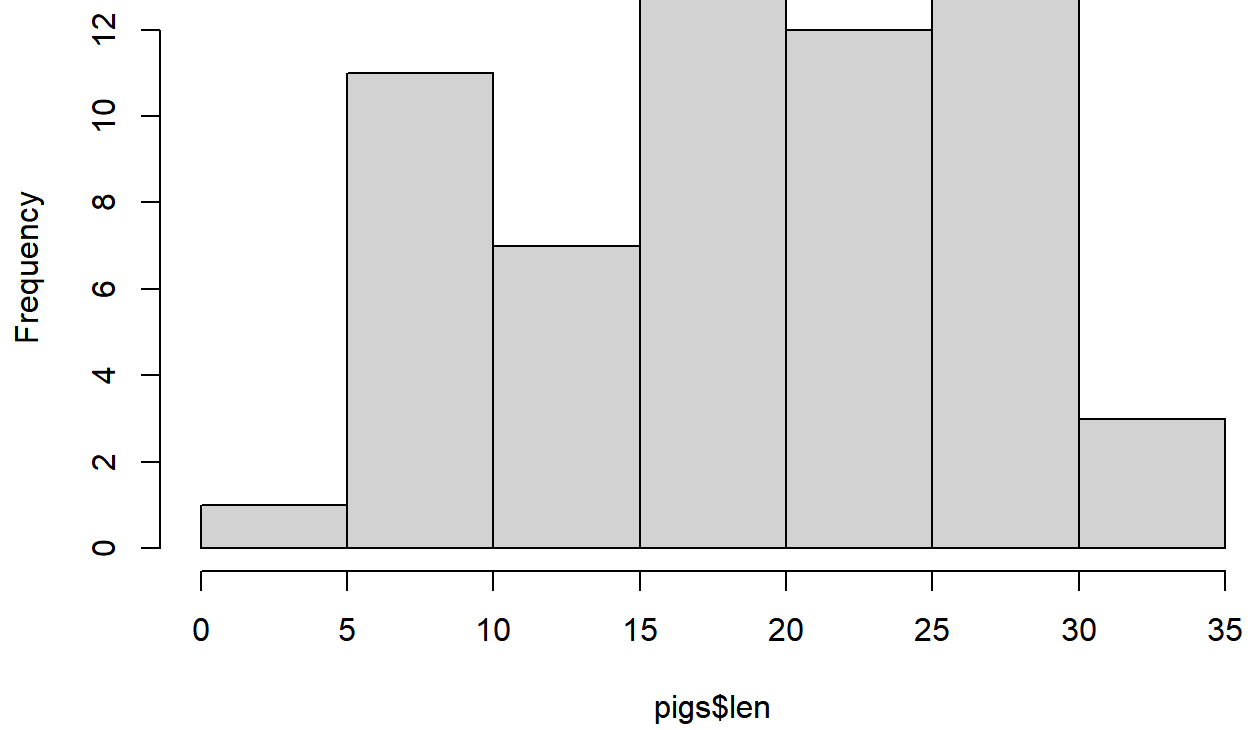
Answer: The variances look equal. If I have an issue with homogeneity of variances, square-root or log-transformations can often help.

4. That leaves one assumption left: normality. Check the normality of len using a method of your choosing. Is this assumption met? Write one sentence explaining your reasoning.

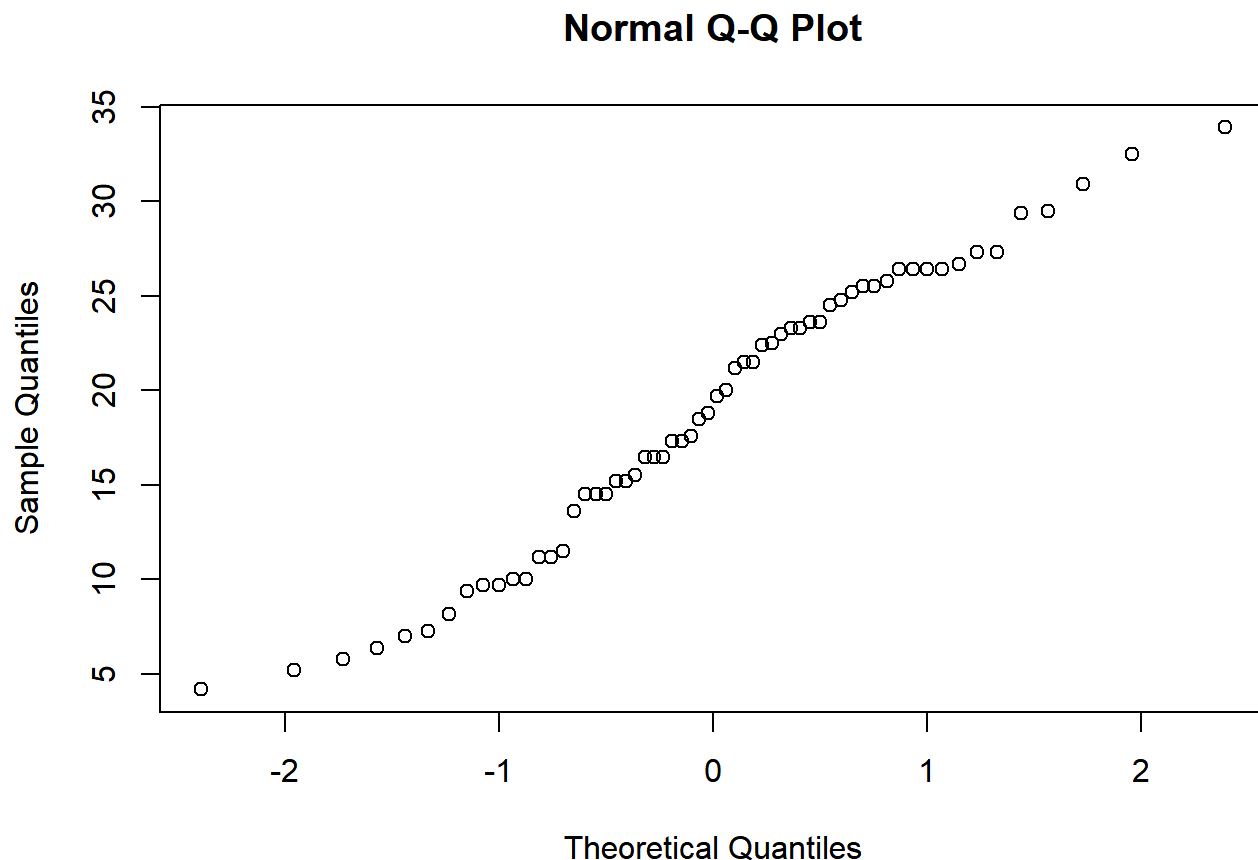
Answer: I would choose Normal Q-Q plot because the Q-Q plot shows a roughly straight line, suggesting the data is approximately distributed.

```
hist(pigs$len)
```

Histogram of pigs\$len



```
qqnorm(pigs$len)
```



5. Conduct a two-sided t-test comparing len between your two levels of dose. Set `var.equal=T` in the t-test function (we are pretending the assumptions are met here, but t-tests are actually pretty robust to violations of normality).

```
?t.test
```

```
## starting httpd help server ... done
```

```
t.test(len ~ dose, data=pigs, alternative = "two.sided", paired=F, var.equal=T)
```

```
##
## Two Sample t-test
##
## data: len by dose
## t = 9.0516, df = 58, p-value = 1.09e-12
## alternative hypothesis: true difference in means between group High and group Low is not equal to 0
## 95 percent confidence interval:
##  9.589641 15.035359
## sample estimates:
## mean in group High mean in group Low
##           22.9175           10.6050
```

6. Write a one sentence conclusion of your t-test and be sure to include means, SEs, and summary statistics.

Answer: The length (pm \pm SE) of odontoblasts (22.92 ± 0.82) of guinea pigs recieved a high (≥ 1 mg/day) dose of vitamin C were significantly longer than those (6.67 ± 1.01) of recieved a low (< 1) dose by an average of 12.31 pm ($t_{58}=9.05$, $p < 0.001$)

```
library(plotrix)
pigs %>% group_by(dose) %>% summarise(means=round(mean(len),2), SE = round(std.error(len),2))
```

```
## # A tibble: 2 × 3
##   dose means    SE
##   <fct> <dbl> <dbl>
## 1 High   22.9  0.82
## 2 Low    10.6  1.01
```