Plotting and NHST

Take-home quiz #1

Assignment

Your mission is to (1) replicate the graphs below as closely as possible and (2) provide statistical analysis/`R` output as requested in a single `Rmarkdown` file.

Make the plots *exactly* as they appear below—colors, sizes, shapes, axis ranges, etc.—as points will be taken off for substantial deviations. The only exceptions are that you are to replace all red text with something specific to your data.

Online reference guides

You will be able to complete all necessary responses from script made available so far in class. But if you get stuck, remember that in addition to simply asking Google and finding relevant questions already on sites like StackOverflow, the folks at `Rstudio` have several references available:

`Rmarkdown` cheatsheet:

https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf

`Rmarkdown` reference guide:

https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf

`ggplot` cheatsheet:

https://www.rstudio.com/wp-content/uploads/2015/12/ggplot2-cheatsheet-2.0.pdf

And of course the complete `ggplot` documentation:

http://docs.ggplot2.org/current/

Getting started

Identify and prepare data

You will need just one of the following datasets available in base `R`.

- Identify which one you will be using by presenting the data structure.
- Describe and show any necessary modifications to variable classes.

```
data("attitude")
data("ToothGrowth")
data("mtcars")
```

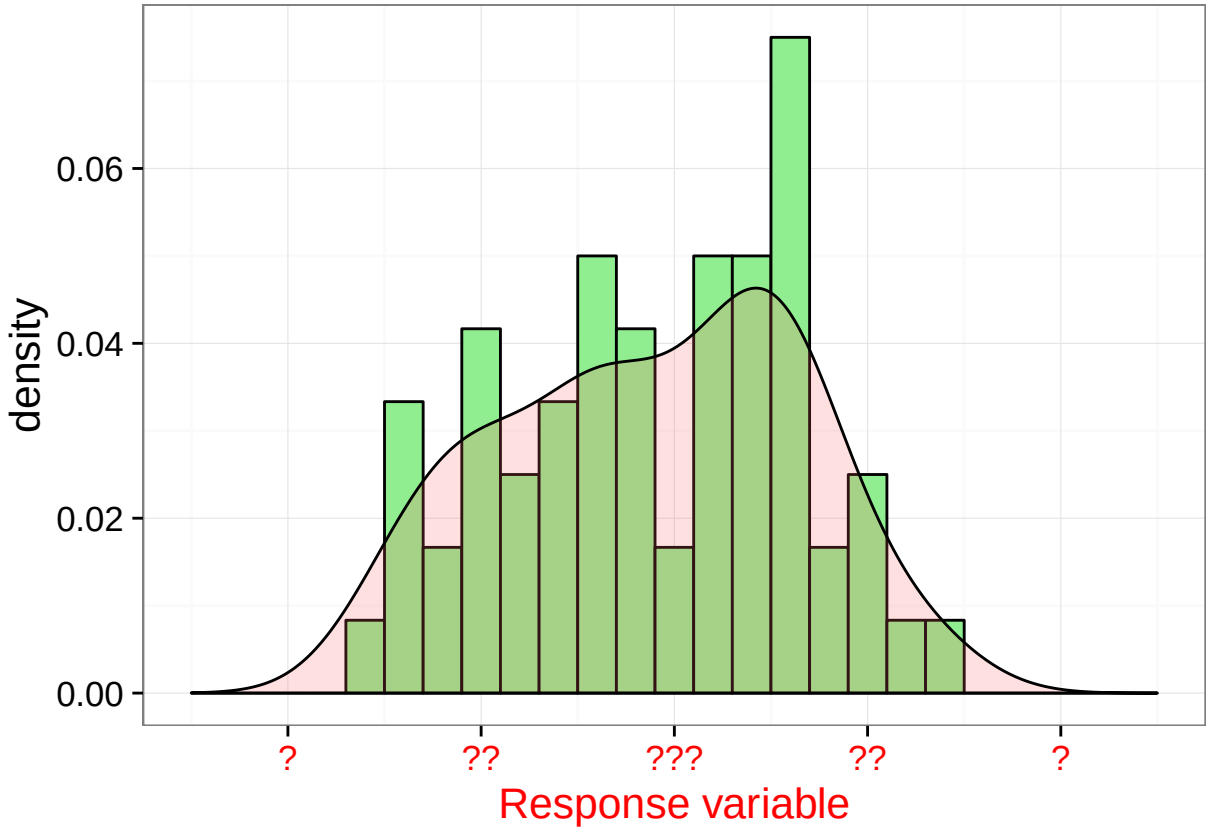
Load required packages

Load only those packages you require.

```
library(?)
.
.
.
library(??)
```

Distribution of response variable

Plot

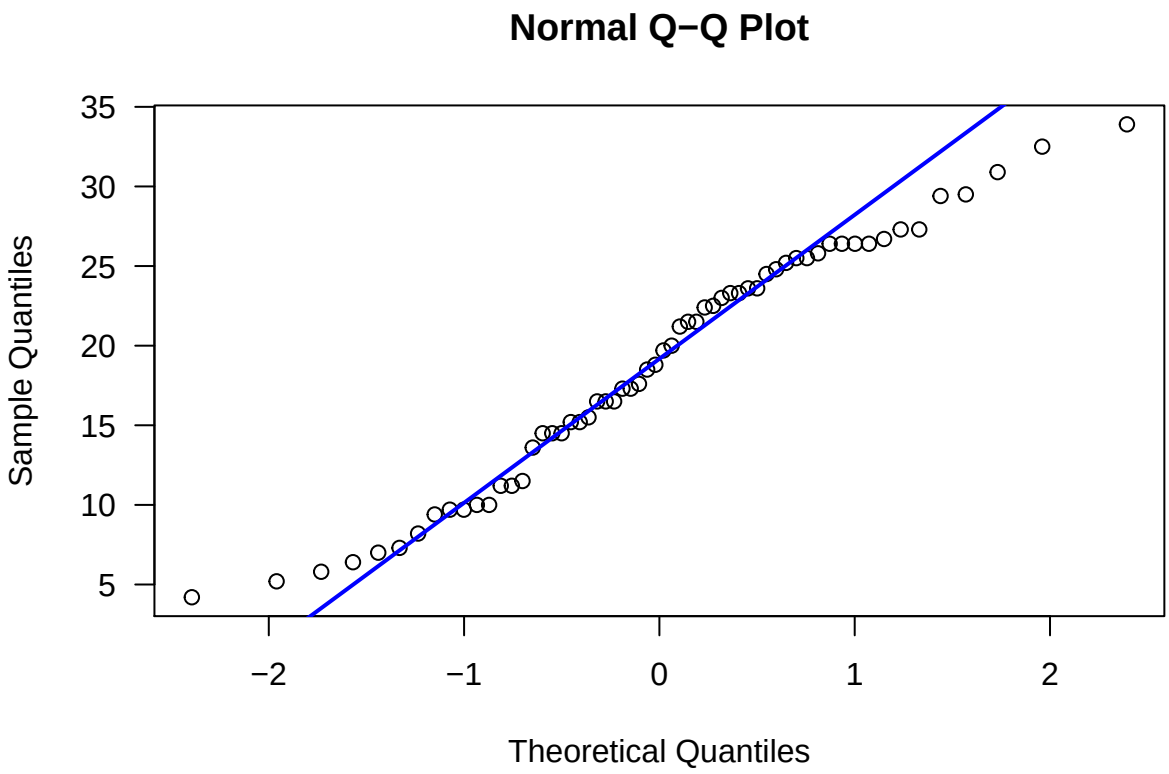


Data model

Text answers:

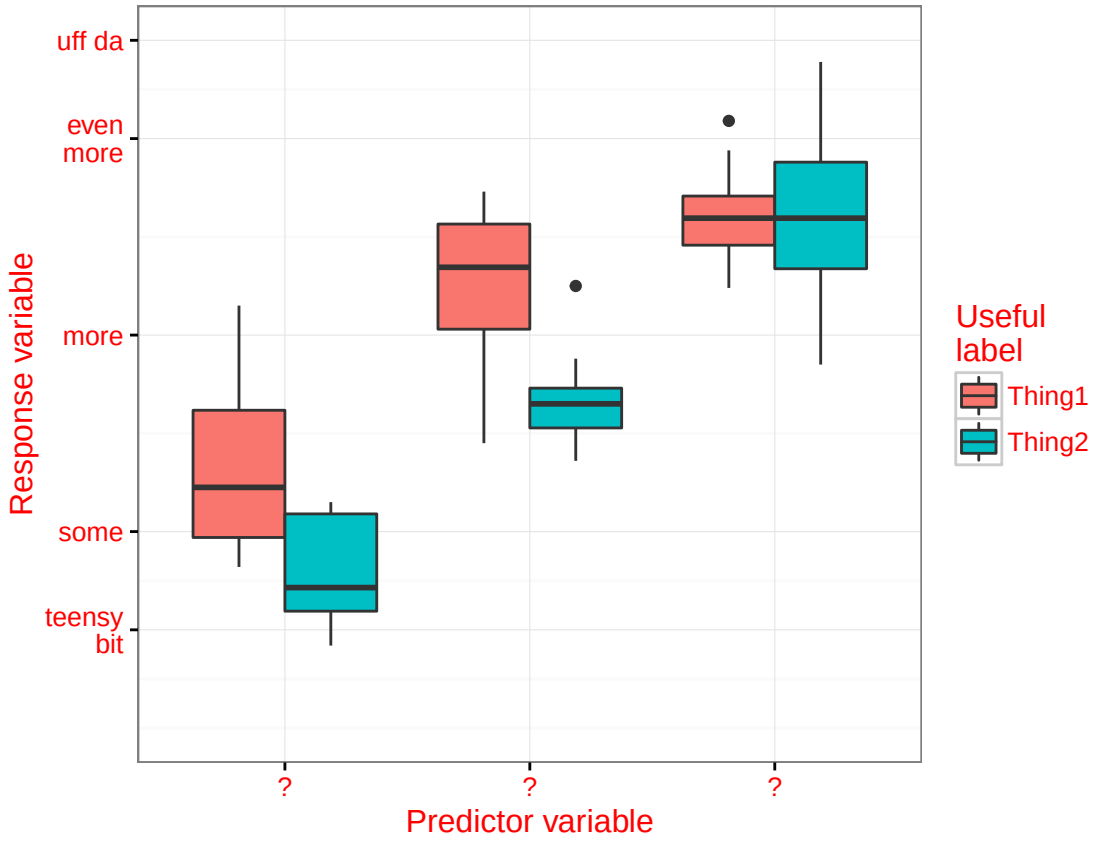
- What is a Probability Distribution Function (PDF)?
- Which PDF seems appropriate for these data?

How well do the data fit the PDF you selected? Provide a graph like this:



Data visualization

Present a boxplot with all relevant variables.



Model fitting

Which statistical model?

Text answers:

- Which general statistical approach (type of test) might you apply to these data?
- State a null and alternative hypothesis.

Run a test

Provide output for the statistical test that fits the following format:

ANOVA table	DF	SS	MS	F	P
x1					
x2					
Residuals					

```
# Provide script for statistical test here
```

Post-hoc model validation

Use a post-hoc test to ensure important model assumptions were met.

```
# Provide script for post-hoc model validation here
```

Post-hoc comparisons

If necessary, conduct a post-hoc test to compare groups within factors.

```
# Provide script for post-hoc group comparison here
```

Conclusions

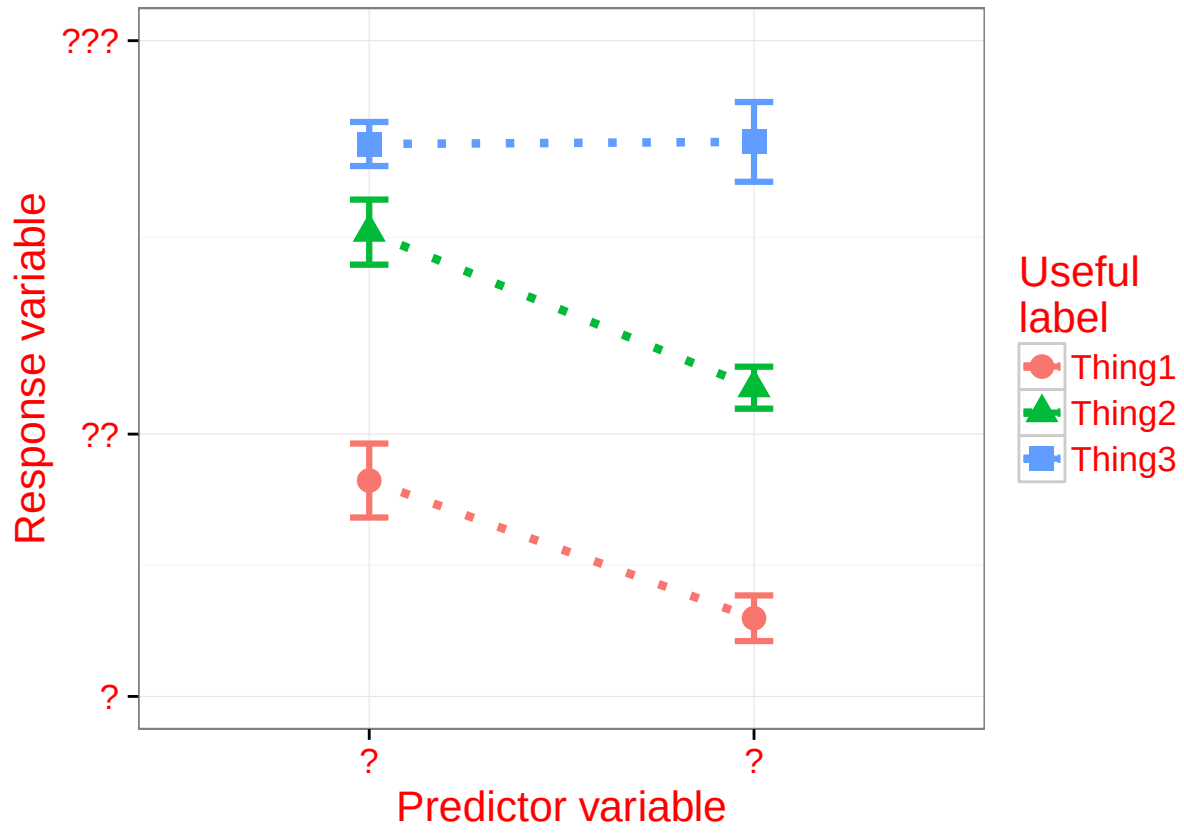
Text answer:

Briefly describe the relationship between the variables in your dataset based on the results of the statistical test(s) conducted so far.

Another statistical test?

Additional data visualiation

Present and respond to the following graph:



Text answers:

- What does the graph signify?
- What changes, if any, does the graph prompt you to make to your statistical model?

Depending on your second answer, as necessary:

- Fit and present a new statistical model
- Conduct and summarize a post-hoc test of model assumptions
- Conduct post-hoc group comparisons, if needed.

Finally, provide an updated conclusion about these data.