2.3.1: SAS - Simple Inference

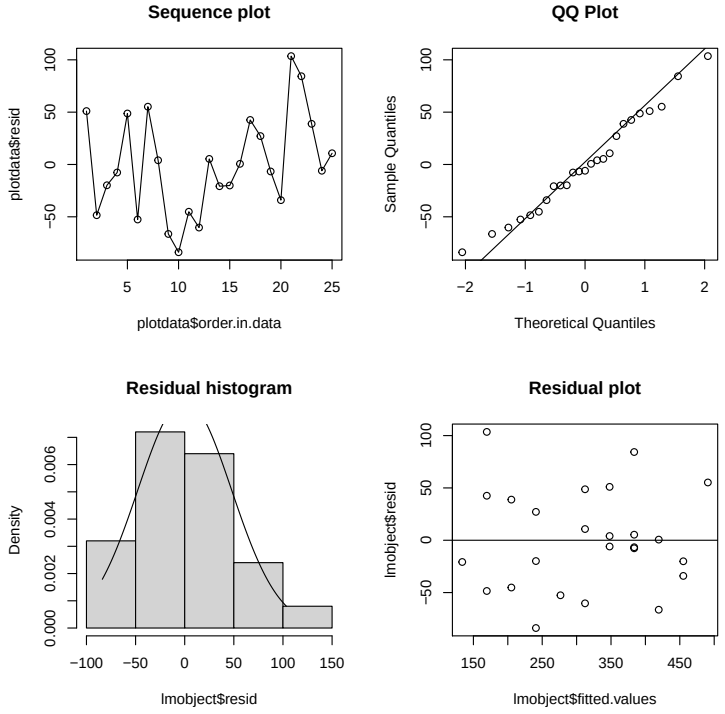
Stat 5100: Dr. Bean

Example: The Toluca Company data from Chapter 1 and Chapter 3 handouts. We really want to say something about *how* lotsize affects workhours. Does it?

```
#Input the Toluca dataset
library(stat5100)
data(toluca)

# Fit a simple linear model
toluca_lm <- lm(workhours ~ lotsize, data = toluca)

# Check model assumptions
visual_assumptions(toluca_lm)
```




```
# Numerical checks
brown_forsythe_lm(toluca_lm)

## [1] "Brown-forsythe test for constant variance in the residuals:"
## [1] "T-statistic: 1.3165, p-value: 0.201"

cor_normality_lm(toluca_lm)

## Correlation test of normality:
```

```
##                   resid expected_norm
## resid         1.0000000     0.9915055
## expected_norm 0.9915055     1.0000000
##
## Total observations: 25
## Make sure to consult with table B.6 for your final result.

coefficient_confidence_lm(toluca_lm)

##              Estimate Std. Error   t value     Pr(>|t|) lower.est  upper.est
## (Intercept) 62.365859 26.1774339  2.382428 2.585094e-02  8.213711 116.518006
## lotsize      3.570202  0.3469722 10.289592 4.448828e-10  2.852435   4.287969

# Get prediction and confidence intervals at all lot sizes less than 50
toluca_small_lot <- toluca[toluca$lotsize < 50, ]

# Prediction intervals
predict(toluca_lm, toluca_small_lot, interval = "predict")

##         fit      lwr      upr
## 2  169.4719 62.54638 276.3975
## 11 205.1739 99.94828 310.3996
## 14 133.7699 24.69771 242.8421
## 17 169.4719 62.54638 276.3975
## 21 169.4719 62.54638 276.3975
## 23 205.1739 99.94828 310.3996

# Confidence intervals
predict(toluca_lm, toluca_small_lot, interval = "confidence")

##         fit       lwr      upr
## 2  169.4719 134.36734 204.5765
## 11 205.1739 175.64938 234.6985
## 14 133.7699  92.58736 174.9524
## 17 169.4719 134.36734 204.5765
## 21 169.4719 134.36734 204.5765
## 23 205.1739 175.64938 234.6985
```