Math 3070 § 1. **Gas Mileage Example: To Pool or Not to Pool in a Two Sample t-Test.** Name: Example
Treibergs June 22, 2011

Data File Used in this Analysis:

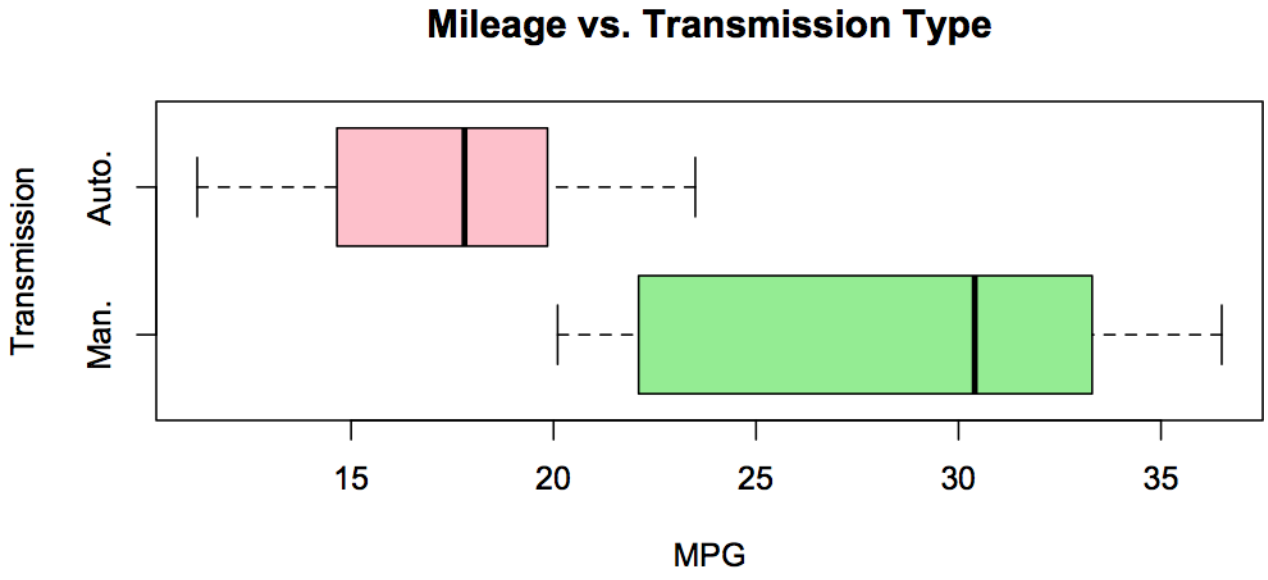
```
# Math 3074 - 1     Gas Data          June 22, 2011
# Treibergs
#
# Taken from Elliott, "Learning SAS in the computer Lab, 2nd ed.,"
# Brooks/Cole 2000. (data set "gas")
#
# From a study by Montgomery & Friedman, "Prediction Using Regression Models with
# Multicollinear Predictor Variables," IIE Transactions, May 1993.
# Test whether the mean of MPG for cars with manual transmission is significantly
# higher than the mean of MPG for cars with automatic transmission. Is a pooled test
# appropriate?
#
# Variables
#
# Disp            Displacement in cubic inches
# HP              Horsepower in ft-lbs
# Torque          in ft-lbs
# ComprRatio
# RearAxleRatio
# NBar            Number of carburetor barresl
# NSpeeds             Number of transmission speeds
# Len             Car Length in inches
# Width           Car width in inches
# Wt              Car weight in pounds
# Tran            Transmission type       0=manual, 1=automatic
# MPG             Gas mileage in miles per gallon
#
Disp HP  Torque ComprRatio RearAxleRatio NBar NSpeeds  Len Width Wt Tran MPG
318.0 140 255 8.5 2.7 2 3 215.3 76.3 4370 1 19.7
440.0 215 330 8.2 2.9 4 3 184.5 69.0 4215 1 11.2
351.0 143 255 8.0 3.0 2 3 199.9 74.0 3890 1 18.3
360.0 180 290 8.4 2.5 2 3 214.2 76.3 4250 1 21.5
140.0  83 109 8.4 3.4 2 4 168.8 69.4 2700 0 20.3
 85.3  80  83 8.5 3.9 2 4 160.6 62.2 2009 0 36.5
350.0 165 260 8.0 2.6 4 3 200.3 69.9 3910 1 18.9
 96.9  75  83 9.0 4.3 2 5 162.5 65.0 2320 0 30.4
351.0 148 243 8.0 3.3 2 3 215.5 78.5 4540 1 13.9
440.0 215 330 8.2 2.7 4 3 231.0 79.7 5185 1 14.9
171.0 109 146 8.2 3.2 2 4 170.4 66.9 2655 1 21.5
302.0 129 220 8.0 3.0 2 3 199.9 74.0 3890 1 17.8
350.0 155 250 8.5 3.1 4 3 196.7 72.2 3910 1 17.8
318.0 145 255 8.5 2.5 2 3 197.6 71.0 3666 1 16.4
231.0 110 175 8.0 2.6 2 3 179.3 65.4 3050 1 23.5
 96.9  75  83 9.0 4.3 2 5 165.2 61.8 2275 0 31.9
500.0 190 360 8.5 2.7 4 3 224.1 79.8 5290 1 14.4
231.0 110 175 8.0 2.6 2 3 179.3 65.4 3020 1 22.1
350.0 170 275 8.5 2.6 4 3 199.6 72.9 3860 1 17.0
```

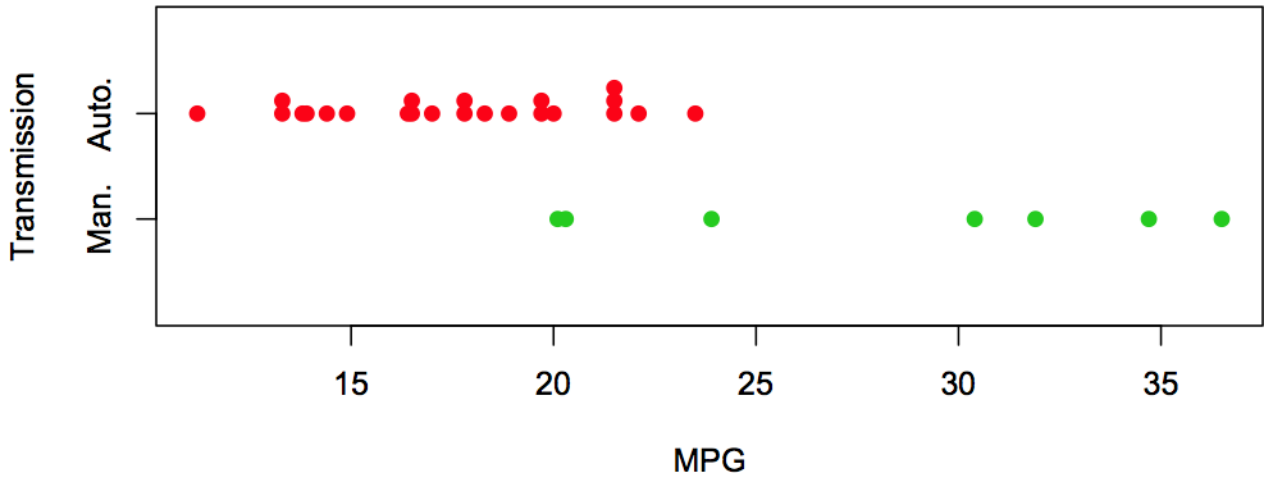
```
250.0 105 185 8.3 2.7 1 3 196.7 72.2 3510 1 20.0
225.0  95 170 8.4 2.8 1 3 194.0 71.8 3365 0 20.1
 89.7  70  81 8.2 3.9 2 4 155.7 64.0 1905 0 34.7
350.0 155 250 8.5 3.1 4 3 195.4 74.4 3885 1 16.5
258.0 110 195 8.0 3.1 1 3 171.5 77.0 3375 1 19.7
460.0 223 366 8.0 3.0 4 3 228.0 79.8 5430 1 13.3
360.0 195 295 8.3 3.2 4 3 209.3 77.4 4215 1 13.8
262.0 110 200 8.5 2.6 2 3 179.3 65.4 3180 1 21.5
350.0 165 255 8.5 2.7 4 3 185.2 69.0 3660 1 16.5
351.0 148 243 8.0 3.3 2 3 216.1 78.5 4715 1 13.3
133.6  96 120 8.4 3.9 2 5 171.5 63.4 2535 0 23.9
```

R Session:

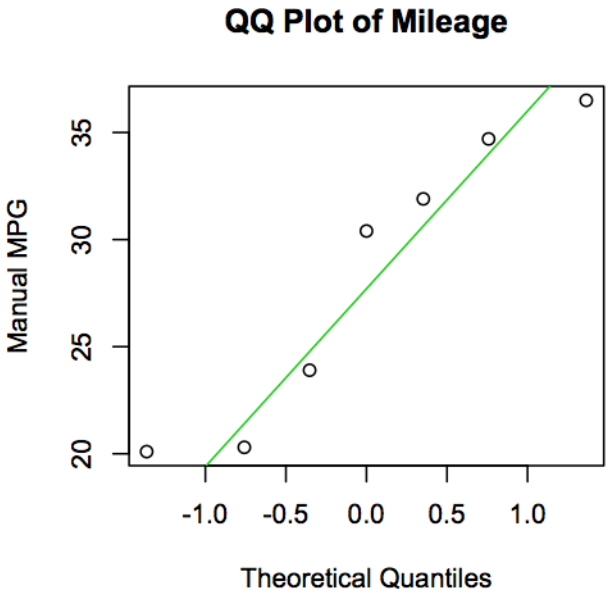
```
R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.34 (5589) i386-apple-darwin9.8.0]

[Workspace restored from /home/1004/ma/treibergs/.RData]


> tt <-read.table("M3074GasData.txt",header=TRUE)
```

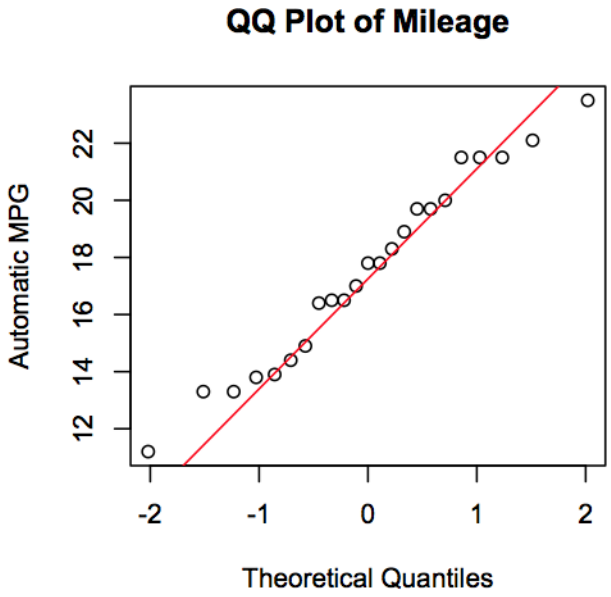
```
> tt
    Disp  HP Torque ComprRatio RearAxleRatio NBar NSpeeds   Len Width   Wt Tran  MPG
1  318.0 140    255        8.5           2.7    2       3 215.3  76.3 4370    1 19.7
2  440.0 215    330        8.2           2.9    4       3 184.5  69.0 4215    1 11.2
3  351.0 143    255        8.0           3.0    2       3 199.9  74.0 3890    1 18.3
4  360.0 180    290        8.4           2.5    2       3 214.2  76.3 4250    1 21.5
5  140.0  83    109        8.4           3.4    2       4 168.8  69.4 2700    0 20.3
6   85.3  80     83        8.5           3.9    2       4 160.6  62.2 2009    0 36.5
7  350.0 165    260        8.0           2.6    4       3 200.3  69.9 3910    1 18.9
8   96.9  75     83        9.0           4.3    2       5 162.5  65.0 2320    0 30.4
9  351.0 148    243        8.0           3.3    2       3 215.5  78.5 4540    1 13.9
10 440.0 215    330        8.2           2.7    4       3 231.0  79.7 5185    1 14.9
11 171.0 109    146        8.2           3.2    2       4 170.4  66.9 2655    1 21.5
12 302.0 129    220        8.0           3.0    2       3 199.9  74.0 3890    1 17.8
13 350.0 155    250        8.5           3.1    4       3 196.7  72.2 3910    1 17.8
14 318.0 145    255        8.5           2.5    2       3 197.6  71.0 3666    1 16.4
15 231.0 110    175        8.0           2.6    2       3 179.3  65.4 3050    1 23.5
16  96.9  75     83        9.0           4.3    2       5 165.2  61.8 2275    0 31.9
17 500.0 190    360        8.5           2.7    4       3 224.1  79.8 5290    1 14.4
18 231.0 110    175        8.0           2.6    2       3 179.3  65.4 3020    1 22.1
19 350.0 170    275        8.5           2.6    4       3 199.6  72.9 3860    1 17.0
20 250.0 105    185        8.3           2.7    1       3 196.7  72.2 3510    1 20.0
21 225.0  95    170        8.4           2.8    1       3 194.0  71.8 3365    0 20.1
22  89.7  70     81        8.2           3.9    2       4 155.7  64.0 1905    0 34.7
23 350.0 155    250        8.5           3.1    4       3 195.4  74.4 3885    1 16.5
24 258.0 110    195        8.0           3.1    1       3 171.5  77.0 3375    1 19.7
25 460.0 223    366        8.0           3.0    4       3 228.0  79.8 5430    1 13.3
26 360.0 195    295        8.3           3.2    4       3 209.3  77.4 4215    1 13.8
27 262.0 110    200        8.5           2.6    2       3 179.3  65.4 3180    1 21.5
28 350.0 165    255        8.5           2.7    4       3 185.2  69.0 3660    1 16.5
29 351.0 148    243        8.0           3.3    2       3 216.1  78.5 4715    1 13.3
30 133.6  96    120        8.4           3.9    2       5 171.5  63.4 2535    0 23.9

> ###################### PICK OFF MPG VS TRAN ########################################
> attach(tt)
> ManualMPG <- MPG[Tran == 0]
> AutomaticMPG <- MPG[Tran == 1]
> Tran <- factor(Tran)
>
> layout(matrix(1:2,ncol=1))
>
> plot(MPG ~ Tran, horizontal = TRUE, xlab = "Transmission",
+ col=c("lightgreen","pink"), main = "Mileage vs. Transmission Type")
> stripchart(MPG ~ Tran, col = c(3,2), method="stack",
+ ylim = c(0.1,2.9), pch = 19, ylab = "Transmission")
> # M3074GasEg.pdf
>
> # The means look different. I'm not sure about if I can assume
> # that the variances are equal.
```


Mileage vs. Transmission Type
Transmission
Auto.
Man.
15
20
25
30
35
MPG


Transmission
Auto.
Man.
15
20
25
30
35
MPG

```
> ################## RUN TEST FOR EQUALITY OF VARIANCES ###################
>
> mean(ManualMPG)
[1] 28.25714
> mean(AutomaticMPG)
[1] 17.54348
> var(ManualMPG)
[1] 46.02619
> var(AutomaticMPG)
[1] 11.14893
> var.test(manualMPG,AutomaticMPG)
Error in var.test(manualMPG, AutomaticMPG) : object 'manualMPG' not found
> var.test(ManualMPG,AutomaticMPG)

F test to compare two variances

data:  ManualMPG and AutomaticMPG
F = 4.1283, num df = 6, denom df = 22, p-value = 0.01264
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
  1.351487 21.221801
sample estimates:
ratio of variances
          4.128305

> # Variances are significantly different. So pooled t-test not appropriate.
>
>
> ###################### QQ PLOTS TO CHECK NORMALITY ASSUMPTION ###############
> layout(matrix(c(1,3,2,4),ncol=2))

> qqnorm(ManualMPG,main="QQ Plot of Mileage",ylab="Manual MPG")
> qqline(ManualMPG,col=3)
> qqnorm(AutomaticMPG,main="QQ Plot of Mileage",ylab="Automatic MPG")
> qqline(AutomaticMPG,col=2)
> # M3074Gas2.pdf
>
> # Small data sets but neither qq plot showsw strong reason to doubt normality.
```






```
> ###################### RUN 2-SAMPLE T-TEST ##############################
> t.test(ManualMPG,AutomaticMPG,alternative = "greater")

Welch Two Sample t-test

data:  ManualMPG and AutomaticMPG
t = 4.0322, df = 6.907, p-value = 0.002561
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 5.669466      Inf
sample estimates:
mean of x mean of y
 28.25714  17.54348

> # Data significantly indicates that MPG for manual transmission is
> # greater than MPG for automatic.
>
> ######## WELCH-SATTERTHWAITE TEST BY HAND ############################
>
> ManBar <- mean(ManualMPG)
> ManV <- var(ManualMPG)
> ManN <- length(ManualMPG)
> AutBar <- mean(AutomaticMPG)
> AutV <- var(AutomaticMPG)
> AutN <- length(AutomaticMPG)
> nu <- (ManV/ManN + AutV/AutN)^2/((ManV/ManN)^2/(ManN-1) + (AutV/AutN)^2/(AutN-1))
> nu
[1] 6.907039
> T <- (ManBar-AutBar)/sqrt(ManV/ManN + AutV/AutN)
> T
[1] 4.032168
> pval <- pt(T,nu,lower.tail=FALSE); pval
[1] 0.002561244
```

```
> alpha<- .05
> se <- sqrt(ManV/ManN + AutV/AutN)
> talpha <- qt(alpha,nu,lower.tail=FALSE)
> LowCI <- ManBar-AutBar-talpha*se
> LowCI
[1] 5.669466
>
> cat("\n\n   Welch-Satterthwaite Two Sample t-test BY HAND\n\n",
+ "data: ManualMPG and AutomaticMPG\n t =", T, ", df =", nu, ", p-value =",
+ pval ,"\n alternative hypothesis: true difference in means > 0\n",
+ "95 percent confidence interval:\n", c(LowCI,Inf),
+ "\n sample estimates:\n mean of manualMPG,  mean of AutomaticMPG\n",
+ c(ManBar,AutBar),"\n\n")


   Welch-Satterthwaite Two Sample t-test BY HAND

 data: ManualMPG and AutomaticMPG
 t = 4.032168 , df = 6.907039 , p-value = 0.002561244
 alternative hypothesis: true difference in means > 0
 95 percent confidence interval:
 5.669466 Inf
 sample estimates:
 mean of manualMPG,  mean of AutomaticMPG
 28.25714 17.54348

>
```