Report on Fuel Consumption of automobiles from 1973-74

Bartek Skorulski

May 13th, 2015

Introduction

We investigate data that extracted from the 1974 *Motor Trend* US magazine. They comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models). We are particularly interested in the following questions:

- Is an automatic or manual transmission better for MPG?
- Quantify the MPG difference between automatic and manual transmissions.

```
ggplot(cs.dt, aes(am, mpg)) + geom_boxplot() +
    theme_tufte() + scale_x_discrete(labels = c("manual",
    "automatic"))
```

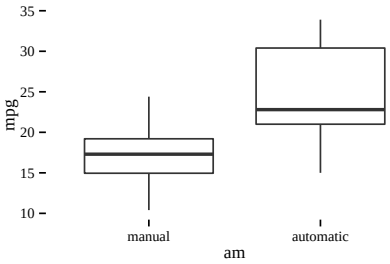


Figure 1: Automatic vs. manual transmission.

```
ggplot(cs.dt, aes(wt, mpg, colour = am)) + geom_point() +
    theme_tufte() + geom_text(aes(label = cyl,
    colour = NULL), vjust = -0.6) + geom_smooth(method = "lm")
```

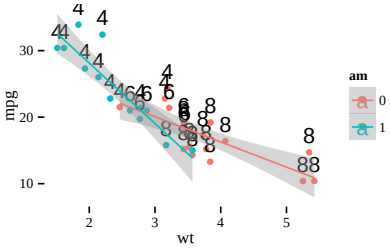


Figure 2: Impact of weight, transmission and number of gears on fuel consumption.

The boxplot on Figure 1 suggests that manual transmission is better for mpg. However, Figure 2 shows that actually weight or number of gears can be the most important factor.

Subset Selection

Let us try to identify the subset of the predictors that can be related to the mpg response. For that we use =regsubsets= function from leaps library.

```
nvmax <- 16
regfit <- regsubsets(mpg ~ ., cs.dt, nvmax = nvmax)

reg.summary <- summary(regfit)
reg.summary
```

```
## Subset selection object
## Call: regsubsets.formula(mpg ~ ., cs.dt, nvmax = nvmax)
## 16 Variables  (and intercept)
##       Forced in Forced out
## cyl6      FALSE      FALSE
```

```
## cyl8      FALSE      FALSE
## disp      FALSE      FALSE
## hp        FALSE      FALSE
## drat      FALSE      FALSE
## wt        FALSE      FALSE
## qsec      FALSE      FALSE
## vs1       FALSE      FALSE
## am1       FALSE      FALSE
## gear4     FALSE      FALSE
## gear5     FALSE      FALSE
## carb2     FALSE      FALSE
## carb3     FALSE      FALSE
## carb4     FALSE      FALSE
## carb6     FALSE      FALSE
## carb8     FALSE      FALSE
## 1 subsets of each size up to 16
## Selection Algorithm: exhaustive
##           cyl6 cyl8 disp hp  drat wt  qsec
## 1  ( 1 )  " "  " "  " "  " " " " "*" " "
## 2  ( 1 )  " "  " "  " "  "*" " "  "*" " "
## 3  ( 1 )  " "  " "  " "  " " " " "*" "*"
## 4  ( 1 )  "*"  " "  " "  "*" " "  "*" " "
## 5  ( 1 )  "*"  " "  " "  "*" " "  "*" " "
## 6  ( 1 )  "*"  " "  " "  "*" " "  "*" " "
## 7  ( 1 )  " "  "*"  " "  "*" " "  "*" "*"
## 8  ( 1 )  "*"  " "  "*"  "*" " "  "*" " "
## 9  ( 1 )  "*"  " "  "*"  "*" " "  "*" " "
## 10  ( 1 ) "*"  " "  "*"  "*" "*"  "*" " "
## 11  ( 1 ) "*"  " "  "*"  "*" " "  "*" " "
## 12  ( 1 ) "*"  " "  "*"  "*" "*"  "*" " "
## 13  ( 1 ) "*"  " "  "*"  "*" "*"  "*" "*"
## 14  ( 1 ) "*"  " "  "*"  "*" "*"  "*" "*"
## 15  ( 1 ) "*"  " "  "*"  "*" "*"  "*" "*"
## 16  ( 1 ) "*"  "*"  "*"  "*" "*"  "*" "*"
##           vs1 am1 gear4 gear5 carb2 carb3
## 1  ( 1 )  " " " " " "   " "   " "   " "
## 2  ( 1 )  " " " " " "   " "   " "   " "
## 3  ( 1 )  " " "*" " "   " "   " "   " "
## 4  ( 1 )  " " "*" " "   " "   " "   " "
## 5  ( 1 )  "*" "*" " "   " "   " "   " "
## 6  ( 1 )  "*" "*" " "   "*"   " "   " "
## 7  ( 1 )  "*" "*" " "   "*"   " "   " "
## 8  ( 1 )  " " " " "*"   "*"   "*"   " "
## 9  ( 1 )  "*" " " "*"   "*"   "*"   " "
```

```
## 10  ( 1 ) "*" " " "*"   "*"   "*"   " "
## 11  ( 1 ) "*" " " "*"   "*"   "*"   "*"
## 12  ( 1 ) "*" " " "*"   "*"   "*"   "*"
## 13  ( 1 ) "*" "*" " "   "*"   " "   "*"
## 14  ( 1 ) "*" "*" "*"   "*"   "*"   "*"
## 15  ( 1 ) "*" "*" "*"   "*"   "*"   "*"
## 16  ( 1 ) "*" "*" "*"   "*"   "*"   "*"
##           carb4 carb6 carb8
## 1  ( 1 )  " "   " "   " "
## 2  ( 1 )  " "   " "   " "
## 3  ( 1 )  " "   " "   " "
## 4  ( 1 )  " "   " "   " "
## 5  ( 1 )  " "   " "   " "
## 6  ( 1 )  " "   " "   " "
## 7  ( 1 )  " "   " "   " "
## 8  ( 1 )  " "   " "   "*"
## 9  ( 1 )  "*"   " "   " "
## 10  ( 1 ) "*"   " "   " "
## 11  ( 1 ) " "   "*"   "*"
## 12  ( 1 ) " "   "*"   "*"
## 13  ( 1 ) "*"   "*"   "*"
## 14  ( 1 ) " "   "*"   "*"
## 15  ( 1 ) "*"   "*"   "*"
## 16  ( 1 ) "*"   "*"   "*"
```

```
names(summary(regfit))
```

```
## [1] "which"  "rsq"    "rss"    "adjr2"
## [5] "cp"     "bic"    "outmat" "obj"
```

```
reg.summary$rsq
```

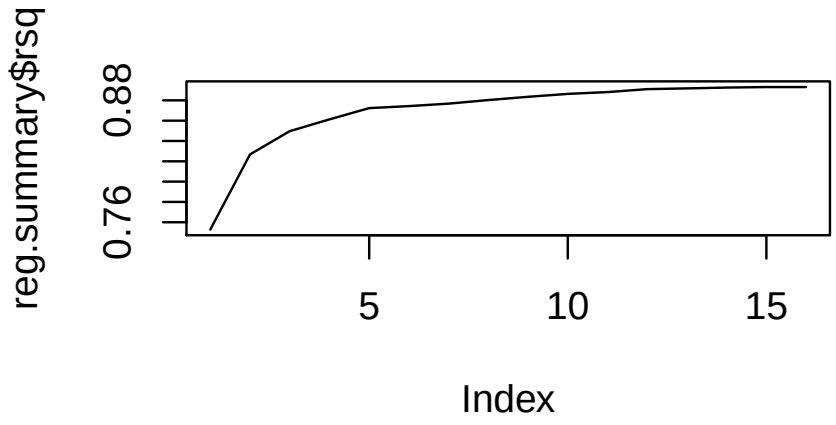
```
##  [1] 0.7528328 0.8267855 0.8496636 0.8612516
##  [5] 0.8723598 0.8743388 0.8767613 0.8802830
##  [9] 0.8834716 0.8863519 0.8881402 0.8910400
## [13] 0.8917557 0.8925800 0.8930592 0.8930749
```

```
reg.summary$which
```

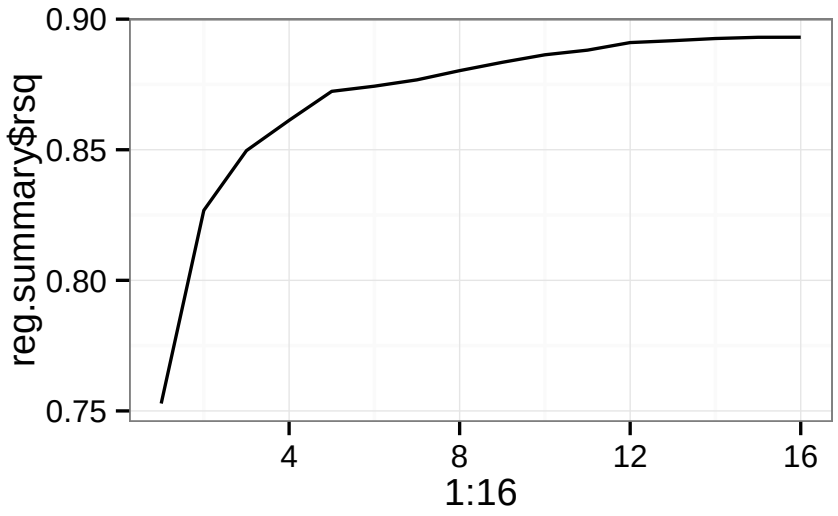
```
##    (Intercept)  cyl6  cyl8  disp    hp  drat
## 1         TRUE FALSE FALSE FALSE FALSE FALSE
## 2         TRUE FALSE FALSE FALSE  TRUE FALSE
## 3         TRUE FALSE FALSE FALSE FALSE FALSE
## 4         TRUE  TRUE FALSE FALSE  TRUE FALSE
## 5         TRUE  TRUE FALSE FALSE  TRUE FALSE
## 6         TRUE  TRUE FALSE FALSE  TRUE FALSE
```

```
## 7         TRUE FALSE  TRUE FALSE  TRUE FALSE
## 8         TRUE  TRUE FALSE  TRUE  TRUE FALSE
## 9         TRUE  TRUE FALSE  TRUE  TRUE FALSE
## 10        TRUE  TRUE FALSE  TRUE  TRUE  TRUE
## 11        TRUE  TRUE FALSE  TRUE  TRUE FALSE
## 12        TRUE  TRUE FALSE  TRUE  TRUE  TRUE
## 13        TRUE  TRUE FALSE  TRUE  TRUE  TRUE
## 14        TRUE  TRUE FALSE  TRUE  TRUE  TRUE
## 15        TRUE  TRUE FALSE  TRUE  TRUE  TRUE
## 16        TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
##      wt  qsec   vs1   am1 gear4 gear5 carb2
## 1  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
## 2  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
## 3  TRUE  TRUE FALSE  TRUE FALSE FALSE FALSE
## 4  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE
## 5  TRUE FALSE  TRUE  TRUE FALSE FALSE FALSE
## 6  TRUE FALSE  TRUE  TRUE FALSE  TRUE FALSE
## 7  TRUE  TRUE  TRUE  TRUE FALSE  TRUE FALSE
## 8  TRUE FALSE FALSE FALSE  TRUE  TRUE  TRUE
## 9  TRUE FALSE  TRUE FALSE  TRUE  TRUE  TRUE
## 10 TRUE FALSE  TRUE FALSE  TRUE  TRUE  TRUE
## 11 TRUE FALSE  TRUE FALSE  TRUE  TRUE  TRUE
## 12 TRUE FALSE  TRUE FALSE  TRUE  TRUE  TRUE
## 13 TRUE  TRUE  TRUE  TRUE FALSE  TRUE FALSE
## 14 TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## 15 TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## 16 TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
##    carb3 carb4 carb6 carb8
## 1  FALSE FALSE FALSE FALSE
## 2  FALSE FALSE FALSE FALSE
## 3  FALSE FALSE FALSE FALSE
## 4  FALSE FALSE FALSE FALSE
## 5  FALSE FALSE FALSE FALSE
## 6  FALSE FALSE FALSE FALSE
## 7  FALSE FALSE FALSE FALSE
## 8  FALSE FALSE FALSE  TRUE
## 9  FALSE  TRUE FALSE FALSE
## 10 FALSE  TRUE FALSE FALSE
## 11  TRUE FALSE  TRUE  TRUE
## 12  TRUE FALSE  TRUE  TRUE
## 13  TRUE  TRUE  TRUE  TRUE
## 14  TRUE FALSE  TRUE  TRUE
## 15  TRUE  TRUE  TRUE  TRUE
## 16  TRUE  TRUE  TRUE  TRUE
```

```
plot(reg.summary$rsq, type = "l")
```




```
ggplot(data = NULL, aes(x = 1:16, y = reg.summary$rsq)) +
    geom_line() + theme_bw()
```




```
## plot(regfit.full ,scale ='r2')
## plot(regfit.full , scale ='adjr2') ##
## plot(regfit.full, scale ='Cp')

fit.full <- lm(mpg ~ ., data = cs.dt)
summary(fit.full)

##
## Call:
## lm(formula = mpg ~ ., data = cs.dt)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5087 -1.3584 -0.0948  0.7745  4.6251
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept) 23.87913   20.06582   1.190
## cyl6        -2.64870    3.04089  -0.871
## cyl8        -0.33616    7.15954  -0.047
## disp         0.03555    0.03190   1.114
## hp          -0.07051    0.03943  -1.788
## drat         1.18283    2.48348   0.476
## wt          -4.52978    2.53875  -1.784
## qsec         0.36784    0.93540   0.393
## vs1          1.93085    2.87126   0.672
## am1          1.21212    3.21355   0.377
## gear4        1.11435    3.79952   0.293
## gear5        2.52840    3.73636   0.677
## carb2       -0.97935    2.31797  -0.423
## carb3        2.99964    4.29355   0.699
## carb4        1.09142    4.44962   0.245
## carb6        4.47757    6.38406   0.701
## carb8        7.25041    8.36057   0.867
##             Pr(>|t|)
## (Intercept)   0.2525
## cyl6          0.3975
## cyl8          0.9632
## disp          0.2827
## hp            0.0939 .
## drat          0.6407
## wt            0.0946 .
## qsec          0.6997
## vs1           0.5115
## am1           0.7113
## gear4         0.7733
## gear5         0.5089
## carb2         0.6787
## carb3         0.4955
## carb4         0.8096
## carb6         0.4938
## carb8         0.3995
## ---
## Signif. codes:
```

```
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.833 on 15 degrees of freedom
## Multiple R-squared:  0.8931, Adjusted R-squared:  0.779
## F-statistic:  7.83 on 16 and 15 DF,  p-value: 0.000124
```

```
fit.wt <- lm(mpg ~ wt * am, data = cs.dt)
summary(fit.wt)
```

```
##
## Call:
## lm(formula = mpg ~ wt * am, data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6004 -1.5446 -0.5325  0.9012  6.0909
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept)  31.4161     3.0201  10.402
## wt           -3.7859     0.7856  -4.819
## am1          14.8784     4.2640   3.489
## wt:am1       -5.2984     1.4447  -3.667
##             Pr(>|t|)
## (Intercept) 4.00e-11 ***
## wt          4.55e-05 ***
## am1          0.00162 **
## wt:am1       0.00102 **
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.591 on 28 degrees of freedom
## Multiple R-squared:  0.833,  Adjusted R-squared:  0.8151
## F-statistic: 46.57 on 3 and 28 DF,  p-value: 5.209e-11
```

```
fit.hp <- lm(mpg ~ hp * am, data = cs.dt)
summary(fit.hp)
```

```
##
## Call:
## lm(formula = mpg ~ hp * am, data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3818 -2.2696  0.1344  1.7058  5.8752
```

```
##
## Coefficients:
##               Estimate Std. Error t value
## (Intercept) 26.6248479  2.1829432  12.197
## hp          -0.0591370  0.0129449  -4.568
## am1          5.2176534  2.6650931   1.958
## hp:am1       0.0004029  0.0164602   0.024
##             Pr(>|t|)
## (Intercept) 1.01e-12 ***
## hp          9.02e-05 ***
## am1           0.0603 .
## hp:am1        0.9806
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.961 on 28 degrees of freedom
## Multiple R-squared:  0.782,  Adjusted R-squared:  0.7587
## F-statistic: 33.49 on 3 and 28 DF,  p-value: 2.112e-09
```

```
fit.wthp <- lm(mpg ~ (hp + wt) * am, data = cs.dt)
summary(fit.wthp)
```

```
##
## Call:
## lm(formula = mpg ~ (hp + wt) * am, data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9873 -1.4467 -0.5355  1.2614  5.5987
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept) 30.70393    2.67515  11.477
## hp          -0.04094    0.01363  -3.004
## wt          -1.85591    0.94511  -1.964
## am1         13.74000    4.22337   3.253
## hp:am1       0.02779    0.01921   1.447
## wt:am1      -5.76895    2.07201  -2.784
##             Pr(>|t|)
## (Intercept) 1.12e-11 ***
## hp           0.00583 **
## wt           0.06034 .
## am1          0.00316 **
## hp:am1       0.15983
```

```
## wt:am1       0.00987 **
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.286 on 26 degrees of freedom
## Multiple R-squared:  0.8793, Adjusted R-squared:  0.8561
## F-statistic: 37.89 on 5 and 26 DF,  p-value: 3.901e-11
```

```
summary(fit.full)$coef
```

```
##               Estimate  Std. Error
## (Intercept) 23.87913244 20.06582026
## cyl6        -2.64869528  3.04089041
## cyl8        -0.33616298  7.15953951
## disp         0.03554632  0.03189920
## hp          -0.07050683  0.03942556
## drat         1.18283018  2.48348458
## wt          -4.52977584  2.53874584
## qsec         0.36784482  0.93539569
## vs1          1.93085054  2.87125777
## am1          1.21211570  3.21354514
## gear4        1.11435494  3.79951726
## gear5        2.52839599  3.73635801
## carb2       -0.97935432  2.31797446
## carb3        2.99963875  4.29354611
## carb4        1.09142288  4.44961992
## carb6        4.47756921  6.38406242
## carb8        7.25041126  8.36056638
##                 t value   Pr(>|t|)
## (Intercept)  1.19004018 0.25252548
## cyl6        -0.87102622 0.39746642
## cyl8        -0.04695316 0.96317000
## disp         1.11433290 0.28267339
## hp          -1.78835344 0.09393155
## drat         0.47627845 0.64073922
## wt          -1.78425732 0.09461859
## qsec         0.39325050 0.69966720
## vs1          0.67247551 0.51150791
## am1          0.37718957 0.71131573
## gear4        0.29328856 0.77332027
## gear5        0.67670068 0.50889747
## carb2       -0.42250436 0.67865093
## carb3        0.69863900 0.49546781
## carb4        0.24528452 0.80956031
```

```
## carb6        0.70136677 0.49381268
## carb8        0.86721532 0.39948495
```

```
ggplot(cs.dt, aes(hp, mpg, colour = am)) + geom_point() +
    geom_text(aes(label = cyl, colour = NULL),
        vjust = -0.6) + theme_tufte() + geom_smooth(method = "lm")
```

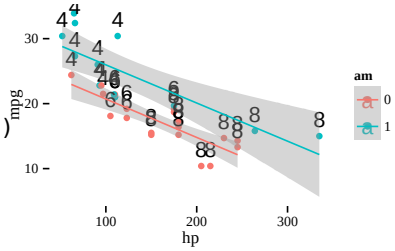


Figure 3: Sepal length vs. petal length, colored by species

```
ggplot(cs.dt, aes(hp, wt, colour = am)) + geom_point() +
    geom_text(aes(label = cyl, colour = NULL),
        vjust = -0.6) + theme_tufte() + geom_smooth(method = "lm")
```

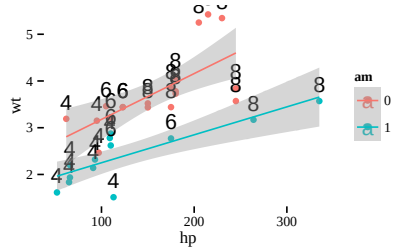


Figure 4: Sepal length vs. petal length, colored by species

```
ggplot(cs.dt, aes(disp, mpg, colour = am)) + geom_point() +
    theme_tufte() + geom_smooth(method = "lm")
```

```
ggplot(cs.dt, aes(wt, disp, colour = am)) + geom_point() +
    theme_tufte() + geom_smooth(method = "lm")
```

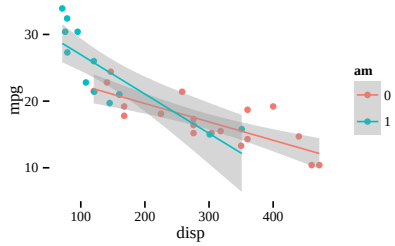


Figure 5: Sepal length vs. petal length, colored by species

Is an automatic or manual transmission better for MPG?

It seems so.

```
fit.wt <- lm(mpg ~ wt * am, data = cs.dt)
summary(fit.wt)
```

```
##
## Call:
## lm(formula = mpg ~ wt * am, data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6004 -1.5446 -0.5325  0.9012  6.0909
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept)  31.4161     3.0201  10.402
## wt           -3.7859     0.7856  -4.819
## am1          14.8784     4.2640   3.489
## wt:am1       -5.2984     1.4447  -3.667
##             Pr(>|t|)
## (Intercept) 4.00e-11 ***
## wt          4.55e-05 ***
## am1          0.00162 **
## wt:am1       0.00102 **
```

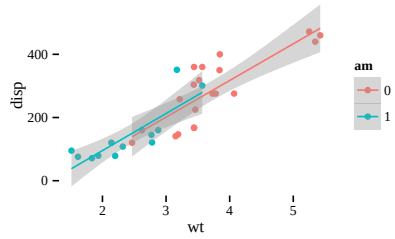


Figure 6: Sepal length vs. petal length, colored by species

```
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.591 on 28 degrees of freedom
## Multiple R-squared:  0.833,  Adjusted R-squared:  0.8151
## F-statistic: 46.57 on 3 and 28 DF,  p-value: 5.209e-11
```

```
fit.hp <- lm(mpg ~ hp * am, data = cs.dt)
summary(fit.hp)
```

```
##
## Call:
## lm(formula = mpg ~ hp * am, data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3818 -2.2696  0.1344  1.7058  5.8752
##
## Coefficients:
##               Estimate Std. Error t value
## (Intercept) 26.6248479  2.1829432  12.197
## hp          -0.0591370  0.0129449  -4.568
## am1          5.2176534  2.6650931   1.958
## hp:am1       0.0004029  0.0164602   0.024
##             Pr(>|t|)
## (Intercept) 1.01e-12 ***
## hp          9.02e-05 ***
## am1           0.0603 .
## hp:am1        0.9806
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.961 on 28 degrees of freedom
## Multiple R-squared:  0.782,  Adjusted R-squared:  0.7587
## F-statistic: 33.49 on 3 and 28 DF,  p-value: 2.112e-09
```

```
fit.hpwt <- lm(mpg ~ (hp + wt) * am, data = cs.dt)
summary(fit.hpwt)
```

```
##
## Call:
## lm(formula = mpg ~ (hp + wt) * am, data = cs.dt)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -2.9873 -1.4467 -0.5355  1.2614  5.5987
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept) 30.70393    2.67515  11.477
## hp          -0.04094    0.01363  -3.004
## wt          -1.85591    0.94511  -1.964
## am1         13.74000    4.22337   3.253
## hp:am1       0.02779    0.01921   1.447
## wt:am1      -5.76895    2.07201  -2.784
##             Pr(>|t|)
## (Intercept) 1.12e-11 ***
## hp           0.00583 **
## wt           0.06034 .
## am1          0.00316 **
## hp:am1       0.15983
## wt:am1       0.00987 **
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.286 on 26 degrees of freedom
## Multiple R-squared:  0.8793, Adjusted R-squared:  0.8561
## F-statistic: 37.89 on 5 and 26 DF,  p-value: 3.901e-11
```

```
fit2 <- lm(mpg ~ wt + hp, data = cs.dt)
summary(fit2)
```

```
##
## Call:
## lm(formula = mpg ~ wt + hp, data = cs.dt)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.941 -1.600 -0.182  1.050  5.854
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept) 37.22727    1.59879  23.285
## wt          -3.87783    0.63273  -6.129
## hp          -0.03177    0.00903  -3.519
##             Pr(>|t|)
## (Intercept)  < 2e-16 ***
## wt          1.12e-06 ***
```

```
## hp           0.00145 **
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.593 on 29 degrees of freedom
## Multiple R-squared:  0.8268, Adjusted R-squared:  0.8148
## F-statistic: 69.21 on 2 and 29 DF,  p-value: 9.109e-12
```

```
fit1 <- lm(mpg ~ wt, data = cs.dt)
fit <- lm(mpg ~ ., data = cs.dt)
summary(fit)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = cs.dt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5087 -1.3584 -0.0948  0.7745  4.6251
##
## Coefficients:
##             Estimate Std. Error t value
## (Intercept) 23.87913   20.06582   1.190
## cyl6        -2.64870    3.04089  -0.871
## cyl8        -0.33616    7.15954  -0.047
## disp         0.03555    0.03190   1.114
## hp          -0.07051    0.03943  -1.788
## drat         1.18283    2.48348   0.476
## wt          -4.52978    2.53875  -1.784
## qsec         0.36784    0.93540   0.393
## vs1          1.93085    2.87126   0.672
## am1          1.21212    3.21355   0.377
## gear4        1.11435    3.79952   0.293
## gear5        2.52840    3.73636   0.677
## carb2       -0.97935    2.31797  -0.423
## carb3        2.99964    4.29355   0.699
## carb4        1.09142    4.44962   0.245
## carb6        4.47757    6.38406   0.701
## carb8        7.25041    8.36057   0.867
##             Pr(>|t|)
## (Intercept)   0.2525
## cyl6          0.3975
## cyl8          0.9632
## disp          0.2827
```

```
## hp            0.0939 .
## drat          0.6407
## wt            0.0946 .
## qsec          0.6997
## vs1           0.5115
## am1           0.7113
## gear4         0.7733
## gear5         0.5089
## carb2         0.6787
## carb3         0.4955
## carb4         0.8096
## carb6         0.4938
## carb8         0.3995
## ---
## Signif. codes:
##   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  ' ' 1
##
## Residual standard error: 2.833 on 15 degrees of freedom
## Multiple R-squared:  0.8931, Adjusted R-squared:  0.779
## F-statistic:  7.83 on 16 and 15 DF,  p-value: 0.000124
```

The MPG difference between automatic and manual transmissions

```
wt.max <- max(cs.dt[am == 1, wt])
wt.min <- min(cs.dt[am == 0, wt])
wt.min
```

```
## [1] 2.465
```

```
cs.dt.res <- cs.dt[wt >= wt.min & wt <= wt.max]
ggplot(cs.dt.res, aes(wt, mpg, colour = am)) +
    geom_point() + theme_tufte() + geom_smooth(method = "lm")
```

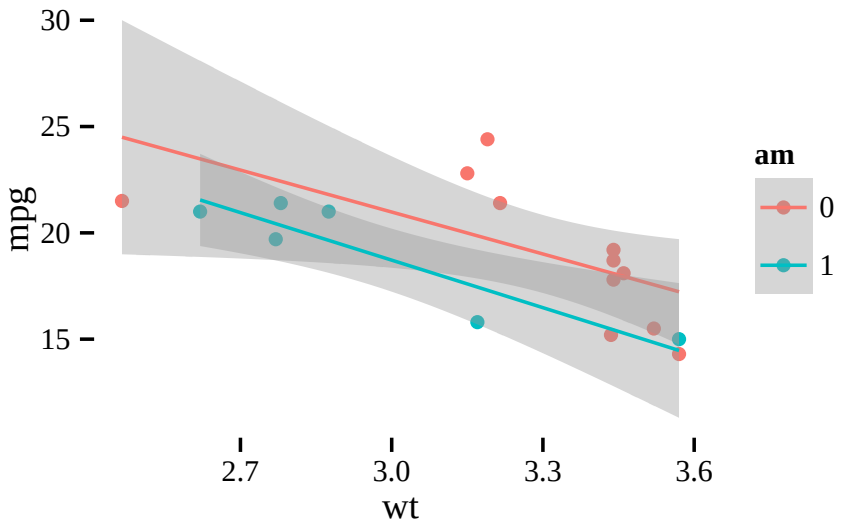


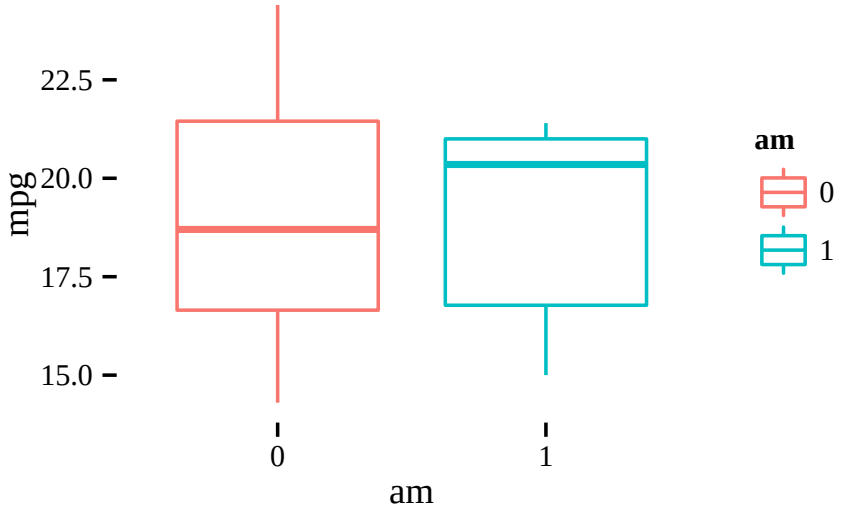
```
ggplot(cs.dt.res, aes(am, mpg, colour = am)) +
    geom_boxplot() + theme_tufte()
```


Appendix

```
pairs(cs.dt[, .(mpg, cyl, disp, hp, drat, wt,
    qsec, vs, am, gear, carb)], panel = panel.smooth)
```


This style provides a- and b-heads (that is, # and ##), demonstrated above. An error is emitted if you try to use ### and smaller headings.


30
25
20
15
mpg
2.7
3.0
3.3
3.6
wt
am
0
1


22.5
20.0
17.5
15.0
mpg
0
1
am
am
0
1

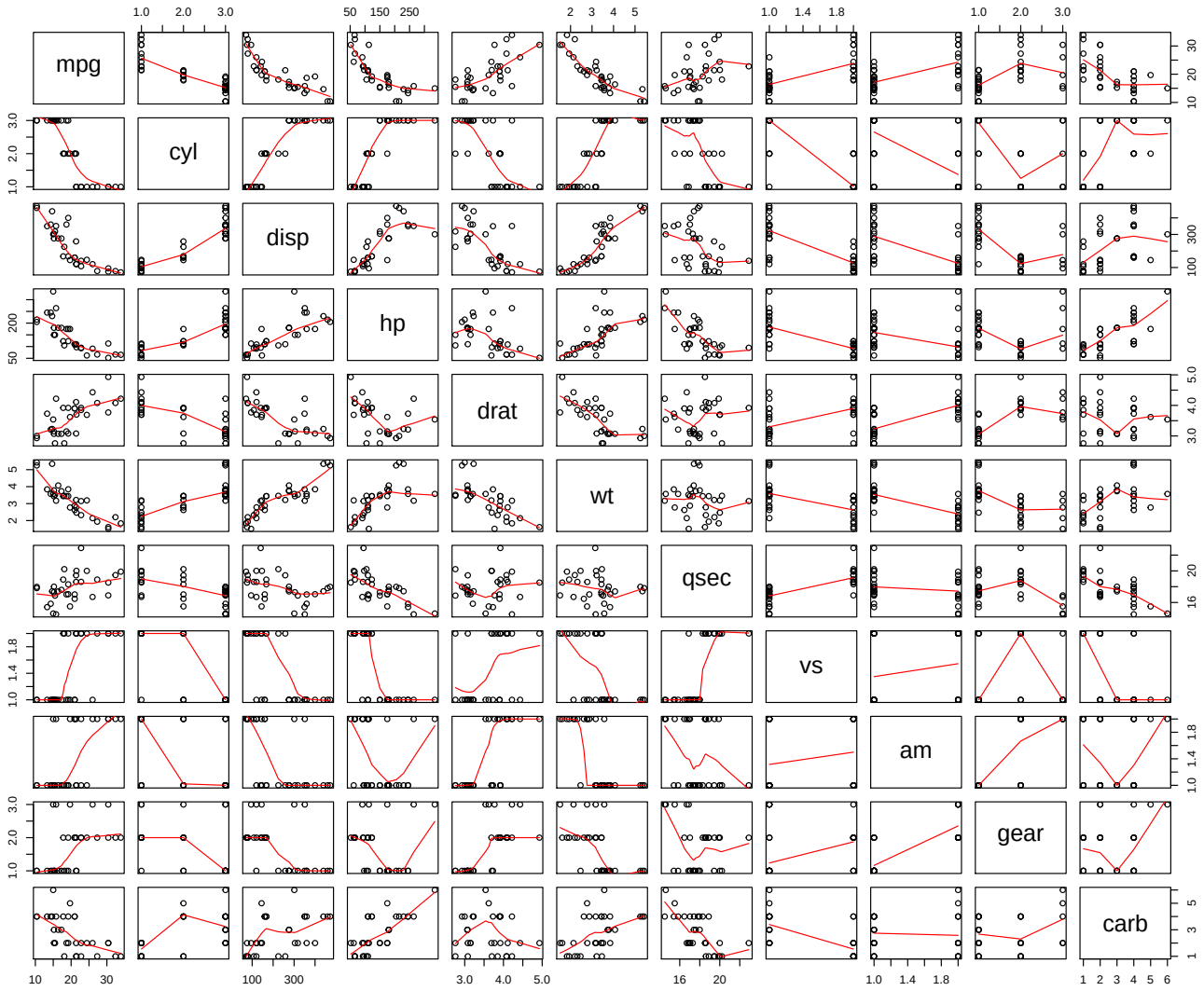


Figure 7: Full width figure

In his later books[1], Tufte starts each section with a bit of vertical space, a non-indented paragraph, and sets the first few words of the sentence in small caps. To accomplish this using this style, use the `\newthought` command as demonstrated at the beginning of this paragraph.

[1] http://www.edwardtufte.com/tufte/books_be

Figures

Margin Figures

Images and graphics play an integral role in Tufte's work. To place figures or tables in the margin you can use the `fig.margin` knitr chunk option. For example:

```
library(ggplot2)
qplot(Sepal.Length, Petal.Length, data = iris,
    color = Species)
```

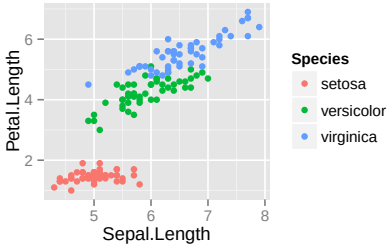


Figure 8: Sepal length vs. petal length, colored by species

Note the use of the `fig.cap` chunk option to provide a figure caption. You can adjust the proportions of figures using the `fig.width` and `fig.height` chunk options. These are specified in inches, and will be automatically scaled down to fit within the handout margin.

Equations

You can also include LaTeX equations in the margin by explicitly invoking the `marginfigure` environment.

$$\frac{d}{dx}\left(\int_0^x f(u)\,du\right) = f(x).$$

Figure 9: An equation

Note the use of the `\caption` command to add additional text below the equation.

Full Width Figures

You can arrange for figures to span across the entire page by using the `fig.fullwidth` chunk option.

```
qplot(wt, mpg, data = mtcars, colour = factor(cyl))
```

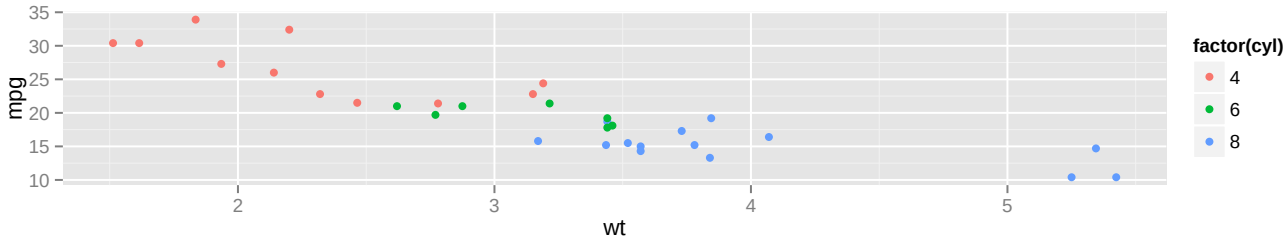


Figure 10: Full width figure

Note the use of the `fig.width` and `fig.height` chunk options to establish the proportions of the figure. Full width figures look much better if their height is minimized.

Main Column Figures

Besides margin and full width figures, you can of course also include figures constrained to the main column.

```
qplot(factor(cyl), mpg, data = mtcars, geom = "boxplot")
```

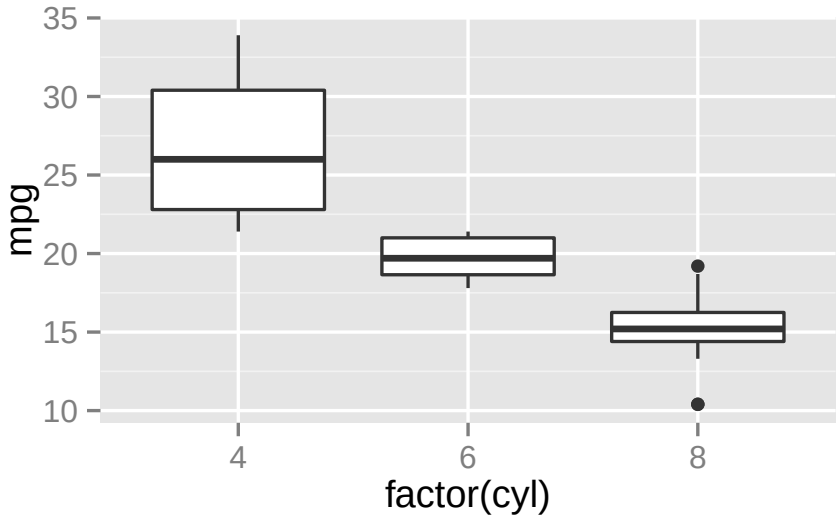


Figure 11: Another figure

Sidenotes

One of the most prominent and distinctive features of this style is the extensive use of sidenotes. There is a wide margin to provide ample room for sidenotes and small figures. Any use of a footnote will automatically be converted to a sidenote. [2]

[2] This is a sidenote that was entered using a footnote.

If you'd like to place ancillary information in the margin without the sidenote mark (the superscript number), you can use the `\marginnote` command.

This is a margin note. Notice that there isn't a number preceding the note.

Note also that the two footnote references (`tufte_latex` and `books_be`, both defined below) were also included in the margin on the first page of this document.

Tables

You can use the **xtable** package to format LaTeX tables that integrate well with the rest of the Tufte handout style. Note that it's important

to set the `xtable.comment` and `xtable.booktabs` options as shown below to ensure the table is formatted correctly for inclusion in the document.

```
library(xtable)
options(xtable.comment = FALSE)
options(xtable.booktabs = TRUE)
xtable(head(mtcars[, 1:6]), caption = "First rows of mtcars")
```

	mpg	cyl	disp	hp	drat	wt
Mazda RX4	21.00	6.00	160.00	110.00	3.90	2.62
Mazda RX4 Wag	21.00	6.00	160.00	110.00	3.90	2.88
Datsun 710	22.80	4.00	108.00	93.00	3.85	2.32
Hornet 4 Drive	21.40	6.00	258.00	110.00	3.08	3.21
Hornet Sportabout	18.70	8.00	360.00	175.00	3.15	3.44
Valiant	18.10	6.00	225.00	105.00	2.76	3.46

Table 1: First rows of mtcars