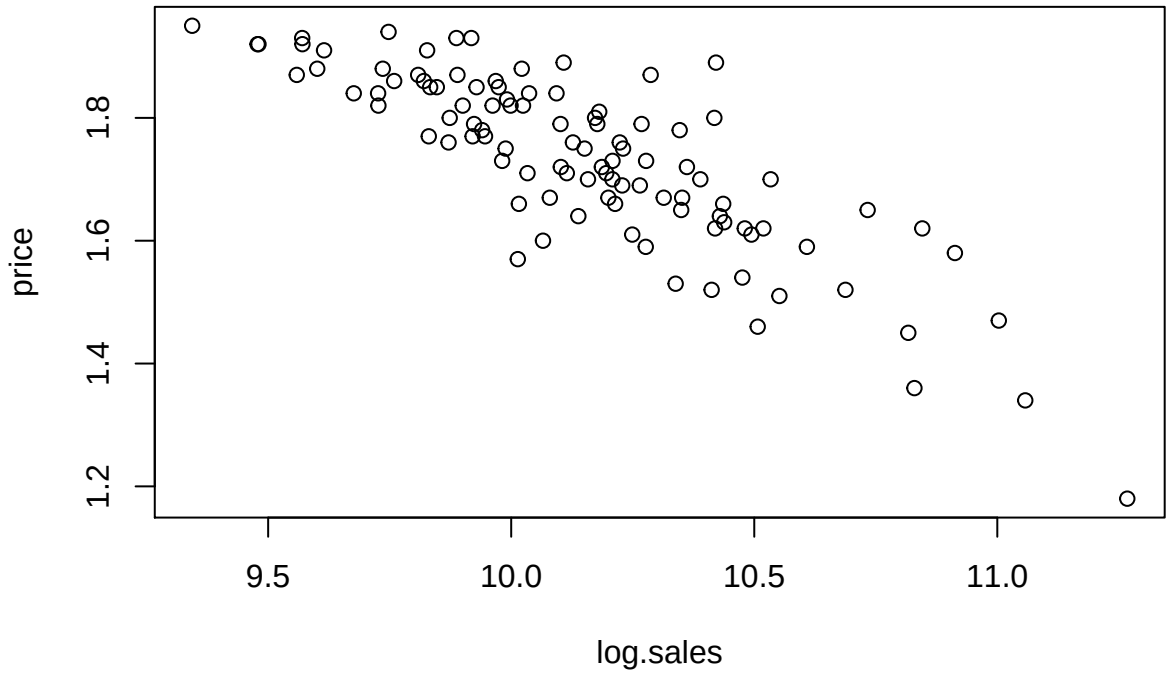
ST565: Time Series HW6

Amirhosein "Emerson" Azarbakht azarbaka@oregonstate.edu

Question 1

The dataset bluebirdlite contains log sales and prices for the "lite" version of bluebird chips. Quantify the relationship between sales and price.

```
plot(bluebirdlite)
```

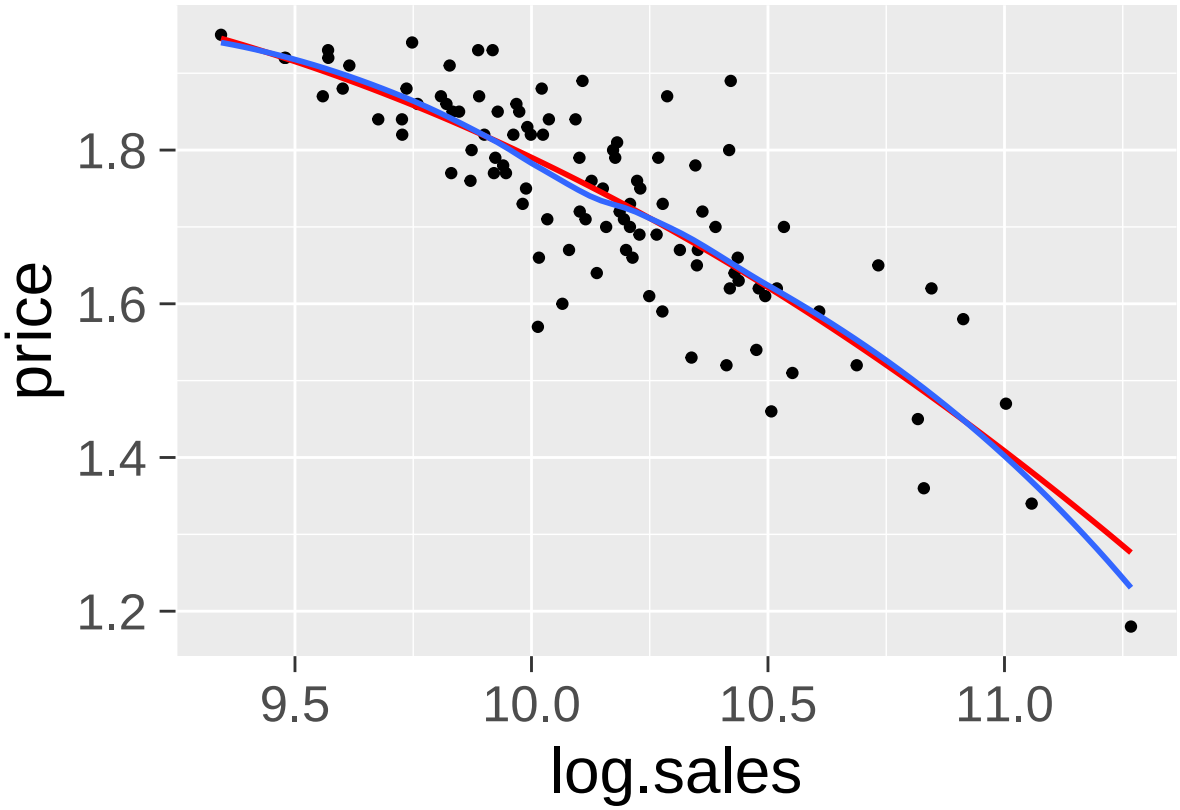



```
big_font <- theme_grey(base_size =  24)


attach(bluebirdlite)

fit_lm <- lm(log.sales ~ price + I(price^2), data = bluebirdlite)
summary(fit_lm)
```

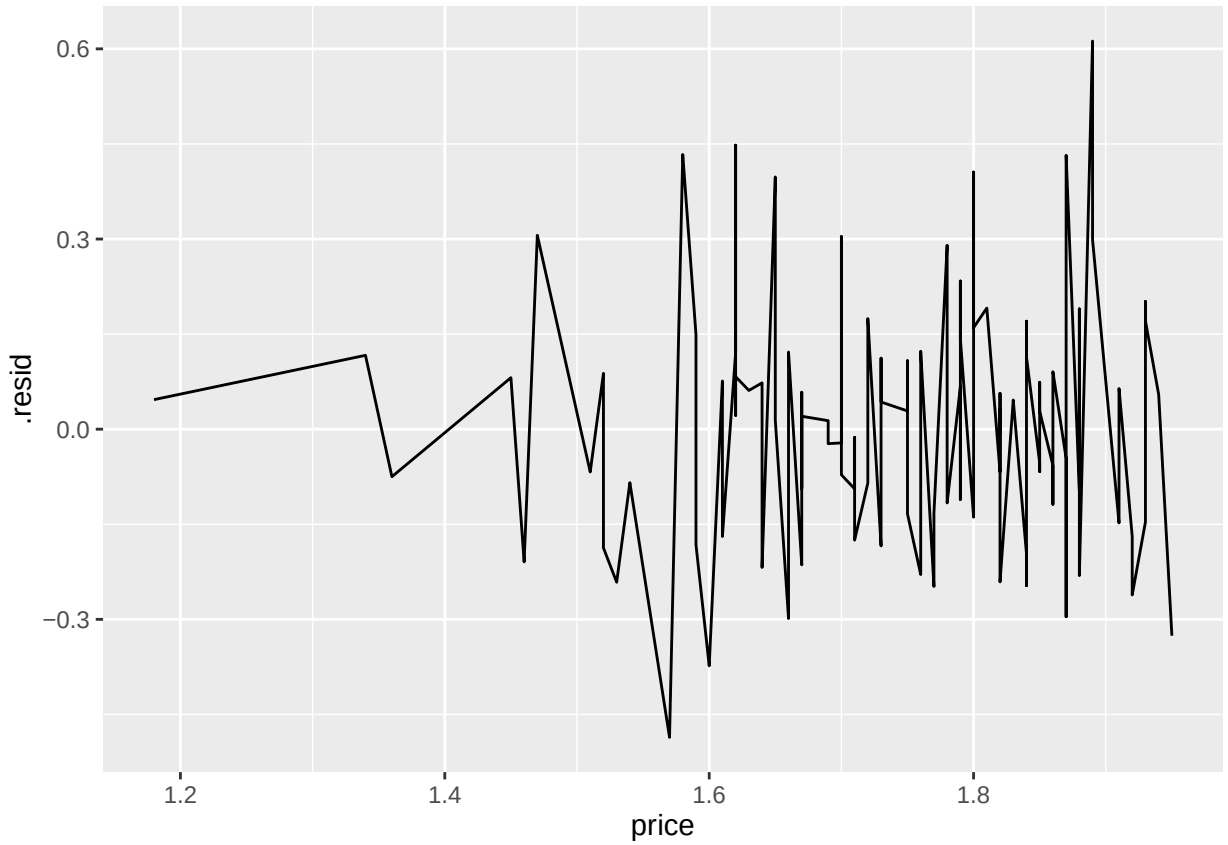
```
##
## Call:
## lm(formula = log.sales ~ price + I(price^2), data = bluebirdlite)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.48609 -0.14069  0.01389  0.11249  0.61228
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.5954     1.6938   7.436 3.48e-11 ***
## price        -0.6507     2.0404  -0.319    0.750
## I(price^2)   -0.4357     0.6121  -0.712    0.478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1968 on 101 degrees of freedom
## Multiple R-squared:  0.7054, Adjusted R-squared:  0.6996
## F-statistic: 120.9 on 2 and 101 DF,  p-value: < 2.2e-16
```

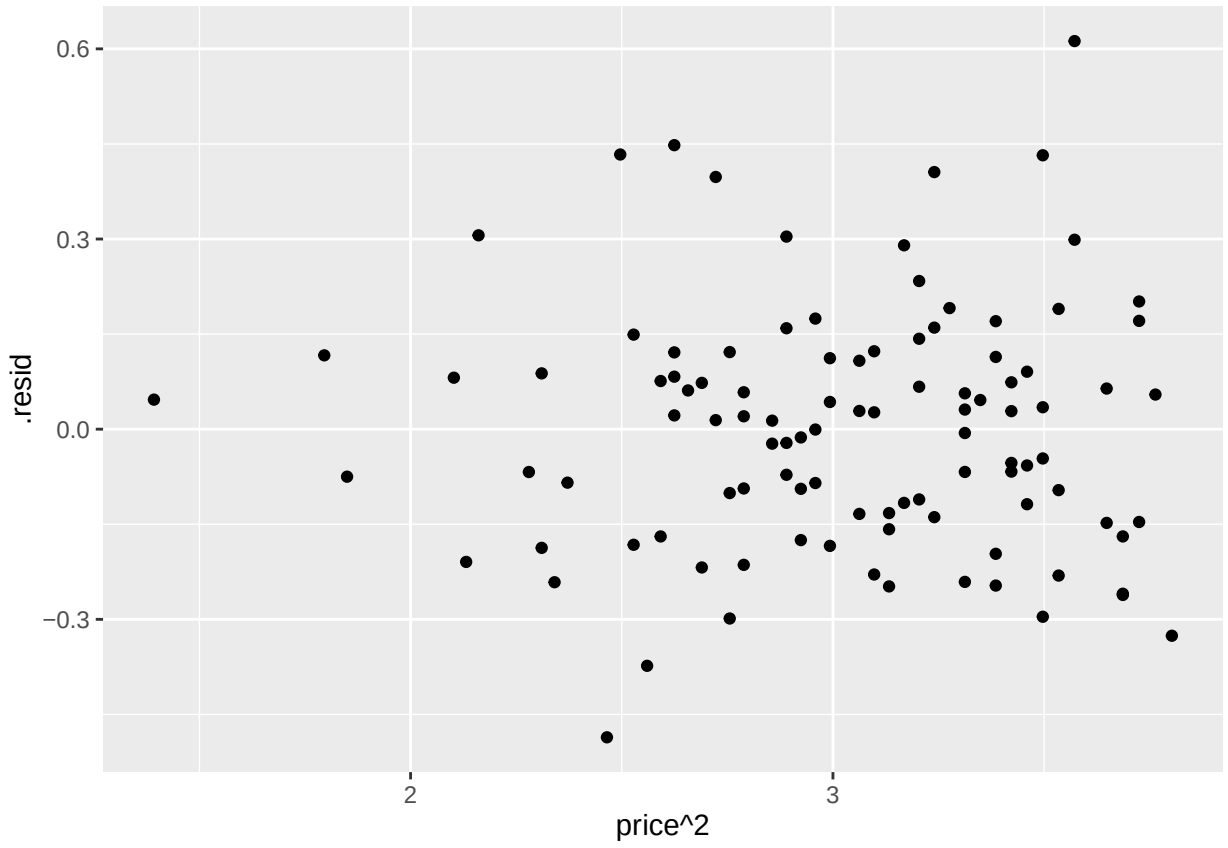
```
qplot(log.sales, price, data = bluebirdlite) +
  geom_smooth(method = "lm", formula = y ~ poly(x, 2), se = FALSE, colour = "red") +
  geom_smooth(se = FALSE)  +
  big_font
```



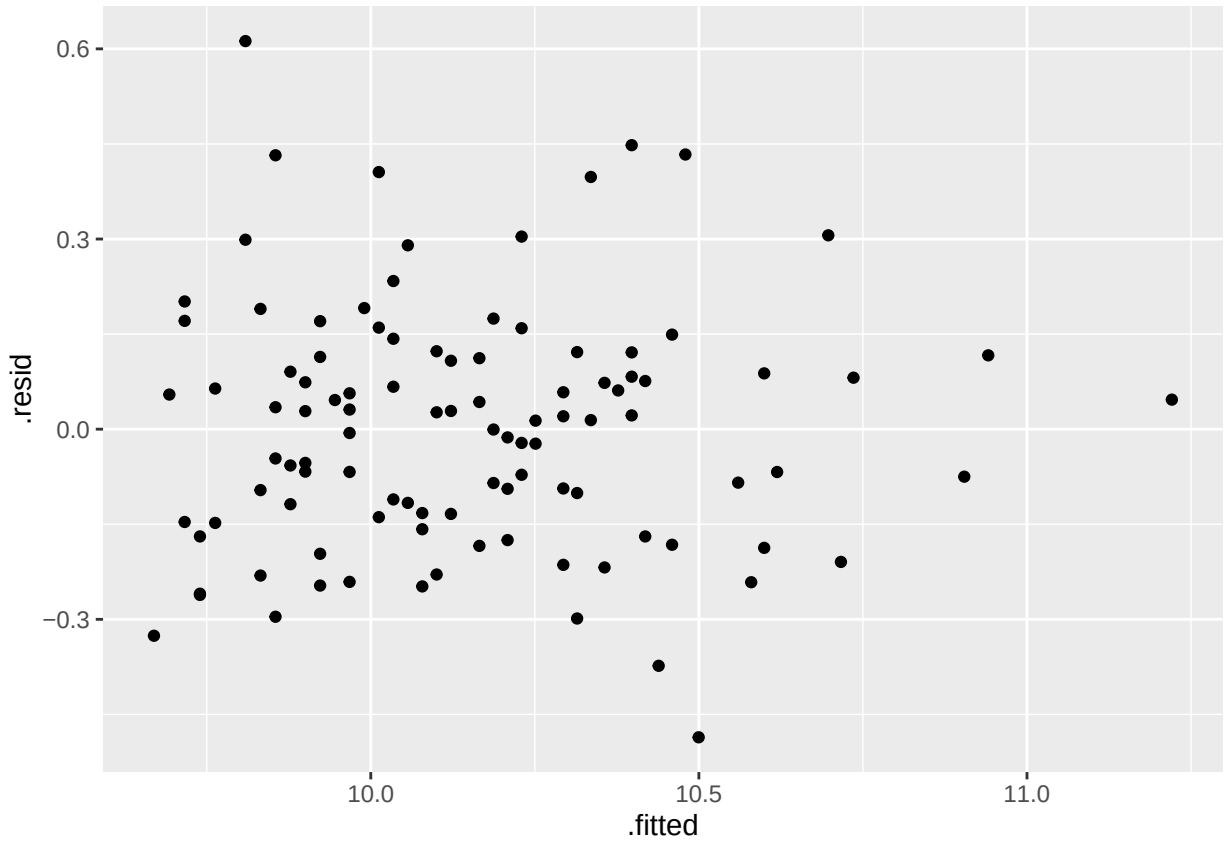

```
# assumptions
# residuals versus covariates
sales_lm <- fortify(fit_lm)
qplot(price, .resid, data = sales_lm, geom= "line")
```



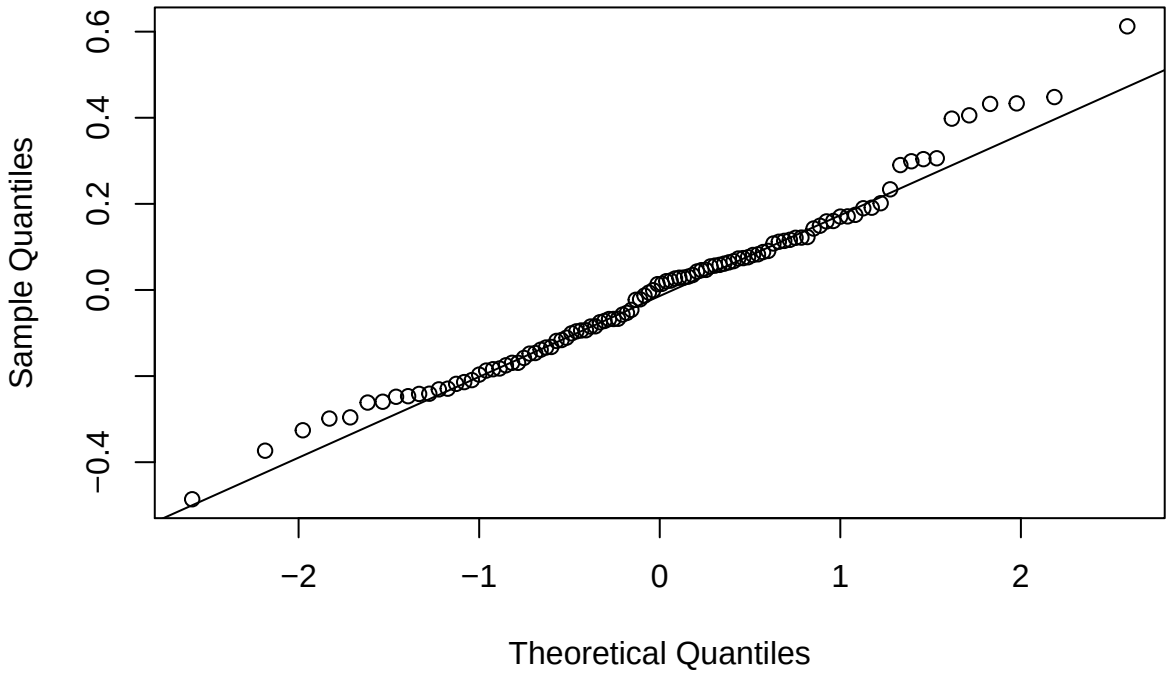

```
qplot(price^2, .resid, data = sales_lm)
```




```
# residuals versus fitted
qplot(.fitted, .resid, data = sales_lm)
```

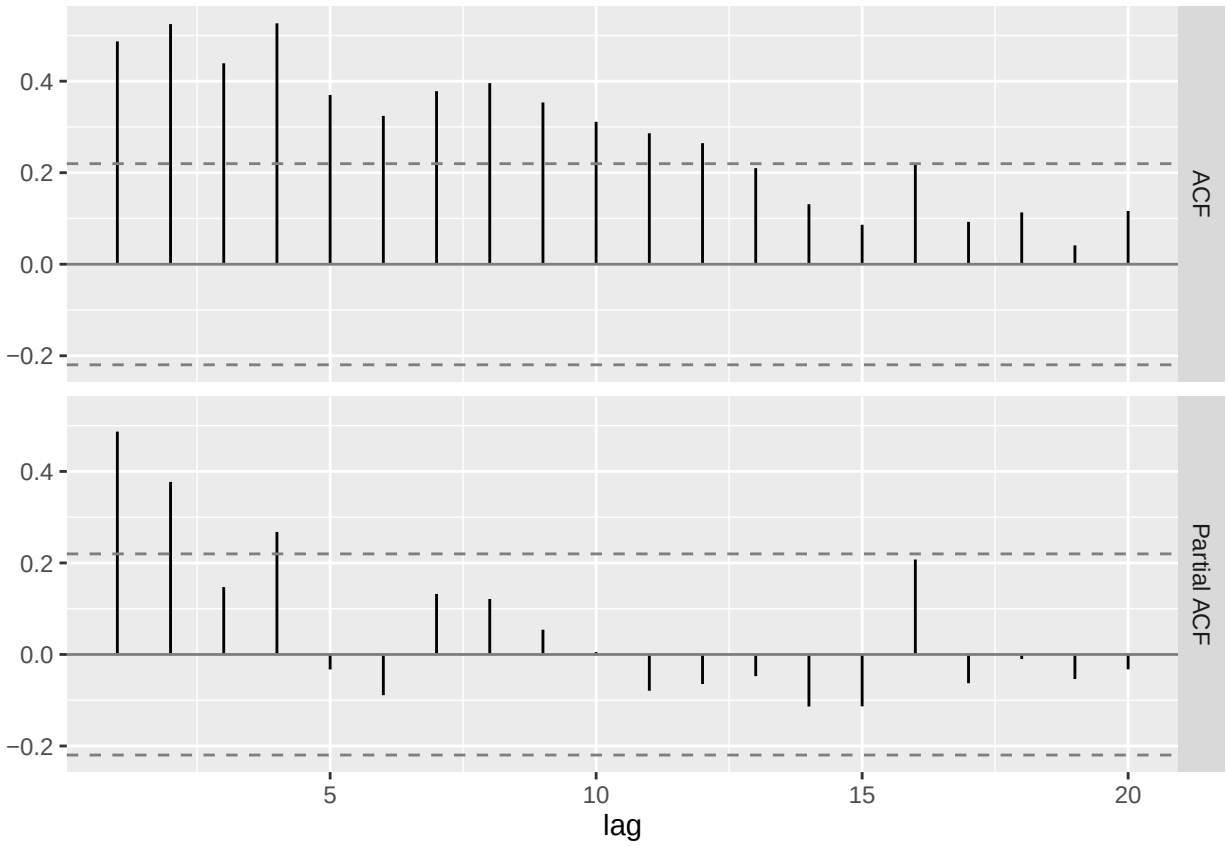



```
# normality of residuals
qqnorm(sales_lm$.resid)
qqline(sales_lm$.resid)
```

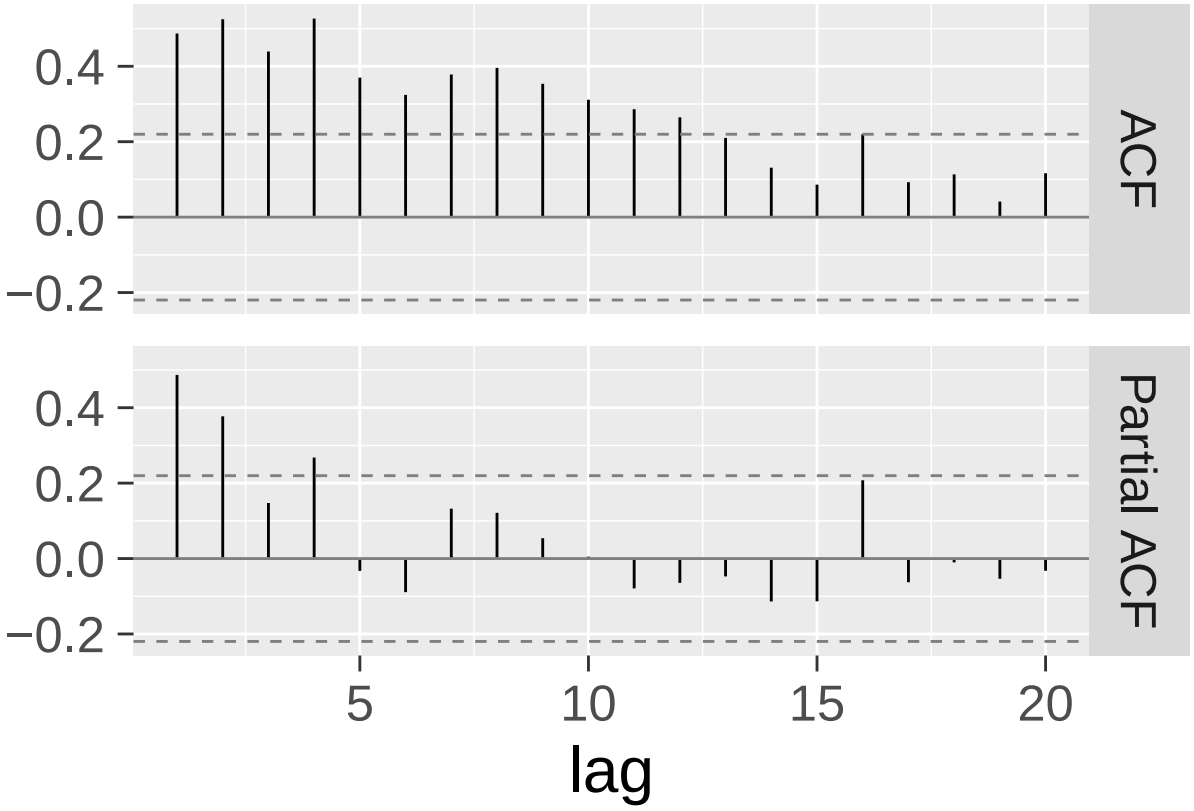



```
# correlation of residuals
source(url("http://stat565.cwick.co.nz/code/get_acf.R")) # my code for examine_corr
examine_corr(residuals(fit_lm))
```

```
## Warning: closing unused connection 5 (http://stat565.cwick.co.nz/code/
## get_acf.R)
```




```
last_plot() + big_font
```

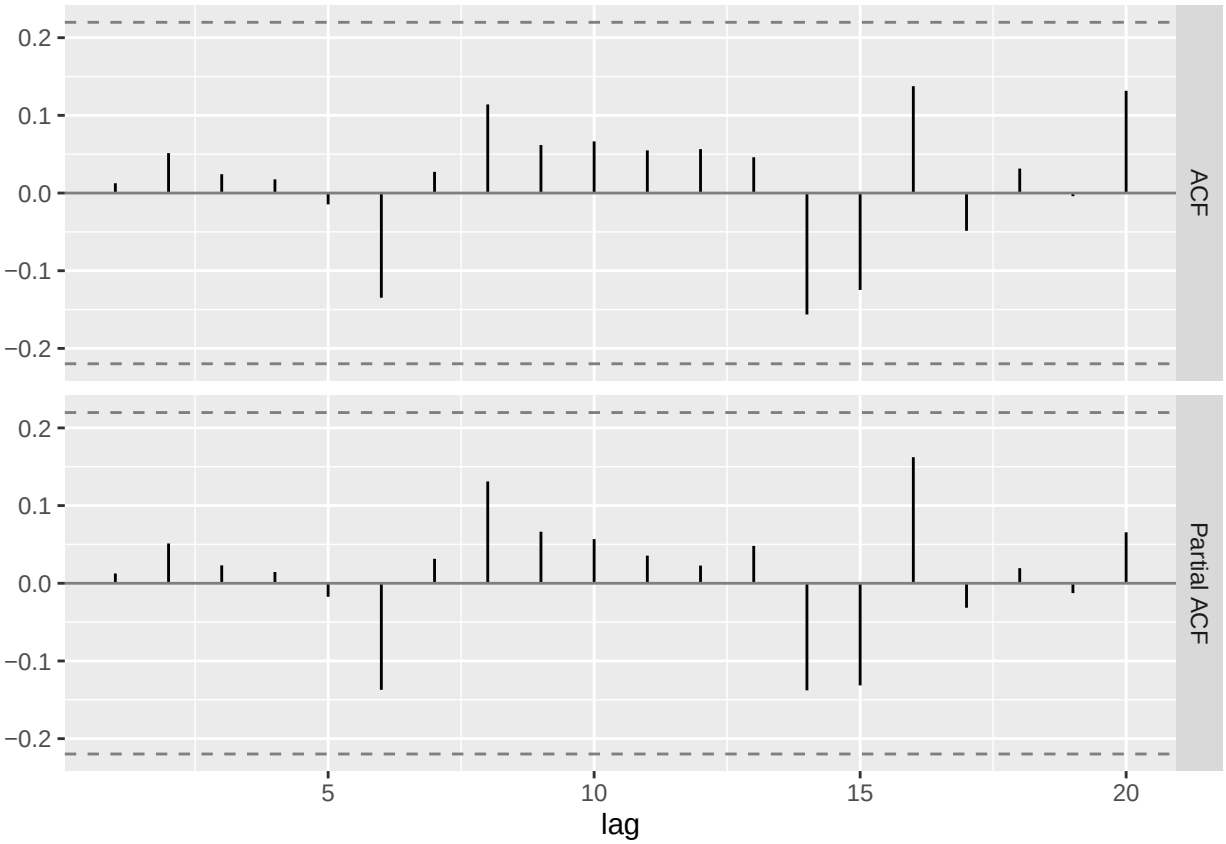



```
# AR (4)? violates regression assumptions

library(nlme)
gls_fit <- gls(log.sales ~ price + I(price^2), data = bluebirdlite,
    correlation = corARMA(p = 4), method = "ML")
summary(gls_fit)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: log.sales ~ price + I(price^2)
##   Data: bluebirdlite
##        AIC       BIC   logLik
##   -92.6015 -71.44638 54.30075
##
## Correlation Structure: ARMA(4,0)
##  Formula: ~1
##  Parameter estimate(s):
##      Phi1      Phi2      Phi3      Phi4
## 0.1851510 0.2132629 0.1345612 0.3039307
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept) 14.899928 1.1651084 12.788448  0.0000
## price       -3.658602 1.3918337 -2.628620  0.0099
## I(price^2)   0.523881 0.4163226  1.258355  0.2112
##
##  Correlation:
```

```
##             (Intr) price
## price       -0.994
## I(price^2)   0.985 -0.997
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -2.23177526 -0.70553535  0.04652695  0.65792580  2.90668252
##
## Residual standard error: 0.1943017
## Degrees of freedom: 104 total; 101 residual
```

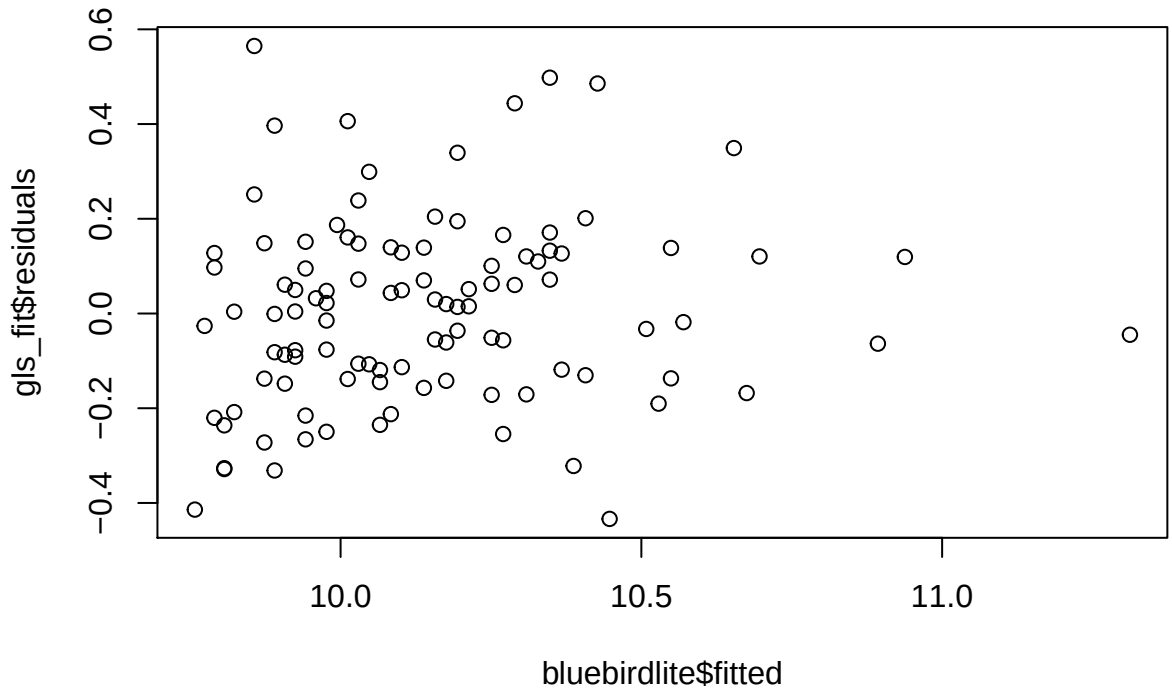
```
# or
# arima_fit <- with(bluebirdlite,
#   arima(log.sales, order = c(4, 0, 1), xreg = cbind(price, I(price^2))))
# arima_fit

# diagnostics
fit_lm$residuals <- residuals(gls_fit, type = "normalized")
bluebirdlite$fitted <- fitted(gls_fit)

examine_corr(fit_lm$residuals)
```



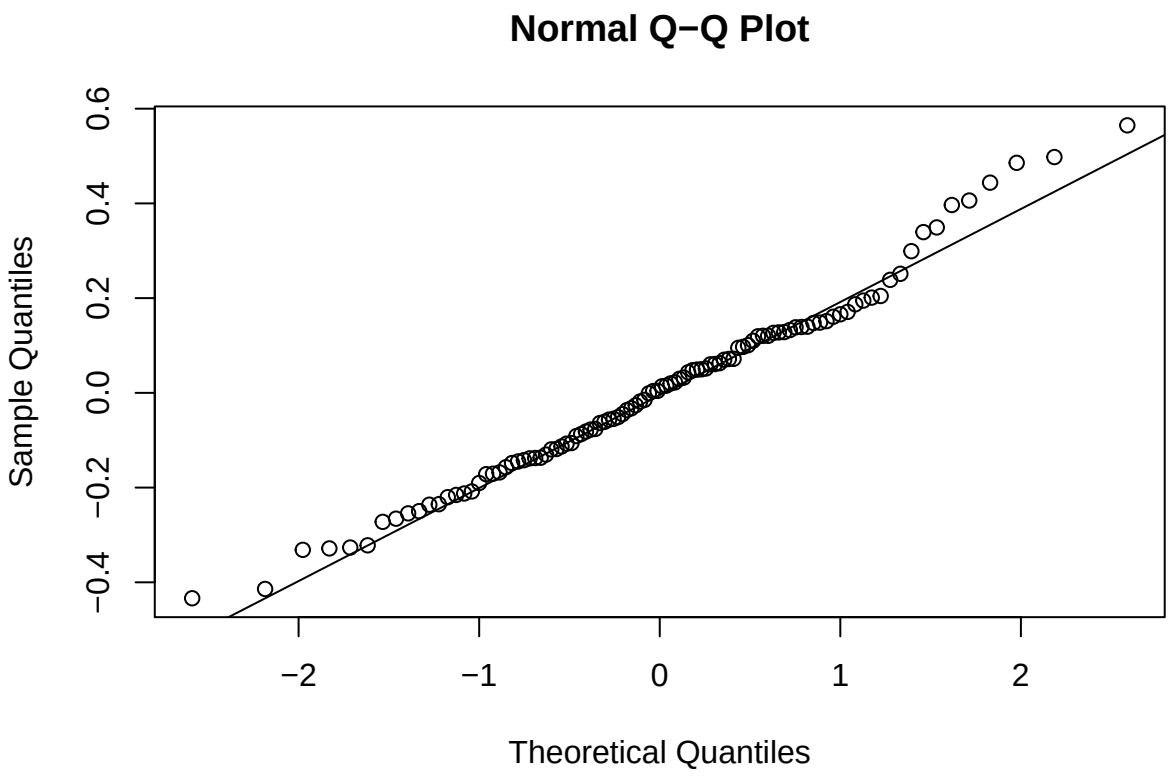

```
plot(bluebirdlite$fitted, gls_fit$residuals, data = bluebirdlite)
```

```
## Warning in plot.window(...): "data" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "data" is not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "data" is not
## a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "data" is not
## a graphical parameter

## Warning in box(...): "data" is not a graphical parameter

## Warning in title(...): "data" is not a graphical parameter
```




```
qqnorm(gls_fit$residuals)
qqline(gls_fit$residuals)
```




```
confint(gls_fit)
```

```
##                  2.5 %     97.5 %
## (Intercept) 12.6163578 17.1834987
## price       -6.3865463 -0.9306585
## I(price^2)  -0.2920958  1.3398587
```

```
confint(fit_lm)
```

```
##                  2.5 %    97.5 %
## (Intercept)  -4.733007 29.923726
## price       -21.524851 20.223381
## I(price^2)   -6.697864  5.826414
```

```
intervals(gls_fit)
```

```
## Approximate 95% confidence intervals
##
##  Coefficients:
##                  lower       est.      upper
## (Intercept) 12.5886666 14.8999282 17.2111898
## price       -6.4196260 -3.6586024 -0.8975787
## I(price^2)  -0.3019906  0.5238814  1.3497534
## attr(,"label")
## [1] "Coefficients:"
##
```

```
##  Correlation structure:
##           lower      est.       upper
## Phi1 0.10104695 0.1851510 -0.06254519
## Phi2 0.12509406 0.2132629  0.23165667
## Phi3 0.01130536 0.1345612  0.32715590
## Phi4 0.11381837 0.3039307  0.47257119
## attr(,"label")
## [1] "Correlation structure:"
##
##  Residual standard error:
##     lower      est.     upper
## 0.1394571 0.1943017 0.2707152
```

```
#' It is estimated that an increase in price of 10 cents is associated with decrease in median sales of

#' Interestingly, accounting for the correlation changes our estimate only slightly here.
```

Question 2

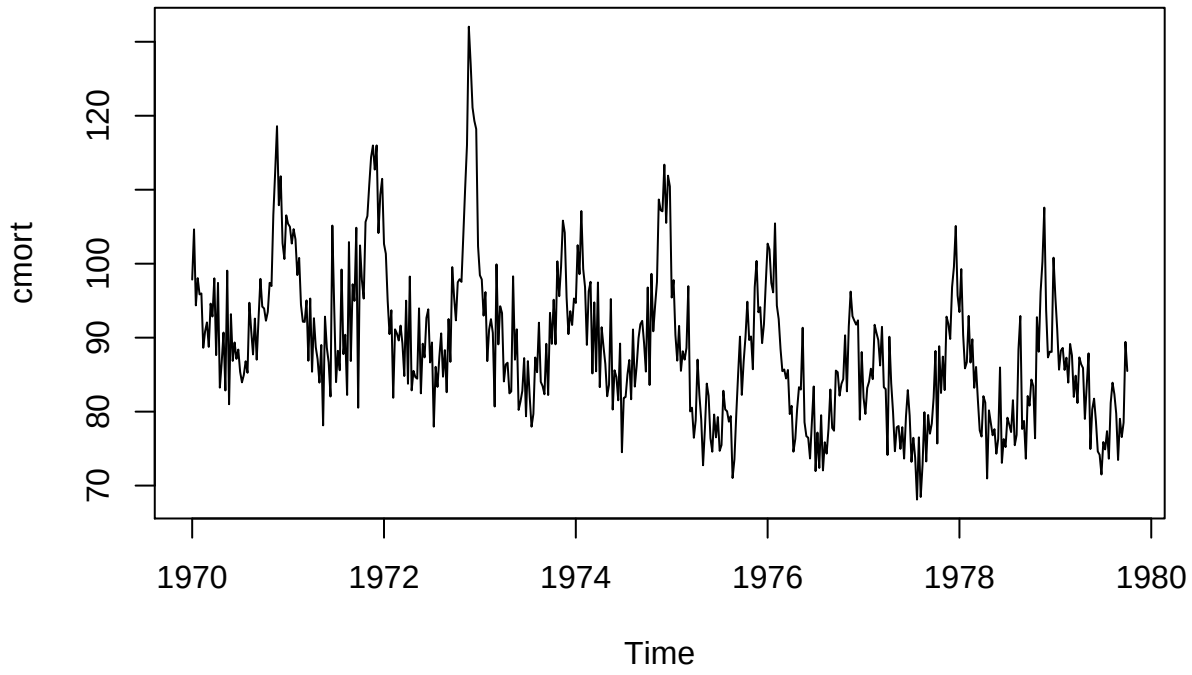
In class we looked at modelling the relationship between mortality, temperature and particulate matter. Repeat the analysis but seasonally difference all three series first. Compare the results.

You can get the cmort, tempr and part times series with:

```
tsa3 <- load(url("http://www.stat.pitt.edu/stoffer/tsa3/tsa3.rda"))
str(cmort)
```

```
##  Time-Series [1:508] from 1970 to 1980: 97.8 104.6 94.4 98 95.8 ...
```

```
mort <- diff(cmort, lag = 52)
plot(cmort )
```


cmort
70
80
90
100
120
1970
1972
1974
1976
1978
1980
Time