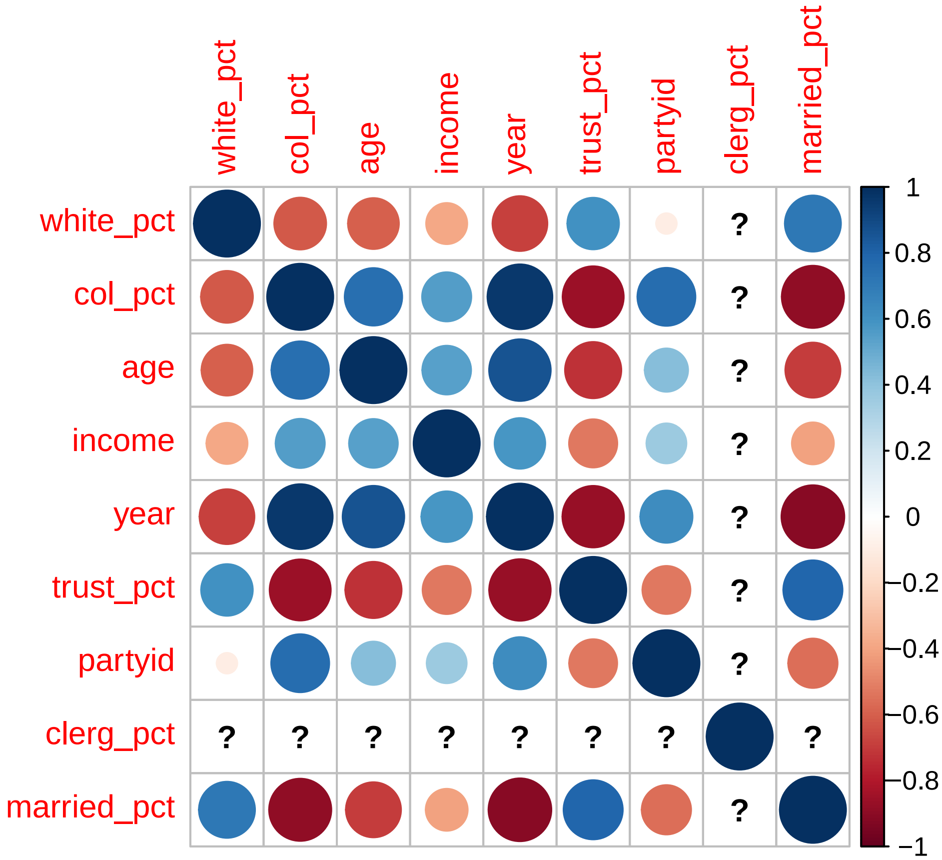
Time Series Lab 3

Kyi Yeung Goh

29/11/2018

1. Create a multivariate time series

```
vars <- c("year", "trust", "conclerg", "sex", "age", "partyid", "marital"
          , "educ", "realinc","race")
sub <- GSS[, vars]

sub <- mutate(sub,
              ntrust = ifelse(trust == 1, 1, 0),
              clerg = ifelse(conclerg == 1,1,0),
              college = ifelse(educ >= 13, 1, 0),
              white = ifelse(race == 1,1,0),
              married = ifelse(marital== 1,1,0),
              income = realinc)


# get means by year
by.year <- aggregate(subset(sub, sel = -year), list(year = sub$year), mean, na.rm = T)


# interpolate for some missing years
# add the extra years
by.year[30:40, "year"] <- c(1979, 1981, 1992, 1995, seq(1997, 2009, 2))
by.year <- arrange(by.year, year)

# make a time series object by.year.ts and interpolate using na.approx
by.year.ts <- ts(by.year)
by.year.ts <- na.approx(by.year.ts)

by.year.ts <- as.data.frame(by.year.ts)
by.year.ts <- mutate(by.year.ts,
                     white_pct = white*100,
                     clerg_pct=clerg*100,
                     trust_pct = ntrust*100,
                     col_pct = college*100,
                     married_pct=married*100)


# correlations
cor.vars <- c("white_pct", "col_pct", "age", "income",
              "year", "trust_pct", "partyid","clerg_pct", "married_pct")
cor.dat <- by.year.ts[, cor.vars]
corrplot(cor(cor.dat))
```

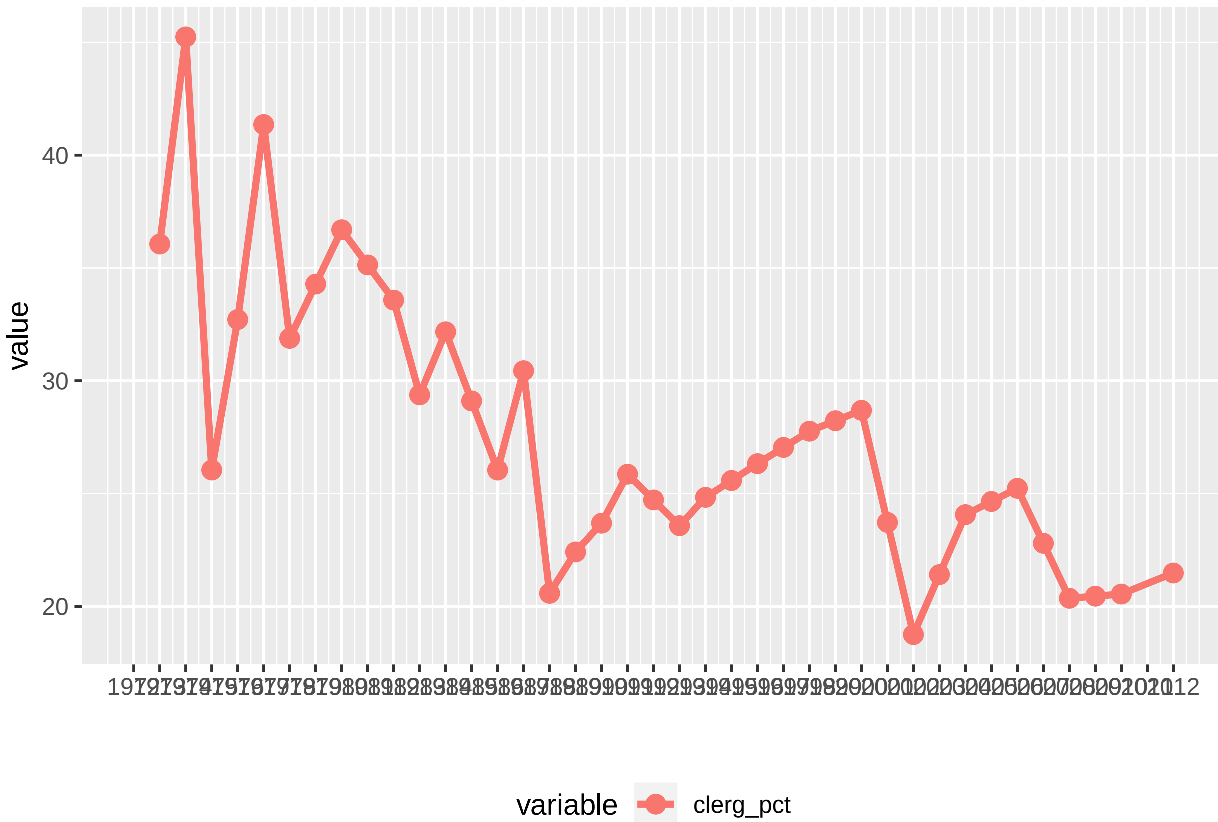


For this, I wanted to look at the relationship between confidence in religious organisations (confclerg) and several predictor variables across time. Here, I looked at trust in members of society (a continuous variable that was recoded to a binary where 1 = TRUST and 0 = NO TRUST), education (where college education of at least 1 year = 1 and below = 0), race (white = 1 and others =0), married (married = 1 and not =0), party identification (continuous variable where 0 is strong democrat and 7 is strong republican), sex and age. The outcome variable, confidence in the clergy, was recoded as a binary variable where strong confidence is 1 and others is 0. For ease of interpretation I recoded all variables except age, sex, income into percentage terms.

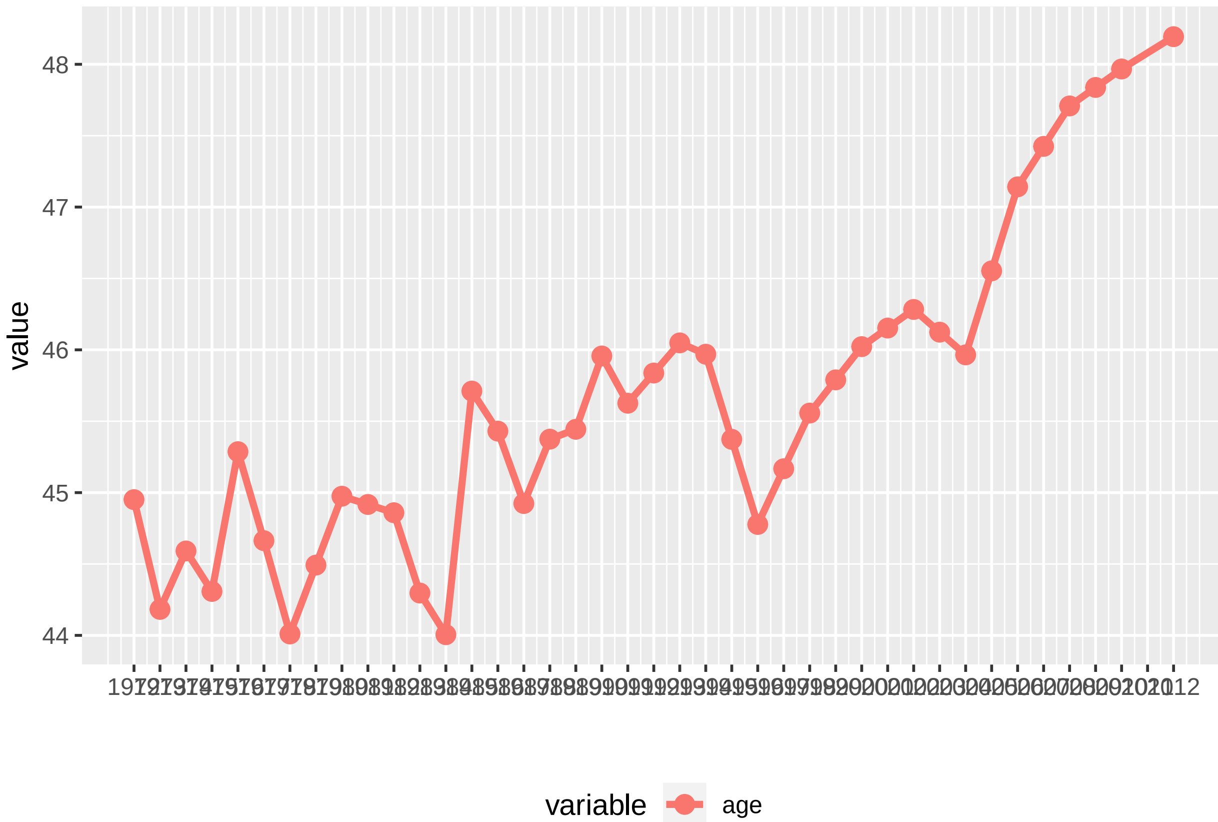
I have also plotted a correlogram in the figure above.

2. Graph the relationships between X and Y.

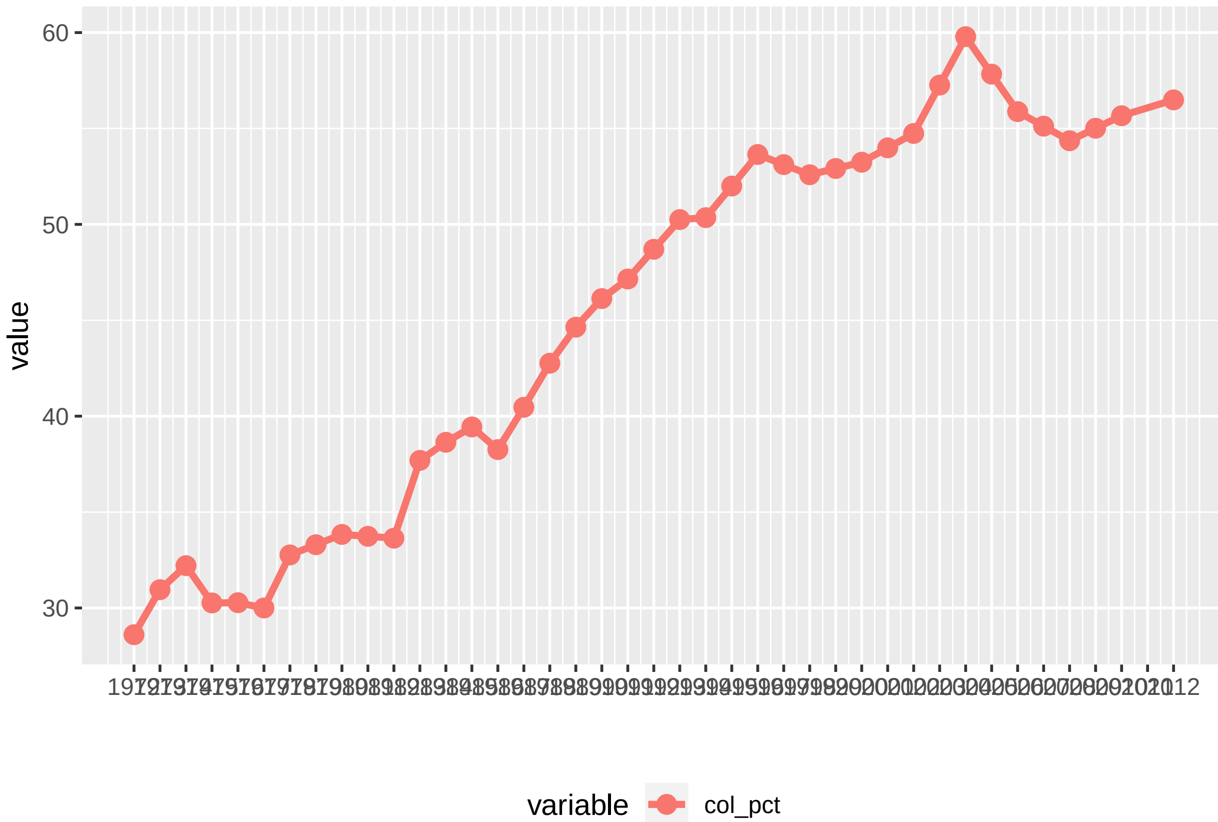
```
keep.vars <- c("white_pct", "col_pct", "age",
               "income", "year", "trust_pct", "partyid",
               "clerg_pct", "married_pct")
plot.dat <- meltMyTS(mv.ts.object = by.year.ts, time.var = "year", keep.vars = keep.vars)

(g_clerg_pct <- ggMyTS(df = plot.dat, varlist =
                          c("clerg_pct"))) #Strong trust in clergy over time
```



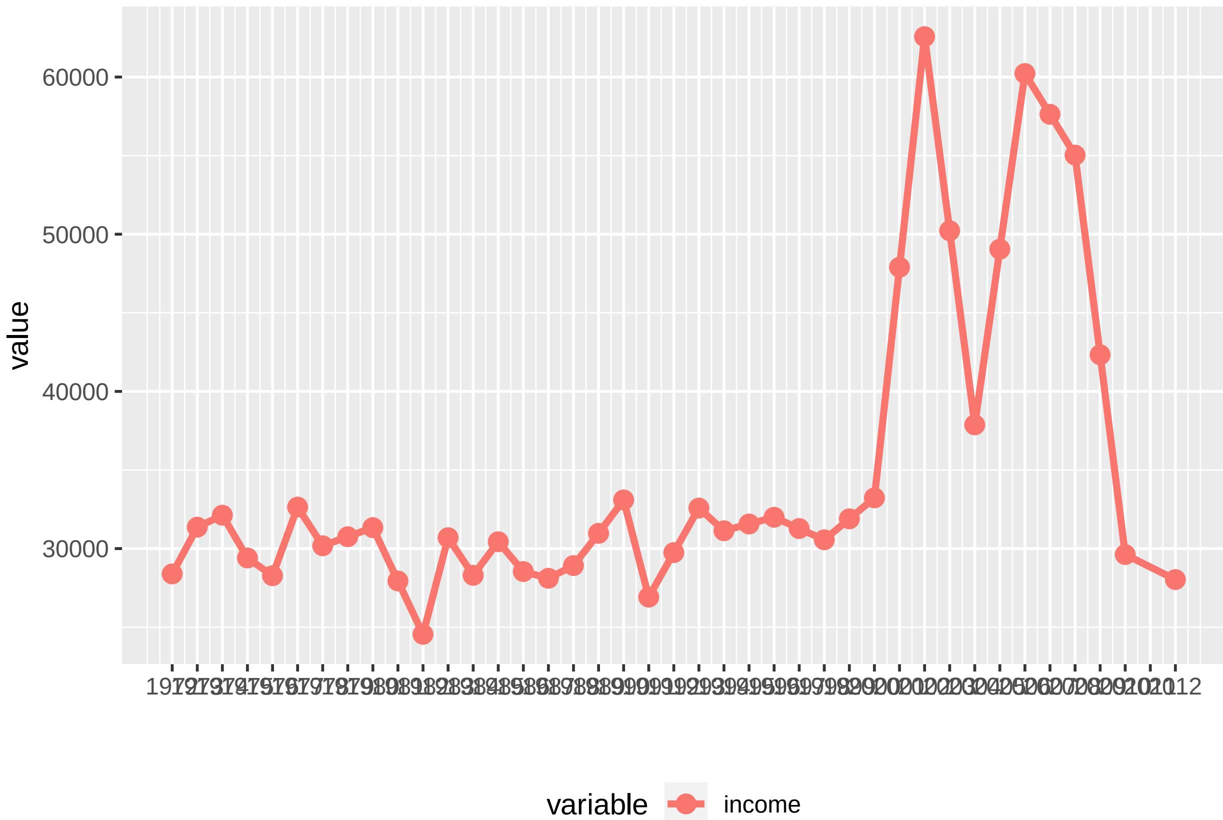

```
(g_age <- ggMyTS(df = plot.dat, varlist =
                    c("age"))) #Average respondent age over time
```



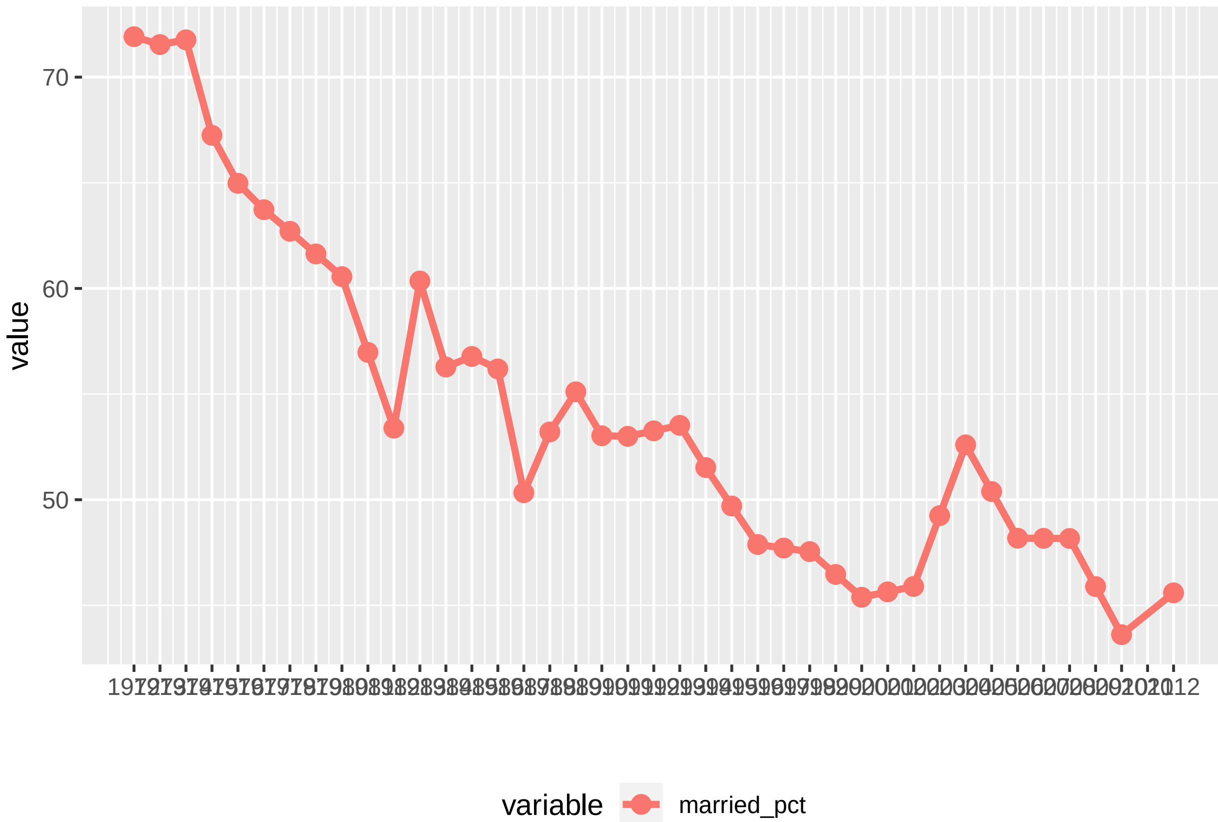

```r
(g_degreelt50_pct <- ggMyTS(df = plot.dat, varlist =
                              c("col_pct"))) #Average percentage of respondents with at least 1 year of
```




```
(g_degreelt50_pct <- ggMyTS(df = plot.dat, varlist =
                              c("income"))) #Average income of respondents over time
```




```
(g_degreelt50_pct <- ggMyTS(df = plot.dat,
                            varlist = c("married_pct"))) #Average percentage of respondents who are mar
```



3. Run a simple time series regression, with one X and no trend. Interpret it.

```
lm.clerg<-lm(clerg_pct ~ col_pct + married_pct + age +
               income + partyid, data=by.year.ts)
summary(lm.clerg, scientific=FALSE)
```

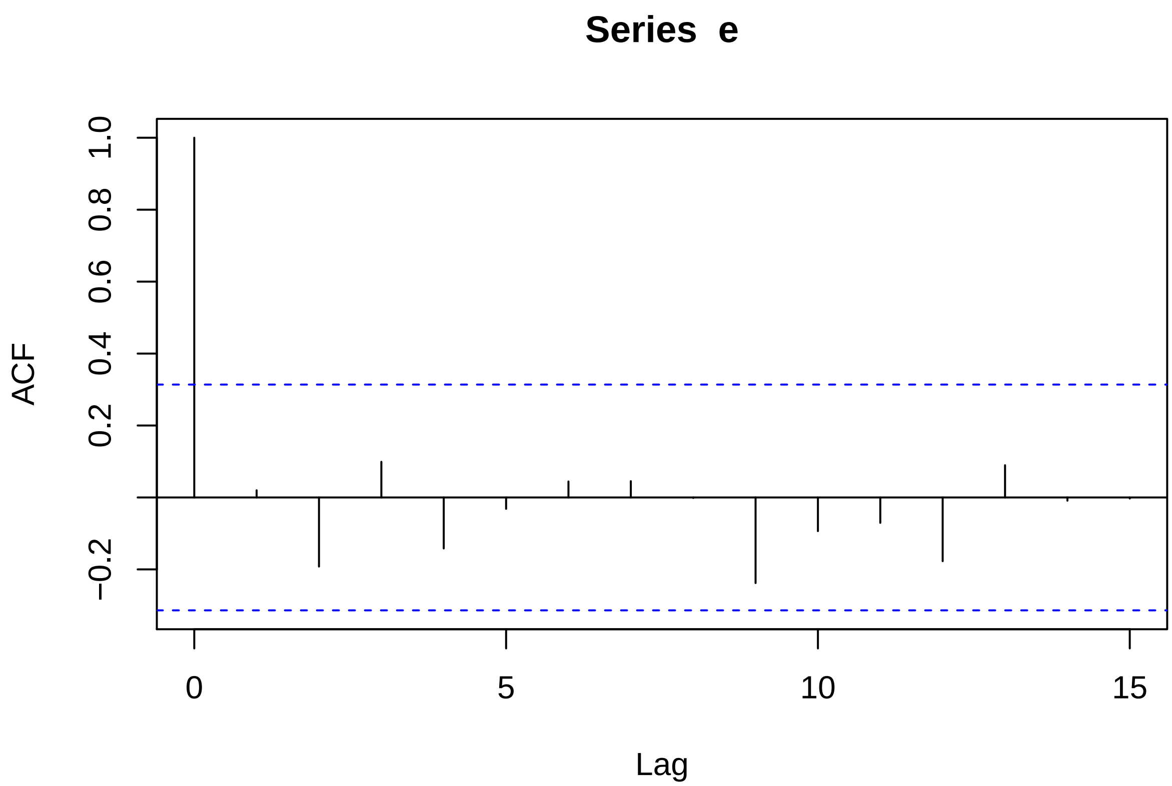
```
##
## Call:
## lm(formula = clerg_pct ~ col_pct + married_pct + age + income +
##     partyid, data = by.year.ts)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.440  -2.047  -0.109   1.384   8.744
##
## Coefficients:
##                  Estimate   Std. Error t value Pr(>|t|)
## (Intercept) 122.30989615  44.81683188   2.729   0.0101 *
## col_pct       0.13594634   0.22201327   0.612   0.5445
## married_pct   0.37976043   0.19378438   1.960   0.0585 .
## age          -1.78909695   0.90081958  -1.986   0.0554 .
## income       -0.00003219   0.00007652  -0.421   0.6767
## partyid     -14.09903125   6.21938907  -2.267   0.0301 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 3.664 on 33 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.6795, Adjusted R-squared:  0.6309
## F-statistic: 13.99 on 5 and 33 DF,  p-value: 0.0000002303
```

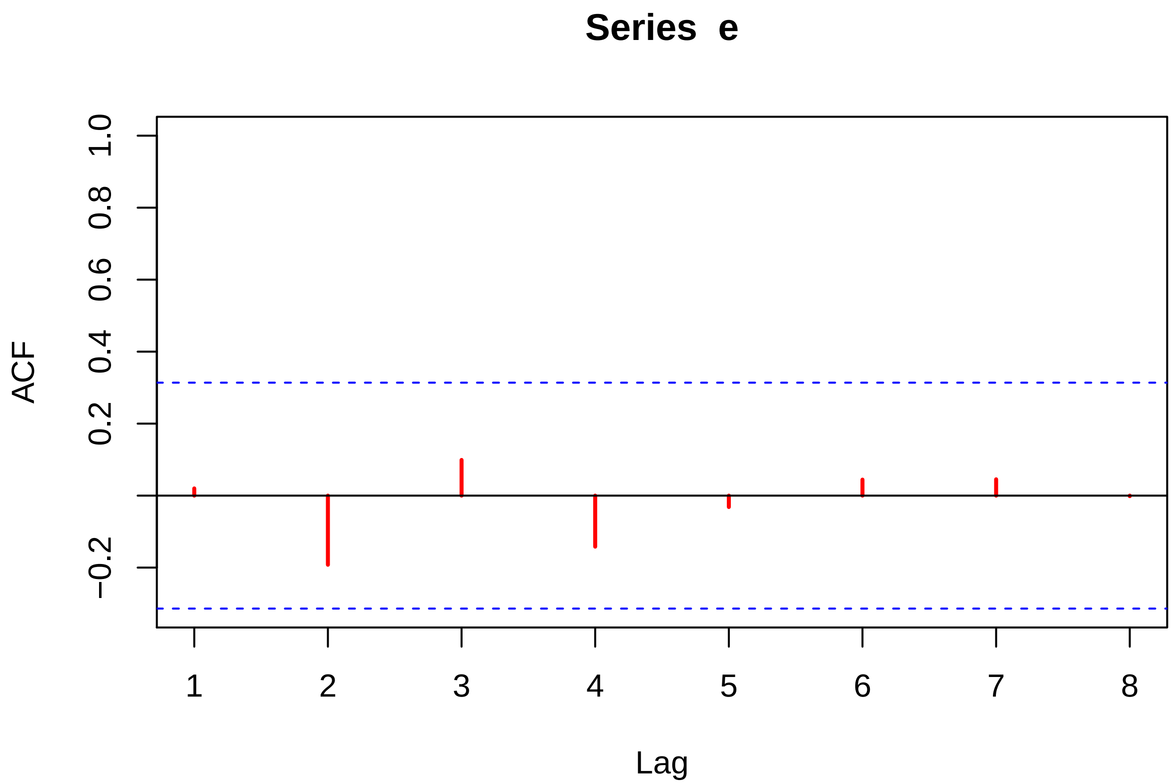
```
bptest(lm.clerg) #no reason to reject null hyppthesis
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm.clerg
## BP = 9.9206, df = 5, p-value = 0.07752
```

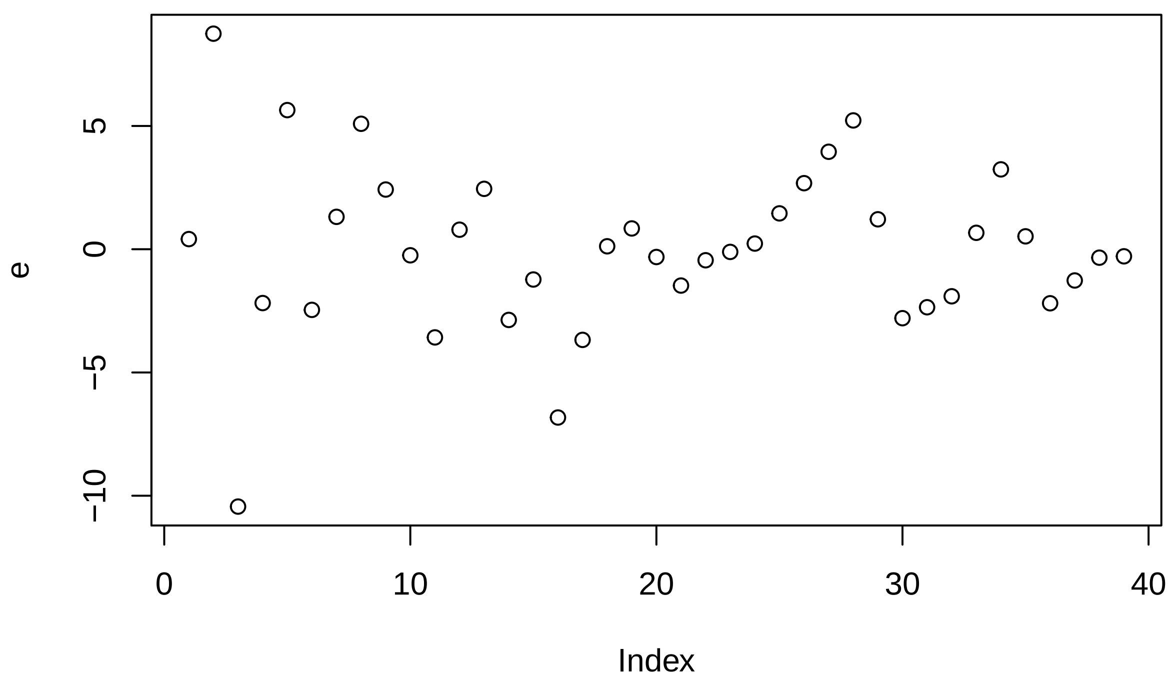
```
# look for autocorrelation in errors
e <- lm.clerg$resid
acf(e)
```




```
acf(e, xlim = c(1,8), col = "red", lwd = 2) # can also customize acf output
```




```
plot(e) # plot residuals over time
```




```
dwtest(lm.clerg) # Durbin-Watson test
```

```
##
##  Durbin-Watson test
##
## data:  lm.clerg
```

```
## DW = 1.9595, p-value = 0.2087
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
bgtest(lm.clerg) # Breusch-Godfrey test
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  lm.clerg
## LM test = 0.016623, df = 1, p-value = 0.8974
```

```r
durbinWatsonTest(lm.clerg, max.lag=2) # Durbin-Watson with more lags
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1      0.01995921      1.959514   0.426
##    2     -0.19210908      2.210768   0.636
##  Alternative hypothesis: rho[lag] != 0
```

A pooled OLS of yearly data suggests that there exists only one significant relationship that between strong confidence in religious institutions and party identification. In this case a one point increase (meaning more Republican) amongst respondents is associated with a 14 percentage point decrease (*) in the percentage of respondents who strongly support the clergy controlling for education, marraige, age and income. There are too other weakly significant relationships. The first is between age and strong support for the clergy with a one year increase in age of respondents resulting in a 1.789 decrease (.) in the percentage of respondents who strongly support the clergy controlling for education, marriage, income and party identification. The other is that between marriage and strong support for the clergy. A one percentage increase in the number of respondent who are married results in a 0.380 increase in the percent of respondents who will strongly support the clergy controlling for all other predictor variables. The Durbin-Watson test suggests that there is indeed autocorrelation and that there needs to be a lag of order 2 to resolve it.

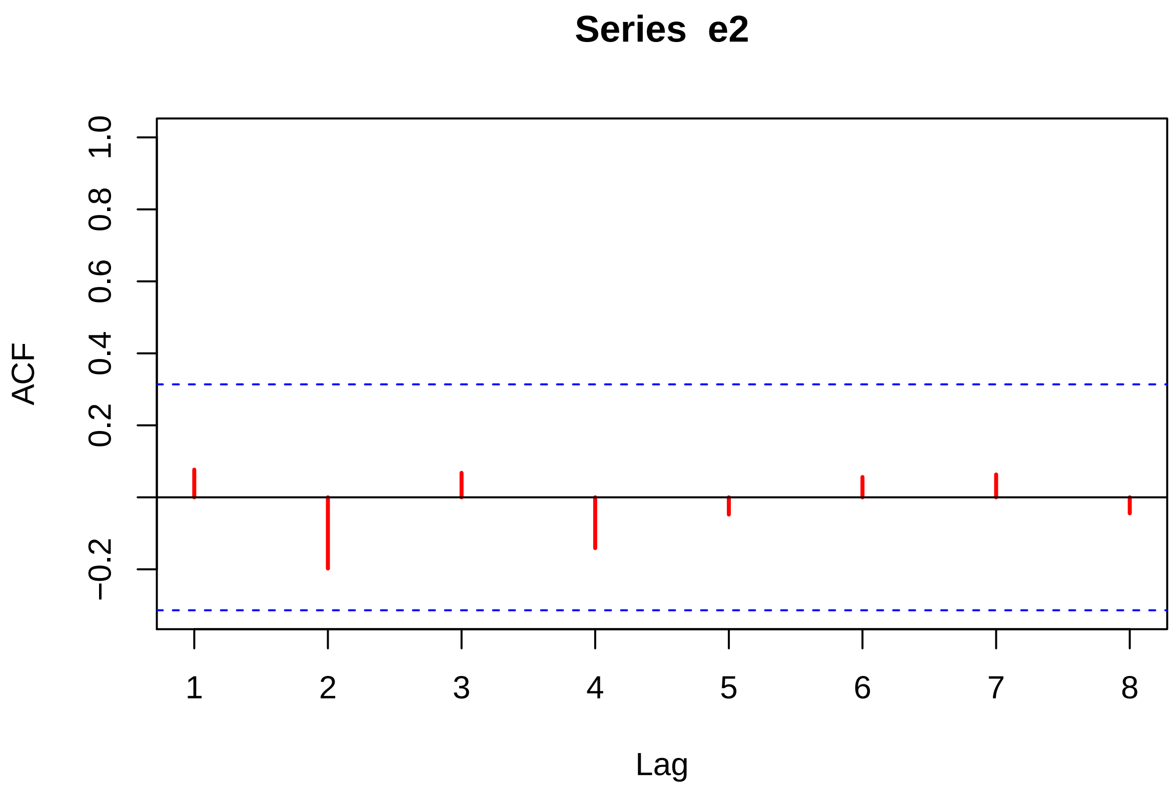
4. Run a time series regression with one X and trend. Interpret it. Perform autocorrelation diagnostics. Explain what you found.

```r
# include year trend
lm.clerg2 <- update(lm.clerg, ~ . + year)
summary(lm.clerg2)
```

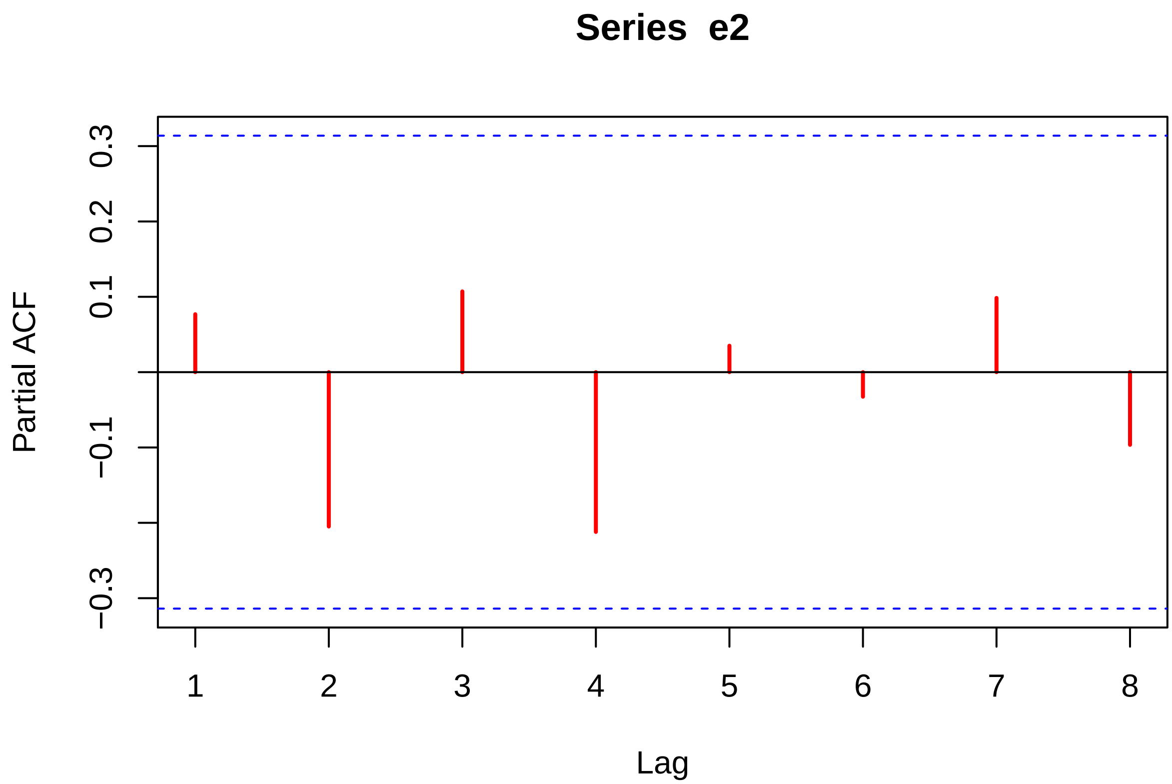
```
##
## Call:
## lm(formula = clerg_pct ~ col_pct + married_pct + age + income +
##     partyid + year, data = by.year.ts)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.3110  -1.9179   0.0481   1.6882   8.3166
##
## Coefficients:
##                 Estimate   Std. Error t value Pr(>|t|)
## (Intercept) 681.21802482 901.97856100   0.755   0.4556
## col_pct       0.39596242   0.47525136   0.833   0.4109
## married_pct   0.32662941   0.21353955   1.530   0.1359
## age          -0.93029656   1.65617540  -0.562   0.5782
## income       -0.00002031   0.00007958  -0.255   0.8002
## partyid     -17.04592808   7.87247615  -2.165   0.0379 *
## year         -0.30094343   0.48505802  -0.620   0.5394
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.698 on 32 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.6833, Adjusted R-squared:  0.6239
## F-statistic: 11.51 on 6 and 32 DF,  p-value: 0.0000007727
```

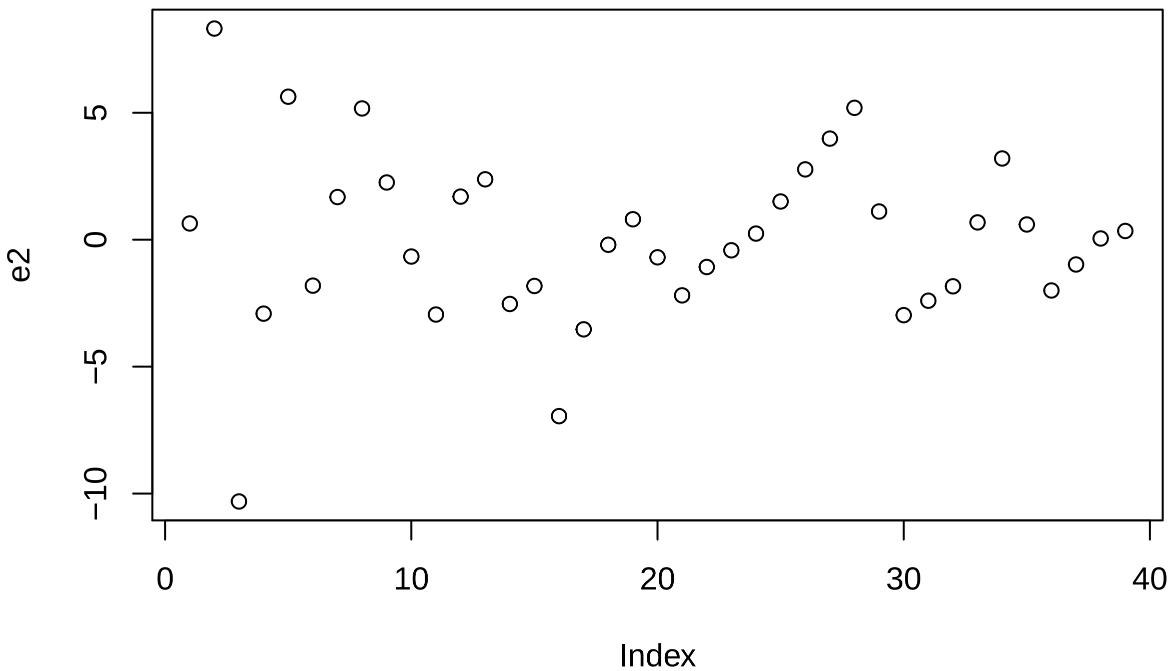
```r
# look for autocorrelation
e2 <- lm.clerg2$resid
acf(e2, xlim = c(1,8), col = "red", lwd = 2)
```




```r
pacf(e2, xlim = c(1,8), col = "red", lwd = 2)
```




```
plot(e2)
```




```
dwtest(lm.clerg2)
```

```
##
##  Durbin-Watson test
##
## data:  lm.clerg2
```

```
## DW = 1.845, p-value = 0.1036
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
bgtest(lm.clerg2)
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  lm.clerg2
## LM test = 0.25801, df = 1, p-value = 0.6115
```

```r
durbinWatsonTest(lm.clerg2, max.lag=3)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1      0.07688129      1.845034   0.188
##    2     -0.19775007      2.236259   0.592
##    3      0.06797605      1.459714   0.122
##  Alternative hypothesis: rho[lag] != 0
```

```r
vif(lm.clerg2)
```

```
##     col_pct married_pct         age      income     partyid        year
##   60.862516    6.891270    9.369651    1.747296    5.181149   85.651443
```

I have included a year trend in this regression and in this case, the year trend has a negative relationship with strong support of religious institutions net of all other predictor variables. It is, however, not statistically significant. Party identification continues to remain statistically significant (*) at the 95% level net of all other factors and any point in time. The Durbin-Watson statistic suggests the presence of autocorrelation and that an order of 2 lags is needed to resolve it. Unsurprisingly, VIF (indicating multicollinearity or the inflation of R2 caused by predictors explaining same variance) is particularly high for percent of college educated persons and year trend. This suggests that the model needs to be simplified or re-specified.

5. Consider running a time series regression with many Xs and trend. Interpret that. Check VIF.

```r
# add some more predictors
lm.clerg3 <- update(lm.clerg2, ~ . + trust_pct + white_pct )
summary(lm.clerg3)
```

```
##
## Call:
## lm(formula = clerg_pct ~ col_pct + married_pct + age + income +
##     partyid + year + trust_pct + white_pct, data = by.year.ts)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.1579 -1.7374  0.1057  1.3951  8.0812
##
## Coefficients:
##                 Estimate    Std. Error t value Pr(>|t|)
## (Intercept) 760.641343897 903.552687850   0.842   0.4065
## col_pct       0.645174354   0.532026106   1.213   0.2347
## married_pct   0.331923389   0.231454140   1.434   0.1619
## age          -0.510893822   1.642431508  -0.311   0.7579
## income       -0.000006529   0.000078486  -0.083   0.9343
```

```
## partyid     -19.907512939  10.574614285  -1.883   0.0695 .
## year         -0.361145757   0.486958372  -0.742   0.4641
## trust_pct     0.578823168   0.284960766   2.031   0.0512 .
## white_pct    -0.059426939   0.224758457  -0.264   0.7933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.567 on 30 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.7238, Adjusted R-squared:  0.6501
## F-statistic: 9.826 on 8 and 30 DF,  p-value: 0.000001383
```

```r
vif(lm.clerg3) # variance inflation factor
```

```
##     col_pct married_pct         age      income     partyid        year
##   81.984616    8.702336    9.904866    1.826767   10.048373   92.788529
##   trust_pct   white_pct
##    4.194179    4.773033
```

```r
durbinWatsonTest(lm.clerg3, max.lag=2)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1       0.1378650      1.721446   0.084
##    2      -0.1561418      2.138368   0.868
##  Alternative hypothesis: rho[lag] != 0
```

This regression adds additional predictor variables *white_pct* (indicating the percentage of respondents who were white) and *trust_pct* (percentage of respondents who trusted people generally) to the model. With more predictor variables added into the model, partyid reduces in statistical significance (.) and is now only at the 90% level. Percentage of trust in respondents is a very weakly significant predictor at the 90% level. Here, a 1 percent increase in trust leads to a 0.579 percent increase in total number of respondents who have strong confidence in the clergy holding all other factors and time constant. The model, while explaining around 65% of total variance, is hampered by the existence of very high VIF scores on college and the year variables. The D-W statistic, meanwhile, suggests a lag of order 2 in order to resolve autocorrelation.

6. Run a first differenced time series regression. Interpret that.

Use the first differences

```r
by.yearFD <- mutate(data.frame(by.year.ts),
                       age = firstD(age),
                       col_pct = firstD(col_pct),
                       income = firstD(income),
                       white_pct=firstD(white_pct),
                       married_pct=firstD(married_pct),
                       partyid=firstD(partyid),
                       clerg_pct=firstD(clerg_pct),
                       trust_pct = firstD(trust_pct),
                       year = year)

lm.clerg4 <- update(lm.clerg3, data = by.yearFD)
summary(lm.clerg4)
```

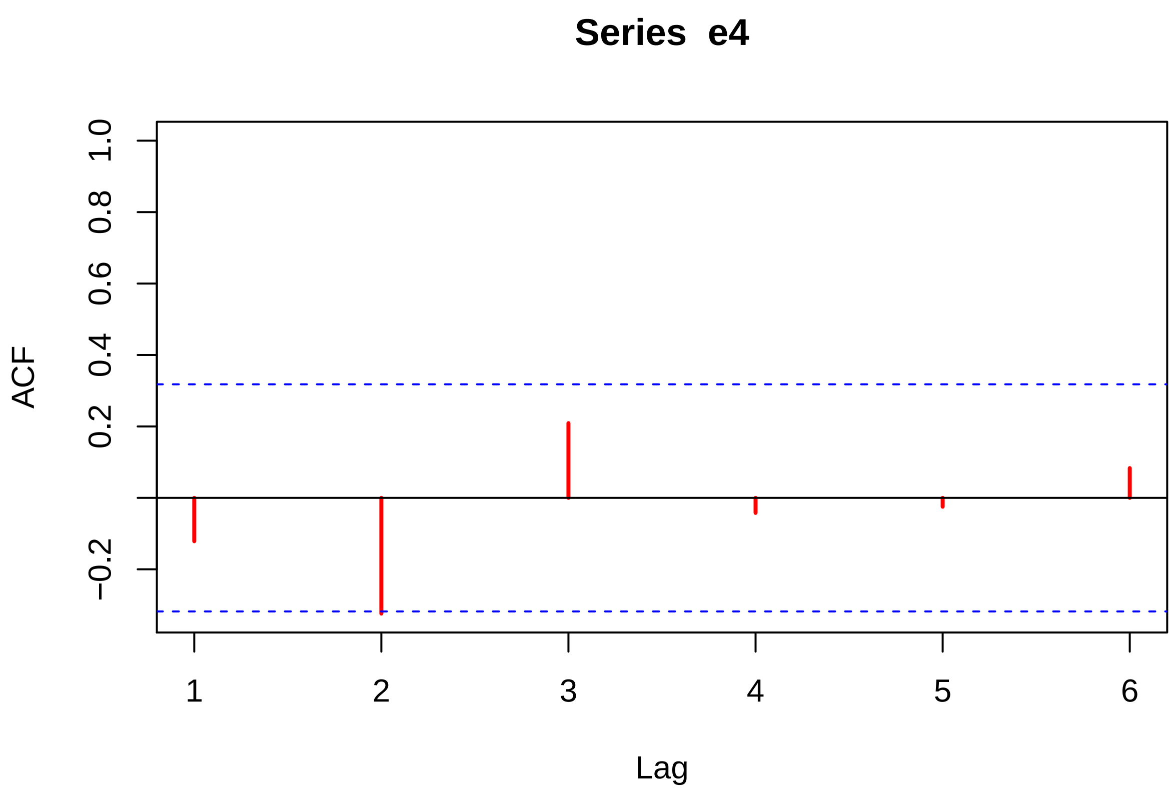
```
##
## Call:
```

```
## lm(formula = clerg_pct ~ col_pct + married_pct + age + income +
##     partyid + year + trust_pct + white_pct, data = by.yearFD)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.0682  -2.1148  -0.7406   2.5996  10.8977
##
## Coefficients:
##                 Estimate   Std. Error t value Pr(>|t|)
## (Intercept)  99.92413499 145.49705210   0.687   0.4977
## col_pct       0.74031996   0.91030878   0.813   0.4227
## married_pct   1.17695132   0.58720496   2.004   0.0545 .
## age           2.48025116   1.64385519   1.509   0.1422
## income        0.00002626   0.00012456   0.211   0.8345
## partyid     -14.22710788  14.35969485  -0.991   0.3300
## year         -0.05028077   0.07282390  -0.690   0.4954
## trust_pct     0.64860664   0.25667001   2.527   0.0172 *
## white_pct    -0.52872023   0.33126819  -1.596   0.1213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.271 on 29 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.4294, Adjusted R-squared:  0.272
## F-statistic: 2.728 on 8 and 29 DF,  p-value: 0.02251
```

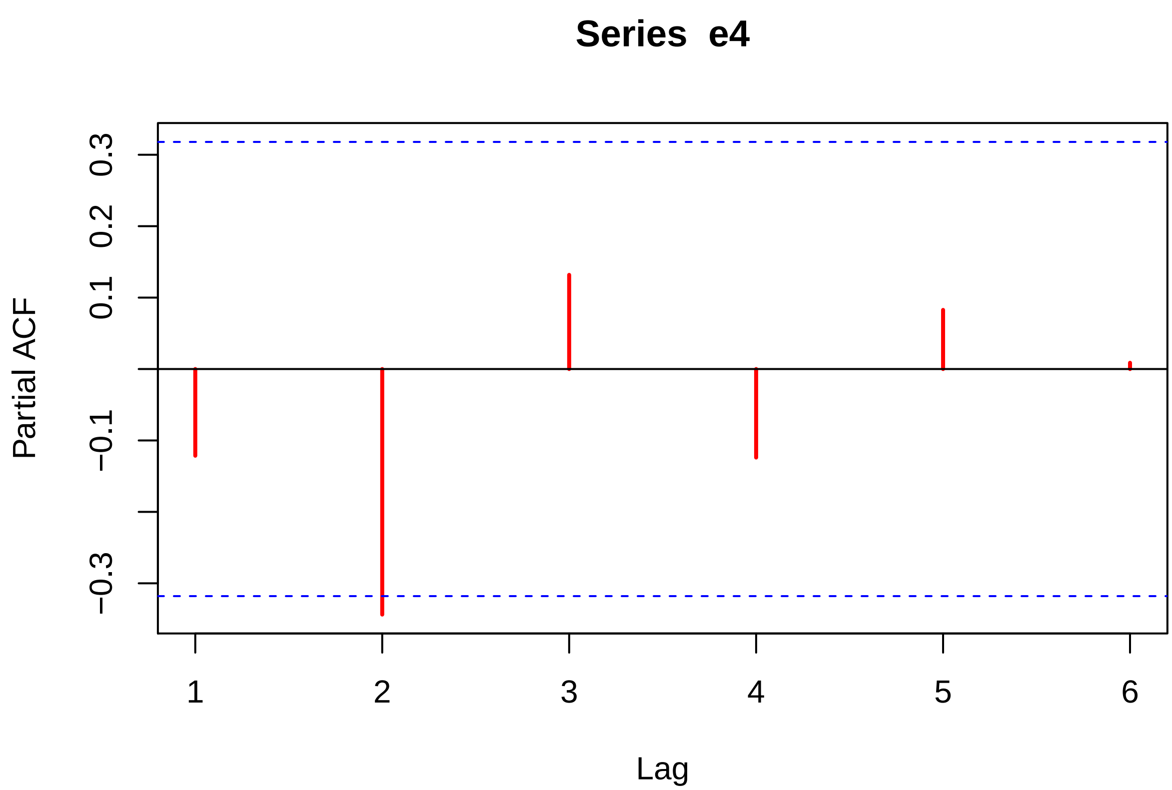
```r
vif(lm.clerg4)
```

```
##     col_pct married_pct         age      income     partyid        year
##    3.089577    4.054956    1.265429    1.279786    5.737554    1.339592
##   trust_pct   white_pct
##    1.126534    6.863363
```

```r
e4 <- lm.clerg4$resid
acf(e4, xlim = c(1,6), col = "red", lwd = 2)
```




```r
pacf(e4, xlim = c(1,6), col = "red", lwd = 2)
```




```r
auto.arima(e4, trace=TRUE)
```

```
##
```

```
##  ARIMA(2,0,2) with non-zero mean : 208.9372
##  ARIMA(0,0,0) with non-zero mean : 212.2498
##  ARIMA(1,0,0) with non-zero mean : 213.9945
##  ARIMA(0,0,1) with non-zero mean : 213.1956
##  ARIMA(0,0,0) with zero mean     : 210.018
##  ARIMA(1,0,2) with non-zero mean : 209.0666
##  ARIMA(3,0,2) with non-zero mean : 211.9607
##  ARIMA(2,0,1) with non-zero mean : 206.8396
##  ARIMA(2,0,1) with zero mean     : 204.1773
##  ARIMA(1,0,1) with zero mean     : 209.9319
##  ARIMA(3,0,1) with zero mean     : 206.4753
##  ARIMA(2,0,0) with zero mean     : 207.8064
##  ARIMA(2,0,2) with zero mean     : 206.1069
##  ARIMA(1,0,0) with zero mean     : 211.6327
##  ARIMA(3,0,2) with zero mean     : 208.9414
##
##  Best model: ARIMA(2,0,1) with zero mean
```

```
## Series: e4
## ARIMA(2,0,1) with zero mean
##
## Coefficients:
##           ar1      ar2     ma1
##       -0.6146  -0.5559  0.7230
## s.e.   0.2214   0.1761  0.2017
##
## sigma^2 estimated as 10.39:  log likelihood=-97.48
## AIC=202.97   AICc=204.18   BIC=209.52
```

I then ran a first difference model, a potential solution for the VIF in the previous models. In this case, there are only 2 statistically significant predictors of strong support for religious institutions. For each percentage change in the percentage of married respondents, the percentage of respondents who have strong confidence in religious institutions increases by 1.177 net of all other Xs and at any point in time. This relationship is significant at the 90% level. Meanwhile, for each percentage change in the respondents who trust persons results in a 0.6486 percentage point positive change in the percentage points of persons who have strong confidence in religious institutions net of all other Xs and time. It is crucial to point out that the adj-R2 has dipped significantly, only explaining around 27% of total variance. This gels well with the fact that the VIFs of all predictor variables are now well below 10, indicating a resolution of the problem of multicollinearity raised in previous parts.

The auto-arima suggests two lags along with 1 lagged average of e-s in this model. I will re-run this model with the suggested ARIMA in part 9.

7. Check your variables for unit roots. Do some tests. Interpret them.

```r
by.yearFD <- mutate(data.frame(by.year.ts),
                       trust_pct = firstD(trust_pct),
                       age = firstD(age),
                       col_pct = firstD(col_pct),
                       income = firstD(income),
                       white_pct=firstD(white_pct),
                       married_pct=firstD(married_pct),
                       partyid=firstD(partyid),
                       clerg_pct=firstD(clerg_pct),
```

```
                       year = year)

lm.clerg4 <- update(lm.clerg3, data = by.yearFD)

adfTest(by.year.ts[,"clerg_pct"], lags = 0, type="ct")
```

```
## Warning in adfTest(by.year.ts[, "clerg_pct"], lags = 0, type = "ct"): p-
## value smaller than printed p-value

##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 0
##   STATISTIC:
##     Dickey-Fuller: -5.0061
##   P VALUE:
##     0.01
##
## Description:
##  Thu Dec  6 01:50:58 2018 by user:
```

```
adfTest(by.year.ts[,"clerg_pct"], lags = 4, type="ct")
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 4
##   STATISTIC:
##     Dickey-Fuller: -2.0781
##   P VALUE:
##     0.5428
##
## Description:
##  Thu Dec  6 01:50:58 2018 by user:
```

```
coeftest(lm.clerg4, vcov = NeweyWest(lm.clerg3, lag = 2))
```

```
##
## t test of coefficients:
##
##                 Estimate    Std. Error t value     Pr(>|t|)
## (Intercept)  99.924134985 373.823078326  0.2673     0.791128
## col_pct       0.740319956   0.339599587  2.1800     0.037521 *
## married_pct   1.176951325   0.183341164  6.4195 0.0000005074 ***
## age           2.480251155   0.898663211  2.7599     0.009915 **
## income        0.000026264   0.000039311  0.6681     0.509353
## partyid     -14.227107883   5.826990099 -2.4416     0.020958 *
## year         -0.050280771   0.204698526 -0.2456     0.807696
## trust_pct     0.648606636   0.183592910  3.5329     0.001398 **
## white_pct    -0.528720231   0.174481111 -3.0302     0.005099 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With 0 lags, there is no reason to reject the null of there being unit root. Even with 4 lags, there continues to be no reason to do so. There is thus a unit root. This suggests that first-differencing is a very compelling way in which to resolve this problem.

8. Perform an Automatic ARIMA on the residuals from one of your earlier models. Tell me what it says.

```r
by.yearFD <- mutate(data.frame(by.year.ts),
                    trust_pct = firstD(trust_pct),
                    age = firstD(age),
                    col_pct = firstD(col_pct),
                    income = firstD(income),
                    white_pct=firstD(white_pct),
                    married_pct=firstD(married_pct),
                    partyid=firstD(partyid),
                    clerg_pct=firstD(clerg_pct),
                    year = year)

lm.clerg4 <- update(lm.clerg3, data = by.yearFD)
auto.arima(e4, trace=TRUE)
```

```
##
##  ARIMA(2,0,2) with non-zero mean : 208.9372
##  ARIMA(0,0,0) with non-zero mean : 212.2498
##  ARIMA(1,0,0) with non-zero mean : 213.9945
##  ARIMA(0,0,1) with non-zero mean : 213.1956
##  ARIMA(0,0,0) with zero mean     : 210.018
##  ARIMA(1,0,2) with non-zero mean : 209.0666
##  ARIMA(3,0,2) with non-zero mean : 211.9607
##  ARIMA(2,0,1) with non-zero mean : 206.8396
##  ARIMA(2,0,1) with zero mean     : 204.1773
##  ARIMA(1,0,1) with zero mean     : 209.9319
##  ARIMA(3,0,1) with zero mean     : 206.4753
##  ARIMA(2,0,0) with zero mean     : 207.8064
##  ARIMA(2,0,2) with zero mean     : 206.1069
##  ARIMA(1,0,0) with zero mean     : 211.6327
##  ARIMA(3,0,2) with zero mean     : 208.9414
##
##  Best model: ARIMA(2,0,1) with zero mean
```

```
## Series: e4
## ARIMA(2,0,1) with zero mean
##
## Coefficients:
##           ar1      ar2     ma1
##       -0.6146  -0.5559  0.7230
## s.e.   0.2214   0.1761  0.2017
##
## sigma^2 estimated as 10.39:  log likelihood=-97.48
## AIC=202.97   AICc=204.18   BIC=209.52
```

The model (from FD model using error from that) that removes all sources of systematic error in the data, whether from non-stationarity, or autoregressive tendencies and/or moving average processes in the errors is a first difference model that contains 2 lags with 1 lagged average of the error term. This means that the final model involves two lags of the error term, 1 difference (in the original FD model) and 1 lagged average of the error term. The interpretation of this model will be done in part 9 below.

9. Run an ARIMA that follows from Step 7. Interpret that, too.

```r
by.yearFD <- mutate(data.frame(by.year.ts),
                        trust_pct = firstD(trust_pct),
                        age = firstD(age),
                        col_pct = firstD(col_pct),
                        income = firstD(income),
                        white_pct=firstD(white_pct),
                        married_pct=firstD(married_pct),
                        partyid=firstD(partyid),
                        clerg_pct=firstD(clerg_pct),
                        year = year)

lm.clerg4 <- update(lm.clerg3, data = by.yearFD)
xvars.fat <- by.yearFD[,c("trust_pct", "age", "col_pct","income",
                          "white_pct", "married_pct","partyid", "year")]

# ARIMA(2,0,1) = OLS
arima.001 <- arima(by.yearFD[,"clerg_pct"], order = c(2,0,1), xreg = xvars.fat)
arima.001
```

```
##
## Call:
## arima(x = by.yearFD[, "clerg_pct"], order = c(2, 0, 1), xreg = xvars.fat)
##
## Coefficients:
##           ar1      ar2     ma1  intercept  trust_pct     age  col_pct
##       -0.8764  -0.7562  1.0000    44.5190     0.1765  1.3934   0.6031
## s.e.   0.1718   0.1493  0.0876    75.5143     0.1903  1.1672   0.5213
##       income  white_pct  married_pct  partyid     year
##       0.0000    -0.5006       0.7180  -5.7701  -0.0226
## s.e.  0.0003     0.1979       0.2825  10.9982   0.0378
##
## sigma^2 estimated as 6.978:  log likelihood = -93.51,  aic = 213.02
```

```r
Box.test(resid(arima.001), lag = 13, type = c("Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Ljung test
##
## data:  resid(arima.001)
## X-squared = 10.3, df = 13, p-value = 0.6692
```

Notice that this is essentially a first difference model with two lags and one lagged average of the error. The interpretation is thus as follows. The two significant predictors in this model is white_pct and married_pct. For white_pct, a one percent change in the number of white persons in the wave results in a 0.7180 percentage point change in the percentage of the sample who are very confident of their religious leaders net of all other Xs, the time trend, the first and second lag of clerg_pct and the previous lag of the error term. Meanwhile, a one percent change in the number of married respondents in the wave results in a 0.5006% negative change in the percentage of the sample who are very confident of their religious leaders net of all other Xs, the time trend, the first and second lag of clerg_pct and the previous lag of the error term.The box-ljung test (parameter of 16 which is half of the total number of points -2) is not significant and we cannot rejec the null that the residuals is simply white noise, meaning autocorrelation appears to not be a problem.