Collaborative Research Project - Assignment 3

N. Alvarado and G. Tarriba

13 November 2015

Assignment 3 - Gathering, cleaning and analyzing data for our collaborative research project

```
## Loading required package: Formula
##
## Please cite as:
##
##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```
##       exp              wks         bluecol          ind          south
##  Min.   : 1.00   Min.   : 5.00   no :2036   Min.   :0.0000   no :2956
##  1st Qu.:11.00   1st Qu.:46.00   yes:2129   1st Qu.:0.0000   yes:1209
##  Median :18.00   Median :48.00              Median :0.0000
##  Mean   :19.85   Mean   :46.81              Mean   :0.3954
##  3rd Qu.:29.00   3rd Qu.:50.00              3rd Qu.:1.0000
##  Max.   :51.00   Max.   :52.00              Max.   :1.0000
##   smsa      married        sex       union            ed        black
##  no :1442   no : 773   female: 469   no :2649   Min.   : 4.00   no :3864
##  yes:2723   yes:3392   male  :3696   yes:1516   1st Qu.:12.00   yes: 301
##                                                 Median :12.00
##                                                 Mean   :12.85
##                                                 3rd Qu.:16.00
##                                                 Max.   :17.00
##      lwage
##  Min.   :4.605
##  1st Qu.:6.395
##  Median :6.685
##  Mean   :6.676
##  3rd Qu.:6.953
##  Max.   :8.537
```

```
##
## Call:
## lm(formula = lwage ~ ed, data = Wages)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.92996 -0.26863  0.00931  0.28453  1.83076
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.838779   0.030997  188.37   <2e-16 ***
## ed          0.065204   0.002358   27.65   <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4243 on 4163 degrees of freedom
## Multiple R-squared:  0.1552, Adjusted R-squared:  0.155
## F-statistic: 764.5 on 1 and 4163 DF,  p-value: < 2.2e-16
```

```
##                 2.5 %     97.5 %
## (Intercept) 5.7780088 5.89954938
## ed          0.0605805 0.06982708
```

```
##
## <table style="text-align:center"><caption><strong>OLS regression of the Percentage of Wages Variatio
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## <tr><td></td><td colspan="2" style="border-bottom: 1px solid black"></td></tr>
## <tr><td style="text-align:left"></td><td colspan="2">lwage</td></tr>
## <tr><td style="text-align:left"></td><td>(1)</td><td>(2)</td></tr>
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## <tr><td style="text-align:left"></td><td>(0.002)</td><td>(0.002)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>
## <tr><td style="text-align:left">exp</td><td></td><td>0.01<sup>***</sup></td></tr>
## <tr><td style="text-align:left"></td><td></td><td>(0.001)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>
## <tr><td style="text-align:left">Constant</td><td>5.84<sup>***</sup></td><td>5.44<sup>***</sup></td><
## <tr><td style="text-align:left"></td><td>(0.03)</td><td>(0.03)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## <tr><td style="text-align:left">R<sup>2</sup></td><td>0.16</td><td>0.25</td></tr>
## <tr><td style="text-align:left">Adjusted R<sup>2</sup></td><td>0.15</td><td>0.25</td></tr>
## <tr><td style="text-align:left">Residual Std. Error</td><td>0.42 (df = 4163)</td><td>0.40 (df = 4162
## <tr><td style="text-align:left">F Statistic</td><td>764.53<sup>***</sup> (df = 1; 4163)</td><td>681.5
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## </table>
```