ps2; **due in 2 weeks**

[version: Tuesday 16th January, 2018 16:18]

Professor X hypothetizes that wage is affected by work experience (estimated): the hypothesis is that the more experience, the higher the wages. We know that experience is not everything and other variables matter as well. As a bonus (extra credit), you may also take into account other variables.

Data are here:
use https://sites.google.com/site/adamokuliczkozaryn/adv_reg/wages.dta**, clear** .

1. produce some relevant descriptive statistics (remember: never run regressions without descriptive stats!)–submit interesting/relevant results, not all the permutations you can think of (probably want to do it in Stata, but can do it by hand too! or both!)–as always don't forget to interpret
2. then for simplicity retain just 4 obs, so that we can see mechanics easier and also calculate by hand–to make sure everyone has the same data, run the following in stata:

```
set seed 1234567
sample 4, count
```

   so that the observations on main variables are:

wage	exp
18.163	14
12.57	12
13.649	11
12	13

3. calculate by hand: regression coefficients, and t-statistics (covered next week); also for at least first two obs calculate predicted values for an average person (mean of X); calculate Rsq
4. check your estimates with Stata (remember: Stata is always right):

```
keep wage exp
reg wage exp
predict yhat
predict resid, r
l
```

---

general directions (always the same):

- if asked to do sth by hand and you subset your data to few obs to do that, say 4 obs, do all the other stuff like descriptive stats for the full sample–it is way more interesting!
- when doing things by hand, show all the work, all the steps
- make it as easy on yourself as possible: round up numbers! simplify!
- if you calculate any meaningful number, say slope coefficient or t-stat, always interpret it!
- preferably use txt or pdf formats; doc(x) often messes up formatting
- do not submit more than 10 pages of the output (12pt font, single spaced)
- submit into the Sakai's dropbox; ps are due by the beginning of the next class unless indicated otherwise, eg "due in 2 weeks"; late ps are not accepted; if your writing is legible you can write by hand and then take a picture and submit that
- we are on the way to developing the final project with these ps: as we progress, your ps should start resembling a coherent and logical project where you use regression to answer interesting questions–say in few sentences why are you doing what you are doing–that is, answer the "so what question": what's the goal of all that, why are you doing this? you need a compelling justification for what you are doing; be brief, say couple sentences
- always submit dofile if you calculate anything in stata; because you are only submitting code (do not submit any datasets), it must load data from Internet–just put your data onto your own website, wordpress, google drive, etc
- always, cite your data (at minimum full name and url (if applicable))