Practice Questions:

Consider the following data on education and number of children collected from women wages 35-54 from the General Social Survey from 1972-1984. Assume the sample is restricted to White women and Black women.

```
. sum year educ meduc feduc age kids black east northcen west farm othrural town smcity
```

Variable	Obs	Mean	Std. Dev.	Min	Max
year	1,129	78.13995	4.091798	72	84
educ	1,129	12.69088	2.640236	0	20
meduc	1,129	9.131975	4.016956	0	20
feduc	1,129	9.715678	3.49515	0	20
age	1,129	43.4845	5.836421	35	54
kids	1,129	2.743136	1.653899	0	7
black	1,129	.085031	.2790514	0	1
east	1,129	.2488928	.4325632	0	1
northcen	1,129	.319752	.4665871	0	1
west	1,129	.1080602	.310594	0	1
farm	1,129	.1984057	.398976	0	1
othrural	1,129	.1018601	.3025982	0	1
town	1,129	.3170948	.4655509	0	1
smcity	1,129	.125775	.3317426	0	1

```
regress kids educ meduc feduc age black east northcen west farm othrural town smcity
```

Source	SS	df	MS	Number of obs	=	1,129
				F(12, 1116)	=	9.34
Model	281.629582	12	23.4691319	Prob > F	=	0.0000
Residual	2803.87972	1,116	2.51243702	R-squared	=	0.0913
				Adj R-squared	=	0.0815
Total	3085.5093	1,128	2.73538059	Root MSE	=	1.5851

kids	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	-.1342054	.0210637	-6.37	0.000	-.1755343	-.0928766
meduc	-.0025845	.0160468	-0.16	0.872	-.0340697	.0289008
feduc	-.0184516	.018414	-1.00	0.317	-.0545816	.0176784
age	.0231654	.0082819	2.80	0.005	.0069155	.0394153
black	.9317919	.1752166	5.32	0.000	.5880008	1.275583
east	.2432724	.1350129	1.80	0.072	-.0216353	.5081801
northcen	.3875773	.1236687	3.13	0.002	.1449279	.6302266
west	.2359492	.170209	1.39	0.166	-.0980166	.569915
farm	-.0760507	.1509348	-0.50	0.614	-.3721987	.2200973
othrural	-.1884569	.179949	-1.05	0.295	-.5415334	.1646195
town	.0835488	.1270928	0.66	0.511	-.1658189	.3329165
smcity	.1964418	.1630284	1.20	0.228	-.1234349	.5163185
_cons	3.335735	.4697198	7.10	0.000	2.414102	4.257369

1. Interpret the coefficient on `educ`.

One year additional education is associated w/ 0.13 fewer children

2. Interpret the coefficient on `black`.

Black women have on average 0.93 more children than White women, holding all else constant

3. Test the null hypothesis that the number of children is equal between White women and Black women.

$H_0: \beta_{black} = 0$

$H_1: \beta_{black} \neq 0$

$t = \frac{0.93 - 0}{0.175} = 5.32 \Rightarrow$ reject at 1% level

if $|t| > 2.58 \Rightarrow$ reject H_0 at all conventional levels

4. Test the null hypothesis that Black women have at least one more child than White women.

$H_0: \beta_{black} \geq 1$

$H_1: \beta_{black} < 1$

$t = \frac{0.913 - 1}{0.175} = -0.4$ (assume $\alpha = 0.05$)

reject if $t < -c_{0.05} = -1.645$

$-0.4 > -1.645 \Rightarrow$ do not reject H_0

5. The correlation coefficient between mother's years of education (`meduc`) and father's years of education (`feduc`) is 0.63. Does this mean that we are likely to experience a problem with multi-collinearity? Explain

No. Perfect multicollinearity only if meduc is a linear function of feduc. Also, 0.63 not that high

6. List the four Gauss-Markov assumptions. For each one, state whether the assumption is likely to hold in this context, and explain why.

(assume we are considering β_{educ})

1) zero conditional mean: likely violated. for example, marital status affects education + number of kids

2) x, y iid → likely ok, repeated cross-section from representative sample

3) no large outliers: probably ok, educ + kids mechanically capped

4) homoskedasticity : hard to say, but if higher variance among less educated women, then violated.

7. Why do the R-squared and adjusted R-squared differ? Explain which you prefer and why.

adjusted R^2 "penalizes" additional controls, might slightly prefer bc reduces temptation to include unnecessary variables. However, b/c few ind. variables, they are similar here

8. Consider the following result, which includes a quadratic term for education, `educ2` = `educ`^{2}. What is the marginal impact of an additional year of education for a high school graduate? For a college graduate (16 years)?

```
. regress kids educ educ2 meduc feduc age black east northcen west farm othrural town smcity

      Source |       SS           df       MS      Number of obs   =     1,129
-------------+----------------------------------   F(13, 1115)     =      9.35
       Model |  303.199016        13  23.3230012   Prob > F        =    0.0000
    Residual |  2782.31028     1,115  2.49534555   R-squared       =    0.0983
-------------+----------------------------------   Adj R-squared   =    0.0878
       Total |  3085.5093      1,128  2.73538059   Root MSE        =    1.5797

------------------------------------------------------------------------------
        kids |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1438818   .0968875     1.49   0.138    -.0462205    .3339841
       educ2 |  -.0107765   .0036654    -2.94   0.003    -.0179684   -.0035846
       meduc |  -.0002613   .0160116    -0.02   0.987    -.0316776     .031155
       feduc |  -.0217223    .018385    -1.18   0.238    -.0577953    .0143508
         age |   .0237972   .0082565     2.88   0.004     .0075972    .0399973
       black |   .9301003   .1746206     5.33   0.000     .5874783    1.272722
        east |   .2018261   .1352893     1.49   0.136    -.0636243    .4672765
    northcen |   .3418143   .1242263     2.75   0.006     .0980706    .5855581
        west |   .1999689     .17007     1.18   0.240    -.1337244    .5336622
        farm |  -.0768713   .1504208    -0.51   0.609    -.3720111    .2182684
    othrural |  -.1718755   .1794245    -0.96   0.338    -.5239232    .1801723
        town |   .0798203   .1266661     0.63   0.529    -.1687105    .3283511
      smcity |   .1855702    .162515     1.14   0.254    -.1332996    .5044399
       _cons |   1.630353   .7453841     2.19   0.029     .1678392    3.092866
------------------------------------------------------------------------------
```

$\frac{\partial kids}{\partial educ} = 0.144 - 2 \cdot 0.011 educ$

HS grad ≈ -0.114 children

Col grad ≈ -0.201 children