* **Regression on a subset of the data set**: in this case, the subset is the freshman (attend['frosh']==1) in the data set “attend”

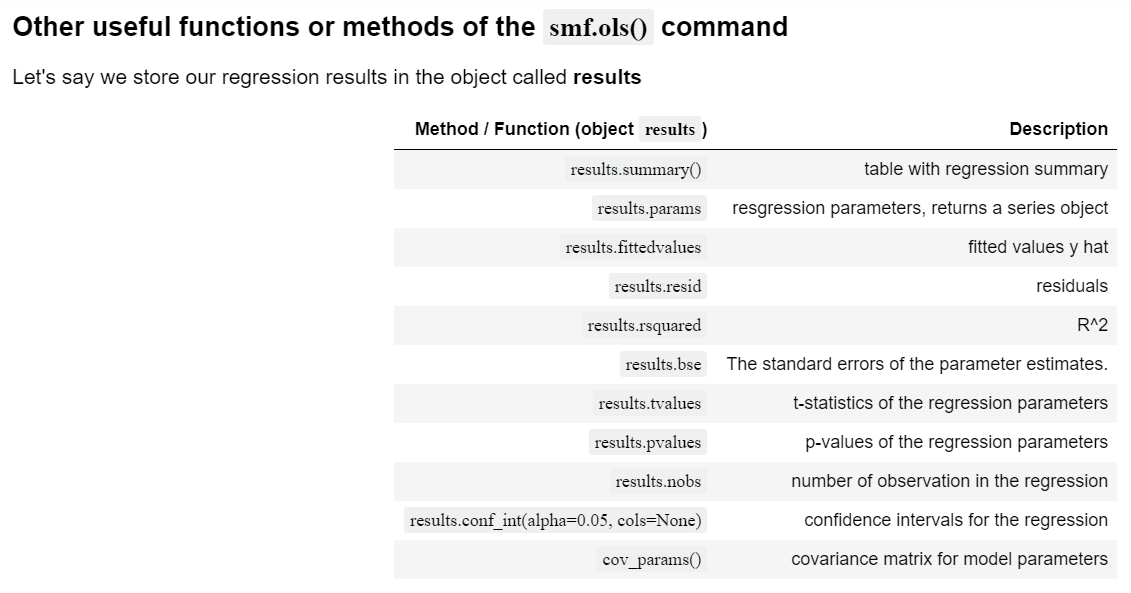
mfre = **smf.ols**(*formula* = 'atndrte ~ priGPA + ACT', *data* = attend[attend['frosh']==1]).fit()

* **Retrieving estimated parameters and summary stats**

yhat2 = mc4.*params[0]* + mc4.*params[1]*\*3.65 + mc4.*params[2]*\*20

mc4.**rsquared**

mc4.**nobs**



* ***A table with your regression results in one line***

**smf.ols(...).fit().summary()**

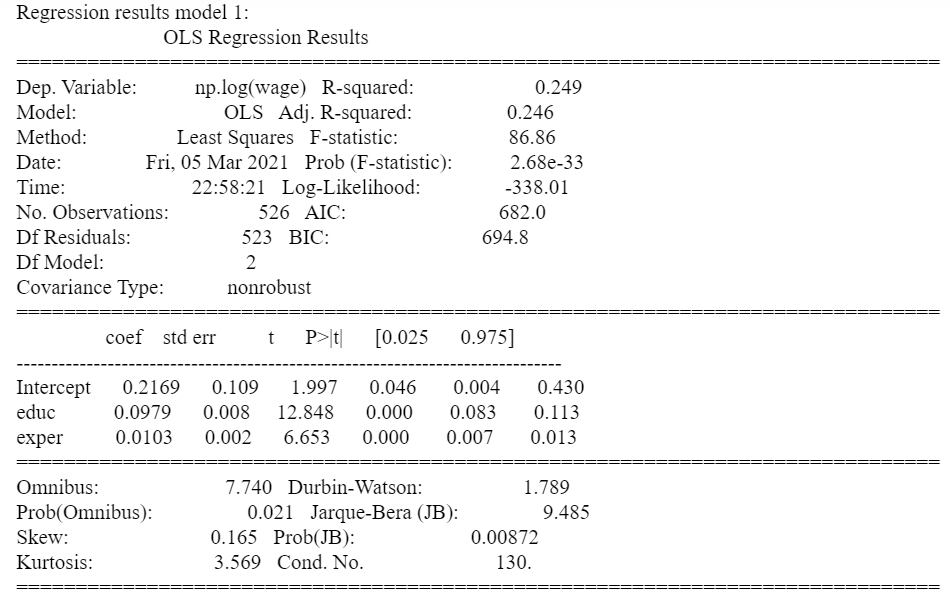
e.g.

wage1 = woo.dataWoo('wage1')

reg = **smf.ols**(formula='np.log(wage) ~ educ + exper ',

data=wage1)**.fit().summary()**

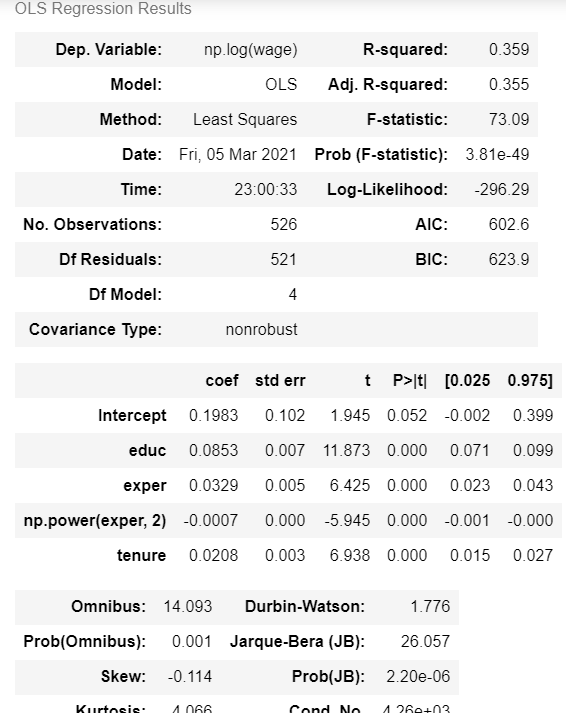
print(f'Regression results model 1: \n{reg}\n')



#Another Model

smf.ols(formula='np.log(wage) ~ educ + exper + np.power(exper,2) + tenure',

data=wage1).fit().summary()

****

* ***A Table for Beta’s***

m1 = smf.ols(formula='np.log(wage) ~ educ + exper + tenure',

data=wage1).fit()

m2 = smf.ols(formula='np.log(wage) ~ educ + exper + np.power(exper,2) + tenure',

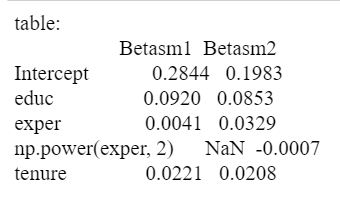
data=wage1).fit()

table = pd.DataFrame({'Betasm1': round(m1.params, 4), #Dictionary ->Data Frame

'Betasm2': round(m2.params, 4)})

table.style.format('{:,.3f}'.format)

print(f'table: \n{table}\n')

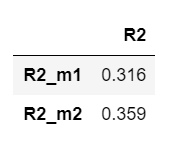


# Make another table comparing the R2

df = pd.DataFrame([round(m1.rsquared, 4),round(m2.rsquared, 4)],

index=['R2\_m1', 'R2\_m2'], columns=['R2'])

df.style.format('{:,.3f}'.format)



* **Scatter plots**

# *Graph the relationship between the predicted education and abil. Set motheduc and fatheduc at their average values in the sample, 12.18 and 12.45, respectively.*

# ***set the proper variables***

htv['yhat'] = m3.fittedvalues

htv['resid'] = m3.resid

htv['yaveduc'] = m3.params[0] + m3.params[1]\*12.18 + m3.params[2]\*12.45 + m3.params[3]\*htv['abil'] + m3.params[4]\*(htv['abil']\*\*2)

#***You need to sort the values by abil so that the line plots nicely****.* htv.sort\_values(by=['abil'], inplace=True)

*#* ***this does the scatter of the abil vs predicted education function***

sns.regplot(x = htv['abil'], y=htv['yhat'], scatter\_kws={"color": "blue", 'alpha': 0.3}, fit\_reg=False, data=htv).set\_title("log(Salary) vs log(Sales)")

# ***this adds the line that shows the quadratic relation between this two*.**

plt.plot(htv['abil'], htv['yaveduc'], color='purple', linewidth=4, label="Pop. Reg Line")

plt.title("A graph of {x,y}, the population regression function, and the sample regression function" )

plt.legend(loc='upper left')

plt.show()

