

IP[y]: Notebook (/tree/Documents/GitHub/CUSP_Coding/DS_HW3)

Data_Science_HW3 (autosaved)

File Edit View Insert Cell Kernel Help



Code

Cell Toolbar: None

Problem 1

```
In [4]: import math
import random
import pandas as pd
import numpy as np
import statsmodels.api as sm
from matplotlib import pyplot as plt
%matplotlib inline
```

```
In [5]: gril = pd.read_stata('griliches.dta')
```

1b

```
In [6]: bVariableList = ['rns', 'mrt', 'smsa', 'med', 'iq', 'kww', 'age', 's', 'expr', 'lw']

for var in bVariableList:
    print str(var)+" summary"
    print gril[var].describe()
    print " "
```

```
rns summary
count      758.000000
mean         0.269129
std          0.443800
min          0.000000
25%          0.000000
50%          0.000000
```