ANLT5030 – Unit 3 Assignment 1 Tutorial

SAS Studio

Instructions

- Use appropriate descriptive statistics to summarize each of the three variables for the 18 No Gulf View condominiums.
- Compare your summary results. Discuss any specific statistical results that would help a real estate agent understand the condominium market.
- Develop a 95% confidence interval estimate of the population mean sales price and population mean number of days to sell for Gulf View condominiums and interpret your results.
- Develop a 95% confidence interval estimate of the population mean sales price and population mean number of days to sell for No Gulf View condominiums and interpret your results.
- Assume the branch manager requested estimates of the mean selling price of Gulf View condominiums with a margin of error of $40,000 and the mean selling price of No Gulf View condominiums with a margin of error of $15,000. Using 95% confidence, how large should the sample sizes be?
- Gulf Real Estate Properties just signed contracts for two new listings: a Gulf View condominium with a list price of $589,000 and a No Gulf View condominium with a list price of $285,000. What is your estimate of the final selling price and number of days required to sell each of these units?

Dataset

- Download the GulfProp.xlsx file from the course datasets or from the Unit 3 Welcome announcement in the course announcements.

Edit Labels in Excel File First

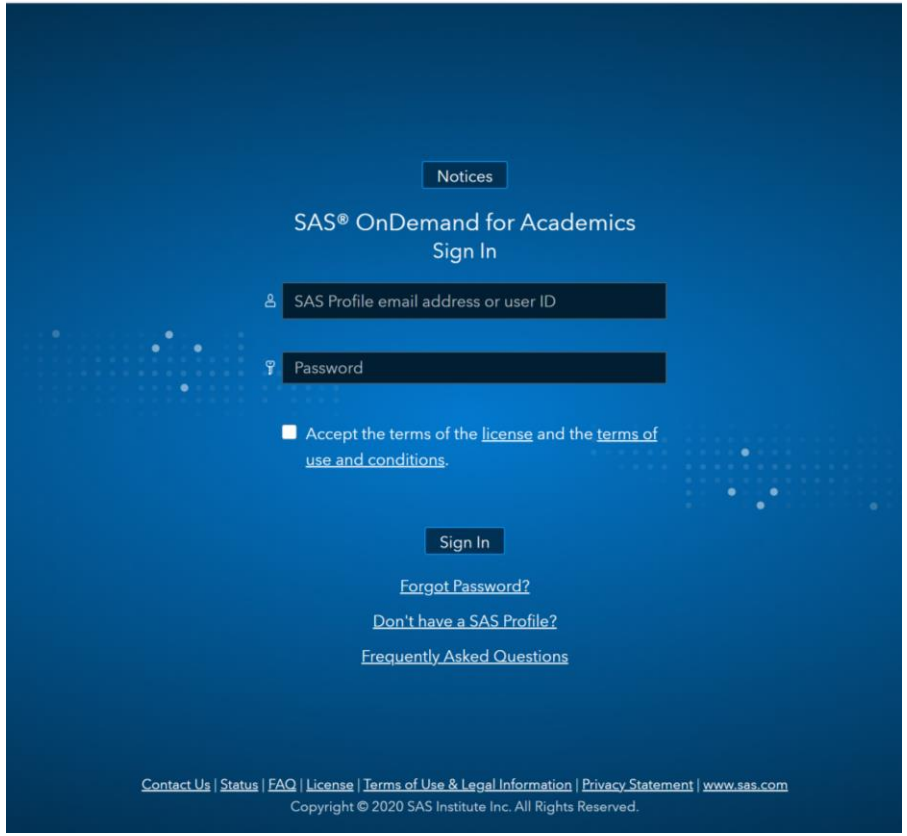
- First we need to edit the labels in the Excel file (using Excel) to have only one row of variable names (instead of two rows of labels).
- Then save the Excel file.

	A	B	C	D	E	F	G
1	Gulf View Condominiums			No Gulf View Condominiums			
2	List Price	Sale Price	Days to Sell	List Price	Sale Price	Days to Sell	
3	495.0	475.0	130	217.0	217.0	182	
4	379.0	350.0	71	148.0	135.5	338	
5	529.0	519.0	85	186.5	179.0	122	

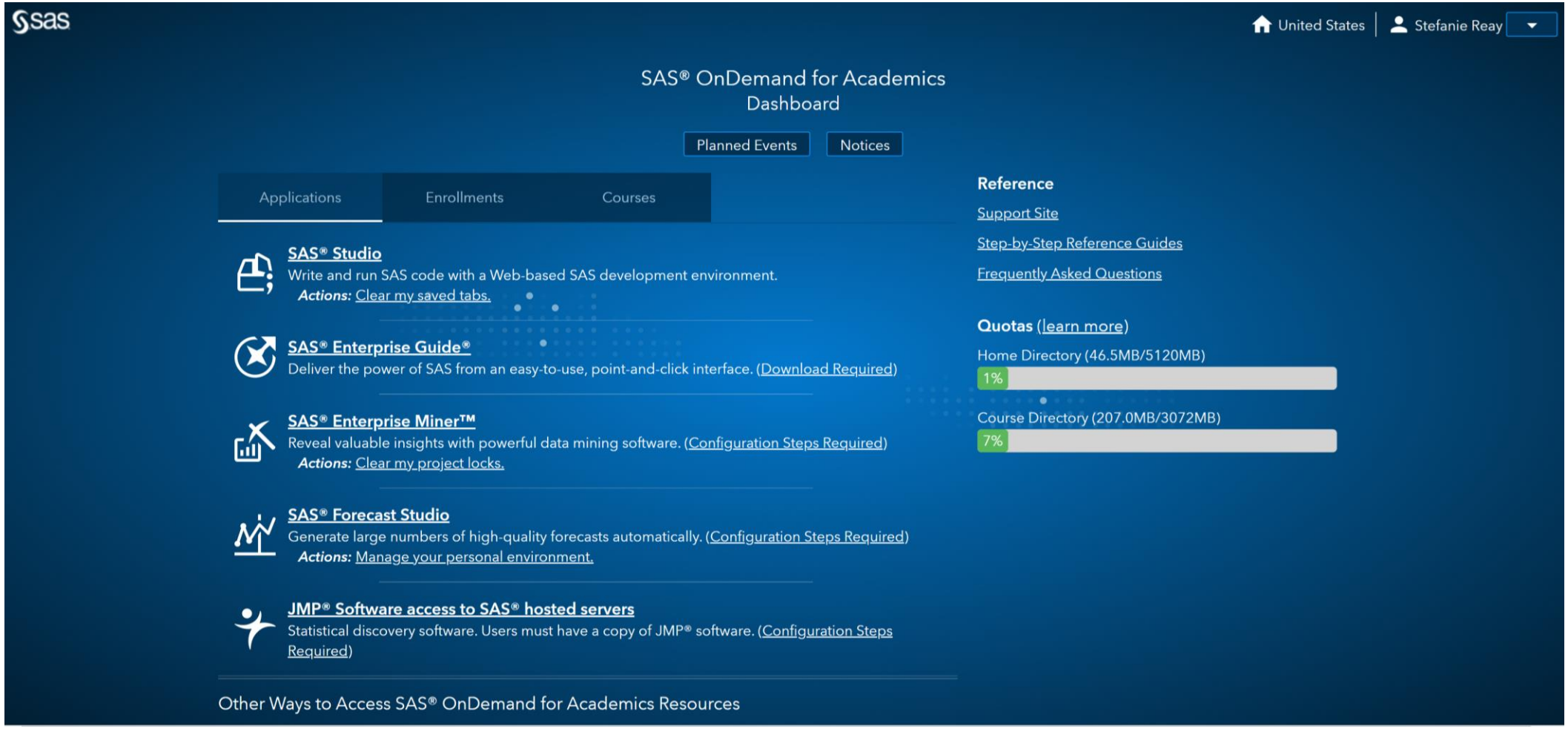
	A	B	C	D	E	F
1	GV_List Price	GV_Sale Price	GV_Days to Sell	NGV_List Price	NGV_Sale Price	NGV_Days to Sell
2	495.0	475.0	130	217.0	217.0	182
3	379.0	350.0	71	148.0	135.5	338
4	529.0	519.0	85	186.5	179.0	122

Access the SAS OnDemand for Academics Control Center

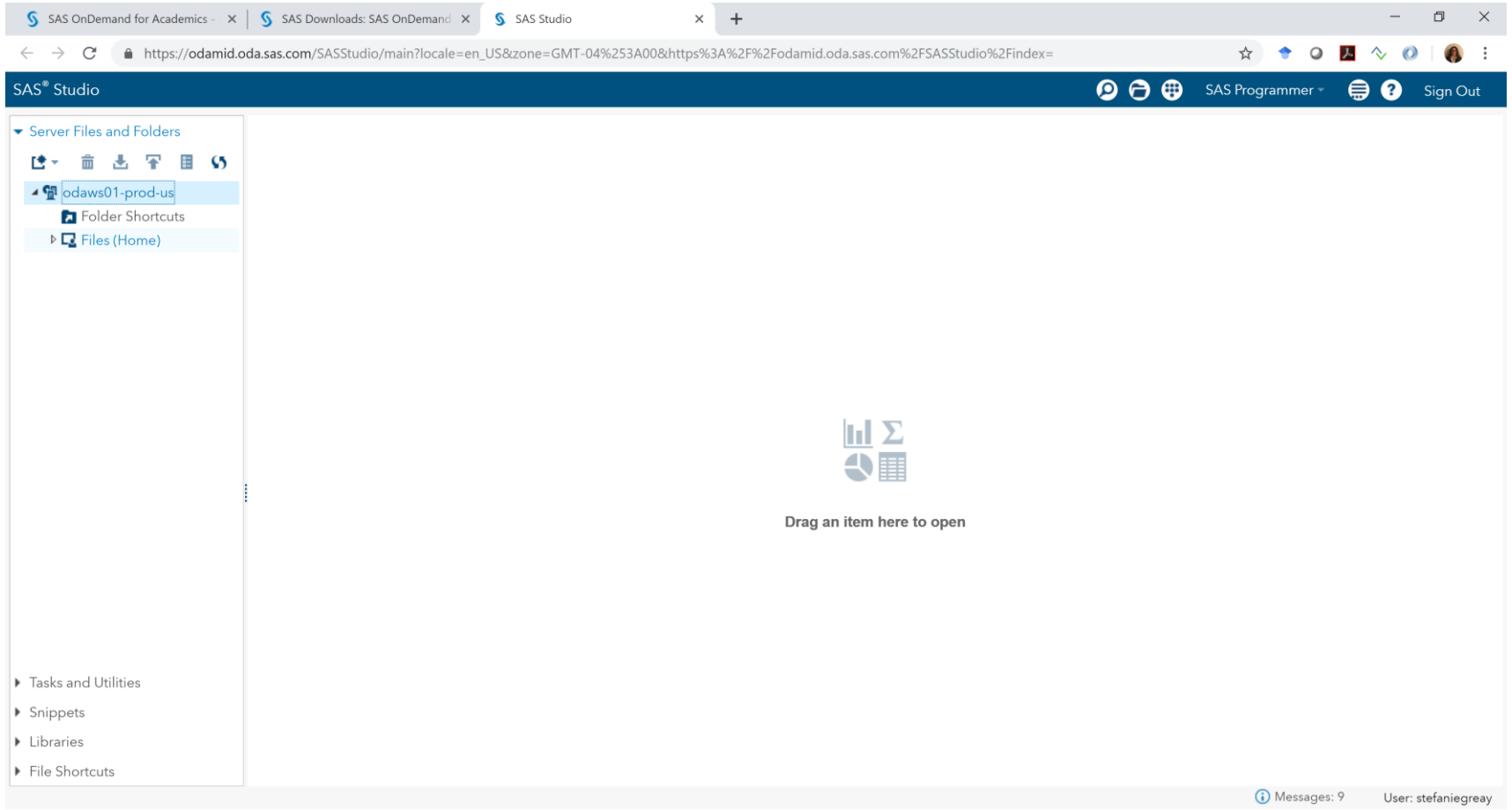
https://odamid.oda.sas.com/SASODAControlCenter



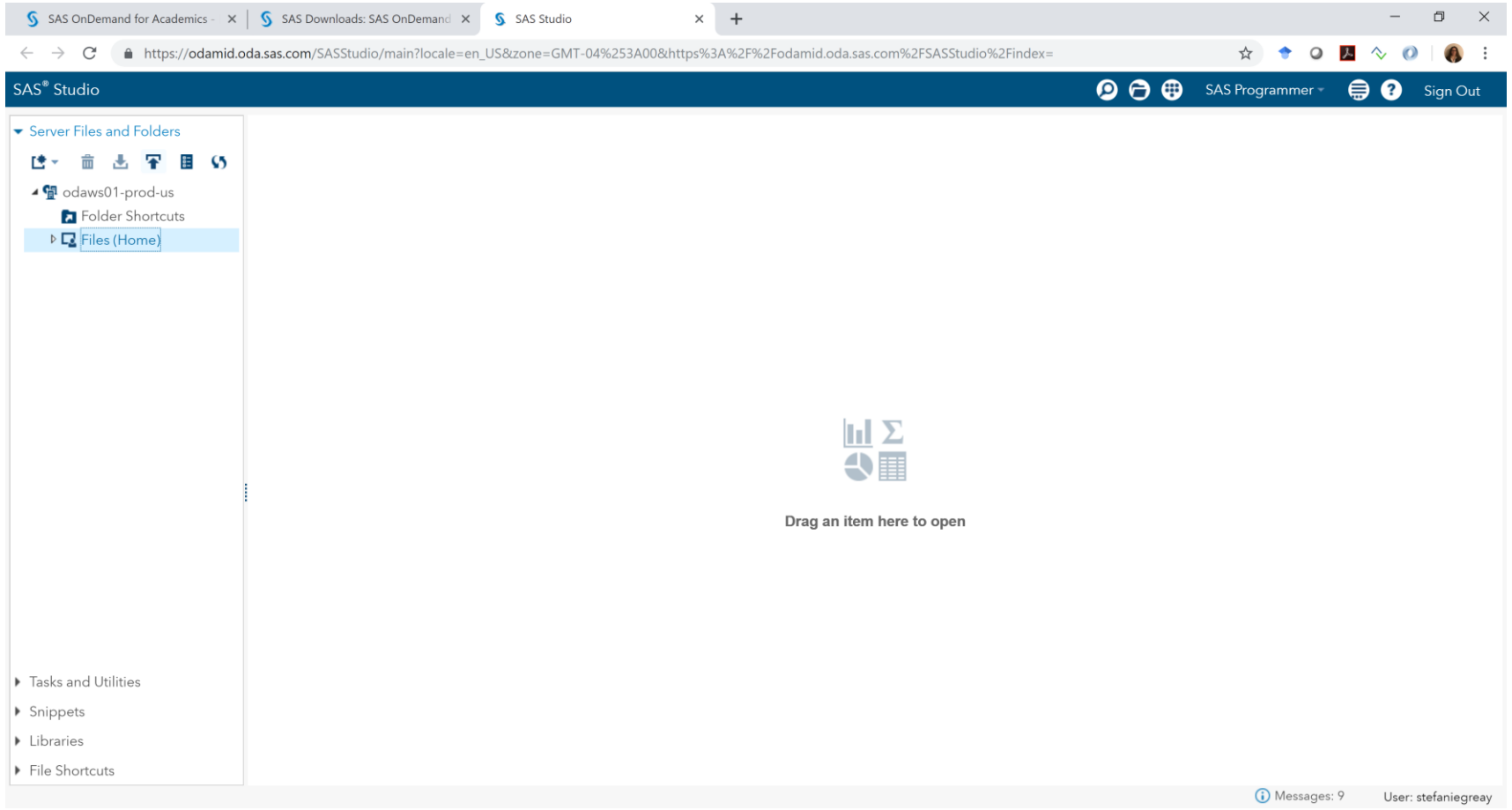
SAS OnDemand for Academics (SODA) Control Center



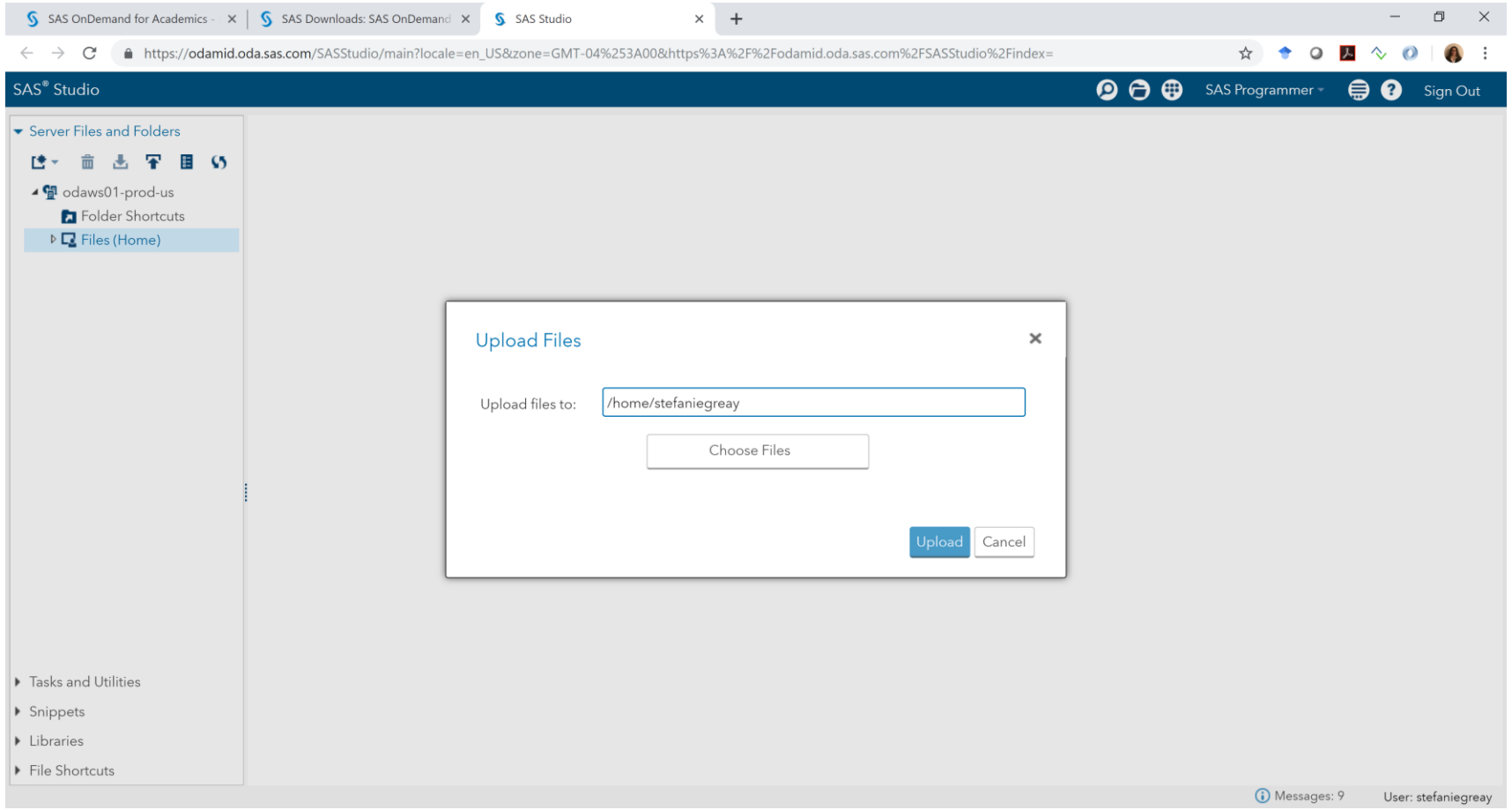
Click on Files(Home)



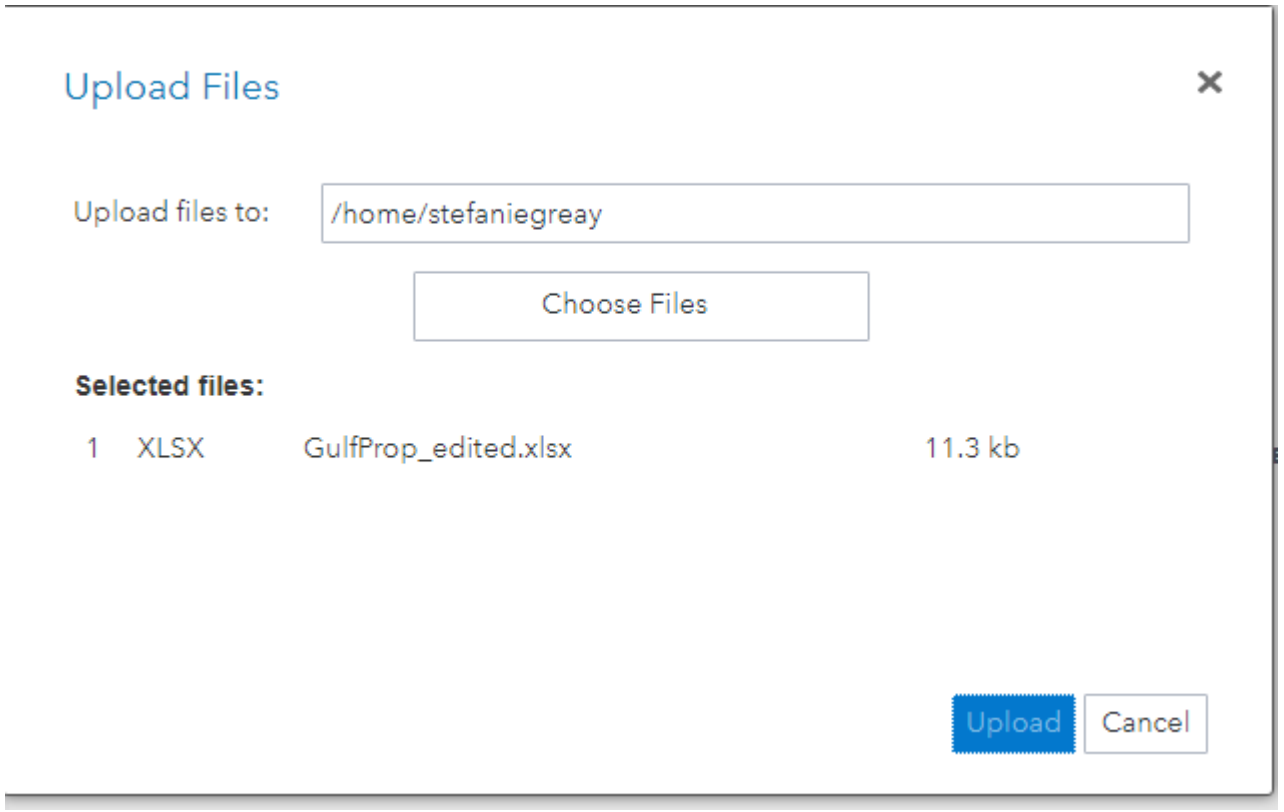
The Upload button will display in dark blue



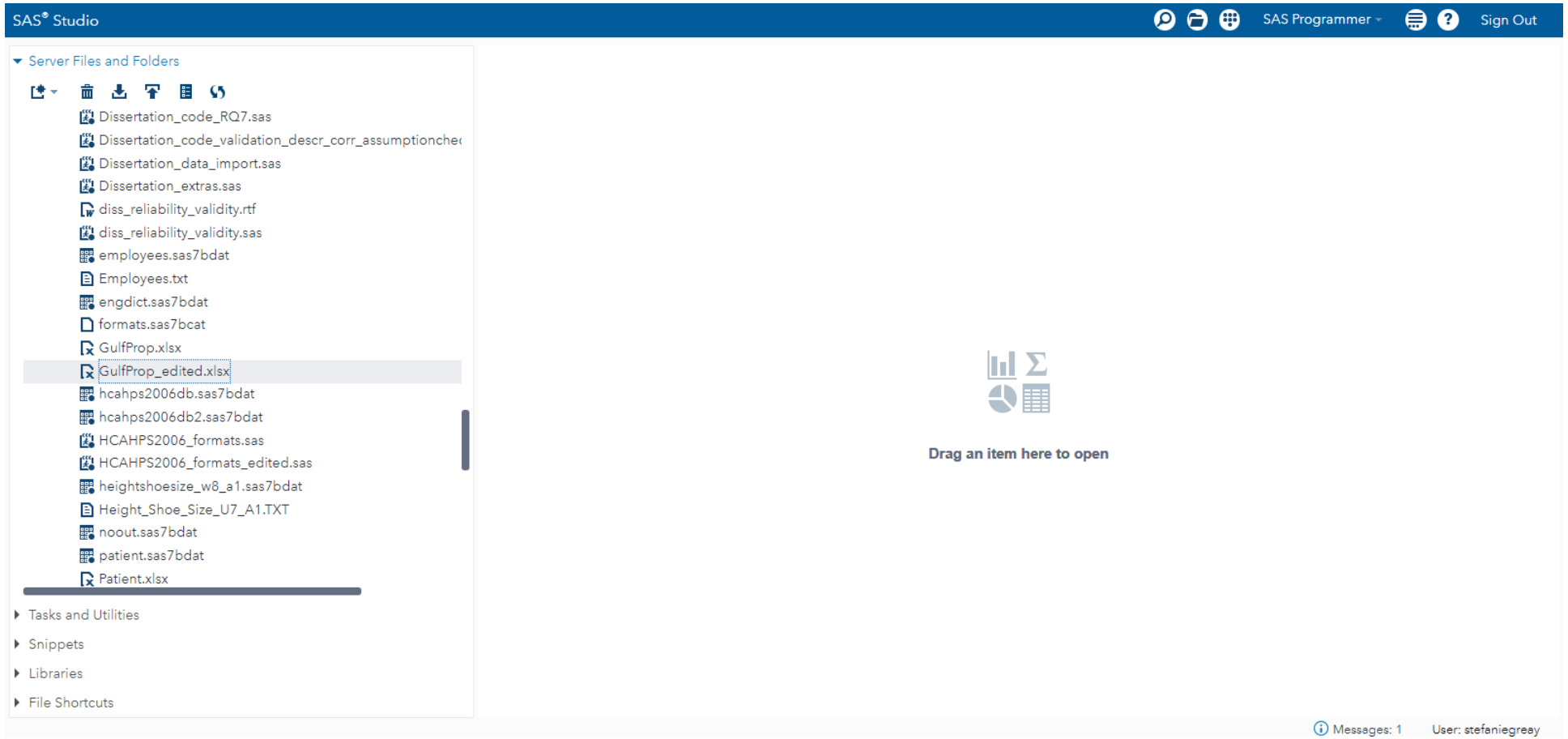
You can create a folder at this point, if you wish, or simply upload to your home directory.



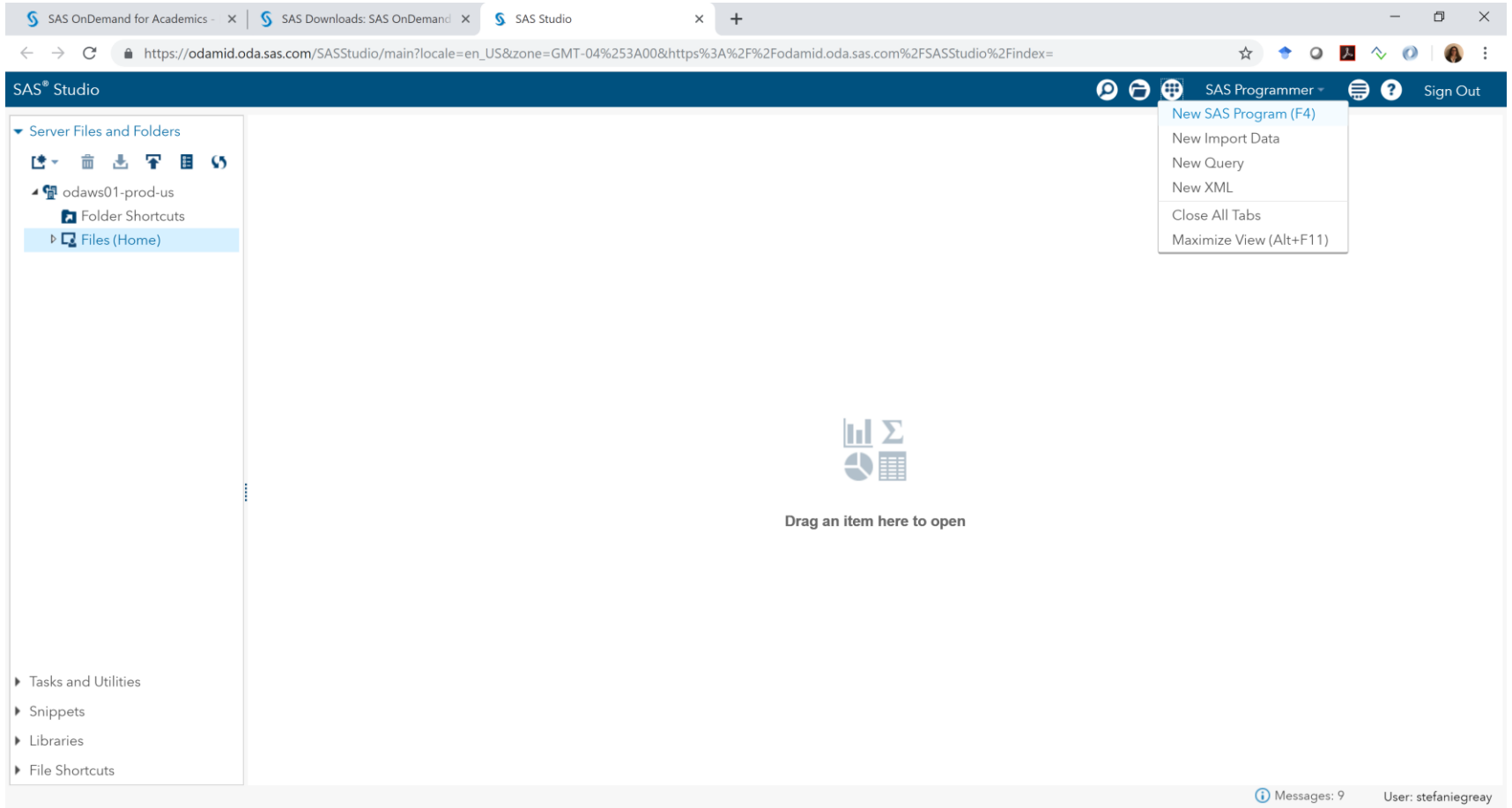
Select “Choose Files” to browse your computer for the dataset you want to upload. Once the dataset has been selected, click “Upload.”



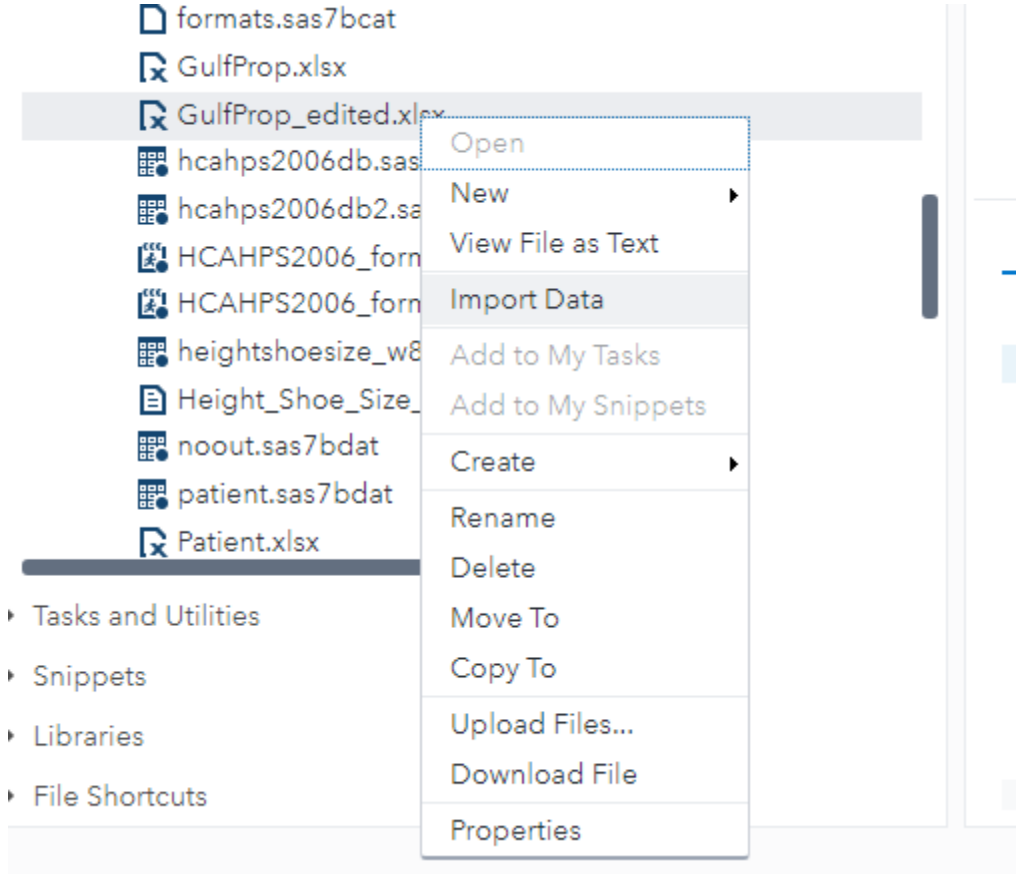
You will be able to view your files by clicking on “Files(Home)” to verify that your file successfully uploaded.



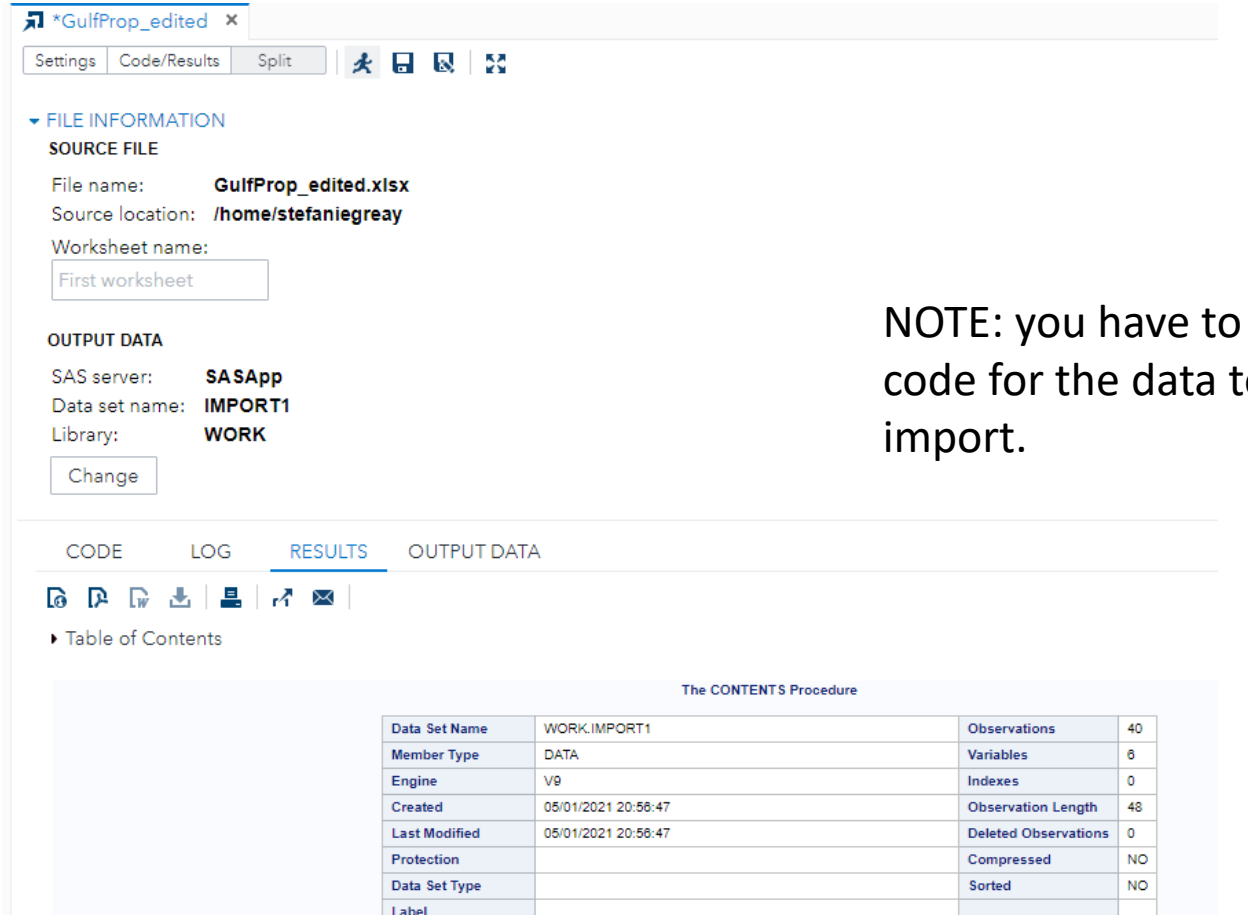
To get started with the SAS portion of the Unit 3 Assignment 1 assignment, start a new SAS program.



Import the dataset into a SAS dataset format (from the current xlsx format)

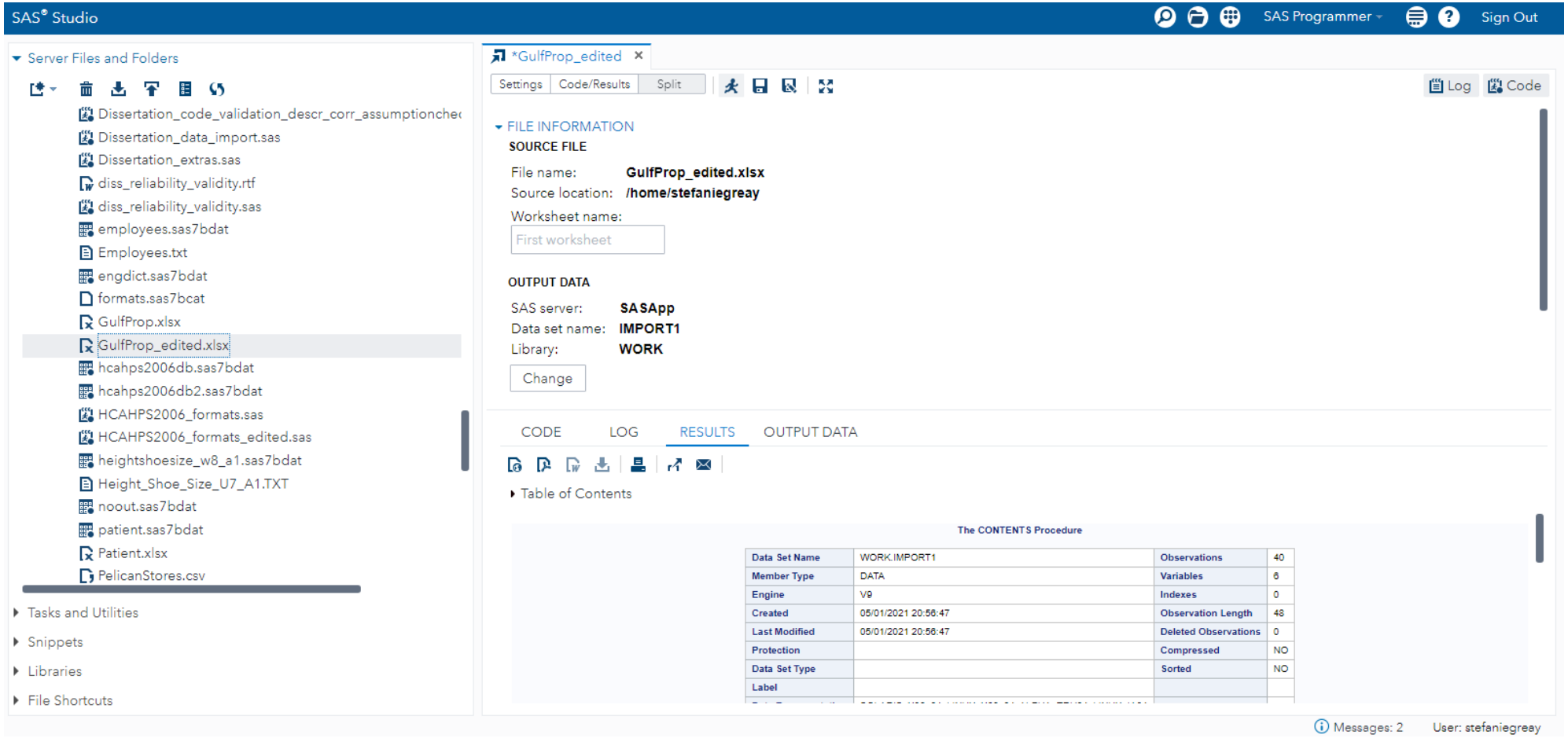


The Proc Import code will be written for you (save this as a template to use for future imports!)

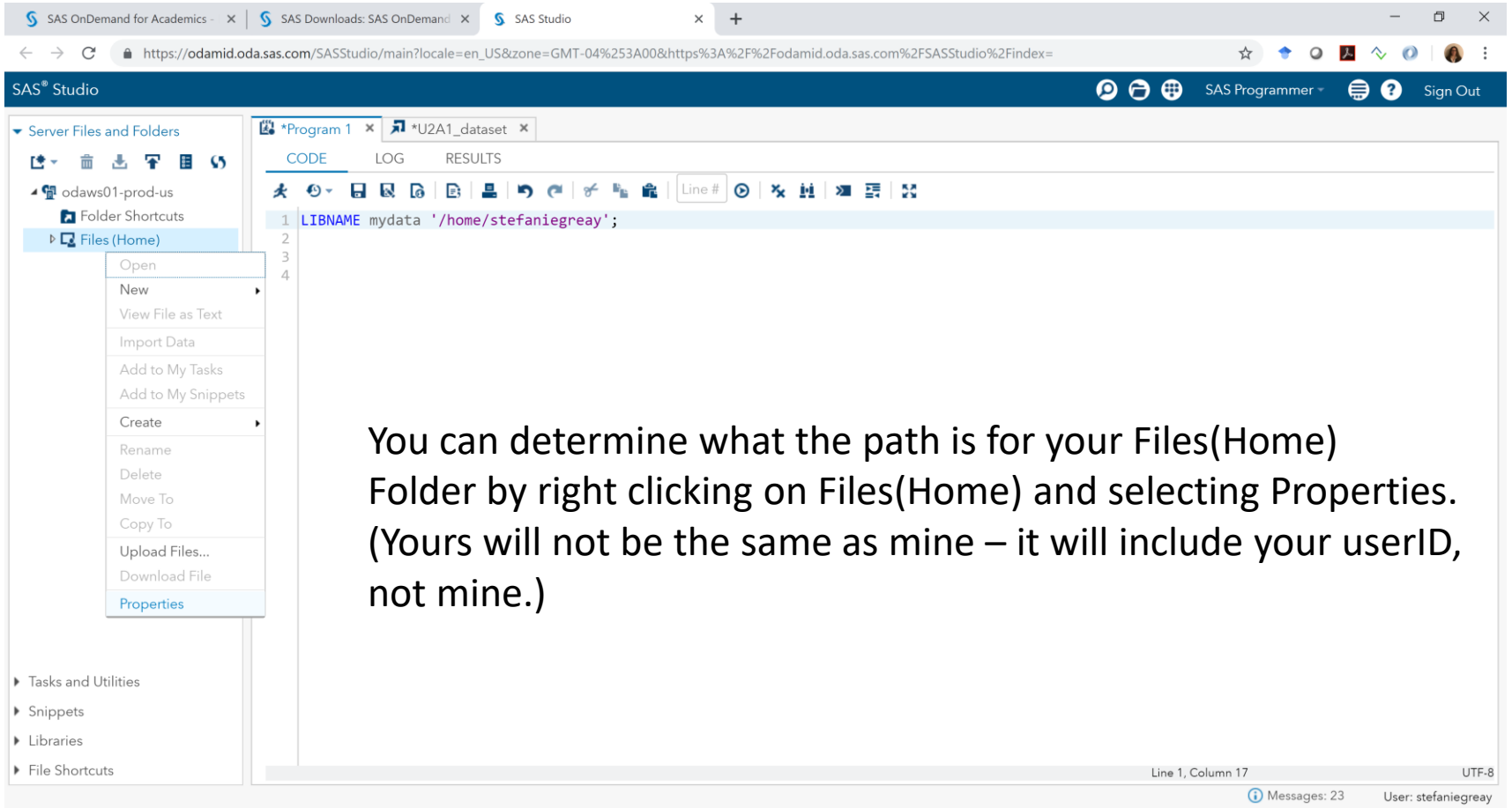


NOTE: you have to run this code for the data to actually import.

To run the code, click the icon that looks like a guy running.

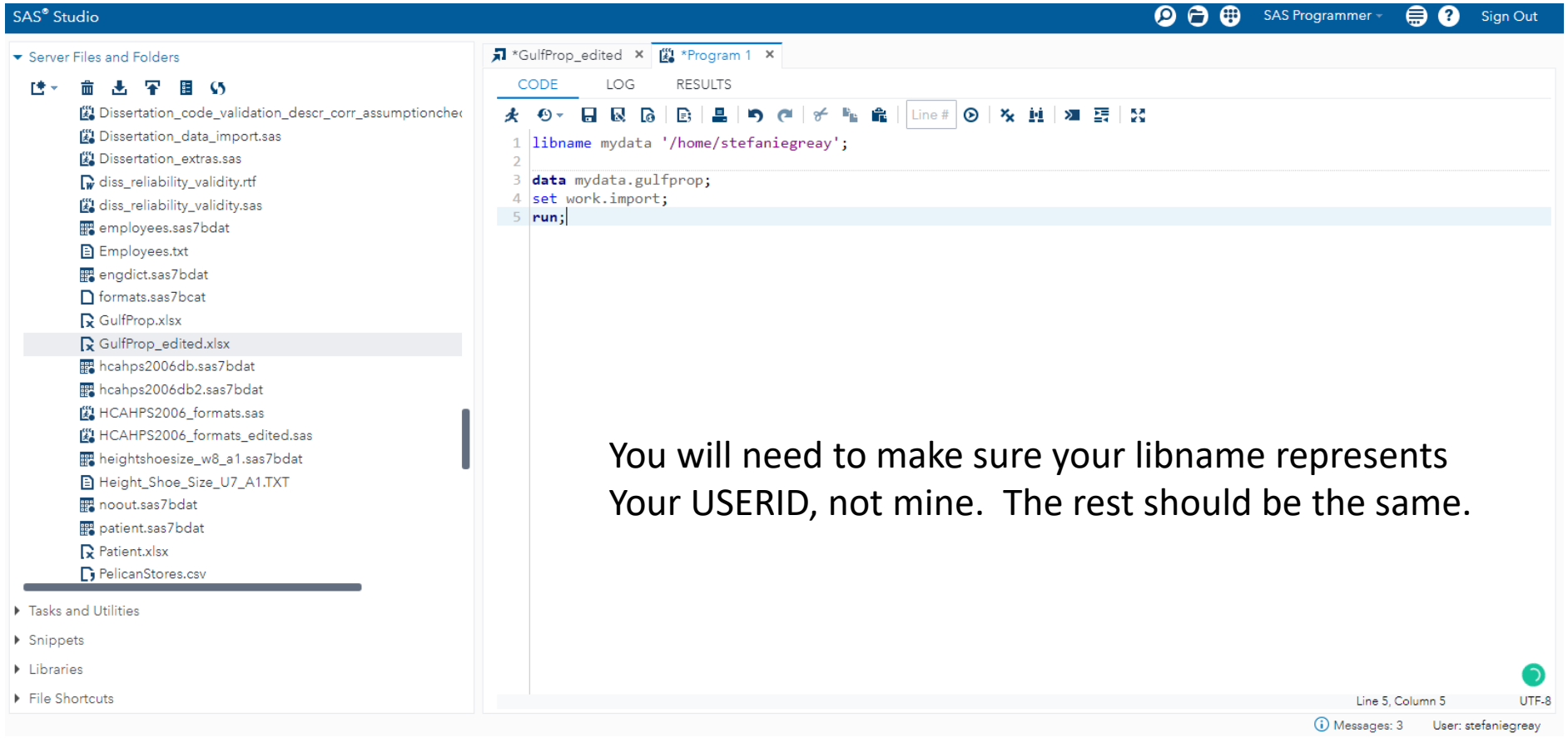


To create a SAS Library for your Files(Home) folder, you need to use a libname statement



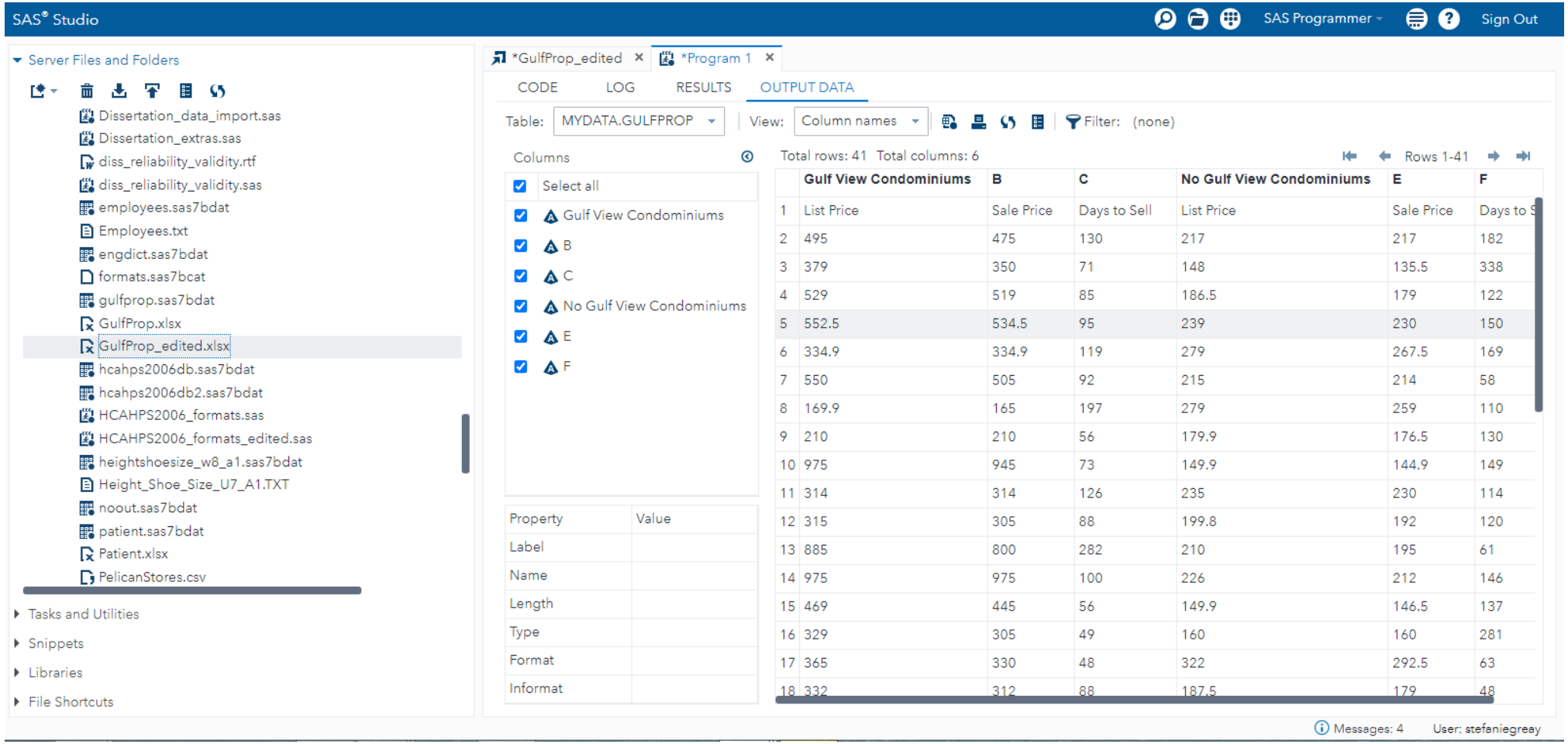
You can determine what the path is for your Files(Home) Folder by right clicking on Files(Home) and selecting Properties. (Yours will not be the same as mine – it will include your userID, not mine.)

Save the temporary SAS dataset created by the import to your library using the following sample code.



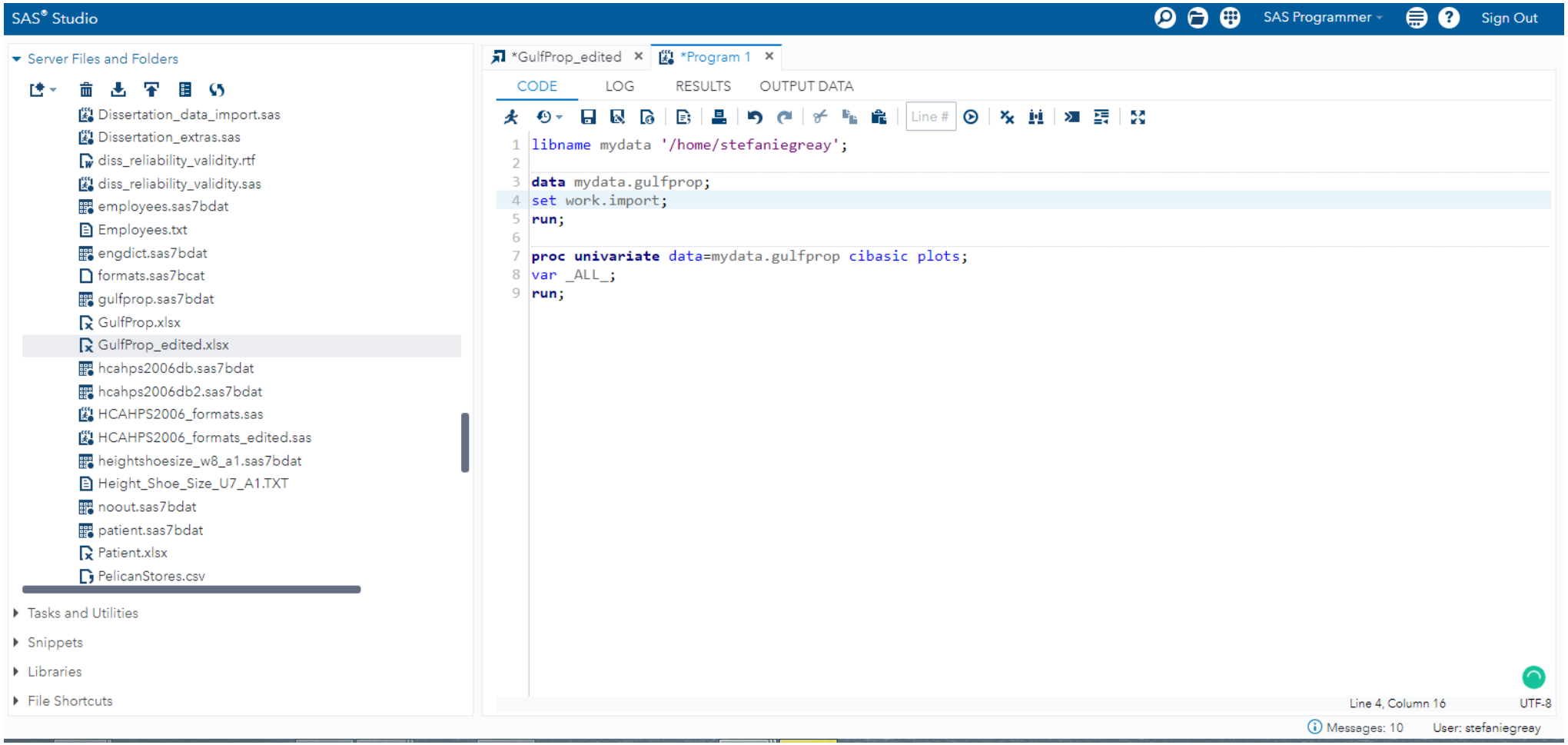
You will need to make sure your libname represents Your USERID, not mine. The rest should be the same.

When you run the code, you will see the dataset in the ouput data window and can verify its success.



	Gulf View Condominiums	B	C	No Gulf View Condominiums	E	F
1	List Price	Sale Price	Days to Sell	List Price	Sale Price	Days to S
2	495	475	130	217	217	182
3	379	350	71	148	135.5	338
4	529	519	85	186.5	179	122
5	552.5	534.5	95	239	230	150
6	334.9	334.9	119	279	267.5	169
7	550	505	92	215	214	58
8	169.9	165	197	279	259	110
9	210	210	56	179.9	176.5	130
10	975	945	73	149.9	144.9	149
11	314	314	126	235	230	114
12	315	305	88	199.8	192	120
13	885	800	282	210	195	61
14	975	975	100	226	212	146
15	469	445	56	149.9	146.5	137
16	329	305	49	160	160	281
17	365	330	48	322	292.5	63
18	332	312	88	187.5	179	48

You can now run any procedures against that dataset via the code window.



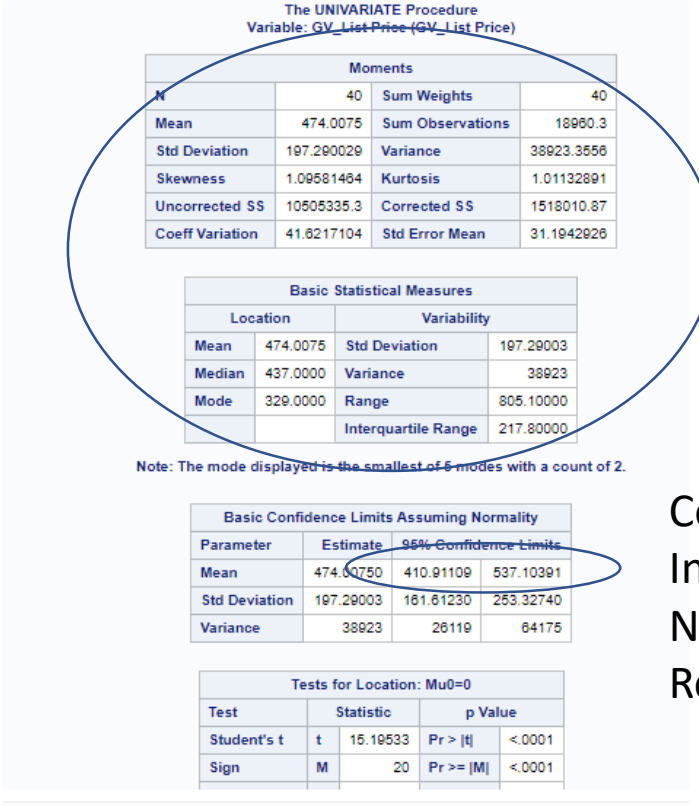
Sample Code for this assignment

```
libname mydata '/home/stefaniegreay';

data mydata.gulfprop;
set work.import;
run;

proc univariate data=mydata.gulfprop cibasic plots;
var _ALL_;
run;
```

Once you run the code, you can review the results to see the descriptive summaries (and graphs) for each of the variables for each type of condominium, and the confidence intervals for each.



The UNIVARIATE Procedure
Variable: GV_List Price (GV_List Price)

Moments			
N	40	Sum Weights	40
Mean	474.0075	Sum Observations	18960.3
Std Deviation	197.290029	Variance	38923.3556
Skewness	1.09581464	Kurtosis	1.01132891
Uncorrected SS	10505335.3	Corrected SS	1518010.87
Coeff Variation	41.6217104	Std Error Mean	31.1942926

Basic Statistical Measures			
Location		Variability	
Mean	474.0075	Std Deviation	197.29003
Median	437.0000	Variance	38923
Mode	329.0000	Range	805.10000
		Interquartile Range	217.80000

Note: The mode displayed is the smallest of 5 modes with a count of 2.

Basic Confidence Limits Assuming Normality			
Parameter	Estimate	95% Confidence Limits	
Mean	474.00750	410.91109	537.10391
Std Deviation	197.29003	161.61230	253.32740
Variance	38923	26119	64175

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	15.19533	Pr > |t|	<.0001
Sign	M	20	Pr >= |M|	<.0001

Numeric Descriptive Summary

Confidence Interval Numeric Results

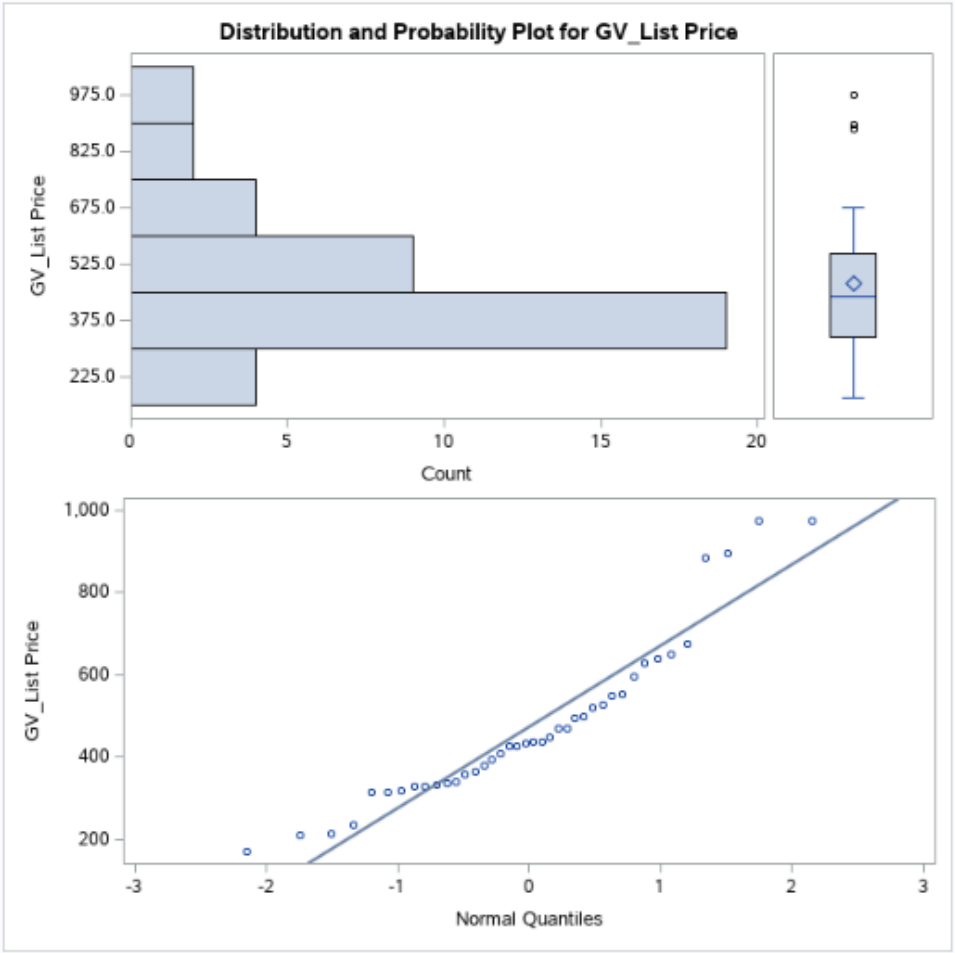
Interpreting Confidence Intervals

Example interpretation of a confidence interval for a 95% confidence interval for the average age of students currently enrolled at college ABCD that results in a point estimate of 20 with a margin of error of 2:

- “We are 95% confident that the true average/mean age of students currently enrolled at college ABCD is 20 years old, with a margin of error of 2 years.”
- OR
- “We are 95% confident that the true average/mean age of students currently enrolled at college ABCD is between 18 and 22 years old.”

Graphical descriptive summaries are below the tabular/numeric summaries

Histogram
(The histogram shows us whether the distribution of the data is symmetric or skewed (to the left or to the right) and whether it is unimodal or not.)



Box Plot
(The dots above or below the box and whiskers are outliers, the box represents the middle 50% of the data.)

Normal Probability Plot
(The closer to the line the dots are, the closer the distribution is to a normal distribution.)

Sample Size calculations

Parameter	Sample Size Estimate to get a (1-α)% confidence interval within a margin of error of E	Notes
μ	$n = \left(\frac{Z_{\left(\frac{\alpha}{2}\right)} * s}{E}\right)^2$ where n is rounded up to the nearest whole number	If σ is known (which is unusual), use σ instead of s.

E is the margin of error given in the problem.

s is the standard deviation from the descriptive summary for the variable of interest (from the software output)

$Z_{\left(\frac{\alpha}{2}\right)}$ is equal to 1.96 for a 95% confidence interval. (This $Z_{\left(\frac{\alpha}{2}\right)}$ value is the absolute value of Z (from the standard normal distribution) where the probability of being less than or equal to that value is equal to $\frac{\alpha}{2}$ where α is found from the 100*(1- α)% confidence interval, which would be α=.05 when the confidence level is 95%.)