

ANLT5010 – Week 4

Assignment 1 Tutorial

SAS Studio



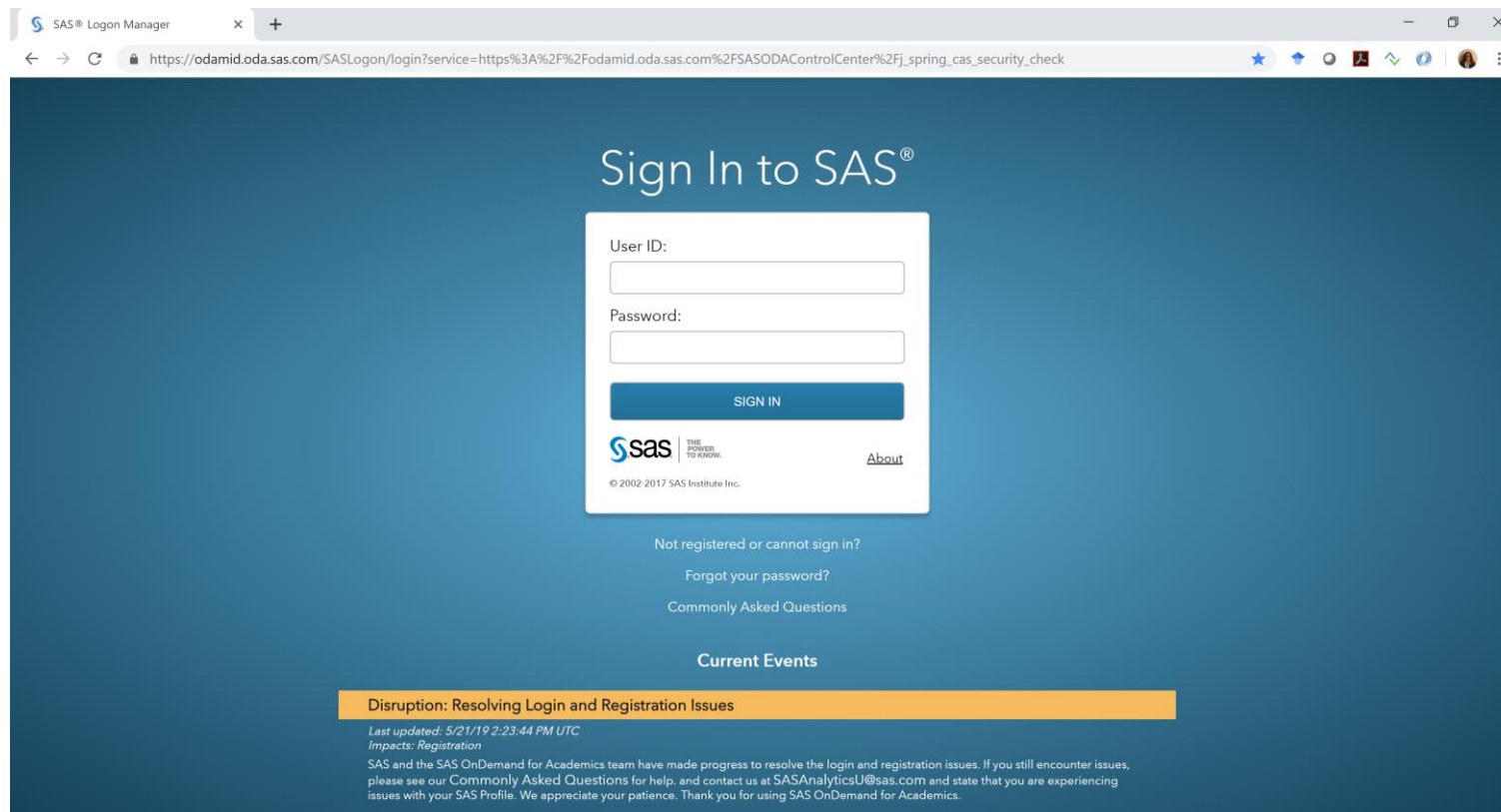
Dataset

- Download the cf_ANLT5010_W4_AirBNB.csv file from the Week 4 Welcome announcement in the course announcements or the Week 4 assignment area.



Access the SAS OnDemand for Academics Control Center

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



The screenshot shows a web browser window with the URL https://odamid.oda.sas.com/SASLogon/login?service=https%3A%2F%2Fodamid.oda.sas.com%2FSASODAControlCenter%2Fj_spring_cas_security_check. The page has a dark blue background and features a white login form in the center. The form is titled "Sign In to SAS®" and contains two input fields: "User ID:" and "Password:". Below these fields is a blue "SIGN IN" button. Under the button is the SAS logo with the tagline "THE POWER TO KNOW." and a link to "About". At the bottom of the form, it says "© 2002-2017 SAS Institute Inc.". Below the login form, there are links for "Not registered or cannot sign in?", "Forgot your password?", and "Commonly Asked Questions". At the bottom of the page, there is a section titled "Current Events" with a yellow banner that reads "Disruption: Resolving Login and Registration Issues". Below the banner, it says "Last updated: 5/21/19 2:23:44 PM UTC" and "Impacts: Registration". The text continues: "SAS and the SAS OnDemand for Academics team have made progress to resolve the login and registration issues. If you still encounter issues, please see our Commonly Asked Questions for help, and contact us at SASAnalyticsU@sas.com and state that you are experiencing issues with your SAS Profile. We appreciate your patience. Thank you for using SAS OnDemand for Academics."



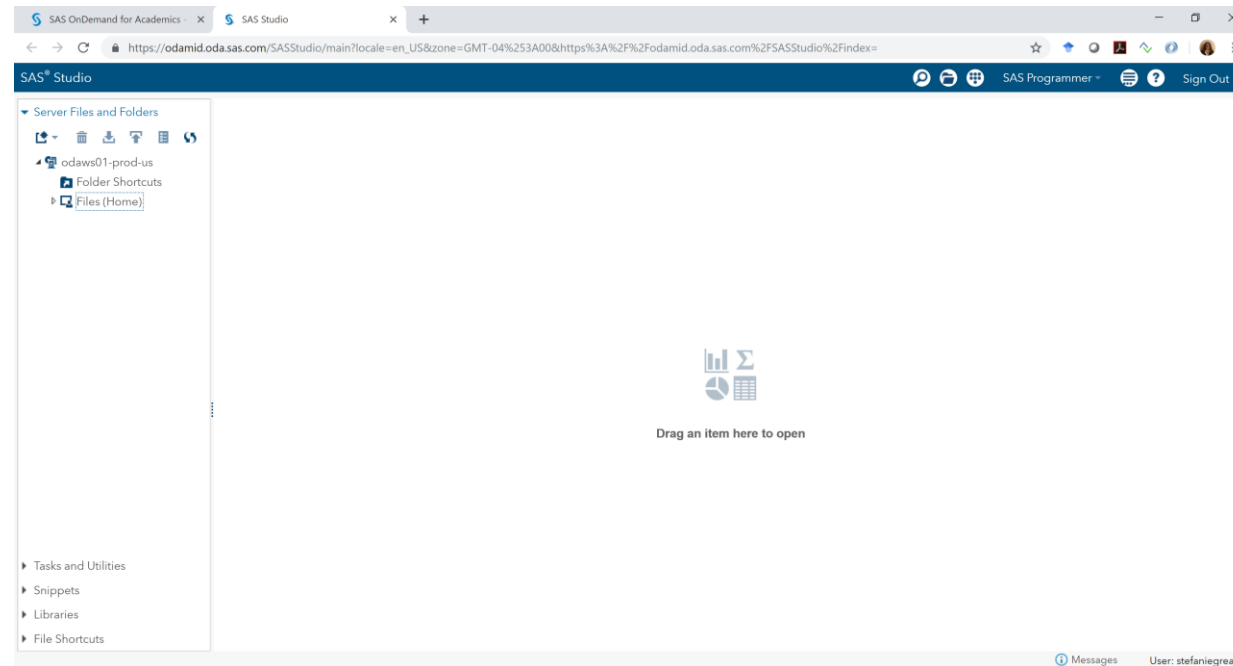
SAS OnDemand for Academics (SODA) Control Center

The screenshot displays the SAS OnDemand for Academics (SODA) Control Center dashboard. At the top right, the user is logged in as 'Stefanie Reay' in the 'United States' region. The dashboard title is 'SAS® OnDemand for Academics Dashboard'. Below the title are buttons for 'Planned Events' and 'Notices'. A message states: 'Your session in the United States region has timed out. You may now sign out to end this session as well or begin another regional session.' The main content area is divided into two columns. The left column has tabs for 'Applications', 'Enrollments', and 'Courses'. Under 'Applications', there are five items: 'SAS® Studio' (Write and run SAS code with a Web-Based SAS development environment. Actions: [Clear my saved tabs.](#)), 'SAS® Enterprise Guide®' (Deliver the power of SAS from an easy-to-use, point-and-click interface. (Download Required)), 'SAS® Enterprise Miner™' (Reveal valuable insights with powerful data mining software. (Configuration Steps Required) Actions: [Clear my project locks.](#)), 'SAS® Forecast Studio' (Generate large numbers of high-quality forecasts automatically. (Configuration Steps Required) Actions: [Manage your personal environment.](#)), and 'JMP® Software access to SAS® hosted servers' (Statistical discovery software. Users must have a copy of JMP® software. (Configuration Steps Required)). The right column has a 'Reference' section with links to 'Support Site', 'Step-by-Step Reference Guides', and 'Frequently Asked Questions'. Below this is a 'Quotas (learn more)' section showing two progress bars: 'Home Directory (22.5MB/5120MB)' at 0% and 'Course Directory (207.0MB/3072MB)' at 7%.

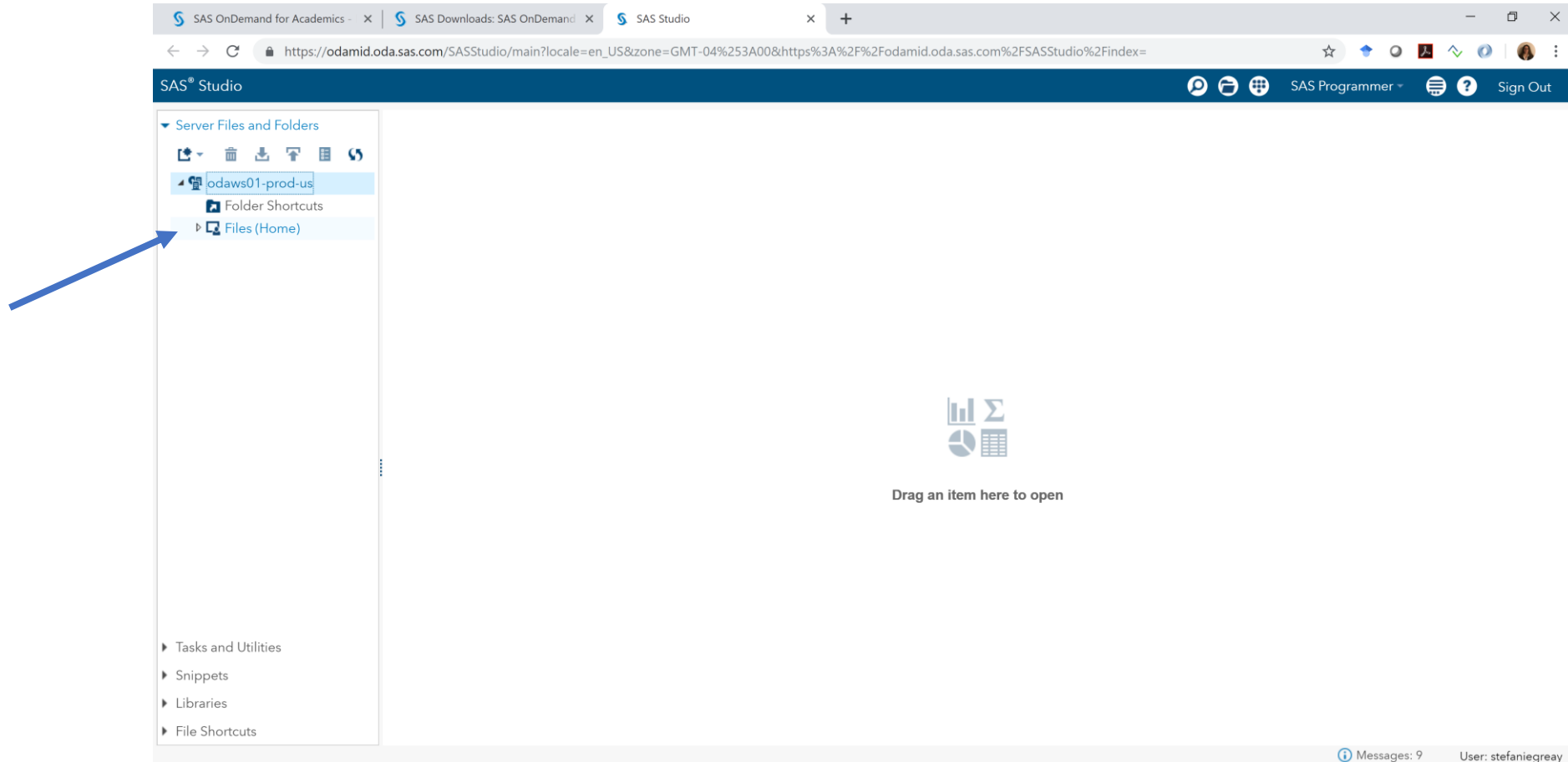


SAS Studio

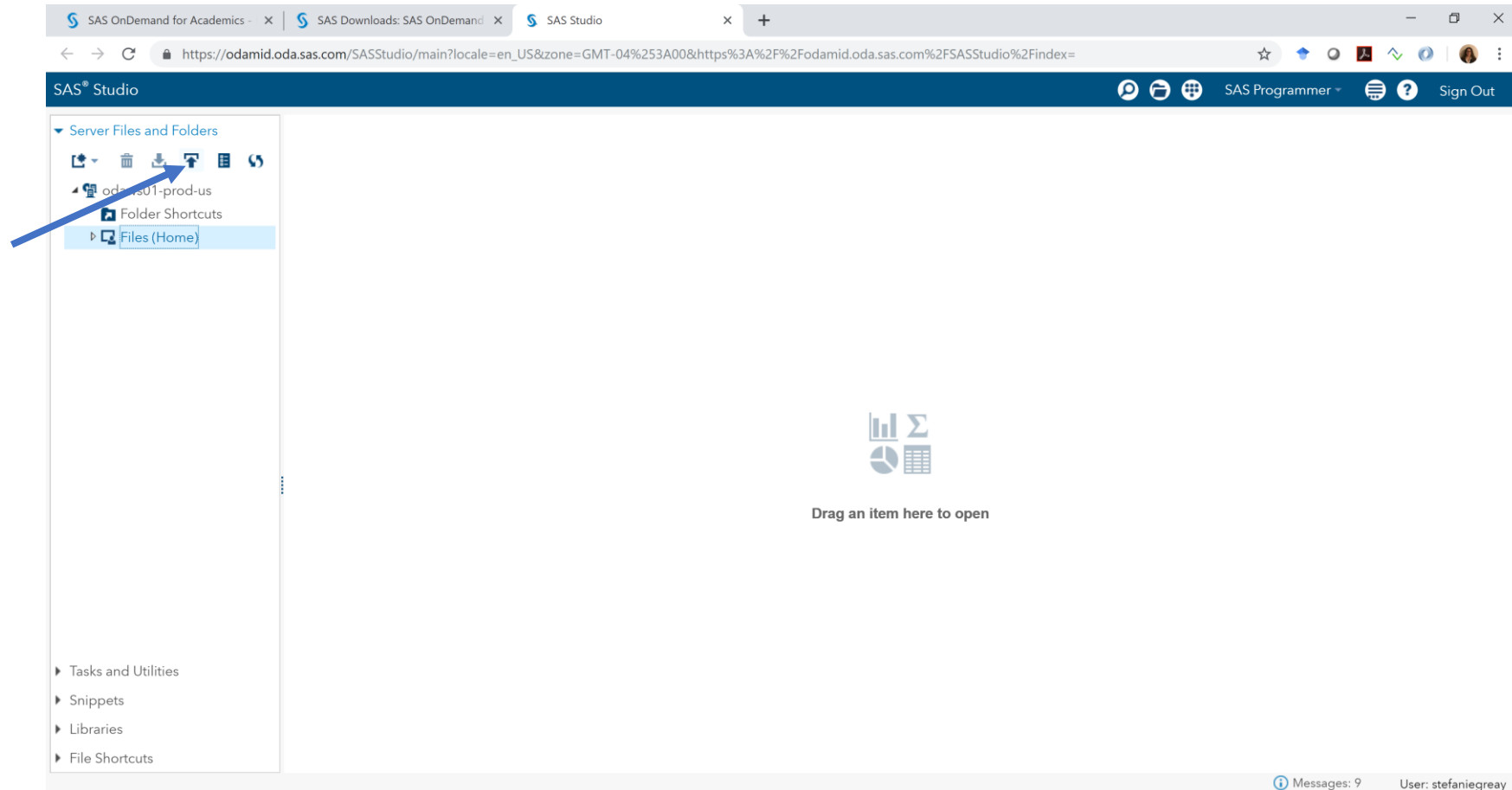
<https://odamid.oda.sas.com/SASStudio/>



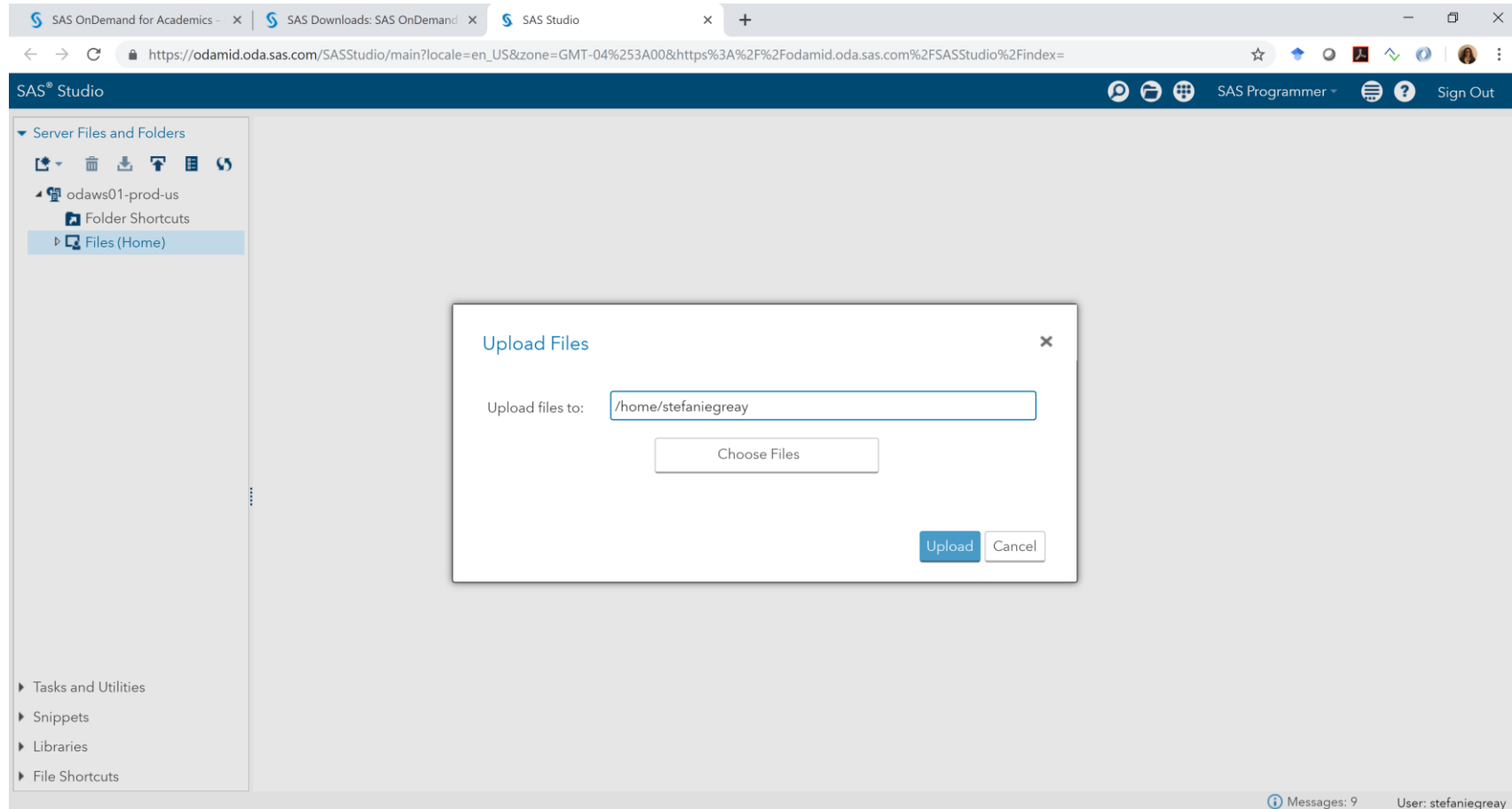
Click on Files(Home)



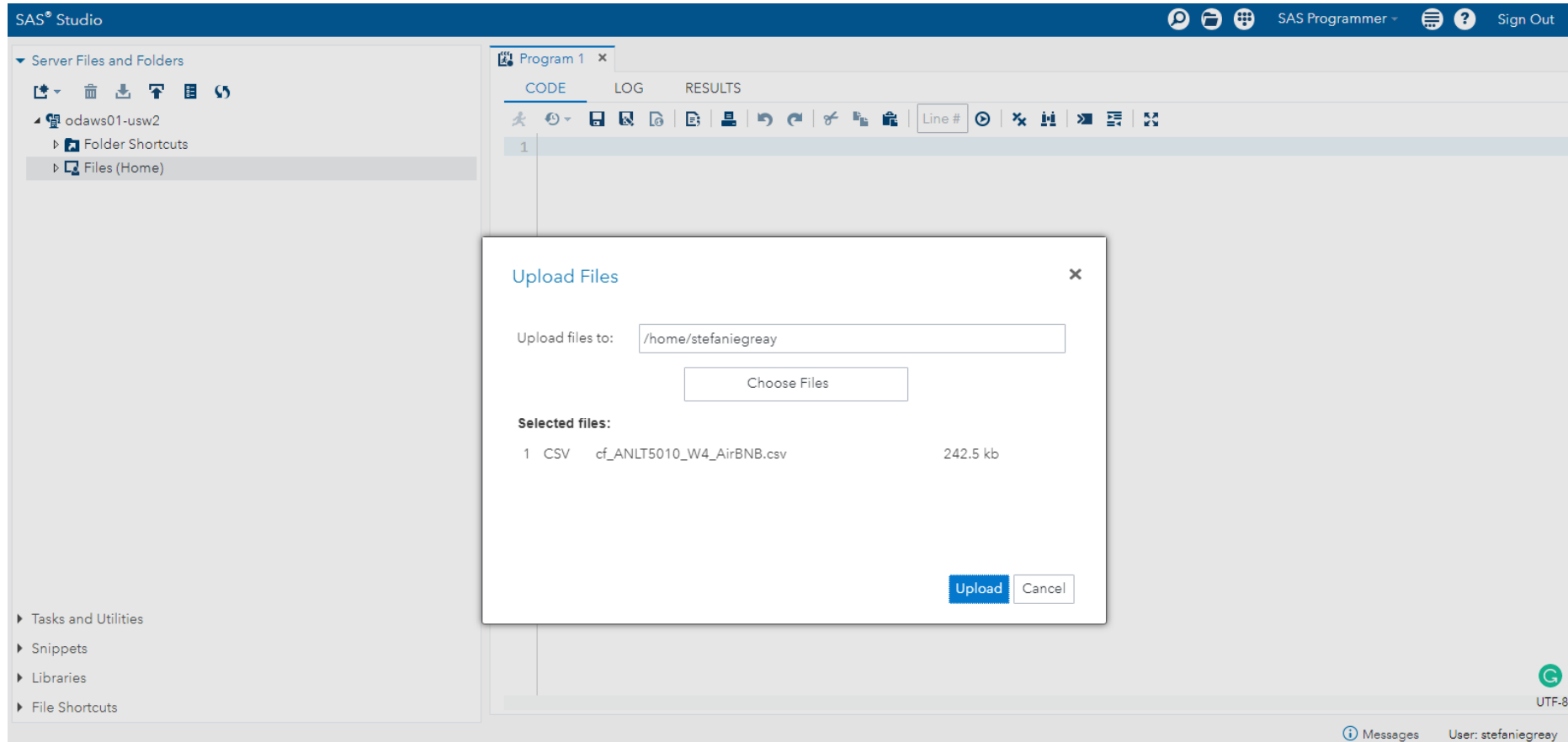
The Upload button will display in dark blue



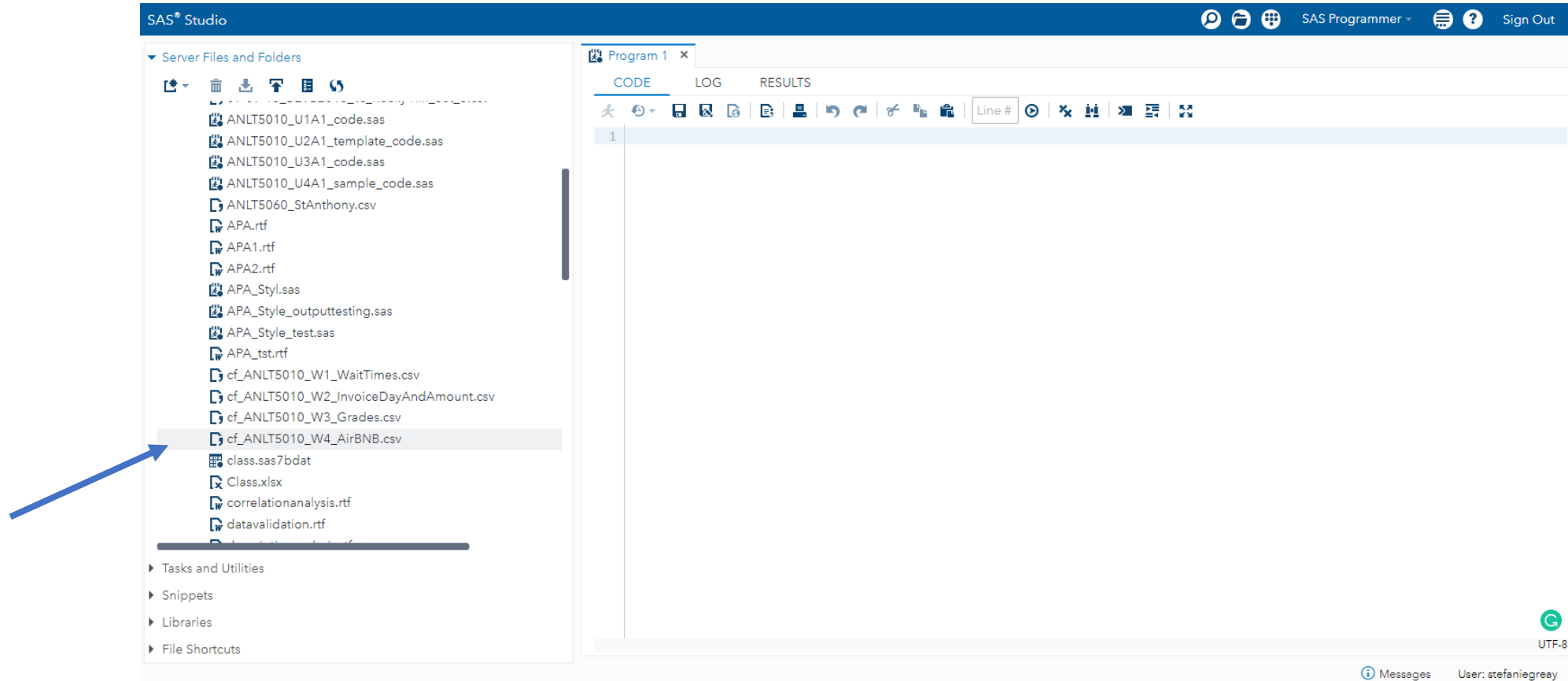
You can create a folder at this point, if you wish, or simply upload to your home directory.



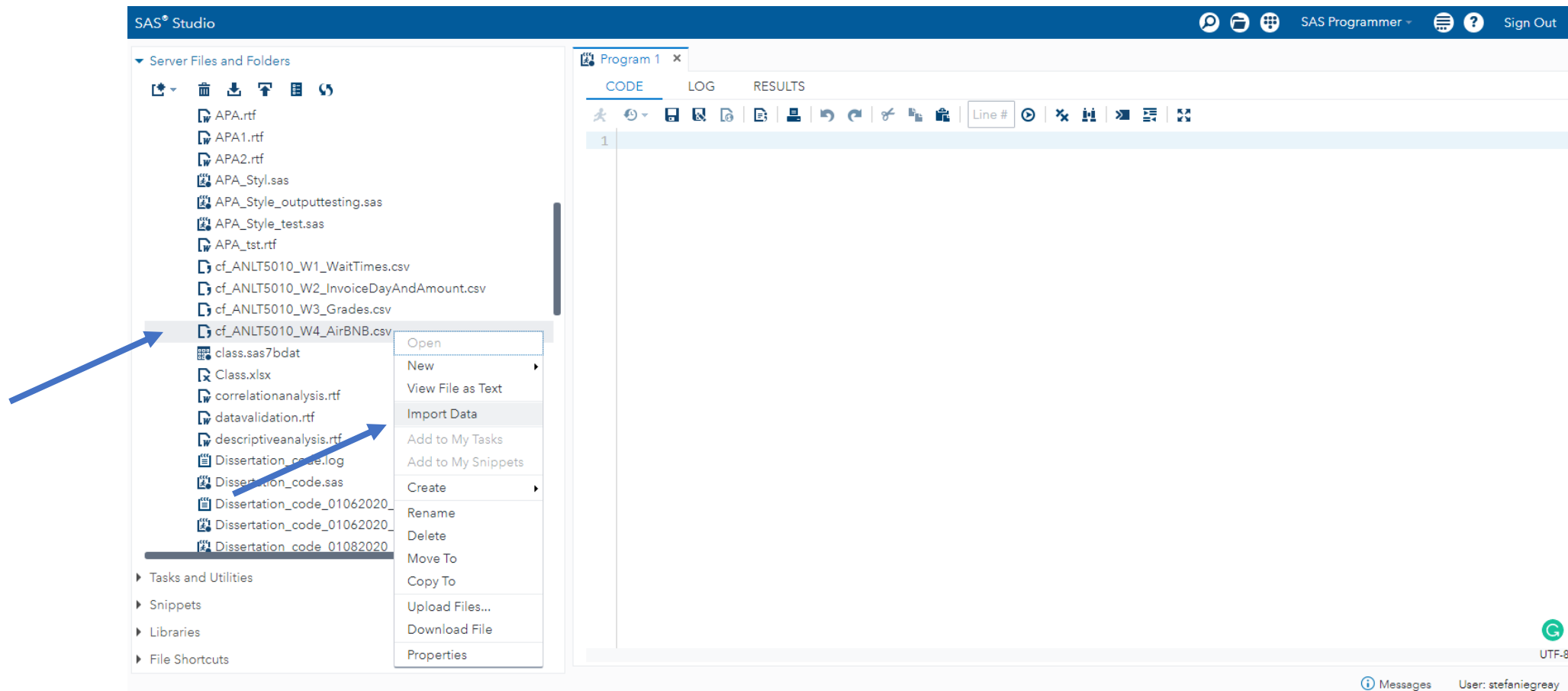
Select “Choose Files” to browse your computer for the dataset you want to upload. Once the dataset has been selected, click “Upload.”



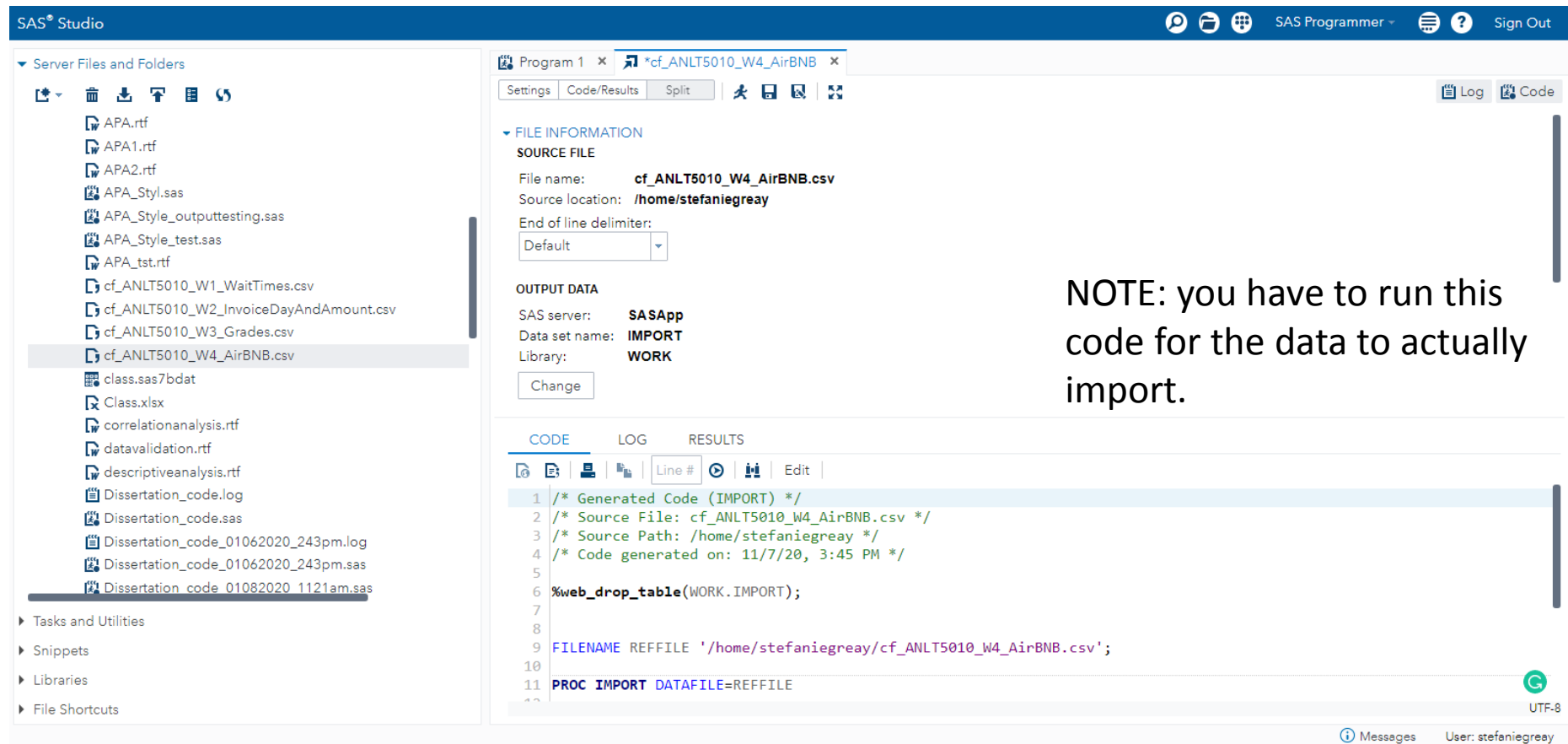
You will be able to view your files by clicking on “Files(Home)” to verify that your file successfully uploaded.



To import the dataset into a SAS dataset format (from the current csv format), right click on the name of the file, and select “Import Data.”



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



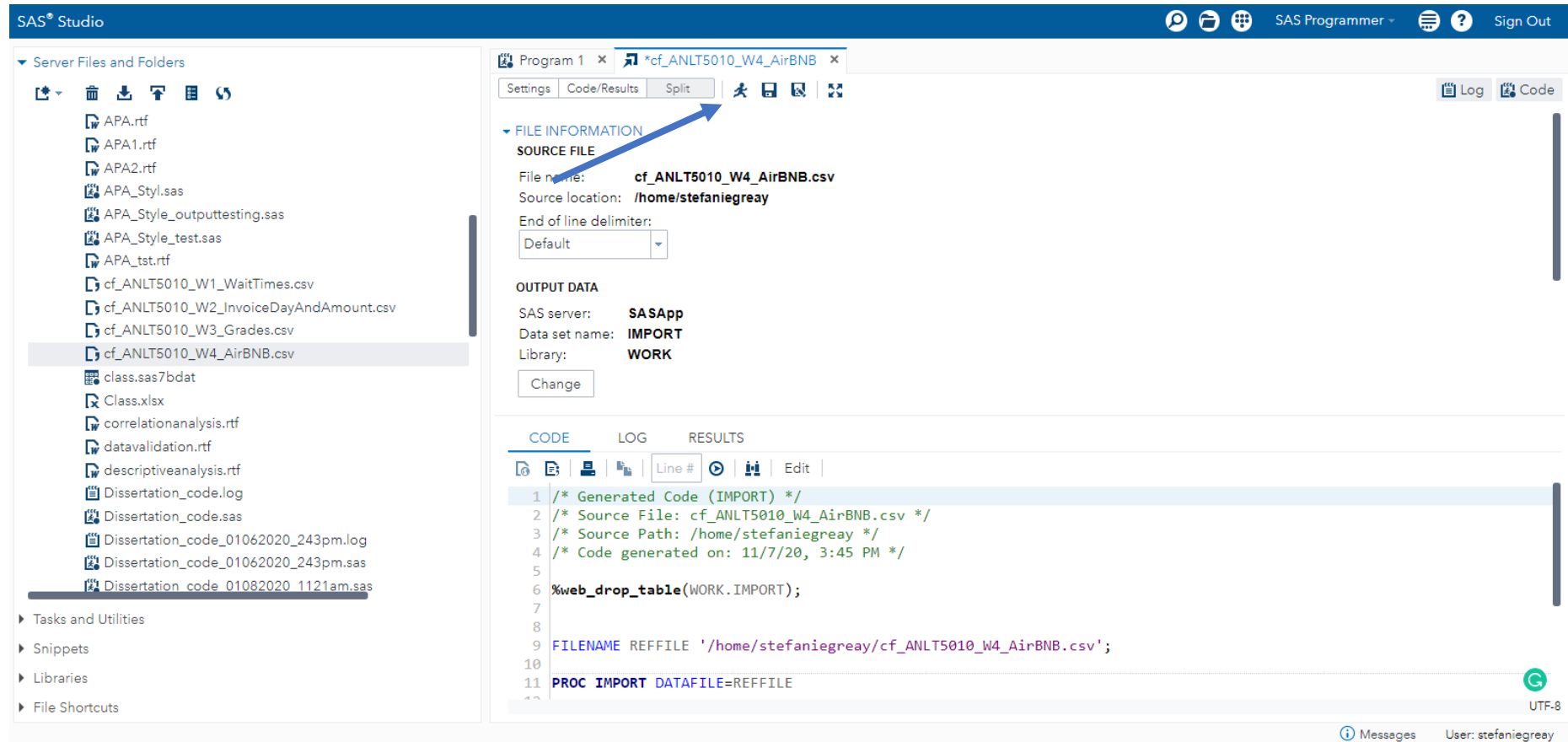
The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'cf_ANLT5010_W4_AirBNB.csv' selected. The main window displays the 'FILE INFORMATION' and 'OUTPUT DATA' sections. The 'FILE INFORMATION' section shows the file name 'cf_ANLT5010_W4_AirBNB.csv' and source location '/home/stefaniegreay'. The 'OUTPUT DATA' section shows the SAS server 'SASApp', data set name 'IMPORT', and library 'WORK'. Below these sections, the 'CODE' pane displays the generated Proc Import code:

```
1 /* Generated Code (IMPORT) */  
2 /* Source File: cf_ANLT5010_W4_AirBNB.csv */  
3 /* Source Path: /home/stefaniegreay */  
4 /* Code generated on: 11/7/20, 3:45 PM */  
5  
6 %web_drop_table(WORK.IMPORT);  
7  
8  
9 FILENAME REFFILE '/home/stefaniegreay/cf_ANLT5010_W4_AirBNB.csv';  
10  
11 PROC IMPORT DATAFILE=REFFILE  
12
```

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.



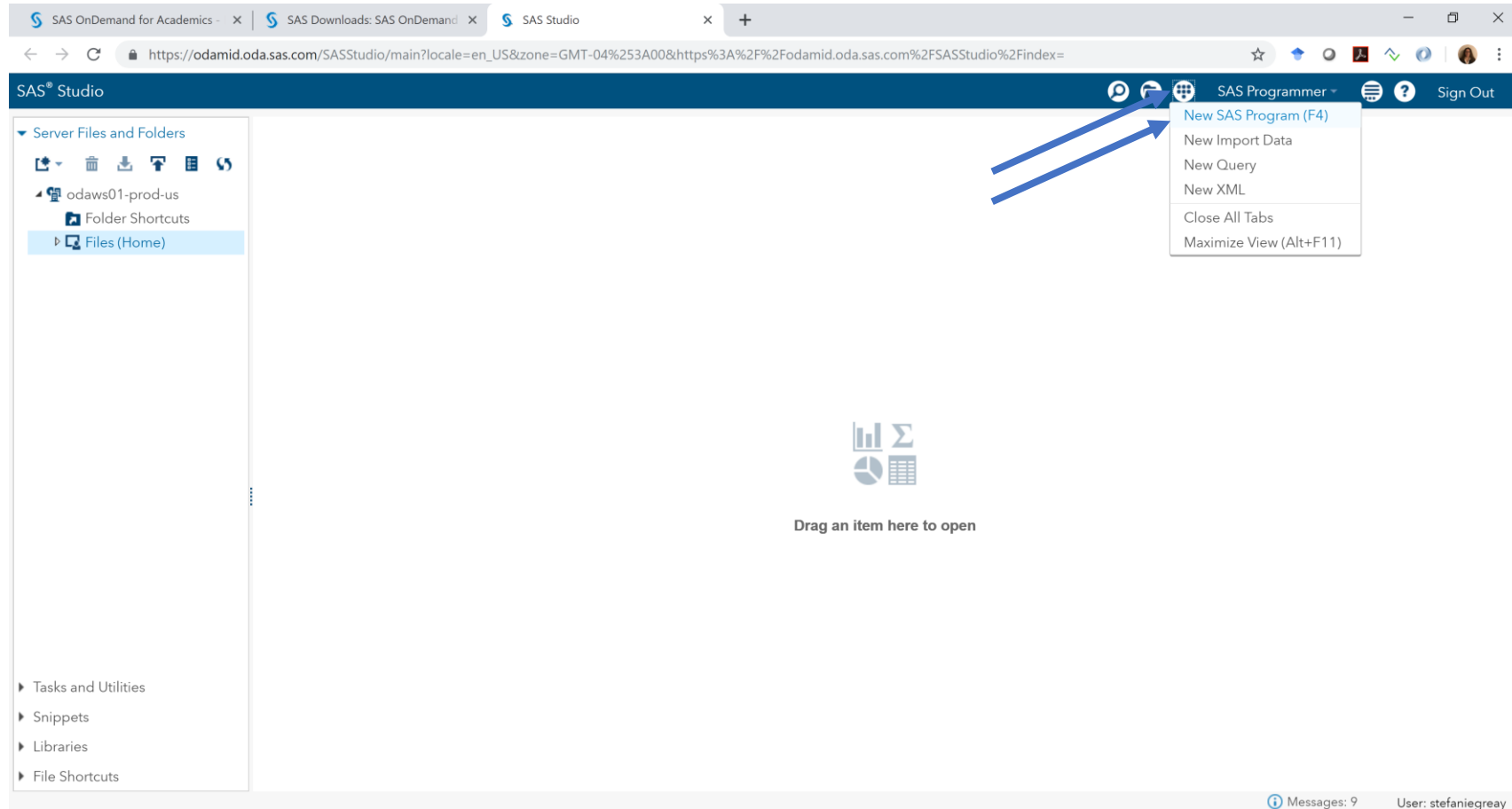
When you run the code, you will see the dataset and summary in the output data window and can verify its success.

The screenshot displays the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'cf_ANLT5010_W4_AirBNB.csv' selected. The main window shows the 'Program 1' tab with a sub-tab for '*cf_ANLT5010_W4_AirBNB'. The 'FILE INFORMATION' section displays the source file details: File name: cf_ANLT5010_W4_AirBNB.csv, Source location: /home/stefaniegreay, and End of line delimiter: Default. The 'OUTPUT DATA' section shows the SAS server: SASApp, Data set name: IMPORT, and Library: WORK. Below this, the 'RESULTS' tab is active, showing a 'Table of Contents' for 'The CONTENTS Procedure'. The table lists metadata for the dataset 'WORK.IMPORT'.

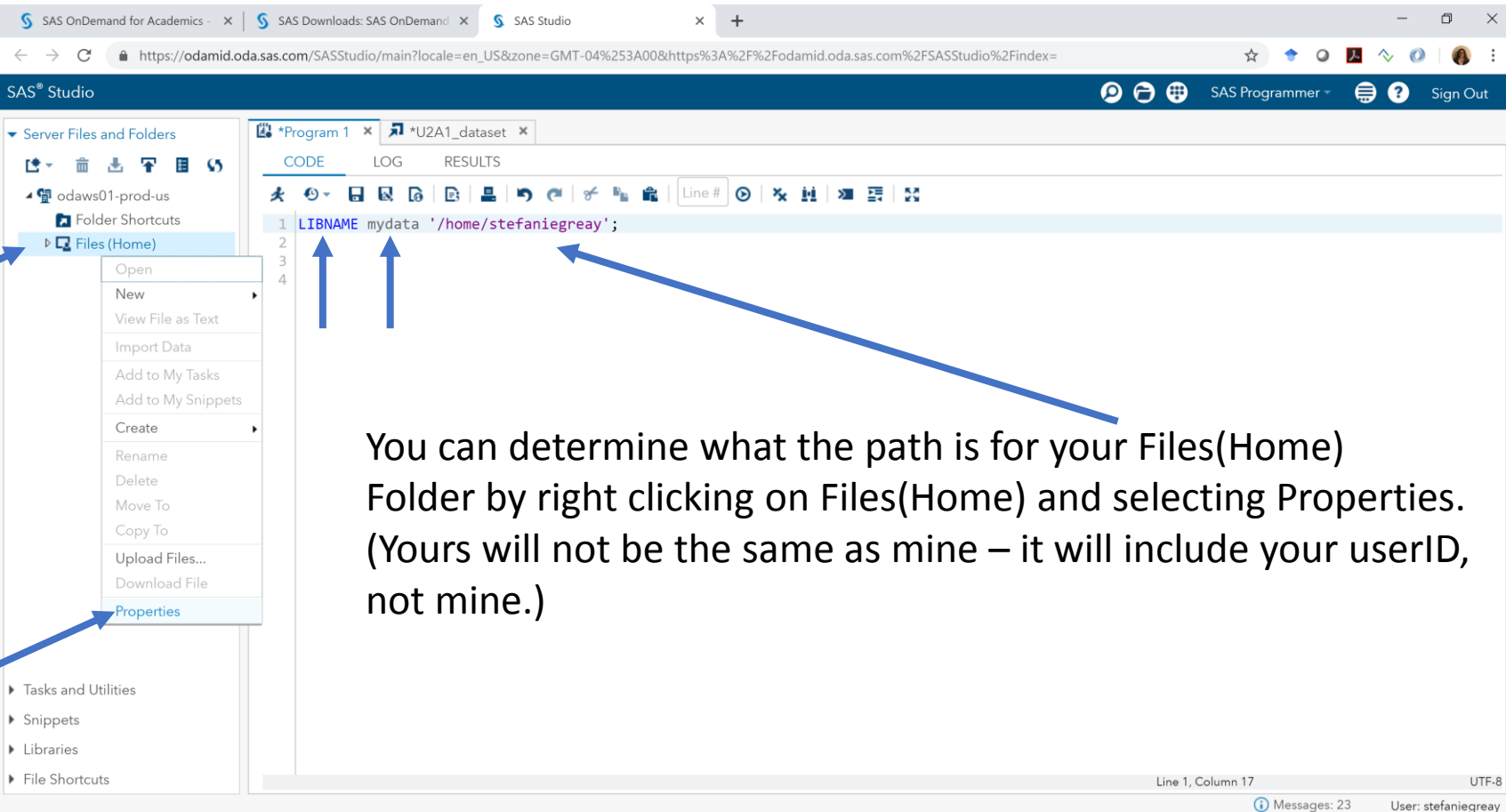
The CONTENTS Procedure			
Data Set Name	WORK.IMPORT	Observations	1746
Member Type	DATA	Variables	16
Engine	V9	Indexes	0
Created	11/07/2020 15:47:24	Observation Length	200
Last Modified	11/07/2020 15:47:24	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			



To get started with the SAS portion of the Week 4 Assignment 1 assignment, start a new SAS program.



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

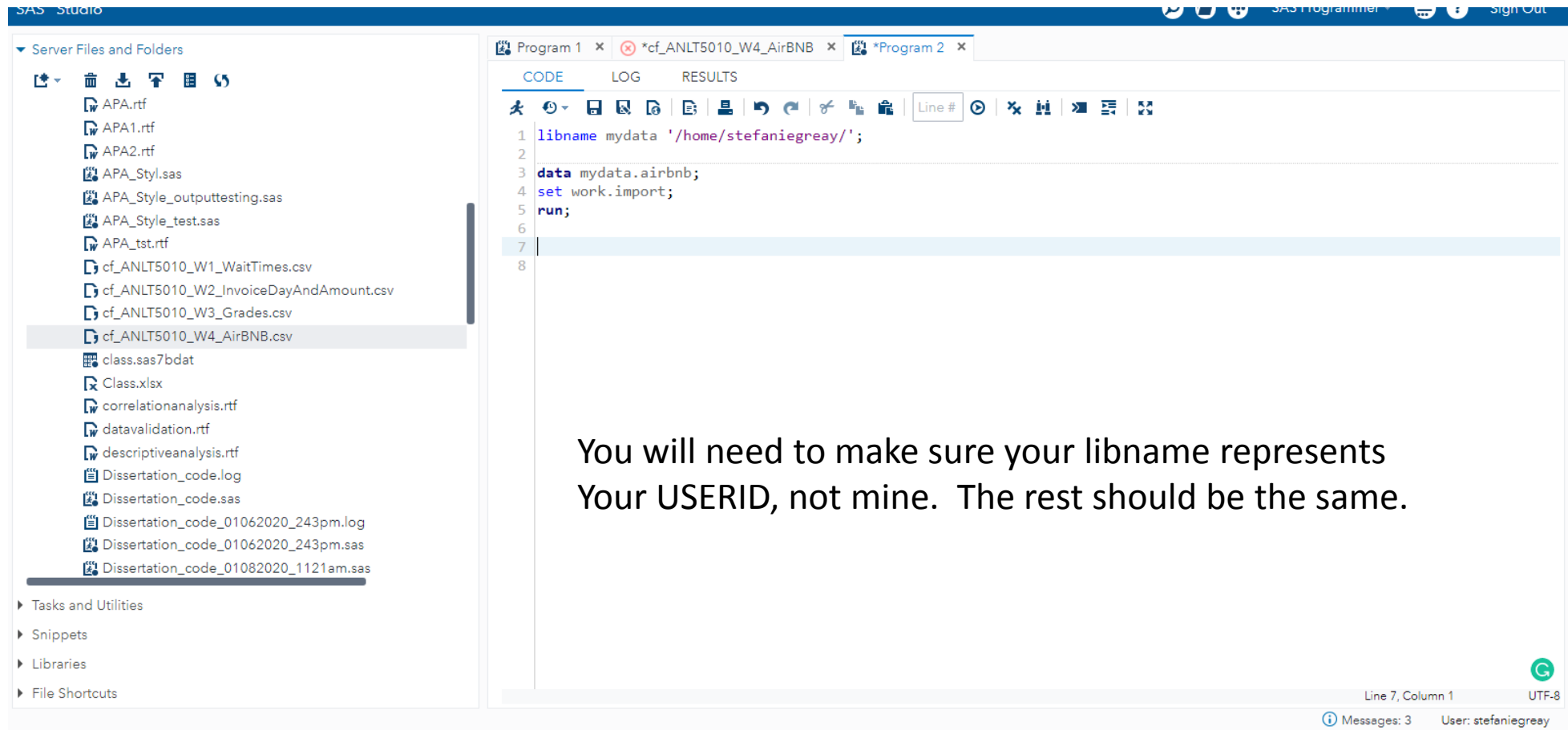


The screenshot shows the SAS Studio web interface. On the left, the 'Server Files and Folders' pane displays a tree structure with 'Files (Home)' selected. A right-click context menu is open over 'Files (Home)', with the 'Properties' option highlighted at the bottom. A blue arrow points from the 'Properties' option to the explanatory text. In the center, the 'CODE' pane shows a SAS program with the following line of code: `LIBNAME mydata '/home/stefaniegreay';`. Two blue arrows point from the text 'You can determine what the path is for your Files(Home) Folder...' to the words 'mydata' and the path string in the code. A third blue arrow points from the text '(Yours will not be the same as mine – it will include your userID, not mine.)' to the path string. The bottom status bar indicates 'Line 1, Column 17' and 'UTF-8'.

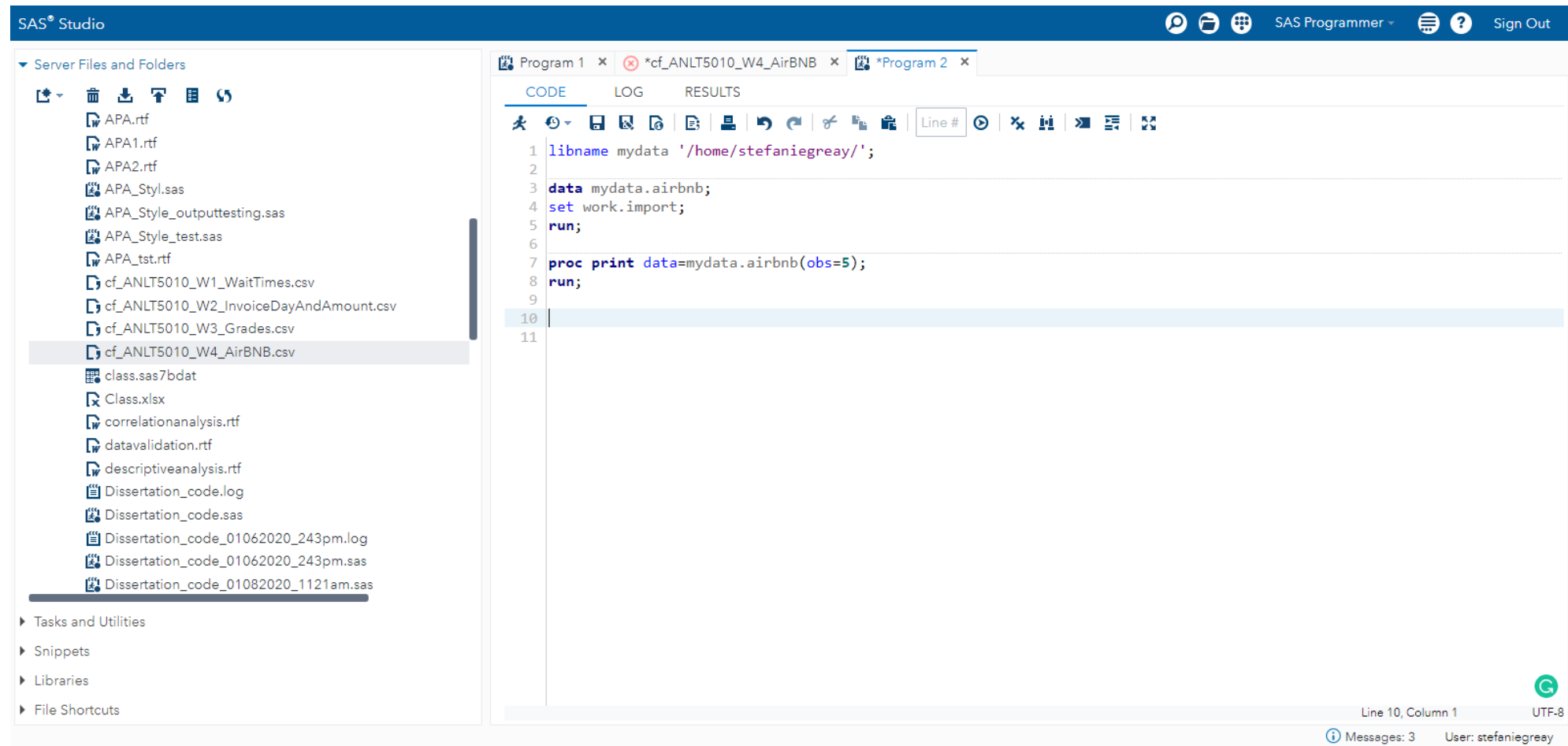
You can determine what the path is for your Files(Home) Folder by right clicking on Files(Home) and selecting Properties. (Yours will not be the same as mine – it will include your userID, not mine.)



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



Use a Proc Print statement to print the first 5 observations



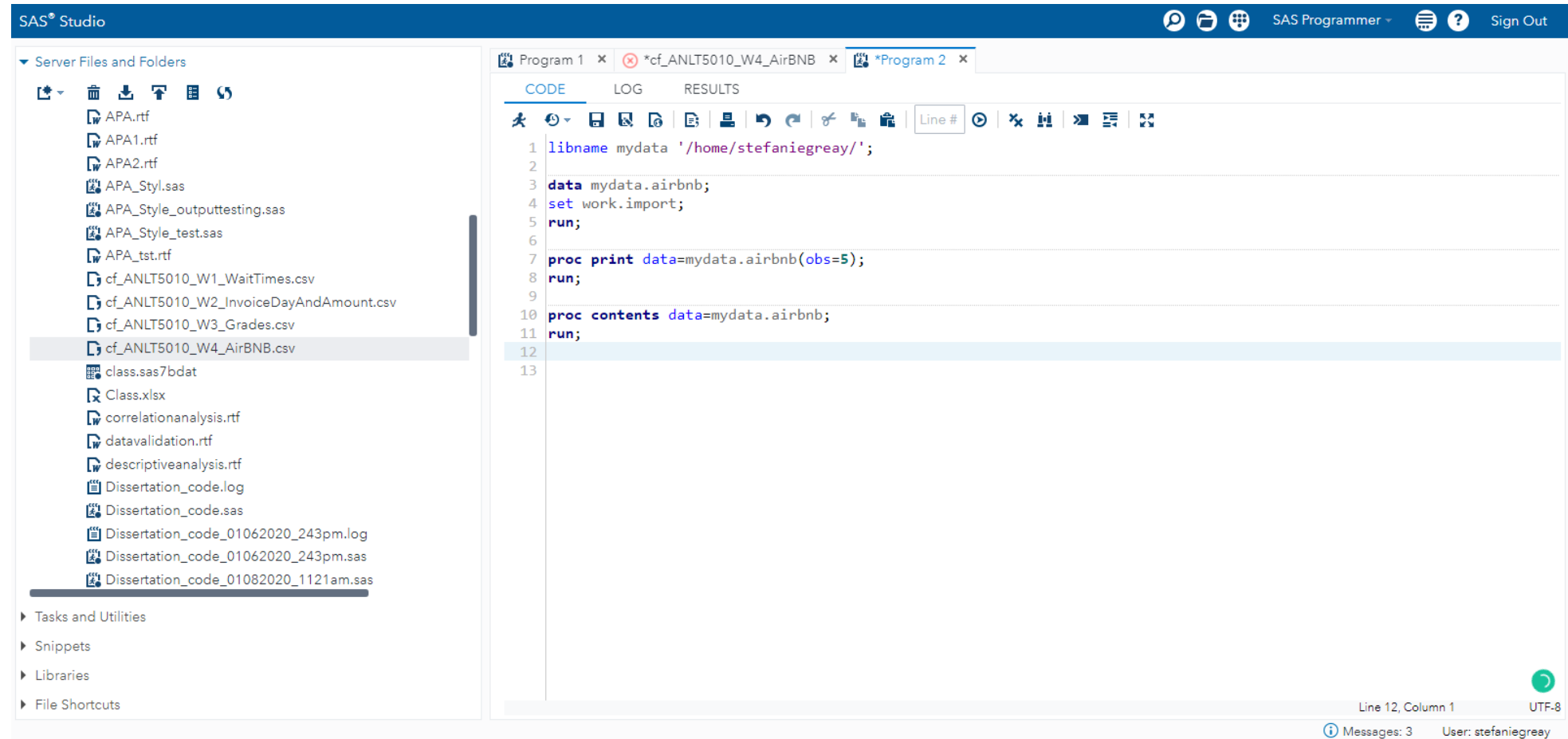
The screenshot displays the SAS Studio web interface. On the left, the 'Server Files and Folders' pane shows a list of files, with 'cf_ANLT5010_W4_AirBNB.csv' selected. The main editor pane shows a SAS program with the following code:

```
1 libname mydata '/home/stefaniegreay/';
2
3 data mydata.airbnb;
4 set work.import;
5 run;
6
7 proc print data=mydata.airbnb(obs=5);
8 run;
9
10
11
```

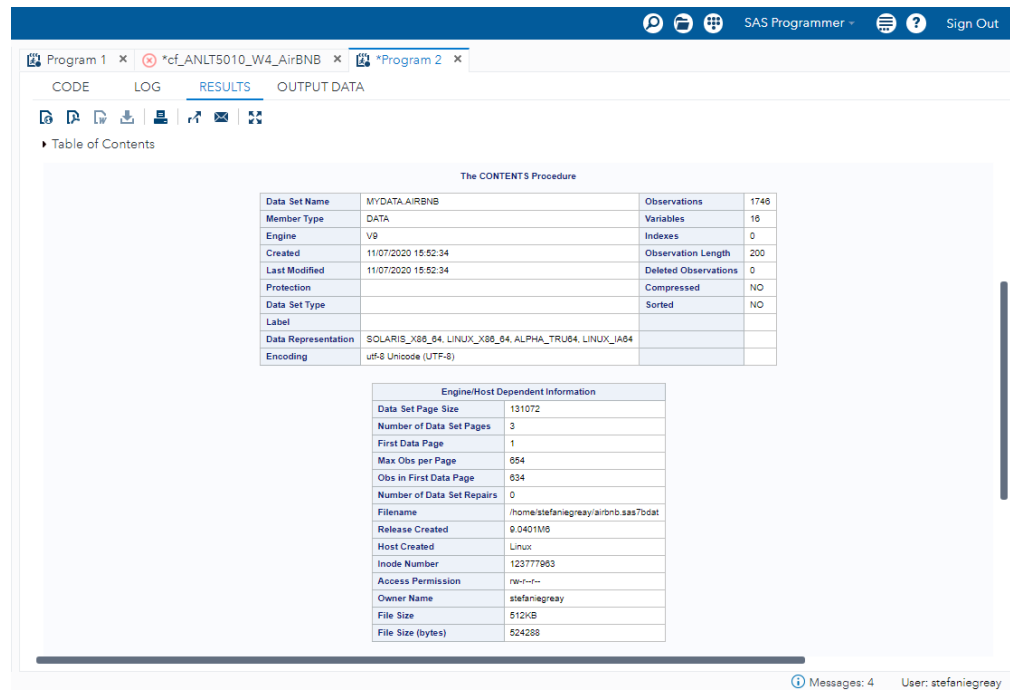
The status bar at the bottom indicates 'Line 10, Column 1' and 'UTF-8' encoding. The user is identified as 'stefaniegreay'.



Use a Proc Contents statement to look at the contents of the dataset



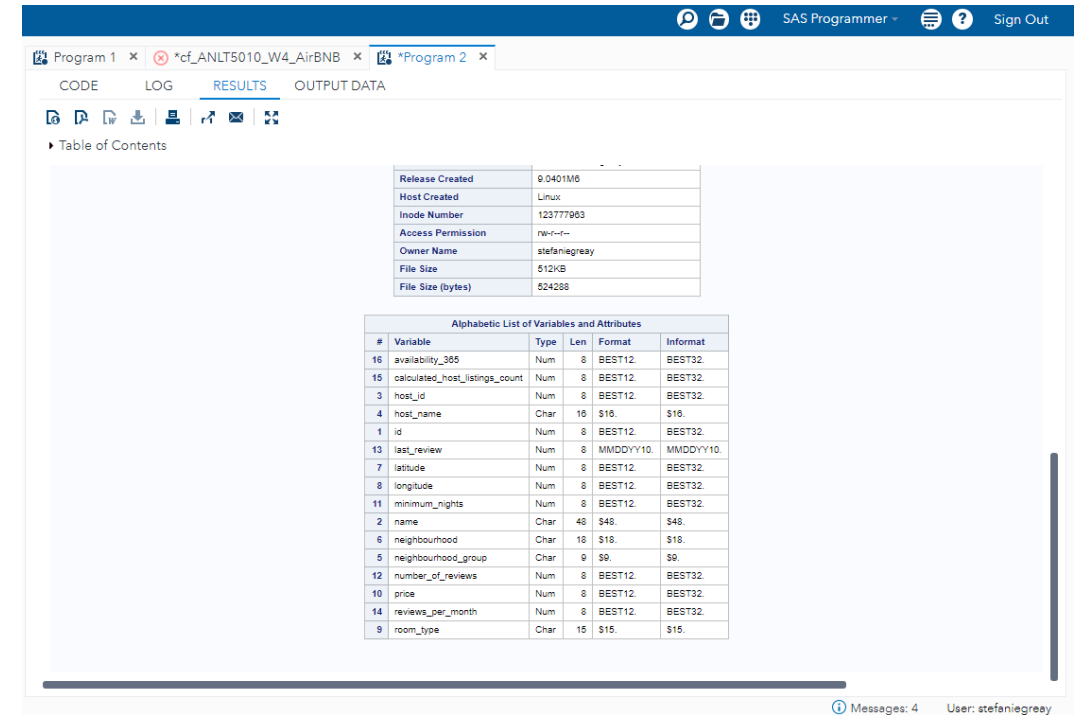
Review the Results pane to select a qualitative and quantitative variable to include in next steps



The screenshot shows the SAS Results pane with the CONTENTS procedure output for the dataset MYDATA.AIRBNB. The output is divided into two sections: a summary table and a detailed Engine/Host Dependent Information table.

The CONTENTS Procedure			
Data Set Name	MYDATA.AIRBNB	Observations	1746
Member Type	DATA	Variables	16
Engine	V9	Indexes	0
Created	11/07/2020 15:52:34	Observation Length	200
Last Modified	11/07/2020 15:52:34	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRUB4, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	3
First Data Page	1
Max Obs per Page	654
Obs in First Data Page	634
Number of Data Set Repairs	0
Filename	/home/stefaniegray/airbnb.sas7bdat
Release Created	9/04/2016
Host Created	Linux
Inode Number	123777963
Access Permission	rw-rw-r--
Owner Name	stefaniegray
File Size	512KB
File Size (bytes)	524288

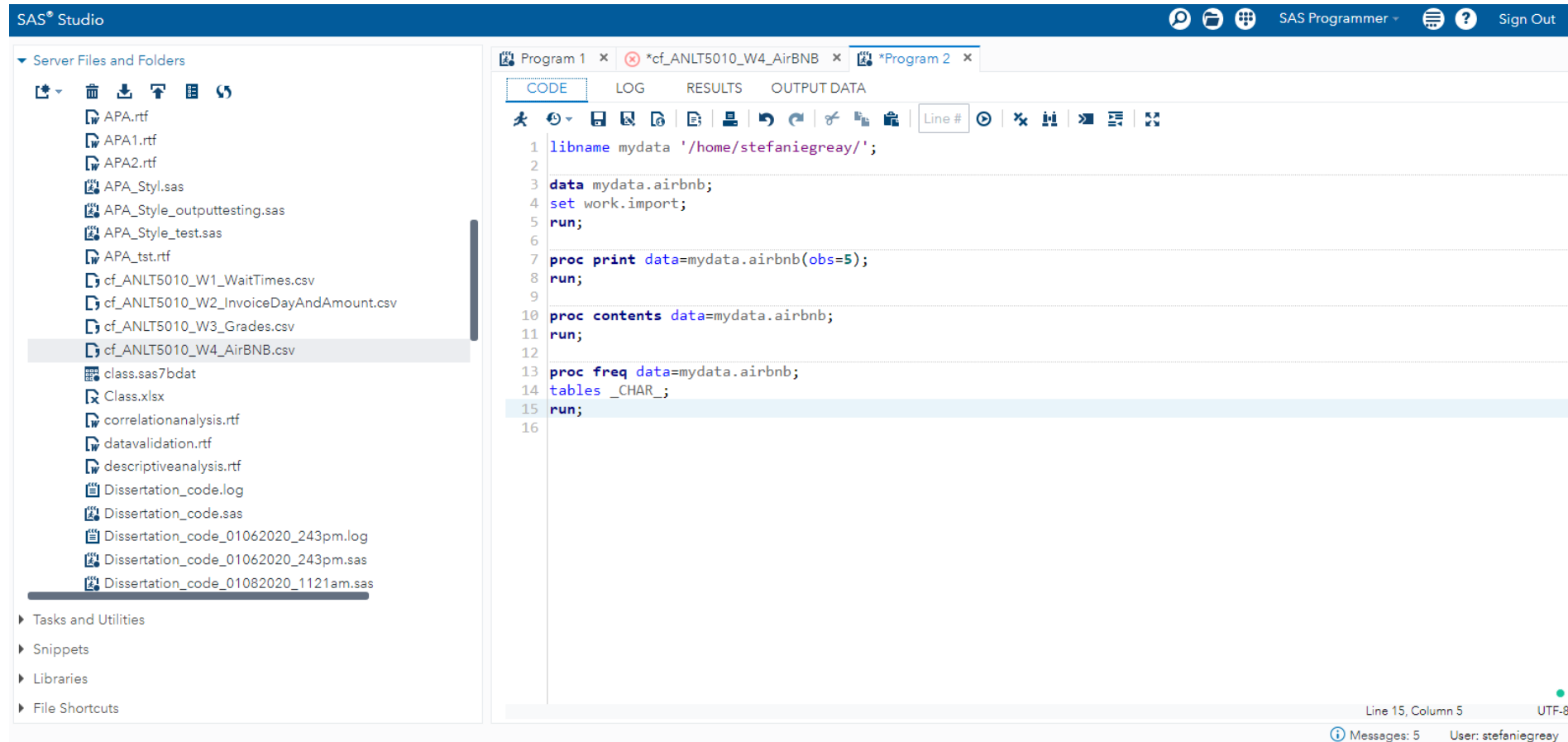


The screenshot shows the SAS Results pane with the Alphabetic List of Variables and Attributes. The output is a table listing all variables in the dataset, including their variable number, name, type, length, format, and informat.

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
16	availability_365	Num	8	BEST12.	BEST32.
15	calculated_host_listings_count	Num	8	BEST12.	BEST32.
3	host_id	Num	8	BEST12.	BEST32.
4	host_name	Char	10	\$10.	\$10.
1	id	Num	8	BEST12.	BEST32.
13	last_review	Num	8	MMDDYY10.	MMDDYY10.
7	latitude	Num	8	BEST12.	BEST32.
8	longitude	Num	8	BEST12.	BEST32.
11	minimum_nights	Num	8	BEST12.	BEST32.
2	name	Char	48	\$48.	\$48.
6	neighbourhood	Char	18	\$18.	\$18.
5	neighbourhood_group	Char	9	\$9.	\$9.
12	number_of_reviews	Num	8	BEST12.	BEST32.
10	price	Num	8	BEST12.	BEST32.
14	reviews_per_month	Num	8	BEST12.	BEST32.
9	room_type	Char	15	\$15.	\$15.



Use a Proc Freq statement to create a frequency table for a qualitative (categorical) variable(s) in the dataset



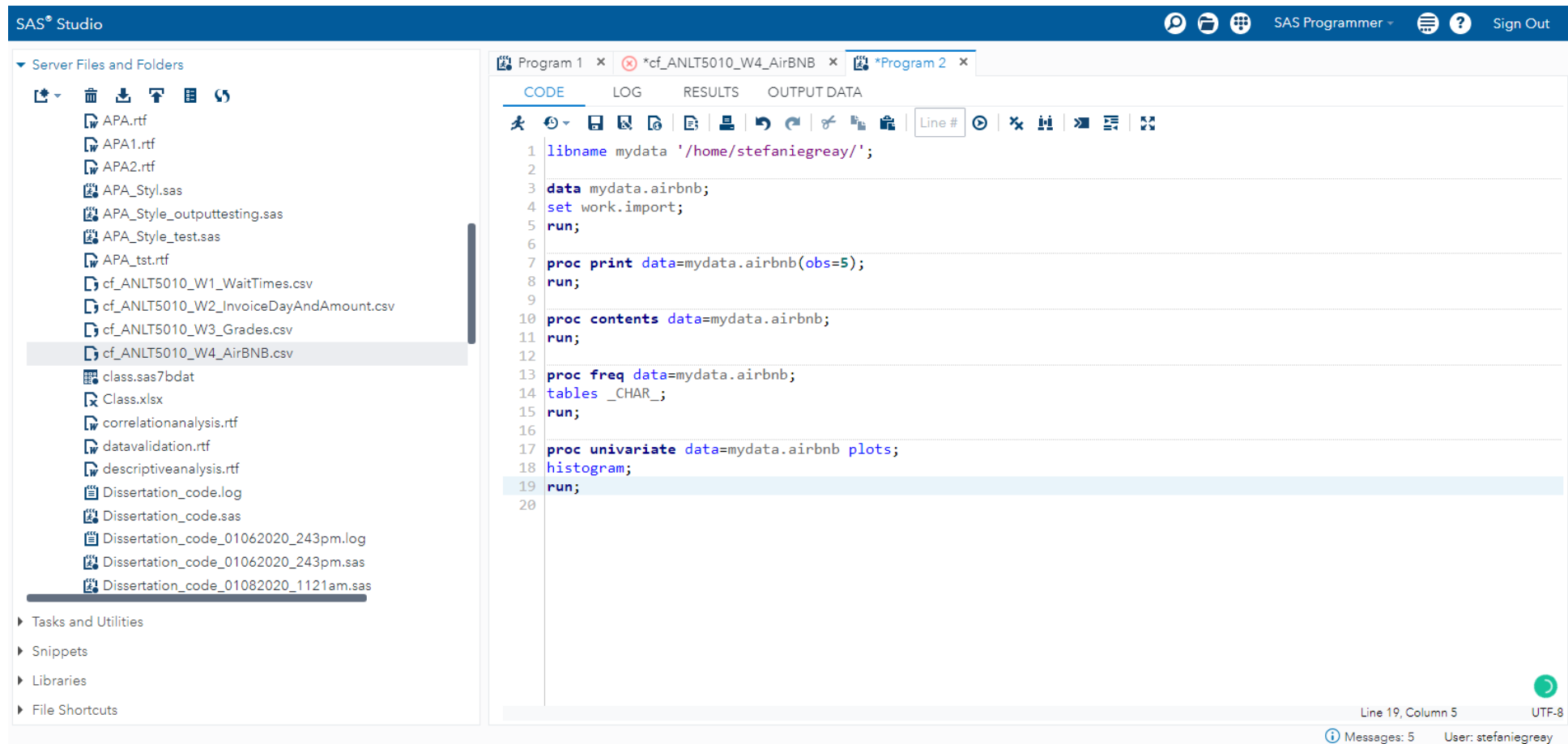
The screenshot displays the SAS Studio environment. On the left, the 'Server Files and Folders' pane lists various files, with 'cf_ANLT5010_W4_AirBNB.csv' selected. The main editor window shows a SAS program with the following code:

```
1 libname mydata '/home/stefaniegreay/';  
2  
3 data mydata.airbnb;  
4 set work.import;  
5 run;  
6  
7 proc print data=mydata.airbnb(obs=5);  
8 run;  
9  
10 proc contents data=mydata.airbnb;  
11 run;  
12  
13 proc freq data=mydata.airbnb;  
14 tables _CHAR_;  
15 run;  
16
```

The status bar at the bottom indicates 'Line 15, Column 5' and 'UTF-8'. The bottom right corner shows 'Messages: 5' and 'User: stefaniegreay'.



Use a Proc Univariate to summarize the quantitative variable(s) in the dataset



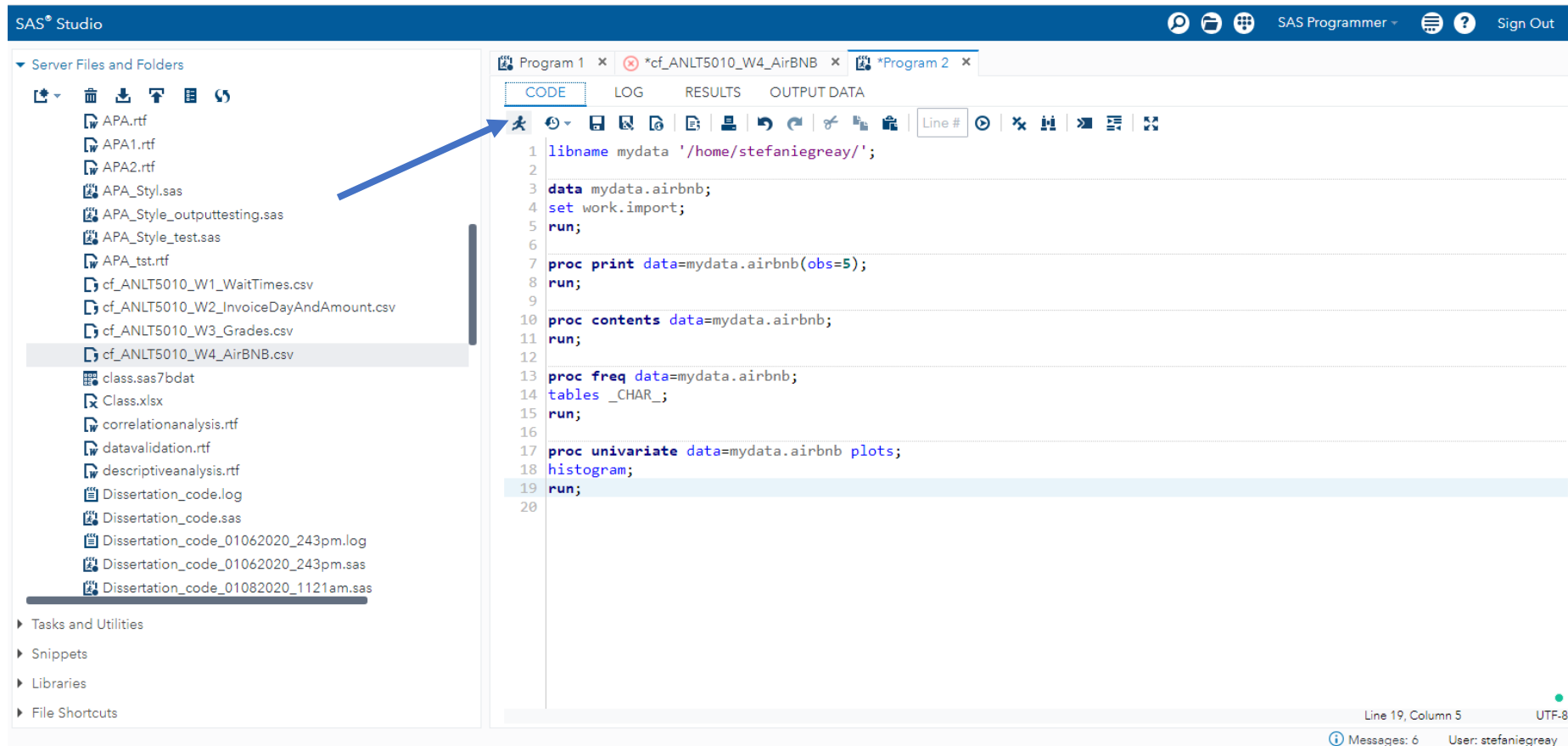
The screenshot displays the SAS Studio interface. On the left, the 'Server Files and Folders' pane shows a list of files, with 'cf_ANLT5010_W4_AirBNB.csv' selected. The main editor window shows a SAS program with the following code:

```
1 libname mydata '/home/stefaniegreay/';
2
3 data mydata.airbnb;
4 set work.import;
5 run;
6
7 proc print data=mydata.airbnb(obs=5);
8 run;
9
10 proc contents data=mydata.airbnb;
11 run;
12
13 proc freq data=mydata.airbnb;
14 tables _CHAR_;
15 run;
16
17 proc univariate data=mydata.airbnb plots;
18 histogram;
19 run;
20
```

The status bar at the bottom indicates 'Line 19, Column 5' and 'UTF-8'. The bottom right corner shows 'Messages: 5' and 'User: stefaniegreay'.



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



Full code (basic)

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

```
data mydata.airbnb;
```

```
set work.import;
```

```
run;
```

```
proc print data=mydata.airbnb(obs=5);
```

```
run;
```

```
proc contents data=mydata.airbnb;
```

```
run;
```

```
proc freq data=mydata.airbnb;
```

```
tables _CHAR_;
```

```
run;
```

```
proc univariate data=mydata.airbnb plots;
```

```
histogram;
```

```
run;
```

