ANLT5020 – Unit 6 Assignment 1 Tutorial

SAS Studio

© Stefanie G. Reay, MS, PhD, Capella University, 2021

Instructions

Complete the following tasks:

- Create a SAS dataset containing this data using a SET statement.
- Clarify the labels on the MOS, DOS, YOS, and DRG by using a LABEL statement to give the columns more meaningful names than the three-letter variable names.
- Use a SAS procedure to reformat the Charges variable to show dollars and cents, including a dollar sign.
- Change the length of the LastName variable to 36 characters using a LENGTH statement to allow for longer last names that may flow through this field/variable in future data.
- Combine the MOS, DOS, and YOS fields into a single date field, manipulate that field so that SAS understands it as a date, and use a FORMAT statement to format the new combined date field like this: 02/14/2016.
- Create a list of at least five best practices for managing big data based on your evaluation of the assigned readings and your work in SAS. Provide a description for each best practice. Be sure to address privacy, security, and ethical concerns in your recommendations.

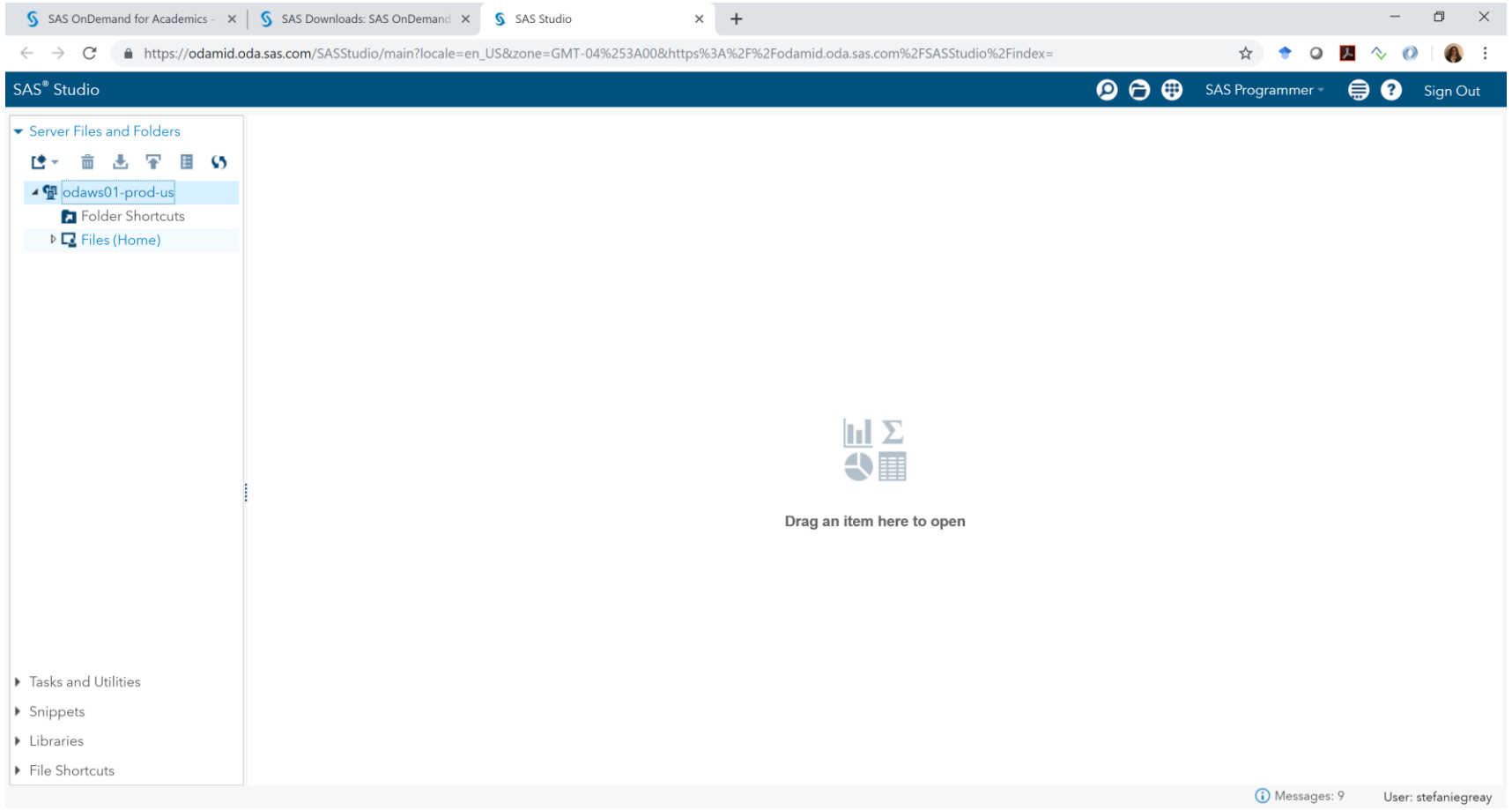
© Stefanie G. Reay, MS, PhD, Capella University, 2021

Dataset

- Download the Claims.txt file from the course datasets zip file or from the Unit 6 Welcome announcement in the course announcements.

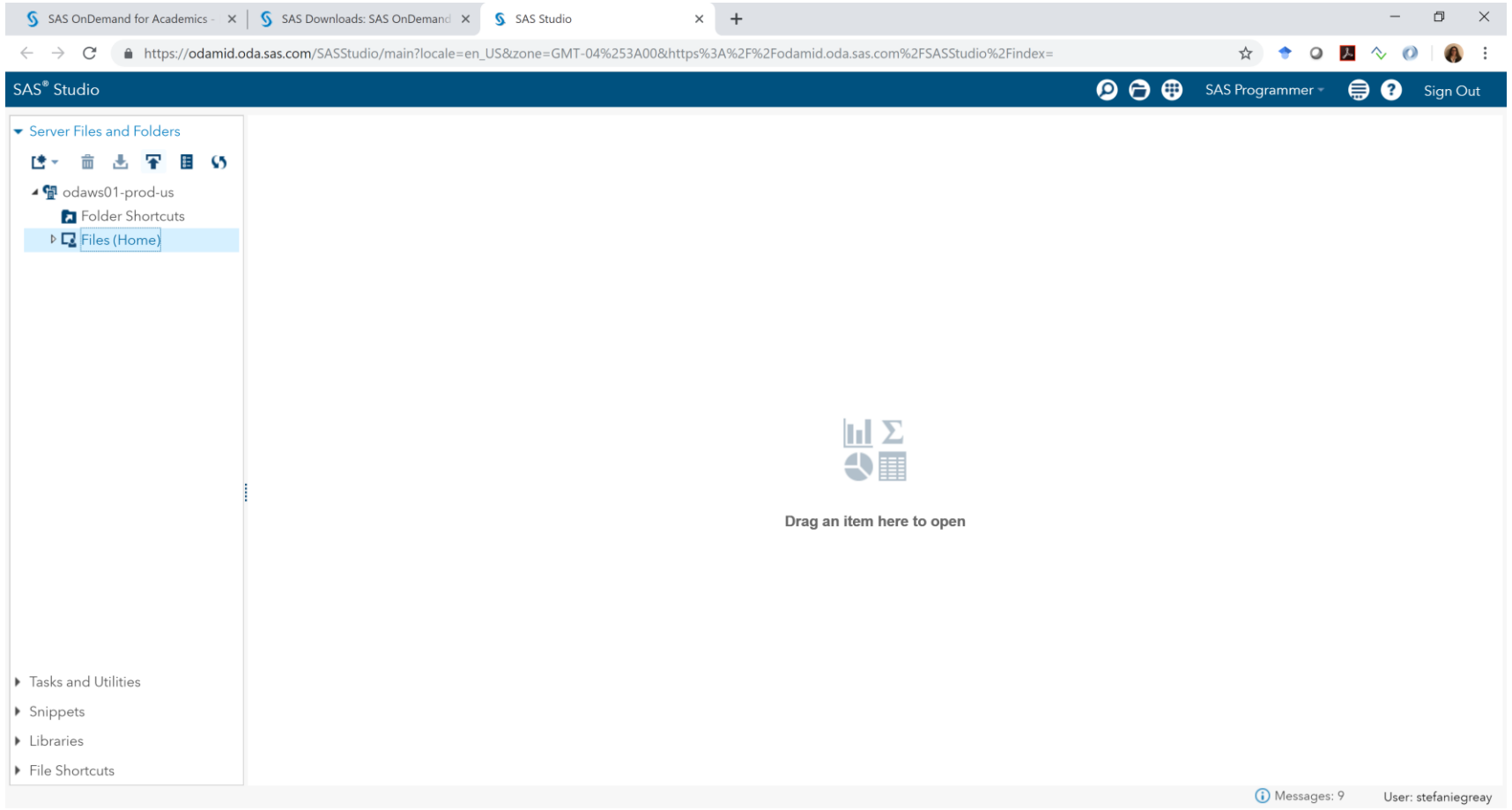
© Stefanie G. Reay, MS, PhD, Capella University, 2021

Click on Files(Home)



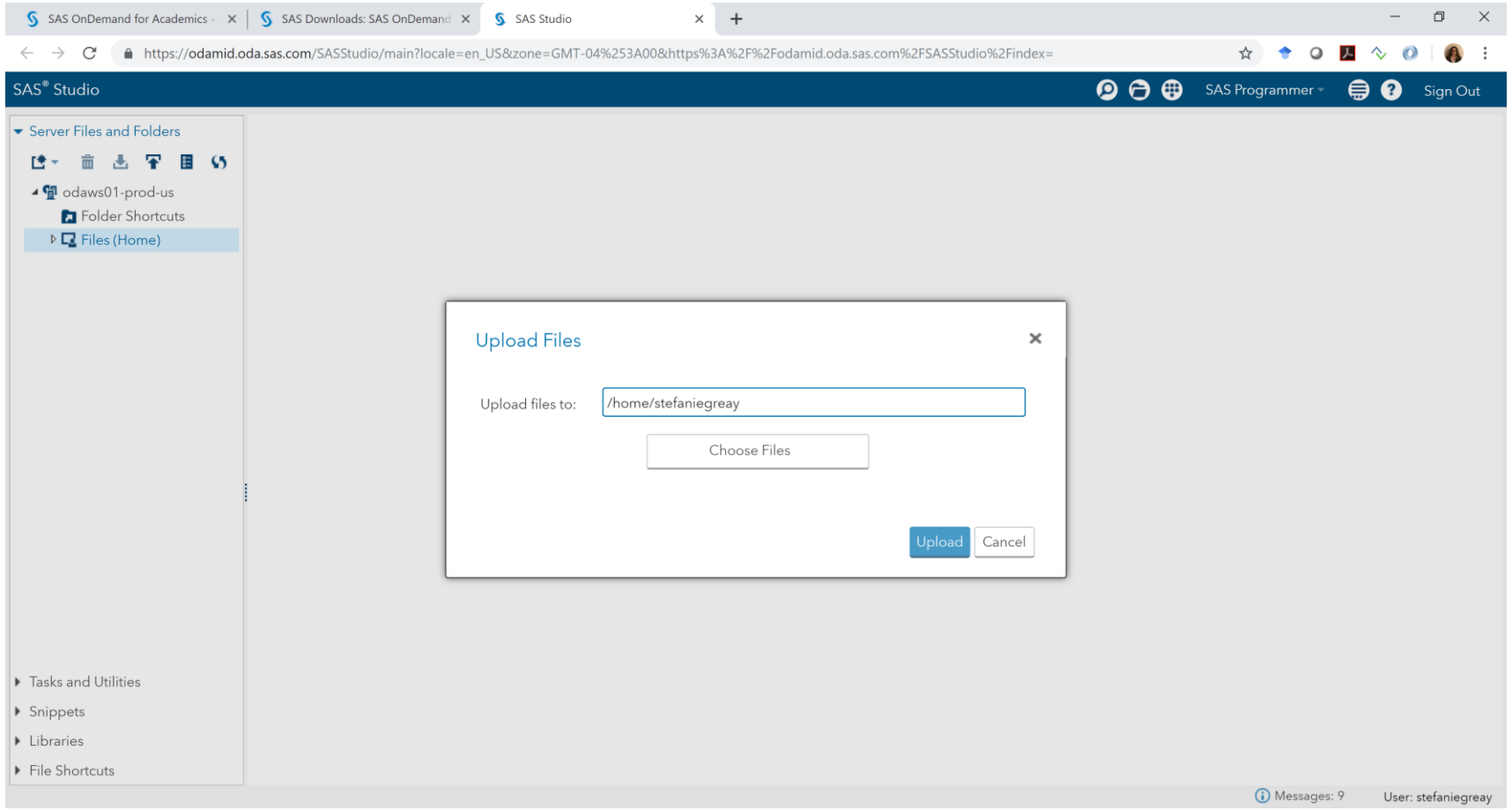
© Stefanie G. Reay, MS, PhD, Capella University, 2021

The Upload button will display in dark blue



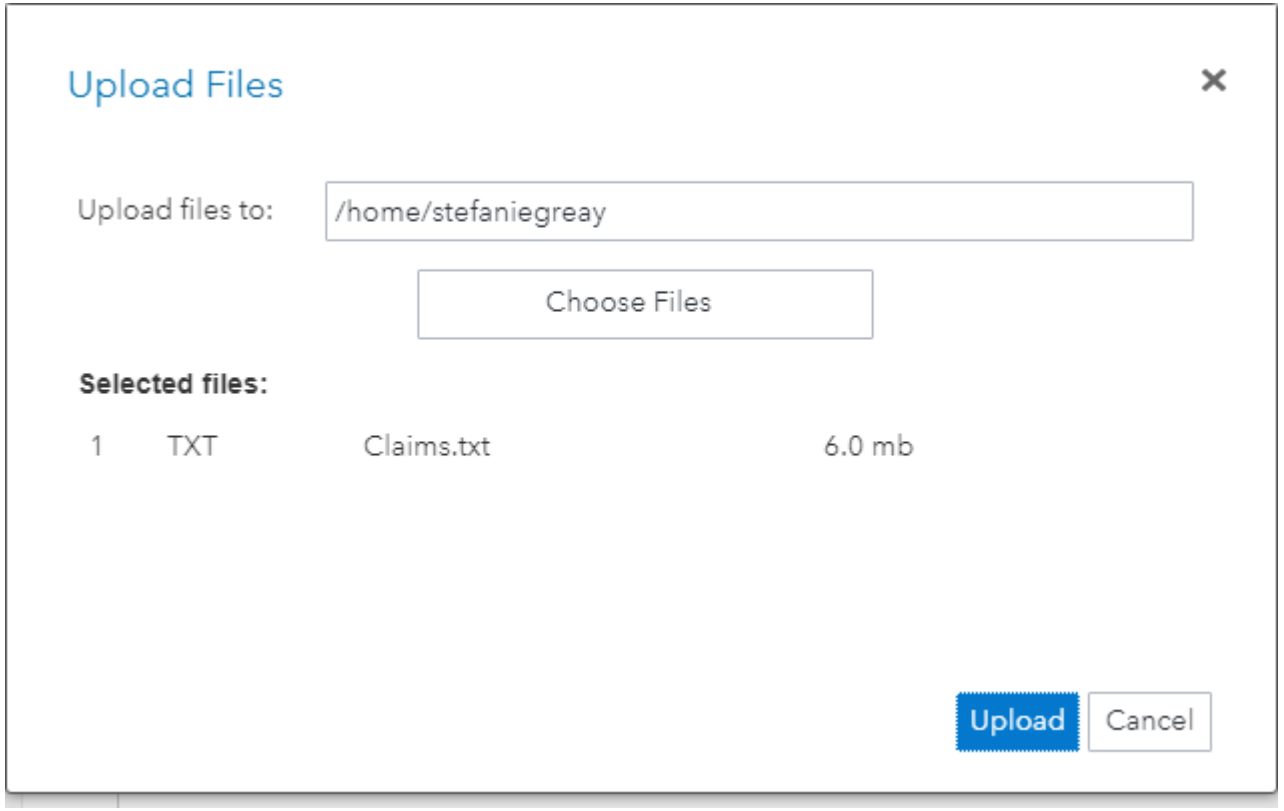
© Stefanie G. Reay, MS, PhD, Capella University, 2021

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



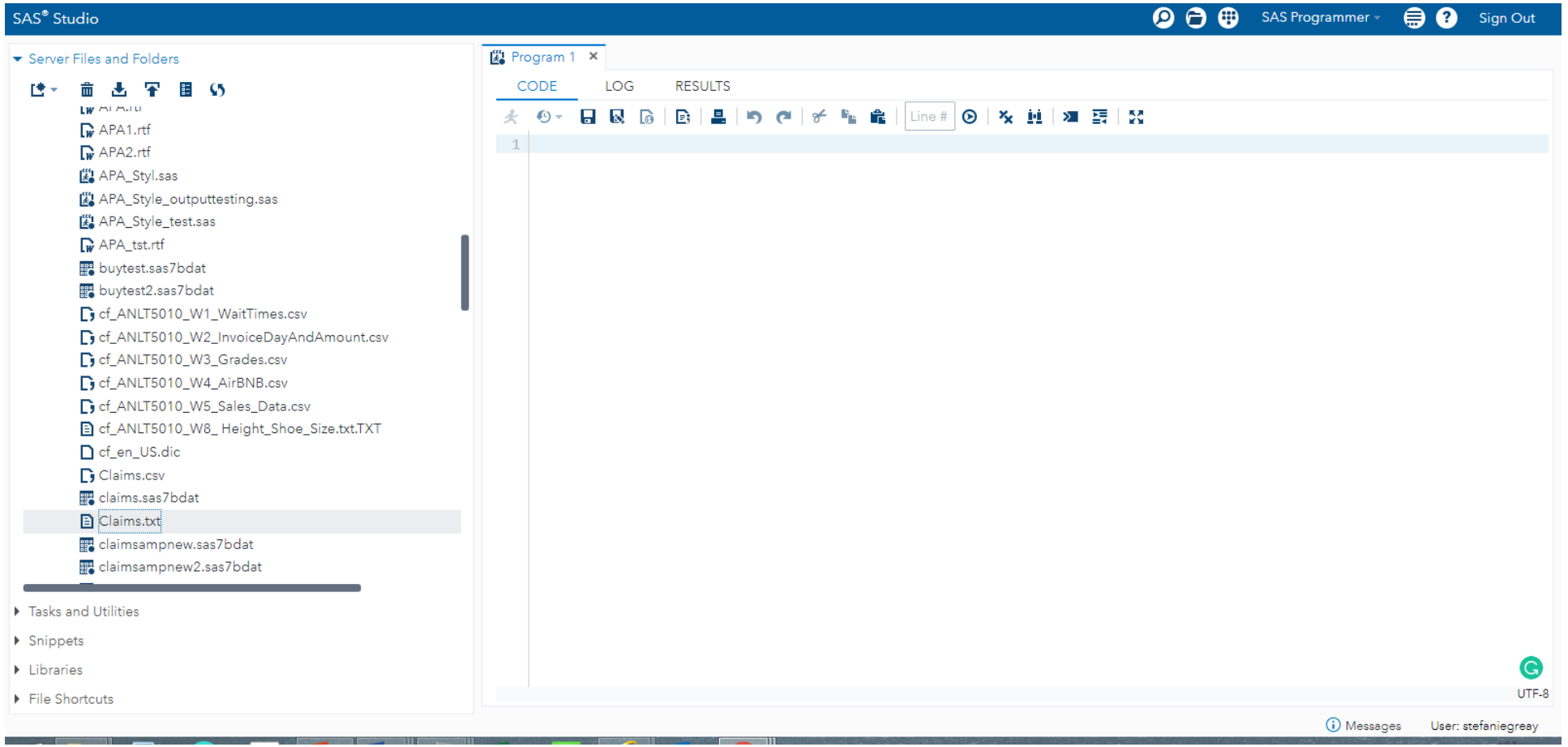
© Stefanie G. Reay, MS, PhD, Capella University, 2021

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



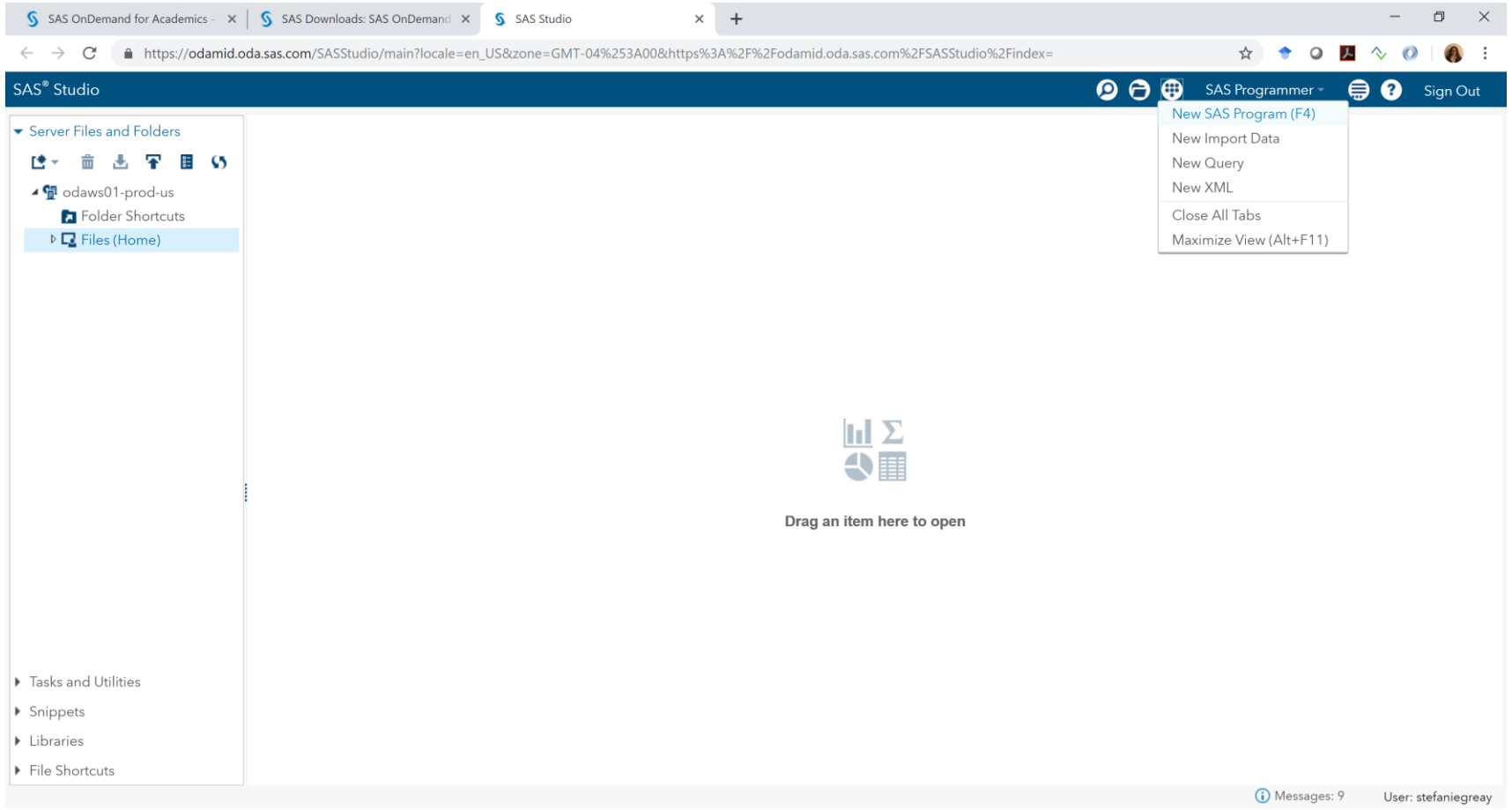
© Stefanie G. Reay, MS, PhD, Capella University, 2021

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



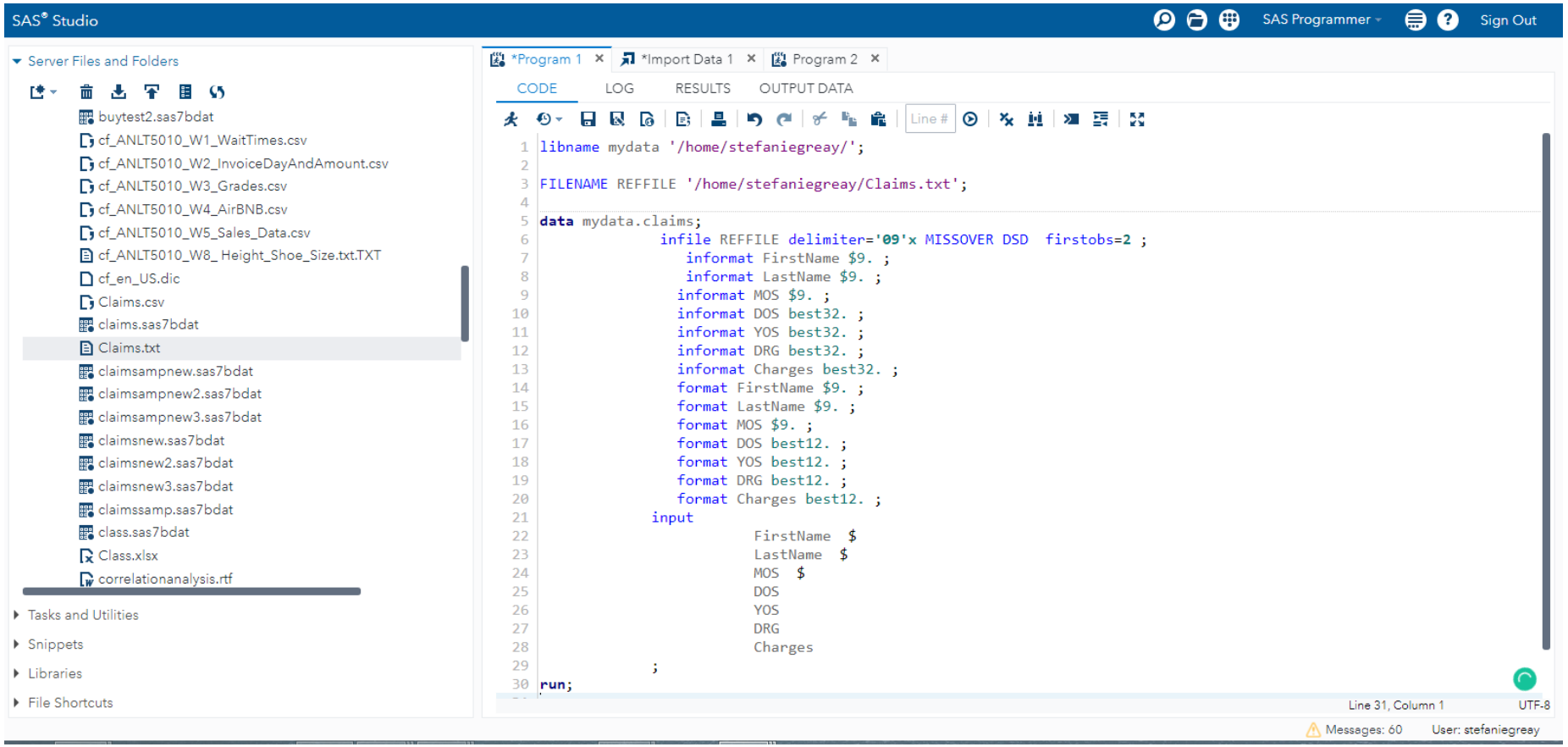
© Stefanie G. Reay, MS, PhD, Capella University, 2021

To get started with the SAS portion of the assignment, start a new SAS program.



© Stefanie G. Reay, MS, PhD, Capella University, 2021

Import the dataset into a SAS dataset format using a SAS data step, as requested in the assignment. (You will need to change the location of your folder and the file you uploaded.)



© Stefanie G. Reay, MS, PhD, Capella University, 2021

Code to import data in tab-delimited text file using data step.

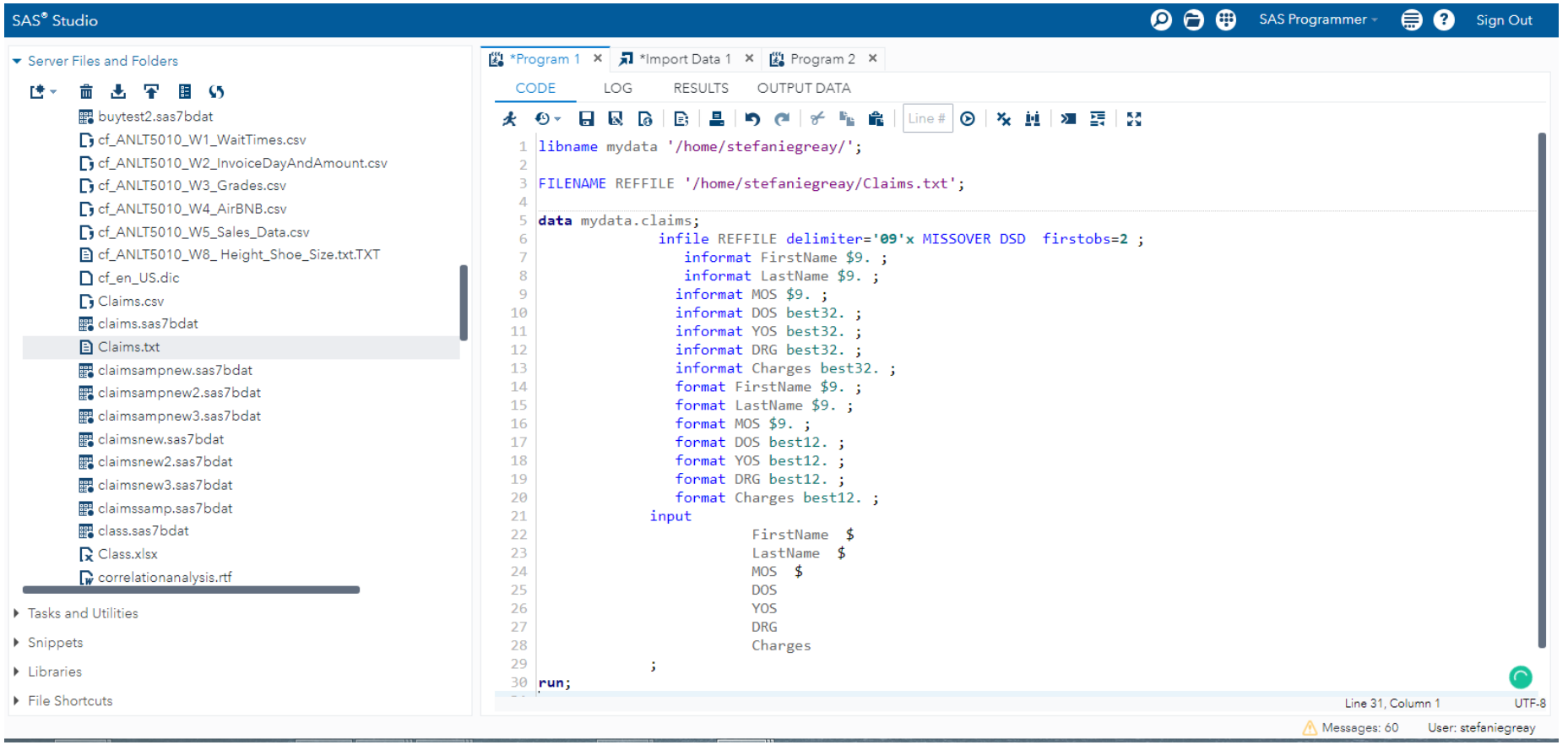
```
libname mydata '/home/stefaniegreay/';

FILENAME REFFILE '/home/stefaniegreay/Claims.txt';

data mydata.claims;
             infile REFFILE delimiter='09'x MISSOVER DSD  firstobs=2 ;
                informat FirstName $9. ;
                informat LastName $9. ;
               informat MOS $9. ;
               informat DOS best32. ;
               informat YOS best32. ;
               informat DRG best32. ;
               informat Charges best32. ;
               format FirstName $9. ;
               format LastName $9. ;
               format MOS $9. ;
               format DOS best12. ;
               format YOS best12. ;
               format DRG best12. ;
               format Charges best12. ;
            input
                        FirstName  $
                        LastName  $
                        MOS  $
                        DOS
                        YOS
                        DRG
                        Charges
            ;
run;
```

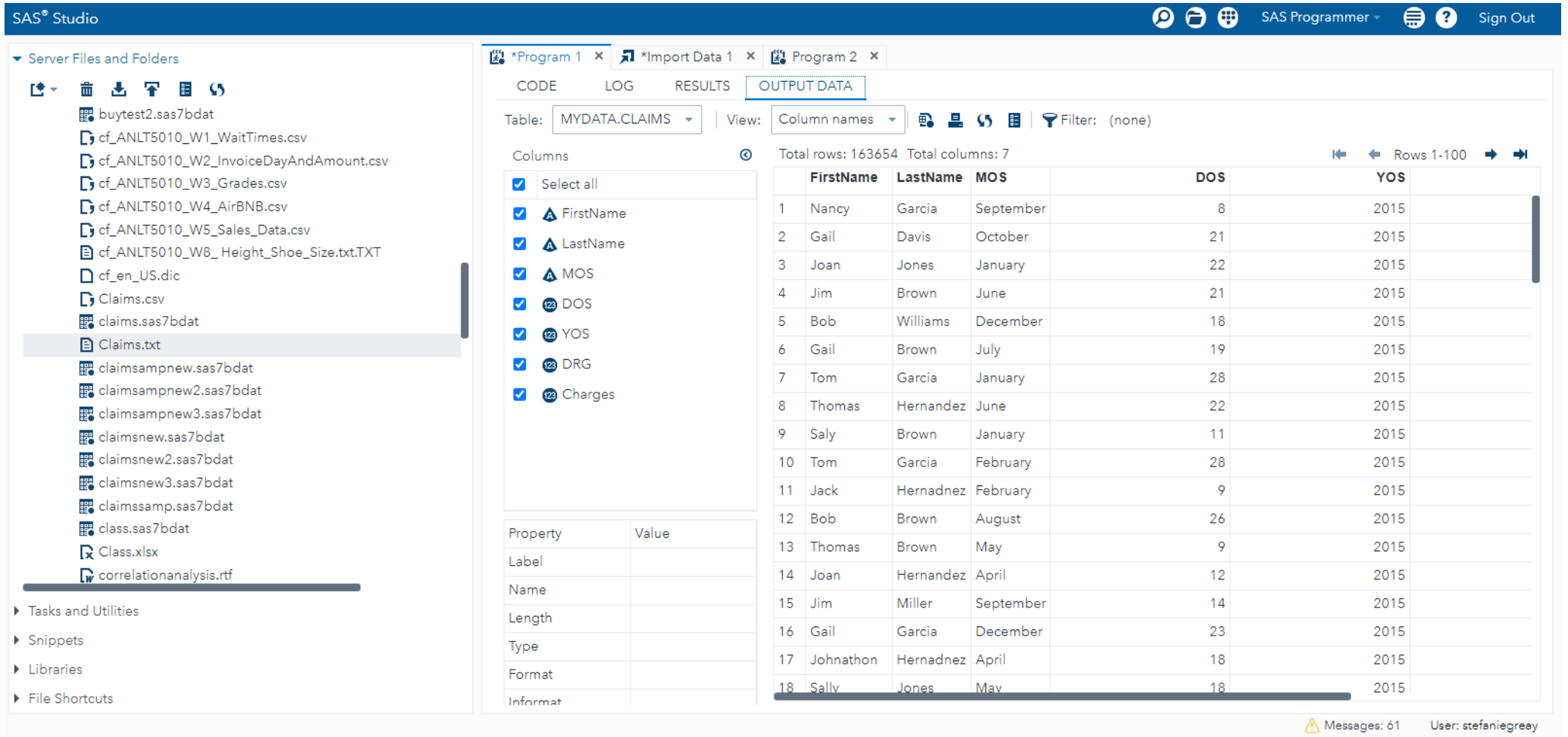
© Stefanie G. Reay, MS, PhD, Capella University, 2021

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



© Stefanie G. Reay, MS, PhD, Capella University, 2021

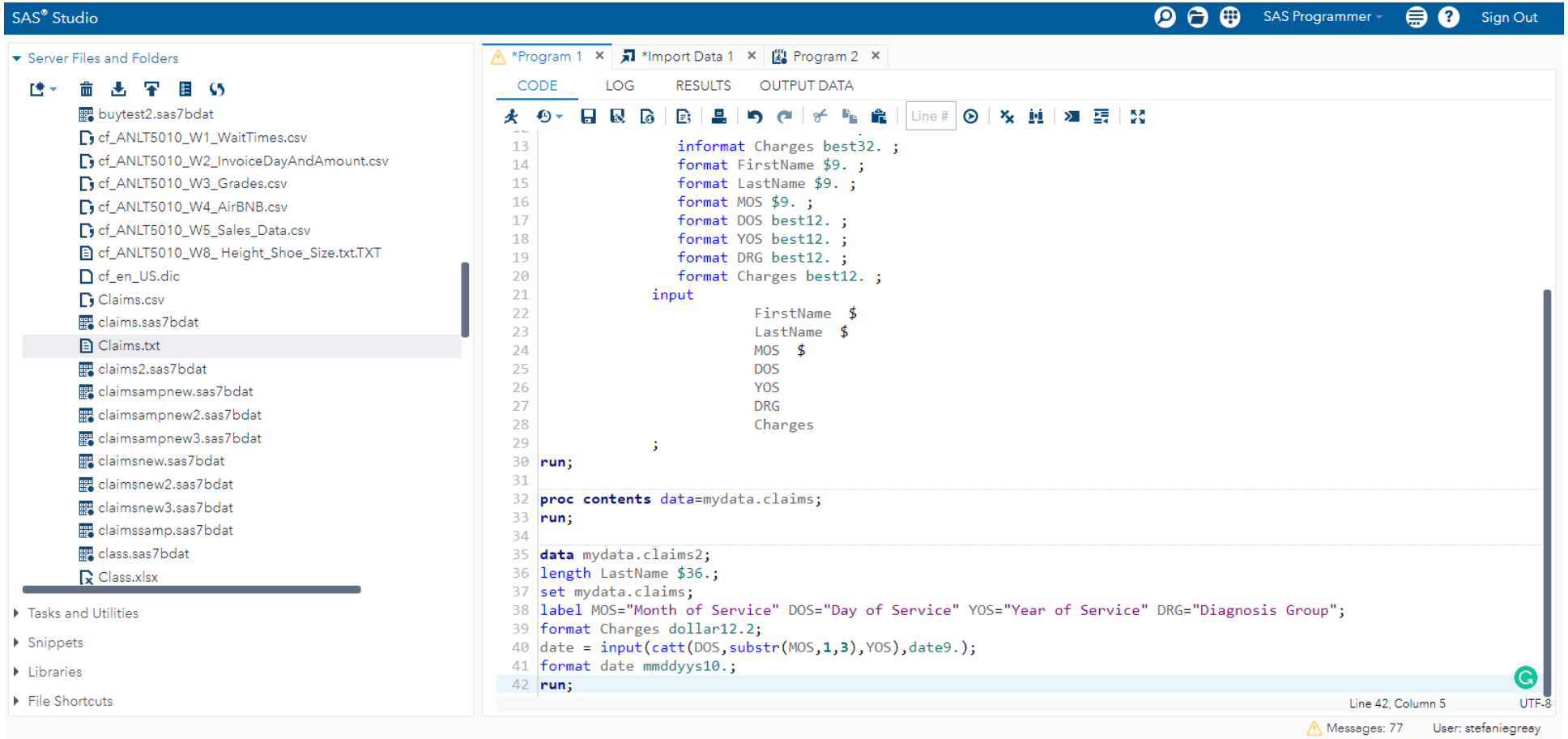
When you run the code, you will see the dataset in the output data window and can verify its success.



	FirstName	LastName	MOS	DOS	YOS
1	Nancy	Garcia	September	8	2015
2	Gail	Davis	October	21	2015
3	Joan	Jones	January	22	2015
4	Jim	Brown	June	21	2015
5	Bob	Williams	December	18	2015
6	Gail	Brown	July	19	2015
7	Tom	Garcia	January	28	2015
8	Thomas	Hernandez	June	22	2015
9	Saly	Brown	January	11	2015
10	Tom	Garcia	February	28	2015
11	Jack	Hernadnez	February	9	2015
12	Bob	Brown	August	26	2015
13	Thomas	Brown	May	9	2015
14	Joan	Hernandez	April	12	2015
15	Jim	Miller	September	14	2015
16	Gail	Garcia	December	23	2015
17	Johnathon	Hernadnez	April	18	2015
18	Sally	Jones	May	18	2015

© Stefanie G. Reay, MS, PhD, Capella University, 2021

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



© Stefanie G. Reay, MS, PhD, Capella University, 2021

Sample Code for completing requested data manipulation tasks

```
data mydata.claims2;
length LastName $36.;
set mydata.claims;
label MOS="Month of Service" DOS="Day of Service" YOS="Year of Service" DRG="Diagnosis Group";
format Charges dollar12.2;
date = input(catt(DOS,substr(MOS,1,3),YOS),date9.);
format date mmddyys10.;
run;
```

© Stefanie G. Reay, MS, PhD, Capella University, 2021