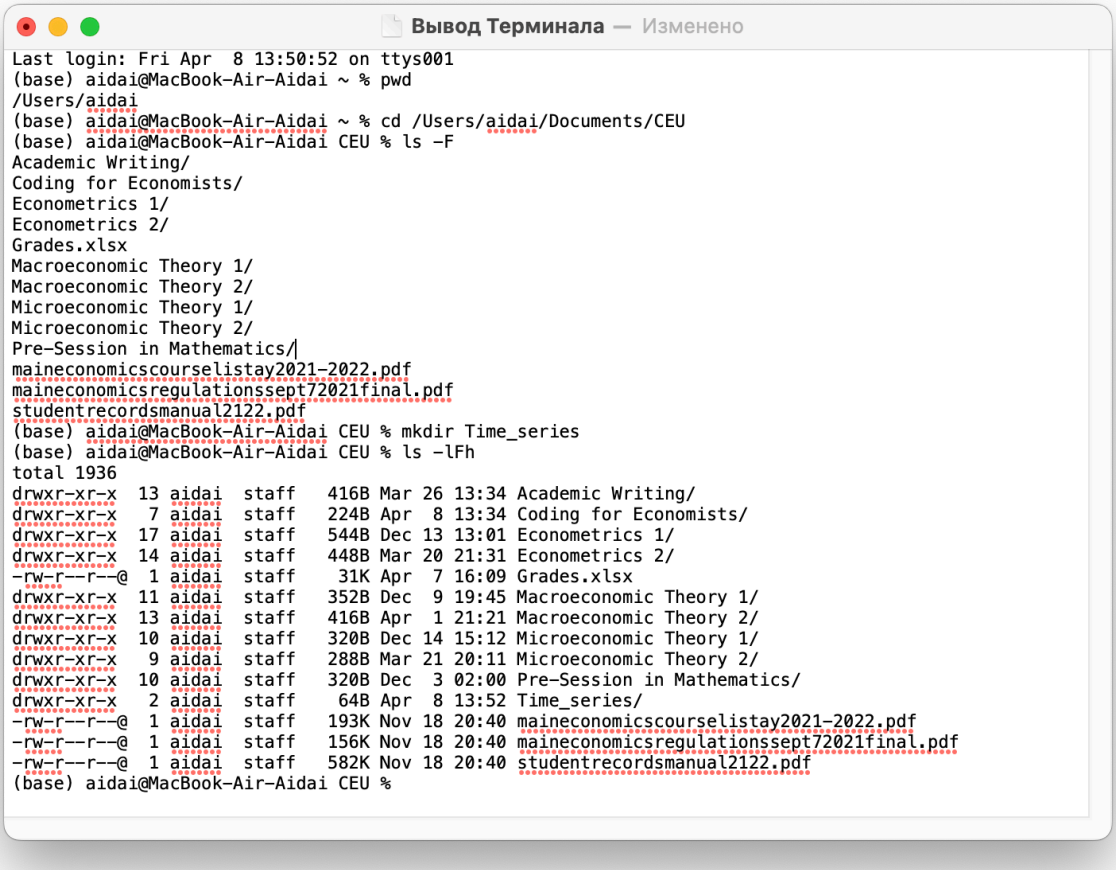
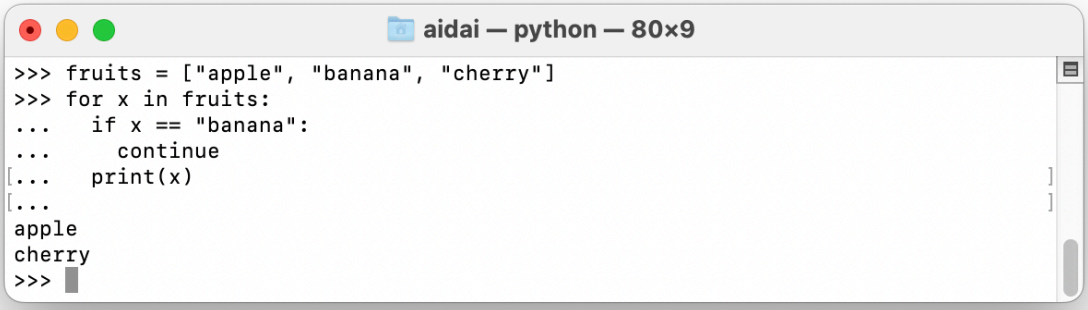
ASSIGNMENT

1. **Understand folder structure. Perform operations on files in different folders**

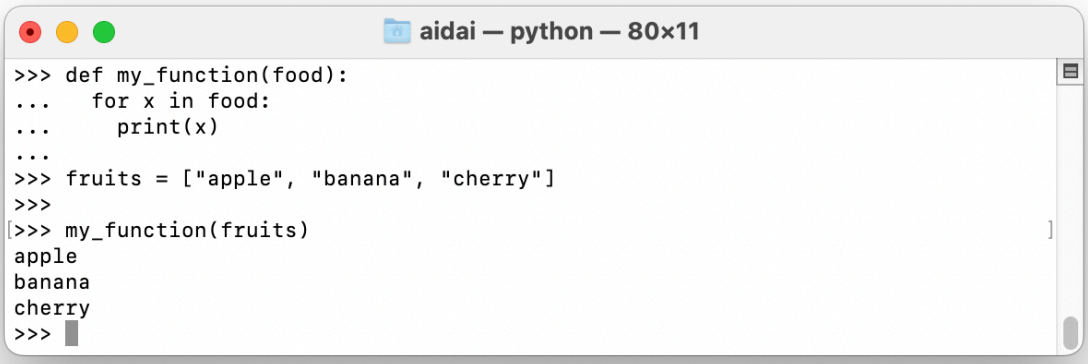
pwd displays the path of the current working directory.
cd changes the current working directory to the specified drive and directory.
ls -F lists the names of files in the specified directory in a short form
mkdir creates a new directory.



2. **Automate repeating tasks using Python “for” loops.**



3. **Break up work into smaller components using Python functions.**



4. **Use Python “lists” and “dictionaries” appropriately. Demonstrate one of the two.**

```
(base) aidai@MacBook-Air-Aidai ~ % python
Python 3.9.7 (default, Sep 16 2021, 08:50:36)
[Clang 10.0.0 ] :: Anaconda, Inc. on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> student_info={}
>>> student_info['A']={'maths':82,'english':90,'science':80}
>>> student_info['B']={'maths':93,'english':80,'science':90}
>>> student_info['C']={'maths':92,'english':86,'science':95}
>>> print(student_info)
{'A': {'maths': 82, 'english': 90, 'science': 80}, 'B': {'maths': 93, 'english':
 80, 'science': 90}, 'C': {'maths': 92, 'english': 86, 'science': 95}}
>>>
```

5. **Automate repeating tasks using Stata "for" loops.**

forvalues is used for loop over consecutive values.

6. **Break up work into smaller components using Stata .do files.**
7. **Read .csv data in in Stata.**

import delimited reads into memory a text-delimited file from disk

8. **Fix common data quality errors in Stata (for example, string vs number, missing value).**

A new variable is generated to get rid of missing values.

9. **Aggregate, reshape, and combine data for analysis in Python or Stata. Demonstrate at least one of these data manipulations.**

collapse converts the dataset in memory into a dataset of means, sums, medians, etc. clist must refer to numeric variables exclusively.

10. **Prepare a sample for analysis by filtering observations and variables and creating transformations of variables. Demonstrate all three.**

use loads into memory a Stata-format dataset previously saved by save.
keep works the same way as drop, except that you specify the variables or observations to be kept rather than the variables or observations to be deleted.
label data attaches a label (up to 80 characters) to the dataset in memory. Dataset labels are displayed when you use the dataset and when you describe it.

11. **Save data in Stata.**

save stores the dataset currently in memory on disk under the name filename. If filename is not specified, the name under which the data were last known to Stata (c(filename)) is used. If filename is specified without an extension, .dta is used. If filename contains embedded spaces, remember to enclose it in double quotes.

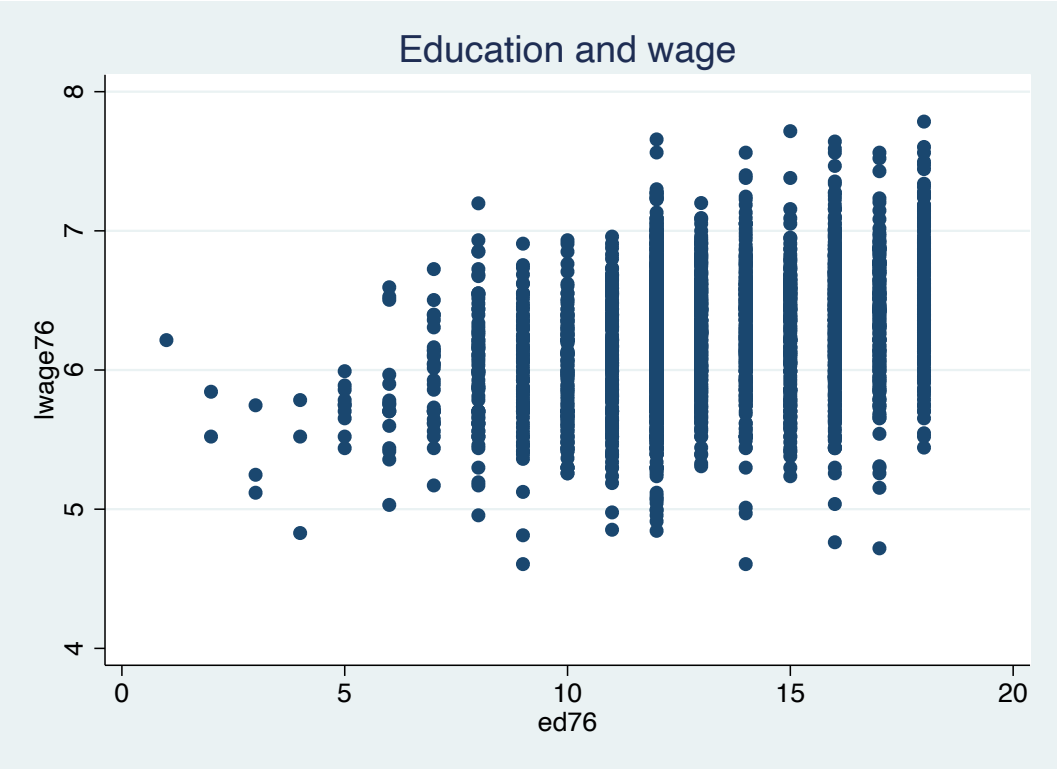
12. **Run ordinary least squares regression in Stata.**

regress performs ordinary least-squares linear regression. regress can also perform weighted

estimation, compute robust and cluster–robust standard errors, and adjust results for complex survey designs.

13. Create a graph (of any type) in Stata.

twoway is a family of plots, all of which fit on numeric y and x scales



14. Save regression tables and graphs as files. Demonstrate both.

outreg2 arranges regression, summary, and tabulation into an illustrative table.

15. Install a Stata package. (Can be the same as we already did in class).

ssc installs and uninstall packages from SSC.