

SEARCH



RESOURCES



CONCEPTS



- ✓ 1. Lesson Overview
- ✓ 2. Analyzing with IPython
- ✓ 3. Scripting Your Analysis

Scripting Your Analysis

Being able to write and run scripts is invaluable for programming tasks and proj

You can write your code in a text editor and then run the file in your terminal. H
printing column names from the census income dataset. If you save your file as
shown on the right, you can run it as shown on the left. Make sure you are in the
saved this file in!

```
[junolee:~/udacity]$ ls
census_income_data.csv demo.py
[junolee:~/udacity]$ python demo.py
age
workclass
fnlwgt
education
education-num
marital-status
occupation
relationship
race
sex
capital-gain
capital-loss
hours-per-week
native-country
income
[junolee:~/udacity]$
```

```
demo.py
1 import pandas as pd
2
3 df = pd.read_csv('census_inco
4
5 for column in df.columns:
6     print(column)
7
```

Running a basic Python script

Ideally, you'd group your analysis into functions and run them in your main func
organize your code and generalize if possible.

```
[junolee:~/udacity]$ python demo.py
age
workclass
fnlwgt
education
education-num
marital-status
occupation
relationship
race
sex
capital-gain
capital-loss
hours-per-week
native-country
income
[junolee:~/udacity]$
```

```
demo.py
1 import pandas as pd
2
3
4 def load_data():
5     df = pd.read_csv('census_
6     return df
7
8
9 def print_columns(df):
10     for column in df.columns:
11         print(column)
12
13
14 if __name__ == '__main__':
15     df = load_data()
16     print_columns(df)
17
```

Functions help organize code

The script below creates a double histogram of ages for people with lower and h