Python for Data Analysis and Visualization (Fall 2024)

Claudia Carroll, TRIADS

Class 1 (Oct 14, 2024)

**Class 1 Homework Exercises**

1. Explain in your own words what the following code is doing:

file = open("gdp\_europe.csv").read()

print(file[2])

1. Write the code to print out all values in a column of index 8, of a csv file ‘test\_data.csv’, greater than 15.

df = pd.read\_csv('test\_data.csv')

filtered\_values = df[df.iloc[:, 8] > 15]

1. Write the code to extract all rows with the value 89 in the column of index 8 from a csv file titled ‘test\_data.csv’ and print those rows to a new csv file. You should copy the header from test\_data.csv to the new csv also.

filtered\_rows = df[df.iloc[:, 8] == 89]

filtered\_rows.to\_csv('filtered\_data.csv', index=False)

1. Write the code the get the minimum value of a row “year” in a dataframe df\_years

df\_years.loc[“year”].min()

1. If you are about to write some code using the pandas library, what is the first line of code you have to enter in your program?

import pandas as pd

1. Below is a table containing the population of major cities in millions by year. Write the code to print which cities in the following table had a population greater than 20 million in the year 2010.

*Note: I have made up these numbers except for 2020!*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | London | Paris | New York | Tokyo |
| 2020 | 9.0 | 10.9 | 18.9 | 37.5 |
| 2010 | 8.9 | 10.4 | 19.2 | 37.4 |
| 2000 | 8.9 | 10.5 | 19.1 | 37.1 |
| 1990 | 8.8 | 10.4 | 19.0 | 36.5 |

countries = df.columns.values

for x in countries:

min = df[x].min()

print(x, min)