

Intro to Data Science - Lab 2

DATA 1501 — Dr. Mihail
Department of Computer Science
Valdosta State University

September 7, 2021

1 Introduction

In this lab, you will create Pandas (Python Data Analysis Library) dataframes with various types of data. The lab submission will consist of a Word document with answers to questions in the following parts.

1.1 Part 1

Create a new Colab notebook and run the following code in a code cell:

```
# Import pandas library
import pandas as pd

# initialize list of lists
data = [['tom', 10], ['nick', 15], ['juli', 14]]

# Create the pandas DataFrame
df = pd.DataFrame(data, columns = ['Name', 'Age'])

# print dataframe.
df
```

- Question 1: How many attributes are there in the dataset you just created and printed?
- Question 2: What are the attribute names?
- Question 3: What are the attribute types?

1.2 Part 2

Create a code cell and run the following code:

```

# Python code demonstrate creating
# DataFrame from dict narray / lists
# By default addresses.

import pandas as pd

# initialise data of lists.
data = {'Name': ['Tom', 'nick', 'krish', 'jack'],
        'Age': [20, 21, 19, 18]}

# Create DataFrame
df = pd.DataFrame(data)

# Print the output.
df

```

Question 1: Describe the difference you see in the critical parts of the code that created the datasets in this part and part 1.

1.3 Part 3

Create a code cell and run the following code:

```

# Python code demonstrate to create
# Pandas DataFrame by passing lists of
# Dictionaries and row indices.
import pandas as pd

# Intitilialise data of lists
data = [{'b': 2, 'c': 3}, {'a': 10, 'b': 20, 'c': 30}]

# Creates pandas DataFrame by passing
# Lists of dictionaries and row index.
df = pd.DataFrame(data, index = ['first', 'second'])

# Print the data
df

```

Question 1: Notice there are 2 records in the dataset just created, however, in one of the records, the programmer did not specify a value for the attribute **a**. What did Pandas do?

1.4 Part 4

Please write two paragraphs about what you've learned in this lab.

Due Date: Before Midnight on Sunday, September 5th.
Submit the Word document via Blazeview/Assignments/Lab 2.