A0_Hall_Devin

January 24, 2022

1 Assignment - A0

1.1 Due: 1/24 @ 11:59pm

1.2 Devin Hall

Assignment Objectives:

Upon **successful completion** of this assignment, a student will be able to:

- Add new text and code to their assignment solutions
- Gain experience in formatting text using *Markdown* or *LaTeX*
- Install or access to programming software to copmlete this assignment and subsequent assignments.
- Load data from an online source and begin to access properties of the data.

This document is designed to be a homework template using Jupyter noteboks.

We can have code chunks throughout our answers. Here we will have a code chunk to load in the packages we expect to use in this assignment.

Some pre-requisites we will need for the project

[1]:
```
import pandas
```

1.3 (Question 2) String Manipulation

1.3.1 Hello World

[2]:
```
print("Hello World")
```

```
Hello World
```

1.3.2 Formatted Strings

[3]:
```
v1 = 'cs4821-cs5831'
v2 = 2022
print(f"Hello {v1}, welcome to {v2}")
```

```
Hello cs4821-cs5831, welcome to 2022
```

1.4 (Question 3) Comments

[4]:
```
# test comment
```

[5]:
```
# This is a larger comment block
# that may span multiple lines
```

1.5 (Question 4) Iris Data

[6]:
```
# Load the iris data from the website
IRIS_URL = 'https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.
 ↪data'

# How many rows should we print?
NUM_ROWS = 7

# Some constants to define for the column names
COLUMN_SEPAL_LENGTH = 'Sepal Length (cm)'
COLUMN_SEPAL_WIDTH = 'Sepal Width (cm)'
COLUMN_PETAL_LENGTH = 'Petal Length (cm)'
COLUMN_PETAL_WIDTH = 'Petal Width (cm)'
COLUMN_CLASS = 'Class'

# Read it using pandas
data = pandas.read_csv(IRIS_URL, names=[COLUMN_SEPAL_LENGTH,␣
 ↪COLUMN_SEPAL_WIDTH, COLUMN_PETAL_LENGTH, COLUMN_PETAL_WIDTH, COLUMN_CLASS])
# For part a
print("The first 7 rows of the data: ")
data.head(n=NUM_ROWS)
```

```
The first 7 rows of the data:
```

[6]:

	Sepal Length (cm)	Sepal Width (cm)	Petal Length (cm)	Petal Width (cm)	\
0	5.1	3.5	1.4	0.2	
1	4.9	3.0	1.4	0.2	
2	4.7	3.2	1.3	0.2	
3	4.6	3.1	1.5	0.2	
4	5.0	3.6	1.4	0.2	
5	5.4	3.9	1.7	0.4	
6	4.6	3.4	1.4	0.3	

	Class
0	Iris-setosa
1	Iris-setosa
2	Iris-setosa
3	Iris-setosa
4	Iris-setosa

```
5  Iris-setosa
6  Iris-setosa
```

[7]:
```
print(f"The number of samples is {len(data)}")
```

```
The number of samples is 150
```

[8]:
```
# Get the columns
sep_length_col = data[COLUMN_SEPAL_LENGTH]
sep_width_col = data[COLUMN_SEPAL_WIDTH]
petal_length_col = data[COLUMN_PETAL_LENGTH]
petal_width_col = data[COLUMN_PETAL_WIDTH]

# Print some information about these columns
print(f"{COLUMN_SEPAL_LENGTH}: min - {sep_length_col.min()}, max -␣
 ↪{sep_length_col.max()}")
print(f"{COLUMN_SEPAL_WIDTH}: min - {sep_width_col.min()}, max - {sep_width_col.
 ↪max()}")
print(f"{COLUMN_PETAL_LENGTH}: min - {petal_length_col.min()}, max -␣
 ↪{petal_length_col.max()}")
print(f"{COLUMN_PETAL_WIDTH}: min - {petal_width_col.min()}, max -␣
 ↪{petal_width_col.max()}")
```

```
Sepal Length (cm): min - 4.3, max - 7.9
Sepal Width (cm): min - 2.0, max - 4.4
Petal Length (cm): min - 1.0, max - 6.9
Petal Width (cm): min - 0.1, max - 2.5
```