DATA1002 Week 2 Tutorial

Monday 11/08/25

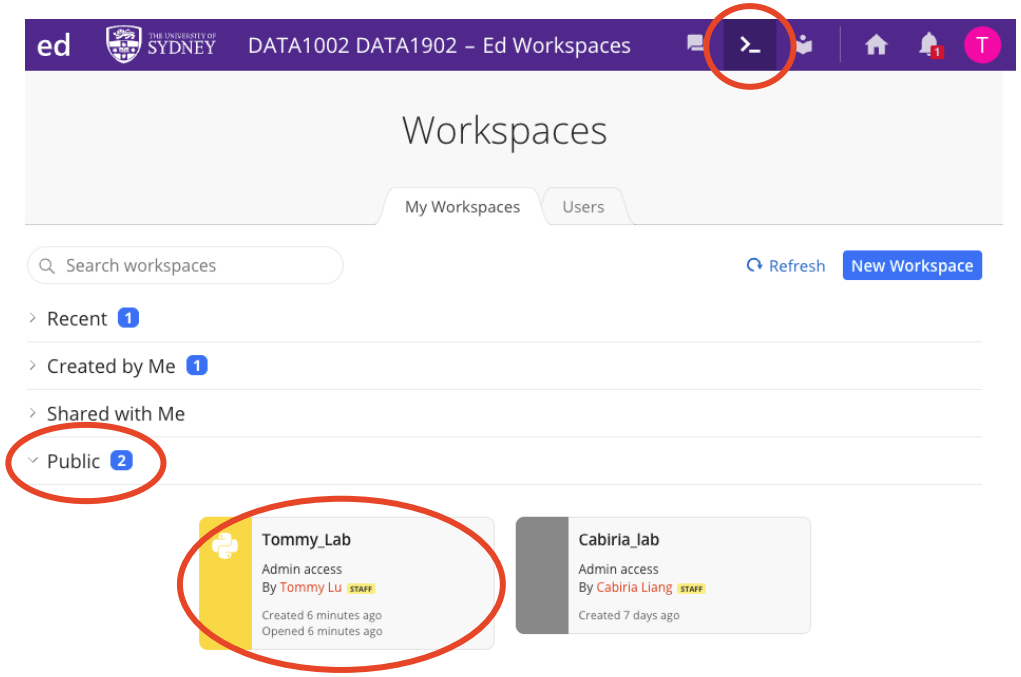
Tutorial Outline

- Introduction – Runsheet & Housekeeping
- Programs & Variables
- Expressions & Assignment
- Notional Machine
- Navigating CSVs

THE UNIVERSITY OF SYDNEY

Tutor: *Tommy Lu*

Access the material for this tutorial through **Ed Workspaces**



Today's Tutorial Runsheet

Time	Heading & Exercise(s)	Exercise Summary	Learning Outcome
12:00-12:10	**Essential Housekeeping**	—	—
12:10-12:15	**Python Overview**	Discussion about Python's applications in data science	Recognise Python's broad applications and relevance.
12:15-12:30	**Programs** (3 mini-tasks)	a) Identify syntax errors in a snippet. b) Correct indentation issues. c) Predict output of a given program.	Understand program structure, indentation rules, and case sensitivity in Python.
	Variables (2 tasks)	a) Declare variables of different types (int, float, str, bool). b) Reassign to different types and predict outcomes.	Declare and reassign variables; understand dynamic typing and naming conventions.
12:30-12:45	**Expressions & Assignment** (2 tasks)	a) Write expressions mixing literals & variables. b) Use parentheses to change calculation order.	Apply arithmetic & logical expressions; control order of operations; correctly assign values.
	Notional Machine (2 tasks)	a) Step-through a given code snippet, tracking variable state in a table. b) Debug incorrect output by tracing variables.	Apply mental model of Python execution to follow code flow and debug issues.
1:00-1:10	**Break**	Short break	—
1:10-1:30	**Navigating CSVs** (3 tasks, using info.csv)	a) Load CSV in Pandas, view first rows. b) Calculate mean, max. c) Filter rows and calculate stats. Excel comparison for filtering.	Read and manipulate CSV data; calculate summary stats; apply filters in Python and Excel.
1:30-2:00	**Ed Discussion Exercises / Exam-Style Questions Discussion**	—	—

Essential Housekeeping

Expectations for tutorials:

- Attendance?
- Material?
- Lab worksheet?
- Recommendations – group work, questions, Ed exercises, consolidating knowledge
- Recommendations – Students with Python experience

Provided material:

- These slides
- Python exercises on Ed

Python Overview

How can **Python** be used in **Data Science**?

Programs

```
# 1. PROGRAMS
# Exercise 1: Fix the errors
print("Hello World"
Print("this line has wrong case")

# Exercise 2: Correct indentation
if True:
print("This needs fixing")

# Exercise 3: Predict the output
x = 5
X = 10
print(x)
print(X)

# BONUS Exercise (if already familiar with Python): Discount Calculator
# Apply a 20% discount if total price is $100 or more, 10% if between $50-$100, otherwise no discount.
# Rearrange the following lines into an executable program

if total > 100:
if total >= 100:
elif total < 99:
elif total >= 50:
elif total > 50 and total < 99:
else:
discount = 0.10
discount = 0.20
discount = 0.0
total = float(input("Enter the total price: "))
final_price = total - (total * discount)
print("Final price:", final_price)
```

For code, go to Tommy_Lab Ed Workspace, Week 2

Variables

```
# 2. VARIABLES
# Exercise 1: Declare variables
age = ___       # integer
price = ___     # float
name = ___      # string
is_active = ___ # boolean

# Exercise 2: Reassign variable type and print
# REASSIGN HERE
print(age)
print(price)
print(name)
print(is_active)

# BONUS Exercise (if already familiar with Python): Write a program that store product details, update stock, and generate a formatted
report.
```

For code, go to Tommy_Lab Ed Workspace, Week 2

Expressions & Assignments

```
# 3. EXPRESSIONS & ASSIGNMENT
# Exercise 1: Writing expressions
x = 4
y = 2
# Write an expression that adds x and y, multiplies by 3, and stores in result
result = ___
print(result)

# Exercise 2: Without changing numbers, only operators, parentheses etc., how many numbers can be calculated?
value1 = 10 + 2 * 5
print(value1)
# assign value 2, 3, etc.

# BONUS Exercise (if already familiar with Python): Simple Bank Account

print("Bank Account Simulator")
balance = 1000
transaction = input("Enter 'deposit' or 'withdraw': ").lower()
amount = # FILL IN HERE #

if transaction == "deposit":
    # FILL IN HERE #
    print("Deposit successful! New balance:", balance)
elif # FILL IN HERE #
    if # FILL IN HERE #
        # FILL IN HERE #
        print("Withdrawal successful! New balance:", balance)
    else:
        print("Insufficient funds.")
else:
    print("Invalid transaction type.")
```

For code, go to Tommy_Lab Ed Workspace, Week 2

Notional Machine

```python
# 4. NOTIONAL MACHINE
# Exercise 1: Trace code
x = 10
y = 5
x = x + y
y = x - y
x = x - y
print(x, y)

# Exercise 2: What is the output?
a = 5
b = a + 3
a = b - 2
c = a + b
a = c - 1
c = a + b
print(a, b, c)

# BONUS Exercise (if already familiar with Python): Predict final values of x, y, and total without running the code. Then run it to check.

x = 1
y = 2
total = 0

for i in range(1, 4):
    x = x + i
    for j in range(i):
        y = y + j
        if (x + y) % 2 == 0:
            total += x
        else:
            total -= y

print(x, y, total)
```

For code, go to Tommy_Lab Ed Workspace, Week 2

Navigating CSVs

```
# 5. Navigating CSVs
# Import Data



# Calculate key representative statistics



# Filter and calculate



# Let's do it in Excel!
```

climate_data_2017.... w2_exercises_stude...

```
Date,State,City,Station Code,Minimum temperature (C),Maximum temperature (C),Rainfall (mm),Evaporation (mm),Sunshine (hours),Direction of maximum wind gust,Speed of maximum wind gust (km/h),9am Temperature (C),9am relative humidity (%),3pm Temperature (C),3pm relative humidity (%)
2017-06-01,NSW,Sydney,066062,8.0,18.3,0,4.8,9.7,WSW,43,10.3,54,17.8,39
2017-06-01,QLD,Brisbane,040913,7.4,23.3,0,2.4,7.7,SW,24,12.1,80,22.5,55
2017-06-01,VIC,Melbourne,086338,7.9,13.9,0.4,,,SSW,20,10.5,82,13.6,57
2017-06-02,VIC,Melbourne,086338,8.0,14.7,0,,,SSW,20,10.6,64,13.3,69
2017-06-02,QLD,Brisbane,040913,8.3,24.2,0,2.2,8.1,W,30,14.7,78,22.5,57
2017-06-02,NSW,Sydney,066062,8.5,18.1,0,3.2,4.8,SSW,41,10.7,61,16.4,50
2017-06-03,QLD,Brisbane,040913,9.1,23.1,0,1.4,7.7,SSW,28,14.7,86,22.9,57
2017-06-03,VIC,Melbourne,086338,9.9,14.3,0,,,SSW,19,10.8,76,13.1,73
2017-06-03,NSW,Sydney,066062,10.7,17.4,0,3.4,2.1,S,41,13.4,64,16.4,66
2017-06-04,VIC,Melbourne,086338,4.4,15.8,0,,,NE,13,5.7,100,15.5,53
2017-06-04,NSW,Sydney,066062,10.5,19.7,1.4,0.6,6.8,W,28,12.2,77,19.0,57
2017-06-04,QLD,Brisbane,040913,12.8,22.6,0,1.8,9.1,ENE,31,16.3,77,22.5,52
2017-06-05,VIC,Melbourne,086338,5.6,17.5,0,,,NW,28,8.7,82,16.8,57
2017-06-05,QLD,Brisbane,040913,10.1,25.3,0,1.8,9.2,ENE,31,15.0,71,24.9,30
2017-06-05,NSW,Sydney,066062,9.4,19.6,0,2.8,9.4,W,28,11.5,78,19.2,41
2017-06-06,VIC,Melbourne,086338,8.7,13.9,6.4,,,SSW,43,11.8,78,12.6,68
2017-06-06,QLD,Brisbane,040913,9.2,25.2,0,2.4,7.7,N,41,17.9,38,24.5,13
2017-06-06,NSW,Sydney,066062,9.1,17.9,0,2.2,7.2,WSW,59,11.2,58,14.7,43
2017-06-07,VIC,Melbourne,086338,5.9,14.3,0,,,SE,28,8.3,91,14.0,54
2017-06-07,NSW,Sydney,066062,8.8,15.5,54.6,2.6,0.0,SSW,56,13.3,82,14.2,85
2017-06-07,QLD,Brisbane,040913,5.0,22.8,0,4.4,9.1,SE,31,12.4,68,22.5,39
2017-06-08,QLD,Brisbane,040913,9.9,22.5,0,2.2,9.2,ENE,48,13.0,75,22.1,25
2017-06-08,VIC,Melbourne,086338,3.6,15.0,0,,,SSW,30,6.6,86,14.0,56
2017-06-08,NSW,Sydney,066062,12.4,18.1,61.0,1.8,1.3,SE,44,13.3,91,16.9,84
2017-06-09,NSW,Sydney,066062,10.4,18.9,3.4,1.6,4.8,ESE,52,12.5,90,18.2,64
2017-06-09,VIC,Melbourne,086338,6.6,15.0,1.8,,,S,22,11.0,98,13.9,66
2017-06-09,QLD,Brisbane,040913,7.7,24.1,0,4.8,9.3,NE,37,15.8,27,23.8,14
2017-06-10,NSW,Sydney,066062,12.5,18.1,38.8,3.6,0.2,SSE,50,15.8,85,17.5,70
2017-06-10,QLD,Brisbane,040913,7.7,20.8,0,3.8,1.9,NNE,33,14.7,33,20.4,25
2017-06-10,VIC,Melbourne,086338,7.0,15.8,0,,,SSW,17,7.5,100,14.0,68
```

For code, go to Tommy_Lab Ed Workspace, Week 2

Ed Discussion Exercises / Exam-Style Questions Discussion

Question 1: Importance of File Formats in Data Science Projects:

- Explain the difference between structured and unstructured text files in data science projects.
- Describe the benefits of using a comma-separated values (CSV) format for data exchange.
- Discuss the role of metadata in data files and the importance of a header row.

Question 2: Ethical Considerations When Working with Text Data:

- Discuss some ethical considerations associated with working with text data in data science projects.
- Explain the importance of anonymisation and data privacy when dealing with personal information in text files.
- Describe potential biases that might be present in text data and their impact on analysis.