

Introductory Programming Using Python

Day 2

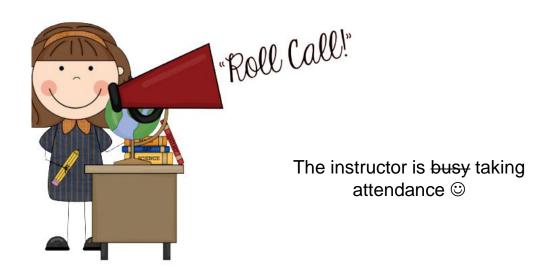
By Jason Lim & Derek Lee Republic Polytechnic





Administrative matters

Please change your log in screen name to resemble the registered name for the course



Perform software installation (if haven't done so):

- Python 3.8.x
- Wing IDE

Course Material Link: https://bit.ly/py-june21



Trainers

Day 1

Derek Lee

derek_lee@rp.edu.sg



Day 2
Jason Lim
jason_lim@rp.edu.sg





Programme Day Two

Morning	Afternoon
 Read and writing files Copying, moving and deleting files and folders Working with Excel Image Processing 	 Connecting to the Web Sending emails Creating Chart Generating PDF



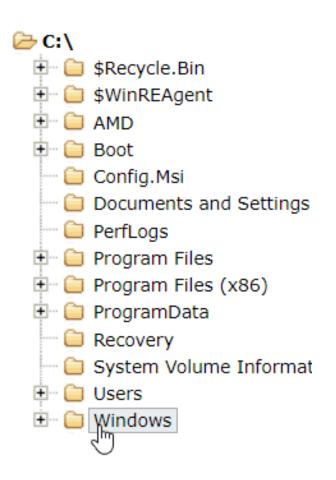
Outline for the day

Time	Agenda
9.00am	Welcome and admin matters
9.15am – 10.30am	Read and writing filesCopying, moving and deleting files and folders
10.30am – 10.45am	Break
10.45am – 12.30pm	Working with ExcelImage Processing
12.30pm – 1.30pm	Lunch
1.30pm – 3.15pm	Connecting to the WebSending emails
3.15pm – 3.30pm	Break
3.30pm – 4.30pm	Creating ChartGenerating PDF
4.45pm – 5.00pm	Wrap up, Q&A

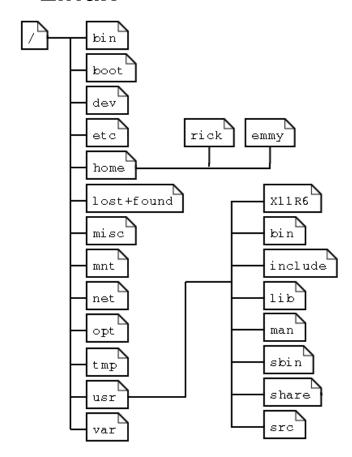


File Tree

Windows



Linux



File Paths



An absolute file path describes how to access a given file or directory, starting from the root of the file system.

 A relative file path is interpreted from the perspective your current working directory.

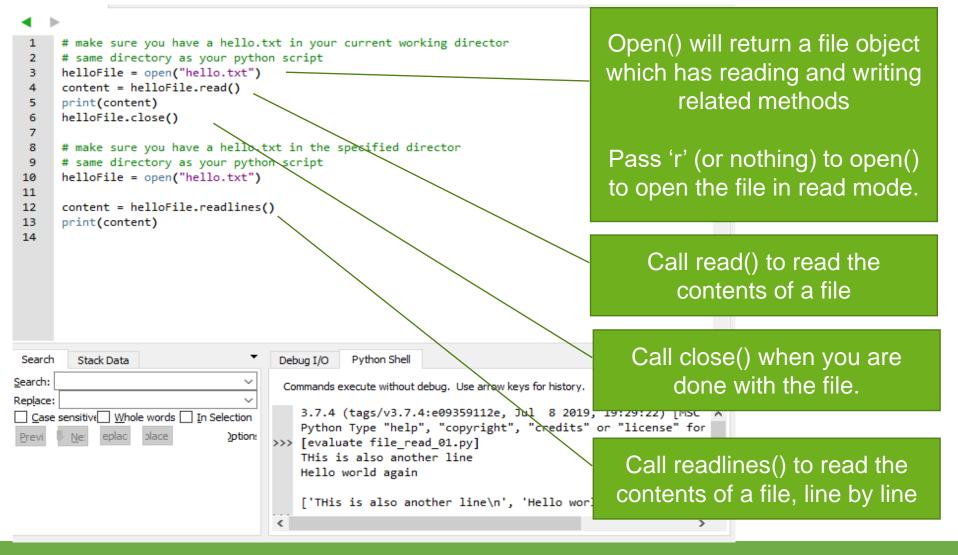
Relative Paths	Absolute Paths
\	C:\
.\	C:\bacon
.\fizz	C:\bacon\fizz
.\fizz\spam.txt	C:\bacon\fizz\spam.tx
.\spam.txt	C:\bacon\spam.txt
\eggs	C:\eggs
\eggs\spam.txt	C:\eggs\spam.txt
\spam.txt	C:\spam.txt

Absolute file paths are notated by a **leading forward slash or drive label**.

Relative file paths are notated by a lack of a leading forward slash.

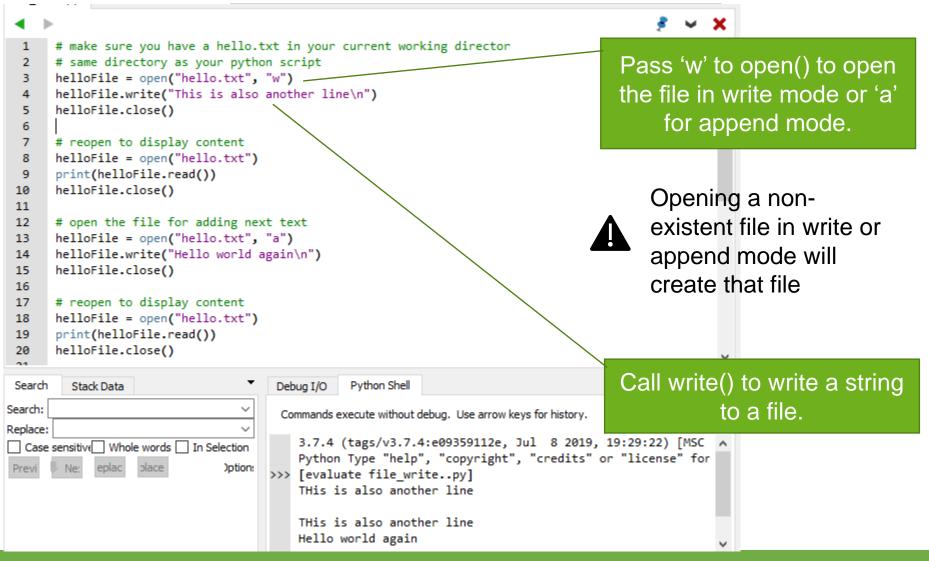


Read files









file_write.py



Copy and moving files

```
import shutil

# copy file, destination must exist
shutil.copy("hello.txt", "folder1")

# recursively copy an entire directory
# error if the destination folder already exist
shutil.copytree("folder1", "folder_to_delete")
shutil.copytree("folder1", "folder_to_delete2")

# move file, destination must exist
shutil.move("folder1/hello.txt", "folder2")

# move and rename file
shutil.move("folder_to_delete/hello.txt", "folder_to_delete2/newhello.txt")
```

https://docs.python.org/3/library/shutil.html

- shutil.copy(src, dst) Copy the file src to the file or directory dst
- shutil.copytree(src, dst) -Recursively copy an entire directory tree rooted at src.
- shutil.move(src, dst) Recursively move a file or
 directory (src) to another
 location (dst).



Deleting files

```
import os

# error if file do not exist
os.unlink("folder2/hello.txt")

# get current working directory
print(os.getcwd())

# delete directory (can only delete empty folder)
os.rmdir("folder_to_delete")

import shutil
# delete directory (with content)
# error if folder is not found
shutil.rmtree("folder_to_delete2")
```

e.g. To delete all .docx file in the current folder

```
import os

for filename in os.listdir():
    if filename.endswith(".docx"):
        print(filename)
        os.unlink(filename)
```

- os.unlink() will delete a file
- os.rmdir() will delete a folder (but folder must be empty)
- shutil.rmtree() will delete a folder and all its contents

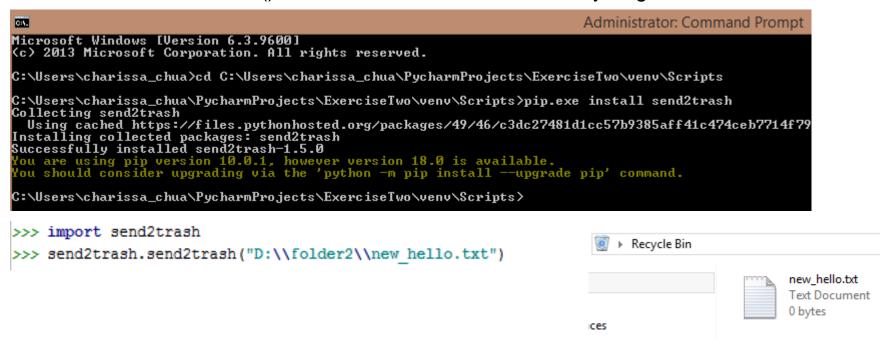


Deleting can be dangerous, so do a dry run first



send2Trash module

- Install send2trash module using pip.exe
 -> pip install send2trash
- send2trash.send2trash() will send a file or folder to the recycling bin

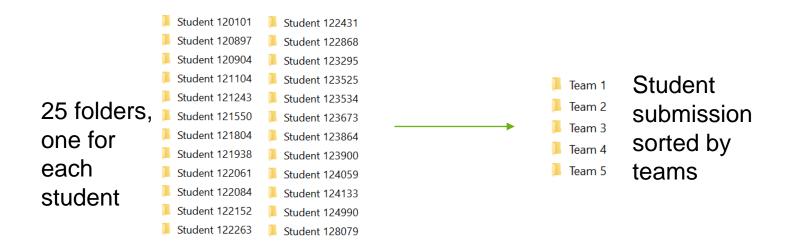


file_trash.py



Use Case Sharing

- Organizing students' submissions into separate folder
 - Class of 25 students





Other Use Cases

- System administrators can use these commands to
 - Copy and backup files to other hard-disks
 - Delete folders/ files at fixed schedules
 - End of financial year?
 - End of semester?
- Others use
 - Check timestamp of files, and delete all files created before a certain date



Exercise

Write code to achieve the following:

- 1. Create a file named: "myfile.txt".
- 2. Write the following line of text into the file:
 - Programming is fun!
- 3. Close the file
- 4. Create a folder called "myfolder"
 - Use os.mkdir() command
- 5. Copy myfile.txt to myfolder



Python Package Index

- https://pypi.org/
- A repository of software for the Python Programming Language
- Python Installation provides the core libraries needed for the common tasks
 - Additional packages can be found at the website and installed as extension
 - E.g. send2trash, openpyxl, pillow etc
- Installation is easy done with the following command
 - pip install <software_package>
- Installed packages can be found at:
 - C:\python38\Lib\site-packages



Using pip install

For all windows users by default

- Open command prompt
- pip install <package_name>

For Mac User

- Open terminal
- pip3 install <package_name>

For RP staff using RP issued laptop

- Open command prompt
- pip install --user <package_name>



Double-Dash

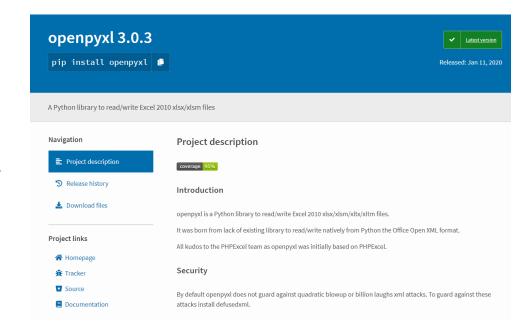


Excel Spreadsheet Manipulation with Python



Working with Excel

- Install openpyxl module using "pip install openpyxl"
- Full openpyxl documentation: <u>https://openpyxl.readthedocs.io/en/stable/index.html</u>





Typical Workflow for Excel Automation

You are given some data in a spreadsheet

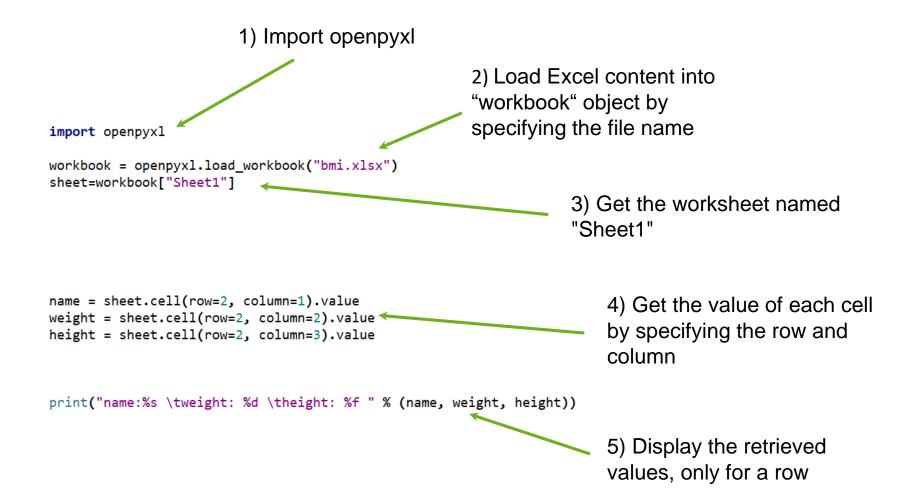
You want to do some or all of the following

- Analyze the data
- Manipulate the data
- Create visualization (Charts, Pivot Table etc)

Output the processed data in another spreadsheet



Reading Excel file





Reading Excel file

The typical workflow for reading excel file is to use a *for* loop to go through each row to read the data

```
import openpyxl
                                                              1) Get the number of rows
                                                              and columns
workbook = openpyxl.load workbook("bmi.xlsx")
sheet=workbook["Sheet1"]
max row = sheet.max row # get number of rows
                                                                    2) Use For loop to go
#loop through every row
                                                                    through every row
for i in range(2, max row + 1):
   #read cell
   name = sheet.cell(row=i, column=1).value
   weight = sheet.cell(row=i, column=2).value
                                                                          3) Display the retrieved
   height = sheet.cell(row=i, column=3).value
                                                                          values, for all rows
   print("name:%s \tweight: %d \theight: %f " % (name, weight, height))
```



Update Excel file

```
import openpyxl
workbook = openpyxl.load_workbook("bmi.xlsx")
                                                       1) Load file into memory & get
sheet=workbook["Sheet1"]
                                                       the sheet
max row = sheet.max row # get number of rows
# add a column header for bmi
                                                         2) Adds a header at the 4<sup>th</sup>
sheet.cell(row=1, column=4).value = "bmi"
                                                         column
#loop through every row
for i in range(2, max_row + 1):
   #read cell
                                                            3) Perform calculation with
   name = sheet.cell(row=i, column=1).value
                                                            values taken from the excel
   weight = sheet.cell(row=i, column=2).value
   height = sheet.cell(row=i, column=3).value
                                                            files
   bmi = weight / (height * height)
   sheet.cell(row=i, column=4).value = bmi 	
                                                          4) Add calculated value
                                                          to cell
   print("name:%s \tBMI: %f" % (name, bmi))
#save the file
workbook.save("bmi update.xlsx")
                                                 5) Save the spreadsheet
```

Create Excel file



If you have data in nested python list, you can write the data into an excel file.

```
import openpyxl
workbook = openpyxl.Workbook()
#get the default sheet
sheet=workbook["Sheet"]
#create a list of tuples as data source
data = [
   [225.7, 'Gone with the Wind', 'Victor Fleming'],
                                                                 1) Some data in nested list
   [194.4, 'Star Wars', 'George Lucas'],
   [161.0, 'ET: The Extraterrestrial', 'Steven Spielberg']
for row in data:
                                               2) Using for loop to add each row
   sheet.append(row)
                                               of data into the excel sheet
#save the spreadsheet
workbook.save("movies.xlsx")
                                              3) Save the spreadsheet
```



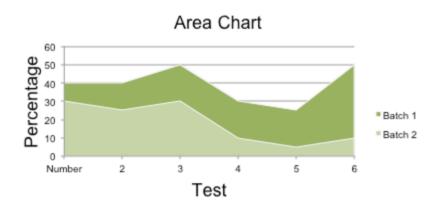
Format Excel

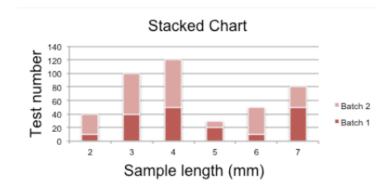
```
import openpyxl
from openpyxl.styles import Font, PatternFill, Border, Side
                                                                                1) Import necessary
workbook = openpyxl.load_workbook("bmi.xlsx")
sheet=workbook["Sheet1"]
                                                                                functions
#define the colors to use for styling
BLUE = "0033CC"
LIGHT BLUE = "E6ECFF"
WHITE = "FFFFFF"
                                                                              2) Setup colors and styles
#define styling
header font = Font(name="Tahoma", size=14, color=WHITE)
header fill = PatternFill("solid", fgColor=BLUE)
# format header
                                                                       3) Loop through cell and set
for row in sheet["A1:c1"]:
 for cell in row:
                                                                       properties
   cell.font = header font
   cell.fill = header_fill
#define styling
white side = Side(border style="thin", color=WHITE)
blue side = Side(border style="thin", color=BLUE)
alternate fill = PatternFill("solid", fgColor=LIGHT_BLUE)
border = Border(bottom=blue_side, left=white_side, right=white_side)
# format rows
for row_index, row in enumerate(sheet["A2:C3"]):
 for cell in row:
   cell.border = border
   if row index %2:
     cell.fill = alternate_fill
```



Excel Charts

- Openpyxl supports the creation of many types of charts
 - Area Charts
 - Bar and Column Charts
 - Bubble Charts
 - Line Charts
 - Scatter Charts
 - etc









Create Excel Chart

```
1) Import necessary
import openpyxl
                                                                  functions
from openpyxl.chart import BarChart, Reference, Series
workbook = openpyx1.load_workbook("bmi.xlsx")
sheet=workbook["Sheet1"]
                                                        2) Load the data from excel
                                                        file
chart = BarChart()
labels = Reference(sheet, min col=1, min row=2, max row=3)
data = Reference(sheet, min_col=3, min_row=1, max_row=3)
                                                               3) Specify the label and
chart.add_data(data, titles_from_data=True)
                                                               the data range
chart.set categories(labels)
chart.title = "Height"
sheet.add chart(chart, 'E1')
                                             4) Add the chart to the
workbook.save('bmi_chart.xlsx')
                                             sheet, and save the file in
                                             another excel file.
```

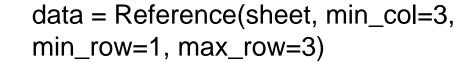


Create Excel Chart

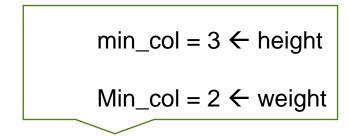
```
import openpyxl
                                                                         1) Import necessary
from openpyxl.chart import BarChart, Reference, Series
                                                                         functions
workbook = openpyxl.load workbook("bmi.xlsx")
sheet=workbook["Sheet1"]
                                                                2) Load the data from excel
                                                                file
chart = BarChart()
# first column is used as label, starting from row 2
labels = Reference(sheet, min col=1, min row=2, max row=3)
                                                                      3) Specify the label and
# first row is used for header, that is why min row is 1
                                                                      the data range
data = Reference(sheet, min col=3, min row=1, max row=3)
chart.add data(data, titles from data=True)
chart.set_categories(labels)
                                                    4) Specify values for title x-axis
chart.title = "Bar Chart"
                                                    and y-axis for the chart
chart.x axis.title = "Name"
chart.y axis.title = "Height"
chart.series[0].SeriesLabel = "height"
                                                           5) Add the chart to the
                                                           sheet, and save the file
sheet.add chart(chart, 'E1')
workbook.save('bmi_chart.xlsx')
                                                           in another excel file.
```



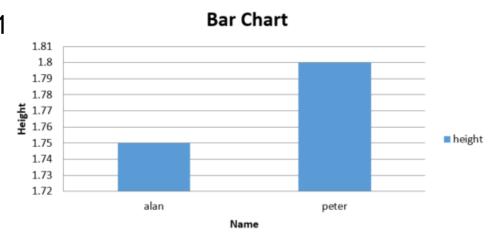
More Explanation on Chart Reference







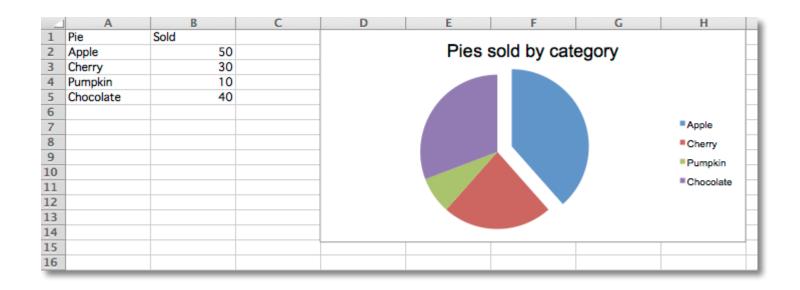
labels = Reference(sheet, min_col=1
min_row=2, max_row=3)





Exercise

- Write code to achieve the following:
 - 1. Refer to the documentation at https://openpyxl.readthedocs.io/en/stable/charts/pie.html#id1
 - 2. Follow the instruction to create the following pie chart.





Use Case Sharing

- RPA (Robotic Process Automation) can be used to send WhatsApp messages to students individually
- Write script
 - to split contact list for students based on intake year
 - to add country code: 65 to all the numbers

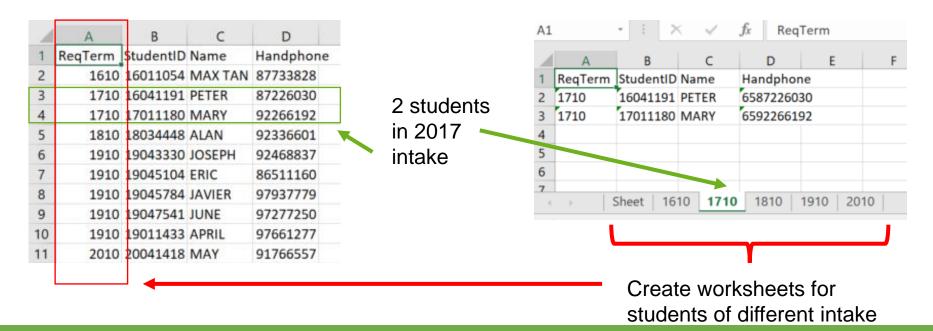
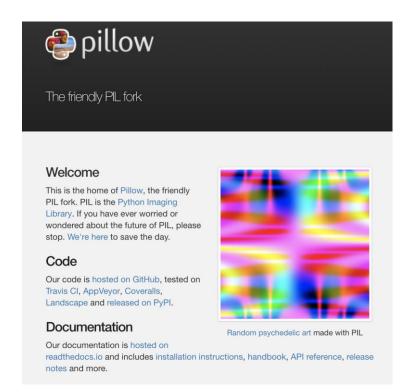




Image Processing with Python



Image Processing



For the next section we are going to use the Python Image Library, or in short Pillow.

Install using the following command: pip install pillow

The documentation is at: https://pillow.readthedocs.io/en/stable/ha ndbook/overview.html



Image Processing

As a start we need to import it:

```
import os
                                                 import Image
from PIL import Image
                                                We can open images with
filename = "img/clungup.jpg"
                                                 im = Image.open(fullname)
im = Image.open(filename)
print ("%s - %s" % (im.size, im.mode))
                                                 Then we can display the image using
                                                 im.show()
# show the image
im.show()
                                                 Print some info about the image using
# close the file
                                                 im.size and im.mode
im.close()
                                      Debug I/O
                                                   Python Shell
                                                                            OS Commands
                                                                Messages
                                     Debug I/O (stdin, stdout, stderr) appears below
                                     (800, 600) - RGB
                                         Size: 800 x 600, Mode: RGB
```



Image Processing

```
import os
from PIL import Image, ImageFilter

filename = "img/clungup.jpg"

im = Image.open(filename)

out = im.filter(ImageFilter.BLUR)

im.show()
out.show()
```





Pillow has many conversion and filters, to use filters we need to extend our import:
from PIL import Image,

The way you can apply filters is:
out = im.filter(ImageFilter.BLUR)

Try other different filters!

ImageFilter

Image processing - filters





image = ImageOps.grayscale(image)



image = image.filter(ImageFilter.FIND EDGES)



image = image.filter(ImageFilter.CONTOUR)



image = ImageOps.solarize(image)



* Remember to include ImageOps in your import statement



Image processing - filters

```
import os
from PIL import Image, ImageFilter, ImageOps
filename = "img/clungup.jpg"
im = Image.open(filename)
# Filter
#out = im.filter(ImageFilter.BLUR)
#out = im.filter(ImageFilter.FIND_EDGES)
#out = im.filter(ImageFilter.CONTOUR)
# ImageOps
out = ImageOps.grayscale(im)
#out = ImageOps.solarize(im)
im.show()
out.show()
```

* Remember to include ImageOps in your import statement



Image Processing - Rotating

```
#Flipping the image horizontally or vertically
out = im.transpose(Image.FLIP_LEFT_RIGHT)
                                                       Flip images
out = im.transpose(Image.FLIP_TOP_BOTTOM)
#Rotating the image
out = im.transpose(Image.ROTATE_90)
out = im.transpose(Image.ROTATE_180)
                                                       Rotate images
out = im.transpose(Image.ROTATE_270)
#Contrast
#First add ImageEnhance to our imports:
from PIL import Image, ImageFilter, ImageEnhance
#Then:
                                                       make image brighter by
enh = ImageEnhance.Contrast(im)
out = enh.enhance(1.3)
                                                       changing the contrast
```



Image Processing - Writing

```
import os
from PIL import Image, ImageFilter, ImageOps

filename = "clungup.jpg"

src_folder = "img/"
out_folder = "out/"

im = Image.open(src_folder + filename) # img/clungup.jpg
out = im.filter(ImageFilter.BLUR)

outFilename = out_folder + filename # out/clungup.jpg
out.save(outFilename)
```

You can see the image, but it's not being saved!

All you need to do to save the images in the "out" folder is: out.save(the name of the output file)

image_save.py



Image processing – Converting

Sometimes, we may want to keep all your photos in the same format.

Or, we obtained some gif files but want to have bmp or png type.

Pillow understands the output file, and will convert if the output file is different from the input.

fname1 fname2 holiday.jpg

How can we convert the string holday.gif to holiday.jpg?

```
>>> fname1 = "holiday.gif"
>>> fname2 = fname1.split(".")[0] + ".jpg"
>>> print(fname2)
holiday.jpg
>>>

>>> fname1 = "holiday.gif"
>>> f, e = os.path.splitext(fname1)
>>> fname2 = f + ".jpg"
>>> print(fname2)
holiday.jpg
>>>
```



Image processing – Converting

```
import os
from PIL import Image, ImageFilter, ImageOps
filename = "clungup.jpg"
src folder = "img/"
out folder = "out/"
im = Image.open(src_folder + filename) # img/clungup.jpg
out = im.filter(ImageFilter.BLUR)
# split the filename and the extension
f, e = os.path.splitext(filename)
# add the gif extension to the filename
fname2 = f + ".gif"
outFilename = out folder + fname2 # out/clungup.gif
out.save(outFilename)
```

os.path.splitext(file) returns a list. We are only interested in f which is the first item in the list.



Image processing – Watermark

You can reduce the size to 100,100

mark = Image.open("img\\watermark.png")
mark = mark.resize((100,100))

Create a new function called

def watermark(im, mark, position):

It takes the original image, the watermark image and the desired position that we want the watermark to appear. The function will return the result.

We can use this function like:

watermark(im, mark, (0, 50)).show()

or

imOut = watermark(im, mark, (0,50))
imOut.save(fileOut)

There could be cases where you want to leave a small footprint on your images, called watermark.

In this case we can use the \\img\\watermark.png and place it in each image on the bottom right.





Image processing – Watermark

```
from PIL import Image
def watermark(im, mark, position):
    layer = Image.new("RGBA", im.size, (0,0,0,0))
    layer.paste(mark, position)
    return Image.composite(layer, im, layer)
im = Image.open("img\\clungup.jpg")
mark = Image.open("img\\watermark.png")
mark = mark.resize((100,100))
mark.putalpha(128)
out = watermark(im, mark, (0,0))
out.show()
```

First we need to create a new layer with the size of the original image.

Then we paste the watermark image at the desired position and we return the composite.

Finally we merge the image and the layer together and return the result.

Then you can use it like shown here



Use Case I: Batch Resize

- 1. Find all the files in "img" folder with ".jpg" extension
- 2. Resize all the file to 60 x 90.
- 3. Save all the files to the resized folder

```
import os
from PIL import Image, ImageFilter, ImageOps

files = os.listdir('img')
size = 60, 90

for file in files:
    if file.lower().endswith(".jpg"):
        im = Image.open("img/" + file)
        im.thumbnail(size, Image.ANTIALIAS)
        im.save("resized/" + file, "JPEG")
```



Use Case II: Batch Rename

- 1. Find all the files in "img" folder with ".jpg" extension
- 2. Copy all the files to the renamed folder
- 3. Rename all the files with the "s-" prefix.

```
import os
import shutil

files = os.listdir('img')

for file in files:
    if file.lower().endswith(".jpg"):
        shutil.copyfile("img/" + file, "renamed/s-" + file[:-4] + ".jpg")
```



Web Automation with Python



requests – download files and web pages from the Web

pip install requests

import requests

Get the required information from the given URL

```
url="https://api.data.gov.sg/v1/environment/24-hour-weather-forecast"
req=requests.get(url)
print(req.text)
```





Data is in JSON format

import requests

Use a JSON formatter tool to present the data in a nicer form

http://jsonviewer.stack.hu/

```
url="https://api.data.gov.sg/v1/environment/24-hour-weather-forecast"
    req=requests.get(url)
    print(req.text)
                                                                                                                                                                                                                                                            ■ { } JSON
                                                                                                                                                                                                                                                                  ⊟{}0
                                                                                                                                                                                                                                                                                   update timestamp: "2020-07-22T14:51:18+08:00"
{"items":[{"update timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00","timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18+08:00*,"timestamp":"2020-07-22T14:51:18-07-25*,"timestamp":"2020-07-22T14:51:18-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"2020-07-25*,"timestamp":"
                                                                                                                                                                                                                                                                                   timestamp: "2020-07-22T14:25:00+08:00"
"2020-07-22T14:25:00+08:00", "valid period": { "start": "2020-07-22T12
                                                                                                                                                                                                                                                                             "end": "2020-07-23T12:00:00+08:00"}, "general": { "forecast": "Thunder

☐ { } general

"relative humidity": {"low":70, "high":95}, "temperature": {"low":22, '
                                                                                                                                                                                                                                                                                        forecast: "Thundery Showers"
"wind":{"speed":{"low":10, "high":20}, "direction":"ESE"}}, "periods'

→ { } relative humidity
"start": "2020-07-22T12:00:00+08:00", "end": "2020-07-22T18:00:00+08:
                                                                                                                                                                                                                                                                                   :{"west":"Moderate Rain", "east":"Moderate Rain", "central":"Light I
                                                                                                                                                                                                                                                                                   "Light Rain", "north": "Light Rain"}}, { "time": { "start": "2020-07-22T
                                                                                                                                                                                                                                                                             □ [ ] periods
,"end":"2020-07-23T06:00:00+08:00"},"regions":{"west":"Partly Clou
                                                                                                                                                                                                                                                                                   ⊞ {} 0
"east": "Partly Cloudy (Night)", "central": "Partly Cloudy (Night)",
                                                                                                                                                                                                                                                                                   ⊞{}1
"Partly Cloudy (Night)", "north": "Partly Cloudy (Night)"}}, {"time":
"2020-07-23T06:00:00+08:00", "end": "2020-07-23T12:00:00+08:00"}, "re
                                                                                                                                                                                                                                                                                  ⊞ { } 2
                                                                                                                                                                                                                                                                  ■ { } api info
                                                                                                                                                                                                                                                                              status : "healthy"
```



- To work with JSON data, import json first
- Use json.loads() to load the data in JSON format
- Extract and retrieve the required data

```
import json
import requests

url="https://api.data.gov.sg/v1/environment/24-hour-weather-forecast"
req=requests.get(url)

data = json.loads(req.text)

# print update timestamp
update_time = data["items"][0]["update_timestamp"]
print("Update time: " + update_time)

# print forecast
forecast = data["items"][0]["general"]["forecast"]
print("Forecast: " + forecast)
```

Update time: 2020-07-22T14:51:18+08:00

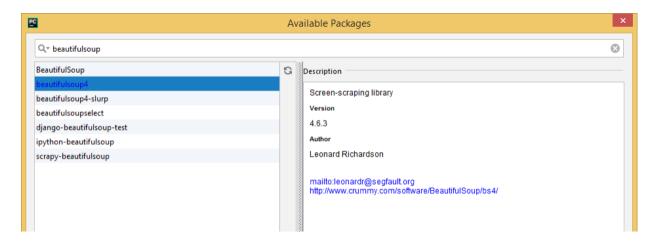
Forecast: Thundery Showers



Beautiful Soup – a third party module that parses HTML (web pages)

Web Scraping – download and process Web content

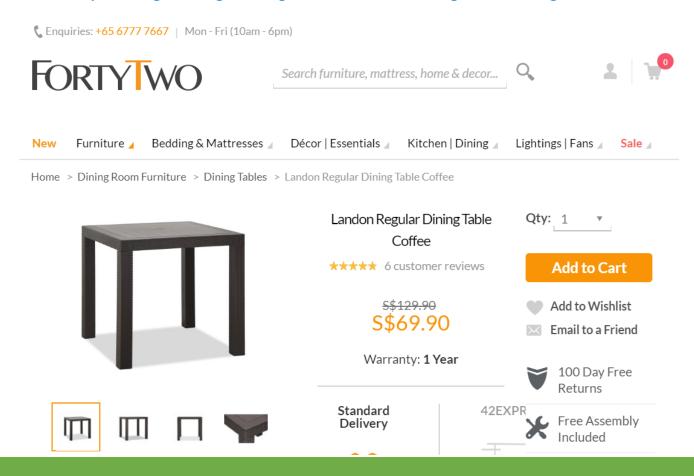
Install Beautiful Soup 4 - pip install beautifulsoup4





What's the URL?

https://www.fortytwo.sg/dining/dining-tables/landon-regular-dining-table-coffee.html

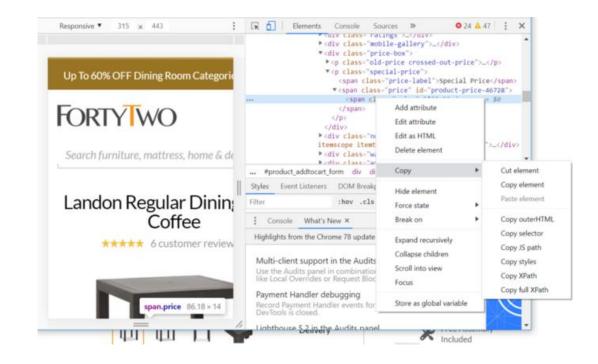




Get the url

https://www.fortytwo.sg/dining/dining-tables/landon-regular-dining-table-coffee.html

- Select the element to extract.
 - right-click -> "Inspect"
 - hover to "Copy"
 - click on "Copy selector"

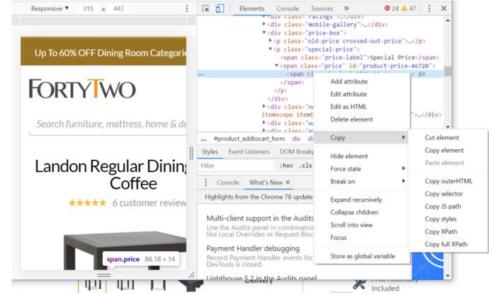




Get the url

import bs4

- Select the element to extract
 - right-click -> "Inspect"
 - hover to "Copy"
 - click on "Copy selector"



```
import requests
requestObj = requests.get("https://www.fortytwo.sg/dining/dining-tables/landon-regular-dining-table-coffee.html")
requestObj.raise_for_status()
soup = bs4.BeautifulSoup(requestObj.text, 'html.parser')
                                                                                              Debug I/O
                                                                                                         Python Shell
                                                                                                                    Messages
                                                                                                                              OS Cc
elements = soup.select("#product-price-46728") # $69.90
                                                                                              Debug I/O (stdin, stdout, stderr) appears below
print("Current Price: " + elements[0].text)
                                                                                              Current Price: S$69.90
elements = soup.select("#old-price-46728") # $129.90
                                                                                              Old Price: S$129.90
print("\nOld Price: " + elements[0].text)
                                                                                              Delivery Date:
elements = soup.select("div.earliest-delivery-date") # Earliest by Sunday, 31 May 2020
                                                                                              Earliest by
print("\nDelivery Date: " + elements[0].text)
                                                                                              Sunday, 31 May 2020
```

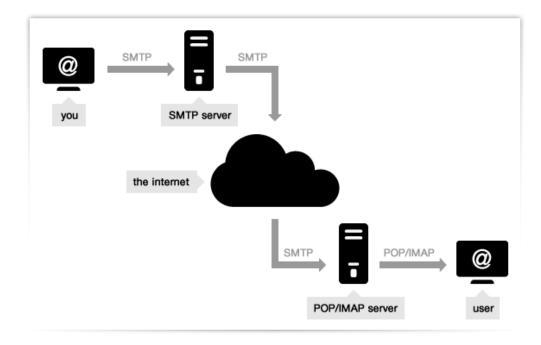
web_scrap.py



Email Automation with Python



Send Email



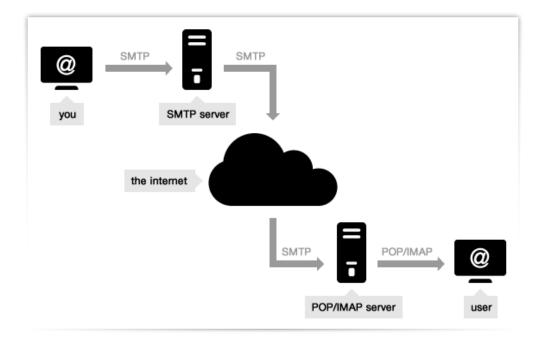
- SMTP (Simple Mail Transfer Protocol) is used for sending and delivering from a client to a server via port 25, 465 or 587: it's the **outgoing server**.
- IMAP and POP are two methods to access email.
 IMAP is the recommended method when you need to check your emails from several different devices, such as a phone, laptop, and tablet.

https://www.mailgun.com/blog/which-smtp-port-understanding-ports-25-465-587/

https://serversmtp.com/what-is-smtp-server/



Send Email



- Note: The SMTP servers used when you send your emails- Hotmail, Gmail , Yahoo Mail – are shared among users
- Common providers establish some **strict limits** on the number of emails you can send (e.g. Yahoo's restriction is 100 emails per hour).
- If you plan to send a bulk email or set up an email campaign you should opt for a professional outgoing email server like turboSMTP,
- which guarantees a controlled IP and ensure that all your messages reach their destination.



Incoming Mail (IMAP) Server	imap.gmail.com Requires SSL: Yes Port: 993
Outgoing Mail (SMTP) Server	smtp.gmail.com Requires SSL: Yes Requires TLS: Yes (if available) Requires Authentication: Yes Port for SSL: 465 Port for TLS/STARTTLS: 587
Full Name or Display Name	Your name
Account Name, User name, or Email address	Your full email address
Password	Your Gmail password

Note: If you are using your office network, most port numbers, including 587, may be blocked.



- Import smtplib module
- Specify Gmail email & password, receiver's email address, email title & content
- Connect to SMTP server using Port 587
- Call starttls() to enable encryption for your connection
- Login using email and password
- Call sendmail()
- Call quit() to disconnect from the SMTP server

```
import smtplib

sender_email_address = "your_email_address@gmail.com"
sender_email_password = "xxxxxxxxxxxxxxx"
receiver_email_address = "another_email_address@gmail.com"
email_title_content = "Subject: Sending Email Using Python\nThis is a test email."
email_title_content = "Subject: Sending Email Using Python\nThis is a test email."
```

➤ The start of the email body must begin with "Subject: " for the subject line. The "\n" newline character separates the subject line from the main body content.

```
print("Trying to connect to Gmail SMTP server")
smtpObj = smtplib.SMTP("smtp.gmail.com", 587)
smtpObj.starttls()

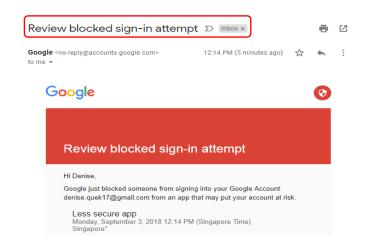
print("Connected. Logging in...")
smtpObj.login(sender_email_address, sender_email_password)

smtpObj.sendmail(sender_email_address, receiver_email_address, email_title_content)
print("Email sent successfully...")
smtpObj.quit()
```

58



 Google may block attempted sign-in from unknown devices that don't meet their security standards!



```
C:\Users\denise_quek\AppData\Local\Programs\Python\Python37\python.exe D:/CET_Python/Denise/TestEmail.py
Trying to connect to Gmail SMTP server
Connected. Logging in...
Traceback (most recent call last):
    File "D:/CET Python/Denise/TestEmail.py", line 13, in <module>
        smtpObj.login(sender_email_address, sender_email_password)
    File "C:\Users\denise quek\AppData\Local\Programs\Python\Python37\lib\smtplib.py", line 730, in login
        raise last_exception
    File "C:\Users\denise quek\AppData\Local\Programs\Python\Python37\lib\smtplib.py", line 721, in login
        initial_response_ok=initial_response_ok)
    File "C:\Users\denise quek\AppData\Local\Programs\Python\Python37\lib\smtplib.py", line 642, in auth
        raise SMTPAuthenticationError(code, resp)
smtplib.SMTPAuthenticationError: (534, b'5.7.9 Application-specific password required. Learn more at\n5.7.9
```



Steps To Create Google App Password

- Step 1: Login to Gmail. Go to Account → Signing in to Google
- Step 2: Make sure that 2-Step Verification is on
- Step 3: Create an App password



Generated app password

•••••	•	
Password		
securesally@gmail.com		
mail address		
inter the information below to connect to	your Google account.	
Add your Google accou	gle account	

Your app password for Windows Computer

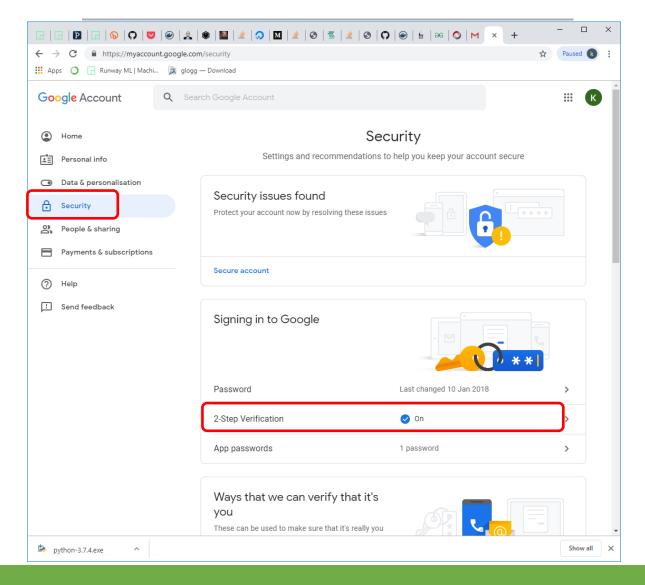
How to use it

- 1. Open the "Mail" app.
- 2. Open the "Settings" menu.
- Select "Accounts" and then select your Google Account.
- Replace your password with the
 16-character password shown above.

Just like your normal password, this app password grants complete access to your Google Account. You won't need to remember it, so don't write it down or share it with anyone. Learn more

DONE







Replace your actual password with the App password

```
import smtplib

sender_email_address = "your_email_address@gmail.com"

sender_email_password = "xxxxxxxxxxxxxxxx"

receiver_email_address = "another_email_address@gmail.com"

email_title_content = "Subject: Sending Email Using Python\nThis is a test email."
```

Run your email program

```
C:\Users\denise_quek\AppData\Local\Programs\Python\Python37\python.exe D:/CET_Python/Denise/TestEmail.py
Trying to connect to Gmail SMTP server
Connected. Logging in...
Email sent successfully...
Process finished with exit code 0
```



Use Case: Send emails to students

Send email to students who were absent

```
Student Email
                                 Status
          code.musically@gmail.com Present
2 Alicia
          code.musically@gmail.com Present
  Bryan
          code.musically@gmail.com Absent
  Carol
          code.musically@gmail.com Absent
5 David
          code.musically@gmail.com Present
  Evelyn
                            #! python3
            Sheet1
                            import openpyxl, smtplib
                            def sendEmail(name, emailTo):
                                email body = "Subject: Your attendance. \nDear %s, \nYou were absent for class.\n" %(name)
                                 smtpObj = smtplib.SMTP("smtp.gmail.com", 587)
                                 smtpObj.starttls()
                                smtpObj.login("code.musically@gmail.com", "xxxxxxxxxxxx")
                    10
                                smtpObj.sendmail('code.musically@gmail.com', emailTo, email body)
                    11
                    12
                    13
                                 smtpObj.quit()
```



Use Case: Send Emails to Students

- Open an Excel file
- Send email to students who were absent.

```
workbook = openpyxl.load workbook("D:\CET Python\students attendance.xlsx")
16
        sheet = workbook["Sheet1"]
17
18
        max row = sheet.max row
19
        max column = sheet.max column
20
21
        for i in range(1, max row+1):
22
23
            attendance = sheet.cell(row=i, column=3).value
24
25
            if attendance == "Absent":
26
27
                name = sheet.cell(row=i, column=1).value
                email = sheet.cell(row=i, column=2).value
28
29
                print(name + " is absent.")
30
                sendEmail(name, email)
31
32
                print("Email sent to " + email)
                print()
33
34
```



Sharing other Use Cases

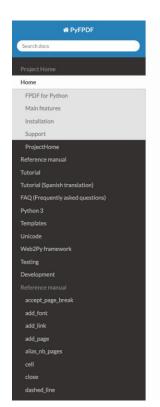
- Sending Emails using Outlook
- Create Appointment using Outlook

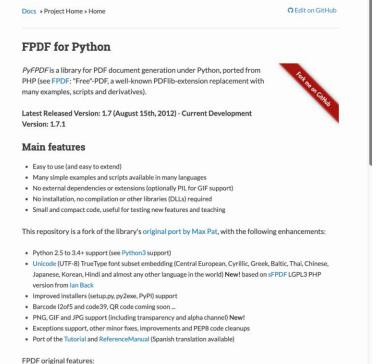


Generate PDF Report with Python









- Install fpdf
 - pip install fpdf

https://pyfpdf.readthedocs.io/en/latest/Tutorial/index.html



PDF – Basic document

```
import fpdf
#create a new pdf
document = fpdf.FPDF()
#define font and color for title and add the first page
document.set font("Times", "B", 14)
document.set_text_color(19,83,173)
document.add page()
#write the title of the document
document.cell(0,5,"PDF Test Document")
document.ln()
#write a long paragraph
document.set_font("Times", "", 11)
document.set_text_color(0)
document.multi_cell(0,10, "This is an example of a long paragraph. \n" * 10)
document.ln()
#save the document
document.output("pdf report.pdf")
```

- Import fpdf
- Create a new pdf document
- Add page
- Add text
- Save file

PDF Test Document

This is an example of a long paragraph.



PDF – adding images

```
import fpdf
#create a new pdf
document = fpdf.FPDF()
#define font and color for title and add the first page
document.set_font("Times", "B", 14)
document.set_text_color(19,83,173)
document.add_page()
#add a image
document.image("rp_logo.png", x=10, y=5, w=23)
document.set_y(40);
#write the title of the document
document.cell(0,5,"PDF Test Document")
document.ln()
#write a long paragraph
document.set_font("Times", "", 11)
document.set_text_color(0)
document.multi_cell(0,5, "This is an example of a long paragraph. " * 10)
document.ln()
#save the document
document.output("pdf_report.pdf")
```

- Import fpdf
- Create a new pdf document
- Add page
- Add text, logo
- Save file



PDF Test Document

This is an example of a long paragraph. This is an example of a long paragraph.

https://pyfpdf.readthedocs.io/en/latest/reference/image/index.html

pdf_image.py



PDF – Adding password

```
import fpdf
import PyPDF2
```

#create a new pdf

pip install PyPDF2

```
document = fpdf.FPDF()
#define font and color for title and add the first page
document.set font("Times", "B", 14)
document.set_text_color(19,83,173)
document.add page()
#write the title of the document
document.cell(0,5,"PDF Test Document")
document.ln()
#save the document
document.output("pdf report before pw.pdf")
#save the document into a new password protected/encrypted pdf
pdffile = open(r"pdf report before pw.pdf", "rb")
pdfReader = PyPDF2.PdfFileReader(pdffile)
pdfWriter = PvPDF2.PdfFileWriter()
for pageNum in range(pdfReader.numPages):
    pdfWriter.addPage(pdfReader.getPage(pageNum))
pdfWriter.encrypt('123') 
resultPDF = open(r"pdf_report_after_pw.pdf", "wb")
pdfWriter.write(resultPDF)
resultPDF.close()
pdffile.close()
```

Password is 123

https://pythonhosted.org/PyPDF2/



Use Cases

Automation:

- Generation of reports with data from spreadsheet or database
- Generation of Course Certificates in PDF format
- Password protection of banking statement in PDF file



Charting/Visualisation with Python



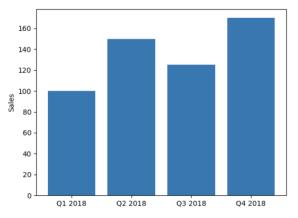


pip install matplotlib

Full documentation: https://matplotlib.org/



```
import matplotlib.pyplot as plt
 1
2
 3
     #set up values
 4
     VALUES = [100, 150, 125, 170]
 5
     POS = [0,1,2,3]
 6
     LABELS = ['Q1 2018','Q2 2018','Q3 2018','Q4 2018']
7
 8
     #set up the chart
9
     plt.bar(POS, VALUES)
10
     plt.xticks(POS, LABELS)
11
     plt.ylabel('Sales')
12
13
     #to display the chart
14
     plt.show()
```



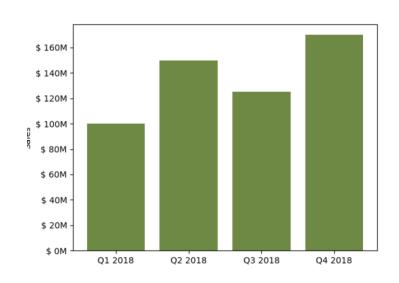
- Install matplotlib
- Prepare data
- Create bar graph
- Display the chart

https://matplotlib.org/api/_as_gen/matplotlib.pyplot.bar.html



```
import matplotlib.pyplot as plt
 2
     from matplotlib.ticker import FuncFormatter
 3
 4
     def value_format(value, position):
 5
              return '$ {}M'.format(int(value))
 6
 7
     # set up values
 8
     VALUES = [100, 150, 125, 170]
 9
     POS = [0,1,2,3]
10
     LABELS = ['Q1 2018','Q2 2018','Q3 2018','Q4 2018']
11
12
     # set up the chart
13
     # Colors can be specified in multiple formats, as
14
     # described in https://matplotlib.org/api/colors_api.html
15
     # https://xkcd.com/color/rgb/
16
     plt.bar(POS, VALUES, color='xkcd:moss green')
17
     plt.xticks(POS, LABELS)
18
     plt.ylabel('Sales')
19
20
     # retreive the current axes and apply formatter
     axes = plt.gca()
21
22
     axes.yaxis.set_major_formatter(FuncFormatter(value_format))
23
24
     # to display the chart
25
     plt.show()
```

- Install matplotlib
- Prepare data
- Customise graph options
- Create bar graph
- Display the chart





```
import matplotlib.pyplot as plt
 2
 3
     #set up values
     VALUESA = [100, 150, 125, 170]
 4
     VALUESB = [0, 15, 40, 80]
     VALUESC = [3,5,7,3]
     POS = [0,1,2,3]
 8
     LABELS = ['Q1 2018','Q2 2018','Q3 2018','Q4 2018']
 9
10
     # Create the first plot
     plt.subplot(2,1,1)
11
12
13
     #creata a bar graph with informaton about VALUESA
14
     plt.bar(POS, VALUESA)
15
     plt.ylabel('Sales')
16
17
     #create a different Y axis, and add information
18
     #about VALUESB as a line plot
19
     plt.twinx()
20
     plt.plot(POS, VALUESB, 'o-', color='red')
     plt.xticks(POS, LABELS)
21
22
     plt.ylabel('Sales B')
23
     plt.xticks(POS, LABELS)
24
     #create another subplot and fill it iwth VALUESC
25
26
     plt.subplot(2,1,2)
27
     plt.plot(POS, VALUESC)
28
     plt.gca().set_ylim(bottom=0)
     plt.xticks(POS,LABELS)
29
30
     plt.show()
31
```

Multiple charts



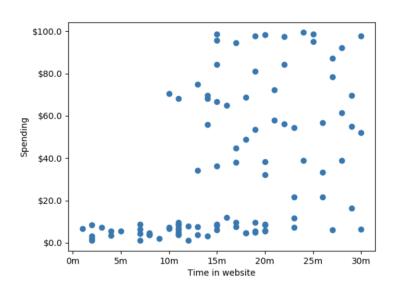
https://matplotlib.org/api/_as_gen/matplotlib.pyplot.subplot.html



Charting – Scatter Plot

```
import csv
 1
      import matplotlib.pyplot as plt
 2
      from matplotlib.ticker import FuncFormatter
 3
 4
 5
     def format minutes(value, pos):
 6
              return '{}m'.format(int(value))
 7
 8
     def format dollars(value, pos):
 9
              return '${}'.format(value)
10
      # read data from csv
11
     fp = open("scatter.csv","r", newline='')
12
      reader = csv.reader(fp)
13
14
     data = list(reader)
15
16
      data x=[]
17
      data y=[]
18
      for x, y in data:
19
              data x.append(float(x))
              data y.append(float(y))
20
21
22
     plt.scatter(data x, data y)
23
24
      plt.gca().xaxis.set major formatter(FuncFormatter(format minutes))
25
      plt.xlabel('Time in website')
      plt.gca().yaxis.set_major_formatter(FuncFormatter(format dollars))
26
      plt.ylabel('Spending')
27
28
29
      plt.show()
```

- To save a plot: plt.savefig(filename)
- Save the plot before you display





End of Day 2

This concludes the Introduction to Python, I hope you enjoyed it.

QUESTIONS?



Where to go from here?



Getting started step by step http://www.python.org/about/gettingstarted/

Run through the python tutorials: http://docs.python.org/tutorial/index.html

Keep the API doc under your pillow: http://docs.python.org/library/index.html

Advanced examples:

http://www.diveintopython.org/toc/index.html

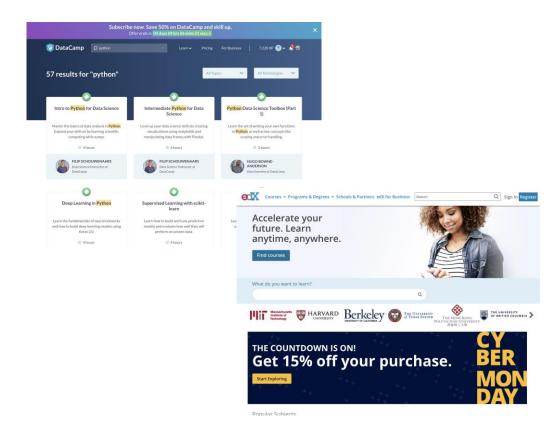
Where to go from here?



MOOC: DataCamp https://www.datacamp.com/

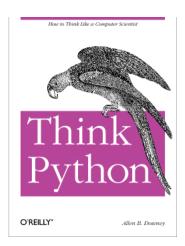
Edx https://www.edx.org/

Udemy (freemium course) https://t.me/freecourse



Where to go from here?



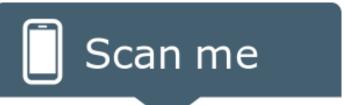


Think Python is an introduction to Python programming for beginners. It starts with basic concepts of programming, and is carefully designed to define all terms when they are first used and to develop each new concept in a logical progression. Larger pieces, like recursion and object-oriented programming are divided into a sequence of smaller steps and introduced over the course of several chapters.

Think Python is a Free Book. It is available under the <u>Creative</u> <u>Commons Attribution-NonCommercial 3.0 Unported License</u>, which means that you are free to copy, distribute, and modify it, as long as you attribute the work and don't use it for commercial purposes. http://greenteapress.com/thinkpython/thinkpython.pdf



Lifelong Learning



https://www.rp.edu.sg/soi/lifelong
 -learning

Short Courses



Z REPUBLIC E

SOI offers an extensive variety of short, industry-relevant courses for ICT skills upgrading and skills acquisition. Our courses are categorized under different areas, ranging from Artificial Intelligence (AI), Business Intelligence/Business Analytics (BI/BA), Business Processes (BP), Unmanned Aerial Vehicle (UAV), IT Security, New/ Digital Media, Software Development to the Internet of Things (IoT). To view our short course offerings, click on the relevant tab below.

Al Data Analytics IT Security DevOps Software Development New/ Digital Media UAV RPA

- + Artificial Intelligence for Everyone A Practical Experience (1 day Beginner)
- + Artificial Intelligence for Techies A Hands-On Approach (1 day Beginner)
- + An Introduction to Code-Free Machine Learning (1 day Beginner)





https://forms.office.com/r/NCmMA7rtUx

Email:

derek_lee@rp.edu.sg jason_lim@rp.edu.sg

Thank you



Source code: https://bit.ly/py-june21