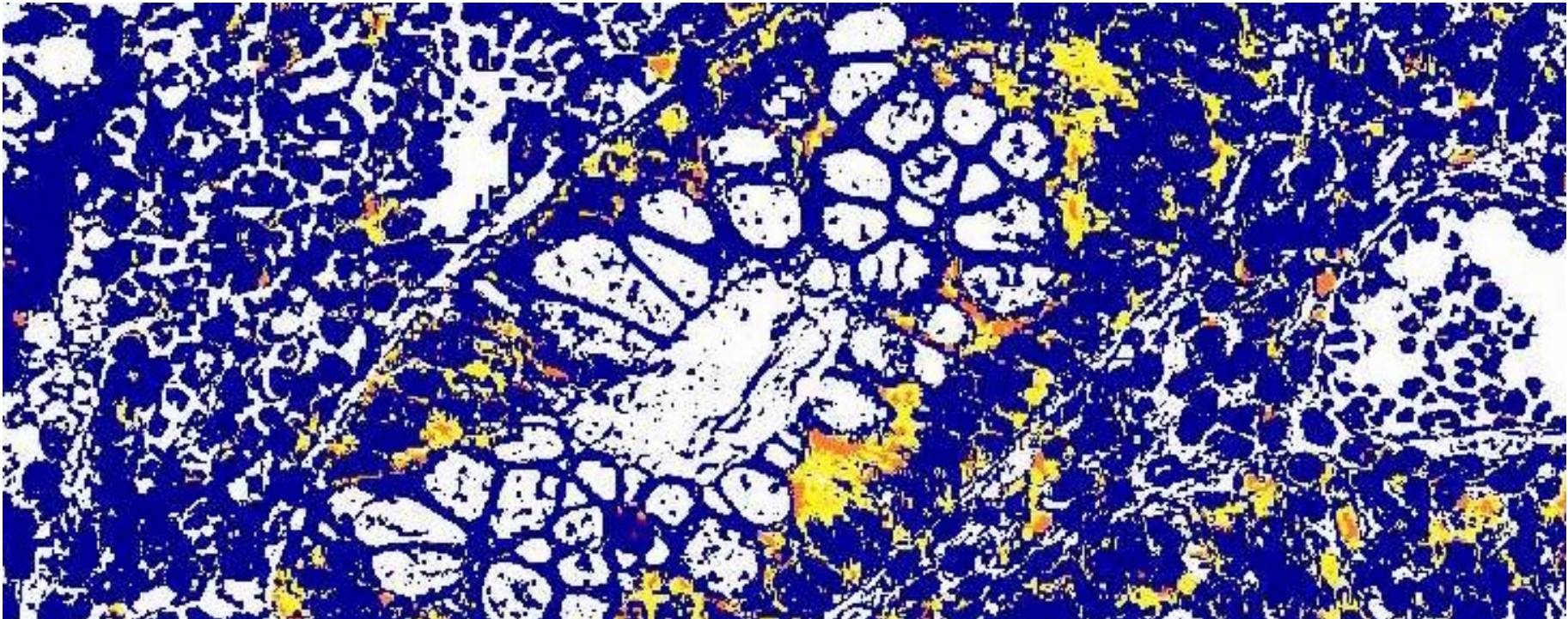




PD04: Teaching an Old Dog New Tricks with Python

Professional Development (PD)

Donald Paterson

Associate Director, Portfolio Lead Statistical Programming
Janssen R&D
High Wycombe, UK

Ulcerative colitis at 40x magnification

Agenda

PD04: Teaching an Old Dog New Tricks with Python

Professional Development (PD)

This presentation is about:

- An Old Dog;
- New Tricks; and
- Python

Agenda

Don

PD04: Teaching an Old ~~Dog~~ New Tricks with Python

Professional Development (PD)

This presentation is about:

- An Old Dog;
- New Tricks; and
- Python

Objective





The Old Dog Don

Skin cells at 20x magnification

My Birthday



Parkrun



Placement Students



Placement Students



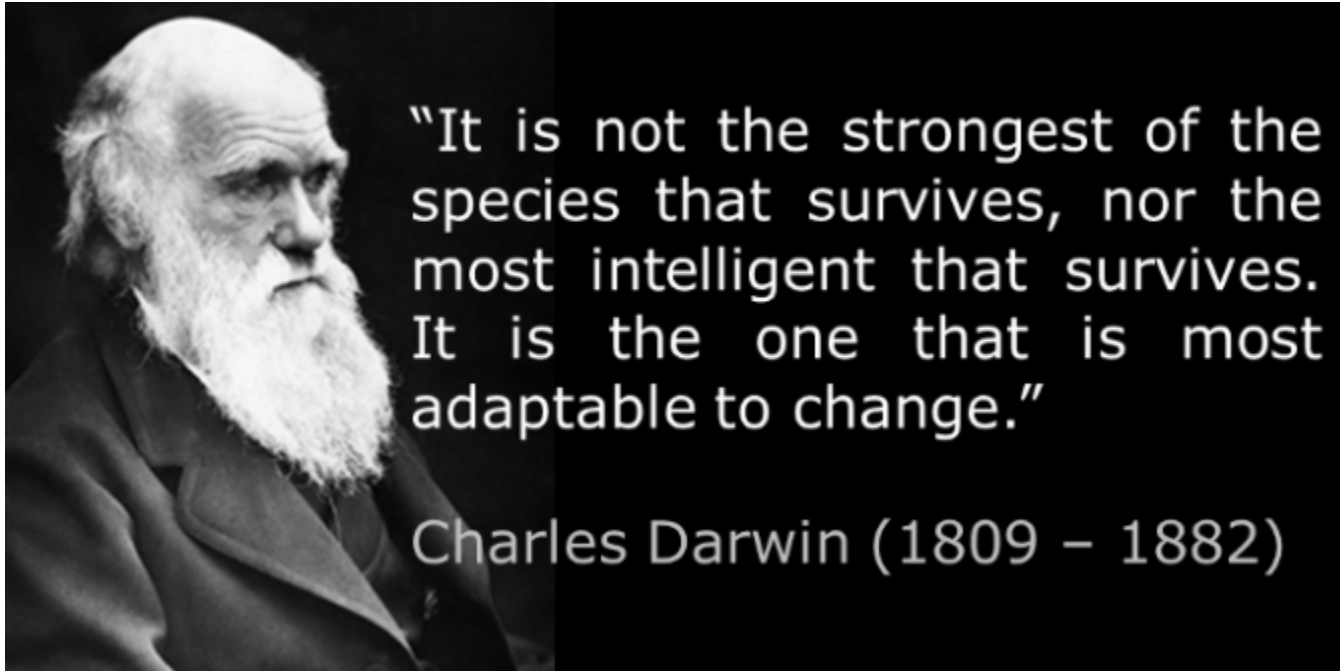
New Tricks

Skin cells at 20x magnification

Change



Change



Awareness



If 2020 was a slide...

Desire



Desire

Energy for Performance®

You manage your time and
your finances...

What would change in your life if you also started managing and expanding your energy?

ENERGY
FOR PERFORMANCE®



Desire



10th–13th November
RAI Amsterdam
The Netherlands

EU 2019

The Clinical Data
Science Conference



9th–13th November

EU Connect 2020

The Clinical Data
Science Conference

Desire



Python

Skin cells at 20x magnification

Automate the boring stuff



Problem

“To maximise the benefits of our development partnership we need to give our partner more visibility of the study data as it accrues”



Dropbox

← → ↕ ↑ > OneDrive - JNJ > Python > PhUSE2020 > dropbox ↕ ↺ Search dropbox 🔍

Python

- .ipynb_checkpoints
- AIT training
- PhUSE2020
 - analysis
 - docs
 - dropbox
 - macros
 - programs
 - rawdata
 - statdata
 - utilities

Name	Status	Date modified	Type
alphabet.sas7bdat	✓	30/08/2020 00:50	SAS Data Set
example.txt	✓	10/08/2020 18:29	Text Document
transfer.sas7bdat	✓	10/08/2020 20:45	SAS Data Set
transfera.sas7bdat	✓	23/08/2020 20:43	SAS Data Set

```
import os

abs_path = "C:/Users/dpaterso/OneDrive - JNJ/Python/PhUSE2020/"
drop_path = "dropbox/"

os.chdir(abs_path + drop_path)
contents = os.listdir('.')
print(contents)
```

['alphabet.sas7bdat', 'example.txt', 'transfer.sas7bdat', 'transfera.sas7bdat']

Dropbox

← → ↕ ↑ > OneDrive - JNJ > Python > PhUSE2020 > dropbox

Search dropbox 🔍

Python

- .ipynb_checkpoints
- AIT training
- PhUSE2020
 - analysis
 - docs
 - dropbox
 - macros
 - programs
 - rawdata
 - statdata
 - utilities

Name	Status	Date modified	Type
alphabet.sas7bdat	✓	30/08/2020 00:50	SAS Data Set
example.txt	✓	10/08/2020 18:29	Text Document
transfer.sas7bdat	✓	10/08/2020 20:45	SAS Data Set
transfera.sas7bdat	✓	23/08/2020 20:43	SAS Data Set

```
mdlist=[]  
  
for x in contents:  
    md = time.strftime('%Y%m%dT%H%M%S', time.gmtime(os.path.getmtime(abs_path + drop_path + x)))  
    mdlist.append(md)  
  
else:  
    print (mdlist)
```

['20200829T235059', '20200810T172917', '20200810T194552', '20200823T194338']

Rawdata

← → ↕ ↑ > OneDrive - JNJ > Python > PhUSE2020 > rawdata > 20200829T235059

Name	Status	Date modified	Type
alphabet.sas7bdat	✓	30/08/2020 00:50	SAS Data Set
example.txt	✓	10/08/2020 18:29	Text Document
transfer.sas7bdat	✓	10/08/2020 20:45	SAS Data Set
transfera.sas7bdat	✓	23/08/2020 20:43	SAS Data Set

```
else:
    print(mdlist)

    ld=max(mdlist)
    print('Last Modified=', ld)

    ldpath=(abs_path + raw_path + ld)
    print('Last Modified Path=', ldpath)

    if not os.path.exists(ldpath):
        os.mkdir(ldpath)

for x in contents:
    dst = ldpath + '/' + x
    copy2(x, dst)
```

```
['20200829T235059', '20200810T172917', '20200810T194552', '20200823T194338']
```

```
Last Modified= 20200829T235059
```

```
Last Modified Path= C:/Users/dpaterso/OneDrive - JNJ/Python/PhUSE2020/rawdata/20200829T235059
```

Run SAS or R or....

←

→

⌵

⬆

📁

>

OneDrive - JNJ

>

Python

>

PhUSE2020

>

rawdata

>

20200829T235059

⌵

📁

PhUSE2020

>

📁

analysis

>

📁

docs

>

📁

dropbox

>

📁

macros

>

📁

programs

⌵

📁

rawdata

⬆

Name	Status	Date modified	Type
 alphabet.sas7bdat	✓	30/08/2020 00:50	SAS Data Set
 example.txt	✓	10/08/2020 18:29	Text Document
 transfer.sas7bdat	✓	10/08/2020 20:45	SAS Data Set
 transfera.sas7bdat	✓	23/08/2020 20:43	SAS Data Set

```
import subprocess
subprocess.call([r"C:/Users/dpaterso/OneDrive - JNJ/Python/PhUSE2020/utilities/adxx.bat"])
import subprocess
subprocess.call([r"C:/Users/dpaterso/OneDrive - JNJ/Python/PhUSE2020/utilities/adsumxx.bat"])
```

📄 adxx.bat - Notepad

File Edit Format View Help

```
cd "C:\Users\dpaterse\OneDrive - JNJ\Python\PhUSE2020\analysis\adam"
"C:\Program Files\SASHome\SASFoundation\9.4\sas.exe" -sysin adxx.sas
```

📄 adsumxx.bat - Notepad

File Edit Format View Help

```
cd "C:\Users\dpaterse\OneDrive - JNJ\Python\PhUSE2020\analysis\adam"
"C:\Users\dpaterse\Anaconda3\envs\rstudio\Scripts\Rscript.exe" adsumxx.R
```

Check for Errors

Log - (Untitled)

```
14
15 proc print data=anal.adxx&lstsnap
NOTE: Writing HTML Body file: sashtml.htm
SYMBOLGEN: Macro variable LSTSNAP resolves to 20200829T235059
16 run;
---
22
202
NOTE: Data file ANAL.ADXX20200829T235059.DATA is in a format that is native to another host, or
the file encoding does not match the session encoding. Cross Environment Data Access will
be used, which might require additional CPU resources and might reduce performance.
ERROR 22-322: Syntax error, expecting one of the following: ;, (, BLANKLINE, CONTENTS, DATA,
DOUBLE, GRANDTOTAL_LABEL, GRANDTOT_LABEL, GRAND_LABEL, GTOTAL_LABEL, GTOT_LABEL,
HEADING, LABEL, N, NOOBS, NOSUMLABEL, OBS, ROUND, ROWS, SPLIT, STYLE, SUMLABEL,
UNIFORM, WIDTH.

ERROR 202-322: The option or parameter is not recognized and will be ignored.
```

```
def search_string_in_file(file_name, string_to_search):
    """Search for the given string in file and return lines containing that string,
    along with line numbers"""
    line_number = 0
    list_of_results = []
    # Open the file in read only mode
    with open(file_name, 'r') as read_obj:
        # Read all lines in the file one by one
        for line in read_obj:
            # For each line, check if line contains the string
            line_number += 1
            if string_to_search in line:
                # If yes, then add the line number & line as a tuple in the list
                list_of_results.append((line_number, line.rstrip()))

    # Return list of tuples containing line numbers and lines where string is found
    return list_of_results

search_string_in_file('adxx.log', 'ERROR')
```

```
[(131,
'ERROR 22-322: Syntax error, expecting one of the following: ;, (, BLANKLINE, CONTENTS, DATA, '),
(135,
'ERROR 202-322: The option or parameter is not recognized and will be ignored. '),
(136, 'ERROR: Errors printed on page 3. ')]
```

Email

```
import smtplib
from email.message import EmailMessage
```

```
msg=EmailMessage()
msg.set_content('I hope that this works')

msg['From']='donald.g.m.paterson@gmail.com'
msg['To']='dpaterso@its.jnj.com'
msg['Subject']="Hello from Python"

mailserver=smtplib.SMTP(host='smtp.gmail.com', port=587)
mailserver.ehlo()
mailserver.starttls()
mailserver.login(MY_ADDRESS, PASSWORD)
mailserver.send_message(msg)
mailserver.quit()
```

[EXTERNAL] Hello from Python



donald.g.m.paterson@gmail.com

To ✓ Paterson, Donald [JRDGB]

I hope that this works

Share

Microsoft Teams

Software



Microsoft Teams is a business-oriented communication and collaboration platform that combines wo... +



Shiny is an open source **R** package that provides an elegant and powerful web framework for building web applications using **R**. **Shiny** helps you turn your analyses into interactive web applications without requiring HTML, CSS, or JavaScript knowledge.

BOOM...we have lift-off



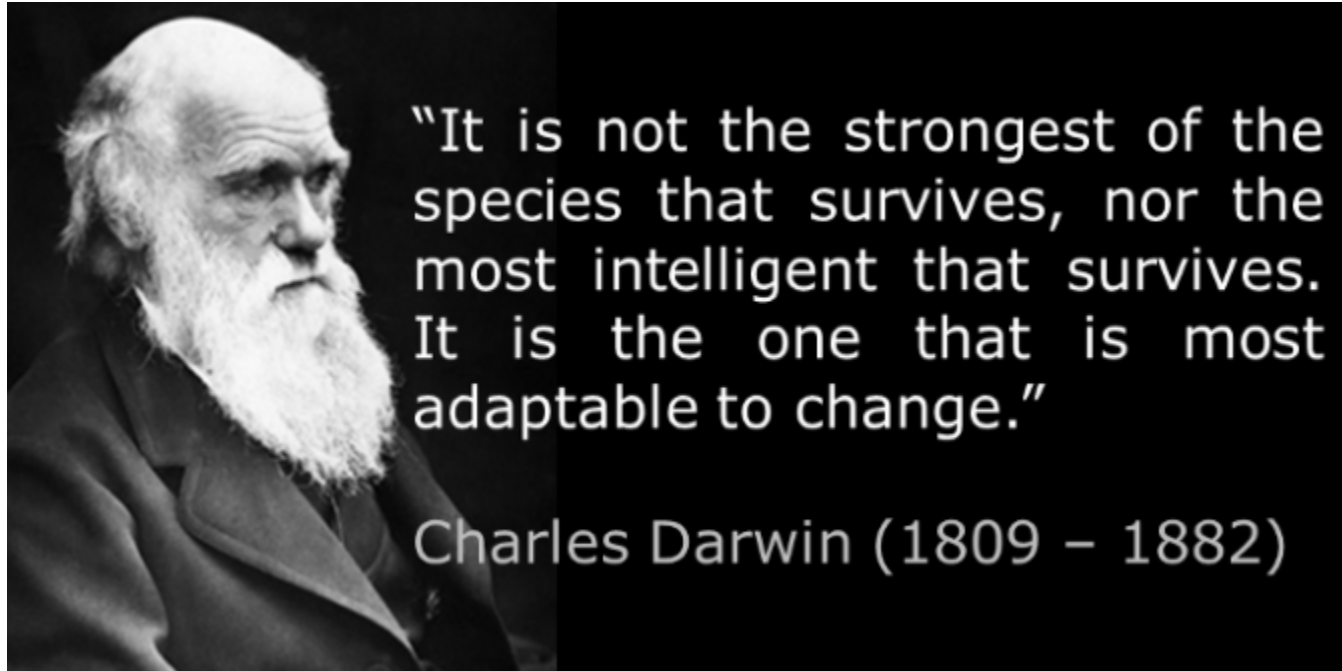
Conclusion

Skin cells at 20x magnification

Placement Students



Change



Desire







**Creating a future where
disease is a thing of the past**