Python, Cloud and Automation

1. Introduction

Jack Minchin

Tourism Economics

2022

Table of Contents

Outline

1. **3rd March - Morning**
   - Part 1: Introduction
     - What is Python?
     - How can it be used?
     - How can I learn?
   - Part 2: Getting Started
     - Downloading
     - Writing
     - Managing Projects
     - Some basic examples
2. **3rd March - Afternoon**
   - Part 3: Automation, Cloud and Examples
     - What are cloud technologies and how are they used?
     - What are pipelines?
     - 3 Examples of automation

What Python is not

Some common misconceptions

- A statistical package
- A program
- A self contained framework

What is Python?

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small- and large-scale projects. (wikipedia)

How can we use it?

1. **Extraction, Transform and Load (ETL) Tasks**
   Reading data from (virtually) any source, transforming into required format and exporting.
2. **Automation of repetitive tasks.**
   Interacting with file system, using libraries to create PowerPoints, interacting directly with excel, read emails - the possibilities really are endless.
3. **Data Visualisation, Analysis and Econometrics**
4. **Full scale applications**
   Larger projects will span multiple categories (e.g Country Profile Report automation)[1]

[1] Current implementation uses another language called JavaScript (Node.js) for the PDF generation.

Where can we use Python, Automation & Cloud

- Automating and creating outputs in PDF, Excel and PowerPoint formats.
  e.g Whitbread Quarterly, Jessie's regular outputs, Sweden Regional TSA, GTS Country Profiles, GCT City Reports
- Python as a 'glue code'
  e.g small scripts to glue wider projects / models together.
- Automating regular data inputs and reading on to the model.
  e.g IATA data processing, STR Data ingestion etc.
- Automating test and checks on model outputs
  e.g APF and GTS output checks (no negatives etc.)
- Storing data in the cloud
  e.g Latest model always uploaded to cloud storage to allow cloud extracts in pipelines, all data extract stored in database to allow quick pulls in model workbooks.

A general note

While Python is great for some use cases, particularly **data processing** and **analysis**, the main objective over the next few days is to bring about more awareness for automation in general, which is language agnostic - but Python is a good place to start.

How to learn Python

1. **Learning-by-doing**
   The **best** way to learn Python is to start writing it. There are multiple courses online, from free YouTube videos to paid courses but in-person courses will not help unless you are spending hours working alone figuring out.
2. **DataCamp.com**
   Combines very short videos with active excercices, it teaches concepts in a concise way and then forces you to get involved. It is also designed for ETL, statistics and data science.