

Algoritmos - Lab1

Santiago Álvarez Sepúlveda

October 1 2018

1 Introducción al Entorno de Desarrollo

En esta lectura se presenta el entorno de desarrollo, el cual está basado en iPython Jupyter Notebook, el cual permite realizar análisis de datos y validaciones estadísticas muy sencillamente. Los primeros objetos descritos en esta lectura pertenecen a los elementos principales de iPython Jupyter Notebooks, algunos comandos y atajos, por ejemplo celdas de código, celdas de texto, ejecución de comandos, simbología de la GUI, cómo usar el lenguaje de programación Python, herramientas de codificación y de documentación para este lenguaje, uso de librerías como Numpy, Pandas y Matplotlib como una formula para trabajar simultaneamente con bases de datos, procesamiento de datos y visualización de los mismos. Por supuesto hay aspectos de Quantopian que se mencionan, los ejemplos relacionados con el uso del lenguaje de programación Python usan datos de los precios y los retornos del stock "MSFT" (Microsoft) en función del tiempo, luego con esta información estimamos algunos valores estadísticos de los precios como su promedio y su desviación estándar y algunos promedios móviles, que son promedios con una ventana de un tamaño dado y que se aplica a una señal como un filtro pasa bajos.

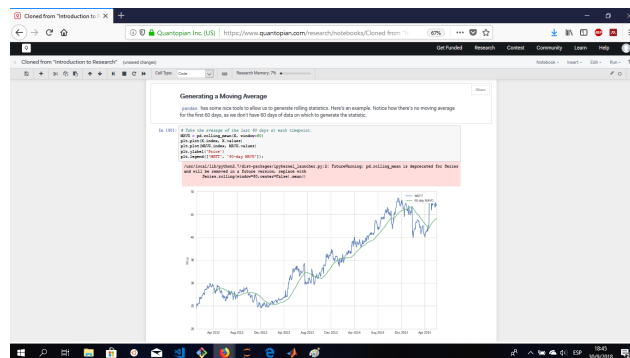
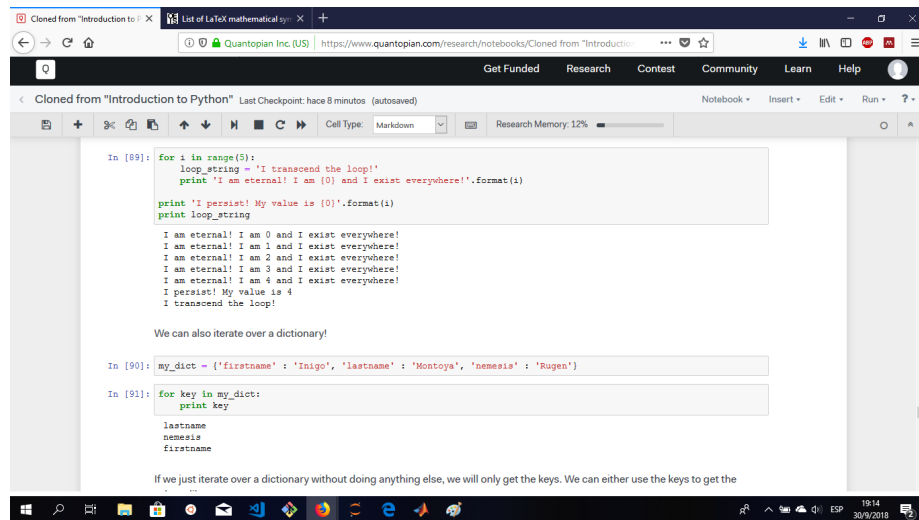


Figure 1: Evidencia de la lectura 1 clonada en Quantopian.

2 Introducción a Python

Ya que todo en Quantopian está programado usando Python, esta resulta ser una lectura muy importante a la cual hay que prestar atención. Como muchas otras introducciones al lenguaje de programación Python, esta muestra las principales sintaxis y las semánticas detrás de cada comando básico que Python entiende. Algunos de estos comandos básicos son los comentarios de una y multiples lineas, como definir e imprimir una variable, el uso de diferentes tipos de datos nativos como enteros, floats, strings, booleanos, algunas operaciones matemáticas como suma resta multiplicación, división, módulo y operaciones de punto flotante, y algunas de las más notables funciones que trae por defecto (built-in) el lenguaje. También se presenta la librería math dónde se definen multiples valores como números irracionales y funciones que permiten realizar procesamientos más extensos de datos en Python. También se muestran ejemplos relacionados con inicialización y manejo de listas, tuples, conjuntos y diccionarios, que son estructuras de datos que permiten organizar información de forma intuitiva y que contienen muchos métodos que permiten procesar y acceder a esta información. Entre muchas otras cosas es importante recalcar que se mencionan las sintaxis correctas para las sentencias If-Else y para los ciclos de repetición, también a la declaración y calls de funciones y como obtener los valores que retornan.



```
In [89]: for i in range(5):
        loop_string = 'I transcend the loop!'
        print 'I am eternal! I am {} and I exist everywhere!'.format(i)

        print 'I persist! My value is {}'.format(i)
        print loop_string

I am eternal! I am 0 and I exist everywhere!
I am eternal! I am 1 and I exist everywhere!
I am eternal! I am 2 and I exist everywhere!
I am eternal! I am 3 and I exist everywhere!
I am eternal! I am 4 and I exist everywhere!
I persist! My value is 4
I transcend the loop!

We can also iterate over a dictionary!

In [90]: my_dict = {'firstname': 'Inigo', 'lastname': 'Montoya', 'nemesis': 'Bugeis'}

In [91]: for key in my_dict:
        print key

lastname
nemesis
firstname

If we just iterate over a dictionary without doing anything else, we will only get the keys. We can either use the keys to get the
```

Figure 2: Evidencia de la lectura 2 clonada en Quantopian.

3 Introducción a Numpy

Numpy adiciona un apoyo para tratamiento de arreglos multidimensionales y con una colección de implementaciones de funciones matemáticas capaces de procesar estos arreglos multidimensionales. Lo que se busca con esto es realizar cálculos a partir de los axiomas del álgebra lineal, por esta razón esta lectura es una colección de ejemplos del uso de elementos de los espacios vectoriales por medio del álgebra lineal a través de las implementaciones de Numpy.

Para empezar es necesario poder definir arreglos multidimensionales con Numpy, los principales atributos, métodos y modos de indexación que esta clase de arreglos tiene. i.e. `x.shape`, `len(x)`, `x[:-1]`, etc. También se presentan las sintaxis para hacer uso de las funciones matemáticas multidimensionales, esto resulta ser relevante a la hora de trabajar con un portafolio de inversión, donde se deben comparar indicadores financieros similares de un conjunto de compañías, por lo que conviene agrupar esta información en matrices.

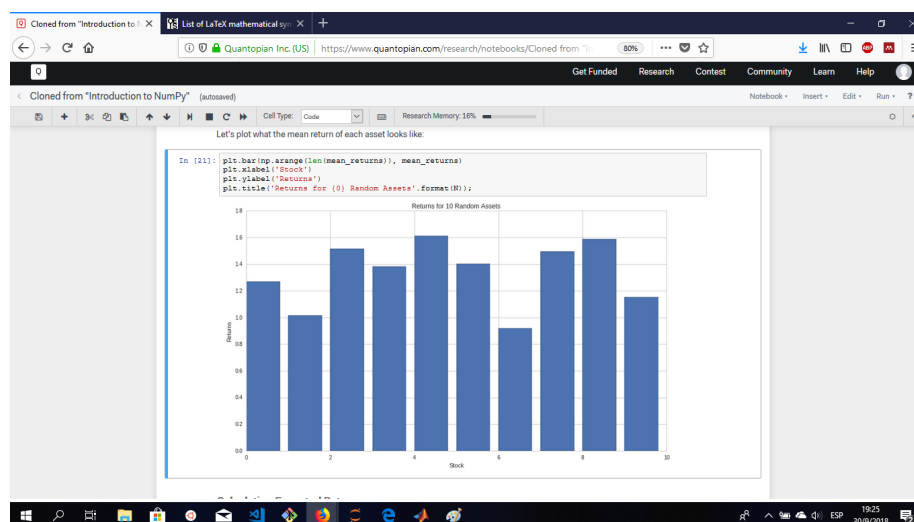


Figure 3: Evidencia de la lectura 3 clonada en Quantopian.

4 Introducción a Pandas

Esta lectura da una introducción a la librería Pandas, la cual es una poderosa colección de estructuras de datos, implementadas para el lenguaje de programación Python, que facilitan el manejo de datos, principalmente son de gran utilidad las **Series**, arreglos 1-Dimensionales que pueden almacenar cualquier tipo de dato, y los **DataFrames**, que son arreglos 2-Dimensionales de tamaño mutable y puede entenderse como un diccionario para ordenar Series. Otro aspecto importante sobre esta librería es que mantiene una relación muy cercana con Numpy, por lo que es posible tratar estos objetos como arreglos multidimensionales y realizar procesamientos lógicos y aritméticos a estos sets de datos.

En la lectura se presentan algunos ejemplos de como definir estos tipos de datos propios de Pandas, sus principales métodos y atributos, y algunos ejemplos del procesamiento estadístico y matemático, filtrado, indexado, preprocesamiento, *slicing* (en español revanado), preprocesamiento y visualización de los conjuntos de datos suministrados por Quantopian almacenados en estructuras de datos de Pandas.

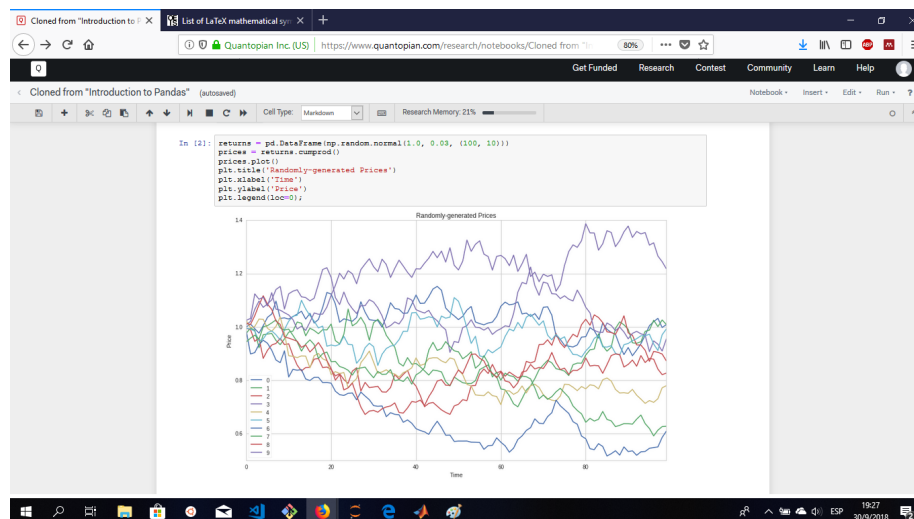


Figure 4: Evidencia de la lectura 4 clonada en Quantopian.

5 Visualización de Datos

En esta lectura se presentan una gran colección de ejemplos sobre el uso de la librería Matplotlib y la manera de presentar de diferentes maneras la información presente en un conjunto de datos almacenados en un arreglo multidimensional. Se evidencia como calcular parámetros estadísticos y probabilísticos como histogramas, histogramas acumulados, grafico de puntos, graficos de barras, interpolaciones lineales, entre otros aspectos relacionados.

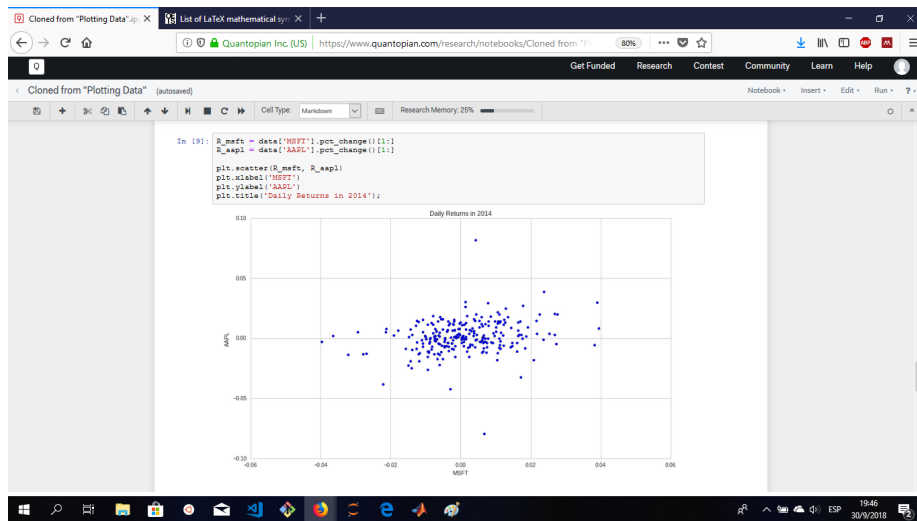


Figure 5: Evidencia de la lectura 5 clonada en Quantopian.