Ya has visto el listado de lenguajes y tipologías de softwares que podemos utilizar en el mundo de los datos. Te los recuerdo aquí:

- Hoja de cálculo: Excel
- Software de BI: powerBI y Tableau
- Softwares estadísticos sin programación: SPSS, JASP (gratuito), Minitab
- **Software de ciencia de datos** (Estadística, ML y DL ; programación): R y Python
- Herramientas de bases de datos: SQL
- Herramientas cloud: Azure, Google cloud, Amazon web services
- Herramientas de Big Data: Spark, Hadoop

Para empezar a dominar la ciencia de datos la clave es dominar R y/o Python para poder manipular datos, explorarlos y extraer patrones de los datos.

Aunque Excel es un software transversal que es importante conocer y siempre usarás :)
(si eres investigador sobre todo de proyectos de ciencias de la vida puede que con softwares estadísticos sin programación puedas avanzar bastante bien).

En esta hoja de trabajo vamos a tomar la decisión de empezar con R o con Python.

Es una pregunta muy recurrente entre mis alumnos y clientes y mi opinión es empezar por una y después dar el salto a la otra ya que son herramientas que se pueden comunicar, y complementarias.

Voy a tratar de darte información para que tomes la decisión en esta hoja de trabajo.

¡Mira!



- Gran flexibilidad ya que te permite programar y crear tus propias rutinas
- Paquetes de gran potencia en ML y DL
- Muy demandado en la industria
- Gran versatilidad
- Mejor documentación y recursos que R
- Flojo en estadística comparado con R
- Gran potencia en machine learning y Deep learning, series de tiempo y computer vision
- El futuro de la ciencia de datos en la industria



- Gran flexibilidad ya que te permite programar y crear tus propias rutinas
- Paquetes muy buenos de estadística y específicos
- Ideal para investigaciones científicas
- Muy buenos paquetes de machine learning
- Se queda corto en Deep learning, series temporales y computer vision
- Software de gran potencia en investigación científica

Ya conoces la herramienta y sus características principales que las definen. Ahora es momento de definirte a ti para poder tomar la mejor decisión.

	Preguntas sobre tu orientación profesional:	**SI**	**NO**
1	Me dedico a investigar parte de mi tiempo	X	
2	Me considero científico	X	
3	¿Trabajo o quiero trabajar en el mundo de la academia (universidad)?		X
4	Mi intención es convertirme en científico de datos para trabajar para proyectos de empresas		X
5	Quiero trabajar para proyectos reales y aplicables	X	
6	Mi perfil es científico más que técnico	X	
7	Quiero dedicarme plenamente como científico de datos en la industria		X

Teniendo en cuenta que:
R tiene un perfil más científico y de estadística.
Python tiene un perfil más técnico y de implementación en la industria / empresa

Mira tus respuestas, si son más científicas utiliza R. Si son más técnicas y a la práctica, utiliza Python

¿Qué herramienta vas a utilizar durante el programa? ¿R o Python?

Voy a utilizar : R
ACTUALMENTE TRABAJO EN UNA CONSULTORÍA AMBIENTAL HACIENDO INVENTARIOS E INFORMES DE FAUNA Y ASÍ MISMO ESTOY EN MEDIO DE UNA INVESTIGACIÓN CIENTIFICA (etapa de análisis de videos y posteriormente análisis de resultados) CON LO QUE CREO QUE PARA AMBAS COSAS LA MEJOR OPCIÓN SERÍA R, ¿NO?
Para la segunda sé que sí e imagino que ¿para la primera también?

¡Bien!

Ahora que ya sabes que herramientas de ciencia de datos vas a empezar a trabajar.

El siguiente paso es tener en mente en qué otra herramienta quieres especializarte ya que R y Python son muy flexibles y poderosas pero siempre es bueno combinarla con otras herramientas o aplicaciones.

Te propongo una tabla resúmen con las herramientas y decides en qué tipo de herramienta quieres especializarte:

Tipo de herramienta	**Herramienta escogida**	**Obligatorio para ser científico de**	**Me voy a especializar en**

		datos	(marca con una X)
Hoja de Cálculo	Excel	SI	X
Ciencia de datos	Elige R o Python según la decisión anterior	SI	R
Bases de datos SQL	SQL	SI (primero aprende R o Python y después SQL)	
BI - Business Intelligence	Elige Power BI o Tableau	NO	(puedes marcar con una X)
Softwares estadísticos sin programar	Elige SPSS, Minitab, JASP	NO	(puedes marcar con una X) X ¿SPSS?
Herramientas BIG DATA	Hadoop, docker, pySpark	NO	(puedes marcar con una X)
Herramientas Cloud	Elige Amazon Web Services, Google Cloud, Azure	NO	(puedes marcar con una X)

¡Bien!

Ahora ya tienes mucho más claro tu recorrido en tema de ciencia de datos.

Ya has visto que lo más importante es dominar R y/o Python y después pasar a usar bases de datos con SQL ya que en muchas ocasiones necesitamos trabajar la consulta en bases de datos. Evidentemente, Excel es una herramienta importante ya que se utiliza como comodín para pequeños reportes, tablas y gráficos rápidos.

El siguiente paso es especializarte en alguna rama concreta según los grupos planteados.

¡Genial!

Ahora ya estás mucho más sitiado con temas de herramientas :). A seguir el camino.

Es una gran decisión y lo acabas de hacer.

Ahora ya no pienses más en herramientas sino céntrate en R o Python y aprende paso a paso con la ayuda del programa.

¡Celébralo como te guste! Una infusión, una cerveza, un zumo de zanahorias o lo que quieras :)

1 abrazo y hasta pronto.
Brindo por tus éxitos :)

Jordi Ollé