



ES



Empleos

Mis Cursos

Blog

Agenda



Buscar en Platzi



Curso de Big Data y Ciencia de Datos

Artículo

Flujo de trabajo de un Data Scientist



Yeison Daza 23 PlatziRank ⌚ Oct. 11, 2016

La labor principal de un data scientist es encontrar valor para las empresas o personas a través de datos, para esto se debe seguir un flujo de trabajo que logre convertir estos datos en información significativa.

El flujo de trabajo se divide en seis pasos, veamos qué se realiza en cada uno.

1. Decide los objetivos

Como primer paso, debemos definir los objetivos de nuestro proceso, estos generalmente se definen como preguntas.

Estos objetivos deben ser claros, medibles y concisos. Debemos tener cuidado en estos debido a que todo el proceso depende de este primer paso.

La naturaleza de los datos y el proceso va a diferir del tipo de objetivos o preguntas que te plantees.

2. Establece prioridades de medición

Ahora que has establecido tus objetivos, debemos definir qué deberíamos medir basados en el objetivo que tenemos.

Debemos tener claro qué tipo de datos necesitamos para lograr el objetivo.

Decide cómo medir: Es muy importante que decidas qué parámetros vas a usar para medir tus datos antes de comenzar a recolectarlos, cómo mides tus datos juega un rol importante en el análisis de los datos.

3. Recolección de datos

Ahora que ya tienes prioridades y parámetros para medir, será más fácil recolectar datos, recuerda que tienes muchas formas de recolectar estos datos y es importante tener claro los datos que se necesitan tener definidos en los dos pasos anteriores.

4. Limpieza de datos

SIGUE CON:

[Workflow Típico de un Data Scientist](#)

> Los datos que has recolectado no son necesariamente útiles, en este proceso debemos asegurarnos que los datos menos útiles no se encuentren en la fase de análisis.

9 Cuando tienes datos no deseados en el sistema, esto afectará en la calidad de las decisiones
10 Tener datos de calidad es importante para tomar mejores decisiones, este proceso requiere
11 bastante tiempo es una buena idea automatizar este proceso.

12 **5. Análisis de datos**

13 Ahora que tenemos los datos necesarios, es tiempo para procesarlos. Existen diferentes
14 técnicas para analizar los datos.

15 **6. Interpretación y comunicación de resultados**

16 Una vez que has analizado los datos, es tiempo de interpretar los resultados y comunicarlos a
17 las personas involucradas dentro de la empresa.

18 Estos seis pasos representan el flujo normal de un data scientist. Durante el transcurso de esta
unidad, veremos cada paso a profundidad y las herramientas necesarias para hacer un trabajo profesional.

