



Herramientas – Ciencia de Datos

DESCRIPCIÓN GENERAL

Contexto general

Para tomar una buena decisión en cuanto a las herramientas a utilizar para cada caso, es importante tomar en cuenta el alcance de los diversos lenguajes de programación. El campo de la ciencia de datos no es la excepción y las herramientas de tratamiento, análisis, clasificación y visualización de datos abundan.

<https://medium.com/@danielVer91/las-6-herramientas-que-dominan-la-ciencia-de-datos-de-c%C3%B3digo-abierto-457757b3c6d6>

La función principal de los lenguajes de programación es escribir programas que permiten la comunicación usuario-máquina.

Herramientas

R

- R es un sistema para análisis estadísticos y gráficos.
- El diseño de los programas se basa más en los datos y su estructura.
- R es un lenguaje **orientado a Objetos**: bajo este complejo término se esconde la simplicidad y flexibilidad de R.
- Orientado a Objetos significa que las variables, datos, funciones, resultados, etc., **se guardan en la memoria activa del computador en forma de objetos con un nombre específico**. El usuario puede modificar o manipular estos objetos con diferentes operadores y funciones (que a su vez son objetos).
- Principal ventaja: Simple e intuitivo.
- Es un lenguaje interpretado, lo que significa que no se necesita compilar el código fuente para poder ejecutarlo, lo que ofrece ventajas como la rapidez de desarrollo.

El código fuente de un programa **está escrito en un lenguaje de programación determinado**, sin embargo este tipo de lenguaje no puede ser ejecutado directamente por el computador, sino que debe ser traducido a otro lenguaje que el ordenador pueda ejecutar más fácilmente (compiladores).

Herramientas

Python

- Es un lenguaje interpretado.
- Python dispone de un intérprete por línea de comandos en el que se pueden introducir sentencias. Cada sentencia se ejecuta y produce un resultado visible, que puede ayudarnos a entender mejor el lenguaje y probar los resultados de la ejecución de porciones de código rápidamente.
- Ventajas:
 - La cantidad de librerías que contiene
 - Tipos de datos
 - Funciones que ayudan a realizar muchas tareas habituales sin necesidad de tener que programarlas desde cero (tensorflow, keras, sklearn, etc.)
 - Sencillez y velocidad para crear programas
 - Gratuito

Herramientas

AWS / IBM Cloud

Colección de servicios de computación en la nube pública que en conjunto forman una plataforma de computación en la nube.

Herramientas

Github

- GitHub es una plataforma de desarrollo colaborativo de software para alojar proyectos utilizando el sistema de control de versiones Git.

Git tiene como propósito es llevar registro de los cambios en archivos de computadora y coordinar el trabajo que varias personas realizan sobre archivos compartidos.

- GitHub aloja tu repositorio de código y te brinda herramientas muy útiles para el trabajo en equipo, dentro de un proyecto.
- Además de eso, puedes contribuir a mejorar el software de los demás. Para poder alcanzar esta meta, GitHub provee de funcionalidades para hacer un **fork** y solicitar **pulls**.
- Realizar un fork es simplemente es generar una copia del repositorio en tu cuenta, para eliminar algún bug o modificar cosas de él. Una vez realizadas tus modificaciones puedes enviar un pull al dueño del proyecto. Éste podrá analizar los cambios que has realizado fácilmente, y si considera interesante tu contribución, adjuntarlo con el repositorio original.

<https://conociendogithub.readthedocs.io/en/latest/data/introduccion/>

Muchas más...





Storytelling

IMPORTANCIA

¿Qué es storytelling?

- **Storytelling es el arte de contar, desarrollar y adaptar “historias”.**
- Al contar buenas “historias”, garantizas que estás creando un material único. Por más que la publicación hable de un tema desgastado o de conocimiento general, el contenido abordará una perspectiva única: la tuya.
- **El modelo Pixar de contar historias**
 - 1° acto: presentación**
 - * Érase una vez...
 - * Los personajes son presentados en su mundo, con toda la rutina sucediendo normalmente hasta que...
 - * **¡Tenemos el evento que anuncia el conflicto!**
 - 2° acto: la jornada**
 - * Debido al conflicto, tenemos **una serie de otros sucesos** que se convierten en obstáculos. Cada obstáculo genera uno nuevo y esto sigue sucediendo hasta llegar al conflicto final.
 - * En esta jornada, **acompañamos la transformación del personaje principal**, que llega al fondo del pozo antes de lograr su transformación, para luego, **solucionar el conflicto**.
 - 3° acto: el cambio**
 - * Son presentados los personajes en su nueva rutina, ahora transformados por la **resolución del conflicto**.
 - * Teniendo como base este cambio, el **mensaje es transmitido** emocionando e impactando a la audiencia.

Prácticas de Storytelling en Proyectos de Ciencia de Datos

Todo dependerá de lo que ofrezcas para la audiencia. Es lógico que, para producir una buena narración, necesitas de un tema que sea importante, que tenga problemas reconocidos por tu público y que vean en ti un camino para solucionarlos.

- **Llevar al lector de un punto A hasta un punto B:** tenemos un objetivo claro (contestar nuestra pregunta) y un mensaje final (conclusiones, propuestas, resultados, análisis, etc) que debe ser fácilmente transmitido y reproducido.
- **Desarrollar tu creatividad:** Haciendo análisis interesantes, generando propuestas interesantes, visualizaciones, presentaciones, etc.
- **Transmite sensaciones positivas con el contenido**
De acuerdo con un artículo publicado por la Scientific American, las historias que estimulan emociones positivas son más compartidas que las publicaciones que ocasionan sentimientos negativos (¿Qué se logró?).
- **Apuesta por la interactividad:** El principal objetivo es presentar información que supuestamente sería aburrida de una manera mucho más entretenida y original, que logre conducir al usuario hasta el final de un relato, estimulando incluso su deseo de compra.

Tips

<https://www.youtube.com/watch?v=f2PbOX0LpLo>

From Numbers to Narrative: Data
Storytelling – Isaac Reyes | ODSC East 2018