

Comunicación de datos utilizando herramientas interactivas

Visualización: Comunicación y Presentación de Resultados

Descripción

Actualmente la capacidad de cómputo y almacenamiento de datos está aumentando a un ritmo vertiginoso. Esto unido a la gran cantidad de información que se puede obtener sobre cualquier área de aplicación hace que el análisis, sintetización y presentación de resultados este cobrando mucha importancia. Un ejemplo de ello es el actual boom en el uso de infografías para la presentación de información en el periodismo de datos. Por ejemplo:

- http://www.nytimes.com/interactive/2012/08/05/sports/olympics/the-100-meter-dash-one-race-every-medalist-ever.html?_r=0
- <http://www.informationisbeautiful.net/>

Otro ejemplo, es la iniciativa Open Data que posibilita que los datos y la información, especialmente aquellos provenientes de administraciones públicas, se publiquen de forma abierta, regular y reutilizable para todo el mundo, sin restricciones de acceso, copyright, patentes u otros mecanismos de control. Esto aumenta la cantidad de datos disponibles, fomentando su análisis, transparencia, eficiencia y desarrollo económico. En base a los datos publicados, su análisis y comunicación puede enriquecer la información con nuevos datos (aportar valor añadido) generando nuevas aplicaciones y servicios y hasta nuevos modelos de negocio.

En definitiva, no solo se trata de almacenar y analizar los datos en bruto, sino saber tratarlos correctamente y presentarlos adecuadamente para obtener el máximo beneficio. En concreto, en esta práctica se abordará el problema de analizar, presentar y comunicar datos utilizando herramientas interactivas.

Objetivos parciales

1. El primer paso es seleccionar la fuente de datos a analizar. Se recomienda la utilización de fuentes de datos Open Data, pudiendo elegir los datos concretos interesantes para el alumno. Por ejemplo:
 - a. <http://datos.gob.es/datos/>
 - b. <http://datos.madrid.es/portal/site/egob>
 - c. <http://opendata.emtmadrid.es>
 - d. <http://www.ine.es/inebmenu/indice.htm>

- e. <http://www.usaspending.gov/>
- f. <https://data.cityofnewyork.us/>

Adicionalmente, se puede utilizar cómo fuente de datos una propia perteneciente a alguna empresa del sector de los alumnos implicados, en caso de interés particular.

2. En caso de ser necesario, limpieza de datos, obtención de missing values, inclusión de metadatos y realización de estadísticas básicas para conocer el conjunto de datos de partida.
3. Diseño de visualización interactiva incluyendo un amplio rango de datos, por ejemplo datos multidimensionales, series temporales, datos espaciales (mapas), etc. El objetivo es conseguir un mapeo visual efectivo que refleje un análisis del conjunto de datos de partida, **obteniendo por lo menos tres titulares que se puedan sacar de los datos**. Como base para su realización se recomienda utilizar una herramienta de visualización interactiva como Tableau, Bokeh, R Shiny, ggplot, Plotly, etc. tratando de incluir:
 - a. Mapping intuitivos
 - b. Diferentes tipos de vistas (o permitirlos mediante interacción)
 - c. Leyendas o claves
4. Incorporación de interacción a la visualización, incluyendo diferentes técnicas como filtering, brushing, panning, linking, etc.
5. Exposición del trabajo realizado. Una vez entregado el trabajo se realizará una sesión en la cual se presentaran los resultados a todos vuestros compañeros focalizando los resultados en los tres titulares o datos obtenidos más importantes durante el análisis. Se abrirá un turno de preguntas en la que se espera que el resto de compañeros contribuyan a la discusión. La calidad de las intervenciones, será también base de la evaluación. Una exposición y discusión excelente se caracteriza, entre otras cosas, por:
 - a. Realizar paralelismos y aplicar conocimientos conceptuales.
 - b. Análisis profundo de los conceptos y casos, apoyado por hechos y experiencia.
 - c. Crítica constructiva.
 - d. Integración de los comentarios de los compañeros.

Nota: No se debe hacer un programa separado para cada objetivo, sino un único programa genérico que cumpla con todos los objetivos simultáneamente.

Entrega de prácticas

La entrega de prácticas se hará a través de aulaVirtual en las fechas anunciadas en el mismo. Se debe entregar un fichero .zip incluyendo todo el proyecto realizado (archivos .twbx para Tableau, o archivos .py para Bokeh, etc.). Las leyendas, cuadros de texto y código deben estar debidamente comentados.

Evaluación de la práctica

La práctica se evaluará comprobando el cumplimiento de los distintos objetivos, y valorando la visualización en su conjunto. Adicionalmente, se valorará la defensa del trabajo realizado así como las intervenciones en la discusión.

Autoría de la práctica

La práctica se puede realizar individualmente, aunque se permiten la existencia de grupos de 2 personas siempre y cuando incluyan a gente de diferentes perfiles (p.ej. perfil informático y matemático).