

PRÁCTICA 5.8 - ANÁLISIS ORIGINAL DE DATOS

1.1.- Objetivo

Identificar qué visualizaciones son las más apropiadas para transmitir un mensaje y seleccionar las visualizaciones que describan mejor las relaciones en los datos. También identificar, usando el mismo conjunto de datos, como trabajar en aplicaciones diferentes.

1.2.- Tarea

Realizar un análisis original sobre un conjunto de datos de su elección. El conjunto de datos puede proceder de cualquier fuente, siempre y cuando no hayan trabajado con ellos en este curso. Las posibles fuentes de datos son:

- <https://ec.europa.eu/eurostat/web/main/data/database>
- <https://datos.gob.es/es/catalogo>
- Fuentes de datos gubernamentales abiertos como Census.gov, Data.gov, OpenDataPhilly.org, Washington DC Open Data (<http://opendata.dc.gov>), NYC Open Data (<https://nycopendata.socrata.com/>), Iowa Data (<https://data.iowa.gov>) o Socrata (<https://opendata.socrata.com/>).
- Conjuntos de datos del Pew Research Center (<http://www.pewresearch.org/data/>)
- Estadísticas deportivas, como las de la Major League Baseball (<http://www.seanlahman.com/baseball-archive/statistics>) o la Liga Nacional de Fútbol Americano (<http://nflsavant.com/>)
- Datos electorales, como los de la FEC (<http://www.fec.gov/pubrec/electionresults.shtml>)
- Conjuntos de datos sanitarios de <https://data.medicare.gov>, <https://www.healthdata.gov/>, o cualquier otra fuente
- Conjuntos de datos del Banco Mundial (<http://data.worldbank.org/>)
- Conjuntos de datos de FiveThirtyEight (<http://fivethirtyeight.com/datalab/>)
- Cualquier conjunto de datos que puedas encontrar en la web.

Tenéis que recopilar los datos en bruto (probablemente en un formato de Excel o csv) por vosotros mismos y realizar vuestros propios análisis. Aunque **no es obligatorio**, la **integración de datos de más de una fuente de datos** dará lugar a una **mayor calificación**.

Tu análisis debe demostrar claramente las herramientas y técnicas que has conocido en esta parte del curso. Puede adoptar cualquier forma que desee (por ejemplo, comparación de los promedios a través de categorías, mapeo de datos geográficos, resultados financieros, desarrollo y visualización de KPI, ...). **Debéis utilizar Tableau** para realizar este mismo análisis.

1.3.- Entregables

Esta tarea debe realizarse de forma individual. Complete las instrucciones de este documento y envíe a Moodle:

- un archivo **PDF** que conste de lo siguiente:
 - **Título** de su análisis
 - **Escenario** o establecimiento de objetivos - ¿Qué pregunta quiere responder y por qué es importante?
 - **Recopilación de datos** - ¿Dónde se consiguieron los datos? ¿Son fiables?
 - **Limpieza y transformación** - ¿Se puede mejorar los datos? ¿Es preciso la creación de nuevos indicadores? ¿Existen valores atípicos?

- **Descripción de los datos** - ¿Cuáles son los elementos clave y dónde/cómo los ha obtenido?
- **Análisis y resultados con visualización** - Cuanto más, mejor. Visualícelos de manera que resulten informativos y visualmente atractivos. Haga capturas de las imágenes e introdúzcalas en el informe.
- **Conclusiones** - ¿Qué ideas o decisiones podemos tomar a partir de tus conclusiones?
- El **archivo de Tableau** utilizado para el análisis.
- Los **archivos originales** (csv, Excel, ...) de los que se parte para el análisis.

1.4.- Instrucciones de envío

Envíe el archivo PDF, el archivo Tableau y los **archivos originales** a través de Moodle **antes de la fecha límite**.

La entrega tardía está permitida, pero habrá una **penalización del 10% por cada 12 horas** hasta un máximo de 48 horas de demora.