

Visualización de la Información

Proyecto de Medio Curso

5/abril/2022

Introducción

Muchas actividades se pueden modelar por medio de una gráfica: entidades que se relacionan entre ellas. Por ejemplo, la Internet se puede modelar como una gráfica en donde cada computadora es un nodo y sus conexiones como aristas que conectan a los nodos. Otro ejemplo lo demuestran las rutas de vuelo entre aeropuertos: los nodos o vértices son los aeropuertos y las aristas conectándolos son las rutas de vuelo; en este esquema, se deben preservar las relaciones geográficas entre diferentes vértices. Cuando se trata de redes complicadas, es importante ilustrar los atributos de la red de manera efectiva de tal forma que su interpretación sea rápida y que se pueda reconocer la información importante sin confusión. En el caso de aeropuertos es importante lograr que la audiencia capte rápidamente la red en su contexto geográfico.

Descripción

Los datos para el proyecto provienen del Informe de Estadísticas de Carga y Pasajeros Aéreos Internacionales de EE.UU. Como parte del programa T-100, el USDOT recibe informes de tráfico de las aerolíneas estadounidenses e internacionales que operan desde y hacia los aeropuertos estadounidenses. Estos datos cubren los años 1990 a 2020 (US Air Traffic 1990-2020).

Hay dos conjuntos de datos disponibles:

- Pasajeros (*International Report Passengers*): datos sobre el número total de pasajeros para cada mes y año entre un par de aeropuertos, según los servicios de una aerolínea en particular. Los datos de carga y pasajeros aéreos internacionales de EE. UU. son confidenciales durante un período de 6 meses, después del cual se pueden divulgar. Como resultado, los informes trimestrales y los archivos de datos sin procesar del año hasta la fecha/año calendario disponibles aquí siempre tendrán un retraso de dos trimestres.
- Salidas (*International Report Departures*): datos sobre todos los vuelos entre puertas de enlace de EE. UU. y fuera de EE. UU., independientemente del origen y el destino. Cada observación brinda información sobre una aerolínea específica para un par de aeropuertos, uno en los EE. UU. y el otro fuera. Existen tres columnas principales registran el número de vuelos a) Programados, b) Chárter y c) Totales.

Objetivo

El proyecto consiste en realizar una visualización de los datos para lo cual se deben presentar varios métodos de visualización, pero es indispensable que uno de ellos sea que visualice redes con posicionamiento geoespacial. Hay muchos datos en la base de datos y por ello, quizás, se deben filtrar o limpiar además de que puede ser útil buscar alguna correlación para poder hacer una visualización más informativa.

Requisitos

- a) Reporte que debe seguir una estructura similar a la de una tesis o tesina:
 - Introducción (sección que debe relatar el contexto del conjunto de datos y el problema o fenómeno que se desea analizar).
 - Metodología (descripción detallada de los métodos que se han usado para analizar y visualizar los datos).
 - Resultados y Discusión (presentación de resultados y discusión de ellos).
 - Conclusiones.
 - Bibliografía.

El reporte no debe incluir ni las instrucciones ni los archivos del código.

- a) Realizar el trabajo en equipo con tres o cuatro integrantes.
- b) Código fuente en archivo(s) separado(s).
- c) Cualquier archivo adicional necesario para poder obtener sus resultados.

Herramientas Sugeridas

- **NetworkX** es un paquete de *Python* para la creación, manipulación y estudio de la estructura, dinámica y funciones de redes complejas.
- El kit de herramientas de **Basemap matplotlib** es una biblioteca para trazar datos 2D en mapas en *Python*.

Fecha de Entrega

Martes 3 de mayo, 2022.

PD. Cualquier duda o asunto no descrito en este documento se puede consultar por correo electrónico.