

Iniciativa en Ciencia de Decisiones

Solución a la Prueba - 2024-1

Jesús González Godoy, PhD | 23-02-2024

DESCRIPCIÓN

Nos interesa obtener y analizar información acerca del Consumo doméstico de energía de transporte por modo y tipo de combustible.

Particularmente, la proporción de tipo de combustible (**Gasoline, Diesel fuel, Liquified petroleum gas, Natural gas, Electricity, etc.**) por modo de transporte:

1. Light vehicles: Cars, Light trucks, Motorcycles
2. Buses: Transit, Intercity, School
3. Medium/heavy trucks
4. Air: General aviation, Domestic air carriers, international air carriers
5. Water: Freight, Recreational
6. Pipeline
7. Rail: Freight (Class I), Passenger, Passenger Transit, Passenger Commuter, Passenger Intercity

Necesitamos conocer esta información de 2008 a 2019 según la información contenida en el Transportation Energy Data Book (TEDB) de la edición 28 a la edición 40.

ENTREGABLES

1. **Repositorio:** Un repositorio en GitHub o GitLab para los programas de descarga, procesamiento y análisis de la información y salidas de los programas. con una estructura de directorios intuitiva y ordenada. Todos los programas deber ser replicables. El README debe ser autocontenido y tener la claridad suficiente para poder replicar los resultados. **(10 puntos)**
2. **Descarga de la información:** Un programa que descargue la información directamente de la página donde se alojan los recursos (Estos recursos se encuentran en archivos Excel o PDF). **(30 puntos)**
3. **Preprocesamiento de la información:** Un programa que obtenga la información de los recursos descargados y genere las razones de tipo de combustible por modo de transporte de 2008 a 2019. La salida de este programa son dos archivos csv, el primero contendrá la información en formato long y el segundo en formato wide. **(30 puntos)**
4. **Visualización y análisis de la información:** Un programa que genere las visualizaciones para analizar las salidas de la tarea anterior. Analizar los cambios registrados en el consumo de combustible por cada tipo de transporte. Particularmente analiza los cambios en las proporciones de consumo de gasolina y diésel. Este análisis, debe estar incluido en algún lugar del repositorio. **(30 puntos)**
5. **Extra:** Utilizar la información correspondiente en Motor Vehicle Registrations by State and Vehicle Type de la edición 40 de *TEDB* para estimar el consumo de combustible por estado, y realizar algunas visualizaciones. **(20 puntos)**

SOLUCIÓN

Repositorio

Con el fin de almacenar, administrar y compartir los recursos que conforman esta solución, se creó un repositorio en GitHub.

https://github.com/JesusGonzalezMX/prueba_2024_1_solucion.git

Entorno de desarrollo

Para tener una solución exitosa se usó un “Developer container” o “Dev Container”, una herramienta que asegura consistencia, reproducibilidad y facilita la colaboración. Los Dev containers son una forma de usar Docker para crear ambientes de desarrollo contenidos en un solo espacio de trabajo organizado y colaborativo, con el cual tenemos las siguientes ventajas:

- **Consistencia:** Estos contenedores aseguran que todos los desarrolladores trabajan con el mismo ambiente sin importar sus configuraciones locales, reduciendo problemas de compatibilidad.
- **Aislamiento:** Estos contenedores están encapsulados y aislados del sistema principal y de otros contenedores, por lo que los cambios hechos en el contenedor de un proyecto no afectan a los demás proyectos, es decir que se pueden tener diferentes bibliotecas, aplicaciones y herramientas en cada contenedor.
- **Reproducibilidad:** Se facilita reproducir el ambiente de trabajo en diferentes máquinas y tener distintas etapas de desarrollo.

con una estructura de directorios intuitiva y ordenada. Todos los programas deben ser replicables. El README debe ser autocontenido y tener la claridad suficiente para poder replicar los resultados.