**Descripción de las Variables del Dataset**

El dataset recopila información relacionada con el monitoreo y gestión de viajes de transporte terrestre. Las variables incluidas reflejan aspectos clave del desempeño logístico, la ubicación geográfica y los tiempos asociados a cada viaje. A continuación, se detallan las variables originales:

**Tipo Cátegorico:**

* **GpsProvider**: Proveedor del sistema GPS que proporciona los datos de ubicación.
* **BookingID**: Identificador único para cada viaje.
* **Market/Regular**: Tipo de viaje.
  + Regular: Viajes realizados con proveedores bajo contrato.
  + Market: Viajes realizados con proveedores sin contrato.
* **Vehicle\_no**: Número de identificación del vehículo.
* **Origin\_Location**: Lugar de inicio del viaje.
* **Destination\_Location**: Lugar de destino del viaje.
* **Org\_lat\_lon**: Coordenadas de latitud y longitud del lugar de inicio.
* **Des\_lat\_lon**: Coordenadas de latitud y longitud del lugar de destino.
* **Current\_Location**: Ubicación en tiempo real del vehículo.
* **DestinationLocation**: Repetición del lugar de destino.
* **Ontime**: Indica si el vehículo llegó a tiempo (calculado con base en el ETA planificado y real). G on time o Nan faltante
* **Delay**: Indica si el vehículo llegó con retraso (calculado con base en el ETA planificado y real). R retraso o Nan Faltante.
* **OriginLocation\_Code**: Código del lugar de origen.
* **DestinationLocation\_Code**: Código del lugar de destino.
* **VehicleType**: Tipo de vehículo utilizado.
* **Driver\_Name**: Nombre del conductor.
* **CustomerID**: Identificador único del cliente.
* **CustomerNameCode**: Nombre y código del cliente.
* **SupplierID**: Identificador del proveedor del vehículo.
* **SupplierNameCode**: Código del nombre del proveedor.
* **Material Shipped**: Materiales transportados.

**Tipo Fecha y Tiempos:**

* **BookingID\_Date**: Fecha en la que se creó la reserva del viaje.
* **Data\_Ping\_timeActual\_eta**: Tiempo real de llegada del vehículo: Hora en la que se recibe un ping del GPS.
* **Planned\_ETA**: Tiempo estimado de llegada planificado.
* **Actual\_ETA**: Tiempo real de llegada del vehículo
* **Trip\_start\_date**: Fecha y hora de inicio del viaje.
* **Trip\_end\_date**: Fecha y hora de finalización del viaje (según documentación, no se considera para calcular retrasos).

**Tipo Númericas:**

* **Curr\_lat**: Latitud actual que se actualiza con cada ping del GPS.
* **Curr\_lon**: Longitud actual que se actualiza con cada ping del GPS.
* **Transportation\_distance\_in\_km**: Distancia total del viaje en kilómetros.
* **Minimum\_kms\_to\_be\_covered\_in\_a\_day**: Mínimo de kilómetros que debe cubrir el conductor en un día.
* **Driver\_MobileNo**: Número de contacto del conductor.

**Limpieza del Dataset:**

Durante la etapa de limpieza de datos, se identificaron variables redundantes, irrelevantes o que no aportaban al objetivo del análisis. Por lo tanto, se decidió eliminar las siguientes variables:

* GpsProvider
* BookingID
* Vehicle\_no
* Origin\_Location
* Destination\_Location
* Current\_Location
* DestinationLocation
* OriginLocation\_Code
* Driver\_MobileNo
* DestinationLocation\_Code
* CustomerNameCode
* SupplierNameCode

Estas eliminaciones permitieron simplificar el dataset y enfocar el análisis en las variables más relevantes para el estudio de la inteligencia artificial aplicada al transporte.