



Descripción de tabla de datos

Vicente González D.

18/septiembre/2024

Agenda

1. Origen y fuentes de los datos
2. Formato de los datos
3. Definición de variables
4. Observaciones finales

1. Origen y fuentes de los datos

La información proporcionada para esta competencia corresponde a **datos operativos reales de trenes** en alguna línea de transporte de pasajeros en el mundo. Esta información es cargada de forma continua a servidores de Alstom por una computadora a bordo de cada tren (20 para este paquete de datos), acorde a estándares de la **UIC**, *Union internationale des chemins de fer* (por sus siglas en francés, *Unión Internacional de Ferrocarriles*)



2. Formato de los datos



Feather File Format

La tabla de datos proporcionada en este paquete es .feather; internamente responde al formato columnar **Arrow IPC**, proporcionado compatibilidad con múltiples librerías para el análisis de datos.

Cada columna de la tabla responde a los siguientes tipos:

Nombre	Tipo	Nombre	Tipo
1. timestamp	<i>timestamp[ns, tz=America/Panama][*]</i>	9. PCE2_ER	<i>int32[*]</i>
2. ACE1_EC	<i>int32[*]</i>	10. PCE3_EC	<i>int32[*]</i>
3. ACE2_EC	<i>int32[*]</i>	11. PCE3_ER	<i>int32[*]</i>
4. ATC_CS	<i>bool[*]</i>	12. PCE4_EC	<i>int32[*]</i>
5. ATC_DS	<i>bool[*]</i>	13. PCE4_ER	<i>int32[*]</i>
6. PCE1_EC	<i>bool[*]</i>	14. UMC_LBA	<i>int32[*]</i>
7. PCE1_ER	<i>int32[*]</i>	15. UMC_ODO	<i>bool[*]</i>
8. PCE2_EC	<i>int32[*]</i>	16. asset	<i>int32[*]</i>
<i>*pyarrow</i>	<i>int32[*]</i>		<i>string[*]</i>

3. Definición de variables

Independiente del tipo de dato en las columnas del archivo suministrado, estas responden a una **unidad de medida** en función de su objetivo, mostrado a continuación:

Variable	Descripción	Unidad
<i>timestamp</i>	Fecha-hora de captura del dato	<i>Timestamp(ns)</i>
<i>ACEN_{EC}</i>	Energía consumida por convertidor; 2 unidades por tren.	<i>KWh</i>
<i>ATC_CS</i>	Estación Actual (ATC**). ID numérico para nodo actual en línea ferroviaria.	---
<i>ATC_DS</i>	Estación Destino (ATC**) ID numérico para nodo destino en línea ferroviaria.	---
<i>PCEN_{EC}</i>	Energía consumida por unidad de tracción; 4 unidades por tren.	<i>KWh</i>
<i>PCEN_{ER}</i>	Energía regenerada por unidad de tracción; 4 unidades por tren.	<i>KWh</i>
<i>UMC_LBA</i>	Rastreo de Misión de Línea (ATC**). Modo automático o manual	Bool
<i>UMC_ODO</i>	Distancia registrada por odometría (UMC***).	<i>dam (decámetro)</i>
<i>asset</i>	ID tren	---

** ATC: Automatic Train Control

*** UMC: Universal Media Controller

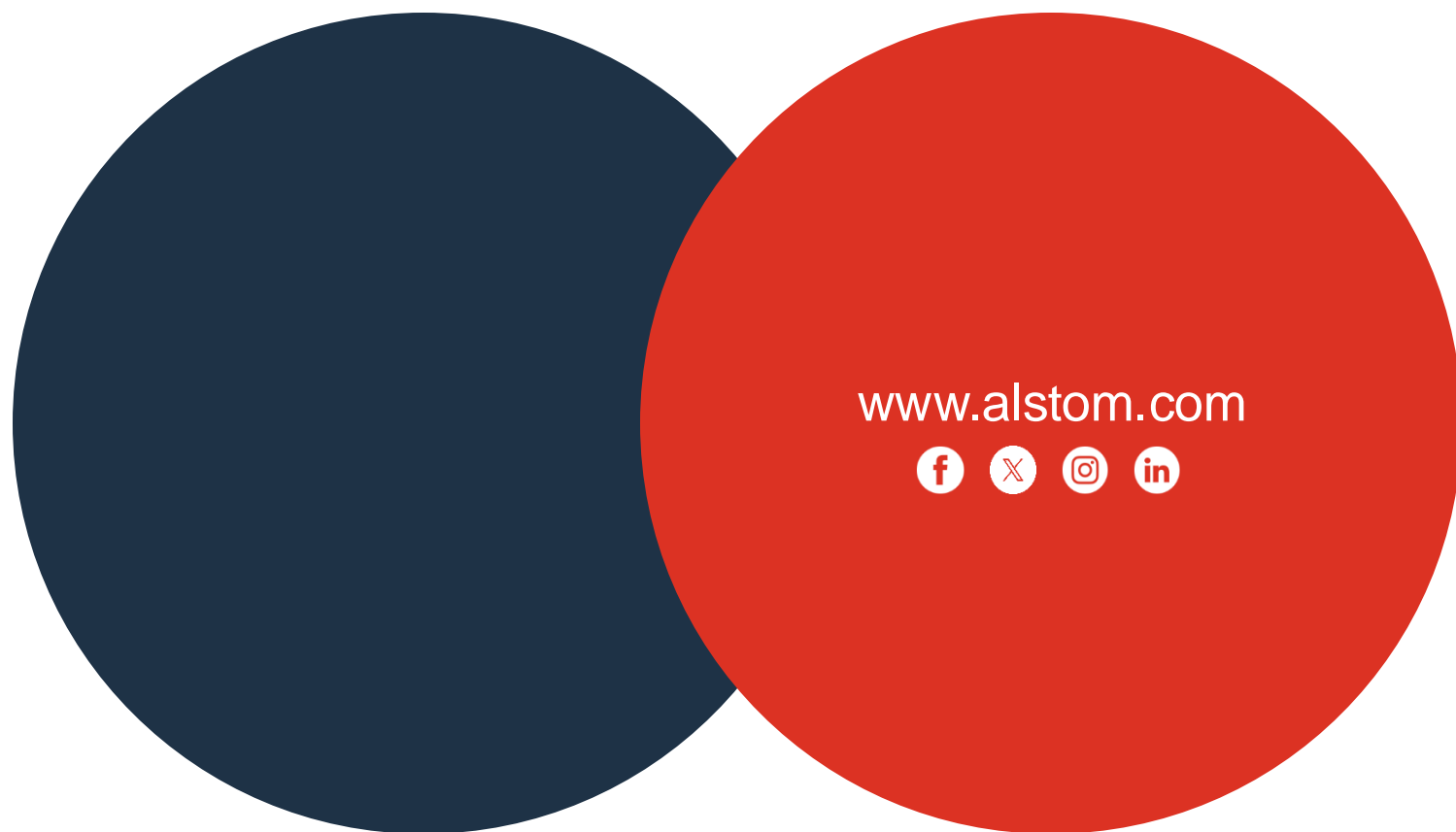
4. Observaciones finales

Como se mencionó anteriormente, dado que los datos proceden de fuentes reales, estos contendrán tendencias que pueden ser interrumpidas, o diferentes en algunos trenes, por situaciones como las siguientes:

- a) Ruido eléctrico en redes internas de comunicación.
- b) Diferente versionado de software en computadora a bordo.
- c) Problemas de conexión a la red.
- d) Sensores con intermitencia en su funcionamiento.
- e) Reinicio periódico de variables debido a desbordamiento, encendido/apagado del tren, reemplazo de partes, etc.

*****Deberá considerar este ruido al momento de procesar los datos*****

Los resultados serán recibidos hasta máximo las 10:00am del día viernes 20/septiembre/2024, luego de ese momento no serán considerados los correos recibidos.



ALSTOM
• mobility by nature •