

Prueba 1

- Para realizar esta prueba debes haber estudiado previamente todo el material disponibilizado correspondiente al módulo.
- Una vez terminada la prueba, comprime la carpeta que contiene el desarrollo de los requerimientos solicitados y sube el `.zip` en el LMS.
- Puntaje total: 10 puntos.
- Desarrollo prueba:
 - La prueba se debe desarrollar de manera Individual.
 - Realiza screenshots de la verificación de instalación de cada herramienta.
 - (El alumno debe tener un total de 7 puntos para aprobar)

Dataset

- Para esta prueba se utilizará el mismo dataset en **todos** los ejercicios.
- Cada registro del dataset corresponde a una transacción realizada con una tarjeta Bip!.
- Las columnas son como se describe a continuación:
 - `fechahoratrnx`: Fecha y hora a la en la que se realiza la transacción.
 - `codigoentidad`: Código del operador.
 - `nombrentidad`: Nombre del operador.
 - `codigositio`: Código del lugar en el que se realiza la transacción.
 - `nombresitio`: Nombre del lugar en el que se realiza la transacción.
 - `nrotarjeta`: Hash de la tarjeta Bip!
- La ubicación del dataset es `s3://bigdata-desafio/transantiago/`.
- Los datos se encuentran en formato columnar Parquet.

Muestra

```

+-----+-----+-----+-----+-----+
-----+-----+
|      fechahoratr|x|codigoentidad|
nombreentidad|codigo sitio|nombresitio|      nro tarjeta|
+-----+-----+-----+-----+-----+
-----+-----+
|2019-02-28 21:00:15|      13|      U6 - Redbus|      2596|
CJRC-35|612d6a01a6158dd63...|
|2019-02-28 21:02:39|      5|U5 - Metropolitana|      7660|
FLXG-38|eb7a60b4eb2e7b0a9...|
|2019-02-28 21:11:25|      5|U5 - Metropolitana|     11338|
FLXT-94|b4a501a26eaab4f8a...|
|2019-02-28 21:14:22|     13|      U6 - Redbus|      2596|
CJRC-35|a727f39a7a85cb6c1...|
|2019-02-28 21:17:02|      5|U5 - Metropolitana|     14379|
BJFJ-15|dcb9033eb5920016b...|
|2019-02-28 21:17:07|      5|U5 - Metropolitana|     14379|
BJFJ-15|8647ac944c4fa5928...|
|2019-02-28 21:20:12|     13|      U6 - Redbus|      2596|
CJRC-35|4339814f270bfab3d...|
|2019-02-28 21:21:34|      5|U5 - Metropolitana|      7661|
FLXG-39|e59769c9c65ff87d8...|
|2019-02-28 21:21:42|      5|U5 - Metropolitana|      7661|
FLXG-39|49de7c2ab5dee56a0...|
|2019-02-28 21:21:44|      5|U5 - Metropolitana|      7661|
FLXG-39|64cc53979994b65e1...|
|2019-02-28 21:31:21|      5|U5 - Metropolitana|      7661|
FLXG-39|09c9d7a0b9a78ebb5...|
|2019-02-28 21:37:10|      5|U5 - Metropolitana|     14379|
BJFJ-15|720d1760077201b0a...|
|2019-02-28 22:01:06|      5|U5 - Metropolitana|      7661|
FLXG-39|1b5f4d29ad5c0a75f...|
|2019-02-28 22:01:13|      5|U5 - Metropolitana|      7661|
FLXG-39|98b4b0a4ac617cfb2...|
|2019-02-28 22:01:17|      5|U5 - Metropolitana|      7661|
FLXG-39|1b49300a5ce2c3a45...|
|2019-02-28 22:01:20|      5|U5 - Metropolitana|      7661|
FLXG-39|d9c9e244023c49fe0...|
|2019-02-28 22:01:23|      5|U5 - Metropolitana|      7661|
FLXG-39|0bb31c472e93e5ef9...|

```

```
|2019-02-28 22:01:32|      5|U5 - Metropolitana|      7661|
FLXG-39|82f249120eaed6624...|
|2019-02-28 22:01:35|      5|U5 - Metropolitana|      7661|
FLXG-39|ab1843f66ae99efbe...|
|2019-02-28 22:03:17|      5|U5 - Metropolitana|      7956|
FLXK-32|4ca48c030ba3eec86...|
+-----+-----+-----+-----+
-----+
```

Ejercicio 1: Operaciones básicas (1 Punto)

- Determine:
 - Cantidad de transacciones.
 - Primera y última transacción en función de la columna fechahoratr.

Ejercicio 2: Operadores del servicio (1 punto)

- Liste los distintos operadores.
- Muestre la cantidad de transacciones para cada operador.

Ejercicio 3: Tarjetas Bip! (4 Puntos)

- Muestre las 20 tarjetas con más transacciones. (2 Puntos)
- Cuente todas las tarjetas que posean más de 2 transacciones por día. (2 Puntos)

Ejercicio 4: Fechas y rangos de horario (4 Puntos)

- Determine los 10 días con más transacciones. (1 Punto)
- Realice una comparación de la cantidad de transacciones por hora para cada operador. Muestre sus resultados en un gráfico donde el eje X sea la hora del día y el eje Y la cantidad de transacciones. (3 Puntos)