

## Parte 3 – Pandas

### 1. Persistencia de Datos

#### 1.1 Definición de los datos y justificación

El propósito principal de este sistema es disminuir la incertidumbre al buscar estacionamiento cerca del campus universitario, proporcionando información actualizada sobre la disponibilidad y las reservas. Para lograr esto, es crucial mantener datos reales sobre el uso que nos permitan analizar el comportamiento de los usuarios, identificar las horas de mayor demanda y evaluar cuán efectivas son las reservas. Estos registros no solo enriquecerán nuestros análisis con pandas, sino que también servirán como prueba del funcionamiento del sistema y como base para futuras mejoras.

Cada vez que un usuario interactúa, se crea información valiosa. Se mantendrán los siguientes datos:

Dato	Descripción	Utilidad analítica
<b>ID de evento</b>	Identificador único de cada acción realizada	Evita duplicados y permite reconstruir el historial de uso del sistema
<b>Fecha y hora</b>	Momento exacto en que ocurre la acción	Permite detectar horarios pico, patrones de ocupación y tiempos de respuesta
<b>Usuario</b>	Código anónimo asignado al usuario	Diferencia comportamientos sin comprometer privacidad
<b>Tipo de acción</b>	Puede ser consulta, reserva, cancelación o reinicio	Mide el uso de cada función y la frecuencia de reservas exitosas
<b>Parqueo</b>	Identificador del área o zona de estacionamiento	Permite saber qué zonas son más demandadas
<b>Espacio</b>	Número o código del espacio seleccionado	Facilita análisis de ocupación específica por espacio
<b>Resultado</b>	Indica si la acción se completó correctamente	Evalúa la eficiencia del sistema y detecta errores recurrentes
<b>Espacios libres después de la acción</b>	Número de espacios disponibles tras cada operación	Sirve para calcular la ocupación promedio y generar gráficas de disponibilidad

<b>Capacidad total del parqueo</b>	Total de espacios por zona	Permite calcular porcentajes de uso y detectar sobrecarga
------------------------------------	----------------------------	---

Además, se añadirán dos archivos auxiliares complementarios:

Usuarios.csv: que contendrá el user\_id y rol (estudiante, docente o visitante), que permite comparar hábitos por tipo de usuario.

Espacios.csv: incluye el lot\_id. Spot\_id y atributos del espacio como accesible o general, facilitando los análisis más detallados sobre tipos de parqueo.

Guardar estos datos es fundamental porque:

Ayuda a documentar el uso real del sistema, asegurando que los análisis se basen en evidencia y no en datos ficticios.

Permite identificar patrones temporales y de comportamiento, como las horas de mayor actividad, las áreas más utilizadas y la frecuencia de cancelaciones. Contribuye a la toma de decisiones sobre futuras mejoras: expansión de áreas, optimización del flujo de reservas o ajustes en las reglas de uso.

Promueve la transparencia y la trazabilidad, ya que cada acción queda registrada de manera ordenada y verificable.

Facilita la integración con herramientas analíticas como pandas.

Los datos se guardarán de manera automática en formato CSV, con encabezados fijos y un registro por cada evento.

**Archivo principal:** data/eventos.csv

**Campos:** event\_id, timestamp, user\_id, accion, lot\_id, spot\_id,

success, free\_spots\_after, capacity, booking\_id, source, app\_version, error\_code

Ejemplos: e0245f1c,2025-11-03T14:26:12Z,u\_05,reserva,Parqueo\_A,A-15,1,27,80

Esta estructura es fácil de leer y reutilizar, y funciona bien con bibliotecas analíticas. Permite filtrar, agrupar y graficar la información sin complicaciones. Además, evita la redundancia, ya que cada evento se guarda solo una vez. Y lo mejor de todo, se puede expandir en el futuro, añadiendo más columnas sin perder la compatibilidad.

En resumen, el sistema almacenará información sobre cada acción que los usuarios realicen en el proceso de consulta, reserva, cancelación o reinicio. La combinación de datos temporales, de usuario y del estado del parqueo crea una base sólida para analizar el comportamiento y la eficiencia del sistema, garantizando así persistencia, organización y un valor analítico a largo plazo.

## 1.2 Implementación: guardado automático en CSV

Organización de los archivos:

eventos.csv	(log principal, append-only)
espacios.csv	(catálogo de espacios)
usuarios.csv	(catálogo de roles)

Convenciones:

Cada vez que se produce una acción, como una consulta, reserva, cancelación o reinicio, se debe registrar automáticamente en el archivo eventos.csv.

Las cabeceras fijas son: event\_id, timestamp, user\_id, accion, lot\_id, spot\_id, booking\_id, success, free\_spots\_after, capacity, source, app\_version, error\_code. Este es un apéndice único, lo que significa que no se sobrescriben filas.

Se asegura de que los nombres sean consistentes, sin espacios, en minúsculas y con guiones bajos.

En esta fase el código fue totalmente documentado con comentarios explicativos en todas las funciones. Se añadieron docstrings y se aplicaron buenas prácticas de nombrado y modularización para asegurar la legibilidad y mantenimiento.

### Ejemplos de llamadas típicas:

```
log_event("u_01","consulta","Parqueo_A", success=True, free_spots_after=25,  
capacity=80)  
# tras reservar:  
log_event("u_01","reserva","Parqueo_A","A-15","b_123", True,  
free_spots_after=24, capacity=80)  
# si falla:  
log_event("u_01","reserva","Parqueo_A","A-15","", False, 24, 80,  
error_code="YA_OCUPADO")
```

Esta estructura asegura la persistencia automática de la información y la integridad del registro, sentando la base para los análisis posteriores con pandas.

## 1.3 Lectura sencilla y futura reutilización

- **Consistencia:** mantener las mismas cabeceras y evitar las columnas “fantasma”.
- **No repetición:** asegurar que cada event\_id sea único (esto ayuda a eliminar duplicados).
- **Tipos claros:** utilizar timestamps en formato ISO-8601 UTC; booleanos representados como 0/1; y capacity y free\_spots\_after en formato numérico.

- **Catálogos cerrados:** las acciones deben ser una de estas: {consulta, reserva, cancelación, reinicio}.
- **Privacidad:** usar un user\_id seudónimo; no incluyas información personal identificable (PII).

Gracias a estas validaciones, los datos podrán analizarse sin limpieza previa y reutilizarse fácilmente en futuras versiones del sistema.

### 3. Reflexión

**6 Sombreros para pensar**

**Datos y hechos**  
Hemos implementado la persistencia en CSV, realizado análisis con Pandas y creado visualizaciones con Streamlit. Todos los entregables se completaron de acuerdo a la rúbrica y utilizando datos generados directamente desde el sistema..

**Negativo**  
Fue un verdadero desafío depurar el código sin recurrir a estructuras penalizadas y al mismo tiempo mantener la organización entre los miembros del grupo.

**Organización**  
Todo el equipo se aseguró de cumplir con los criterios de la rúbrica, dividimos las tareas en subfases y estas fueron revisadas constantemente. Nuestra coordinación nos permitió corregir errores a tiempo.

**Creatividad**  
Sería útil añadir mapas interactivos de los parqueos o estadísticos en tiempo real. Además de conectar el sistema a una base de datos SQL.

**Optimismo**  
Alcanzamos un nivel de integración muy alto. Aprendimos a usar Pandas y Matplotlib de manera práctica y mejoró nuestro trabajo en equipo.

**Emociones**  
Nos sentimos satisfechos al ver el sistema funcionando con datos reales. Al inicio hubo frustración por los errores al generar los CSV y configurar los filtros, pero la motivación creció al ver los resultados visuales en los gráficos.