



Sistemas Distribuidos I (75.74)

Coffee Shop Analysis

TP Escalabilidad: Middleware y Coordinación de Procesos

Docentes

- | | | |
|---------------------|--------------------|--------------------|
| ■ Pablo D. Roca | ■ Nicolás Zulaica | ■ Camila Sebellin |
| ■ Gabriel Robles | ■ Franco Papa | ■ Máximo Grismondi |
| ■ Franco Barreneche | ■ Manuel Reberendo | |



Requerimientos Funcionales

- Se solicita un sistema distribuido que analice la información de ventas en una cadena de negocios de Cafés en Malasia.
- Se cuenta con información transaccional por ventas (montos, items vendidos, etc), información de los clientes, de las tiendas y de los productos ofrecidos.
- Se debe obtener:
 1. Transacciones (Id y monto) realizadas durante 2024 y 2025 entre las 06:00 AM y las 11:00 PM con monto total mayor o igual a 75.
 2. Productos más vendidos (nombre y cant) y productos que más ganancias han generado (nombre y monto), para cada mes en 2024 y 2025.
 3. TPV (Total Payment Value) por cada semestre en 2024 y 2025, para cada sucursal, para transacciones realizadas entre las 06:00 AM y las 11:00 PM.
 4. Fecha de cumpleaños de los 3 clientes que han hecho más compras durante 2024 y 2025, para cada sucursal.



Requerimientos No Funcionales

- El sistema debe estar optimizado para entornos multicomputadoras
- Se debe soportar el incremento de los elementos de cómputo para escalar los volúmenes de información a procesar
- Se requiere del desarrollo de un Middleware para abstraer la comunicación entre grupos, basándose en las interfaces provistas.
- Se debe soportar una única ejecución del procesamiento y proveer *graceful quit* frente a señales SIGTERM.



Datasets, notebook patrón y librerías

- Para construir una simulación realista, se trabajará sobre el siguiente dataset:
 - <https://www.kaggle.com/datasets/geraldooizx/g-coffee-shop-transaction-202307-to-202506/data>
 - Serán necesario los files asociados a transacciones, a items por transacción, usuarios, tiendas y productos.
- Se usarán los valores del siguiente notebook como resultados patrón:
 - Full: <https://www.kaggle.com/code/gabrielroble/fiuba-distribuidos-1-coffee-shop-analysis?scriptVersionId=260252483>
 - Reduced: <https://www.kaggle.com/code/gabrielroble/fiuba-distribuidos-1-coffee-shop-analysis?scriptVersionId=264040974>
-



Interfaces de Middleware

- Para la implementación del Middleware, se partirá de alguna de las interfaces provistas por la cátedra:
 - <https://github.com/7574-sistemas-distribuidos/middleware-interfaces>
- Se espera que se realicen pruebas unitarias demostrando su funcionamiento cubriendo al menos los siguientes casos:
 - Comunicación por Working Queue 1 a 1
 - Comunicación por Working Queue 1 a N
 - Comunicación por Exchange 1 a 1
 - Comunicación por Exchange 1 a N



Se espera del alumno:

- Empleo del tiempo de consultas en clase para resolver dudas y clarificar el negocio del sistema a construir previo a su diseño
- Exposición y verificación en clase de la arquitectura propuesta antes de iniciar su implementación
- Empleo del grupo de correos para realizar consultas que no pudieran ser resueltas en clase
- Consideración de prácticas distribuidas según lo estudiado en clase para elaborar una arquitectura flexible, escalable y robusta
- Aprobación del cuerpo docente para el uso de cualquier librería.
- Demo del sistema en funcionamiento previamente ensayada



- Fecha de entrega: 02/10/2025
- Formato de entrega:
 - Demo del sistema funcionando con al menos una porción del dataset.
 - Documento de arquitectura **actualizado** 4+1 Views o C4Model incluyendo al menos: diagramas de robustez, despliegue, actividades, paquetes, secuencia y DAG.
 - Pruebas unitarias sobre el Middleware. **No se evaluarán otras secciones de código.**