

Documentación del Proyecto

Este proyecto tiene como finalidad desarrollar un sistema automatizado en lenguaje R que permita identificar productos con riesgo de quiebre de stock en el Supermercado del Neumático (SDN). Este tipo de análisis es fundamental para las áreas de compras, logística y operaciones, ya que una ruptura de inventario afecta directamente la disponibilidad de productos, la satisfacción del cliente y la eficiencia operativa de la empresa.

La solución propuesta actúa como una herramienta preventiva, facilitando la toma de decisiones mediante un análisis riguroso, basado en datos históricos de ventas y stock actual. El sistema permite anticipar posibles faltantes de productos y priorizar acciones correctivas de manera objetiva y sistematizada.

Metodología General

La metodología se sustenta en el procesamiento automatizado de dos archivos CSV como fuentes de entrada:

1. Un archivo con el historial de ventas, desagregado por tienda y producto (SKU).
2. Un archivo con el stock actual disponible, también por producto.

Ambos archivos son cargados desde una carpeta local y tratados de manera secuencial a través de un script estructurado en R, que abarca etapas de limpieza, transformación, cálculo, análisis y visualización.

Descripción del Flujo de Trabajo

El flujo general del sistema, representado en el diagrama adjunto, comprende las siguientes etapas:

1. Carga de archivos
Se importan los archivos CSV de ventas y stock. En esta etapa se asegura que ambos tengan estructuras de datos compatibles para el análisis posterior.
2. Transformación de datos
Se estandarizan los nombres de las columnas y se transforma la columna de fechas al formato Date, lo que permite realizar filtros temporales consistentes y operaciones basadas en tiempo.
3. Filtrado de ventas recientes
Se seleccionan los registros correspondientes a los últimos 7 días disponibles, lo que permite capturar el comportamiento más reciente de la demanda.
4. Cálculo de ventas promedio
Se calcula la venta promedio diaria por cada combinación de tienda y SKU, generando una métrica representativa del ritmo de salida de productos.
5. Cálculo de stock crítico
A partir de la venta promedio, se estima el stock crítico, definido como la cantidad mínima necesaria para cubrir un período estándar de tres días de reposición.

6. Cruce con el stock actual

Se integran los datos anteriores con el stock disponible, lo que permite calcular los días de cobertura actuales de cada producto, considerando su demanda reciente.

7. Clasificación del estado del producto

Con base en los días de cobertura y el stock crítico calculado, cada producto se clasifica en uno de los siguientes estados:

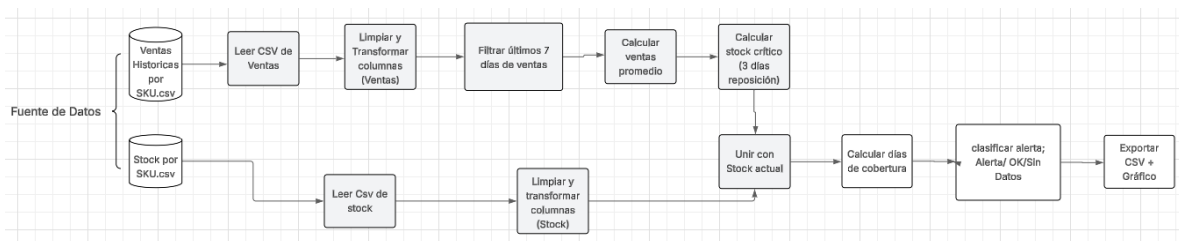
- Alerta: si el stock disponible es igual o inferior al stock crítico, o si los días de cobertura son menores o iguales a 3.
- OK: si el stock es suficiente para cubrir al menos tres días de reposición.
- Sin datos: si no se cuenta con información suficiente para estimar la demanda.

8. Exportación de resultados

Finalmente, el sistema genera dos productos de salida:

- Un archivo CSV que resume por producto las métricas clave (venta promedio, stock crítico, días de cobertura y clasificación de alerta), ordenado por nivel de criticidad para facilitar su revisión operativa.
- Un gráfico en formato PNG, que visualiza los días de cobertura de cada producto utilizando una paleta de colores diferenciadora para los distintos estados (Alerta, OK, Sin datos), lo que permite una rápida interpretación visual.

Ambos archivos se guardan automáticamente con la fecha de ejecución en el nombre, permitiendo generar un historial de ejecuciones diarias. Esta característica facilita su integración con reportes periódicos y sistemas de monitoreo internos.



Despliegue de la Solución

La solución ha sido diseñada para su ejecución en un entorno local utilizando RStudio. Esta decisión responde tanto a la simplicidad de implementación como a la independencia respecto a plataformas corporativas más complejas, permitiendo su rápida adopción incluso por usuarios con conocimientos técnicos intermedios.

Para su funcionamiento, el sistema requiere únicamente los archivos CSV actualizados en una carpeta determinada. Una vez ubicado el script en la misma ruta, este puede ejecutarse manualmente desde RStudio o mediante un proceso automatizado utilizando herramientas nativas del sistema operativo, como el Programador de tareas de Windows o cron en Linux.

El despliegue actual asume un entorno desconectado de sistemas de gestión empresarial (ERP), por lo que la actualización de datos debe realizarse manualmente. No obstante, el diseño modular del código permite que en versiones futuras se integre directamente a sistemas como SAP, mediante conexiones API o consultas a bases de datos relacionales.

En términos de portabilidad, se contempla como evolución natural la utilización del paquete `renv` para el control de versiones y entornos de desarrollo, así como la posibilidad de contenerizar la aplicación mediante Docker, asegurando la replicabilidad del entorno y facilitando el despliegue en servidores u otras estaciones de trabajo.

La salida generada por el script puede ser enviada por correo electrónico, almacenada en carpetas compartidas o integrada en tableros de visualización. En ese sentido, el sistema es lo suficientemente flexible como para adaptarse a los canales y rutinas actuales de supervisión de la empresa.

6. Monitoreo del Modelo

El monitoreo de la solución se abordará desde una doble perspectiva: por un lado, aseguraremos que el script en R se ejecute correctamente y que los resultados generados (archivos CSV y gráficos PNG) sean coherentes; por otro, evaluaremos de forma continua el comportamiento de los indicadores clave, tales como la cantidad de productos en estado crítico, su evolución en el tiempo y la cobertura efectiva del stock.

En una primera etapa, implementaremos un mecanismo de validación básico que permitirá verificar si los archivos de salida se han generado correctamente. Para ello, utilizaremos mensajes en consola y estructuras de control de errores (`tryCatch`) que nos alertarán en caso de fallos durante la ejecución del script.

A nivel operativo, realizaremos un seguimiento de la cantidad de productos clasificados en los estados “Alerta” y “Sin datos”, ya que estos representan los casos más críticos dentro del inventario. El monitoreo de esta métrica nos permitirá evaluar la estabilidad del sistema de abastecimiento y servirá como señal temprana ante posibles deficiencias en los procesos de compras o distribución.

Asimismo, estableceremos una política de respaldo y almacenamiento automático de los archivos generados, manteniendo un historial diario con nombre de archivo fechado. Esta práctica no solo permitirá comparar resultados a lo largo del tiempo, sino que también habilitará futuras extensiones del proyecto, como el entrenamiento de modelos predictivos o la simulación de escenarios de inventario.

En fases posteriores, automatizaremos la notificación de alertas relevantes mediante canales como correos electrónicos, lo que permitirá una respuesta más ágil por parte del equipo operativo. Además, evaluaremos la implementación de una interfaz interactiva en Shiny, que permitirá a los usuarios explorar, visualizar y filtrar los resultados directamente desde un panel web, sin necesidad de interactuar con el código. Esta funcionalidad potenciará el uso operativo del sistema y mejorará la toma de decisiones basada en datos.

Conclusión

El sistema de alertas de stock crítico desarrollado cumple con los objetivos planteados en el inicio del proyecto: entregar una herramienta práctica, automatizada y accionable para apoyar la gestión de inventario en Supermercado del Neumático. Se trata de una solución robusta, que funciona sobre datos reales, y que puede ser implementada y mantenida sin grandes inversiones tecnológicas.

Además, el modelo es escalable y puede ser adaptado a distintos entornos y requerimientos, incluyendo integración con bases de datos, sistemas ERP, o dashboards de visualización avanzada. Gracias a su diseño modular, puede evolucionar progresivamente según las necesidades de la organización, manteniéndose siempre alineado con los principios de la ciencia de datos aplicada a la toma de decisiones empresariales.

