

# Caso Ciencia de datos

1 de noviembre de 2024

## Ejercicio 1: Eco Bicis

- Descarga dos meses de los datos históricos de uso de la ecobici. Escoge los dos meses que quieras.
- Utilizando la herramienta que quieras (R, Pandas en python, u otra que tu conozcas) has un análisis de estos datos:
  - Identifica qué estaciones de ecobici en qué horarios están más saturadas.
  - Identifica los trayectos más recorridos.
  - Para las estaciones más usadas has una gráfica que muestre los tráficoes en cada horario.
  - Reporta otros hallazgos que te parezcan relevantes.
  - Compara los resultados que hayas encontrado entre las tablas de los dos meses que descargaste y escribe tus hallazgos
- Has una pequeña presentación con tus resultados. No es necesario que sea larga, sintetiza los resultados.
- Envía la presentación y el escript en el que trabajaste.

## Ejercicio 2: Bebés

Para cada AGEB de la delegación Álvaro Obregón estima cuántos bebés de 0 a 6 meses de edad habitan ahí el día de hoy. Explica tu razonamiento en menos de 300 palabras. Enlista tus fuentes y presenta los resultados.

## Ejercicio 3: La michoacana

Hace algunos años, la nevería La Michoacana decidió hacer un cambio radical en su forma de operar. En lugar de neverías tradicionales, La Michoacana decidió instalar 4,000 máquinas expendedoras de paletas dentro de la Ciudad de México.

Además de eso, decidió implementar un sistema de suscripción para las máquinas expendedoras. Es decir, en lugar de pagar por cada paleta, los clientes pagan una suscripción mensual fija por tener acceso a cuántas paletas quieran, cuando quieran. Las paletas se hacen en una planta que tiene un congelador central, desde el cual las paletas se distribuyen a las máquinas expendedoras en camiones refrigerados de una capacidad muy grande.

Las máquinas expendedoras tienen la peculiaridad de que el gasto energético depende de la cantidad de paletas que estén en la máquina, pues cada una se congela en un pequeño cajón. El costo de mantener

una paleta por un día es de \$1. El transporte de las paletas desde el congelador central a cada máquina expendedora tiene un costo fijo de \$100 por viaje a una máquina y cada máquina no puede ser surtida más de una vez al día.

Cada mañana, Luis, el jefe de operaciones, debe decidir qué máquinas expendedoras serán surtidas esa misma mañana y cuántas paletas deben ir a cada máquina expendedora que será surtida. Luis quiere mantener los costos de transporte y de energía bajos pero al mismo tiempo no quiere que existan muchos casos de máquinas sin paletas disponibles, pues los suscriptores podrían molestarse y abandonar el programa. Luis tiene el objetivo de mantener los costos de transporte y energía más bajos posibles y la cantidad de días/máquina donde hubo indisponibilidad de paletas menor a 2% al mes. Luis tiene total libertad todas las mañanas para elegir qué máquinas deben ser surtidas y cuántas paletas surtir en cada una. Para tomar esa decisión, Luis tiene, cada mañana, la siguiente información:

- La cantidad de paletas retiradas para cada máquina, cada día desde que empezó la operación (hace 5 años) hasta el día anterior.
- La cantidad de paletas disponibles en cada máquina, cada día (a final del día, medianoche) desde que empezó la operación (hace 5 años) hasta el día anterior.
- La capacidad, en cantidad de paletas, de cada máquina expendedora.
- El costo de surtir cada máquina (\$100).
- El costo de mantener una paleta por un día en cualquier máquina expendedora (\$1)
- No hay límite de carga en los camiones repartidores

Luis toma la decisión día con día de qué máquinas surtir y cuántas paletas surtir a cada una de ellas de forma bastante subjetiva, intuyendo que máquinas se quedarán sin paletas si no las surte y tratando de identificar una cantidad óptima de paletas para surtir cada una de ellas.

Luis se ha dado cuenta de que se trata de un problema bastante complejo, donde existe una gran oportunidad de tomar mejores decisiones. Se ha dado cuenta de que, para cada máquina, la estrategia de llenar a tope la máquina expendedora y surtirla el día que observe que las paletas están apunto de acabarse para llenar de nuevo la máquina a tope tiene un costo muy alto de energía y un costo bajo de transporte. Se ha dado cuenta también de que, para cada máquina, hacer viajes diarios con el contenido de paletas necesarias para el siguiente día tiene un costo de energía bastante bajo pero un costo de transporte muy alto.

Luis observa también que el consumo de paletas cambia mucho de máquina a máquina y día con día. Luis no tiene restricciones como mandar siempre la misma cantidad de paletas a una misma máquina o mandar paletas con frecuencia fija (e.g. cada 7 días). Dada esa libertad, entiende que se trata de un problema complejo pero que es la clave para lograr sus objetivos de mantener los costos de transporte y energía bajos sin superar su cota de indisponibilidad.

- Con la información anterior, diseña una solución al problema.

No existe una única solución válida al problema, busquemos entender cómo tú entiendes el problema y saber cómo lo resolverías. La respuesta puede ser textual, utilizando notación matemática, utilizando diagramas, pseudocódigo, un video explicando en un pizarrón, notas en una servilleta, algún otro medio o una mezcla de los anteriores, lo que tú creas que va a comunicar mejor tu planteamiento.

Puedes guiar tu respuesta por las siguientes preguntas:

- ¿De qué tipo de problema se trata? ¿Tiene elementos comunes como regresión, clasificación, pronósticos de series de tiempo, clustering, optimización, etc?
- ¿Cómo te imaginas una solución funcional al problema? ¿De qué partes está conformada?
- ¿Cómo interactúan esas partes? ¿Qué supuestos y riesgos ves en tu planteamiento?
- ¿Qué métodos o algoritmos utilizarías durante el desarrollo de esa solución?
- ¿Qué métricas evaluarías durante el desarrollo de la solución?
- ¿Cómo te imaginas el despliegue y la operación en producción de la solución?
- ¿Cómo evaluarías si la solución tuvo un impacto positivo y fue exitosa?

Has una pequeña presentación con tus resultados. No es necesario que sea larga, sintetiza los resultados.

Envía la presentación y el escript en el que trabajaste.

Escríbeme cualquier duda que tengas.