

ITBA

# Trabajo Práctico 1

**Caso de Estudio PedidosYa**  
**82.10 – Analítica en Transporte y Logística**

**Dana Nabel – 62197**  
**Paula Ariana Gonzalez – 60784**





# PedidosYa problema

# El Problema

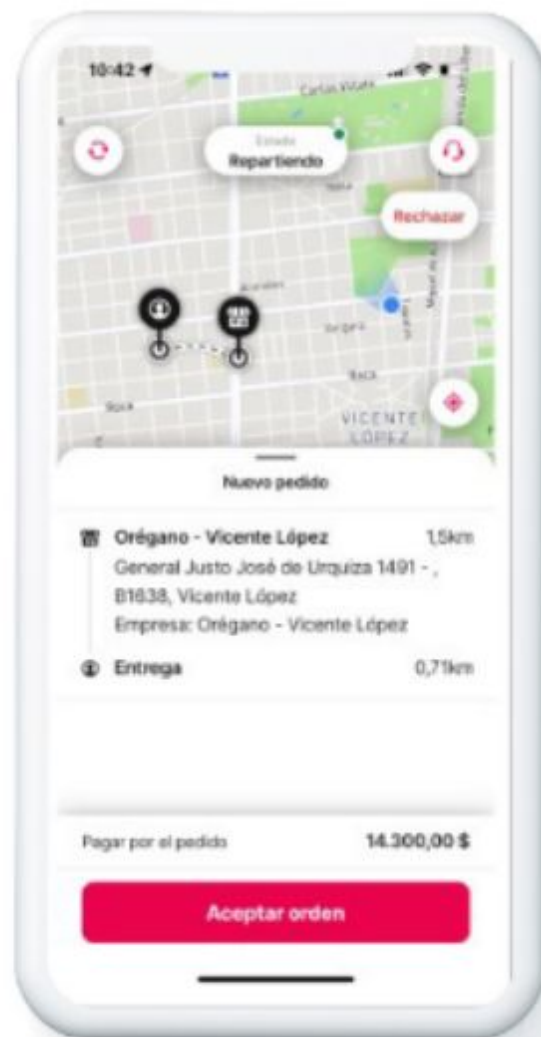
Mediante un algoritmo, se decide en tiempo real la mejor solución repartidor/orden y el repartidor es notificado. El repartidor tiene la capacidad de aceptar o rechazar esa orden, lo cual deriva en una tasa de aceptación de pedidos en el Mercado A de ~70%. A su vez, incluso una vez aceptado el pedido, puede decidir no finalizarlo, lo cual genera una tasa de “un-dispatches” de ~4%. Esto genera dos efectos:

- A) La orden no aceptada (una o más veces) tiene un impacto en Nivel de Servicio.
- B) Se reduce la disponibilidad de repartidores

Supuestos:

- Asumiendo no más de una orden simultánea por repartidor
- Asumir que es viable capitalizar el incremento de ganancias como consecuencia de mejoras de UTR en ahorros de CPO.

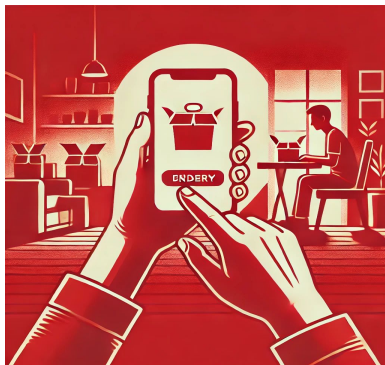
**Considerando los repartidores notificados el día 13/07/24 entre las 20:30 y las 20:45, ¿cuál sería tu propuesta para maximizar la aceptación y ejecución de las órdenes? Estimar el impacto en Nivel de Servicio y Costos asociado a la propuesta.**





# Primeras hipótesis

## Situacion Ideal:



## ¿Que podría estar afectando la tasa de rechazo y la de reasignación ?





# **PedidosYa** problema explicado a partir de la teoría de Logística

## Cadena de Suministro

- Una cadena de suministro es el conjunto de procesos y actividades involucradas en la producción y entrega de un producto (desde el proveedor inicial hasta el cliente final).
- La cadena de suministro tiene un alcance amplio (aprovisionamiento, el almacenaje, la fabricación, la distribución y la gestión de entregas).



- La **última milla** es el eslabón más crítico en la cadena de suministro de PedidosYa. Representa el trayecto final desde el restaurante hasta el cliente.
- La **inmediatez** es parte fundamental del valor que ofrece PedidosYa, donde cualquier retraso impacta directamente en la satisfacción del cliente.

## Lead Time



El lead time en logística es el tiempo total necesario para completar el ciclo de un pedido, desde que se realiza hasta que se entrega. En PedidosYa, un alto lead time puede deberse a varios factores:

- Demoras en la preparación del pedido en el restaurante.
- Recolección tardía del pedido por parte del repartidor
- Retrasos en la entrega debido a condiciones externas (tráfico, clima, distancias largas).
- Mucha cantidad de rechazos del pedido por parte de repartidores

## Sincronización y Coordinación en la Cadena de Suministro

La teoría de logística sugiere que una cadena de suministro eficiente requiere una sincronización óptima entre todos los actores involucrados. En el caso de PedidosYa, esta sincronización incluye:

- Restaurantes: Los restaurantes deben estar preparados para gestionar los pedidos de manera eficiente, reduciendo los tiempos de espera y mejorando el flujo de trabajo.
- Repartidores: La correcta asignación de repartidores según la proximidad y la disponibilidad ayuda a reducir los tiempos de recolección y entrega.
- Plataforma: La plataforma debe actuar como el coordinador central, asegurando que la información fluya de manera eficaz y en tiempo real





## Impacto en el Nivel de Servicio y Costos Operativos

La capacidad de PedidosYa para mantener un nivel de servicio alto, es decir, entregas rápidas y fiables, depende de su capacidad para equilibrar los costos operativos.

- Un servicio más rápido y eficiente suele implicar un mayor costo operativo. Si la cadena de suministro no está optimizada, los costos de cada entrega pueden dispararse, afectando la rentabilidad de la plataforma.
- La mejora en el nivel de servicio no siempre implica mayor rentabilidad, ya que la logística requiere un balance entre la calidad del servicio y el costo

## Soluciones desde la Teoría Logística

- Forecasting: Implementar predicciones de demanda para optimizar la asignación de repartidores y reducir esperas.
- **Lead Time: Mejorar la sincronización entre la preparación de pedidos y la recolección con un algoritmo de dispatch eficiente.**
- Optimización de Rutas: Utilizar un sistema de gestión de transporte (TMS) que optimice rutas, minimizando distancias y tiempos de entrega en zonas congestionadas o de largas distancias.



# Datos pedidos

# Variables que utilizamos:

- Los datos de los **pedidos** efectuados del mes de Julio
  - Datos de los **repartidores** (con su tasa de aceptación, rechazo, tipo de vehículo, pedidos realizados, antigüedad y tiempo en la aplicación)
- Datos de los **comercios** asociados (con su rating y su tiempo de espera promedio)
- **Clima** en el rango de tiempo estipulado del análisis (fecha/hora, temperatura, precipitación, condición y el viento)





# Limpieza de datos



# Pasos en la limpieza de datos:

- 1) → Ordenamos los datos por fecha y hora de transacción.
- 2) → Se eliminan los registros con NAs en las variables de `"transicion_coordenada_latitud"`, `"transicion_coordenada_longitud"` y `"transicion_id_repartidor"`.
- 3) → Con la latitud y longitud de los repartidores a los cuales se les envía una notificación del pedido, se calcula su distancia al comercio y la distancia total del pedido.
- 4) → Se borran los registros que tienen `"tiempo_estimado_entrega"` = NaT
- 5) → Se calcula el tiempo de espera entre que se le notifica a un repartidor y se entrega en el comercio.





# Pasos en la limpieza de datos:

- 6) → De la tabla de REPARTIDORES, se eliminan los `id_repartidor` duplicados, unificandolos calculando el promedio de los valores de los campos, sumando el total de `pedidos_entregados` y eligiendo el `vehículo` más utilizado.
- 7) → Unimos las bases, y eliminamos los pedidos de aquellos `id_repartidor` que no se encuentran en la base de REPARTIDORES.
- 8) → De la tabla de COMERCIOS, eliminamos los 2 que tenían `espera_en_vendor = NA`
- 9) → Unimos las bases, y eliminamos los pedidos de aquellos `id_comercio` que no se encuentran en la base de COMERCIOS.
- 10) → Unimos la base de CLIMA, a través de un campo calculado `'creacion_transicion_redondeada'` y `'date'`.
- 11) → Eliminamos los registros de los días 18/7/2024 y 22/7/2024 ya que no habían datos de clima en estos días.





# **Análisis Exploratorio de los Datos (EDA)**



# PedidosYa: Principales Métricas – Julio 2024

- Se obtienen 2.486.549 registros de datos de pedidos.
- Se identifican 290.692 pedidos de los cuales **288.926 fueron entregados** y 1233 fueron cancelados por los clientes.
- Entregados por **5521 repartidores**.
- Se realizaron pedidos de **2318 comercios** distintos.

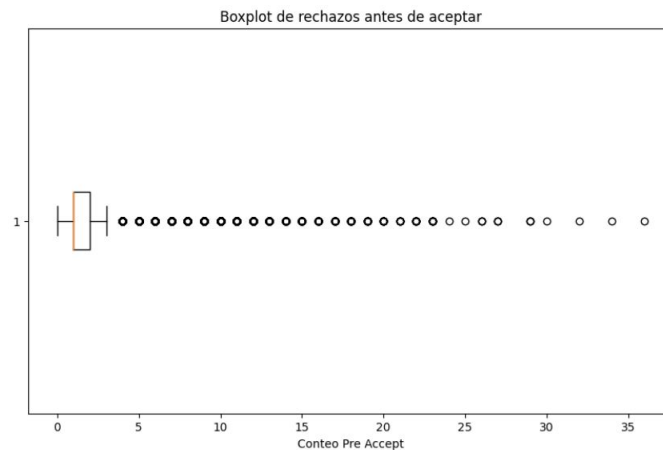
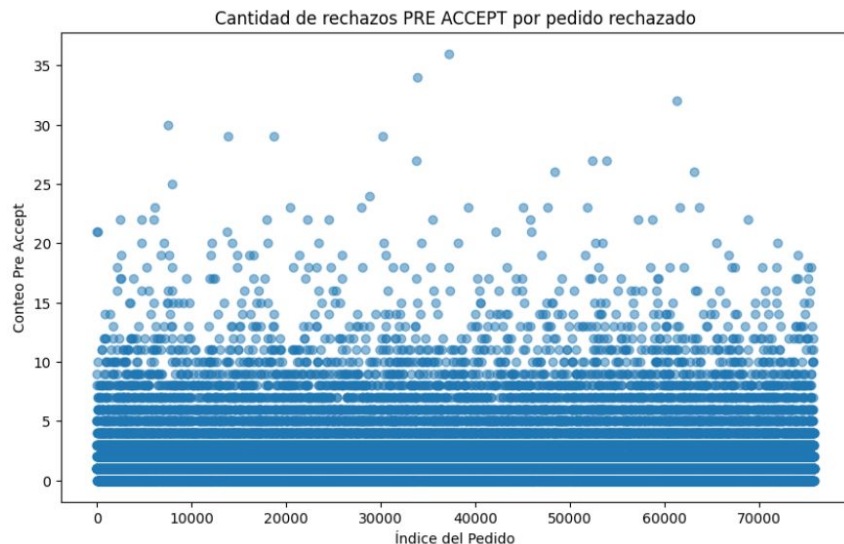
**Tasa de rechazo** =  $\text{sum}(\text{pre-accept undispatch}) / \text{sum}(\text{notificaciones}) = \mathbf{31,23\%}$

**Tasa de reasignación** =  $\text{sum}(\text{post-accept undispatch}) / \text{sum}(\text{accept}) = \mathbf{4,53\%}$





# Primer pantallazo de los pedidos rechazados



Cantidad de pedidos : 81107

Media: 1.88

Desvío Estándar: 1.95

Mínimo: 0

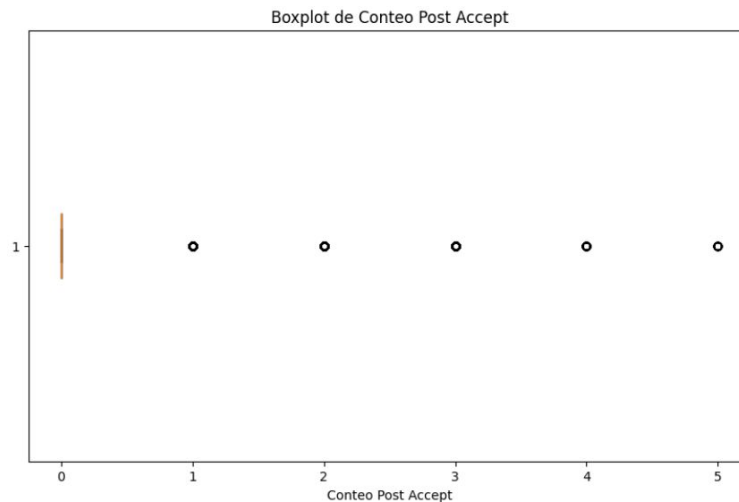
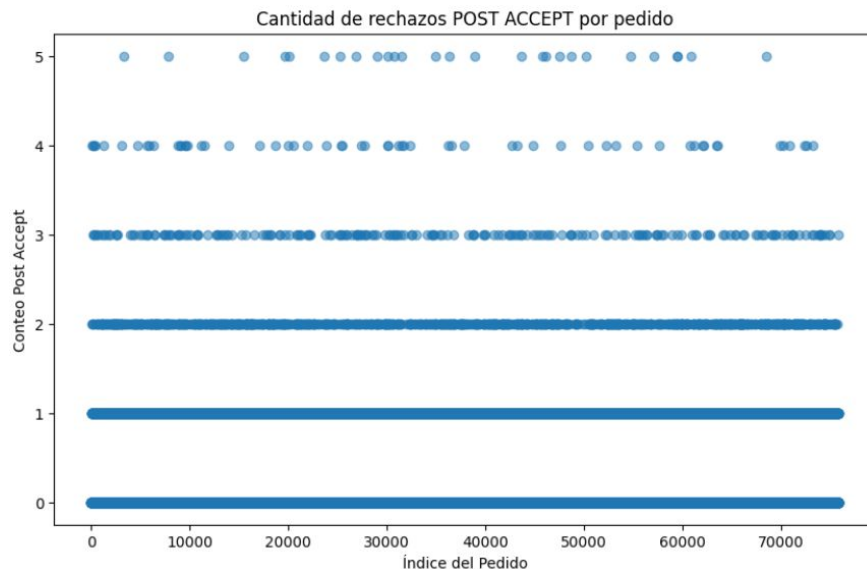
Percentil 25: 1

Percentil 50: 1

Percentil 75: 2

Máximo: 36

# Primer pantallazo de los pedidos rechazados



Cantidad de pedidos : 81107

Media: 0.18

Desvío Estándar: 0.45

Mínimo: 0

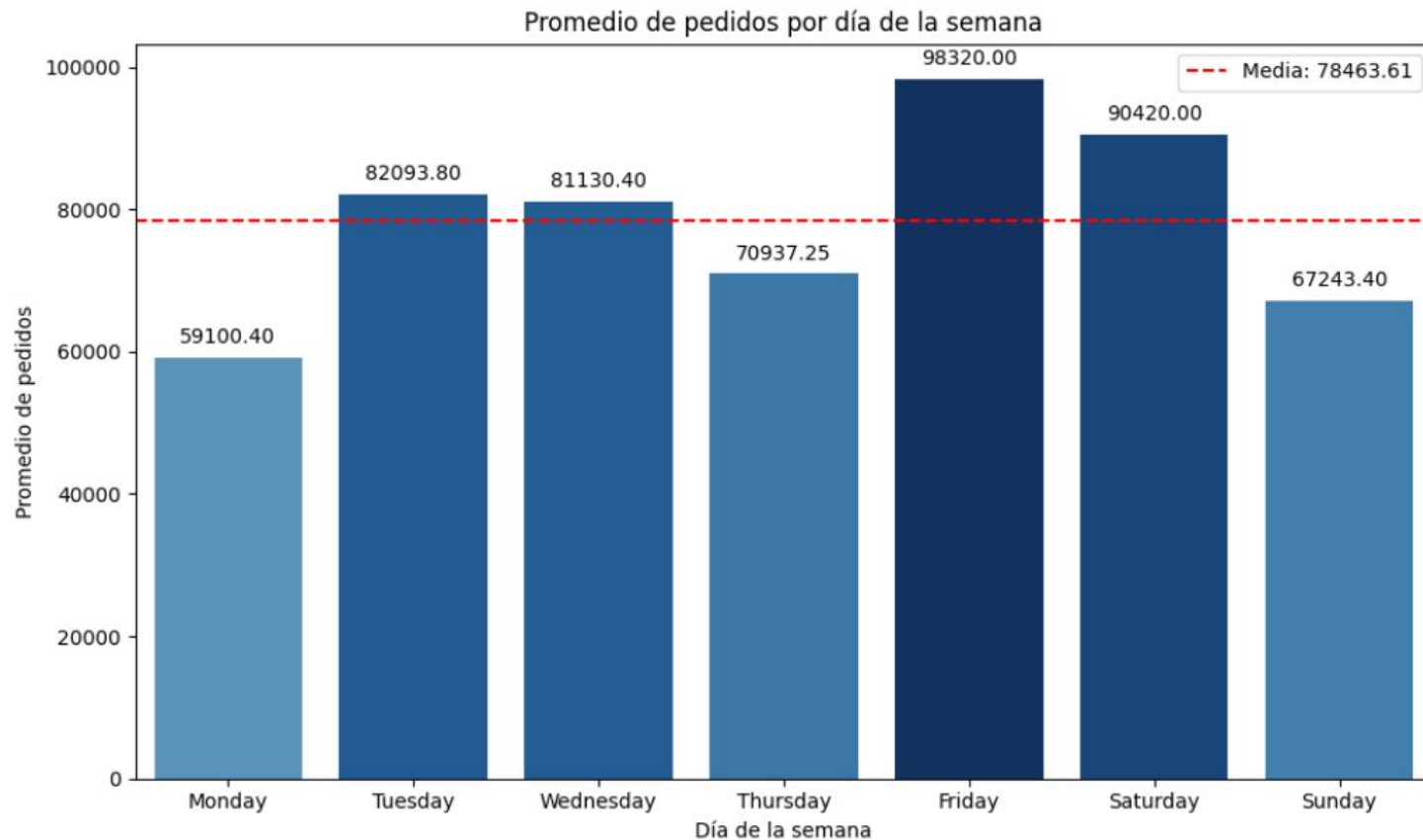
Percentil 25: 0

Percentil 50: 0

Percentil 75: 0

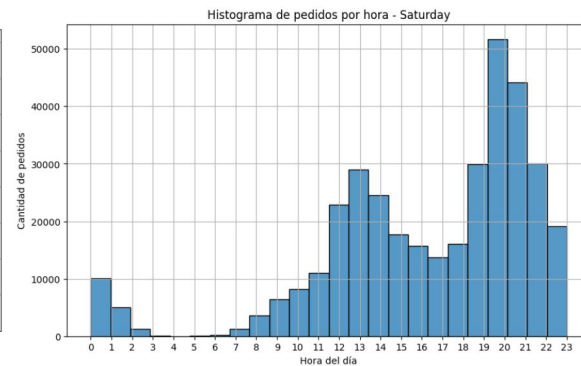
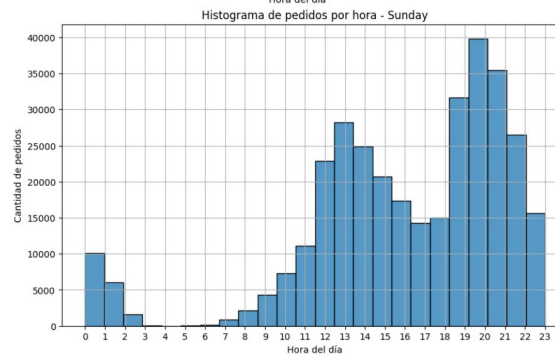
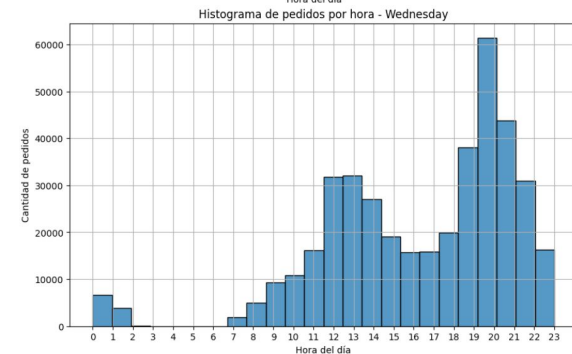
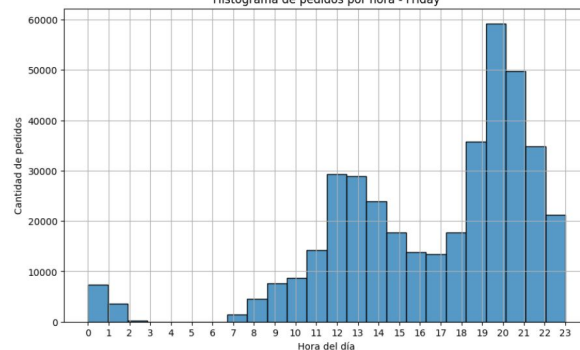
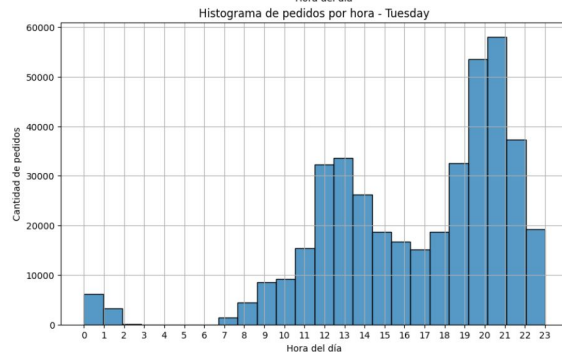
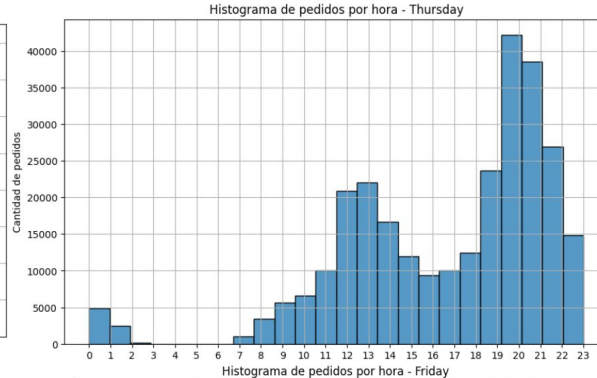
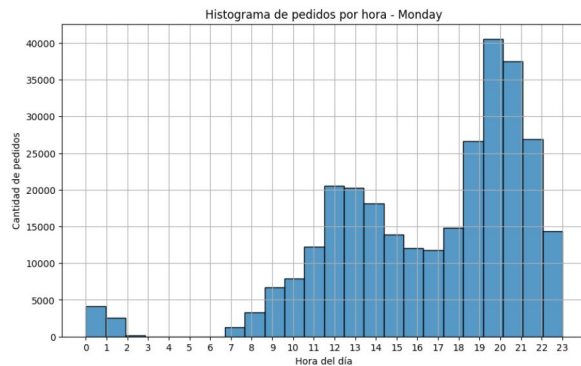
Máximo: 5

# Análisis por Día de Semana

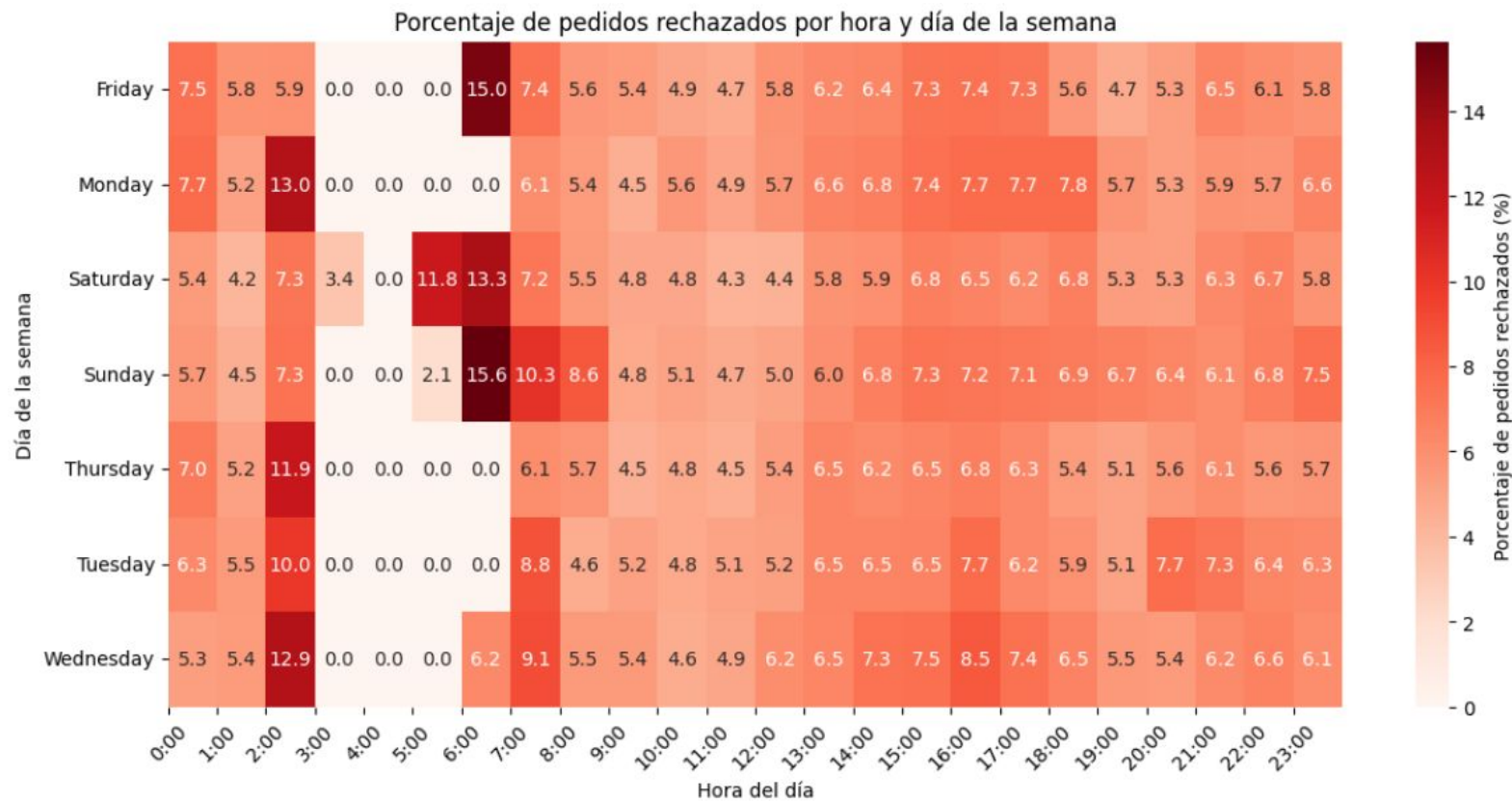




# Frecuencia de pedidos Por día de semana



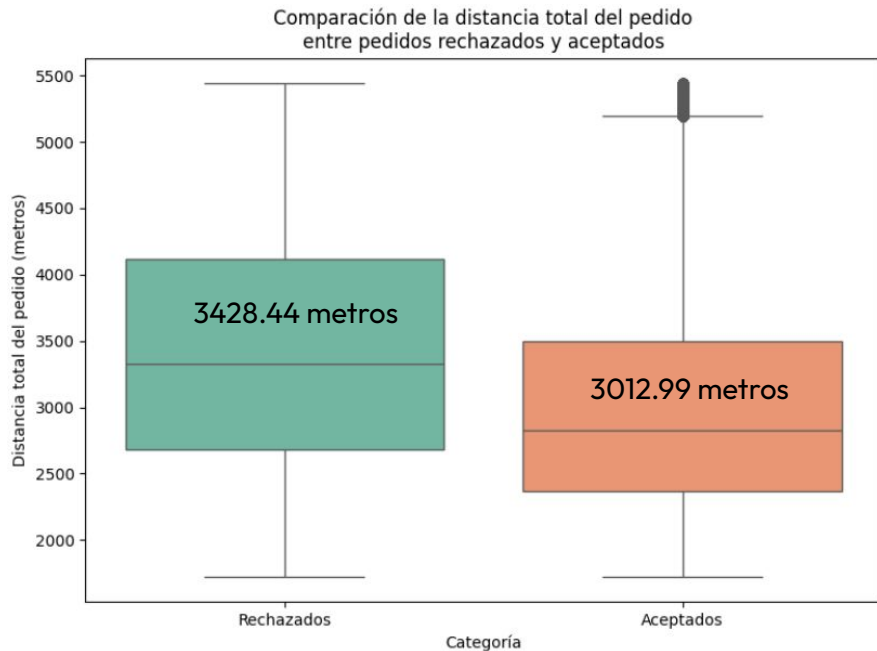
# Horarios donde más se rechazan pedidos



# Comparación de medias en pedidos Aceptados - Rechazados

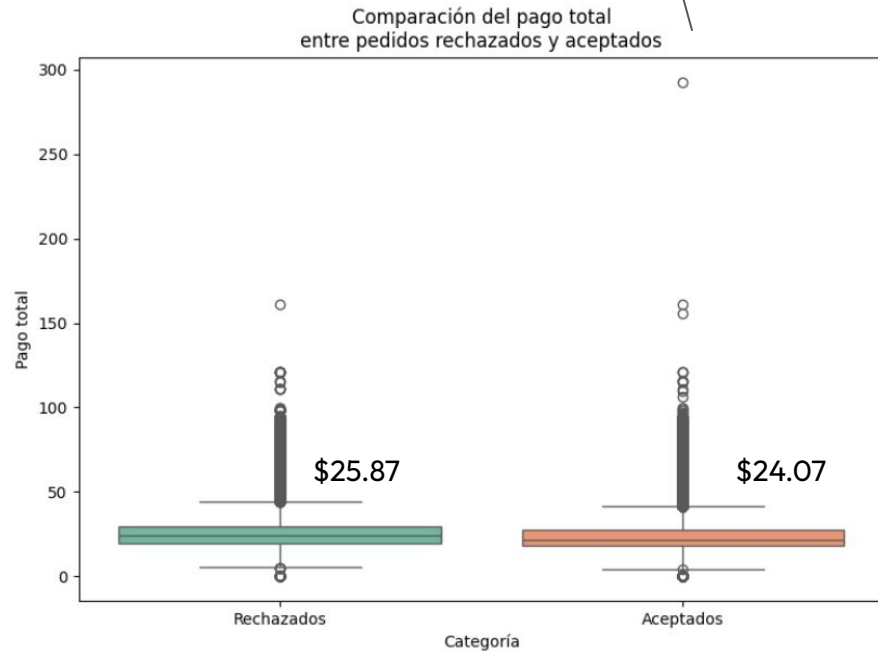


Pensamos que el pago no es suficiente para evitar que los repartidores no rechacen.



ANOVA: P-value: 0.0000

Los pedidos con mayor **distancia total** tienen más rechazos de los pedidos más cercanos para los repartidores.



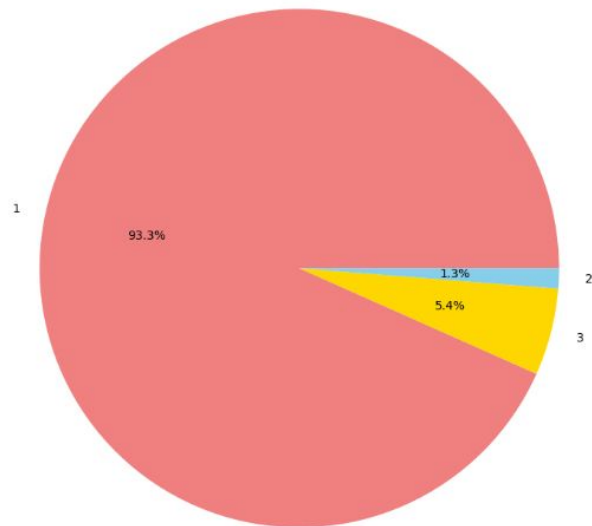
ANOVA: P-value: 0.0000

Los pedidos con mayor **pago total** tienen más rechazos de los pedidos más cercanos para los repartidores.

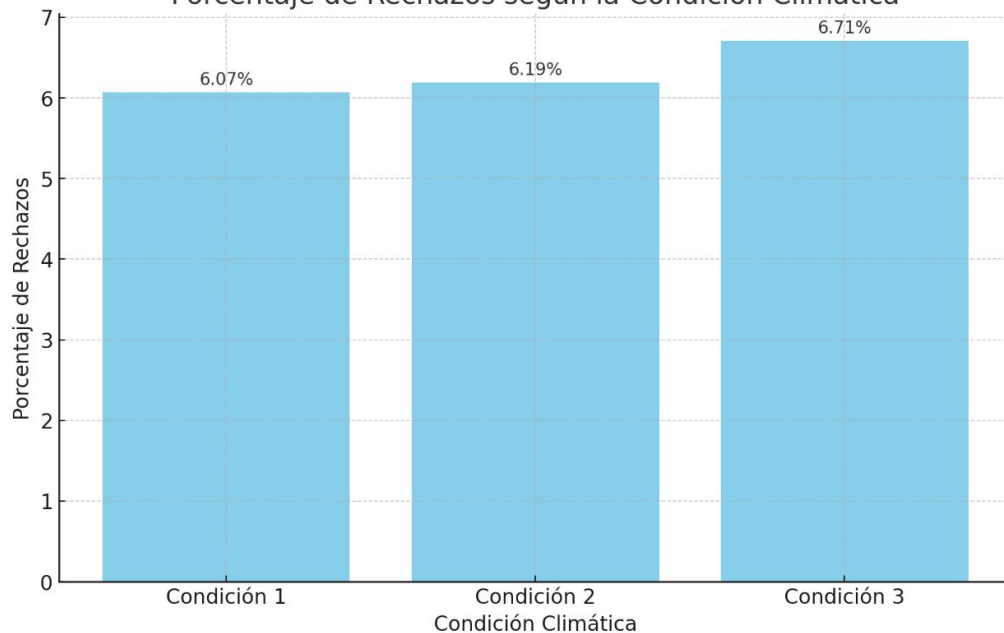
# Análisis de Clima y pedidos rechazados



Proporciones de pedidos por Tipo de clima



Porcentaje de Rechazos según la Condición Climática

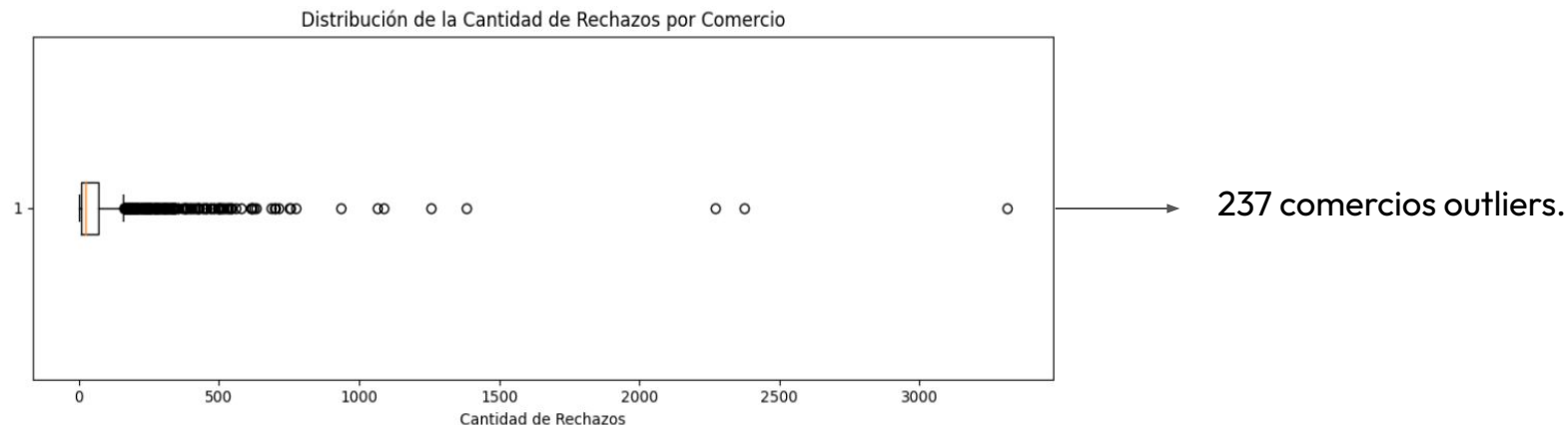


Condición 1: No llueve

Condición 2: Hay llovizna quizás

Condición 3: Llueve

# Análisis de Comercio y pedidos rechazados

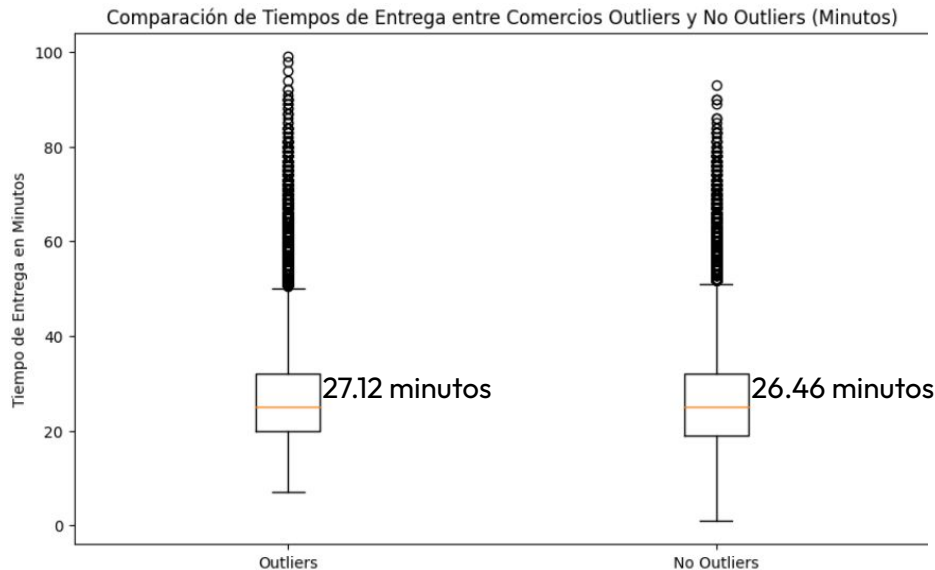


Cantidad de pedidos : 2218  
Media: 68.43  
Desvío Estándar: 147.64  
Mínimo: 1

Percentil 25: 9  
Percentil 50: 26  
Percentil 75: 70  
Máximo: 3311



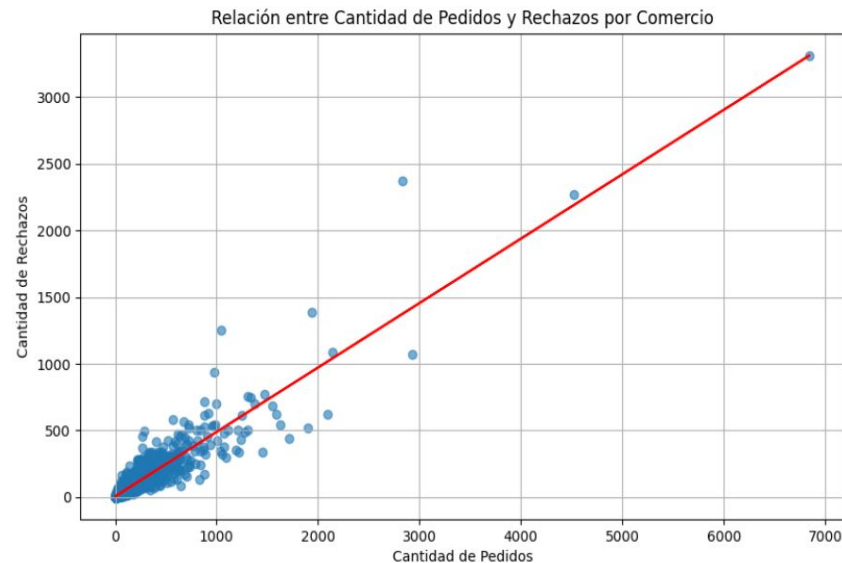
# Análisis de Comercio y pedidos rechazados



Las medias son estadísticamente diferentes ( $p < 0.05$ ).



Los comercios con mayor #rechazos tardan más tiempo en entregar los productos.

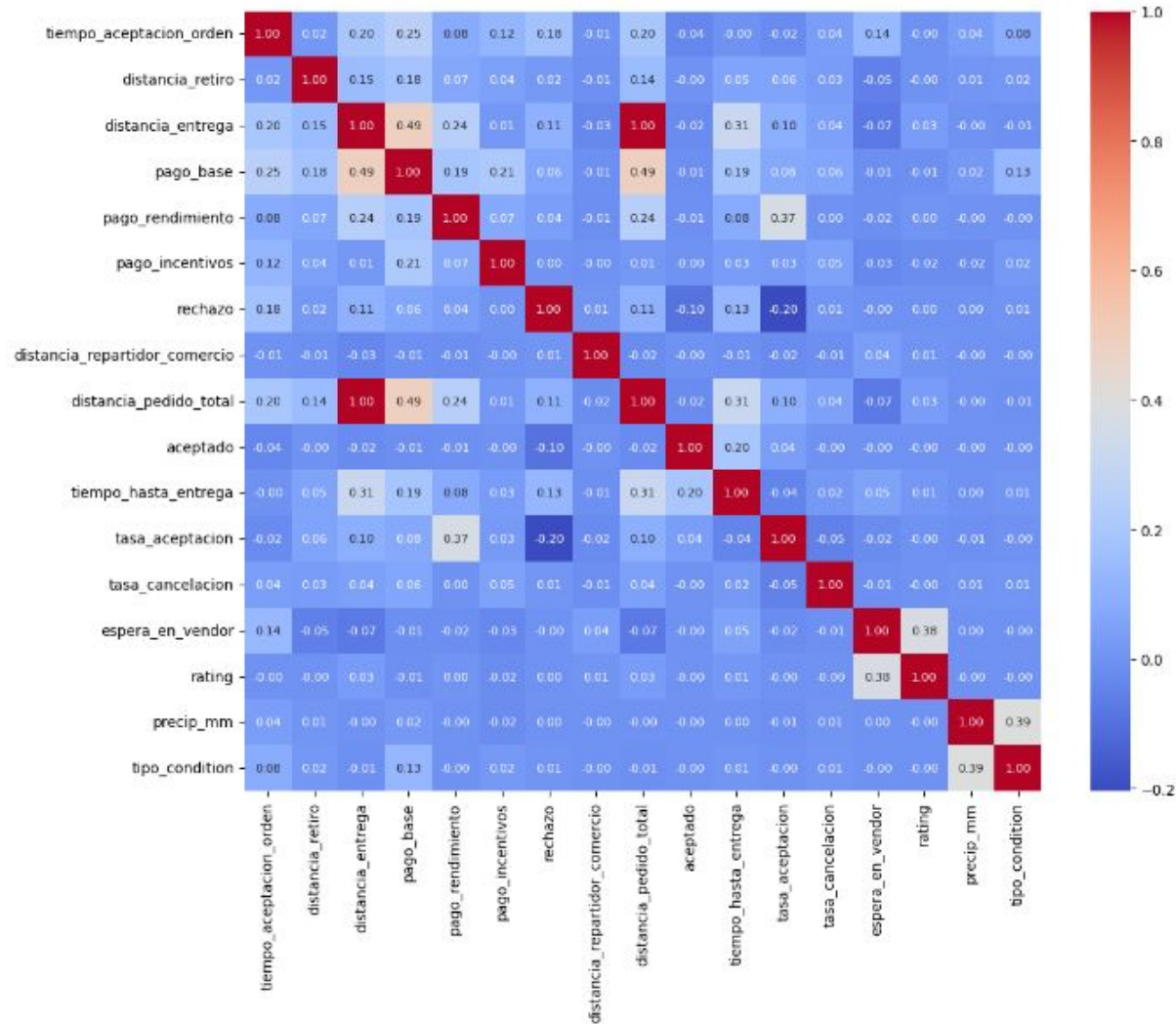


Coefficiente de Pearson: 0.93

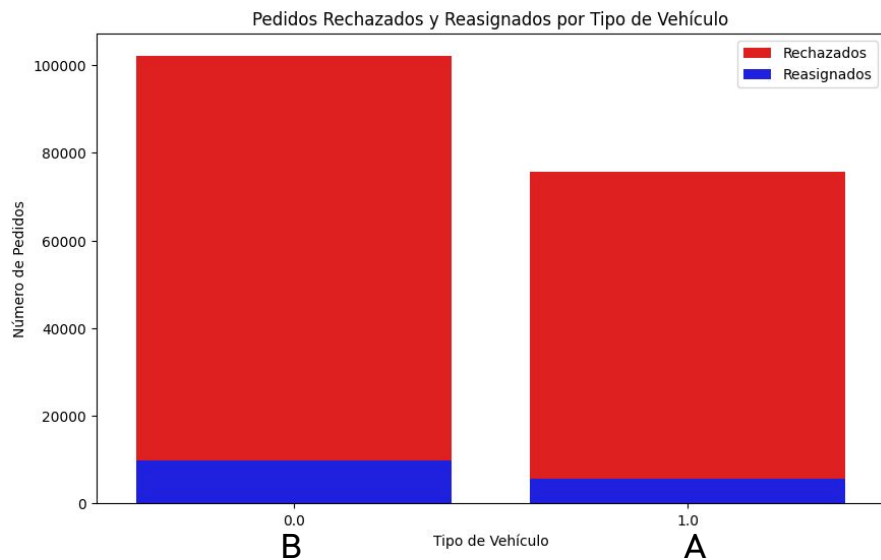


Los comercios con mayor #rechazos tienen proporcionalmente más #pedidos.

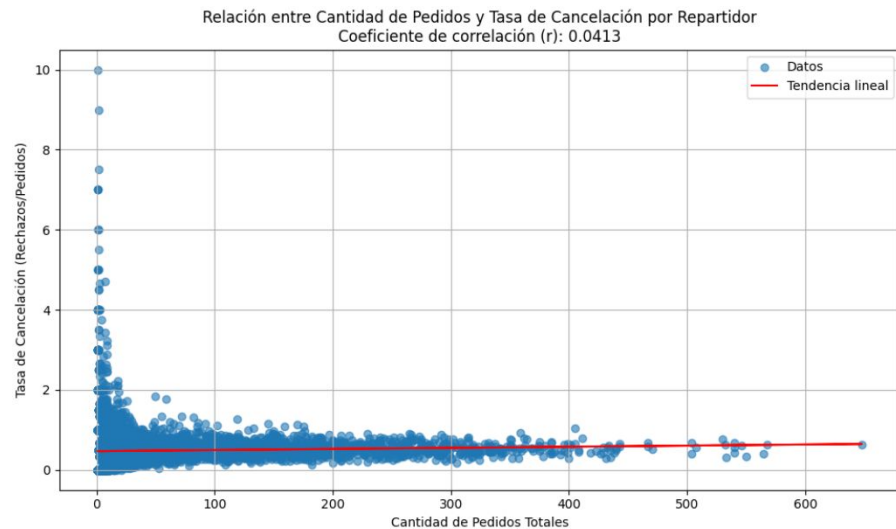
# Análisis de Correlación



# Análisis de repartidores y pedidos rechazados



El tipo de vehículo B tiene una mayor cantidad y proporción de pedidos rechazados y reasignados que el vehículo A



A mayor #pedidos totales no aumenta la tasa de cancelación.



# Modelo de Regresión Logística

Variables que a **mayor valor disminuyen** la probabilidad de **aceptación** de un pedido:

- **distancia total del pedido**
- **tiempo de espera en el comercio**
- **condiciones climáticas**
- **pago total**



Un mayor pago no siempre garantiza la aceptación del pedido. Otras variables influyen más fuerte en los rechazos.

Variables que a **mayor valor aumentan** la probabilidad de **aceptación** de un pedido:

- **rating** de un comercio
- **tiempo estimado para la entrega**



Al tener un mayor margen de tiempo les motiva más a los repartidores a aceptar ese pedido.



# Análisis Predictivo

# Features para el modelo predictivo



- 1) Se genera una partición del día 13/7/2024 en el **rango horario** de 20:30 hasta 20:45.

**Tasa de rechazo inicial = 30,42%**

**Tasa de reasignación inicial = 4,17%**

- 2) Generamos una tabla con valores numéricos de los siguientes features:

Id\_repartidor\_x  
id\_comercio  
Transicion\_id\_repartidor  
Tiempo\_aceptacion\_orden  
Distancia\_retiro  
Distancia\_entrega  
Pre\_accept\_undispatch  
Post\_accept\_undispatchs  
Pago\_base  
Pago\_rendimiento

Pago\_incentivos  
Distancia\_repartidor\_comercio  
Distancia\_pedido\_total  
Aceptado  
Tiempo\_hasta\_entrega  
Tasa\_aceptacion  
Tasa\_cancelacion  
Pedidos\_entregados  
Antigüedad\_semanas

Tiempo\_medio\_entrega  
Rating  
Espera\_en\_vendedor  
Temperature  
Precip\_mm  
Wind  
Tipo\_condition  
Creacion\_pedido\_num  
Creacion\_transicion\_num  
Vehiculo\_binario

Id\_pedido\_num  
Transicion\_accepted  
Transicion\_cancelled  
Transicion\_completed  
Transicion\_courier\_notified  
Transicion\_left\_pickup  
Transicion\_near\_dropoff  
Transicion\_near\_pickup  
Transicion\_picked\_up  
Transicion\_queued



# Modelo predictivo

## 1. Entrenamos el modelo:

- Usamos la base de datos limpia y construida con variables clave
- Se entrenaron modelos de Regresión Logística para predecir las tasas de rechazo y reasignación.

## 2. Modificamos las variables y predecimos:

- Se identificaron y modificaron las variables que afectaban significativamente las tasas de rechazo y reasignación, como la distancia del pedido y el tiempo de espera.
- Se aplicó el modelo para predecir las variables binarias: pre\_accept\_undispatch y post\_accept\_undispatch, relacionadas con las decisiones de los repartidores antes y después de aceptar los pedidos.

Espera\_en\_vendedor **-20%**

Distancia\_pedido\_total **-15%**

Pago\_incentivos **+8%**

## 3. Recalculamos las tasas de rechazo y reasignación:

- Con el modelo ajustado, se recalcularon las tasas de rechazo y reasignación para ver si nuestra propuesta era efectiva.



# Resultados



# Resultados

Propuesta Efectiva



## Impacto en el nivel de servicio:

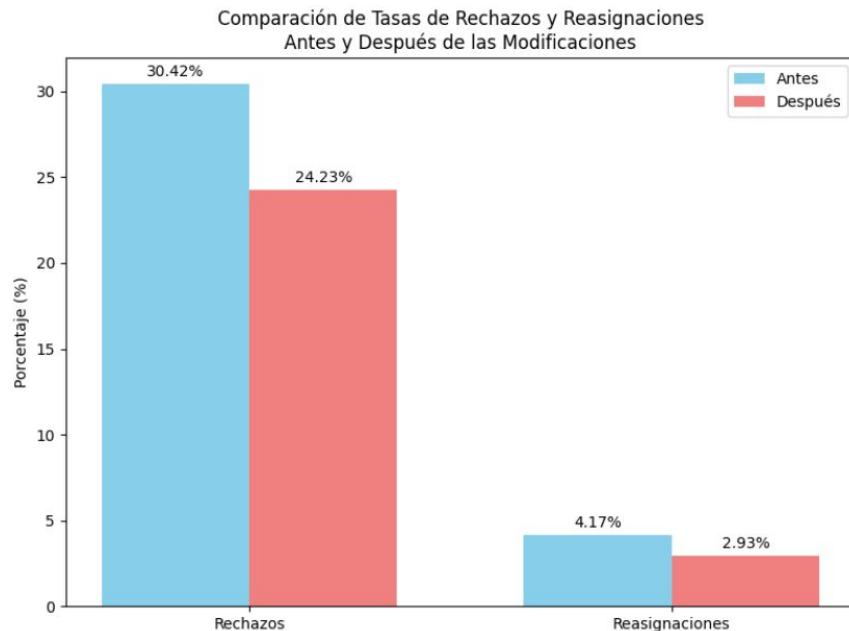
Tasa de rechazo = 24.23%

Tasa de reasignación = 2.93%

## Estimación de costos:

Costos iniciales +\$348,68

Suma total de pago\_incentivos antes del cambio: 2324.54  
Suma total de pago\_incentivos después del cambio: 2673.22





# Conclusiones



# PedidosYa: Conclusiones

**Objetivo del TP:** Identificar las variables que influyen en la tasa de rechazo y reasignación de pedidos en PedidosYa y proponer soluciones para reducirlas.

## Hallazgos Principales:

- Los incentivos que se le dan a los repartidores actualmente no ayudan a reducir los rechazos.
- Variables como tiempo de espera y distancias largas incrementan significativamente la tasa de rechazo.
- El clima afecta la aceptación.
- La **última milla** es crítica en la cadena de suministro de PedidosYa, donde cualquier retraso afecta directamente la satisfacción del cliente.



# PedidosYa: Conclusiones

## Propuestas principales

- **Reducir la espera en vendor un 20%**
  - Haciendo que el comercio priorice las órdenes de PEYA (logística)
  - Capacitación (RRHH) y estandarización de procesos (Logística)
  - Ofrecer bonificaciones por tiempos rápidos
- **Reducir la distancia entre comercio y repartidor un 15%** → Ajustando el algoritmo de asignación
- **Aumentar el pago de incentivos en un 8%** → A los repartidores que se encuentran más lejos del comercio



**¡Muchas gracias  
por su atención!**