



AMBIENTANDO FIESTAS

# Ciencia y análisis de datos de los Pedidos de Ambientando Fiestas

# Agenda

- 1) Motivación
- 2) Objetivos
- 3) Acerca de Ambientando Fiestas
- 4) Problemas de negocio
- 5) Obtención y guardado de datos
- 6) Data wrangling
- 7) Análisis de datos en base a los problemas de negocio
- 8) Conclusiones

# 1) Motivación

- ▶ Caso real de organización con datos obtenidos de diferentes fuentes
- ▶ Aplicación necesaria de contenidos vistos en el curso para manipular los datos
- ▶ Expectativa de crecimiento del dataset y mejora de modelos de ML

## 2) Objetivos

- ▶ Adquisición de datos: E-commerce y App gestión interna
- ▶ Data wrangling: Limpieza y exploración de datos de los pedidos
- ▶ Generar respuestas a problemas de negocio
- ▶ Generación de insights: Obtener información estratégica de la organización

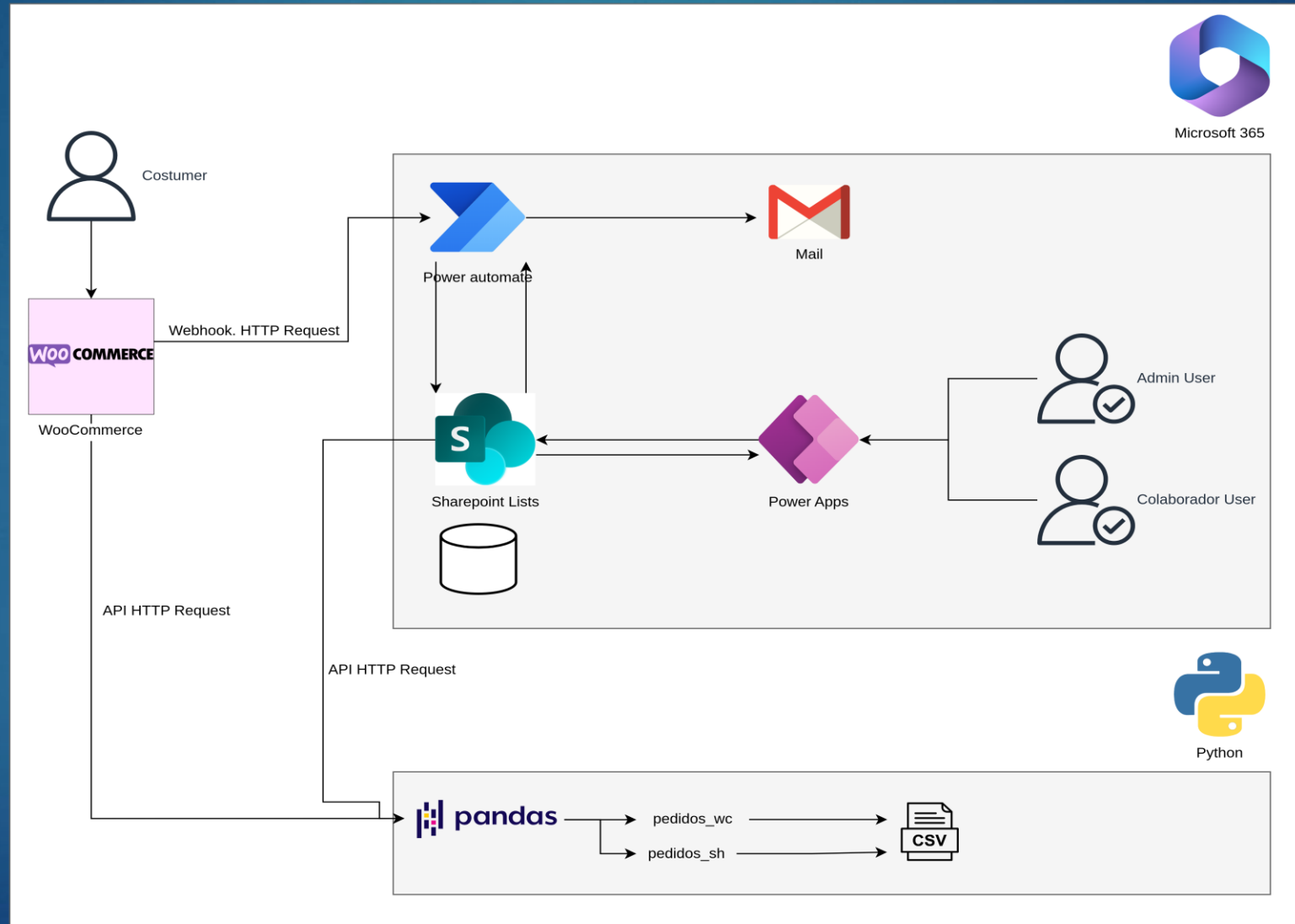
### 3) Acerca de Ambientando Fiestas

Proveedor de servicios de decoración de eventos y candy bar.

Cada evento es denominado como un Pedido

### 3) Acerca de Ambientando Fiestas

## Arquitectura de datos



### 3) Acerca de Ambientando Fiestas

#### ¿Qué es un pedido para el sistema?

Evento contratado que requiere ambientación:

Cumpleaños

Casamientos



Baby shower





## 4) Problemas de negocio

- 1) ¿Cuánto tiempo antes al día del evento reserva la gente?
- 2) Se requieren visualizar cantidad de fiestas por mes
- 3) ¿Cómo determino cuál es el momento que más gente necesito contratar?
- 4) ¿Cómo determinar cantidad de personal por zona
- 5) ¿La gente reincide con la organización? ¿Compra los mismos productos?



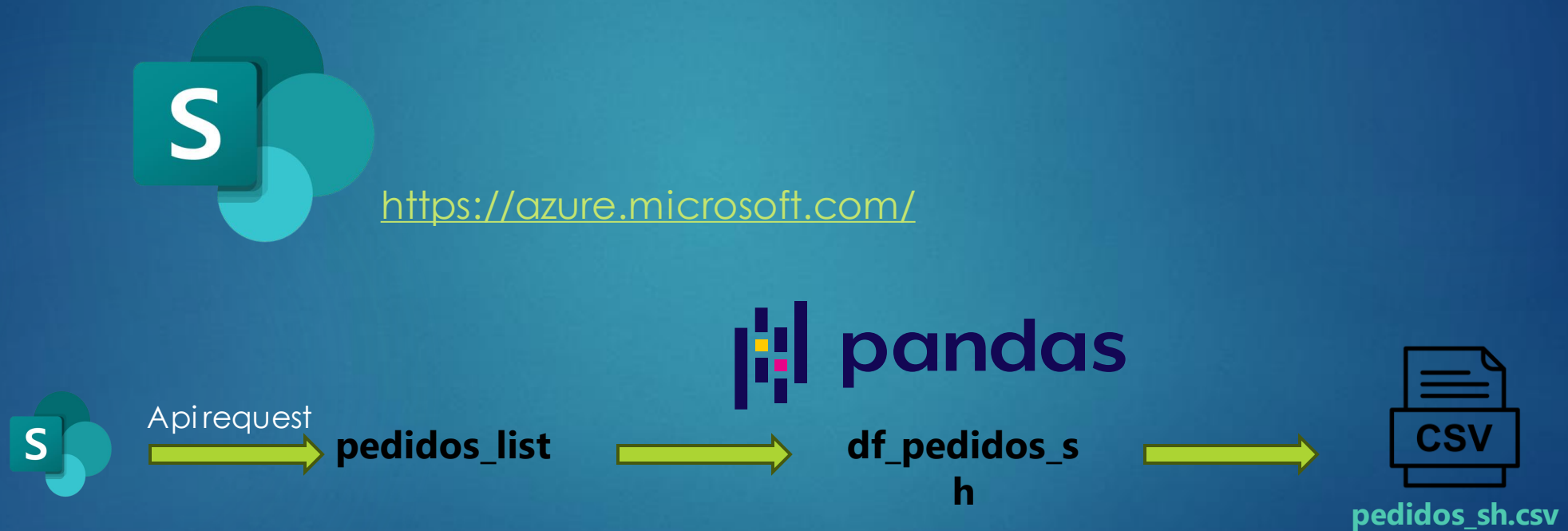
# 5) Obtención y guardado de datos

Fuentes de datos



# 5) Obtención y guardado de datos

## Esquema general



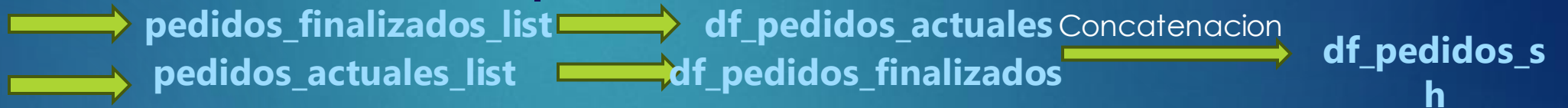
# 5) Obtención y guardado de datos

## Esquema detallado

class SharePointList

Api request (token azure)

 pandas



`df_pedidos_sh`

Exportación de datos



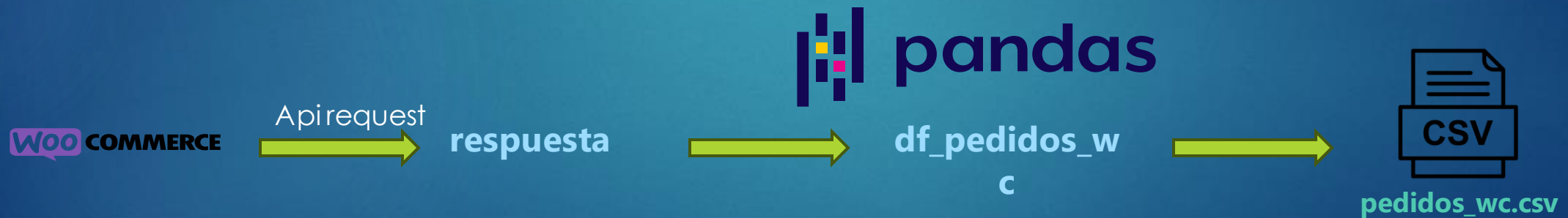
`pedidos_sh.csv`

## 5) Obtención y guardado de datos

### Esquema general



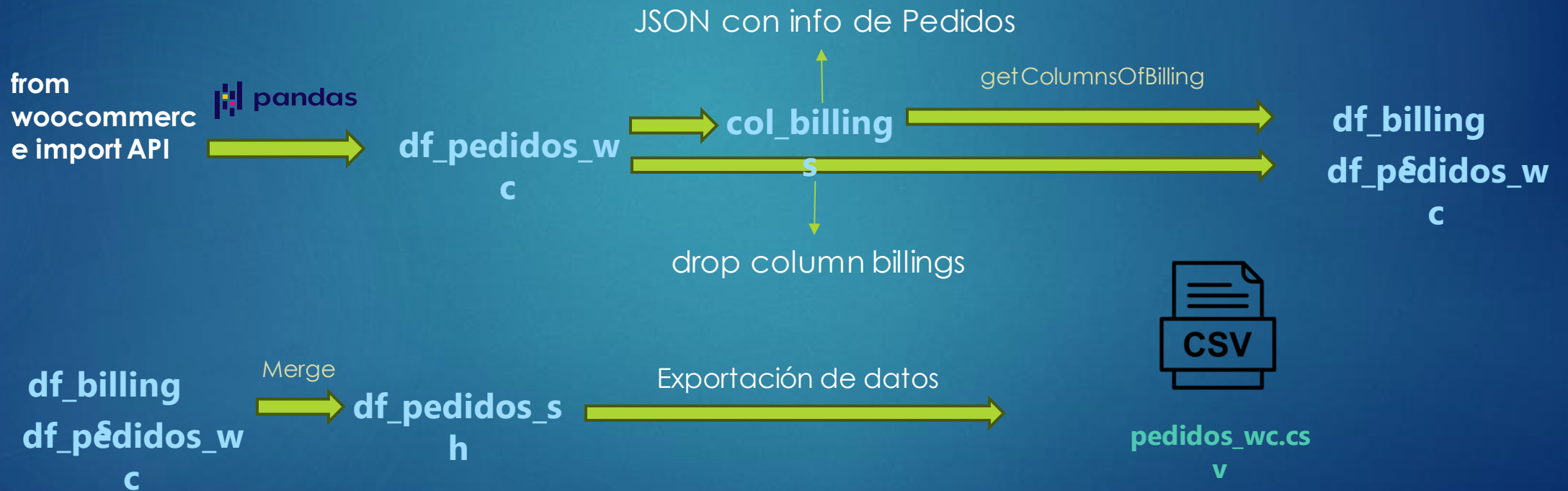
<https://woocommerce.com/document/woocommerce-rest-api/>



# 5) Obtención y guardado de datos

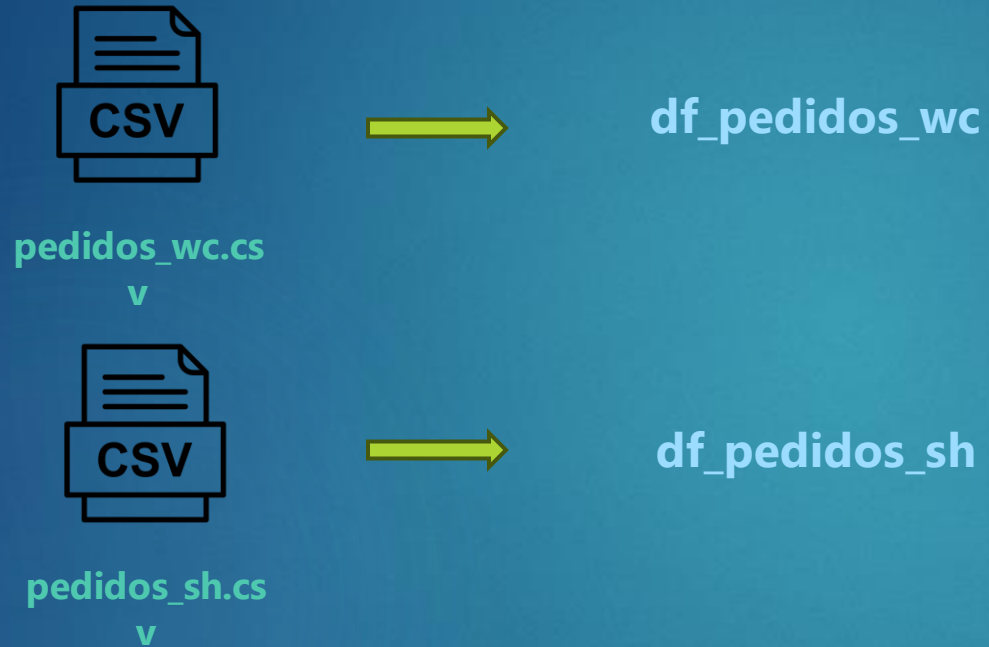
## Esquema detallado

### Api Woo-Commerce



## 6) Data wrangling

- Descargar los archivos .csv





- Combinar los archivos



## 6) Data wrangling

- Data set resultante
  - 3617 registros – 75 columnas
  - Fecha creación primer pedido: 10/02/2022
  - Fecha creación último registro: 12/06/2024

Key  Title (número de pedido)  **df\_pedidos\_wc\_s**  
h



## 6) Data wrangling

- Limpieza de datos
  - Eliminar columnas innecesarias
  - Detección de valores erróneos
  - Conversión de datos correspondientes en campos de fecha
  - Análisis de datos nulos
- Data set resultante:
  - 3594 registros – 32 columnas

## 7) Análisis de datos en base a los problemas de negocio

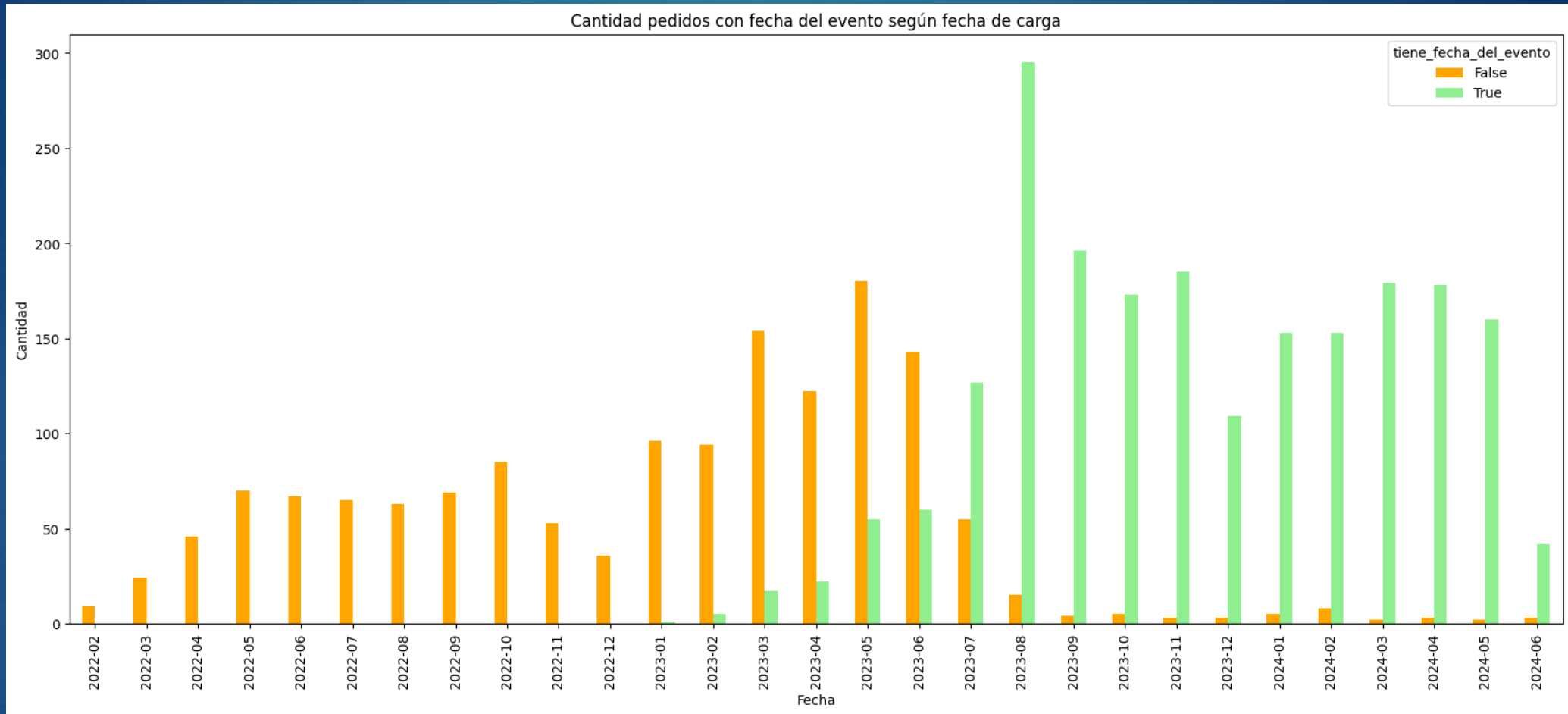
### 1) ¿Cuánto tiempo antes al día del evento reserva la gente?

Pedidos sin dato de fecha del evento -> 1484 de 3594 - (41.29 %)

(Para avanzar se busca solucionar los missing values)

## 7) Análisis de datos en base a los problemas de negocio

¿Por qué tenemos eventos que no tienen dato de fecha?



El sistema interno (App gestión interna) comenzó el 08-2023. Lo cual coincide con la fecha en que se cargaron mayor cantidad de fechas de los pedidos

## 7) Análisis de datos en base a los problemas de negocio

¿Dónde podemos conseguir los datos faltantes?

Pedido: 9989

Columna: costumer\_note

Nombre de del homenajeado/a: Victoria\r\n\* Cuantos años cumple:  
5\r\n•\t**Fecha: 27/10/24**\r\n•\tHora de inicio: 13.30\r\n\* Hora de finalización:  
16.30\r\n\* Dirección del lugar: Gral. Bartolomé Mitre 132, Lomas de Zamora,  
Buenos Aires.\r\n\* Es casa, SUM, salón de fiestas o otro?: Salón de  
fiestas\r\n•\tTemática: La Sirenita (pelirroja)\r\nQué servicio quieres: Inflable  
burbuja (Con piso inflable)

Pedido: 9403

Columna: costumer\_note

Cuantos años cumple: 1 año\r\n•\t**Fecha 21/04**\r\n•\tHora de inicio 13hs\r\n.  
Hora de finalización 18hs\r\n. Dirección del lugar: Barrio privado Pilara. Lote 291(  
La diligencia)\r\n. Es casa, SUM, salón de fiestas o otro? Casa\r\n•\tTemática  
Bambi

## 7) Análisis de datos en base a los problemas de negocio

Aplicación regex para encontrar las fechas en columna: **costumer\_note**

**El patrón que se puede ver es:**

- 28 de junio ({nro\_dia} de {nombre\_mes})
- 10/08/2024
- 16 marzo ({nro\_dia} {nombre\_mes})
- 26.11.23 ({nro\_dia}.{nro\_mes}.{nro\_año})

```
pattern = r'\b(\d{1,2} de [a-zA-Z]+(?!s*HS) | \d{1,2}(:/\d{1,2})?/\d{2,4} | \d{1,2}\.\d{1,2}\.\d{2,4})\b'
```

Fechas encontradas con regex: 1433

Fechas NO encontradas con regex:

2161

Cantidad de datos  
que se podrían  
recuperar

## 7) Análisis de datos en base a los problemas de negocio

### Datos recuperados solo aplicando el regex

Pedidos sin dato de fecha del evento -> 1226 de 3594 - (34.11 %)

Cantidad de eventos agregados: 258

### Ejemplos datos no recuperados:

- 28 de junio



## 7) Análisis de datos en base a los problemas de negocio

### Datos recuperados solo aplicando el regex

Pedidos sin dato de fecha del evento -> 1226 de 3594 - (34.11 %)

Cantidad de eventos total recuperados: 258

### Ejemplos datos no recuperados:

- 28 de junio



## 7) Análisis de datos en base a los problemas de negocio

### Datos recuperados solo aplicando el regex

Pedidos sin dato de fecha del evento -> 1226 de 3594 - (34.11 %)

Cantidad de eventos agregados: 258

### Ejemplos datos no recuperados:

- 28 de junio

## 7) Análisis de datos en base a los problemas de negocio

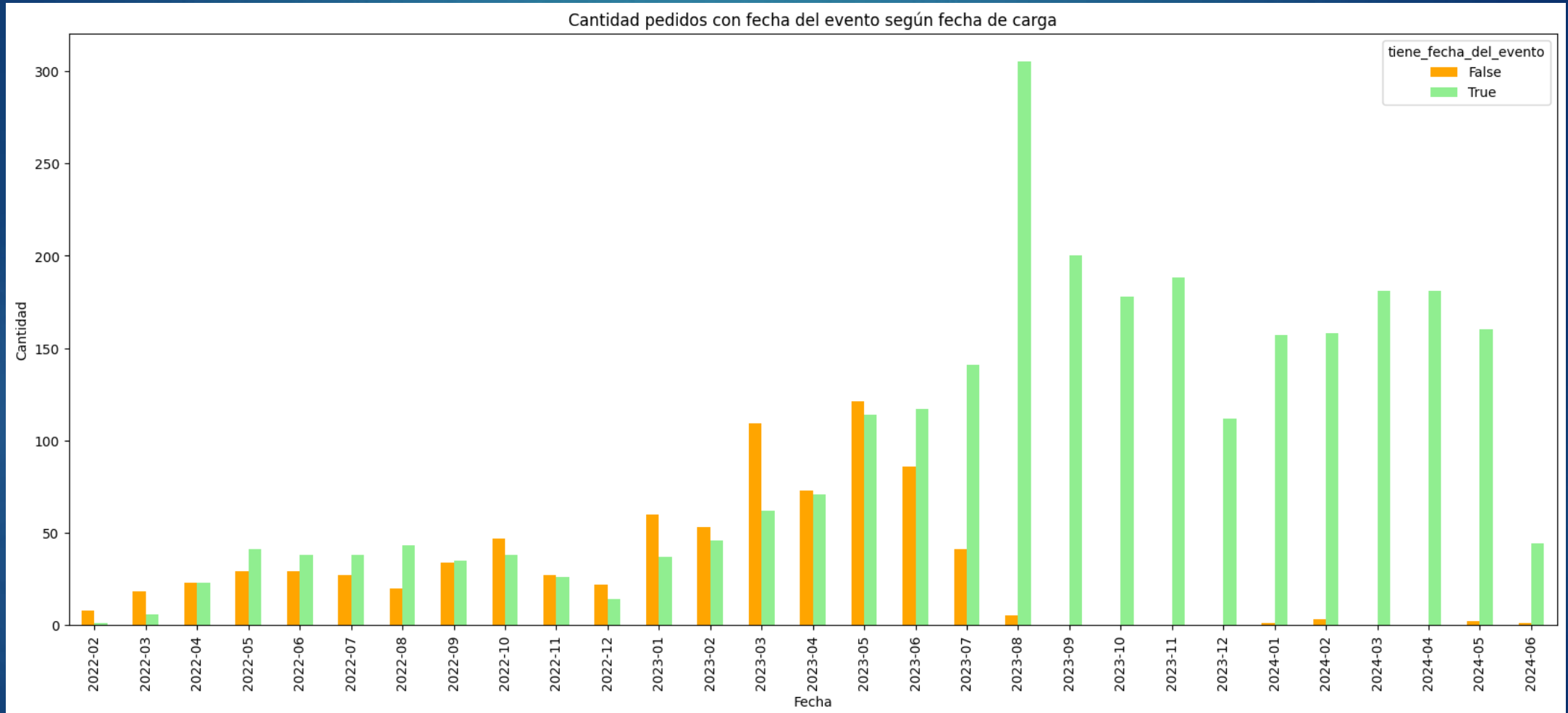
### Datos recuperados solo aplicando funciones

Pedidos sin dato de fecha del evento -> 839 de 3594 - (23.34 %)

Cantidad de eventos agregados: 645

# 7) Análisis de datos en base a los problemas de negocio

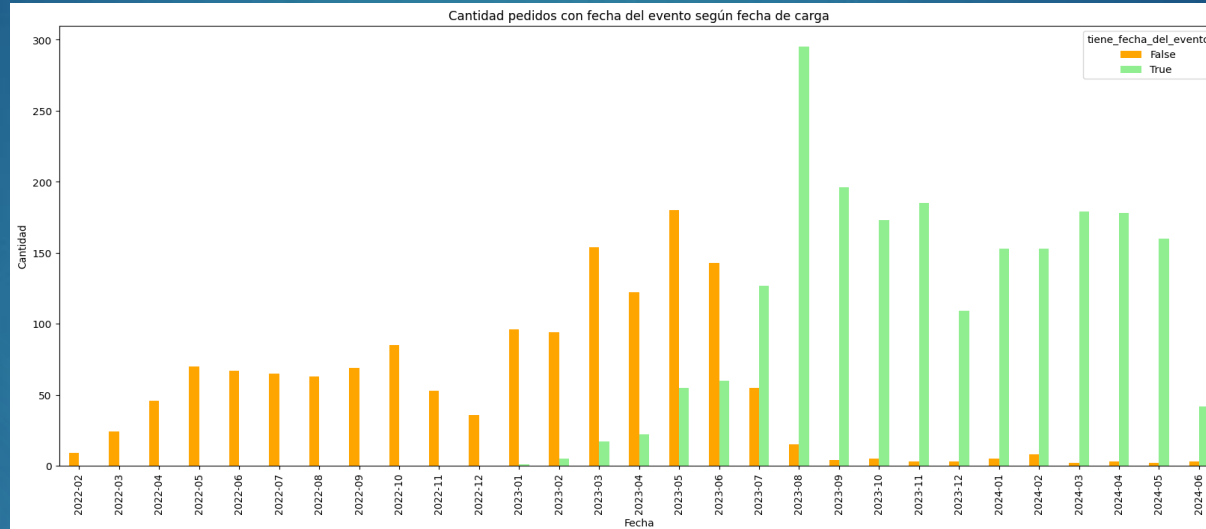
Estado actual de pedidos con y sin fecha del evento



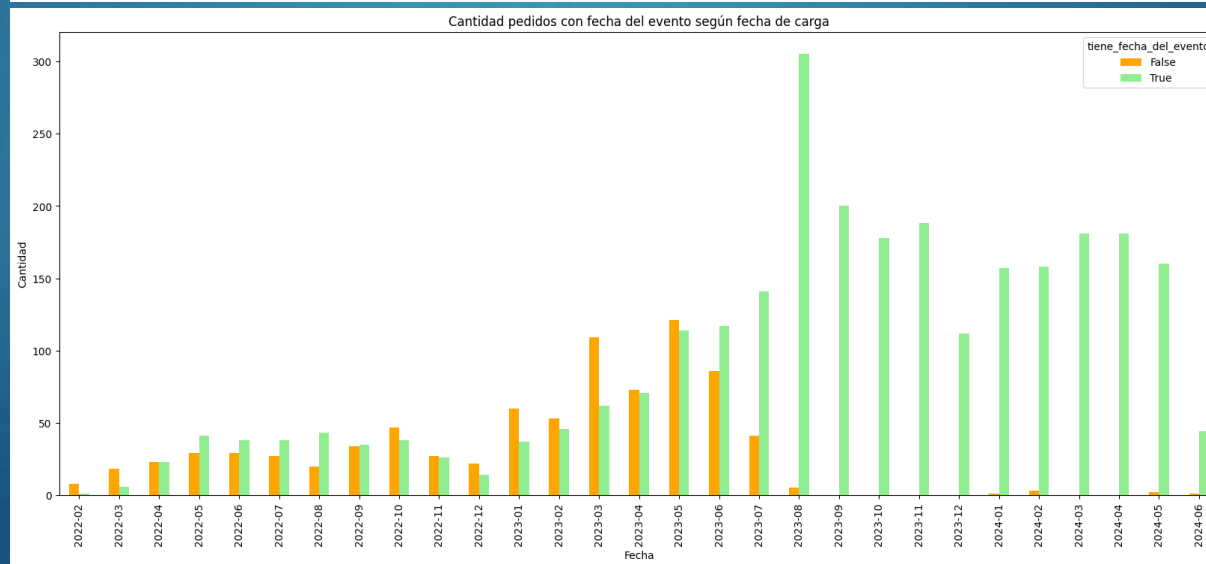
# 7) Análisis de datos en base a los problemas de negocio

Comparación estado inicial y actual de pedidos con y sin fecha del evento

Estado inicial



Estado actual



## 7) Análisis de datos en base a los problemas de negocio

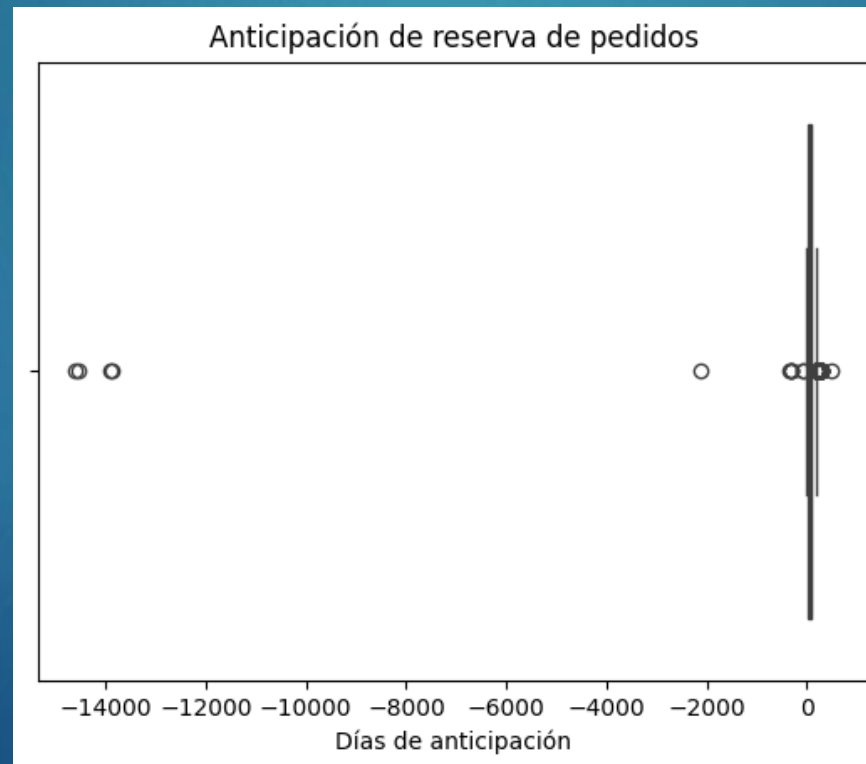
### 1) **¿Cuánto tiempo antes al día del evento reserva la gente?**

Utilizamos solo los datos que tienen inicio\_evento

## 7) Análisis de datos en base a los problemas de negocio

### 1) ¿Cuánto tiempo antes al día del evento reserva la gente?

Utilizamos solo los datos que tienen inicio\_evento

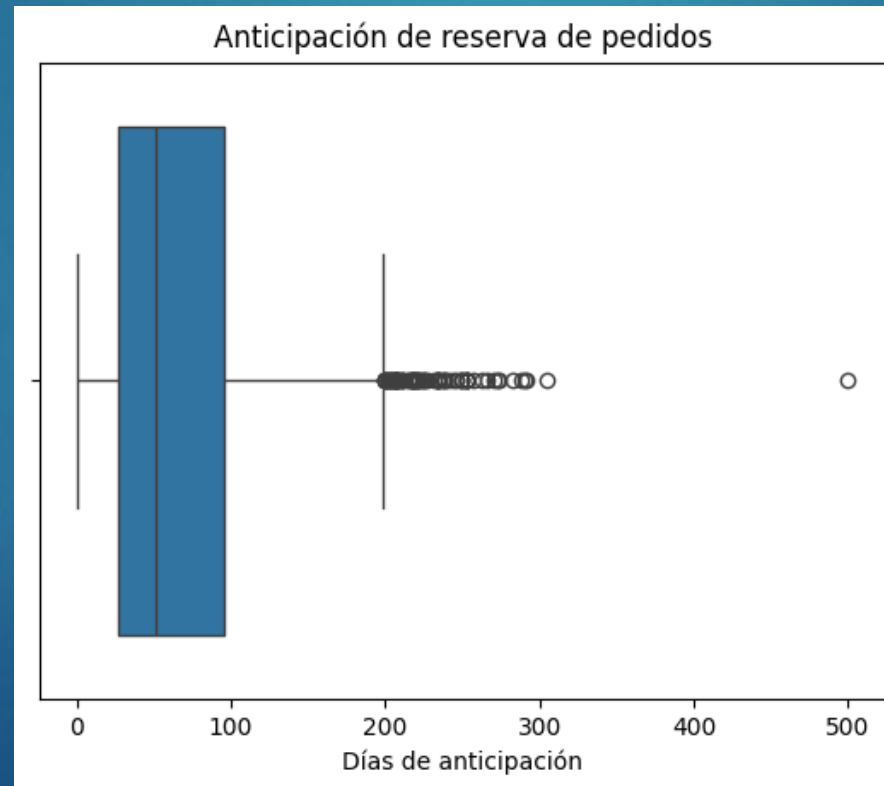


Tenemos outliers

## 7) Análisis de datos en base a los problemas de negocio

### 1) ¿Cuánto tiempo antes al día del evento reserva la gente?

Utilizamos solo los datos que tienen inicio\_evento



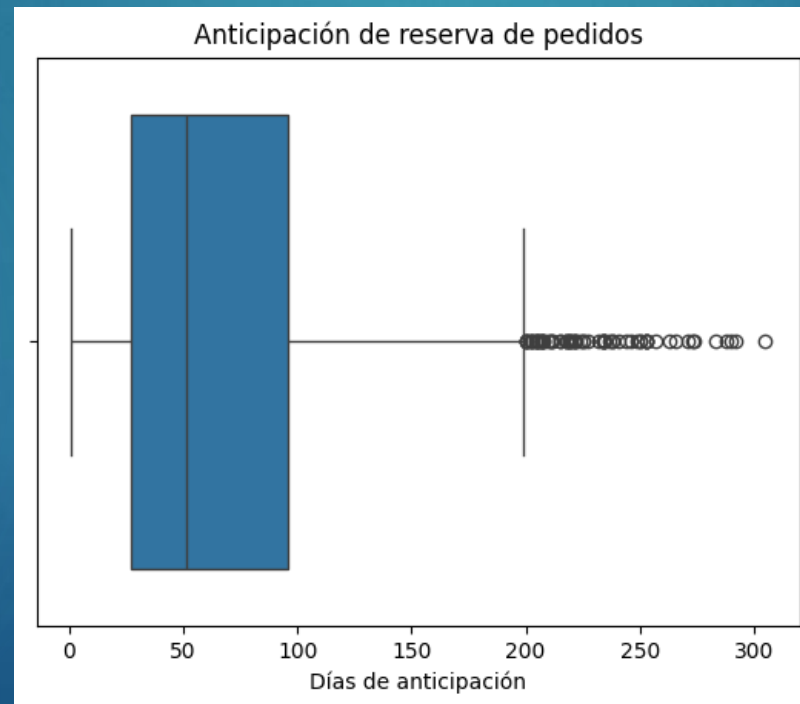
Tenemos outliers



## 7) Análisis de datos en base a los problemas de negocio

### 1) ¿Cuánto tiempo antes al día del evento reserva la gente?

Utilizamos solo los datos que tienen inicio\_evento



## 7) Análisis de datos en base a los problemas de negocio

### 1) ¿Cuánto tiempo antes al día del evento reserva la gente?

Utilizaremos solo los datos que tienen inicio\_evento

