

# Sweet, sweet Data Science

Caso de estudio real:

Predicción de la demanda de un obrador mediante series temporales

Miguel Ruiz Nogues  
Illán Lois Bermejo



# Presentación

- 1) Situación actual
- 2) Propuesta de mejora
- 3) Modelo
- 4) Project planning
- 5) Conclusiones



## **1) Situación actual:**

- El obrador de una cadena pastelera de Madrid, requiere predecir el número de cada producto que las tiendas le solicitan durante los períodos críticos
- Actualmente el proceso se hace con criterio experto y estimaciones sencillas

## **2) Propuesta de mejora:**

- Implementar un modelo predictivo que realice estimaciones en base al histórico de pedidos
- Beneficios esperados:
  - Mejorar la predicción que hacen a mano
  - Mejorar la exactitud de la predicción



### 3) Modelo:

#### 1) Datos:

Datos mínimos

Histórico de pedidos  
Mayor antigüedad posible

- Fecha
- Id producto
- Cantidad

Datos adicionales

En el histórico, incluir la tienda

Datos de los productos (ingredientes, etc)

Datos de campañas comerciales



Todavía no hemos tenido acceso a los datos



## **2) Tecnologías a emplear**

Con seguridad

Series temporales  
Google Notebooks / jupyter  
Github – Repository  
Github - Wiki  
Bash  
Tableau / Google dashbord

Previsión

Web-Scraping  
API  
Clustering



## **3) Presentación y análisis de resultados**

Comprobación del modelo en Campaña de Navidad

## 4) Project planning:

### 1) Datos

Recolección

Preparación (Data Engineering)

Exploratory Data Analysis

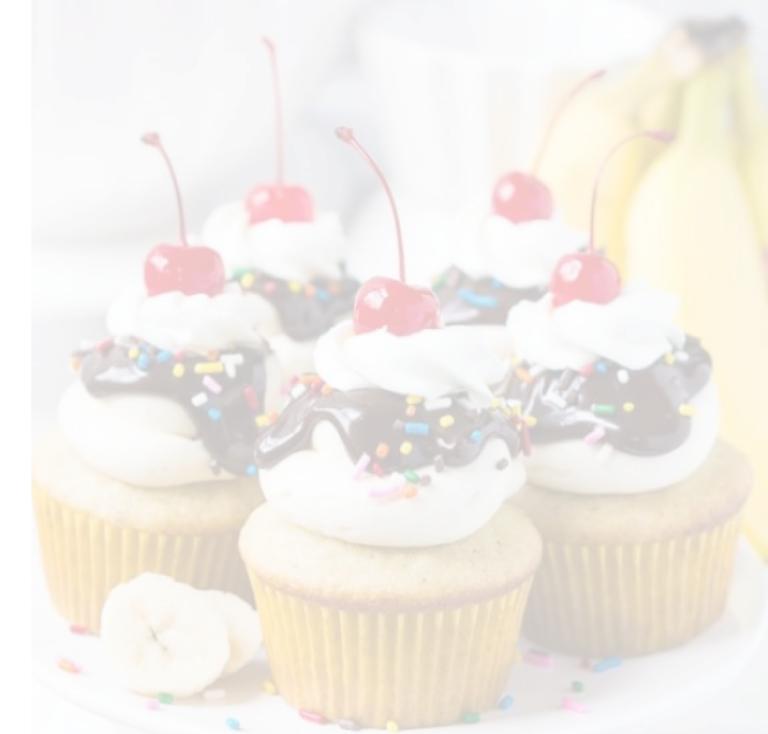
### 2) Modelo Predictivo

Time series – Arima to EDA+Prophet+ MLP Neural Network

Forecasting

Clustering de productos

### 3) Industrialización

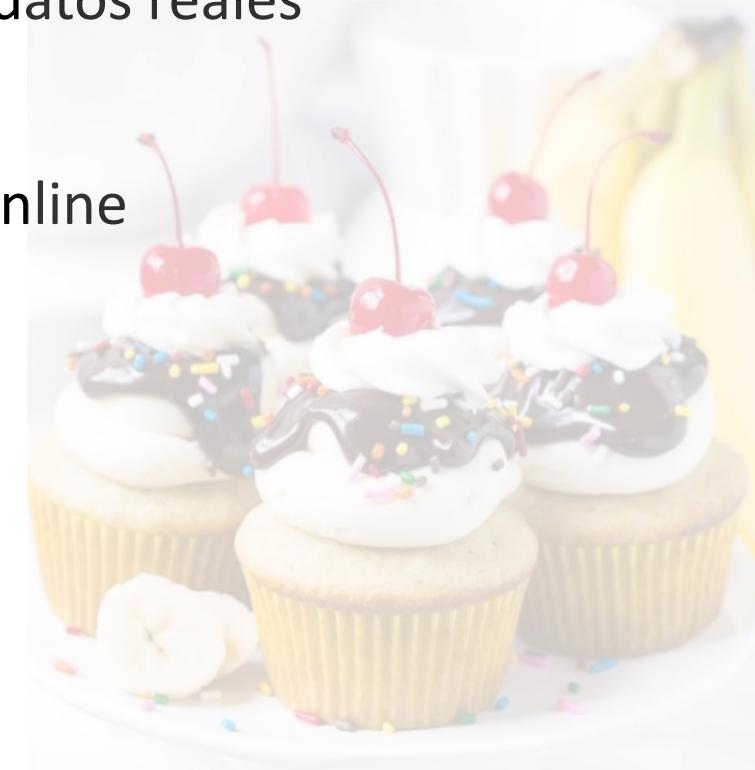


## 5) Conclusiones:

1) Proyecto real

Datos reales

Problemas de tratamiento de datos reales



2) Fuerte componente  
de herramienta

Desarrollo de una aplicación online

3) Análisis con series temporales

4) Trata toda la cadena  
de valor del dato

5) Inconvenientes

Ausencia de modelos de ML

Dificultad de implementar web-scraping o APIs

6) Riesgos

Ausencia de datos, falta de calidad de los datos

Dificultad para llegar a conclusiones