

# ANÁLISIS DE RESTAURANTES

---

ALUMNOS:

✚ ALEJANDRO FERNÁNDEZ TRIGO

✚ JUAN DIEGO VILLALOBOS QUIRÓS



HITO III – INFORME DE RESULTADOS  
INTELIGENCIA EMPRESARIAL – CURSO 2021/2022

## ÍNDICE

1. RESUMEN DE OBJETIVOS
2. HERRAMIENTAS UTILIZADAS
3. API DE YELP
4. PROCESADO – MICROSOFT POWER BI
5. CONCLUSIONES EXTRAIDAS
6. BIBLIOGRAFÍA, ADJUNTOS Y NOTAS

## RESUMEN DE OBJETIVOS

Retomando el contexto de las entregas previas, el presente trabajo abarca el estudio, mediante las herramientas vistas en la asignatura, de un 'dataset' que recoge todos los restaurantes de alta cocina según la prestigiosa Guía Michelin.

El primer adelanto pretendía analizar el 'dataset' para determinar algunas cuestiones básicas (que tipos de comida son más premiados, que países tienen mejor cocina, etc.), pero fue evolucionando hasta el día presente e implementamos por el camino algunas técnicas nuevas como la extracción de datos complementarios mediante una API para hacer un análisis más complementado.

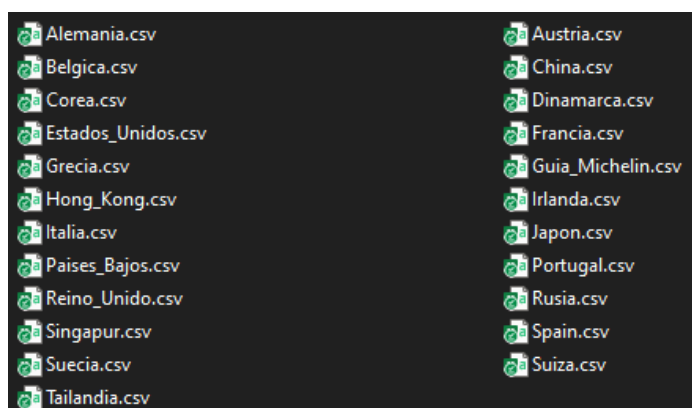
## HERRAMIENTAS UTILIZADAS

Para la realización del trabajo pues, tal y como se comentó en el hito II de la asignatura, se ha hecho uso principalmente de la herramienta Microsoft Power BI, así como del lenguaje de programación Python, la API de servicios de terceros (concretamente Yelp) y lógicamente, la fuente de datos original, el 'dataset' descargado desde Kaggle (presente en [ESTA](#) dirección).

## API DE YELP

Se procedió a generar (mediante filtros desde la herramienta de MS Power BI), múltiples ficheros .CSV correspondientes a cada país. Para ello, primeramente, se determinó un conjunto de países significativos que serían nuestra muestra de estudio.

Cabe destacar que la Guía Michelin no incluye datos de todos los países del mundo y en el caso de algunos, los datos están poco actualizados por lo que se realizó esta criba inicial, quedando así pues 22 países significativos de Europa, América y el continente asiático.



Esta fue la entrada del 'script' desarrollado en lenguaje Python para hacer llamadas a la API del servicio Yelp tomando como parámetros el nombre, longitud y latitud.

```
# Dada la información de un restaurante, obtener datos de Yelp.
def obtenerInformaciónRestaurante(term, latitude, longitude, pais):

    query = {'term':term, 'latitude':latitude, 'longitude':longitude}
    my_headers = {'Authorization' : 'Bearer .....'}
    response = requests.get('https://api.yelp.com/v3/businesses/search', headers=my_headers, params=query)
    path = r'datasets/' + pais + '_yelp.json'

    with open(path, 'w') as outfile:
        outfile.write(response.text)
    return response


# Dado un país, obtener sus nombres de restaurantes, latitudes y longitudes.
def procesarPaís(p):
    rows = []
    path = r'datasets/' + p + '.csv'
    pais = open(path, encoding='Latin1')
    csvreader = csv.reader(pais)
    header = next(csvreader)
    headers = header[0].split(";")

    term = []
    lat = []
    lon = []
    rows = []

    for row in csvreader:
        restaurante = row[0].split(";")
        if(len(restaurante) > 7):
            if(restaurante[0] != "" and restaurante[5] != "" and restaurante[6] != ""):
                term.append(restaurante[0])
                lat.append(restaurante[5])
                lon.append(restaurante[6])

    pais.close()
    return (term, lat, lon, p)
```

Copia completa del código puede encontrarse [AQUÍ](#). Este script permitió generar un fichero CSV resultante a modo de segunda tabla que contenía nuevos datos para cada restaurante seleccionado, siendo estos datos:

-  Delivery
-  Pickup
-  City
-  Zip code
-  Country
-  State
-  Price
-  Reviews
-  Rating
-  Name

## PROCESADO – MICROSOFT POWER BI

Sumado al proyecto original que se explicó en el hito II, se procedió a añadir el nuevo CSV generado como una segunda entrada (tabla) al ya existente proyecto de Microsoft Power BI.

Nuevamente, este proceso requirió de ciertos arreglos:

### ▲ PASOS APLICADOS

Origen ✱  
Dividir columna por delimi... ✱  
Tipo cambiado  
Filas superiores quitadas ✱  
Valor reemplazado ✱  
Tipo cambiado1  
Columnas con nombre ca...  
Errores quitados  
Columnas quitadas  
Columnas con nombre ca...  
Columnas quitadas1  
Valor reemplazado1 ✱  
Valor reemplazado2 ✱  
Valor reemplazado3 ✱  
Valor reemplazado4 ✱  
Valor reemplazado5 ✱  
Valor reemplazado6 ✱  
Valor reemplazado7 ✱  
Valor reemplazado8 ✱

- Dividir las columnas del fichero importado (los ficheros CSVs generados, separados por comas necesitan de este paso intermedio).
- Cambiar los tipos numéricos.
- Cambiar los tipos de monedas a una única variable, esto es, cambiar los símbolos de Euro, Yuan, etc. por \$, \$\$, \$\$\$, \$\$\$\$ y \$\$\$\$\$ en cada caso.
- Eliminar datos innecesarios como el código postal o el estado (provincia).
- Cambiar los tipos naturalmente booleanos (0/1) por tipos True/False reales que MS Power BI entienda.

Una vez limpio, se procedió a relacionar ambas tablas mediante la clave 'Nombre' para poder realizar gráficas que mezclasen atributos de ambas tablas ya que datos tan relevantes como el nº de estrellas Michelin solo están presente en la tabla origen.



Datos originales.csv



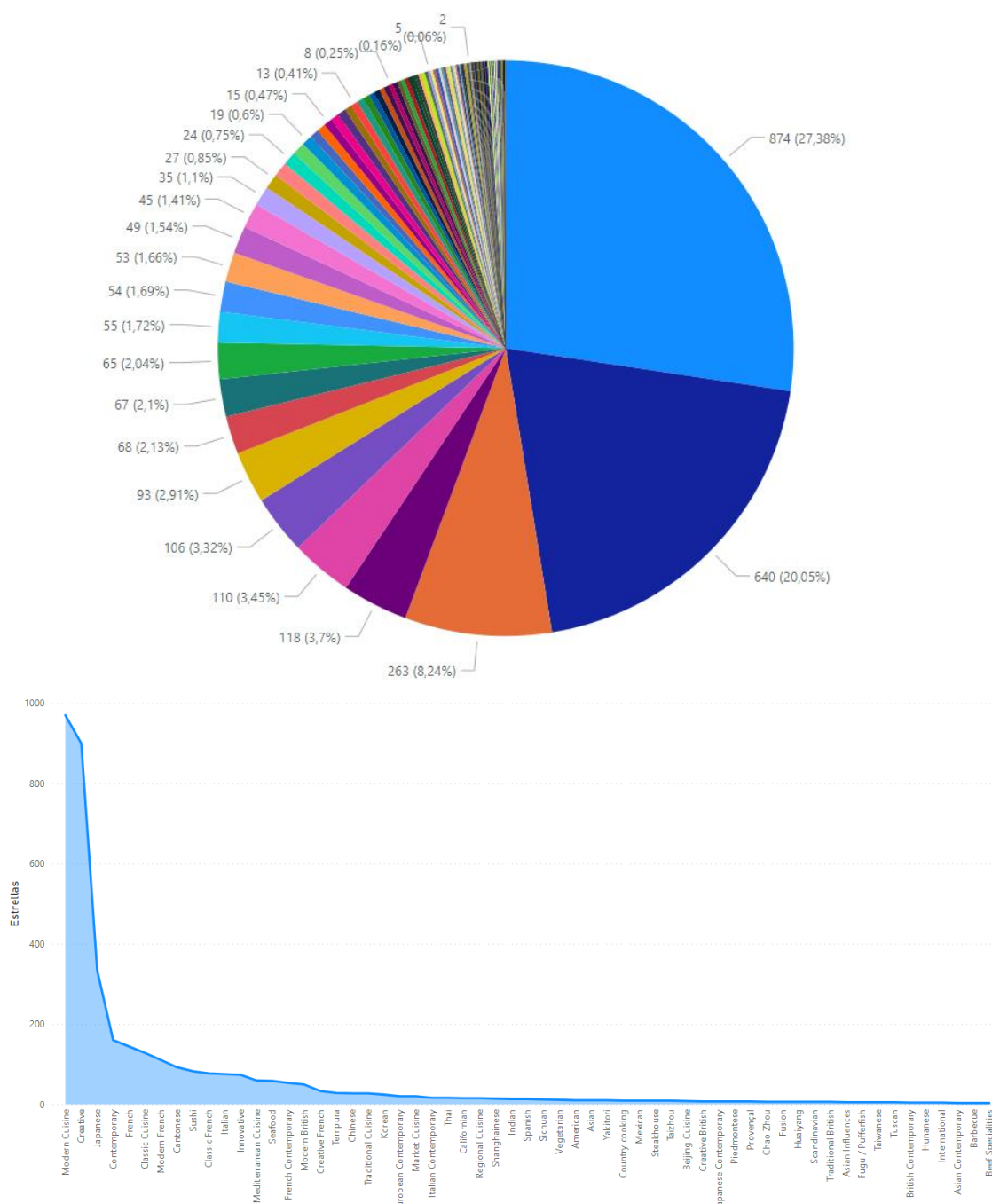
Datos de Yelp.csv

## CONCLUSIONES EXTRAIDAS

Hecho este procesamiento, se comenzaron a extraer conclusiones mediante la generación de gráficas en la misma herramienta de Microsoft Power BI, siendo estas las siguientes:

### Cocina

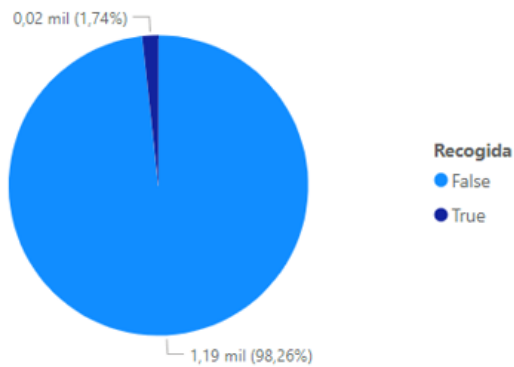
- Modern Cuisine
- Creative
- Japanese
- Contemporary
- French
- Classic Cuisine
- Modern French
- Cantonese
- Italian
- Sushi
- Classic French
- Mediterranean Cuisine
- Innovative
- Seafood
- Modern British
- French Contemporary
- Creative French
- Traditional Cuisine
- Chinese
- Tempura
- Market Cuisine
- European Contemporary
- Korean
- Italian Contemporary
- Regional Cuisine
- Thai



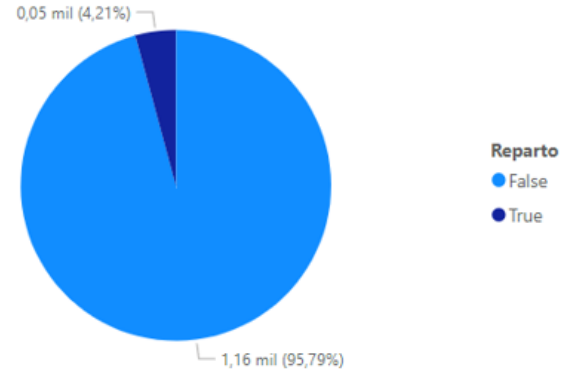
Se puede apreciar a simple vista cuales son los tres tipos de cocina más recurrentes, la cocina moderna, la cocina creativa y la cocina japonesa son las tres modalidades más comunes y así mismo, las más premiadas.

Seguidamente, podemos ver como los restaurantes de alta cocina no son propensos en ningún caso al modelo de negocio del reparto a domicilio, mucho menos al sistema de 'takeaway' o pedir para llevar.

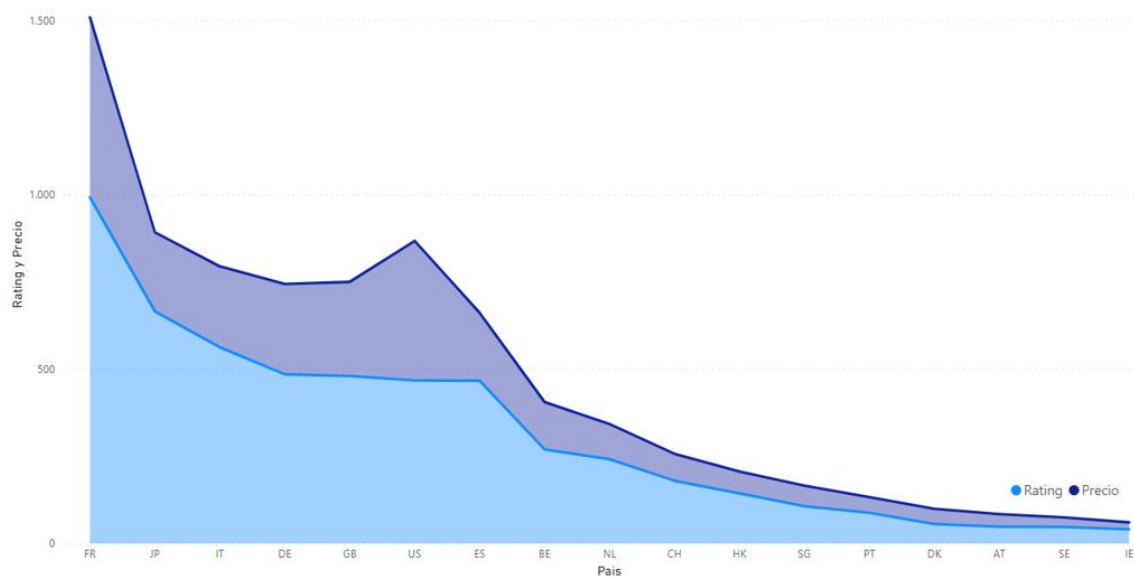
Recuento de Recogida por Recogida



Recuento de Reparto por Reparto



No obstante, hemos observado como esta tendencia se incumple en la mayoría de los casos cuando nos referimos a restaurantes estadounidenses.



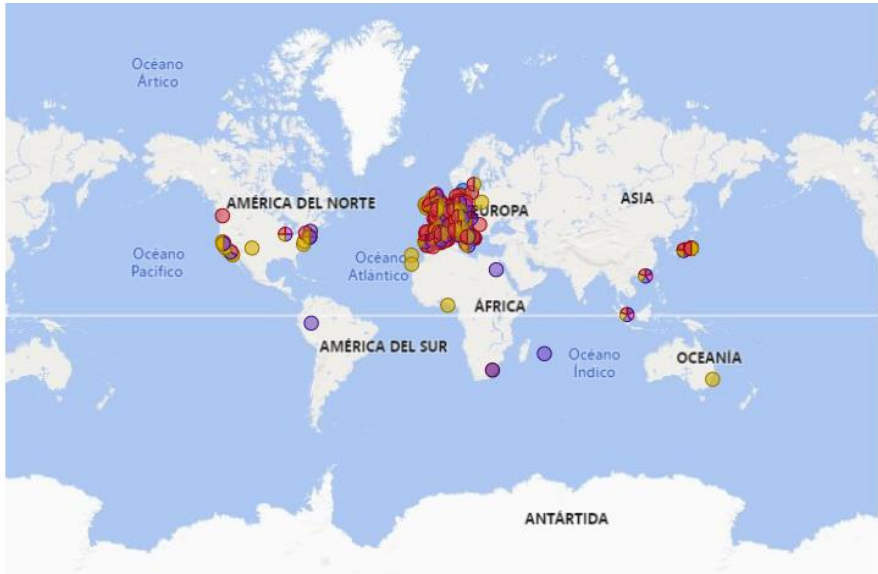
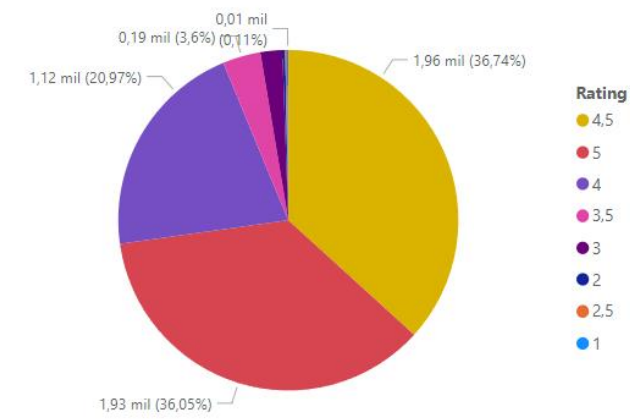
Podemos ver a continuación como el país mejor valorado por los usuarios, de acuerdo con los datos provistos por Yelp, es Francia. Sin embargo, es también con diferencia el país más caro de todos. Es seguido por Japón e Italia.

Continuando con el rango de precios, en el siguiente gráfico vemos la distribución de la puntuación que otorgan los usuarios frente al precio del restaurante. Podemos observar como los restaurantes más baratos (\$) tienen una alta puntuación, aquellos en un rango medio (\$\$, \$\$\$) acumulan las peores calificaciones y, sin embargo,

los más caros (\$\$\$\$ , \$\$\$\$\$) son los que gozan de las notas más altas según los comensales.



Podemos ver como es el reparto general de calificaciones otorgadas por los usuarios a lo largo del mundo (restringido a los 22 países que tomamos de muestra):





Finalmente, cerramos el estudio destacando el ‘podio’ de chefs mejor valorados del mundo, compuesto por:



**Mathieu Pacaud**



**Arkadiusz Zuchmanski**



**Florian Pansin**

No obstante, dado que este trabajo se origina en la Universidad de Sevilla, nos gustaría destacar los tres mejores restaurantes del entorno español según la Guía Michelin y la opinión de los comensales, siendo estos:

- 🚦 **ABaC (Barcelona)**
- 🚦 **Akelarre (San Sebastián, Guipúzcoa)**
- 🚦 **Aponiente (El Puerto de Santa María, Cádiz)**

## BIBLIOGRAFÍA, ADJUNTOS Y NOTAS

Se puede encontrar una copia completa del proyecto en [ESTE](#) enlace disponible para usuarios autenticados de la Universidad de Sevilla.

Se han usado como fuentes los datos originales extraídos de Kaggle, la documentación de la API de Yelp, material vario de la asignatura referente a Microsoft Power BI, así como indicaciones provistas por el profesorado.