

## Actividad 1. Business Intelligence a partir de las opiniones de Trip Advisor – Análisis de Texto [1]

Jorge -> Big Ben + Privilege  
Marcos -> Alhambra + Parque Natural  
Iñigo -> Plaza de la Virgen Blanca + Restaurante La Casa del Patrón  
Julen -> Cataratas del Niagara + Palacio de Urgoiti  
Iñigo -> Templo de Bayon + Karlovy Lazne  
Rubén -> Buckingham palace + Hotel Carlton  
Iñigo -> Notredam + Restaurante Begoña  
Unai- Torre Eiffel + La Viña  
Jon -> Coliseo + Mac Donalds de la Fontana de Trevi  
Yeray -> Basílica de la Sagrada Familia + Port Aventura

**Paso 0.** 2 Datasets (.csv) de opiniones de atracciones turísticas/restaurantes/lugares de ocio/ hoteles...

**Portaventura** → PA-OP.csv

**Basílica Sagrada Familia** → SF-OP.csv

**Paso 1.** Hacer una nube de palabras de cada los comentarios de cada uno de los datasets

¿Qué conclusiones podemos sacar?

## PORT AVENTURA



---

## Conclusiones

De esta nube de palabras podemos sacar que los usuarios muestran especial interés por:

- **Atracción:** Al ser un parque de atracciones ,podríamos decir que es un pilar fundamental la opinión de los usuarios acerca de el y por lo general vemos opiniones positivas acerca de los mismos.
- **Espectáculos:** Podríamos decir que es el segundo plato fuerte después de las atracciones y es que Port Aventura ofrece un gran numero de espectáculos



para ya sea acompañarlos junto a una buena dosis de adrenalina o simplemente porque no disfruta de las atracciones y puede gozar de los espectáculos. Por lo general, los espectáculos reciben una valoración positiva.

- **Colas, express, pagar:** Los usuarios por lo general reciben negativamente las altas colas que se producen, sintiéndose obligados a adquirir un pase rápido.
- **Niños:** Un parque de atracciones suele ser un entorno familiar por lo que la adaptación para los niños, los usuarios valoran positivamente la zona de sésamo aventura.
- **Comida:** La comida es otro punto negativo que destacan los usuarios, ya que según su opinión se ofrece una comida cara y de mala calidad.
- **Ferrari:** La gran novedad de Port Aventura es Ferrari Land y ha causado opiniones bastantes contrastadas, por un lado de gente que le gusta mucho la tematización de la marca de coches italiana pero por otra gente que ve un parque pequeño e incompleto.

## BASILICA SAGRADA FAMILIA



---

**Conclusiones:**

- **Obra:** Al llevar los años que lleva construyéndose, la gente tiene cierto interés acerca de la obra de esta basílica.
- **Barcelona:** Lugar donde se sitúa la basílica, una zona alta en turismo.
- **Arte:** Considerada para muchos usuarios como una obra de arte monumental.
- **Tiempo:** Necesidad de planificar con tiempo debido a que las entradas se agotan con facilidad.

**Paso 2.** Añade más columnas de información relevante a tus datasets

¿Por qué añades esta información? ¿Qué conclusiones podrías sacar con ella?

He añadido las siguientes columnas:

- Fecha del comentario
- Título del comentario
- Puntuación del comentario

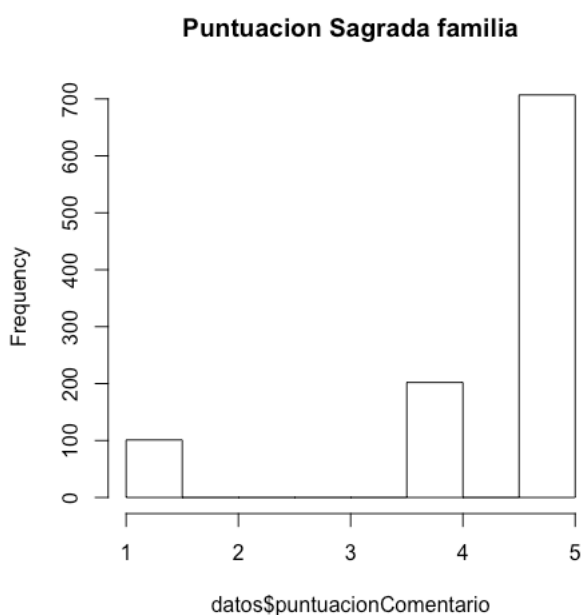
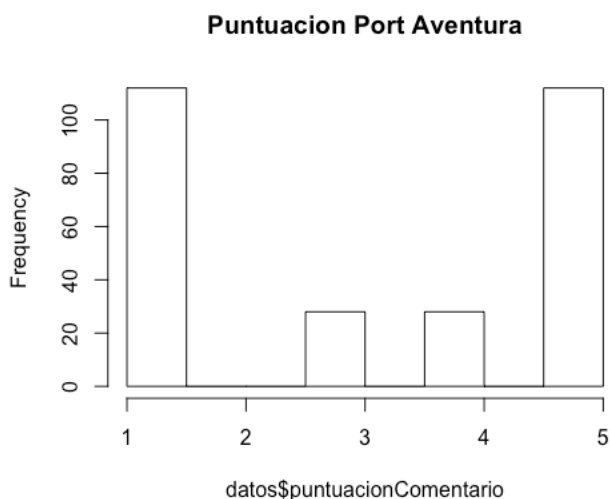
Son 3 parámetros en cuenta a tener a la hora de analizar los comentarios, como por ejemplo en que fechas se nos visita puede indicarnos la orientación de la opinión de un usuario (ej: no había donde taparse en caso de lluvia), el título nos da un breve resumen de si el usuario está contento o no y la puntuación nos permite saber de manera cuantitativa lo que el usuario opina y obtener una media de la opinión de todos los usuarios.

**Paso 3.** Estudia las fechas de los comentarios

¿Qué conclusiones podemos sacar?

Nos permiten ver opiniones según las novedades del último año, por temporadas, por las temáticas de ciertos eventos especiales, si según pasa el tiempo van mejorando o disminuyendo las expectativas... etc

**Paso 4.** Sería posible sacar un gráfico con la puntuación de los visitantes/reviewers por año?



He conseguido solamente obtener el grafico de el numero de personas que ha votado cierta puntuación ,como podemos observar en **Port Aventura**:

Vemos dos polos opuestos claramente, por un lado los grandes fanáticos del parque y por otro aquellos no se han sentido comodoss en el parque ya sea por precios,colas,comidas...

En la **basilica** vemos que una gran mayoría lo valora de una forma muy positiva, solo un pequeño porcentaje lo valora negativamente los cuales aunque sean pocos siempre puede ser interesante leer.

**Paso 5.** Crear datasets independientes dependiendo de las valoraciones (excelente-bueno- malo). Hacer una nube de palabras de cada uno de ellos.

## Port aventura

**Datasets creados:** 1-punt.csv, 3-punt.csv, 4-punt.csv, 5-punt.csv

Malo (1 puntos)



Bueno (3 ,4 puntos)

3 puntos



# Deusto

Facultad de Ingeniería  
Ingeniaritza Fakultatea



4 puntos



Excelente (5 puntos)



Facultad de Ingeniería  
Ingeniaritza Fakultatea



**BASILICA SAGRADA FAMILIA**

MALA (1 PUNTO)



BUENA ( 4 PUNTOS)





Facultad de Ingeniería  
Ingeniaritza Fakultatea



EXCELENTE (5 PUNTOS)

