

DESTINOS TURÍSTICOS PARECIDOS A BCN

Maria Ortuño Serra

Abril 25, 2021

1. Introducción

1.1. Antecedentes

El turismo, tal y como lo conocemos hoy en día, nace en el siglo XIX, como consecuencia de la revolución industrial, con desplazamientos cuya intención principal es el ocio, descanso, cultura, salud, negocios o relaciones familiares.

Desde entonces, se han ido detectando varias modalidades de turismo a causa de las distintas intenciones que las personas persiguen en este viaje. De este modo podemos identificar el turismo de negocios, urbano, natural, gastronómico, de aventura, cultural, deportivo, de sol y playa, o de lujo, entre otros.

Barcelona, por ejemplo, desde la apertura de sus fronteras durante la dictadura franquista, se caracteriza por atraer un gran número de turistas de sol y playa y, en menor medida, turistas gastronómicos que quieren descubrir la dieta mediterránea y sus platos estrella.

Sin embargo, cada ciudad y rincón del mundo tienen un tipo de turismo distinto según sus características.

1.2. Problema

En este proyecto, se desarrolla un análisis comparativo de distintas ciudades del mundo para identificar sus semejanzas y diferencias con el objetivo de determinar que ciudades se parecen a Barcelona en relación a su atractivo turístico.

Algunas de las variables que se tendrán en cuenta son: el número de habitantes, el número de turistas en 2018, la ratio entre estas dos variables, la presencia de playa o costa marítima y los lugares más visitados en cada una de las ciudades.

1.3. Interés

Los principales agentes interesados en los resultados de este proyecto serán las agencias turísticas, las aerolíneas y las aplicaciones de reserva de hoteles y vuelos.

Con esta información, estos agentes tendrán la capacidad de hacer una campaña de marketing más personalizada a sus clientes ya que les propondrá hacer viajes a destinos turísticos parecidos a los que ya hicieron en el pasado. Del mismo modo, si el algoritmo de la aplicación detecta que este cliente busca un viaje completamente distinto (aventura, cultural, etc.) también le podrá recomendar los destinos que cumplan estos requisitos.

2. Importación y limpieza de datos

Primero de todo, se decidieron las 11 ciudades a comparar e investigar. En el proceso, aunque fuese aleatorio se intentó que hubiera destinos de distintos continentes y que hubiera ciudades con climas y dimensiones distintas.

A continuación, se buscaron datos básicos de estos destinos como sus coordenadas (para poder usar Foursquare), el número de habitantes en 2020 y el número de turistas que recibió en 2018.

Figura 1: Base de datos

CITY	COUNTRY	LAT	LON	INHABI- TANTS (M)	TURISTS 2018 (M)	RATIO	BEACH/COAST
BARCELONA	SPAIN	41,39	2,17	1,66	15,80	9,49	YES
MADRID	SPAIN	40,42	- 3,70	3,33	7,10	2,13	NO
CHICAGO	USA	41,85	- 87,65	2,69	58,00	21,58	YES
NUEVA YORK	USA	40,71	- 74,01	8,34	65,20	7,82	YES
SIDNEY	AUSTRALIA	- 33,87	151,21	4,14	16,10	3,89	YES
HELSINKY	FINLAND	60,17	24,94	0,66	1,24	1,89	YES
OSLO	NORWAY	59,91	10,75	0,70	4,80	6,88	YES
PARÍS	FRANCE	48,85	2,35	2,14	40,00	18,70	NO
BERLIN	GERMANY	52,52	13,41	3,77	13,50	3,58	NO
LONDRES	UK	51,51	- 0,09	9,43	20,42	2,17	NO
ROMA	ITALY	41,89	12,51	4,28	15,20	3,55	NO

FUENTE: Creación propia

Se usó el registro turístico de 2018 ya que se considera que los datos obtenidos para los años 2019 y 2020 estarían sesgados por la crisis del Covid-19 y, por lo tanto, no representaría la realidad.

Una vez obtenida toda esta información se inició el proceso de investigación y análisis que consistió en usar la herramienta de Foursquare para explorar cada una de las ciudades y ver cuáles son los lugares más transitados.

3. Metodología

3.1. Obtención de datos – Foursquare

Tal y como se ha comentado anteriormente, para poder explorar cada una de las ciudades seleccionadas se ha utilizado la aplicación Foursquare.

Foursquare es una popular aplicación de geolocalización para social media que invita a los usuarios a dar “Check-ins”, con sus teléfonos celulares, en los negocios y lugares que visitan (Ignacio Santiago, n.d.). Es una aplicación muy interesante ya que permite a los

negocios a crear campañas para fidelizar sus clientes y a ofrecer promociones especiales para ser más competitivo y atraer a los clientes de la competencia.

Además, ser usuario no solo te permite beneficiarte de estas promociones, sino que también te permite explorar distintas ciudades, conocer sus lugares más populares y acceder a las opiniones que otros usuarios han publicado sobre estos sitios.

Después de extraer los datos sobre los lugares a visitar de cada ciudad utilizando el método explore, ordenamos esta información en variables ficticias que representaran cada categoría. Algunas de las categorías que encontramos fueron restaurantes, museos, panaderías, tiendas, jardines, gimnasios, hoteles, etc.

A continuación, se seleccionaron las 10 categorías con más frecuencia de cada ciudad, se ordenaron de forma descendente y se juntó esta información con el marco de datos inicial (df) para poder proceder a hacer el análisis de datos.

3.2. Modelo predictivo

El método utilizado para analizar los datos fue el de los vecinos más cercanos (KNN) que es un modelo de aprendizaje automático supervisado que consiste en clasificar cada dato nuevo (en nuestro caso, cada ciudad) en el grupo que corresponda, según tenga k vecinos más cerca de un grupo o de otro (Merkle, 2020).

Para determinar la cercanía, se utiliza la métrica de similitud. Es decir, al final, el algoritmo agrupará los distintos datos en grupos de ciudades que sean parecidas entre si según las variables que le hayamos dado.

Considero que este modelo es el más apropiado para el objetivo de este proyecto ya que se utiliza para aplicaciones de sistemas de recomendación tales como plataformas de contenido digital, procesos de venta cruzada en eCommerce, etc. (Merkle, 2020).

4. Análisis exploratorio de datos

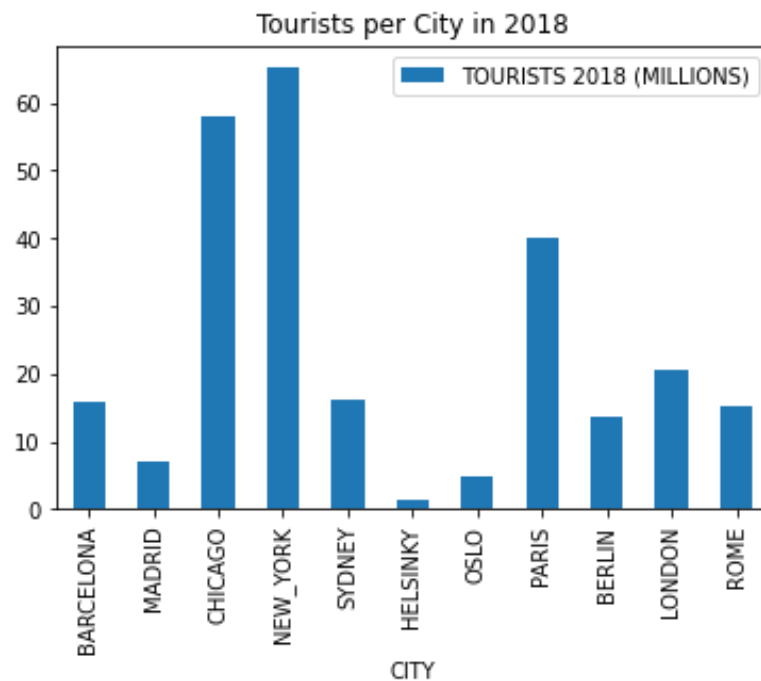
Entre las 11 ciudades analizadas se puede observar que forman parte de distintos continentes como Europa (que tenemos 9 ciudades de 7 países distintos), América (dos ciudades de Estados Unidos) y Oceanía (una ciudad de Australia).

También se puede destacar que el 54,5% de las ciudades se encuentran en la costa, mientras que el 45,5% restante son poblaciones continentales.

Además, por lo que respecta al número de habitantes, los datos demográficos muestran que las ciudades escogidas tienen datos muy variables. Por ejemplo, en el marco de datos utilizado, se observan desde ciudades que no alcanzan el millón de habitantes como Oslo o Helsinki, a otras que superan los 8 millones como Nueva York o Londres.

A nivel de turistas, estas ciudades reciben un nivel de turistas muy variado tal y como se puede ver en el siguiente gráfico:

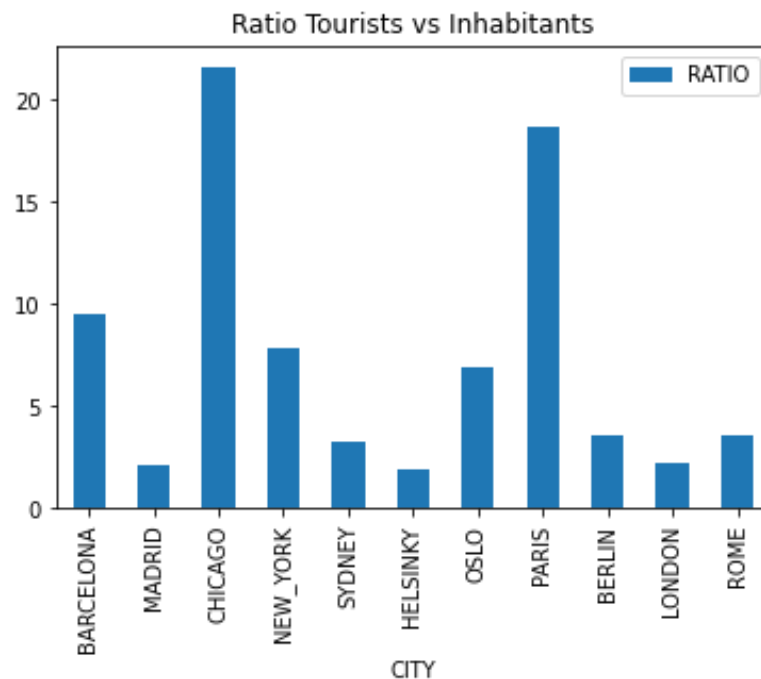
Figura 2: Turistas en 2018



FUENTE: Creación propia con Python

Las ciudades que recibieron más turistas en 2018 fueron Nueva York, Chicago y París. Sin embargo, considero que es importante reparar en el hecho de que hay ciudades que recibieron un porcentaje mayor de turistas respecto a su población.

Figura 3: Ratio Turistas



FUENTE: Creación propia con Python

Gracias a estos datos, se puede observar que, aunque Chicago y París siguen siendo de las ciudades con mayor ratio de turistas/habitantes, Oslo y Barcelona han mejorado considerablemente su posición respecto al total ya que reciben 6,88 y 9,49 veces más de turistas que su población habitual, respectivamente.

5. Resultados

Después de hacer este análisis inicial, y añadirle los datos adquiridos de los resultados de la búsqueda en Foursquare, obtenemos un marco de datos que recoge ambos, los datos demográficos y los datos sobre las 10 categorías principales de los lugares que se encuentran en cada una de las ciudades.

A continuación, se aplicó el modelo de K vecinos más cercanos (KNN) con el objetivo de agrupar las ciudades según su grado de similitud en 4 grupos. Los grupos que el modelo estableció son los siguientes:

Grupo 1: Chicago

Grupo 2: Barcelona, Madrid, Nueva York, Sídney, Helsinki, Oslo, Berlín y Londres.

Grupo 3: Roma

Grupo 4: París

Figura 4: Mapa del Mundo



FUENTE: Creación propia con Python

A diferencia de lo esperado en un inicio, el modelo ha agrupado en un mismo conjunto todas las ciudades a excepción de 3 a pesar de sus distintas características climatológicas, geográficas y culturales.

Si observamos con más detalle la clasificación de las ciudades, podemos ver que entre los 10 lugares más frecuentados de cada grupo se pueden destacar los siguientes lugares:

- Grupo 1: gimnasios, helipuertos, estaciones de tren, pubs y restaurantes mejicanos.
- Grupo 2: hoteles, cafeterías, restaurantes y tiendas.
- Grupo 3: restaurantes, hoteles, heladerías y plazas.
- Grupo 4: restaurantes, cafeterías, plazas y jardines.

Por lo que respecta al caso concreto de Barcelona, y a la futura aplicación de este modelo para hacer recomendaciones a turistas que les gustó esta ciudad, se puede deducir que estos turistas buscaran lugares donde predomine el turismo urbano y gastronómico ya que así pueden conocer la ciudad, su cocina y aprovechar para ir de tiendas.

6. Debate y recomendaciones

A pesar de haber conseguido una división en grupos de las ciudades escogidas, considero que es importante destacar que es muy probable que este modelo, tal y como se ha planteado en este proyecto, esté incompleto para hacer predicciones reales sobre los próximos destinos de los clientes de aerolíneas o agencias de viajes.

Por este motivo, propongo para futuros estudios ampliar el número de variables que se tengan en cuenta como el clima, la cultura y las actividades imprescindibles para un buen viajero, entre otras.

Además, también será necesario recopilar datos a partir de encuestas o historiales para poder obtener un conjunto de entrenamiento y conjunto de prueba para mejorar la eficiencia y eficacia del modelo.

7. Bibliografía:

MERKLE (septiembre 1, 2020). El algoritmo K-NN y su importancia en el modelado de datos. Recuperado de: <https://www.merkleinc.com/es/es/blog/algoritmo-knn-modelado-datos>

IGNACIO SANTIAGO (n.d.) FOURSQUARE, Qué es, Cómo Funciona y Para Qué Sirve. Recuperado de: <https://ignaciosantiago.com/foursquare-que-es-como-funciona/>