

PRÁCTICA BICIMAD

OBJETIVO Y OBTENCIÓN DE DATOS

En esta práctica sobre BiciMad nos centramos en el estudio del flujo de las bicicletas públicas en Madrid durante los meses de verano, primavera y Navidad, además, lo comparamos con el flujo global de un año entero. Dicho estudio consistirá en ordenar las estaciones según su exceso de bicicletas para posteriormente realizar un trasvase de ellas. Trabajamos con datos de la web de BiciMad que hemos obtenido desde [https://opendata.emtmadrid.es/Datos-estaticos/Datos-generales-\(1\)](https://opendata.emtmadrid.es/Datos-estaticos/Datos-generales-(1)).

PROCEDIMIENTO E IMPLEMENTACIÓN

Para ello, unimos los rdd de cada apartado que queremos estudiar en uno solo, por ejemplo, si queremos estudiar los meses de verano, unimos en un mismo rdd los datos de junio, julio y agosto del 2017, año con el cual trabajamos (abril de 2017 a marzo de 2018). Posteriormente, obtenemos los datos que nos interesan del rdd total a través de la función mapper. Dicha función consiste en sacar los datos del usuario, tanto su código como su tipo, y las estaciones donde recoge y aparca la bicicleta en un trayecto. Nos centramos en los usuarios con tipo 1, ya que los otros dos tipos son especiales, por ejemplo, mecánicos de BiciMad. Creamos una lista con los rdd que nos aportan información sobre el número de bicicletas recogidas o aparcadas en una estación en concreto, y unimos esa lista en un único rdd. Después, agrupamos el rdd por estaciones, y como resultado obtenemos el exceso o déficit en una determinada estación. Ordenamos los resultados en función del exceso de bicis y tenemos las estaciones ordenadas de menor a mayor número de bicicletas durante los meses evaluados.

INSTRUCCIÓN DE EJECUCIÓN

Para ejecutar el programa habrá que introducir en la función `ordenar_bicis` el rdd (unión de meses) que queramos estudiar, y tras el proceso se imprimirá la lista de estaciones correspondiente. En el archivo .py aparecen todas las temporadas, si queremos probar solo una (introduciendo los archivos correspondientes en FILES), es importante que comentemos el resto de líneas, pues si no están sus archivos darán error.

RESULTADOS Y CONCLUSIONES

Cómo hemos dicho vamos a centrarnos en las estaciones con mayor déficit o exceso de bicis al final de una temporada concreta. Como hay más de 200 estaciones, (el programa muestra los datos sobre todas las estaciones disponibles en cada periodo de tiempo) presentamos los resultados obtenidos para las 10 estaciones con mayor pérdida de bicis y las 10 con mayor ganancia, siendo estas las más representativas de cara a nuestro análisis:

TEMPORADA DE VERANO: meses de junio, julio y agosto

[(99, -1082), (1, -965), (2, -950), (102, -948), (154, -947), (23, -840), (6, -816), (109, -815), (158, -775), (28, -775), [...], (136, 756), (83, 758), (175, 821), (43, 1449), (133, 1531), (168, 1627), (82, 1936), (129, 2347), (132, 2412), (135, 3681)]

Se ve claramente que en verano hay una gran cantidad de usuarios que parten de la estación 99, siendo la que más déficit sufre al final de la temporada, seguida por las estaciones 1, 2, 102 y 154 que se mantienen cerca, en torno a -950. Por otro lado, llama la atención el exceso acumulado en la 135, triplicando el valor de bicis, 3681, en comparación al déficit de 1082 de la 99. La siguiente estación con exceso de bicis, la 132, disminuye de forma notable a 2412.

TEMPORADA DE PRIMAVERA: meses de marzo, abril y mayo

[(99, -901), (102, -856), (109, -774), (111, -767), (112, -690), (154, -670), (2, -656), (158, -643), (24, -573), (1, -568), [...], (134, 522), (83, 696), (136, 764), (168, 1022), (133, 1331), (82, 1589), (43, 1624), (129, 1773), (132, 2131), (135, 2926)]

Es necesario destacar que las estaciones con mayor y menor número de bicis coincide con las del mes de verano, aunque en este caso la estación 99 tiene menos déficit y la 135 menos exceso. Al igual que en verano las estaciones más usadas como llegada son la 135, 132 y 129, y además las 10 estaciones con mayor exceso de verano y primavera coinciden, aunque no todas en el mismo orden, y en primavera presentan menor llegada de bicis. Esto no ocurre con las 10 de mayor déficit, aunque también se aprecia que las pérdidas son menores, lo que nos lleva a concluir que en primavera se usan menos bicis que en verano.

TEMPORADA DE NAVIDAD: meses de diciembre y enero

[(154, -557), (109, -486), (102, -465), (1, -386), (2, -370), (99, -362), (140, -356), (101, -339), (153, -302), (161, -299), [...], (168, 383), (128, 458), (136, 474), (133, 485), (43, 502), (175, 582), (82, 949), (129, 1088), (132, 1344), (135, 1849)]

Estos datos son prueba de la disminución del uso de la bici en esta temporada, pues el mayor déficit es de 557 bicis, siendo prácticamente la mitad que en los casos anteriores (i.e. menos gente coge bicis); y el mayor exceso de 1849, es decir llegan unas 1000 bicis menos a la estación 135, que es con diferencia la más utilizada como destino en todas las épocas del año.

BALANCE ANUAL

[(154, -3614), (99, -3414), (109, -3330), (102, -3187), (1, -3112), (2, -3047), (111, -2612), (23, -2489), (24, -2455), (158, -2438), [...], (175, 2863), (83, 2926), (136, 3289), (168, 4397), (43, 5316), (133, 5321), (82, 6639), (129, 8128), (132, 9239), (135, 13431)]

En concordancia con lo anterior, las estaciones que más déficit presentan son la 154 (como en Navidad) y la 99 (primavera y verano). Aun así, no hay comparación con la cantidad de bicis de más que llegan a la estación 135, común a lo largo de todo el año como la estación de llegada más concurrida. Solo con las 13 431 bicis que llegan a esta estación se podrían reponer 4 de las estaciones con mayores pérdidas.