

## 1. Respuesta

Las estaciones con mayor arribo a una hora determinada son las 267 y 266 (están una al lado de la otra) entre 18 y 20 horas. Están cerca del Metrobús Buenavista. Me imagino que una gran cantidad de gente que utiliza el tren suburbano hacia el norte de la zona metropolitana llega a esas horas después del trabajo. En la mañana las estaciones con más arribos son la 16,18 y 23. Las tres se encuentran en Reforma, a distancias cortas del Ángel de la Independencia, y la hora de arribo es entre las 8 y 9. Es una densa zona de oficinas y corporativos, y la hora de arribo es típica de horarios de oficinista.

Por otro lado, todos los retiros muestran una gran acumulación entre las 0:00 y 0:59, es decir, pasando medianoche. Es de notar que es la última hora a la que se permiten retiros, así que esto acumula muchos. Las estaciones con más retiros son la 18,271 y 27. Por la 271 me imagino que hay gente que llega de trabajar en el Estado de México a esa hora y las utiliza para aproximarse a su casa. La 18 y la 27 se encuentran sobre Reforma, ambas relativamente cerca de la Glorieta de Insurgentes. Es posible que gente que trabaje o busque esparcirse en la Zona Rosa y aledaños las use frecuentemente.

## 2. Respuesta

Son relativamente pocas las estaciones que muestran un alza haciendo un ajuste lineal sobre estos últimos tres meses. En retiros, 27 estaciones muestran un coeficiente de ajuste lineal positivo, aunque la mayoría resultan casi planos. En arribos, son 23 solamente. De ellas, las que tienen la tendencia más clara en ambos rubros son la 61, 176, 365, y 273. Como nota: la estación 365 (Mixcoac) parece no funcionar muy seguido realmente. La ilustración está en la figura 1.

## 3. Respuesta

Lo correcto hubiera sido hacer una correlación con un función de convolución con un retraso temporal, de unos 20 minutos, para obtener algo más realista. Desgraciadamente, he perdido mucho tiempo aprendiendo a ordenar datos con fecha. Pero en el espíritu de obtener buenas aproximaciones rápidamente, usaremos la correlación de Pearson entre las Retiros y los Arribos acumulados por hora por cada estación. Aunque esto dará un poco de correlaciones elevadas, nos servirá para tener una buena aproximación en intervalos razonables, ya que el tiempo máximo de uso permitido de una unidad es de 45 minutos. De todas maneras, esto nos permite agrupar ya las estaciones con posible alta correlación y posteriormente hacer un estudio más puntal. Véase la figura 2.

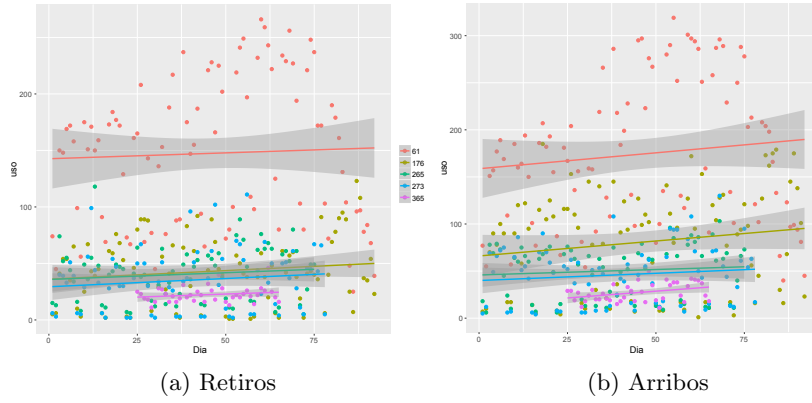


Figura 1: Uso por retiro y arribo en las cinco estaciones con tendencia ligeramente al alza. Se muestra un ajuste lineal por mínimos cuadrados con intervalos de confianza. Nótese que la estación 61 muestra una variación muy grande en sus datos. Los datos de final de año puede que estén causando un descenso espurio en la tendencia.

## 4. Respuesta

Honestamente, no se a que se refieren con método de aprendizaje no supervisado. Yo lo que hubiera hecho (con más tiempo) es implementar alguna forma de análisis de componentes independientes, para ver si las distintas estaciones se agrupaban en grupos con comportamiento independiente (probablemente por horarios no correlacionados dentro del intervalo de uso posible), y con eso, usando datos geográficos y de horarios laborales y lúdicos, descubrir si los grupos de estaciones se agrupan por tipo de uso.

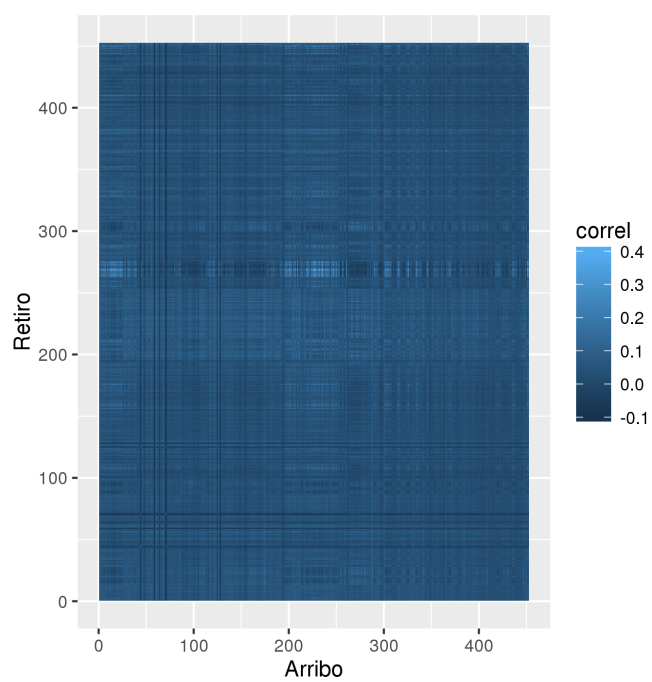


Figura 2: Mapa de color de la correlación directa de retiros y arribos por estación, datos acumulados por hora.