Distribuidora del Sur, S.A.

El dataset que se nos fue proporcionado de Distribuidora del Sur, S.A. nos dan 2180 observaciones con 11 variables distintas. Entre las variables tenemos COD\_VIAJE, CLIENTE, UBICACIÓN, CANTIDAD, PILOTO, Q, CREDITO, UNIDAD, MES y ANIO.

El COD\_VIAJE es un código único que identifica cada una de las 2180 observaciones. En CLIENTE tenemos varias informaciones juntas: tenemos el nombre del cliente y en algunos casos también hay un comentario diciendo si en la entrega hubo un faltante, o si el viaje fue hecho por una devolución, o si fue por un simple despacho al cliente. Se decidió separar toda esta información que nos daba la variable CLIENTE en varias columnas. La primera fue una columna FALTANTE que tiene un 1 si en esa entrega hubo un faltante, y un 0 si no lo hubo. La segunda fue la columna DEVOLUCION que de igual modo tiene un 1 si el viaje fue por una devolución y un 0 si no lo fue. Otra llamada CLIENT que contiene únicamente el nombre del cliente. Por último, una llamada COMENTARIO, en la que clasifica si cada viaje fue por un despacho, un faltante o por devolución.

UBICACIÓN tiene codificado el código del lugar de donde salió el despacho. CANTIDAD nos dice la cantidad de producto que se entregó en cada viaje. PILOTO nos dice el nombre y apellido de quien manejó y entregó el pedido. Q es el monto en Quetzales que pagó el cliente por esa entrega. CREDITO es la cantidad de días que se le permite al cliente pagar posterior a la entrega del producto. UNIDAD es el tipo de vehículo con el cual se hizo la entrega. Cada tipo de vehículo tiene diferente tamaño y, por lo tanto, diferente límite de productos que puede transportar. MES nos dice el número del mes en el que se hizo la entrega, y por último ANIO nos da el año en el que se hizo la entrega.