Instacart Operations Research Challenge

Data Sets

deliveries.csv

stores.csv

Problem

We have a number of deliveries to be delivered at different locations, each due at a certain time (due_at).

There are five stores. A delivery can be shopped at any of these.

We have 10 shoppers. A shopper will drive to a particular store, shop for the items, and drive to the delivery location(s) to deliver - this is a trip. A shopper can work on up to 3 deliveries during each visit to a store. Once a shopper is finished with the current set of deliveries (trip), he or she can start another trip.

All deliveries in a trip should be shopped at the same store.

A trip starts when a shopper starts to drive towards the store and ends after delivering all deliveries (maximum of 3) in the trip.

The **goal** is to minimize the sum of absolute errors of delivery times: sum(abs(delivered_at - due_time)).

No delivery should be left out.

Additional Info

The formula for shopping time is (5 + number of items) minutes.

The formula for driving time is (straight line distance in miles * 5) minutes.

Assume all shoppers are ready to start at time *2014-03-13 15:00:00* and are located at store_id = 49.

Performance is important as this needs to be run very frequently.

Finding the optimal solution to this problem may not be practical.

Use **Python, R or Julia** and any open source librairies you'd like.

At the end, we'd like:

The plan output as a CSV in the following format **(you generate trip_id, shopper_id)**

```
delivery_id,trip_id,shopper_id,trip_started_at,trip_ended_at,store_id,delivered_a
```

```
240254,1,1,2015-11-12 15:14:09,2015-11-12 17:14:09,60,2015-11-12 16:14:09
240369,1,1,2015-11-12 15:14:09,2015-11-12 17:14:09,60,2015-11-12 17:14:09
...
# note that trip_id, shopper_id, trip_started_at, trip_ended_at, store_id fields
# will be the same for a given trip_id
```

- Sum of absolute errors of delivery times achieved
- An explanation of the solution
- The code

Please ask if you have any questions. Have fun, and good luck!