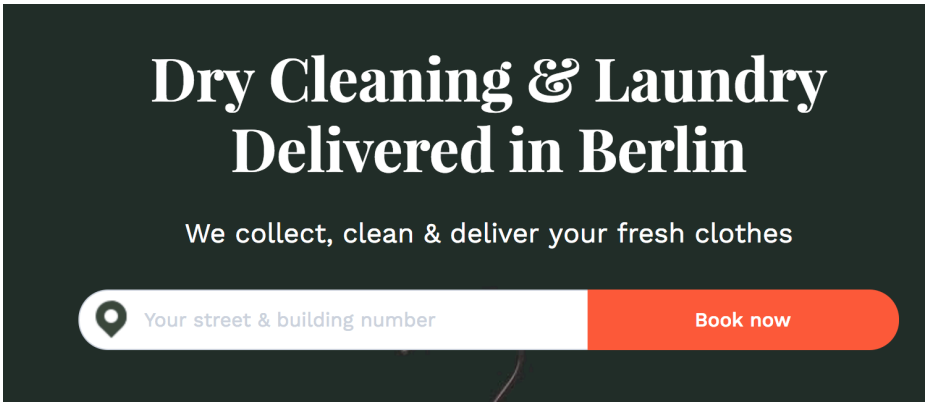
Goal

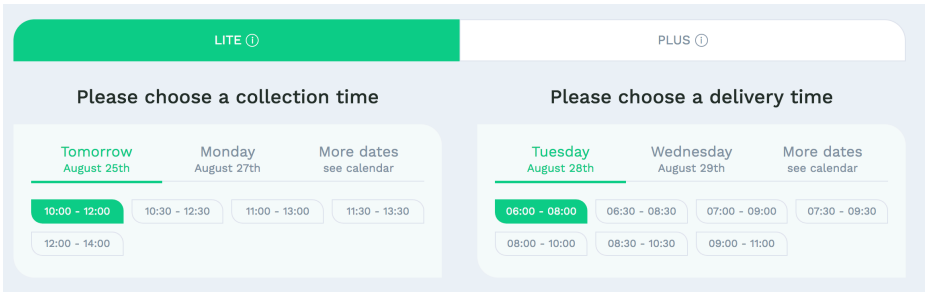
In Zipjet we aim to make our service accessible to everyone just in time customers need it. For this it is extremely important to know behavior of our customers, predict demand in advance and provide supply when and where necessary.

In this test case you are asked to forecast demand using historical demand data, let's call it **D**. Obviously, historical values of **D** alone don't provide a true picture of demand as we could loose orders when not providing supply. Therefore, as a features you will use supply data, **S,** to be precise, sequence of supply development.

Description and Definitions

To understand the problem from the customer's view point, please go to our website or application and play with placing an order. You will see that you are asked to provide your location and then based on this and service class (you can read about this on website/app too) you are offered collection and delivery date and time slots like shown in figure below:



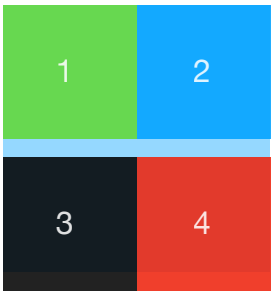


These available time slots are calculated by our in-house routing tool for every lat/long L_j and time slot t_i in any time t_{back} before slot starts. The problem arises when time provided is no way suitable for you and then you choose to use laundry in the neighborhood.

For this test case you are provided with demand and supply data in a format described below.

Demand, D is number of pickup (collection) tasks completed on the scale of day, shift and city. Shift here is an operational term meaning that operational working day comprises two shifts: morning shift **MS** (until 11 am) and evening shift **ES** (after 4 pm). You will find it in file **"demand.csv"**.

Supply, S is ratio of area available for every desired time slot t_i in any time t_{back} before slot starts. This is calculated for every unit of grid which covers operating city. For this test case the city is covered by 31 units grid. To make things more clear, consider below a toy example, where city (enclosing square) is divided into 4 units grid.



Let's consider unit 4 for which supply sequence data is shown below:

	day	timeslot_from	calculated_datetime	grid_unit	avail_area_ratio
41	2018-01-08	07:00	2018-01-05 14:42:53	4	1.00
42	2018-01-08	07:00	2018-01-05 14:54:25	4	0.94
43	2018-01-08	07:00	2018-01-05 15:59:54	4	0.00
44	2018-01-08	07:00	2018-01-05 16:02:08	4	0.37
45	2018-01-08	07:00	2018-01-05 16:29:22	4	0.56
46	2018-01-08	07:00	2018-01-05 16:32:34	4	0.62
47	2018-01-08	07:00	2018-01-05 17:54:24	4	0.00

<*day*> and <*timeslot_from*> here refer to the date and time of the slot start time (t_i) for which supply is calculated at <*calculated_datetime*> (t_{back}). From this example you can see that in the beginning 100% of area was available for grid unit 4, then few hours later same day, at 17:54, it was not available at all, 0%.

Problem Statement

Using provided supply and demand data:

1. model demand using supply, **D(S), don't change format of demand,** whereas for supply you need to engineer features.
2. make demand forecast for 7 days horizon starting from 2018-07-10 given that we provide sufficient supply for this day.

Delivery Format

You are free to use Python/R to solve the test case.

Please, deliver results as a notebook and attach working code for all steps of your work.