**Dataset instruction**

The primary dataset comes from a start-up company where one of our team member is currently working. The data recorders are extracted from a time card system. The system is listed in the company’s SharePoint site and it requests employees to input their working hours on daily basis. The data has 3,834 records and is dated from April 1st 2015 to September 26th 2015.

Some of the business specific confidential information has been transformed or modified per the company’s requests and has been verified and approved by the company for our project.

**Field Information**

* Task Name:

The name of the task that employees should work on within a period of time. Predefined and pre-assigned to employees.

* Billable:

Billable type of time recorded. Predefined, can be any one of the four types – Billable, Non-billable, Strategic, Biz Dev

* Date Entry:

The time stamp of each record, systematically saved from SharePoint

* Time Entry:

Hours spent on a specific task/task group. Entered by employees. Number format with 2 decimal places

* Level:

Task level. Predefined value and drop down list in SharePoint for employee to choose.

* Task Notes:

Free text box for employees to input the detail of the time cards.

* T-Client (for filtering):

Client Name. Values changed in data set to protect confidentiality.

* T-Created by:

Employee’s’ name. Values changed in data set to protect confidentiality.

Supporting the above set, we have calculated few more columns in the tool, which facilitate in creating visualizations.

The second dataset is the working schedule of Company U; it is an Excel data table which contains information about the calendar date, Month, Quarter, Year, day of the week, working hours, work day type (business day/holiday/weekend).

This secondary dataset provides the working hour baseline and weekday information for the first dataset. The relationship has been established in Power BI using columns [T-date] (transformed from date, simple date type; as we had to remove the time stamp from [date entry] field) and [date] from the secondary dataset. The snapshot below shows the relationship in Power BI. 