**Project proposal**

**Background**

Oure goal of this project is Advertising new product for biscuit company that produced new biscuit and must to be real tested from people so we can from MTA dataset know the entries will accrue in which time in the day and which day of the week, and in before that time must most of people are waiting and distribute company products for free for people

**Question:**

- How many people we can offer our product?

-How can improve our product after looking at selling chart at this month we distribute our product?

**Data Description:**

**MTA dataset**

From this dataset of subway of New York we can now number of entries and exits by station

The dataset from 3 months ago has a lot of rows and 11 columns

Columns are

C/A: control area (A002)

UNIT: Remote Unit for a station (R051)

SCP: Subunit Channel Position represents a specific address for a device

(002-00-00)

Station: Represents the station name the device is located at

Line Name: Represents the line originally the station belonged to BMT, IRT or IND

Date: Represents the Date(MM-DD-YY)

Time: Represents the time (hh:mm:sss) for a scheduled audit event>

DESC: represent the "REGULAR" scheduled audit event (Normally occurs every 4 hours)

Entries : the cumulative entry register value for a device

Exits: The cumulative exit register value for a device

The dataset is available on the following website :

<http://web.mta.info/developers/turnstile.html>

**Tools:**

The technical used

The dataset available as text file

The programming language is Python

We can import data from python language as text file

SQL

We can use sql in python to manipulate and get data.