



EDA MTA

29.09.2021

—

Abdulrahman amer alshehri

T5

Problem Statement

The Sweets Factory is a rapidly expanding chain of stores and currently has two locations in New York City, and they plan to open one or two more locations in new neighborhoods over the next year, we'll focus on other municipalities for placement recommendations.

Goals

One of the goals of data modeling is to create the most efficient way to store information while still providing complete access and reporting. Data modeling is a critical skill for every data scientist, whether you are designing a research or designing a new data warehouse for your business.

Data Description

Field Name	Description
C/A	Control Area (A002)
UNIT	Remote Unit for a station (R051)
SCP	Subunit Channel Position represents an specific address for a device (02-00-00)
STATION	Represents the station name the device is located at
LINENAME	Represents all train lines that can be boarded at this station
DIVISION	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:ss) 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 Metropolitan Transportation Authority is North America's largest transportation network, serving a population of 15.3 million people across a 5,000-square-mile travel area surrounding New York City through Long Island, southeastern New York State, and Connecticut.

Tools

1. Python
2. Panda
3. Matplotlib
4. SQL