from: **Joe Jones** <jjones@muny.mta.gov>

to: Data Dudes <info@datadudes.com>

date: Sun, Jan 11, 2015 at 1:23 PM

subject: Checking In

David, Nelson, and Michael --

It was great to meet with you and chat at the event where we recently met and had a nice talk. We'd love to take some next steps to see if working together is something that would make sense for both parties.

As we probably remember, we work for Music Under New York (MUNY). MUNY is a program of the Metropolitan Transportation Authority that schedules musical performances in the New York City Subway. We are interested in harnessing the power of data and analytics to optimize the scheduling and placement of our performers.

We currently have about 30 stations where we host performances, but we are open to adding (or removing) stations and maybe even setting up performances in parks or other areas near subway entrances/exits.

Where we'd like to solicit your engagement is to use MTA subway data, which as I'm sure you know is available freely from the city, to help us optimize the placement and timing of our performances. We would also be interested in optimizing performance types (e.g. music vs. dance) at locations, and maybe even choosing these performance types based on the likely demographics at each location (e.g. adults vs. families)?

The ball is in your court now do you think this is something that would be feasible for your group? From there we can explore what kind of an engagement would make sense for all of us.

Best,

Joe and Linda, MUNY

Problem Statement

MUNY is looking for advice on optimizing the placement and timing of our performances throughout the MTA. The key goals for MUNY are to increase performance revenues and audience visibility. SInce the number of performances they can schedule is limited they would like to know the best locations within the New York Subway system to focus their efforts.

Preliminary Analysis of MTA Data

From preliminary analysis of MTA data, we are able to determine passenger volume by:

- Stations
- Times of day
- Days of week
- Seasons (and tourists!)

Deeper Analysis of MTA Data

Interaction effects: We know that there are interactions effects between the variable aboves. For example: some stations have large difference between weekday and weekend traffic (i.e. Wall St), while the difference is not as pronounced for others (i.e. 15th St - Prospect Park).

Diminishing returns on station expansion: We can also see from the data that a relatively small number of stations receive a relatively large proportion of traffic. For example, the 22 highest volume stations capture 25% of the total traffic, and the remaining ~400 stations capture the remaining 75%. This insight can tell us that focusing on the highest volume stations will increase exposure quickly, but there are diminishing returns to expanding into additional (and less-traveled) stations.

What we can do with more data and analysis

Between gathering data from other sources, researching “willingness to give”, and collecting data on current MTA performances, we could also gain insights into:

- Choosing the best venue for talent based on:
    - Station dimensions (e.g. bigger stations for dance groups)
    - Station demographics (e.g. placing magicians or jugglers near stations with more families)
- Determining the highest grossing times of day on both weekdays, weekends, and holidays
- Pursuing additional revenue streams via promos/ads, audience incentives/engagement
- Increasing brand recognition via advertising