|  |
| --- |
| [SADAIA] |
| Proposal of data science project |
| Traffic stops at subways in NYC during Rian |

|  |
| --- |
| Amooll alshahrani  ‏18‏/2‏/1443 |

**PRPOPSAL :**

## **Data selection**

**The problem:** there was a big subway congestion on the stop when it was raining**.**

**The tasks:** visualize the traffic stops during unexpected weather .

**The added value to the organization:** (general vision**)**

**-**empower the data science in dealing subway issues

-focusing on the stops that more exposure to traffic than others **.**

**-** help decision makers to take Precaution procedure before the occurrence of rain/snow. etc.

The essential challenge: talking permeative procedures to avoid the traffic in unexpected weather condition at subway stations. The data science will visualize the traffic time -during unexpected weather -based on number of entries and exist. With the help of weather data.

**Data set:**

This data consisted of entry and exit counters for all NYC subway station turnstiles, captured after every 4 hours.In addition, structured data of weatheras we needed, we were able to perform thorough data cleaning to make it useful for our analysis.

**I obtained the data from** [MTA Turnstile Data](http://web.mta.info/developers/turnstile.html)

And [National Weather Service : Observed Weather for past 3 Days : New York City,](https://w1.weather.gov/data/obhistory/KNYC.html) **.**

We used this dataset to analyze the impact of weather on the traffic.

Data analysis:

* Prepared MTA turnstile data from **January to April 2017.**
* Cleaning the data by **renaming columns**, **adding calculated fields**, **removing erroneous data**and **aggregating by station and date.**
* Found that **weekdays had the highest traffic**