

Data Analysis Plan

Air Quality Citizen Science Project

BCCHE



Area of Study

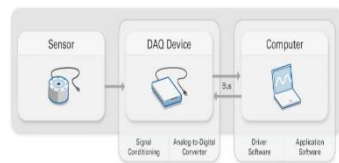
Williamsburg – 11 locations

The Bronx – 13 locations

Objectives

- Stationary Network and Personal Monitoring to measure ambient air concentrations of $PM_{2.5}$ and personal exposure of $PM_{2.5}$
- Study does increase spatial and temporal resolution of data pattern
- Cover 'hot spot' areas like major highways and heavy industrial and commercial area

Step
1



Data Acquisition

Step
2



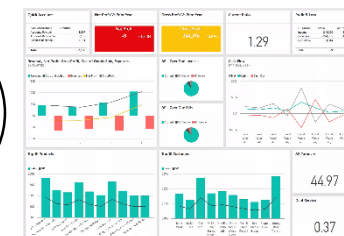
Data Preprocessing
and Data QAQC

Step
3



Data Analysis

Step
4



Data Visualization

Sources and Data Frequency

- Stationary Network – High Precision, High Cost
 - DEC Sites – FEM, FRM (Hourly)
 - NYCCAS Sites (2 weeks)
 - RT Sites (every 15 mins)
- Stationary Network and Personal Monitoring – Low cost
 - Airbeam2 (every minute)

Database Setup

PostgreSQL

Data (.xlsx and .csv files)

- Missing Data Analysis
- Filtering Noisy Data
- Outlier Analysis
- Data Cleaning
- Data Normalization
- Data Integration

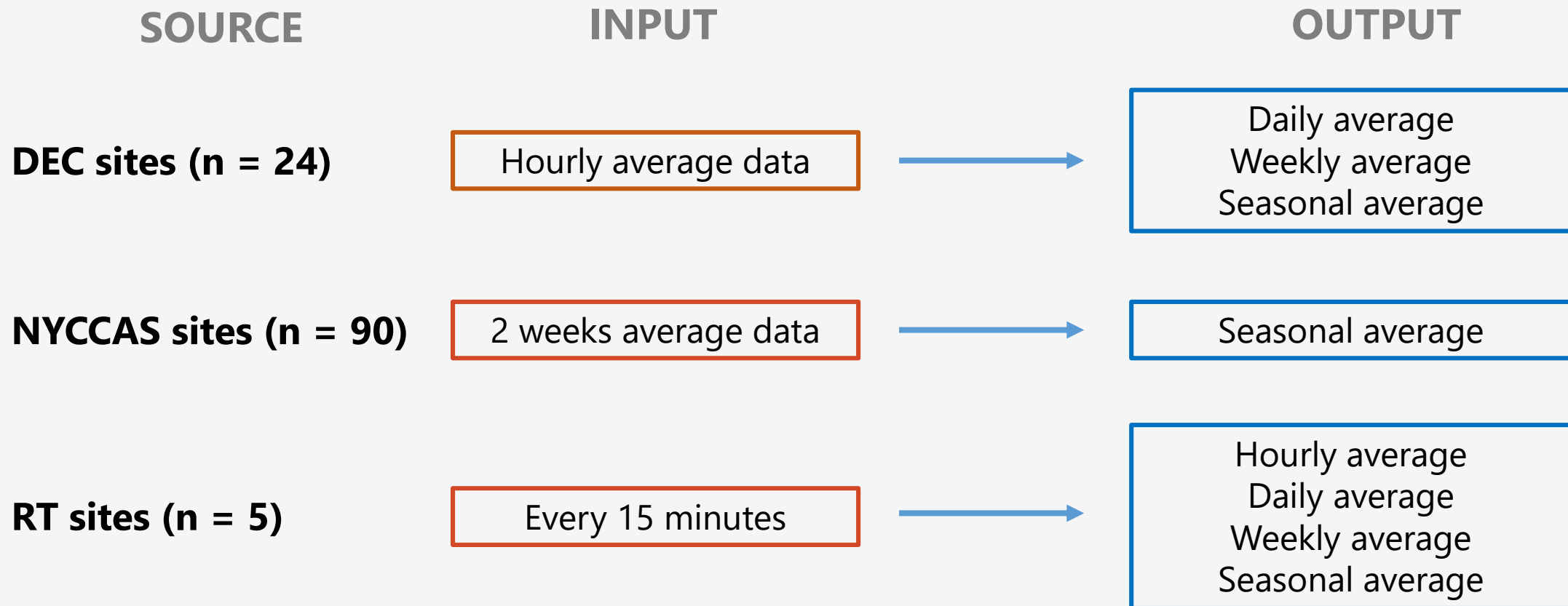
Exploratory Data Analysis (EDA)

- Descriptive Statistics (Time based estimates of mean, standard deviation)
- Correlation Mapping (Between AirBeam2 and NYCCAS, DEC, RT Units)
- Time Series Based Analysis
 - Trend
 - Seasonality
 - Noise
- Regression Analysis
- Plots – Scatter, Pie Charts, Histograms and Bar Charts

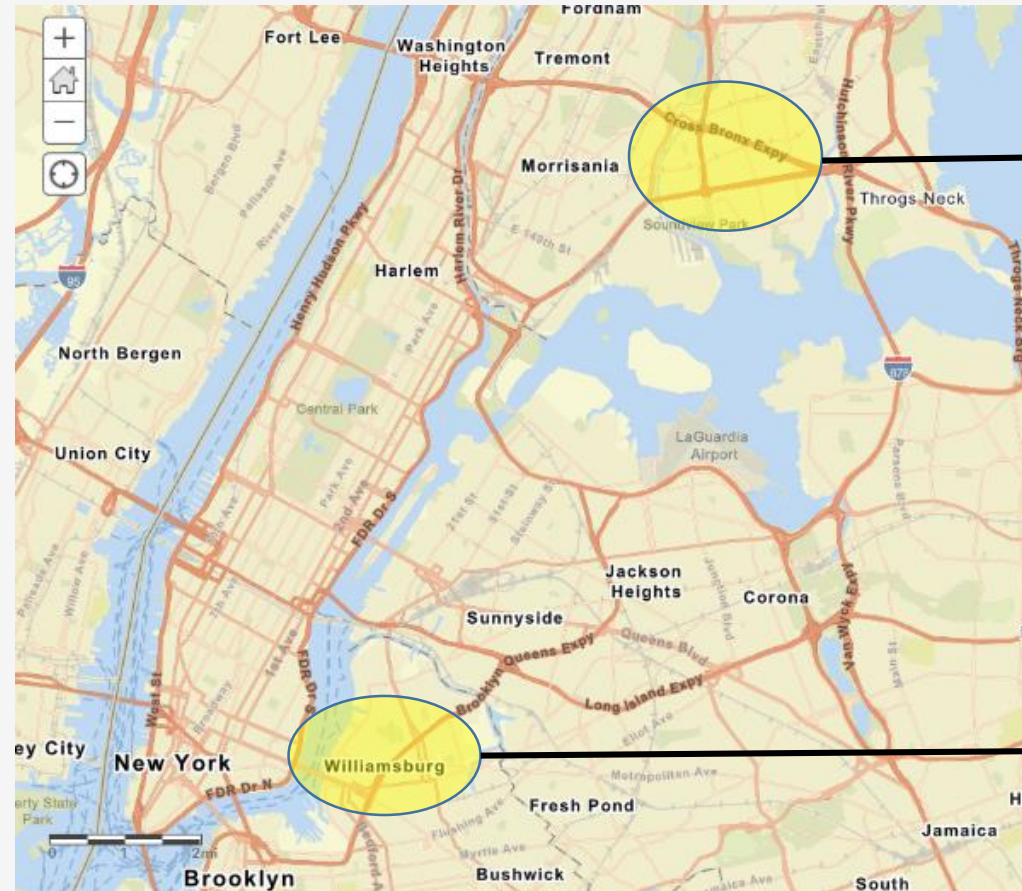
Visualization and Results

- A visual overview at the air quality over time (NYCCAS, DEC)
- Identify PM patterns in data (daily, weekly, monthly, weather patterns)
- Communication of results through visual tools and dashboard, using Tableau (work in progress)
- Determine best methods to present insights and conclusions
- Provide recommendations

What do we know about Air Quality in New York City?



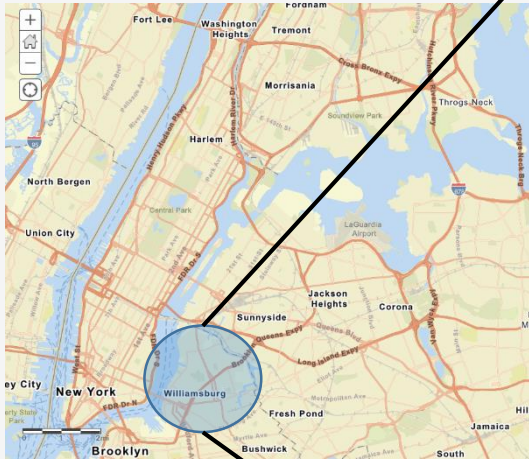
Which neighborhood is of our interest?



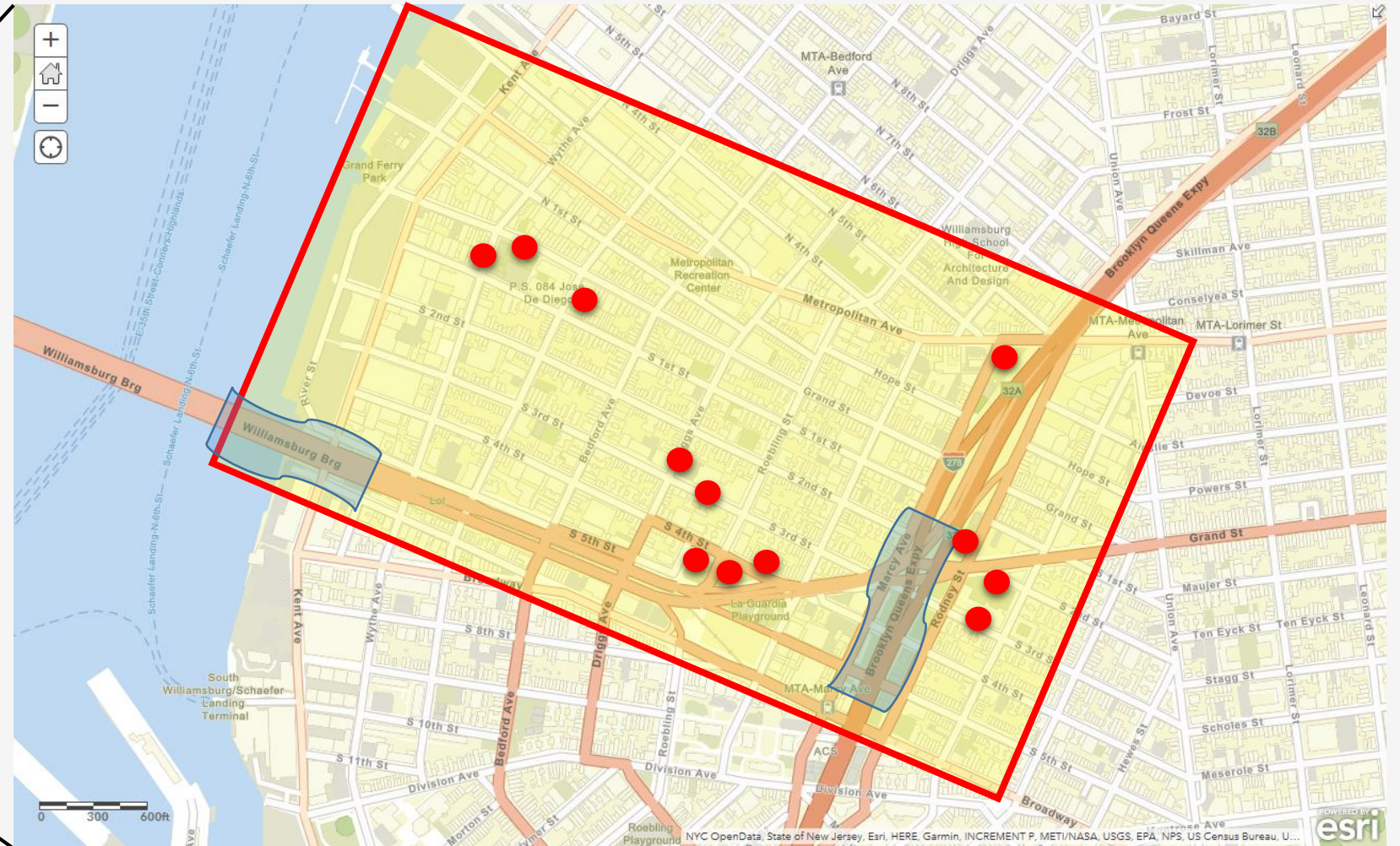
The Bronx

Williamsburg

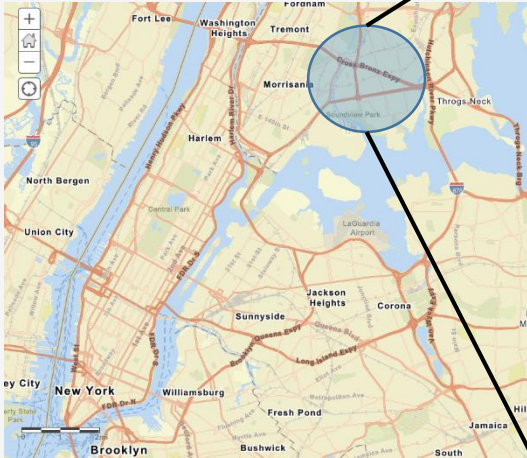
Williamsburg – Locations of Interest



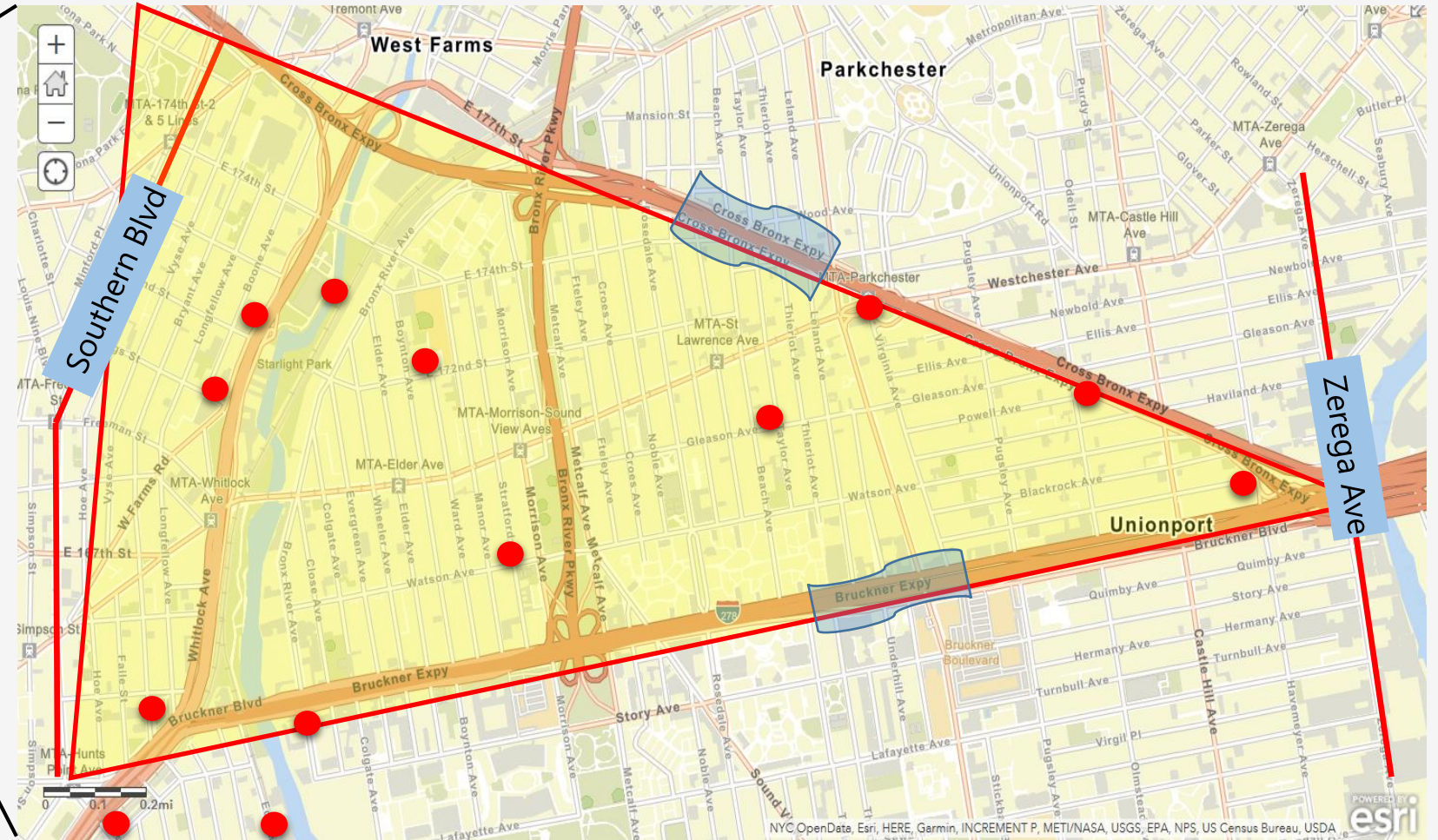
● 11 locations



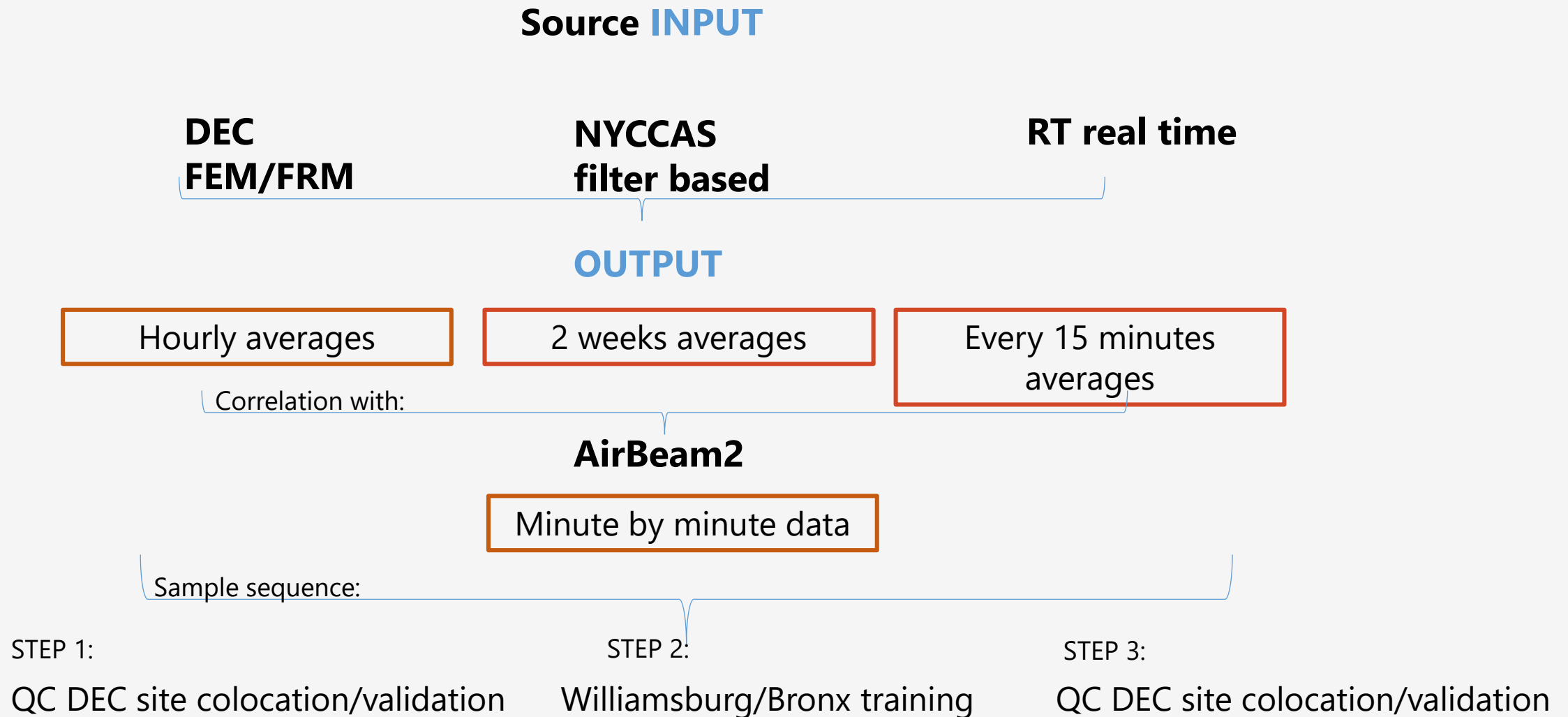
The Bronx – Locations of Interest



● 13 locations



QA/QC Validation based on data collected at the QC DEC site – Training



AirBeam2 in Project areas

INPUT

Minute by minute data

AirBeam2
Stationary Network

OUTPUT

Hourly average
Daily average
Weekly average
Seasonal average

PERFORMANCE, R^2

AirBeam2

FEM, FRM, TEOM

NYCCAS

RT

Temperature

Humidity

Pressure

COLOCATION, R^2

AirBeam2

Research Question

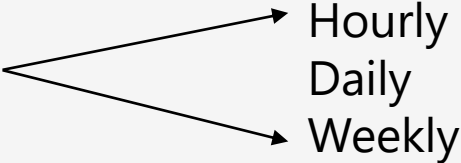
- 1. What are the results of the QA/QC assessment of the AirBeam instruments compared to FEM/FRM and NYCCAS instruments (before and after field campaigns)?**
- 2. How do the AirBeam data compare to the pdR data in the Project area (El Puente only)?**
- 3. How do the AirBeam data pattern compare to the closest DEC monitoring sites?**
- 4. How do the AirBeam data compare to the average citywide concentrations (all DEC sites)?**
- 5. How do the AirBeam data (seasonal averages) compare to the NYCCAS LUR, pdR and DEC data sets?**
- 6. What are the diurnal/weekly/weekdays vs weekends/monthly/seasonal patterns in the Project areas? (based on mean PM2.5 concentrations)**
- 7. Based on AirBeam Stationary Network Data can additional hotspots be identified?**
- 8. How do the data pattern in Williamsburg compare to the data pattern in the Bronx?**
- 9. How useful are the collected personal monitoring data for characterizing exposure pattern and how do the findings compare to the stationary network?**
- 10. Do the data pattern provide answers to the study objectives for the two project areas?**

Data Analysis

- Statistics Summary – To be obtained for each site
 - Daily average
 - Weekly average
 - Monthly average
 - Seasonal average
- Correlation between all sites (PM), R^2 —————> Bivariate plots (Wind speed, direction, PM)
 - Daily
 - Weekly
 - Monthly
 - R^2 with F, RH, rainfall

Weather station information: wunderground.com

Data Analysis

- Average
 - Daily
 - Weekly
 - Monthly
 - Seasonal

```
graph LR; A[Seasonal] --> B[Hourly]; A --> C[Daily]; A --> D[Weekly]
```
- Summary statistics
 - During high traffic flow
 - During low traffic flow

Potential Sources: DOT traffic counters, TripAdvisor, key informants
- Meteorology (T, RH, wind) (Source: wunderground.com)
 - Statistics summary
 - Wind rose diagram for each season
 - Weekly cycles of PM versus Wind speed/direction