Data:

Following datasets are used for the analysis

1. **Borough and Neighborhood Data:**
   We use a free online dataset that has the latitude and longitude coordinates for all neighborhoods in the 5 New York borough. The source of the dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572

2. **Population Projection by Age Group:**
   The target demographic for the pre-school and after-school care business is children under 9 years of age. We need to understand what are the growth projections for this target demographic for each borough to focus our analysis in the areas where there is growth. The dataset for the population projection is obtained from the following source: https://data.cityofnewyork.us/City-Government/Projected-Population-2010-2040-Total-By-Age-Groups/97pn-acdf

3. **New York City Pre-Schools:**
   We need to map the existing pre-schools and child care services in New York City. We will use the Foursquare API query pre-schools in each neighborhood and use the information to create clusters. Following is an example of the data obtained from the Foursquare API:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Anna's Group Family Daycare and Preschool	40.8875	-73.8507	Daycare
1	Wakefield	40.894705	-73.847201	Tender Tots Day Care, Preschool & After School...	40.877	-73.8728	Daycare
2	Wakefield	40.894705	-73.847201	The Lifeskills Preschool	40.8636	-73.8341	Nursery School
3	Co-op City	40.874294	-73.829939	The Lifeskills Preschool	40.8636	-73.8341	Nursery School
4	Co-op City	40.874294	-73.829939	Anna's Group Family Daycare and Preschool	40.8875	-73.8507	Daycare