

## Methods

### Variable selection and data sources

This study proposes multiple models to describe risk factors associated with COVID-19 within selected Manhattan neighborhoods in New York City. An initial set of 41 potential risk indicators were identified based under the broad headings of the built environment, health, and social and economic factors on existing literature on COVID-19 risk factors and comorbidities (Gourevitch et al, Karaye and Horney, Flanagan et al, Roux, Shen and Lung). In searching for data sources, I considered the availability of data for small areal units. Census tract-level information was available for the indicators of interest from the following datasets which were utilized in further analysis: 5-year estimates from the 2014 - 2018 American Community Survey (ACS), the Centers for Disease Control's 2018 Social Vulnerability Index (CDC SVI), the CDC's 500 Cities Project 2017 data, and New York City (NYC) Department of City Planning's MapPluto Release 20v5. Table 1 lists all of the risk indicators assessed in this study and associated data sources.

Table 1. COVID-19 risk indicators and data sources

Construct	Variable	Dataset	Publisher	Year (s)
Built Environment	One & Two Family Buildings	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Multi-Family Buildings	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Mixed Residential & Commercial Buildings	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Commercial & Office Buildings	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Industrial & Manufacturing	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Transportation & Utility	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Public Facilities & Institutions	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Open Space & Outdoor Recreation	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Parking Facilities	MAPPLUTO	NYC Department of City Planning	2020
Built Environment	Vacant Land	MAPPLUTO	NYC Department of City Planning	2020
Social & Economic	Male	American Community Survey	US Census	2014-2018
Social & Economic	Persons aged 50 or older	American Community Survey	US Census	2014-2018
Social & Economic	Civilians in management, business, science, and arts occupations	American Community Survey	US Census	2014-2018
Social & Economic	Civilians in service occupations	American Community Survey	US Census	2014-2018
Social & Economic	Civilians in sales and office occupations	American Community Survey	US Census	2014-2018
Social & Economic	Civilians in natural resources, construction, and maintenance occupations	American Community Survey	US Census	2014-2018

Construct	Variable	Dataset	Publisher	Year (s)
Social & Economic	Civilians in production, transportation, and material moving occupations	American Community Survey	US Census	2014-2018
Social & Economic	Persons below poverty	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Unemployment Rate	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	No high school diploma (age 25+)	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Aged 65 or older	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Civilian non-institutionalized pop with a disability	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Single parent households with children under 18	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Minority (all persons except white, non-Hispanic)	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Persons (age 5+) who speak English "less than well"	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Housing in structures with 10 or more units	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Mobile homes	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Crowding (housing units with more than 1 person per room)	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Households with no vehicle available	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Persons in institutionalized group quarters	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Social & Economic	Percentage uninsured in the total civilian noninstitutionalized population	Social Vulnerability Index	Centers for Disease Control, US Census	2014-2018
Health Status	Arthritis among adults	500 Cities	Centers for Disease Control	2017
Health Status	High blood pressure among adults	500 Cities	Centers for Disease Control	2017
Health Status	Cancer (excluding skin cancer) among adults	500 Cities	Centers for Disease Control	2017
Health Status	Current asthma among adults	500 Cities	Centers for Disease Control	2017
Health Status	Coronary heart disease among adults	500 Cities	Centers for Disease Control	2017
Health Status	Chronic obstructive pulmonary disease among adults	500 Cities	Centers for Disease Control	2017
Health Status	Current smoking among adults	500 Cities	Centers for Disease Control	2017
Health Status	Diagnosed diabetes among adults	500 Cities	Centers for Disease Control	2017
Health Status	Obesity among adults	500 Cities	Centers for Disease Control	2017
Health Status	Stroke among adults	500 Cities	Centers for Disease Control	2017

COVID-19 statistics, including positive cases, hospitalizations, and deaths aggregated by Modified Zip Code Tabulation Areas (MODZCTA), were downloaded from the NYC Department of Health and Mental Hygiene's Github repository on July 7, 2020.

### Geographical scope and data cleaning

Data processing and statistical analyses were completed with QGIS 3.10.6 and in R 3.6.2.

A total of 89 census tracts within the neighborhoods of Chelsea, Flatiron, Midtown South, Murray Hill, and below were included in the analysis. Spatial data was projected within the NAD83 coordinate system. To create the land use variables, additional calculations were performed with the MapPluto spatial data in QGIS to find the percentage area covered by each land use type per census tract. Percentages were also calculated from estimates for people above 50 years old and for occupational groups in the American Community Survey.

### Exploratory analyses

Summary statistics were obtained for risk indicators and COVID-19 outcomes and are presented in Tables 2 and 3, respectively.

Table 2. Percentages across census tracts

Factor	Variable	Average	Min	Max	Standard Deviation
Built Environment	One & Two Family Buildings	1.16	0.00	12.19	2.40
Built Environment	Multi-Family Buildings	18.17	0.00	59.71	14.94
Built Environment	Mixed Residential & Commercial Buildings	19.68	0.00	103.26	17.64
Built Environment	Commercial & Office Buildings	12.64	0.00	59.05	14.14
Built Environment	Industrial & Manufacturing	0.33	0.00	11.25	1.29
Built Environment	Transportation & Utility	3.60	0.00	35.54	7.24
Built Environment	Public Facilities & Institutions	8.16	0.00	92.80	11.59
Built Environment	Open Space & Outdoor Recreation	5.74	0.00	56.28	10.70
Built Environment	Parking Facilities	0.86	0.00	7.39	1.18
Built Environment	Vacant Land	0.98	0.00	11.05	1.84
Social & Economic	Male	48.80	0.00	60.79	7.25
Social & Economic	Persons aged 50 or older	30.86	0.00	56.60	10.47
Social & Economic	Civilians in management, business, science, and arts occupations	64.00	0.00	90.52	16.73
Social & Economic	Civilians in service occupations	11.15	0.00	39.93	9.94
Social & Economic	Civilians in sales and office occupations	19.30	0.00	33.05	5.26
Social & Economic	Civilians in natural resources, construction, and maintenance occupations	1.42	0.00	7.48	1.60
Social & Economic	Civilians in production, transportation, and material moving occupations	3.01	0.00	14.91	3.51
Social & Economic	Persons below poverty	14.91	1.30	51.90	11.04
Social & Economic	Unemployment Rate	4.93	0.30	14.10	2.75
Social & Economic	No high school diploma (age 25+)	10.99	0.00	55.70	14.15
Social & Economic	Aged 65 or older	14.48	1.40	34.30	6.86
Social & Economic	Civilian non-institutionalized pop with a disability	8.96	1.30	25.80	5.44

Factor	Variable	Average	Min	Max	Standard Deviation
Social & Economic	Single parent households with children under 18	3.67	0.00	20.40	3.94
Social & Economic	Minority (all persons except white, non-Hispanic)	42.27	0.00	99.10	24.22
Social & Economic	Persons (age 5+) who speak English "less than well"	8.24	0.00	52.50	11.97
Social & Economic	Housing in structures with 10 or more units	86.67	0.00	100.00	13.12
Social & Economic	Mobile homes	0.06	0.00	1.30	0.20
Social & Economic	Crowding (more than 1 person per room)	5.58	0.00	20.90	4.00
Social & Economic	Households with no vehicle available	79.59	57.80	91.80	7.51
Social & Economic	Persons in institutionalized group quarters	5.91	0.00	49.60	9.98
Social & Economic	Percentage uninsured in the total civilian noninstitutionalized population	4.41	0.00	13.50	2.93
Health Status	Arthritis among adults	14.44	5.60	26.00	4.44
Health Status	High blood pressure among adults	20.50	8.60	40.30	6.85
Health Status	Cancer (excluding skin cancer) among adults	5.12	2.10	9.50	1.46
Health Status	Current asthma among adults	8.23	6.60	10.70	0.65
Health Status	Coronary heart disease among adults	4.02	1.00	11.40	2.16
Health Status	Chronic obstructive pulmonary disease among adults	3.75	1.30	10.10	1.80
Health Status	Current smoking among adults	11.92	7.20	23.60	3.93
Health Status	Diagnosed diabetes among adults	7.12	1.90	22.80	4.60
Health Status	Obesity among adults	14.68	9.60	27.40	3.34
Health Status	Stroke among adults	2.08	0.60	6.70	1.24

Table 3. COVID-19 statistics across MODZCTA

	Average	Min	Max	SD
Positive COVID case per 100k residents (May 18, 2020)	1004.61	447.05	2098.32	392.21
COVID death rate per 100k residents (May 18, 2020)	72.55	0.00	235.70	60.43
Percent positive COVID tests (May 18, 2020)	22.95	16.18	33.80	4.74
Positive COVID case per 100k residents (August 6, 2020)	1288.63	644.58	2601.92	464.54
COVID death rate per 100k residents (August 6, 2020)	85.73	0.00	271.96	69.39
Percent positive COVID tests (August 6, 2020)	6.14	4.14	9.60	1.57

Pearson correlation coefficients were utilized to identify significant associations (at a confidence level of 95%) between the risk construct variables for further analysis. Variables were ranked by the number of significant associations in an initial screening for model factors. Findings have been reproduced below in Table 4.

Table 4. Ranking of variables by number of significant statistical correlations (95% CI)

Counts	Percentile rank	Variable
0	0	Transportation & Utility
1	0.025	Public Facilities & Institutions
2	0.05	Mixed Residential & Commercial Buildings
2	0.05	Vacant Land
3	0.1	Industrial & Manufacturing
6	0.125	Persons in institutionalized group quarters
6	0.125	Parking Facilities
10	0.175	One & Two Family Buildings
12	0.2	Mobile homes
14	0.225	Crowding (housing units with more than 1 person per room)
15	0.25	Housing in structures with 10 or more units
16	0.275	Multi-Family Buildings
17	0.3	Cancer (excluding skin cancer) among adults
18	0.325	Households with no vehicle available
18	0.325	Civilians in sales and office occupations
19	0.375	Civilians in natural resources, construction, and maintenance occupations
20	0.4	Percentage uninsured in the total civilian noninstitutionalized population
20	0.4	Civilians in production, transportation, and material moving occupations
21	0.45	Persons (age 5+) who speak English "less than well"
21	0.45	Percentage Male
22	0.5	Unemployment Rate
22	0.5	Single parent households with children under 18
22	0.5	Open Space & Outdoor Recreation
23	0.575	Persons below poverty
23	0.575	Aged 65 or older
23	0.575	Commercial & Office Buildings
24	0.65	Civilian non-institutionalized pop with a disability
25	0.675	No high school diploma (age 25+)
25	0.675	Current asthma among adults
25	0.675	Civilians in service occupations
28	0.75	Minority (all persons except white, non-Hispanic)

Counts	Percentile rank	Variable
28	0.75	Coronary heart disease among adults
29	0.8	Arthritis among adults
29	0.8	High blood pressure among adults
29	0.8	Chronic obstructive pulmonary disease among adults
29	0.8	Current smoking among adults
29	0.8	Aged 50 or older
30	0.925	Diagnosed diabetes among adults
30	0.925	Obesity among adults
30	0.925	Stroke among adults
31	1	Civilians in management, business, science, and arts occupations