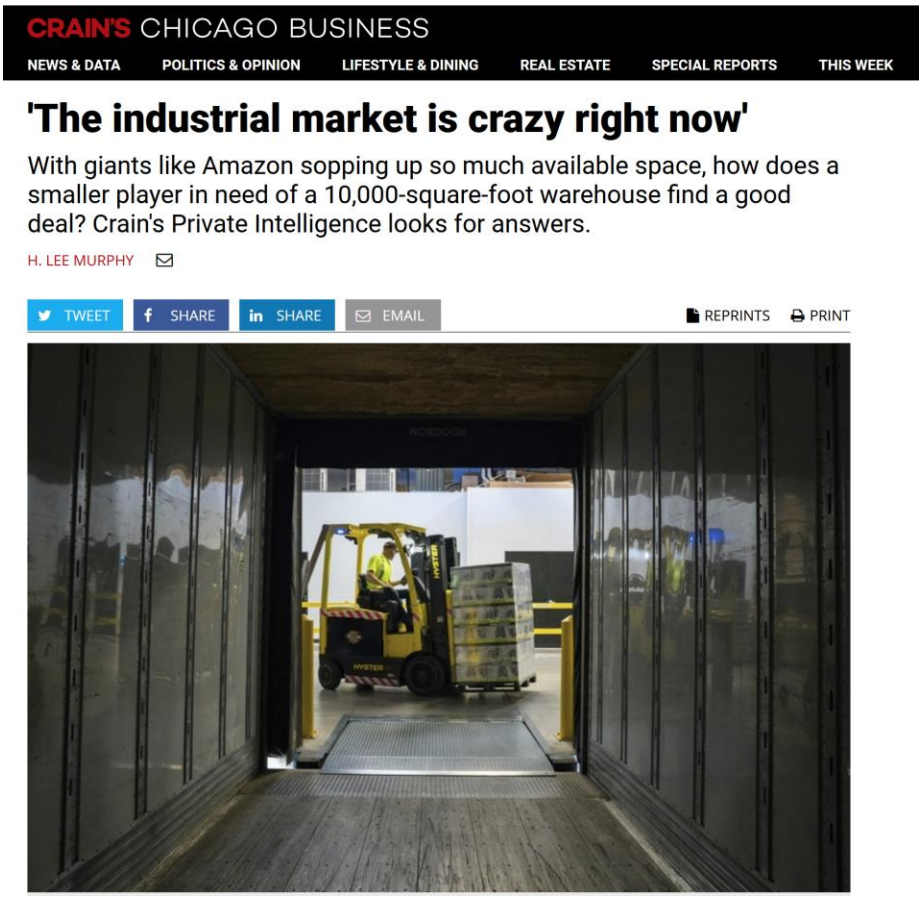
Chicago Business Intelligence for Strategic Planning

Requirements Specification

Atef Bader, PhD



CRAIN'S CHICAGO BUSINESS

NEWS & DATA | POLITICS & OPINION | LIFESTYLE & DINING | REAL ESTATE | SPECIAL REPORTS | THIS WEEK

'The industrial market is crazy right now'

With giants like Amazon sopping up so much available space, how does a smaller player in need of a 10,000-square-foot warehouse find a good deal? Crain's Private Intelligence looks for answers.

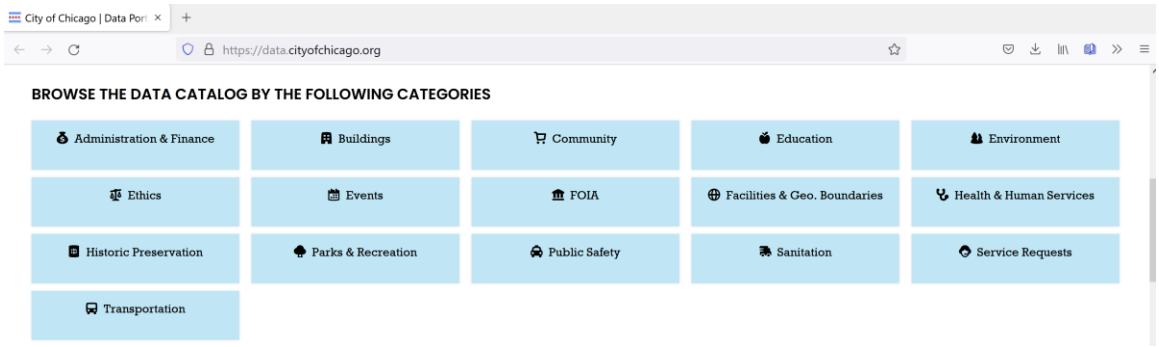
H. LEE MURPHY

TWEET | SHARE | SHARE | EMAIL | REPRINTS | PRINT

1. Overview

In this project, you have been tasked as a full-stack developer to build an application that will be used by data scientists and business analysts for exploratory data analysis and to create different business intelligence reports for the city of Chicago; these reports will be utilized in the strategic planning and the industrial and neighborhood infrastructure investments.

The City of Chicago publishes and updates its datasets on its data portal server (https://data.cityofchicago.org/) in 16 categories. The 3 categories that this project will utilize for exploratory data analysis and creating the business intelligence reports are: Transportation, Buildings, and Health & Human Services.



2. Communities and Businesses Welfare

Requirement 1: The business intelligence reports are geared toward tracking and forecasting events that have direct or indirect negative or positive impacts on businesses and neighborhoods in different zip codes within the city of Chicago. The business intelligence reports will be used to send alerts to taxi drivers about the state of COVID-19 in the different zip codes in order to avoid taxi drivers to be the super spreaders in the different zip codes and neighborhoods. For this report, the taxi trips and daily COVID-19 datasets for the city of Chicago will be used.

The City of Chicago is also interested to forecast COVID-19 alerts (Low, Medium, High) on daily/weekly basis to the residents of the different neighborhoods considering the counts of the taxi trips and COVID-19 positive test cases.

Requirement 2: There are two major airports within the city of Chicago: O'Hare and Midway. And the City of Chicago is interested to track trips from these airports to the different zip codes and the reported COVID-19 positive test cases. The city of Chicago is interested to monitor the traffic of the taxi trips from these airports to the different neighborhoods and zip codes

Requirement 3: The city of Chicago has created the COVID-19 Community Vulnerability Index (CCVI) (https://data.cityofchicago.org/Health-Human-Services/Chicago-COVID-19-Community-Vulnerability-Index-CCV/xhc6-88s9) to identify communities that have been disproportionately affected by COVID-19 and are vulnerable to barriers to COVID-19 vaccine uptake. The city of Chicago is interested to track the number of taxi trips from/to the neighborhoods that have CCVI Category with value HIGH

3. Community Investments and Business Incentives

Requirement 4: For streetscaping investment and planning, the city of Chicago is interested to forecast daily, weekly, and monthly traffic patterns utilizing the taxi trips for the different zip codes.

Requirement 5: For industrial and neighborhood infrastructure investment, the city of Chicago is interested to invest in top 5 neighborhoods with highest unemployment rate and poverty rate and waive the fees for building permits in those neighborhoods in order to encourage businesses to develop and invest in those neighborhoods. Both, building permits and unemployment, datasets will be used in this report.

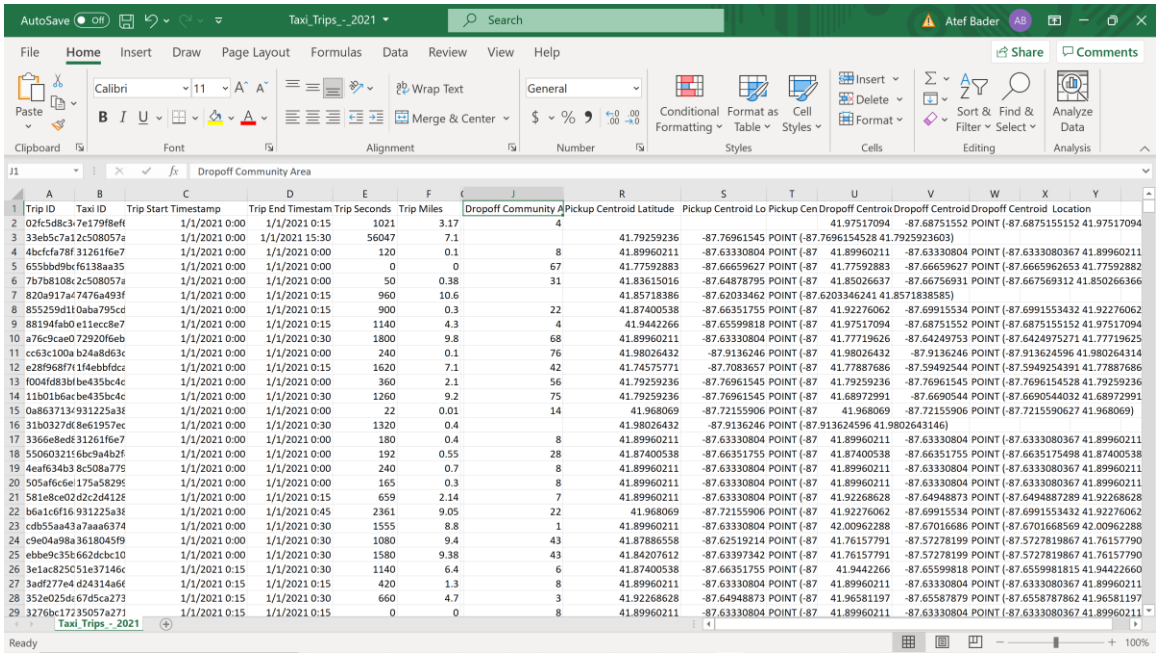
Requirement 6: According to a report published by Crain's Chicago Business (https://www.chicagobusiness.com/private-intelligence/industrial-market-crazy-right-now), The "little guys", small businesses, have trouble competing with the big players like Amazon and Walmart for warehouse spaces. To help small business, assume a new imaginary program has been piloted with the name Illinois Small Business Emergency Loan Fund Delta to offer small businesses low interest loans of up to $250,000 for those applicants with PERMIT_TYPE of PERMIT - NEW CONSTRUCTION in the zip code that has the lowest number of PERMIT - NEW CONSTRUCTION applications and PER CAPITA INCOME is less than 30,000 for the planned construction site. Both, building permits and unemployment, datasets will be used in this report.

4. Data Lake

Requirement 7: Utilize the following data sources to construct the data lake using one of the commonly used database engines like Postgres, MySQL, etc.

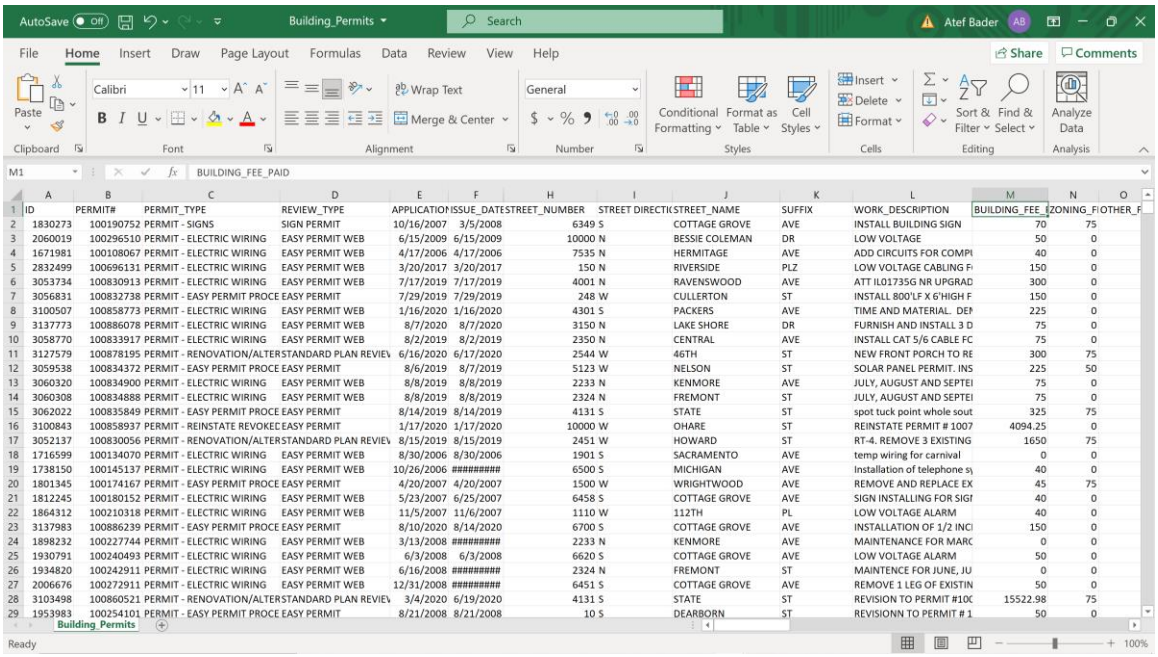
- ## Transportation

  Taxi trips are reported to the City of Chicago in its role as a regulatory agency. The dataset is available for download from the following URL: (https://data.cityofchicago.org/Transportation/Taxi-Trips/wrvz-psew). The dataset for the trips reported by the Transportation Network Providers (also called rideshare companies) is available for download from the following URL: (https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p)



- **Buildings**

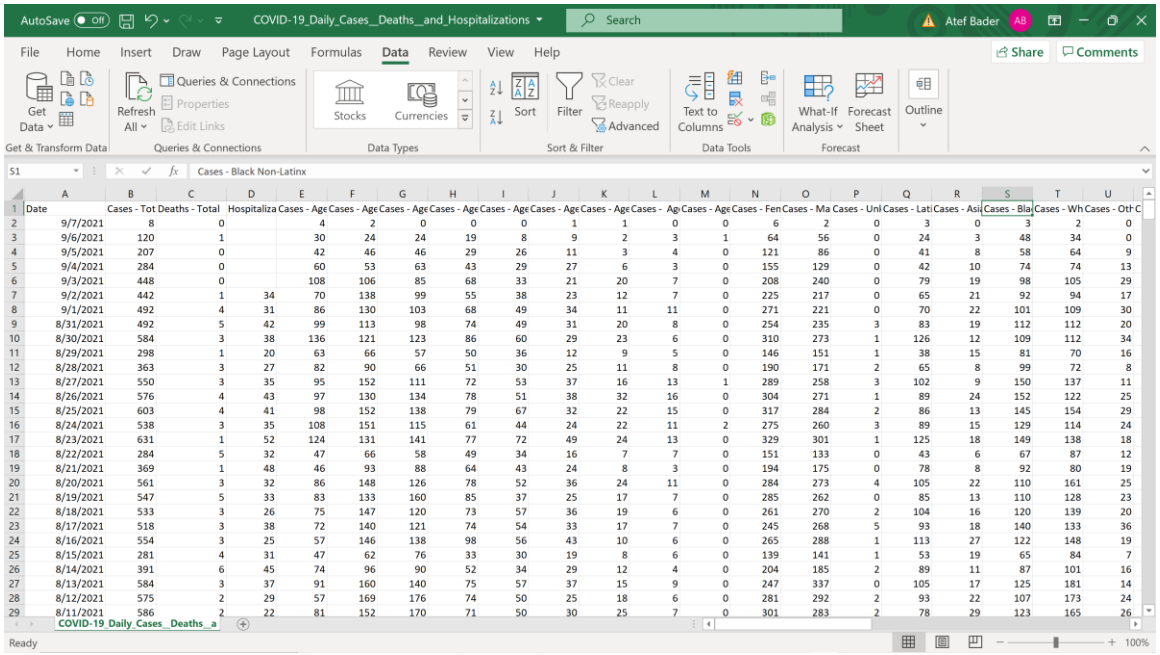
  Permits issued by the Department of Buildings in the City of Chicago. The dataset is available for download from the following URL: (https://www.chicago.gov/city/en/depts/bldgs/dataset/building_permits.html).

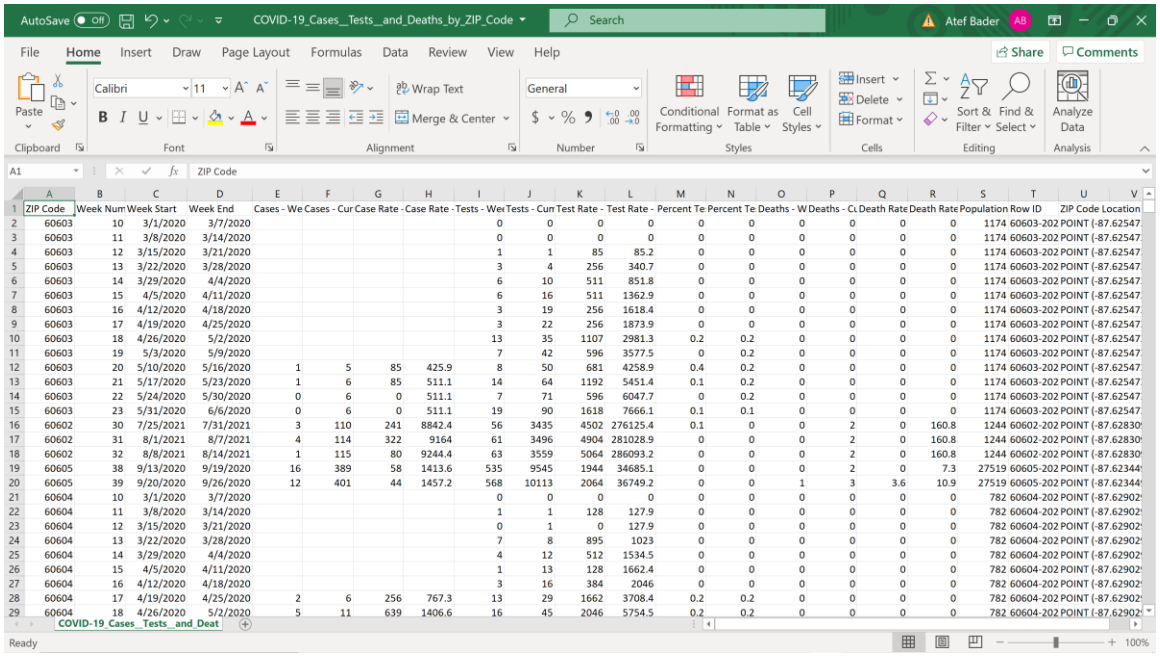


ID	PERMIT#	PERMIT_TYPE	REVIEW_TYPE	APPLICATION	ISSUE_DATE	STREET_NUMBER	STREET DIRECTI	STREET_NAME	SUFFIX	WORK_DESCRIPTION	BUILDING_FEE_	ZONING_FI	OTHER_F
1830273	100190752	PERMIT - SIGNS	SIGN PERMIT	10/16/2007	3/5/2008	6349	S	COTTAGE GROVE	AVE	INSTALL BUILDING SIGN	70	75	
2060019	100296510	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	6/15/2009	6/15/2009	10000	N	BESSIE COLEMAN	DR	LOW VOLTAGE	50	0	
1671981	100108067	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	4/17/2006	4/17/2006	7535	N	HERMITAGE	AVE	ADD CIRCUITS FOR COMPI	40	0	
2832499	100696131	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	3/20/2017	3/20/2017	150	N	RIVERSIDE	PLZ	LOW VOLTAGE CABLING F	150	0	
3053734	100830913	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	7/17/2019	7/17/2019	4001	N	RAVENSWOOD	AVE	ATT IL01735G NR UPGRAD	300	0	
3056831	100832738	PERMIT - EASY PERMIT PROCE	EASY PERMIT	7/29/2019	7/29/2019	248	W	CULLERTON	ST	INSTALL 800'LF X 6'HIGH F	150	0	
3100507	100858773	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	1/16/2020	1/16/2020	4301	S	PACKERS	AVE	TIME AND MATERIAL. DEI	225	0	
3137773	100886078	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	8/7/2020	8/7/2020	3150	N	LAKE SHORE	DR	FURNISH AND INSTALL 3 D	75	0	
3058770	100833917	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	8/2/2019	8/2/2019	2350	N	CENTRAL	AVE	INSTALL CAT 5/6 CABLE FC	75	0	
3127579	100878195	PERMIT - RENOVATION/ALTER	STANDARD PLAN REVIEV	6/16/2020	6/17/2020	2544	W	46TH	ST	NEW FRONT PORCH TO RE	300	75	
3059538	100834372	PERMIT - EASY PERMIT PROCE	EASY PERMIT	8/6/2019	8/7/2019	5123	W	NELSON	ST	SOLAR PANEL PERMIT. INS	225	50	
3060320	100834900	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	8/8/2019	8/8/2019	2233	N	KENMORE	AVE	JULY, AUGUST AND SEPTEI	75	0	
3060308	100834888	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	8/8/2019	8/8/2019	2324	N	FREMONT	ST	JULY, AUGUST AND SEPTEI	75	0	
3062022	100835849	PERMIT - EASY PERMIT PROCE	EASY PERMIT	8/14/2019	8/14/2019	4131	S	STATE	ST	spot tuck point whole sout	325	75	
3100843	100858937	PERMIT - REINSTATE REVOKEC	EASY PERMIT	1/17/2020	1/17/2020	10000	W	OHARE	ST	REINSTATE PERMIT # 1007	4094.25	0	
3052137	100830056	PERMIT - RENOVATION/ALTER	STANDARD PLAN REVIEV	8/15/2019	8/15/2019	2451	W	HOWARD	ST	RT-4. REMOVE 3 EXISTING	1650	75	
1716599	100134070	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	8/30/2006	8/30/2006	1901	S	SACRAMENTO	AVE	temp wiring for carnival	0	0	
1738150	100145137	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	10/26/2006	#########	6500	S	MICHIGAN	AVE	Installation of telephone s	40	0	
1801345	100174167	PERMIT - EASY PERMIT PROCE	EASY PERMIT	4/20/2007	4/20/2007	1500	W	WRIGHTWOOD	AVE	REMOVE AND REPLACE EX	45	75	
1812245	100180152	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	5/23/2007	6/25/2007	6458	S	COTTAGE GROVE	AVE	SIGN INSTALLING FOR SIGI	40	0	
1864312	100210318	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	11/5/2007	11/6/2007	1110	W	112TH	PL	LOW VOLTAGE ALARM	40	0	
3137983	100886239	PERMIT - EASY PERMIT PROCE	EASY PERMIT	8/10/2020	8/14/2020	6700	S	COTTAGE GROVE	AVE	INSTALLATION OF 1/2 INCI	150	0	
1898232	100227744	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	3/13/2008	#########	2233	N	KENMORE	AVE	MAINTENANCE FOR MARC	0	0	
1930791	100240493	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	6/3/2008	6/3/2008	6620	S	COTTAGE GROVE	AVE	LOW VOLTAGE ALARM	50	0	
1934820	100242911	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	6/16/2008	#########	2324	N	FREMONT	ST	MAINTENCE FOR JUNE, JU	0	0	
2006676	100272911	PERMIT - ELECTRIC WIRING	EASY PERMIT WEB	12/31/2008	#########	6451	S	COTTAGE GROVE	AVE	REMOVE 1 LEG OF EXISTIN	50	0	
3103498	100860521	PERMIT - RENOVATION/ALTER	STANDARD PLAN REVIEV	3/4/2020	6/19/2020	4131	S	STATE	ST	REVISION TO PERMIT #10(	15522.98	75	
1953983	100254101	PERMIT - EASY PERMIT PROCE	EASY PERMIT	8/21/2008	8/21/2008	10	S	DEARBORN	ST	REVISIONN TO PERMIT # 1	50	0	

- **Health and Human Services**

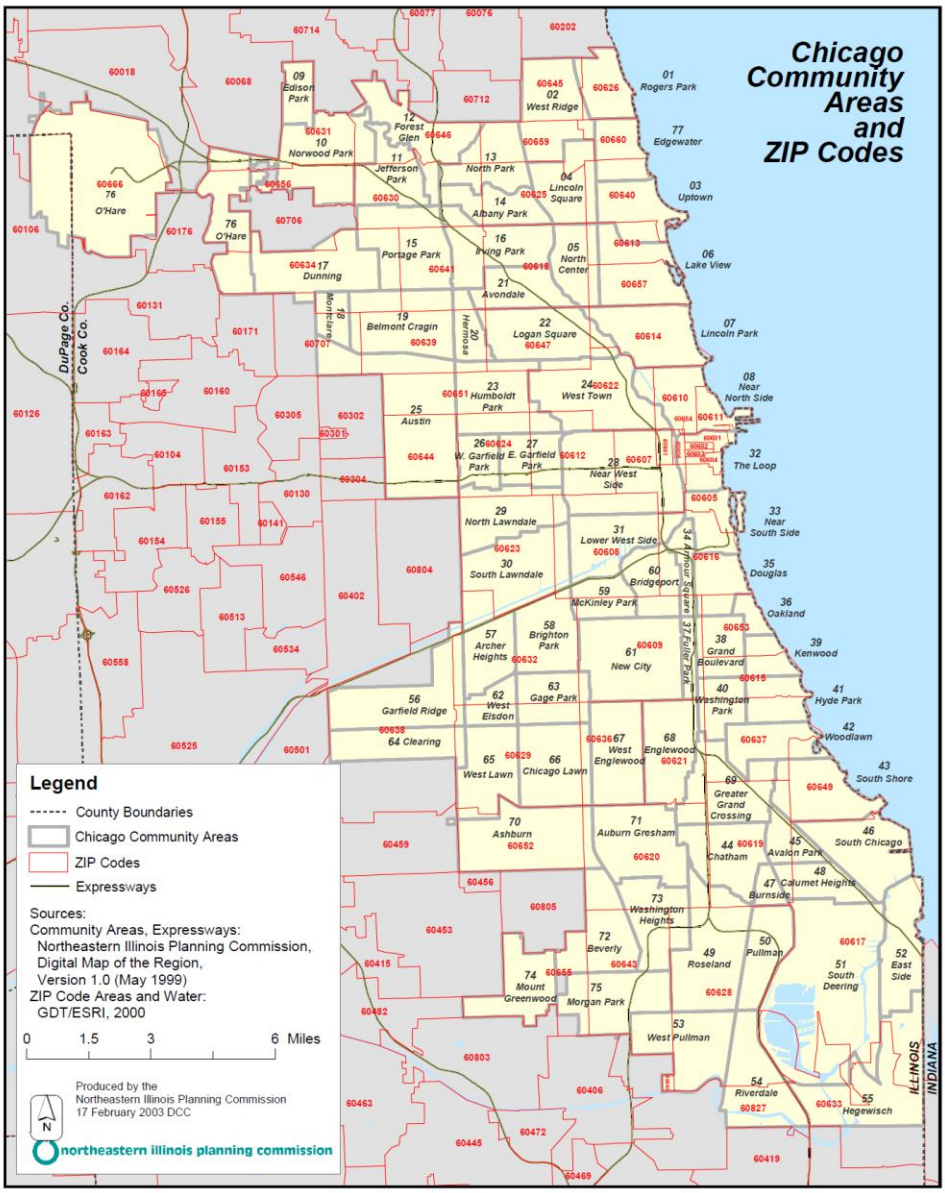
For all datasets related to COVID-19, see the following URL: (https://data.cityofchicago.org/browse?limitTo=datasets&sortBy=alpha&tags=covid-19).





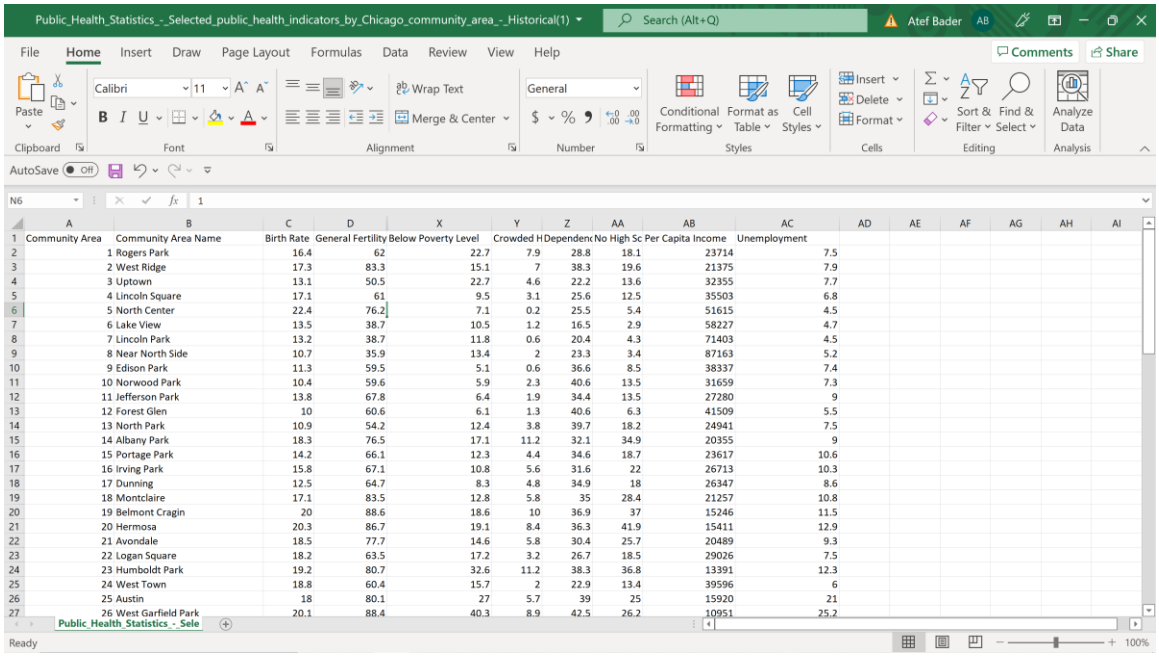
- **Neighborhood Names, Community Areas, and Zip Codes**

  Consider the Chicago Tribune (https://www.chicagotribune.com/chi-community-areas-htmlstory.html) data source and the City of Chicago data portal (https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Community-Areas-current-/cauq-8yn6) to cross-reference Neighborhood Names, Community Areas, and Zip Codes.



- **Unemployment**

  Consider the City of Chicago data portal (https://data.cityofchicago.org/Health-Human-Services/Public-Health-Statistics-Selected-public-health-in/iqnk-2tcu/data) and (https://data.cityofchicago.org/resource/iqnk-2tcu.json) to find the unemployment and poverty level data for the different community areas.



Public_Health_Statistics_-_Selected_public_health_indicators_by_Chicago_community_area_-_Historical(1)

Community Area	Community Area Name	Birth Rate	General Fertility	Below Poverty Level	Crowded H	Dependenc	No High Sc	Per Capita Income	Unemployment
1	Rogers Park	16.4	62	22.7	7.9	28.8	18.1	23714	7.5
2	West Ridge	17.3	83.3	15.1	7	38.3	19.6	21375	7.9
3	Uptown	13.1	50.5	22.7	4.6	22.2	13.6	32355	7.7
4	Lincoln Square	17.1	61	9.5	3.1	25.6	12.5	35503	6.8
5	North Center	22.4	76.2	7.1	0.2	25.5	5.4	51615	4.5
6	Lake View	13.5	38.7	10.5	1.2	16.5	2.9	58227	4.7
7	Lincoln Park	13.2	38.7	11.8	0.6	20.4	4.3	71403	4.5
8	Near North Side	10.7	35.9	13.4	2	23.3	3.4	87163	5.2
9	Edison Park	11.3	59.5	5.1	0.6	36.6	8.5	38337	7.4
10	Norwood Park	10.4	59.6	5.9	2.3	40.6	13.5	31659	7.3
11	Jefferson Park	13.8	67.8	6.4	1.9	34.4	13.5	27280	9
12	Forest Glen	10	60.6	6.1	1.3	40.6	6.3	41509	5.5
13	North Park	10.9	54.2	12.4	3.8	39.7	18.2	24941	7.5
14	Albany Park	18.3	76.5	17.1	11.2	32.1	34.9	20355	9
15	Portage Park	14.2	66.1	12.3	4.4	34.6	18.7	23617	10.6
16	Irving Park	15.8	67.1	10.8	5.6	31.6	22	26713	10.3
17	Dunning	12.5	64.7	8.3	4.8	34.9	18	26347	8.6
18	Montclaire	17.1	83.5	12.8	5.8	35	28.4	21257	10.8
19	Belmont Cragin	20	88.6	18.6	10	36.9	37	15246	11.5
20	Hermosa	20.3	86.7	19.1	8.4	36.3	41.9	15411	12.9
21	Avondale	18.5	77.7	14.6	5.8	30.4	25.7	20489	9.3
22	Logan Square	18.2	63.5	17.2	3.2	26.7	18.5	29026	7.5
23	Humboldt Park	19.2	80.7	32.6	11.2	38.3	36.8	13391	12.3
24	West Town	18.8	60.4	15.7	2	22.9	13.4	39596	6
25	Austin	18	80.1	27	5.7	39	25	15920	21
26	West Garfield Park	20.1	88.4	40.3	8.9	42.5	26.2	10951	25.2

5. Data Collection and Preparation

Requirement 7: Review the available and relevant datasets from the City of Chicago data portal, inspect the data attributes in the different datasets that will be utilized in this project for the construction of the data lake and identify missing attribute(s) or messy data in these datasets, and how these issues will be tackled in the design and implementation of the project.

6. Infrastructure and Technology Stack for Microservices and Cloud-Native Computing

Requirement 8: Review the following technologies/resources and document where they can be utilized in the infrastructure, data lake, microservices and the full-stack development of this project:

1. Postgres (Or any other modern RDBMS engine) to create your data lake.
2. Go language for building microservices. (https://golang.org/dl/)
3. Docker/Container/Kubernetes for deployment of microservices. (https://www.docker.com/get-started)
4. Google Cloud CLI to deploy cloud-native microservices. (https://cloud.google.com/sdk/docs/install)
5. Python packages for time-series forecasting and geospatial queries.
    - Geopy (https://geopy.readthedocs.io/en/stable/)
    - Arcgis (http://www.arcgis.com)
    - Google Maps (https://github.com/googlemaps/google-maps-services-python)
    - Tensorflow/Keras LSTM (https://www.tensorflow.org/api_docs/python/tf/keras/layers/LSTM)
    - Facebook/Prophet (https://facebook.github.io/prophet/)

7. Forecasting and Strategic Planning:

Requirement 9: If you live in the City of Chicago, highly likely you have heard from some of the residents, at some point, to sarcastically state that the city has two seasons: winter season and construction season. To better help in resource allocation, scheduling, and planning for streetscaping projects, the City of Chicago is interested to forecast the volume of traffic

in neighborhoods and zip codes using the taxi trips as a proxy. Use the taxi trip dataset to forecast the daily, weekly, and monthly taxi trips for every zip code and neighborhood.

8. Dashboard For Exploratory Data Analysis and Forecasting

Requirement 10: A web-based frontend dashboard is needed for exploratory data analysis and forecasting of the different datasets and requirements discussed in this document. The frontend can be implemented in any of the modern web-based frameworks/libraries, like Angular, React, etc.