EXAMINING THE LOCAL-SCALE RELATIONSHIP BETWEEN HUMAN MOBILITY AND COVID-19: A CASE STUDY OF SAN DIEGO, CA

A Thesis

Presented to the

Faculty of

San Diego State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Big Data Analytics

---

by

Huong Thi Lien Pham

Spring 2023

SAN DIEGO STATE UNIVERSITY

The Undersigned Faculty Committee Approves the

Thesis of Huong Thi Lien Pham:

Examining the Local-Scale Relationship between Human Mobility and

COVID-19: A Case Study of San Diego, CA

Atsushi Nara, Chair
Department of Geography

Ming-Hsiang Tsou
Department of Geography

Eyal Oren
School of Public Health

Approval Date

DEDICATION

This thesis is dedicated to my beloved family

For their endless love, support, and encouragement.

ABSTRACT OF THE THESIS

Examining the Local-Scale Relationship between Human Mobility
and COVID-19: A Case Study of San Diego, CA
by
Huong Thi Lien Pham
Master of Science in Big Data Analytics
San Diego State University, 2023

A two-way relationship exists between human mobility and COVID-19 cases. While existing studies analyzed this relationship often at country and county levels using the available human mobility and COVID-19 data, fewer studies have focused on the connection at the neighborhood level, partially due to the limited data availability and data integration challenges at the local scale. Nevertheless, to better support decisions, it is essential to understand the connection at a finer geographic and temporal scale, which can provide insights into how human mobility and their behavior and interaction affect disease transmission from the bottom up. This research aims to examine the spatial and temporal dimensions of human mobility and how human movement influences the diffusion of COVID-19 in San Diego, CA.

The human mobility data used in this research is the daily number of mobile devices aggregated from the census block group level to the San Diego Subregional Areas (SRA) level, using the Social Distancing Metrics of the SafeGraph data. The COVID-19 confirmed case data is the number of COVID-19 cases provided by the San Diego Health and Human Services Agency, aggregated to SRAs from the zip code level using the Dasymetric Mapping by HDMA. Dynamic Time Warping (DTW) measures the similarity between human mobility and the COVID-19 time series at the SRA level. The slopes of the time series are then calculated and compared with the DTW values to validate the increase-increase relationship between human mobility and COVID-19 cases.

This study found that the relationship between human mobility and COVID-19 in San Diego depends on time and place. The positive correlation between human mobility and COVID-19 cases happened the most during the winter and holiday season of 2020. The flows of people coming to San Diego County (inflow and netflow) impact the Covid-19 case increase more than those going out of the County (outflow). The human mobility inflow affects the South region more, while the North Central is affected the most by the mobility netflow. Spatial association among SRAs does exist. The revealed patterns are meaningful to local-scale policymaking.

TABLE OF CONTENTS

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

ACKNOWLEDGEMENTS

First and foremost, my thanks go to Dr. Atsushi Nara for his guidance and for supporting me during the past two years and throughout the thesis process. He introduced many methods to approach a data science problem, and my Python skills improved significantly through trial and error and by exploring various approaches he suggested for this thesis.

I would also like to thank Dr. Eyal Oren, Dr. Ming-Hsiang Tsou, and Dr. Atsushi Nara for being my Thesis Committee members. Their invaluable feedback on my thesis helped me look at my study from public health and geographic perspectives and work towards complete academic research.

Finally, I would like to thank my family for being patient and allowing me to find the time needed to finish this thesis.

CHAPTER 1

INTRODUCTION

COVID-19 is a disease caused by SARS-CoV-2, one of the seven types of coronaviruses (Nazario, 2021). Similar to other coronaviruses, such as the Middle East respiratory syndrome (MERS, started 2012) and the sudden acute respiratory syndrome (SARS, 2003), SARS-CoV-2 transmits from person to person through exhaled and inhaled droplets of respiratory fluids. The difference is that SARS-CoV-2 is more contagious and deadlier than the previous types. In addition, the droplets can stay suspended in the air for minutes to hours and spread the virus easier. Since the first COVID-19 case was discovered in Wuhan, China, and spread to almost all other countries, more than six million people have died, and businesses and economies have been significantly impacted (Worldometer, n.d.).

SARS-CoV-2 can transmit from one to another within a six-feet distance in fifteen minutes or longer (according to the CDC). Humans travel to meet one another. When they meet, close contacts increase the risk of infection. Therefore, human mobility plays a crucial role in the spread of this airborne COVID-19 disease (Alessandretti, 2022). As a result, governments place a series of “lock-down” and “stay-at-home” orders to control human movements and the size of indoor gatherings to reduce human contact. At some points, governments lift the COVID-19 restrictions when the growth rate of confirmed cases reduces and flattens, or reasonable COVID-19 vaccination rates are achieved.

Human mobility is one factor that affects COVID-19 case numbers. Besides human mobility, the COVID-19 case numbers depend on vaccination efforts and other standard COVID-19 control measures such as mask-wearing, hand-washing, and physical distancing. Thus, the relationship between human mobility and COVID-19 cases can be diverse, depending on place and time (Linka et al., n.d.; S. Wang et al., 2020; Zeng et al., 2021).

Many studies analyze this relationship at a macro-scale (e.g., country and county level) using available human mobility and COVID-19 data. However, fewer studies focus on the connection at the neighborhood level, partially due to the limited data availability and data integration challenges at the local scale. Therefore, this study selects San Diego, CA as a case study and examines the spatial and temporal dimensions of the human mobility – COVID-19 relationship in 41 San Diego County subregional areas, hoping to contribute to San Diego's COVID-19 policymaking.

This study aims to answer these questions:

1. When does human mobility influence COVID-19 diffusion in San Diego?
2. Where does it happen?
3. What type of human mobility has more effect on San Diego's COVID-19 transmission? For example, people coming to San Diego from other counties and states bring more transmission, or people moving inside local San Diego areas affect the disease more.

This study helps understand the connection between human mobility and COVID-19 at a finer geographic and temporal scale and provides insights into how human mobility affects disease transmission from a bottom-up approach.

CHAPTER 2

LITERATURE REVIEW

A literature review is started by searching the keywords "*human mobility*" and "*COVID-19*" in different sources such as PubMed, Nature, The Lancet Infectious Diseases, and Google Scholar. Up to a hundred studies were published when this literature review was conducted. First, 25 specific studies closest to the COVID-19 – human mobility topic are selected. Their information is synthesized into four key categories – *geographic scale, research findings, data, and methodology* as in Table 2.1.

A correlation between human mobility and COVID-19 case numbers is found in the initial stage of the pandemic (Badr et al., 2020; Cot et al., 2021; Kraemer et al., 2020; Linka et al., n.d.; Nouvellet et al., 2021; B. Wang et al., 2020; Wei et al., 2021; Yechezkel et al., 2021; Zeng et al., 2021), then the correlation reduced and became diverse in the successive stages (Elarde et al., 2021; Li et al., 2021; Nouvellet et al., 2021; Praharaj & Han, 2021; S. Wang et al., 2020; Yuan et al., 2020).

Menkir et al. (2021) find that the number of flights is a driver for the disease importation among countries when COVID-19 first start. The border control orders, including international travel restrictions and airport screening, reduce long-distance travel across countries and regions and drive the case number down. Sharing a similar conclusion, Chinazzi et al. (2020) also point out a decrease in the impact of long travel on COVID-19 diffusion after the disease spread rapidly in communities. In Alessandretti's (2022) study, the national lockdowns or stay-at-home orders help reduce the mobility amount and disease diffusion at local places at the beginning of the pandemic. After the government relaxed the restrictions, a mixed relationship between mobility and case numbers was noticed, likely because people adopted social distancing behaviors at later phases.

Zeng et al. (2021) look at the time lag effect of mobility restriction on the COVID-19 spread and show that the window within seven days has a minor prediction error during the

Table 2.1. Literature Review Summary

No	Area	Scale	Published Date	Mobility Data Source	Mobility Data Type	Mobility Data Aggregate	Mobility Data Duration	COVID-19 Data Source	COVID-19 Data Type	COVID-19 Data Aggregate	COVID-19 Data Duration	Key Findings 1st wave	Key Findings 2nd wave	Methods
1	EU	1	11/8/20	Apple data	air + car + phone	7-day moving average	02/28/20-05/02/20	26 European countries	C	difference b/w today and yesterday's reported cases	n/a	+	-	1
2	USA	3	7/1/20	Teralytics	phone	daily	01/01/20 - 04/20/20	Johns Hopkins University COVID-19 dashboard	C, D	daily growth rate	n/a	+	+	2
3	EU	1	5/5/20	Air travel statistics (Eurostat, 2020)	air	travel frequency b/w countries	n/a	European Centre for Disease Prevention and Control (ECDC)	C, A, D, R	difference between confirmed cases minus recovered cases and deaths	01/22/20-04/13/20	+	-	1
4	AU	2	10/29/20	Google	GPS locations	time	02/15/20-08/15/20	Aus. Health Dept	C	growth rate + doubling time	02/15/20-08/15/20	mixed	mixed	5, 4
5	EU	2	11/30/20	Google, Apple	GPS locations + individual activities	daily incidence	n/a	ECDC	C	daily	01/01/20-05/09/20	mixed	mixed	4, 5
6	USA	3	7/15/20	Google	GPS locations	daily	n/a	NY Times	n/a	daily	n/a	-	-	8
7	ASIA	5	10/28/20	Tencent migration data	n/a	daily	01/21/19-02/09/19	National Health Commission of China	C	daily	01/17/20-02/06/20	+	-	9
8	USA	2, 3	12/16/20	https://data.covid.umd.edu/	trip number/ person; average person-miles travelled; % of residents staying home	daily	01/01/20-04/11/20	Johns Hopkins University COVID-19 dashboard	C	daily	n/a	-	-	10
9	ASIA	5	10/21/20	Baidu Migration Map (https://qianxi.baidu.com/)	population mobility scale index	daily	01/01/20-03/02/20	n/a	C	daily	01/01/20-03/02/20	+	-	10
10	ASIA	4	5/1/20	Baidu Inc.	real-time mobility data	real-time	1/1/20 - 02/10/20	n/a	C	growth rate	1/1/20 - 02/10/20	+	-	2
11	USA	2, 3	11/10/20	SafeGraph, Google data	phone data	hourly	03/01/20-05/02/20	NY Times	C, D	calibrate from cumulative count to daily count	n/a	-	-	1
12	n/a	n/a	9/1/20	n/a	phone (call detail records, GPS traces)	n/a	n/a	n/a	n/a	n/a	n/a	-	-	n/a
13	EU, USA	1	2/18/21	Google, Apple	location, phone	time away from home	03/2020-05/2020	https://ourworldindata.org/ https://covidtracking.com	C	cases per million inhabitants	03/2020-05/2020	+	-	11
14	USA	3	12/23/20	Unacast	phone GPS	index of mobility change	02/09/20-03/08/20	n/a	n/a	n/a	n/a	-	-	8, 3
15	USA	2	6/29/20	n/a	n/a	n/a	n/a	NY Times	C, D, R	n/a	till 06/21/20	-	-	7

(table continues)

Table 2.1. (continued)

No	Area	Scale	Published Date	Mobility Data Source	Type	Aggregate	Duration	COVID-19 Data Source	Type	Aggregate	Duration	Key Findings 1st wave	2nd wave	Methods
16	World	1	4/15/20	n/a	n/a	n/a	n/a	n/a	C	count/ country	01/01/20-03/28/20	-	-	7
17	Israel	4	3/25/21	n/a	cellular data	hourly; daily proportion of people traveling > 1.5 km away from home	02/01/20-05/16/20	n/a	C	daily	02/13/20-05/15/20	+	+	1
18	USA	2, 3	4/13/21	Twitter	number of Twitter users traveling > 0.5 miles	daily	03/06/20-11/11/20	NY Times	C	cumulative	till 11/11/20	+	+	12
19	n/a	n/a	12/13/21	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	mixed	mixed	n/a
20	USA	3	11/2/21	SafeGraph	mobility footprint data	daily	01/01/19-12/31/20	n/a	n/a	n/a	n/a	mixed	mixed	13, 14, 15, 5
21	USA	5	11/25/20	SafeGraph	POI data to map CBG-POI movement networks	weekly	12/30/19-05/11/20	n/a	n/a	n/a	n/a	mixed	mixed	14, 16
22	World	n/a	4/13/21	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	mixed	mixed	17
23	ASIA	2	9/2/21	Google (GCMR)	frequency and length of visits to common places	daily	03/14/20-09/11/20	COVID-19 India Tracker	C	daily/state	03/14/20-09/11/20	mixed	mixed	18, 19
24	World	1, 2, 3	12/18/20	Twitter, SafeGraph, Google, Apple	geotagged Twitter data, phone-based mobility data	year, month, day, hour	5 years before 2020	n/a	C	daily/country	2 years	-	-	20, 21, 22
25	World	1	2/17/21	Apple, Google	data-streams, quantified relative to the max mobility before pandemic	7-day rolling average	01/13/20 (Apple) or 02/15/20 (Google) to 10/25/20	World Health Organization, ECDC, countries	D	daily/country	up to 10/25/2020	+	-	?

Scale: 1 country, 2 state, 3 county, 4 province/region, 5 city/ metropolitan city; COVID-19 Data Type: C confirmed, D death, A active, R recovered; Key Findings: (+) correlated, (-) not correlated or not decided, mixed both correlated and not correlated depending on space and time; n/a: not mentioned first pandemic wave (15 March - 15 June 2020). The study divides mobility into different types and examines the association between each type with the COVID-19 spread. The correlation between mobility to public transit (a kind of mobility) and COVID-19 cases is negative in Victoria, Australia, during the second wave of COVID-19 (15 June to 15 August 2020).

Almost all the 25 studies use COVID-19 confirmed cases instead of death cases, aggregated to a geographic location by daily time unit. A few works use COVID-19 death cases, usually when authors utilize the SEIR method in their research (Chang et al., 2021; Linka et al., 2020; Nouvellet et al., 2021). COVID-19 data are provided by governments and international organizations such as the European Centre for Disease Prevention and Control (ECDC) and the World Health Organization, or Johns Hopkins and New York Times, the two popular COVID-19 data sources of the U.S.

Human mobility data could be air travel statistics as in Linka et al. (2020), GPS traces from phones (Badr et al., 2020; Linka et al., n.d.), or distance traveled and visitations to common places, and log number of interpersonal encounters in McKee et al.'s (2020). Google and Apple's location services are standard in the 25 reviewed studies. Social media is also used for the analysis of mobility and COVID-19. For example, a study about South Carolina (Zeng et al., 2021) measured population mobility by the number of Twitter users traveling more than 0.5 miles in South Carolina. One limit of this data is that it does not present those who do not have Twitter accounts or do not use Twitter, as Twitter users tend to be younger (Statista, 2021).

Out of the 25 findings, ten were studied in the U.S., four in Europe, five in Asia (China, India, Israel), one in Australia, and another in Mexico. A few studies analyzed the relationship between human mobility and Covid-19 cases in multiple countries. The prior works look at the relationship between human mobility and COVID-19 cases at the top-down levels (country, state, county, or province). However, it is essential to see how the relationship evolves from the bottom-up (city, subregion) so that a complete insight into human mobility – COVID-19 interaction can contribute further to the effectiveness of local COVID-19 control policy. In addition to the temporal dimension of the relationship, we also need to look at its spatial dimension to examine whether a geographic spillover effect exists among neighborhood areas for local policy purposes.

CHAPTER 3

STUDY AREA

3.1 San Diego

San Diego is in California's southwestern corner and the far most southwestern corner of the U.S. It borders Tijuana, Mexico, on the South, Orange County and Riverside County on the North, Imperial County on the East, and the Pacific Ocean on the West. Figure 3.1 presents 41 sub-regional areas (SRAs) of the County of San Diego, CA.

Outside the metro areas, the rest of the county is primarily rural (Figure 3.2).

Because of its proximity to Mexico's long border and the Pacific Ocean, San Diego County is a center of international trade, with the Port of San Diego as a commercial port. As a result, the county's economy centers on tourism and international trade, besides military, defense, research, and manufacturing.

San Diego County has a diverse racial and ethnic composition: 42.8% non-Hispanic white, 30.3% Hispanic, 16.7% Asian, 6.4% Black or African American, and 26.1% foreign-born persons (U.S. Census Bureau, n.d.). With a total of 3.3 million residents, San Diego County ranks 5th in the list of the most populous counties in the U.S. and the second most populous in California. According to Statistical Atlas, its population density ranks 43rd in 3,006 U.S. counties and 9th in 58 Californian counties. The high population density is one condition for a virus to spread quickly (Wong & Li, 2020).

3.2 Key Facts about COVID-19 in San Diego

The COVID-19 pandemic started in San Diego County when Gov. Newsom declared a state of emergency on March 04, 2020, and the County reported the first local case on March 09, 2020. The first San Diegan died from COVID-19 on March 22.

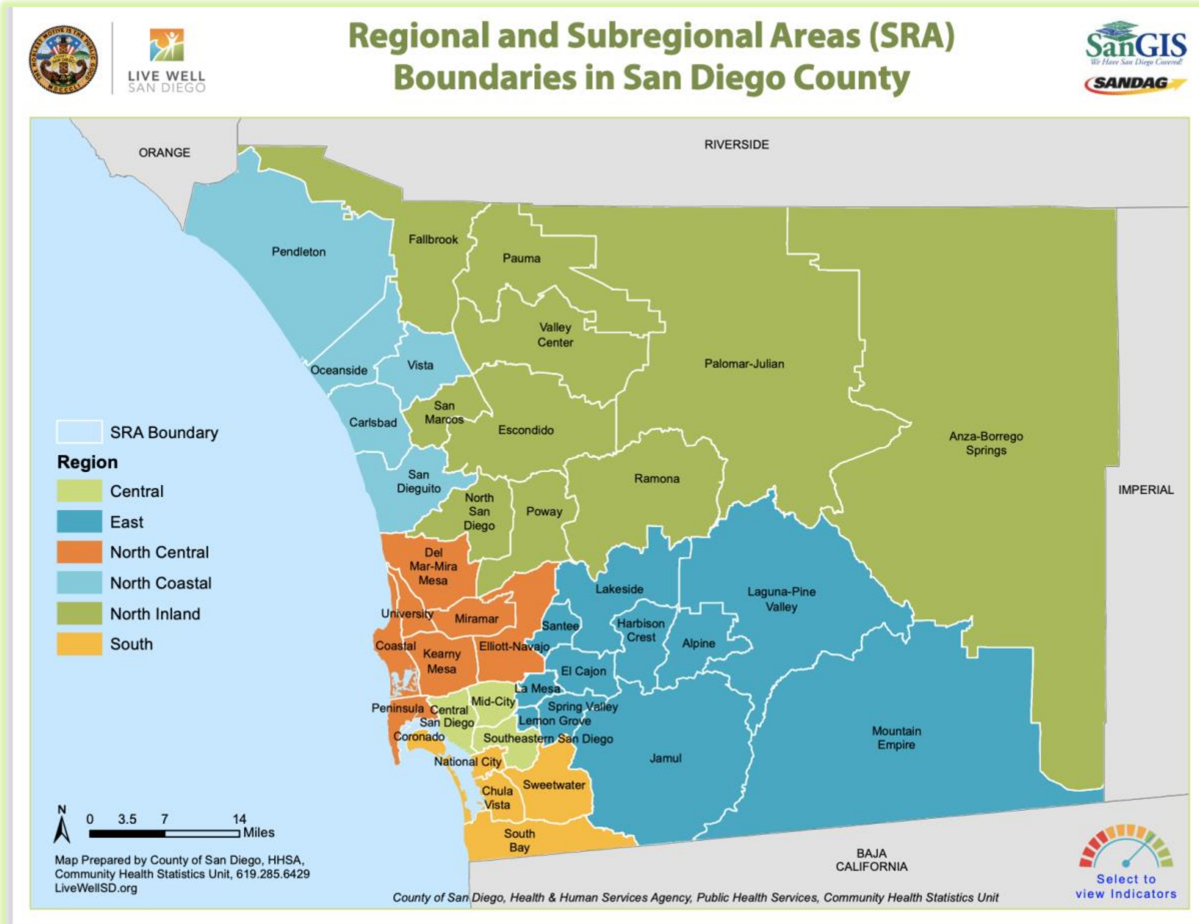


Figure 3.1. San Diego subregional areas (SRAs) (Live Well San Diego, 2018).

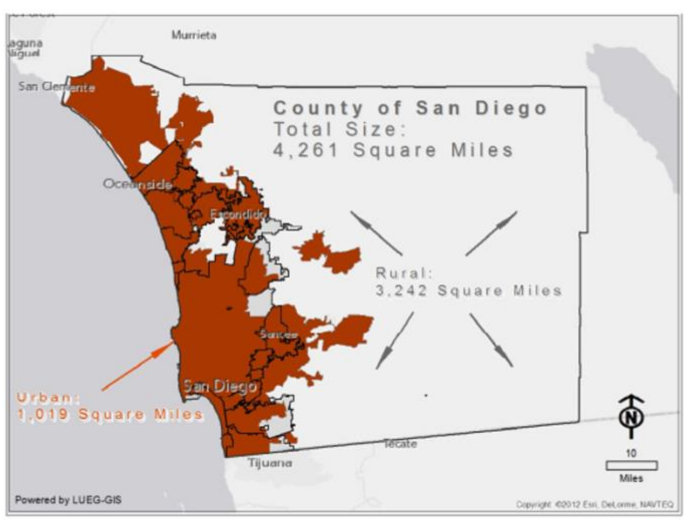


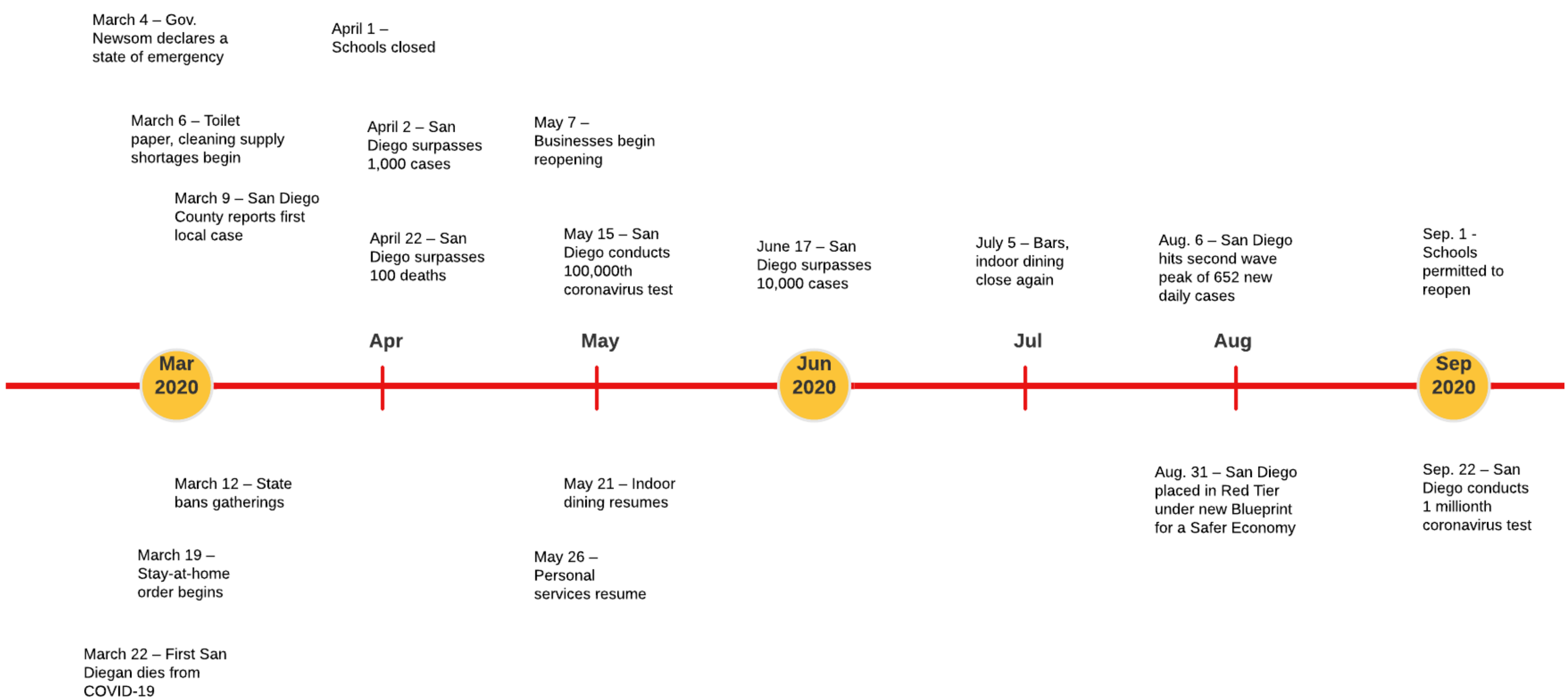
Figure 3.2. Urban vs. rural area in San Diego (Halverstadt, 2013).

On March 12, 2020, California state banned gatherings and imposed the stay-at-home order, followed by the school closure on April 01. Businesses began reopening, and indoor dining and personal services resumed in May 2020. However, the confirmed and death cases increased from thousands to tens of thousands from April to June 2020. The County was marked as Red Tier in the Blueprint for a Safer Economy, implemented by California State in August 2020. (The Blueprint has four scales: Purple, Red, Orange, and Yellow – in the descending order of the COVID spread levels in each County) (covid19.ca.gov, 2022).

Bars and indoor dining closed again in July 2020; the policy was then relaxed when the schools reopened in September of the same year. However, the situation worsened when San Diego moved to the Purple Tier of the Blueprint in November (passing 1,000 deaths and surpassing 100,000 cases in December). As a result, the policy was tightened again in November by the nighttime curfew and placing regional stay-at-home orders again in December 2020.

Pfizer-BioNTech COVID-19 Vaccine was approved by the U.S. Food and Drug Administration (2021) on December 11, 2020. It became available for people 16 years and older, bringing hope and solution for COVID-19 prevention. The first San Diego civilian got the COVID-19 vaccine shot on December 15; the first Vaccination Super Station was opened at Petco Park on January 11, 2021, to administer to 5,000 people per day. A week later, San Diego got 100,000 vaccinations. And by March 05, 2021, the number of San Diegans vaccinated surpassed one million. While pushing vaccination efforts, the County relaxed the policy to control human contact and ended the curfew and stay-at-home order on January 25, 2021. Unfortunately, the COVID-19 case number kept rising, surpassing 200,000 cases and 2,000 deaths in January, then 3,000 in March 2021.

Figure 3.3 is created to summarize the key events related to COVID-19 in a timeline based on the information provided by Lewis (2021).


March 4 – Gov. Newsom declares a state of emergency
March 6 – Toilet paper, cleaning supply shortages begin
March 9 – San Diego County reports first local case
April 1 – Schools closed
April 2 – San Diego surpasses 1,000 cases
April 22 – San Diego surpasses 100 deaths
May 7 – Businesses begin reopening
May 15 – San Diego conducts 100,000th coronavirus test
June 17 – San Diego surpasses 10,000 cases
July 5 – Bars, indoor dining close again
Aug. 6 – San Diego hits second wave peak of 652 new daily cases
Sep. 1 - Schools permitted to reopen
Mar 2020
Apr
May
Jun 2020
Jul
Aug
Sep 2020
March 12 – State bans gatherings
March 19 – Stay-at-home order begins
March 22 – First San Diegan dies from COVID-19
May 21 – Indoor dining resumes
May 26 – Personal services resume
Aug. 31 – San Diego placed in Red Tier under new Blueprint for a Safer Economy
Sep. 22 – San Diego conducts 1 millionth coronavirus test

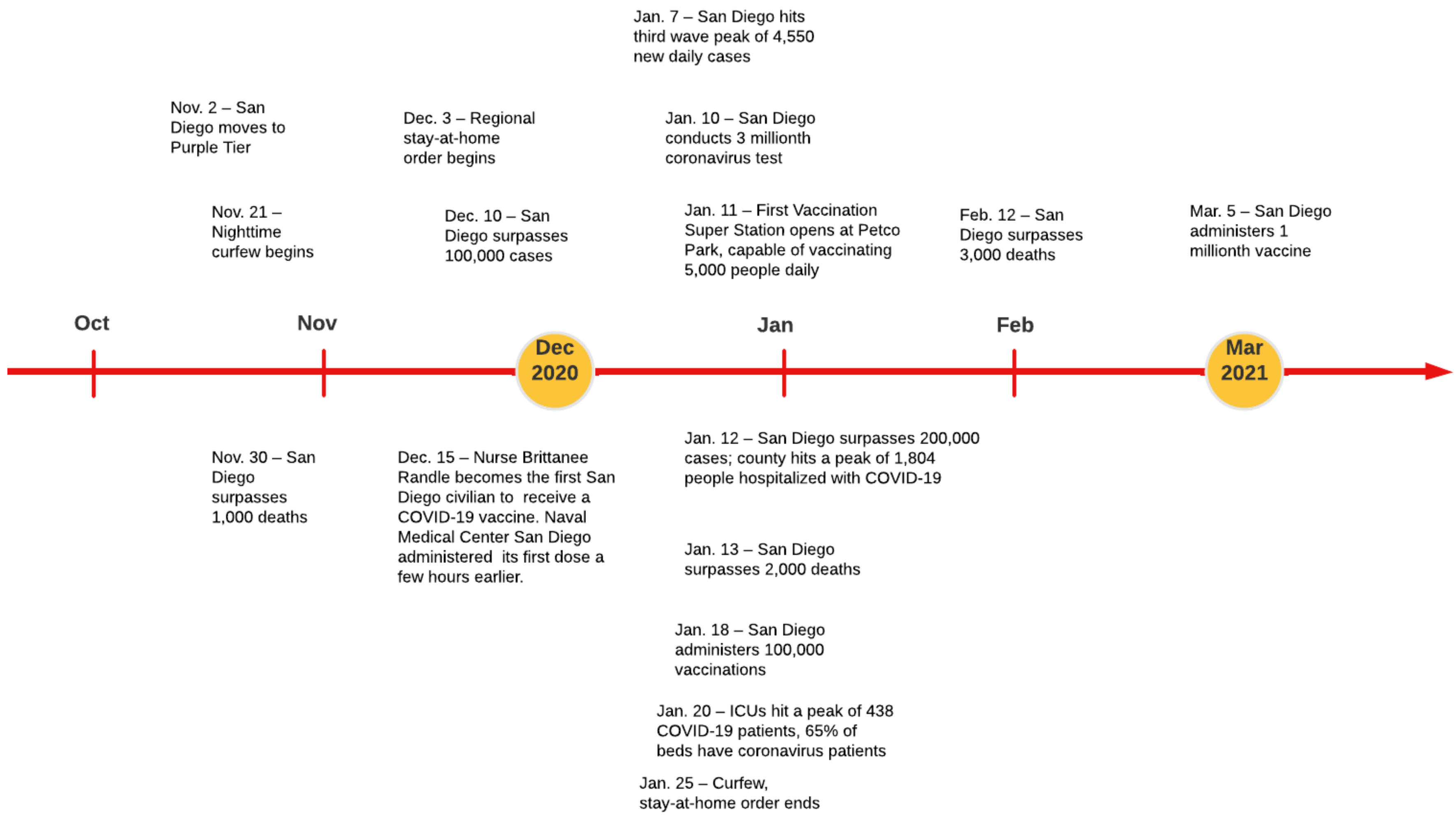


Figure 3.3. A timeline of events relating to COVID-19 in San Diego.

CHAPTER 4

DATA

This chapter introduces two original datasets – human mobility and COVID-19 confirmed new case numbers in San Diego – and how the data are processed, visualized, and combined to prepare for this study's analysis.

Figures 4.1 and 4.7 show the data structures containing San Diego SRA names, dates, values of different flows of human mobility, and COVID-19 case numbers. Figures 4.2-4.5, 4.8, and 4.9 visualize the weekly average human mobility and COVID-19 case time series of each San Diego SRA.

4.1 Human Mobility Data

The human mobility data is collected from the Social Distancing Metrics data on https://www.safegraph.com/, from January 01, 2020, to April 16, 2021. The metrics are generated using a panel of GPS pings from anonymous mobile phones, calculating the mobile device number every day with three day lag from the actual date.

Census Block Group (CBG) is a unit used by the United States Census Bureau to publish sample data. CBG has a population of 600 to 3,000 people, more extensive than a Census Block and smaller than a Census Track. A CBG on SafeGraph dataset is represented by a unique 12-digit FIPS code. Any CBG that a device stayed during the nighttime (6 pm - 7 am) in the prior six weeks is considered the "home" CBG of that device. SafeGraph aggregates the home CBG devices and counts CBGs with at least five devices (in a feature named "device_count"). A destination CBG is the CBG that a device stops by longer than a minute during the considered date range (one day, for example). The "destination_cbgs" feature is in JSON {string: integer} format, in which the string (key) is a destination CBG, and the integer (value) is the number of devices originating from a home CBG to a destination CBG. SafeGraph applies differential privacy to device count metrics to share group patterns and protect the dataset's private information. The HDMA aggregated the device numbers from CBG to SRA levels covering the period from January 01, 2020, to April 16, 2021 (Figure 4.1).

1	sra_id	sra_name	date	inflow	in_sd	in_non_sd	outflow	out_sd	out_non_sd	withinflow	total_in_within	total_out_within
2	1	CENTRAL SAN DIEGO	1/1/20	13927	8287	5640	5800	3612	2188	8116	22043	13918
3	1	CENTRAL SAN DIEGO	1/2/20	14490	9505	4985	5957	4166	1791	8059	22549	14017
4	1	CENTRAL SAN DIEGO	1/3/20	16266	10852	5414	6476	4647	1829	8537	24803	15015
5	1	CENTRAL SAN DIEGO	1/4/20	14882	9466	5416	6126	4251	1875	8325	23207	14451
6	1	CENTRAL SAN DIEGO	1/5/20	12132	8152	3980	5574	3873	1701	8443	20575	14017
7	1	CENTRAL SAN DIEGO	1/6/20	12466	9352	3114	5899	4509	1390	8302	20768	14203
8	1	CENTRAL SAN DIEGO	1/7/20	12868	9976	2892	5847	4578	1269	8280	21148	14127
9	1	CENTRAL SAN DIEGO	1/8/20	13191	10025	3166	6078	4751	1327	8283	21474	14362

Figure 4.1. San Diego human mobility data.

The inflow of an SRA reflects the number of mobile devices that have home CBG in that SRA and the destination CBG outside that SRA, including other U.S. states, other counties in California, or other SRAs in San Diego County. In other words, the inflow numbers represent the human mobility flow from other places (inside the U.S. and outside the SRA) to that SRA.

Similarly, the outflow of each SRA represents the human mobility flow coming from that SRA to other SRAs in San Diego County, other counties in California or other states in the U.S. Withinflow is the human mobility moving inside that SRA. The netflow is calculated by subtracting the outflow from the inflow in the dataset. Inwithin flow is the combination of inflow and within flow.

The in_sd, in_non_sd, out_sd, and out_non_sd features drill down further the amount of human mobility to separate the flow from inside and outside San Diego County. However, these features are not used under the scope of this study.

Each type of human mobility – inflow, netflow, outflow, withinflow, inwithinflow – is converted to a seven-day rolling average. The minimum number of observations in the window required to have a value is 1 to avoid NaN values (Figures 4.2-4.6).

Central San Diego, Carlsbad, Coastal, Kearny Mesa, and Peninsula have the highest and most fluctuated inflows among all 41 SRAs (Figure 4.2). The human mobility outflows in Sweetwater, Southeastern San Diego, Carlsbad, Central San Diego, Del Mar - Mira Mesa, El Cajon, Escondido, Kearny Mesa, Mid-City, Oceanside, North San Diego have the most ups and downs (Figure 4.3). Netflow has a positive value when inflow is more significant

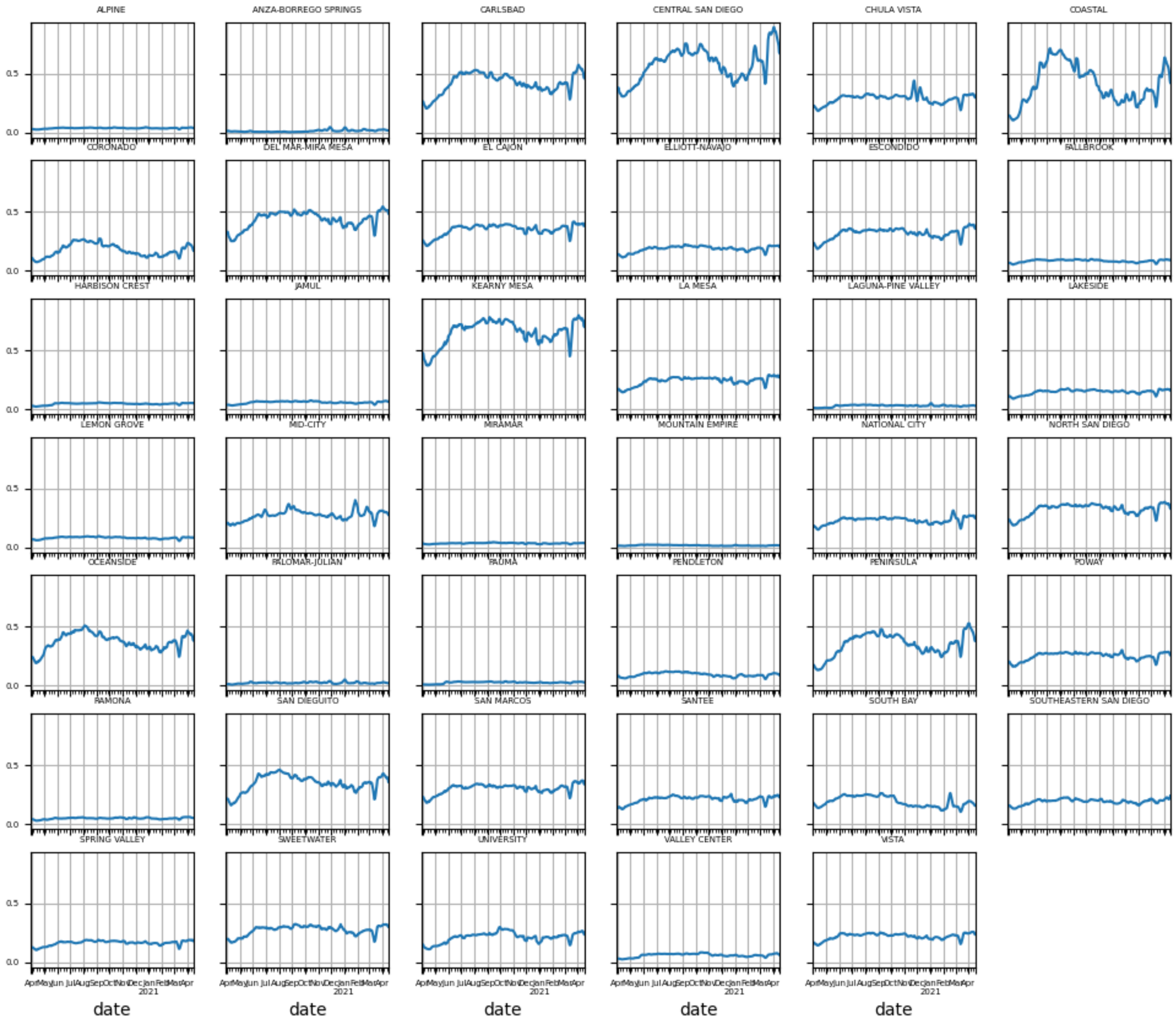


Figure 4.2. Weekly average human mobility inflow of 41 San Diego SRAs.

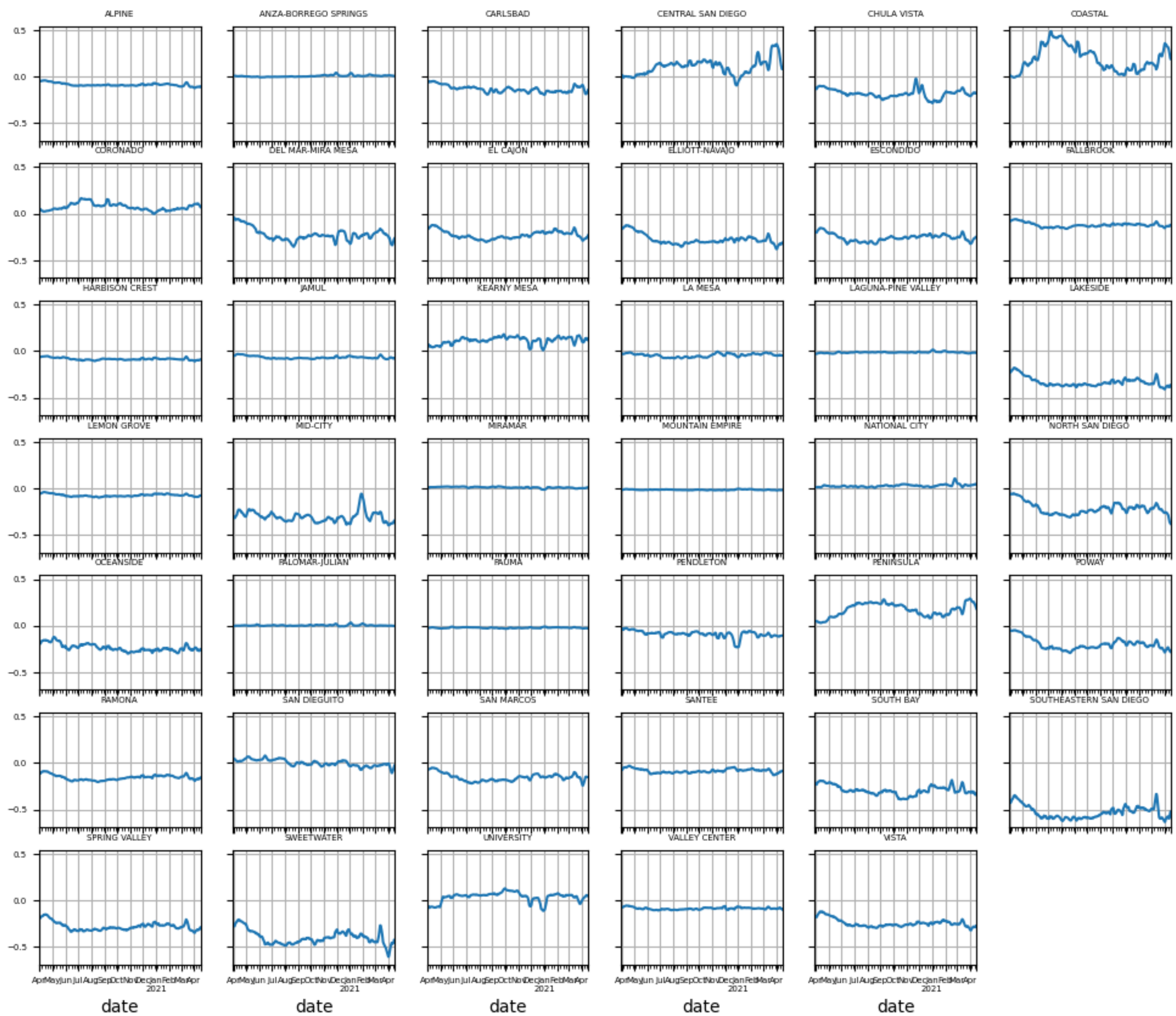


Figure 4.3. Weekly average human mobility netflow of 41 San Diego SRAs.

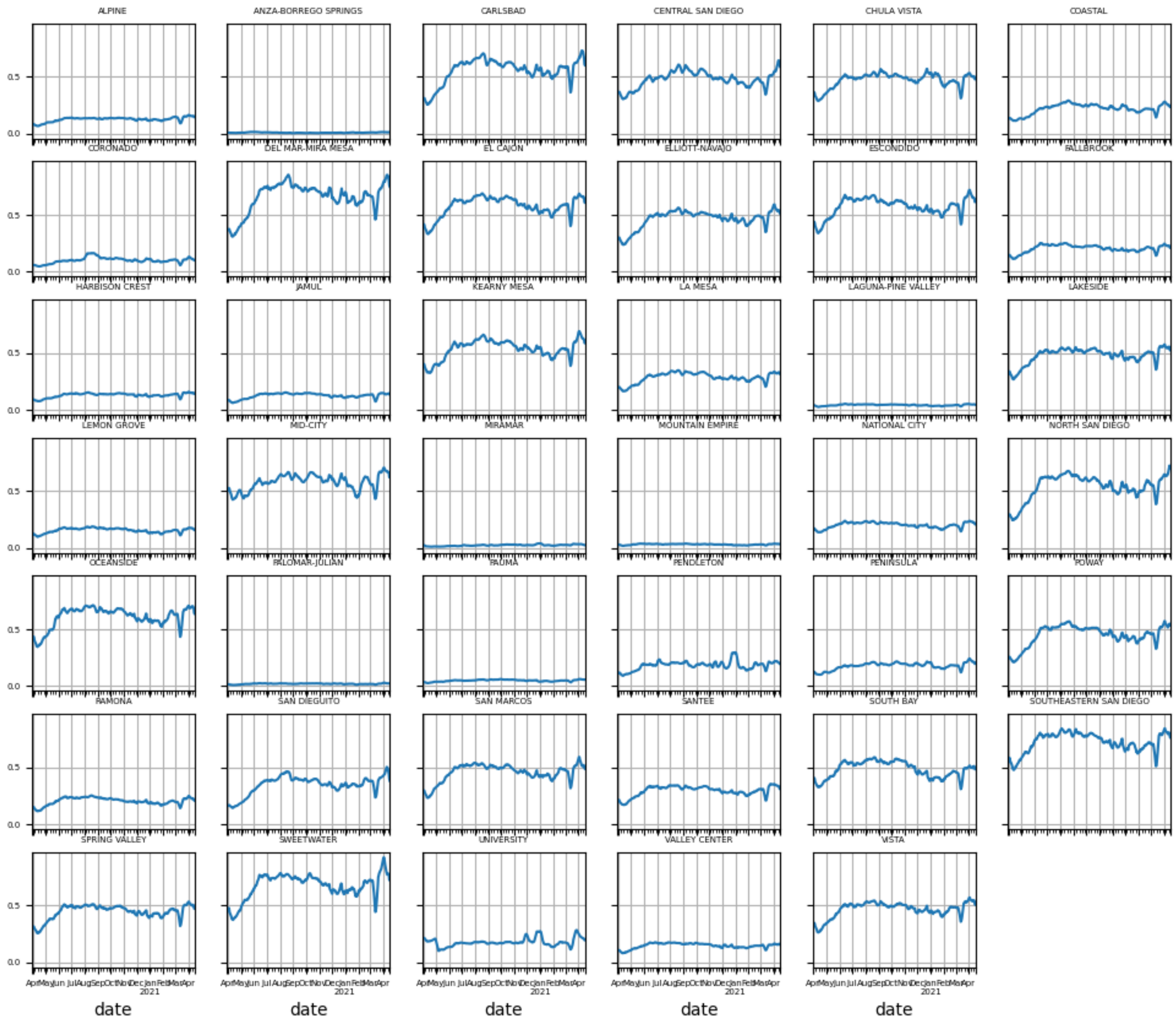


Figure 4.4. Weekly average human mobility outflow of 41 San Diego SRAs.

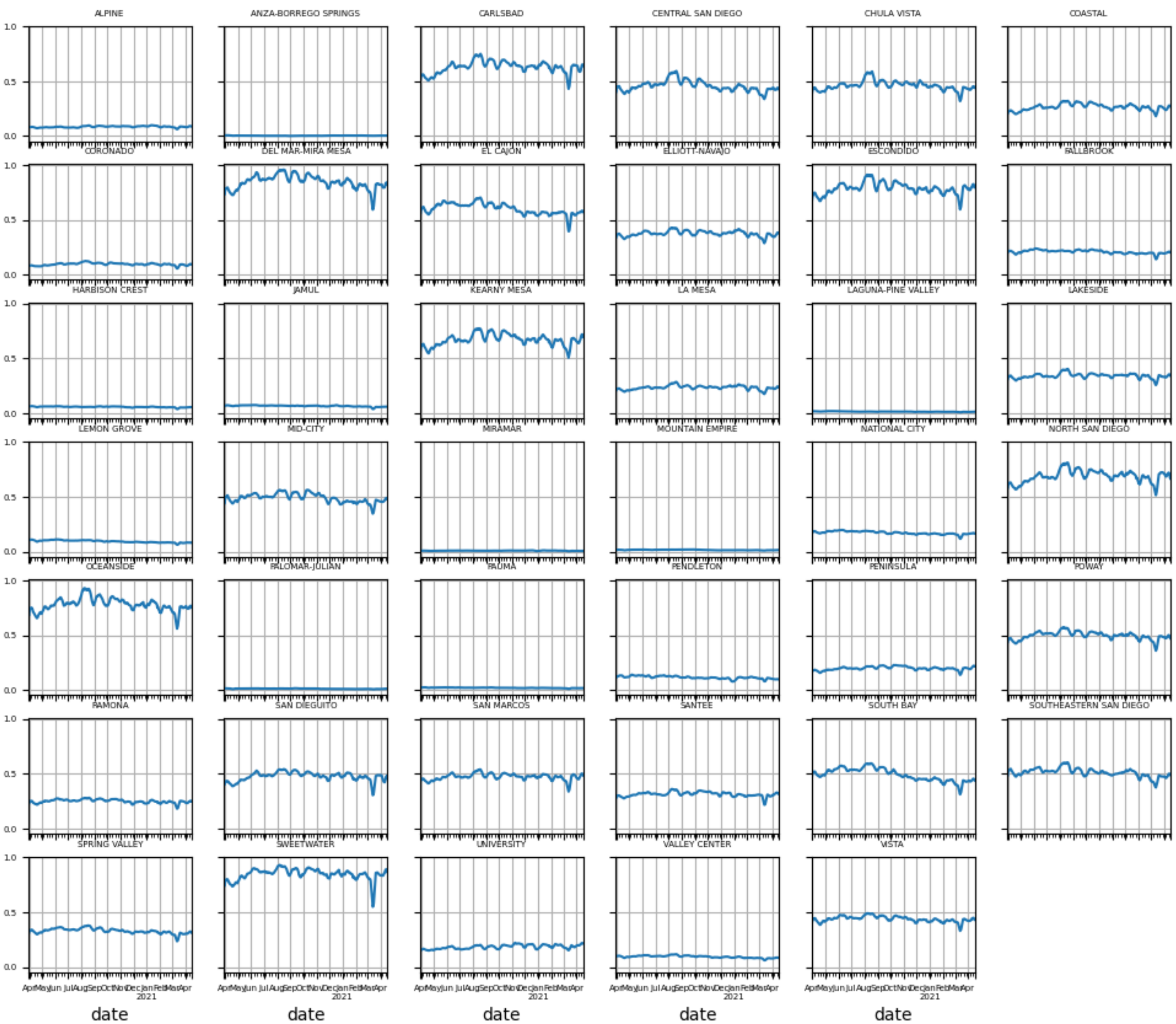


Figure 4.5. Weekly average human mobility withinflow of 41 San Diego SRAs.

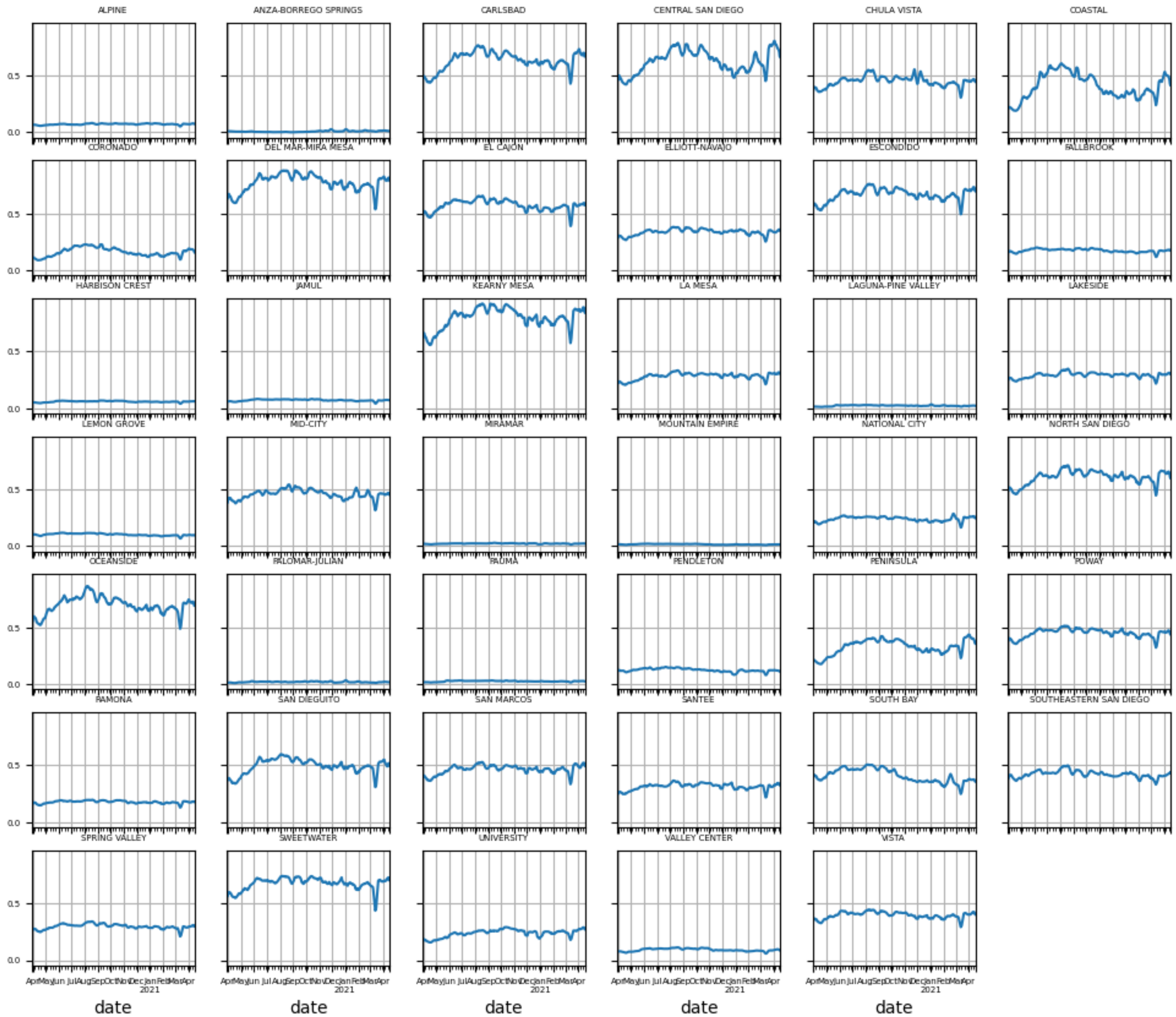
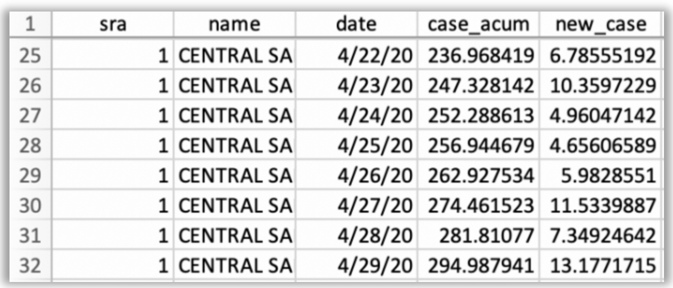


Figure 4.6. Weekly average human mobility total-in-within flow of 41 San Diego SRAs.

than outflow and becomes negative when inflow is lower than outflow. After applying min-max normalization, negative netflow can be pushed down towards the low end of 0.0 while the positive netflow is pushed towards the high end of 1.0 since the range from minimum negative to maximum positive values becomes bigger. The inflow is mostly canceled out with outflow in calculating the netflow, so the ups and downs in netflow reduce significantly compared to inflow and outflow. Figure 4.4 shows that Del Mar – Mira Mesa, Escondido, North San Diego, Oceanside, and Sweetwater have much more inflow than outflow in San Diego. Regarding the withinflow (Figure 4.5), Del Mar – Mira Mesa, Escondido, Kearny Mesa, Oceanside, Sweetwater also have the highest values, just like their inflows and total-in-within flows (Figure 4.6).

4.2 COVID-19 Data

The COVID-19 confirmed cases are provided by the San Diego Health and Human Services Agency (County of San Diego, 2022), from March 30, 2020, to April 16, 2021 (Figure 4.7). The HDMA estimated the number of cases at SRAs from the data at zip codes using the Dasymetric Mapping method.



1	sra	name	date	case_acum	new_case
25	1	CENTRAL SA	4/22/20	236.968419	6.78555192
26	1	CENTRAL SA	4/23/20	247.328142	10.3597229
27	1	CENTRAL SA	4/24/20	252.288613	4.96047142
28	1	CENTRAL SA	4/25/20	256.944679	4.65606589
29	1	CENTRAL SA	4/26/20	262.927534	5.9828551
30	1	CENTRAL SA	4/27/20	274.461523	11.5339887
31	1	CENTRAL SA	4/28/20	281.81077	7.34924642
32	1	CENTRAL SA	4/29/20	294.987941	13.1771715

Figure 4.7. San Diego COVID-19 confirmed case numbers.

The case_acum is the accumulated COVID-19 case number on each day of each SRA. For privacy reasons, the daily new_case is not provided by the County, and so it is calculated by subtracting the previous day's case_acum from a day's case_acum. Negative values appear when the last day's case_acum is bigger than today's case_acum. For this study, all negative and NaN values are converted to zero. The missing days are filled up with the values of the previous days. This study also calculates seven day rolling average for the daily confirmed Covid-19 new case numbers (Figures 4.8 and 4.9)

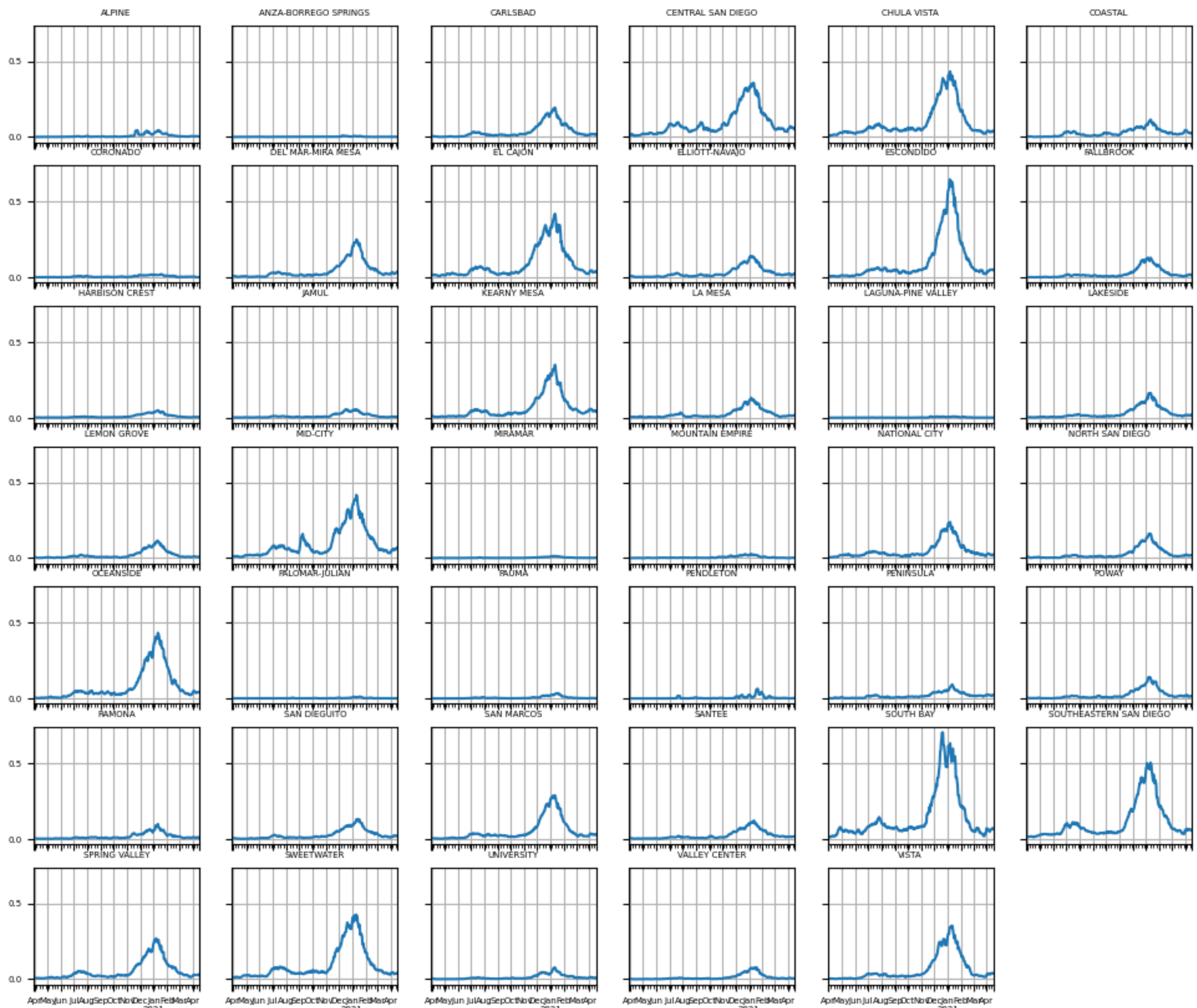


Figure 4.8. Weekly average COVID-19 confirmed new case number of each San Diego SRA.

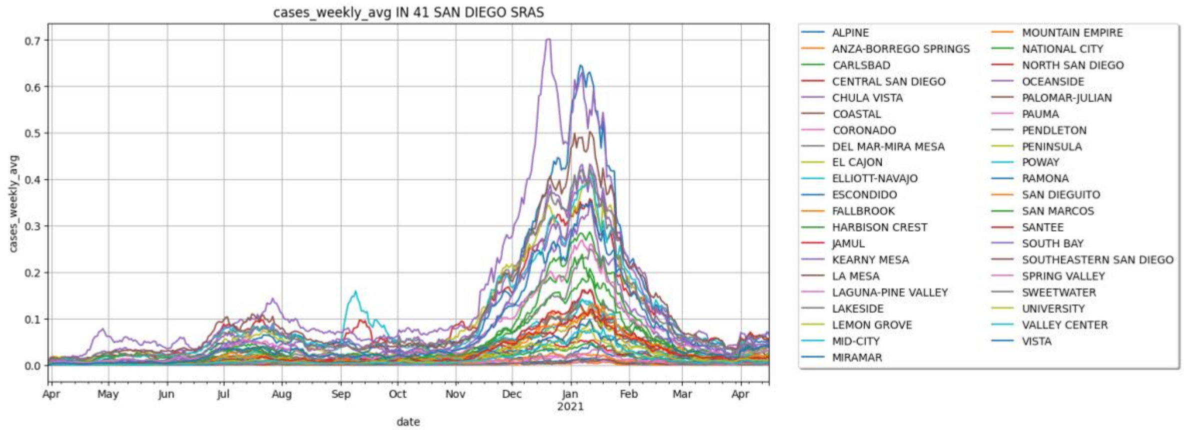


Figure 4.9. Weekly average COVID-19 confirmed new case numbers of all 41 San Diego SRAs.

In Figure 4.8, the SRAs that have the highest peaks of COVID-19 case numbers include Escondido, South Bay, Chula Vista, El Cajon, Mid-City, Southeastern San Diego, Sweetwater, and Oceanside. When putting all COVID-19 case time series of 41 SRAs on the same chart (Figure 4.9), San Diego had a high peak of COVID-19 cases between December 2020 and February 2021. Another peak of COVID-19 cases was in July-August 2020. All 41 time series share similar trends overall, but every time series is different in terms of shapes and maximum points.

4.3 Combine the Two Datasets

Human mobility and COVID-19 data are combined from April 01, 2020, to March 01, 2021. Then, the merged data is normalized by rescaling every feature to the [0, 1] interval (min-max normalization). Min-max normalization is better than standardization (or z-score normalization) in this case for these reasons:

1. Data need to be divided in time or place to check normal distribution before applying the standardization.
2. When the data distribution is unknown, it's safer to use the min-max normalization for data across time and place.
3. Machine learning estimators might misbehave if individual features look different from standard normally distributed data (e.g., Gaussian with 0 mean and unit variance).

CHAPTER 5

METHODOLOGY

This chapter explains how this study analyzes the processed data to achieve the final results. First, it explains the Python programming language utilized in this study. Then it describes the Dynamic Time Warping (DTW) method to compare the similarity between time series and how calculating the slopes of the time series can help validate the DTW method results. Finally, it explains why the 14-day approach is used instead of dividing the data into phases for this study's analysis.

5.1 PROGRAMMING LANGUAGE

Python scripting is used for all the data cleaning, processing, analysis, and visualization.

Standard Python libraries/packages/modules used in this study include Pandas, Geopandas, NumPy, Tslearn (Python package to provide machine learning tools for time series analysis), Scikit-learn (Python module for machine learning), SciPy (open-source software for mathematics, science, and engineering), math, time, datetime, Plotly, Matplotlib, dataframe_image, PYL, Glob, and functools.

5.2 DYNAMIC TIME WARPING (DTW)

To find out when and where the correlation between human mobility and COVID-19 exist in San Diego, this study utilizes DTW to measure the similarity between each human mobility flow and the COVID-19 case time series at the SRA level.

DTW is a standard method for analyzing temporal sequences of finance, audio, video, and graphics data, as well as speech and word recognition (Portilla & Heintz, 2019). The traditional Euclidean method works well for time series having the same length and moving at the same speed and time. Compared with it, the DTW has a better advantage when comparing the two sequences with similar patterns but are out of sync, non-linear, and have different lengths. In the DTW method, multiple points are matched from one line with a point in the other line or vice versa to sync head-to-head and tail-to-tail and match the troughs and peaks between the heads and tails of two lines. The other points between the sequence's heads, tails, lows, and highs, warp accordingly, and thus the two lines match optimally on the map.

To get the distance between the time series, the following warping function (Equations 5.1 and 5.2) finds the best alignment through the grid (warping path) and minimizes the total distance between the points of the lines. The sum of the minimum distances is the similarity between the two sequences.

$$\mathrm{RD}(x_i, y_j) = \mathrm{d}(x_i, y_j) + \min\begin{cases} RD(x_{i-1}, y_{j-1}) \\ RD(x_{i-1}, y_j) \\ RD(x_i, y_{j-1}) \end{cases} \tag{5.1}$$

$\mathrm{RD}(x_i, y_j)$ is the minimal cumulative distance from (0, 0) to (i, j) in matrix D.

$$\mathrm{DTW}(X, Y) = \min\{RD(x_n, y_m)\} \tag{5.2}$$

A distance measure is achieved for the similarity between the human mobility sequence and the Covid-19 case sequence for each San Diego SRA. DTW Distance is explained in Figure 5.1.

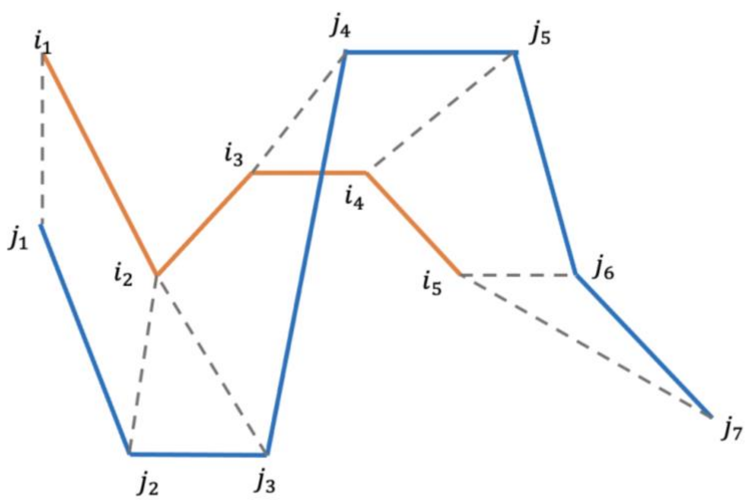


Figure 5.1. DTW Distance between two time series (modified from Alizadeh, 2020)

5.3 Slope calculation

DTW determines which SRAs have more similarity between mobility and Covid-19 cases at which time. Lower DTW distance implies more similarity and vice versa. There is no DTW value cutoff to decide whether the two time series are correlated. Therefore, the slope calculation confirms the correlation between COVID-19 and human mobility.

For those SRAs with low DTW index values, the slopes of the two time series are then compared to examine the increase-increase relationship. The positive slopes of the two time series means that both human mobility and COVID-19 tend to move in the same direction, and the increase in the human mobility flow is positively correlated with the rise of the COVID-19 case number. The DTW method and the slope examination validate the relationship between human mobility and COVID-19 cases in San Diego.

5.4 The Phase Approach and K-Means Clustering for Time Series

To determine whether human mobility and COVID-19 case are correlated in different pandemic phases, the processed data is first divided into phases to apply the DTW and slope calculation methods. The processed data, which range from 04-01-2022 to 03-01-2022, is tried to be divided twice, one time into 5 phases including the peaks (Figure 5.2), and another time into 6 phases (Figure 5.3).

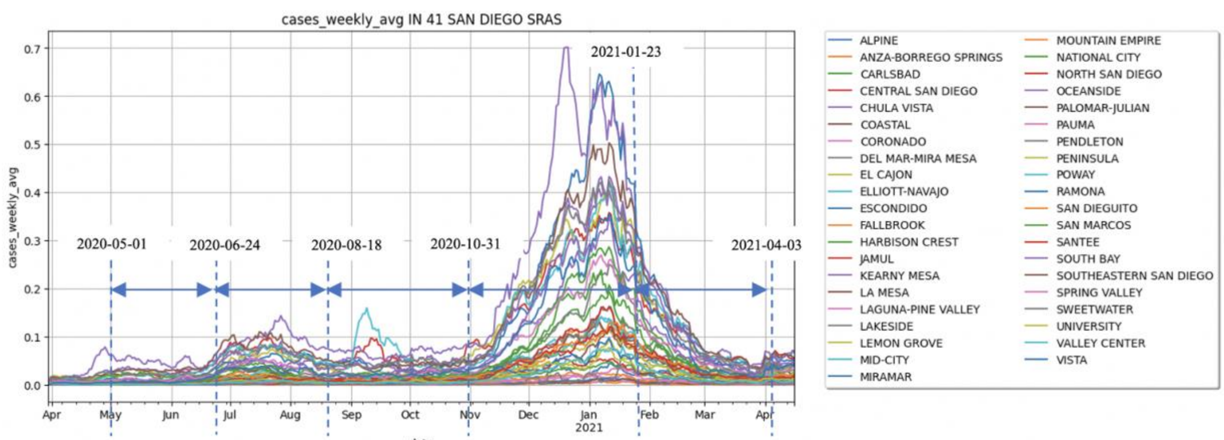


Figure 5.2. COVID-19 data divided into 5 phases including the peaks of the pandemic.

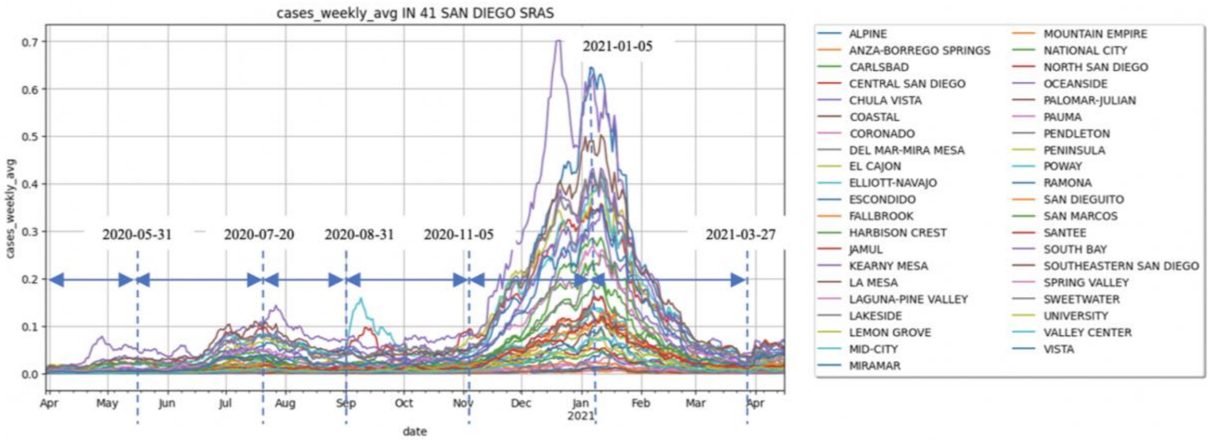


Figure 5.3. COVID-19 data divided into 6 phases.

For each phase, the K-Means Clustering method groups the 41 cases and flow time series (each for each SRA) into sub-groups of similar time series. Each sub-group has a DTW Barycenter Averaging (DBA) line, which is the cluster centroid of that sub-group. Depending on the slope values of each DBA line, the corresponding sub-group can be labeled as "increasing," "decreasing," or "stable." If the slope value > 0, the sub-group is marked as "increasing." If the slope < 0, the sub-group is labeled as "decreasing." And "stable" is labeled for the sub-group when the slope value of its DBA line = 0 (rounding to 3 decimal places).

` Each group of the 41 time series has a different optimal sub-group number. Figure 5.4 shows the K being tried from value 3 to value 17 to look for the best K for the 41 SRA time series group. This step is repeated for each feature (case, inflow, outflow, withinflow, netflow, inwithin flow) in each phase to get all the optimal k values (Figure 5.5). The K value is considered optimal when each sub-group in the K sub-groups presents the most similar time series. For example, the COVID-19 case time series in phase 5 has optimally seven sub-groups, each of which is a collection of the time series having the most similar trends (Figure 5.6). The red lines are the DBA – the cluster centroid of sub-groups in these figures.

After sub-groups of the case and flow time series in each phase are labeled, the individual time series of that sub-group are marked accordingly. Each time series in each phase is now labeled as "increasing," "decreasing," or "stable." Those SRAs with human mobility time series and COVID-19 case time series labeled as "increasing" are selected and compared with their measured DTW values.

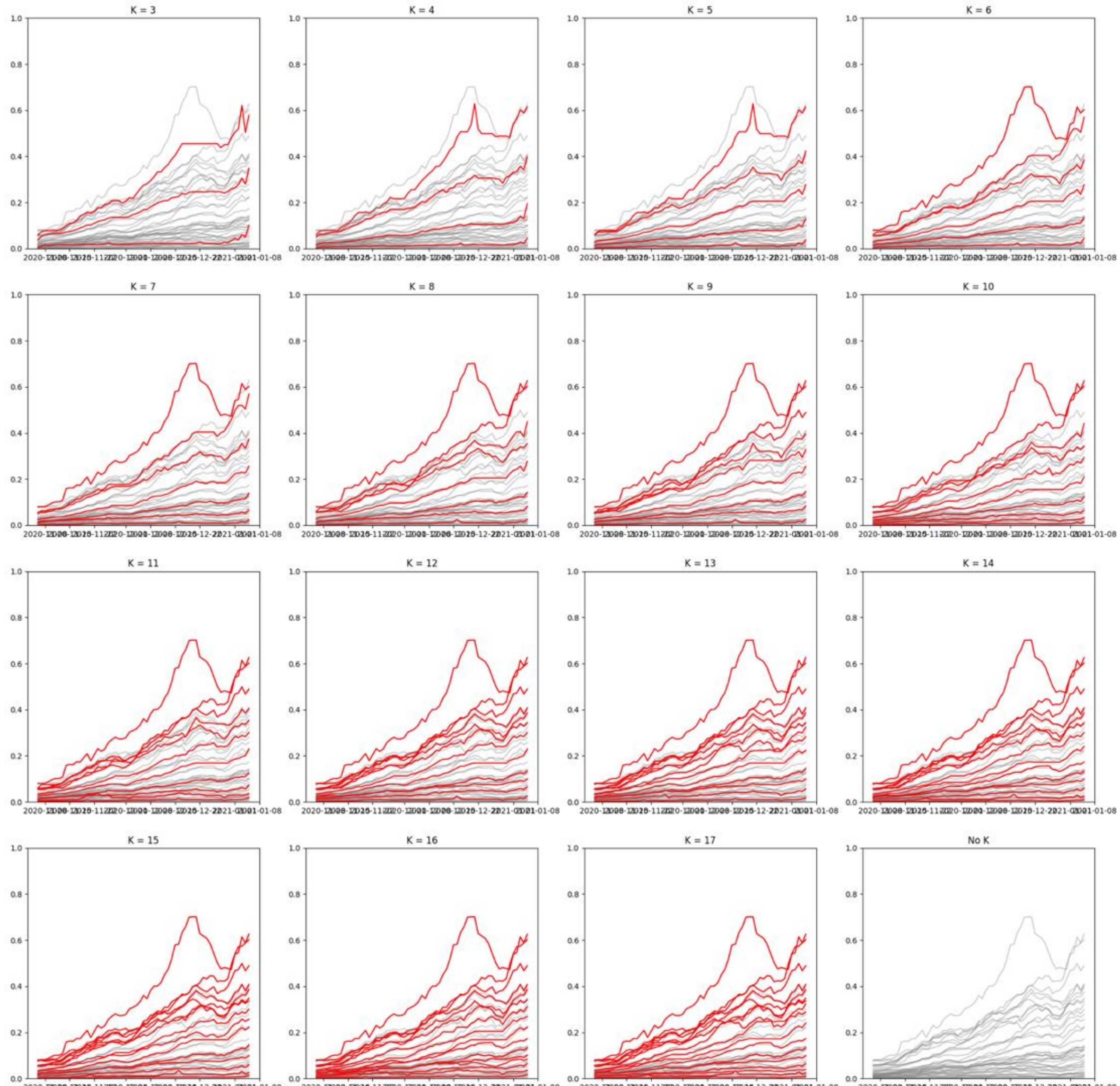


Figure 5.4. Sub-groups of 41 SRA time series group when k = [3, 17]

phase	cases_weekly_avg	inflow_weekly_avg	outflow_weekly_avg	withinflow_weekly_avg	netflow_weekly_avg	total_in_within_weekly_avg
1	3	8	8	10	10	16
2	3	15	10	13	16	13
3	3	12	11	14	13	13
4	5	11	12	14	12	10
5	7	11	11	12	15	13
6	7	10	9	11	17	16

Figure 5.5. Best K values for all features in all phases.

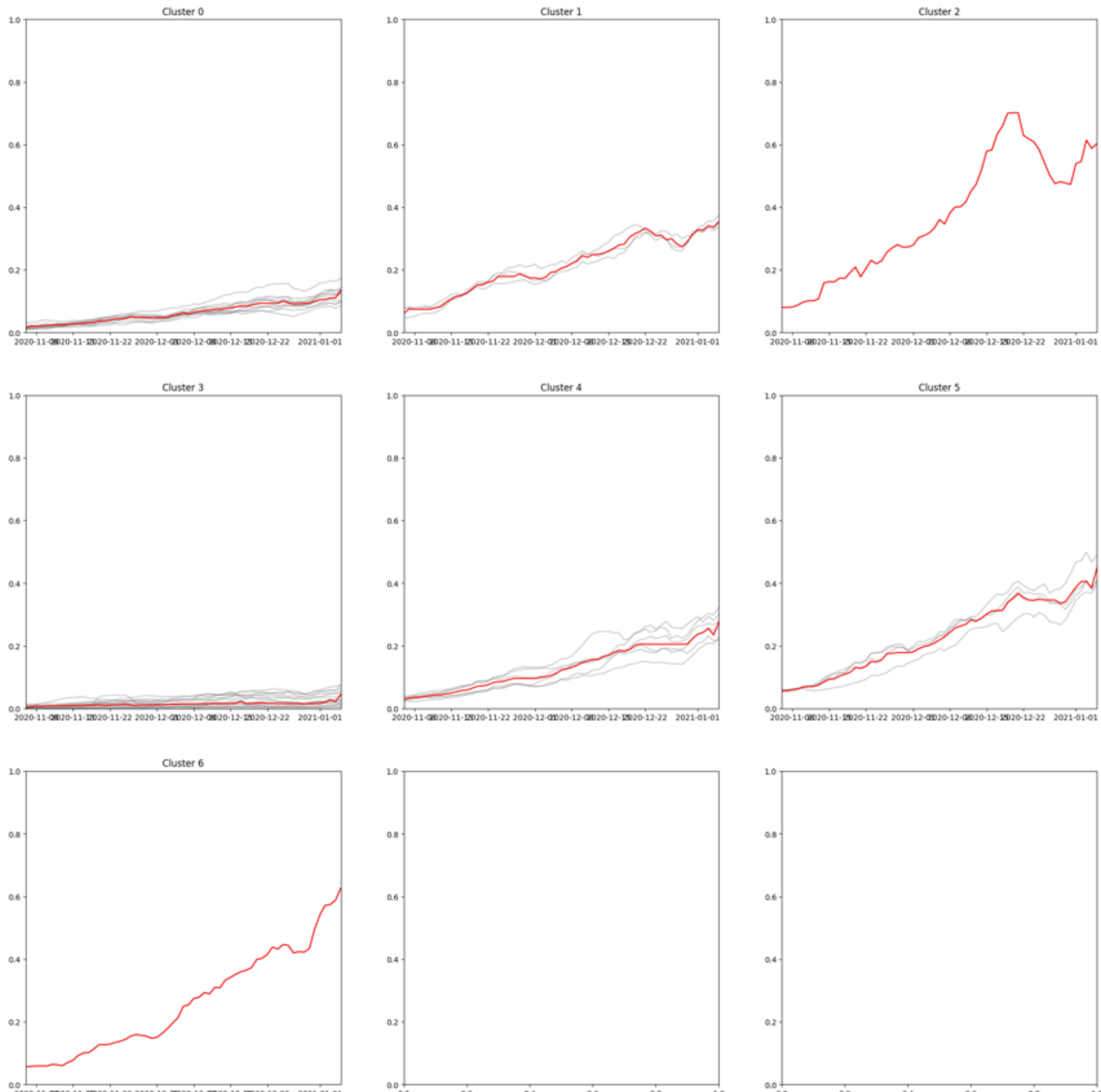


Figure 5.6. Sub-groups of COVID-19 case time series in phase 5 (out of 6 phases).

The phase approach and Time Series K-Mean Clustering approach allow looking at the time series in groups and subgroups. However, they need to give consistent results. K-means is implemented with randomized initialization, so the results change a little each time the TimeSeriesKMeans module is run. It requires running the module many times until good results are stabilized.

The phase approach has some disadvantages. It depends on how the dataset is divided into phases, how the optimal k numbers are subjectively selected, and what thresholds to use (such as the number of decimal places and intercept values to label the time series as "increase high" or "increase low"). When the data is divided into five phases, as in Figure 5.2, the peaks are included in the phases. The peaks have an up slope and a down slope, making up a V shape on top. The up and down slopes are canceled out when calculating a slope for the whole phase; thus, the increase-increase relationship is not revealed in the final results. If the data is divided into six phases, as in Figure 5.3, the division reflects the trend of COVID-19 cases. However, the corresponding division in the flow time series cannot avoid the peaks and has the slope cancellation as in the case of the 5-phase division.

Another issue is that when the cluster's value is applied to its time series, the details about when and where the correlation happens in each SRA on each day can become inaccurate. Although the human mobility movement and COVID-19 case trends can look very similar as a group/sub-group, they can look very different at minor points when zoomed in.

5.5 The 14-Day Approach

Because of the disadvantages of the phase and Time Series K-Mean Clustering method, the 14 days approach is used to examine the human mobility and COVID-19 case in this study. The idea is to look at each day of human mobility and COVID-19 cases instead of a whole phase. To increase the analysis accuracy, this study investigates the historical data of the previous 14 days of each day, applies the DTW and slope calculation on the historical data and the target day (15 days in total), and interpolates the result to the target day. By covering the past 14 days, the DTW method can cover the lag (7 - 10 days) if it exists between human mobility and COVID-19.

For each flow type of human mobility, every day of the dataset is looped over to get DTW and slope values. The dataset starts on 2020-04-01 and ends on 2021-03-01. The first 14 days are skipped. Beginning the 15th day, 2020-04-15 forward, each day has one DTW value and one slope value, measured for that day and its previous 14 days. Precisely, the 2020-04-15's DTW and slope are measured from 2020-04-01 to 2020-04-15. Similarly, 2020-04-16's DTW and slope are measured for 2020-04-02 – 2020-04-16. Similar to getting a moving average, but instead of getting the mean of the 15 days values, this 14-day approach applies the DTW method and calculates the slope for the 15-day time series. One set of DTW and slope value is returned for the last day of the 15 days. The SRAs and days are then filtered out for those having DTW values less than 1.0 and positive slopes of both human mobility and the COVID-19 case time series. These are the SRAs and days that human mobility is positively correlated with COVID-19 in San Diego.

CHAPTER 6

RESULTS

This chapter summarizes the results from the 14-day approach explained in chapter 5. At first, the initial results are collected in Supplemental Material 1 for the DTW, slope, intercept, and label values for all features (COVID-19 case, mobility inflow, mobility outflow, mobility netflow, withinflow, inwithinflow) from 2020-04-15 to 2021-02-28. Next, Supplemental Material 2 is created by filtering the rows that have DTW values < 1.0 and human mobility and COVID-19 time series having positive slopes; other values are converted to blank. After that, each of the human mobility flows has one heatmap created for the filtered data in Figures 6.1, 6.2, 6.3, 6.4, and 6.5.

Based on the color allocation in these heatmaps, "*when the positive correlation between human mobility and COVID-19 in San Diego happens*" is revealed. The original data visualization validates the answers in Figures 6.6-6.11.

To find out "*where is the positive correlation between human mobility and COVID-19*," the study counts the days each SRA has a positively correlated relationship and sorts the SRAs from the highest to the smallest count values. Those SRAs appearing at the top of the list is where the relationship exists the most. The choropleth maps show where those places are located on a map. The color shade of the choropleth maps is based on the count numbers and can explain the spillover effect of the neighboring SRAs (Figures 6.13-6.18).

Figure 6.18 provides examples of the case and flow time series in SRAs having the most positive correlation between human mobility and COVID-19 as a way to validate the results of the "*when*" and "*where*" questions above.

Finally, this study also counts the days showing a positive correlation between mobility and COVID-19 by the types of human mobility flows, including inflow, outflow, netflow, withinflow, and inwithinflow. The statistics of the day count in Table 6.1 confirms that the netflow and inflow have more impact on San Diego COVID-19 case, while the outflow, withinflow, and inwithinflow do not have much impact.

6.1 When Does the Positive Correlation Happen

6.1.1 Netflow

The heatmap in Figure 6.1 visualizes the DTW values between the human mobility netflow and COVID-19 in San Diego. The dates that meet the two conditions (DTW values < 1.0 and positive slopes for both case and flow time series) spread out the whole period, but the darker color focuses more on June-July 2020, in the last quarter of 2020, and in January 2021.

6.1.2 Inflow

The inflow in Figure 6.2 has a significant number of cells with low DTW values in the dark orange color in December 2020 and January 2021. Other cells in orange color happen in late April, early May, mid and late June, late July, and October 2020. This explains that human mobility inflow and COVID-19 correlation in San Diego County exist throughout the year and tend to be positively correlated in winter 2020.

6.1.3 Outflow

The outflow (Figure 6.3) has a vivid positive correlation in late November and December 2020 and the first half of January 2021, but only a little in the earlier period.

6.1.4 Withinflow and Inwithinflow

Similar to the outflow, the within and inwithin flows (Figures 6.4 and 6.5) show a more positive correlation between mobility and COVID-19 in December 2020 and January 2021.

6.1.5 Validate with the Original Data Visualization

COVID-19 case number (Figure 6.6) increased slightly in the summer of 2020 then rose significantly in the winter/holiday of 2020. The trends of the COVID-19 time series and the following human mobility flows match the timing of the findings.

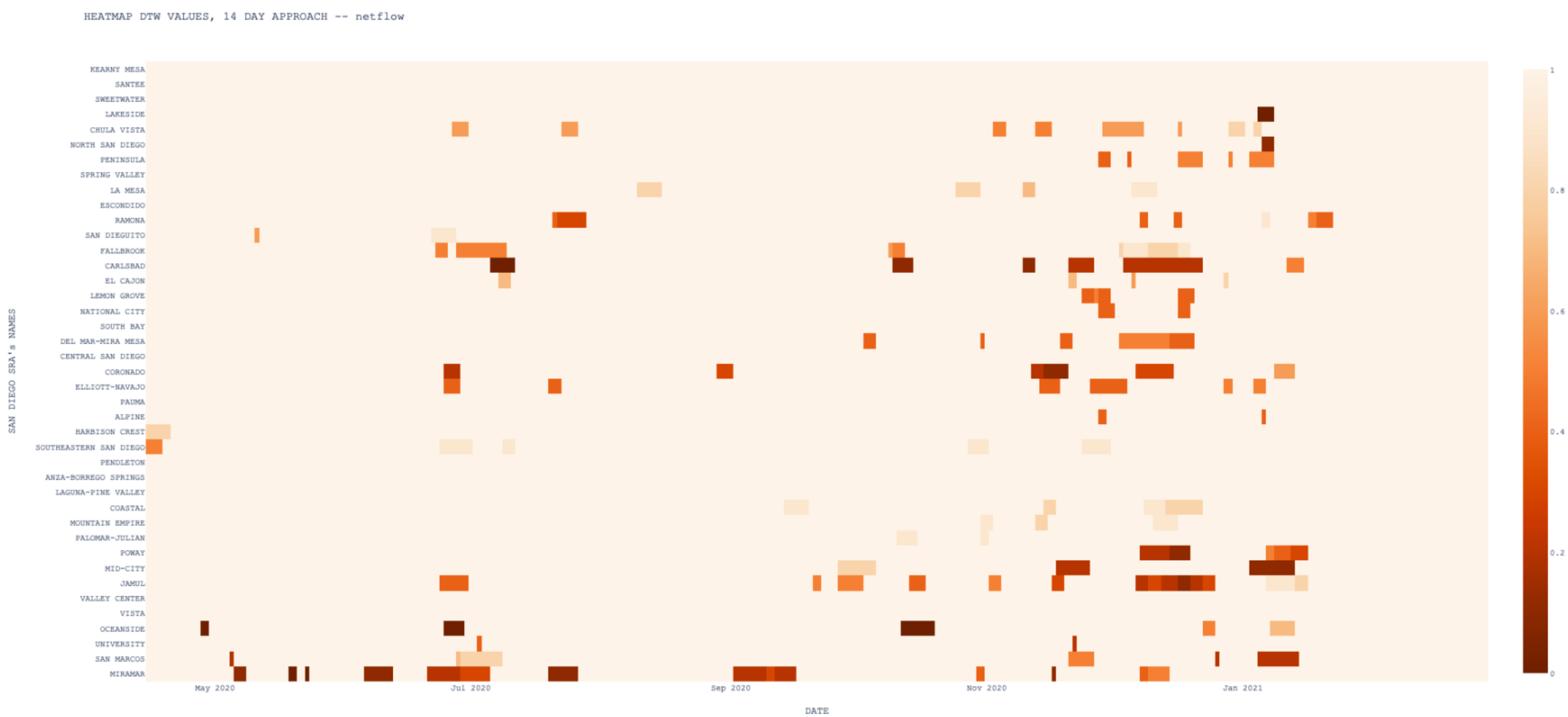


Figure 6.1. Heatmap of DTW values between human mobility netflow and COVID-19.

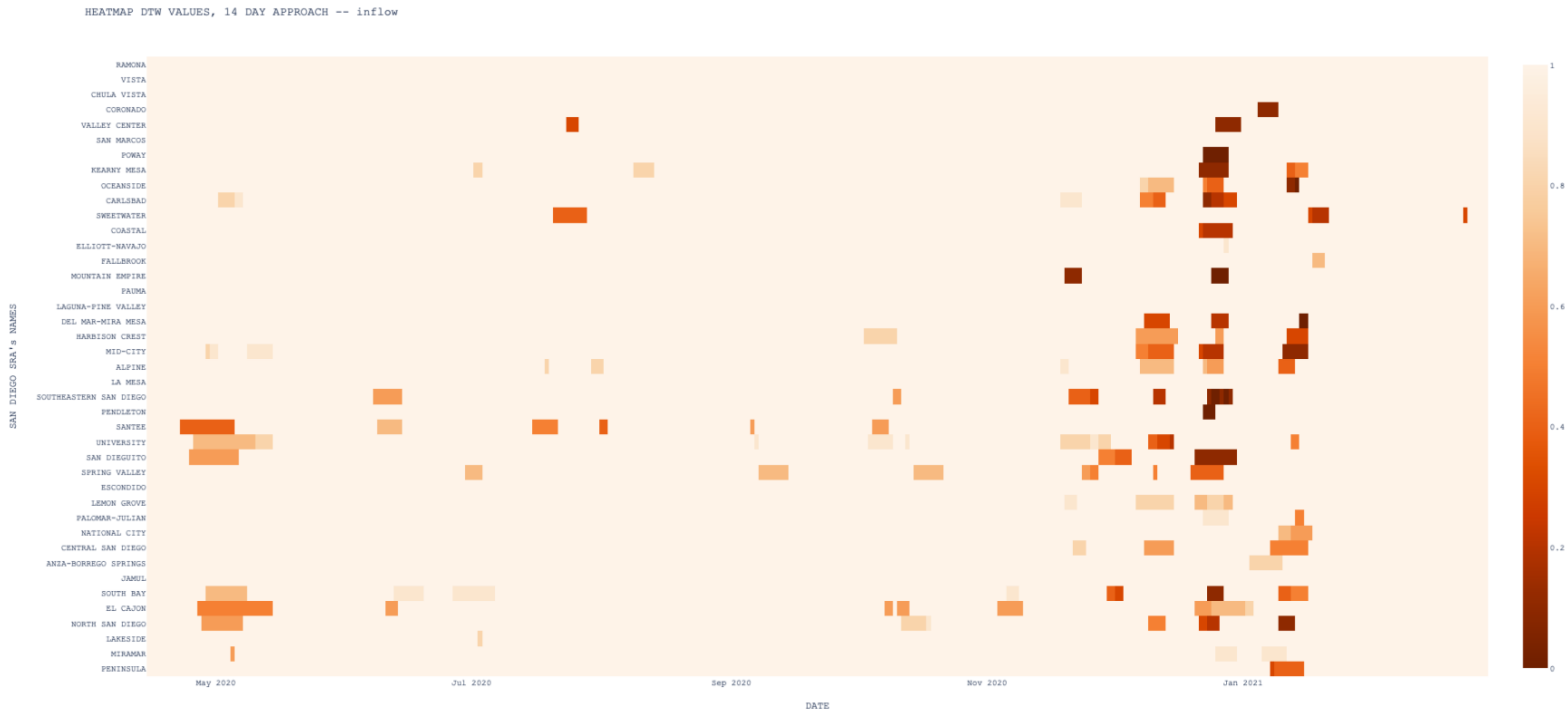


Figure 6.2. Heatmap of DTW values between human mobility inflow and COVID-19.

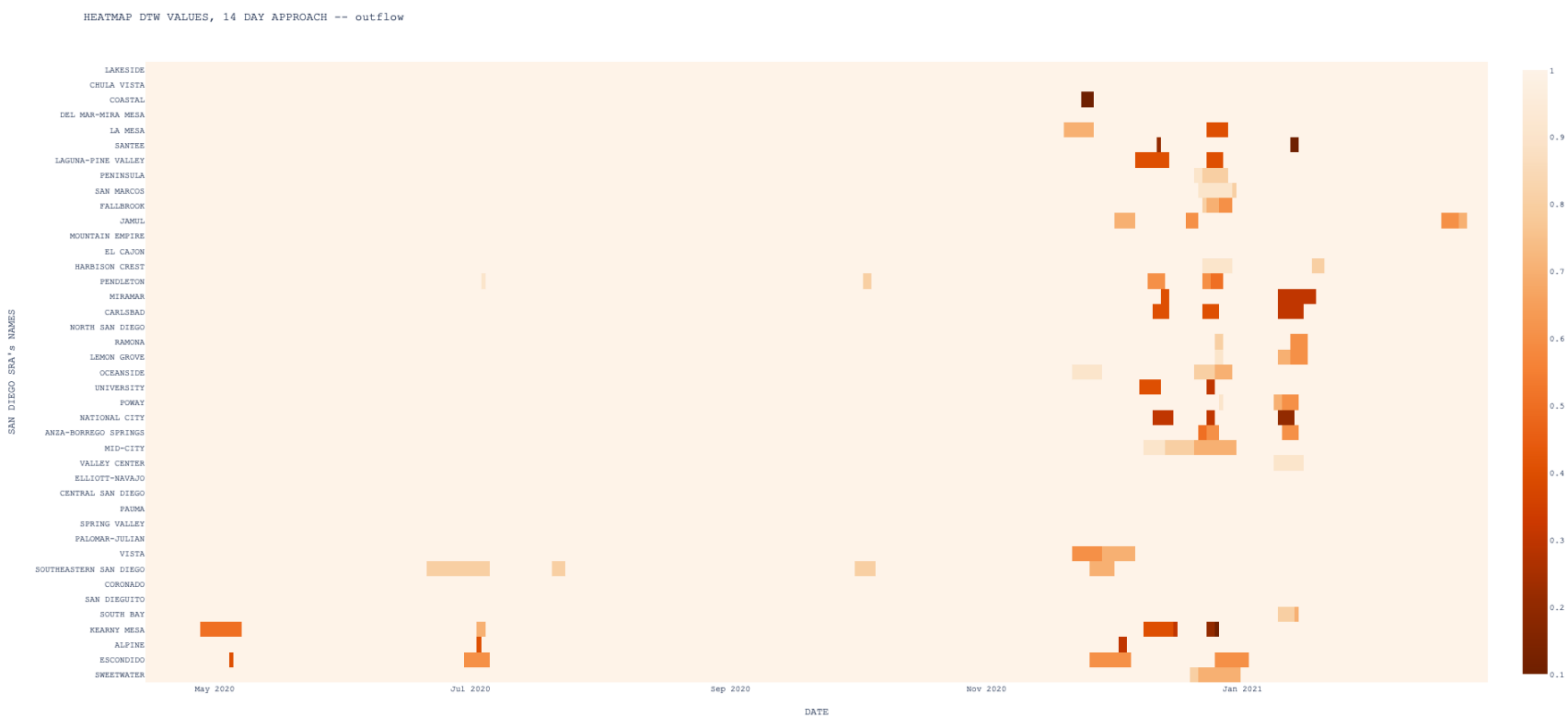


Figure 6.3. Heatmap of DTW values between human mobility outflow and COVID-19.

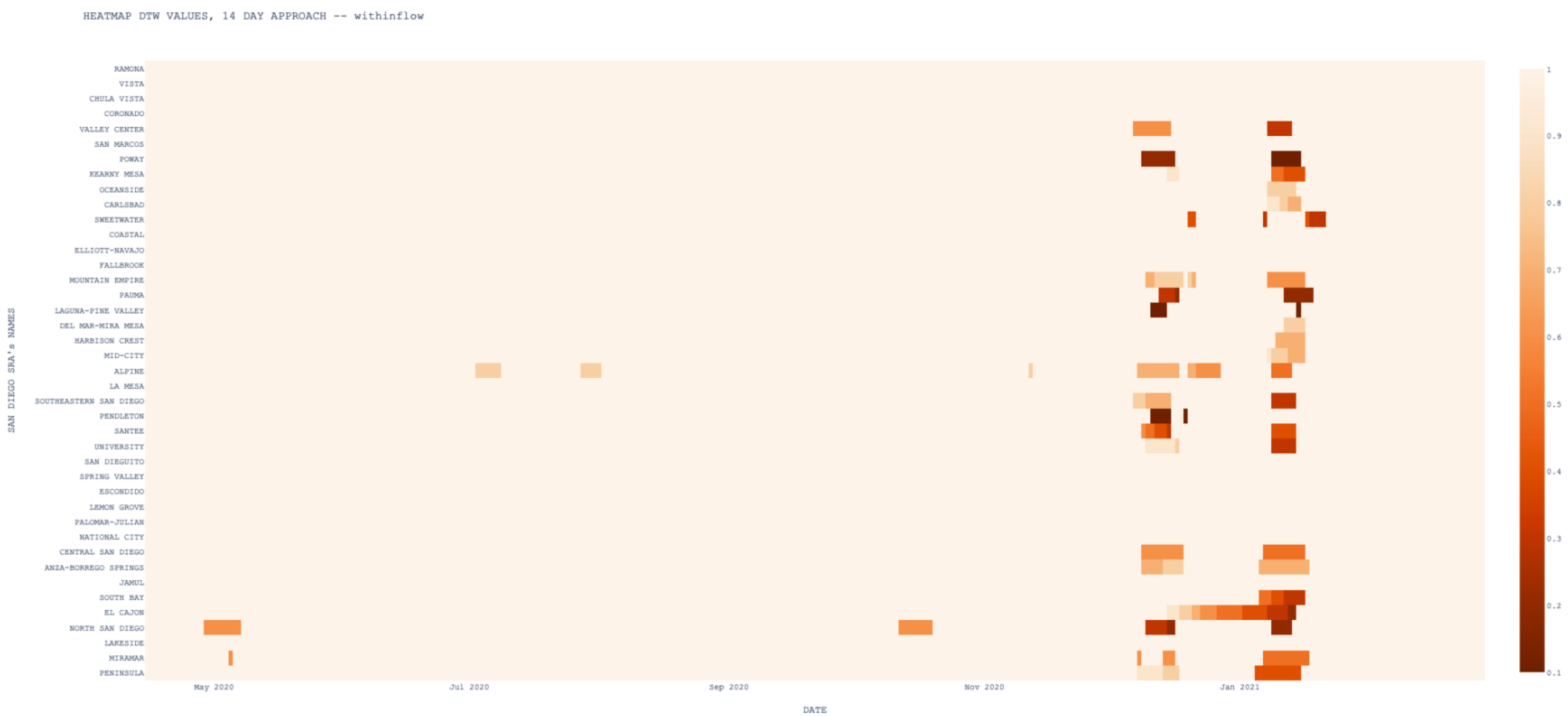


Figure 6.4. Heatmap of DTW values between human mobility withinflow and COVID-19.

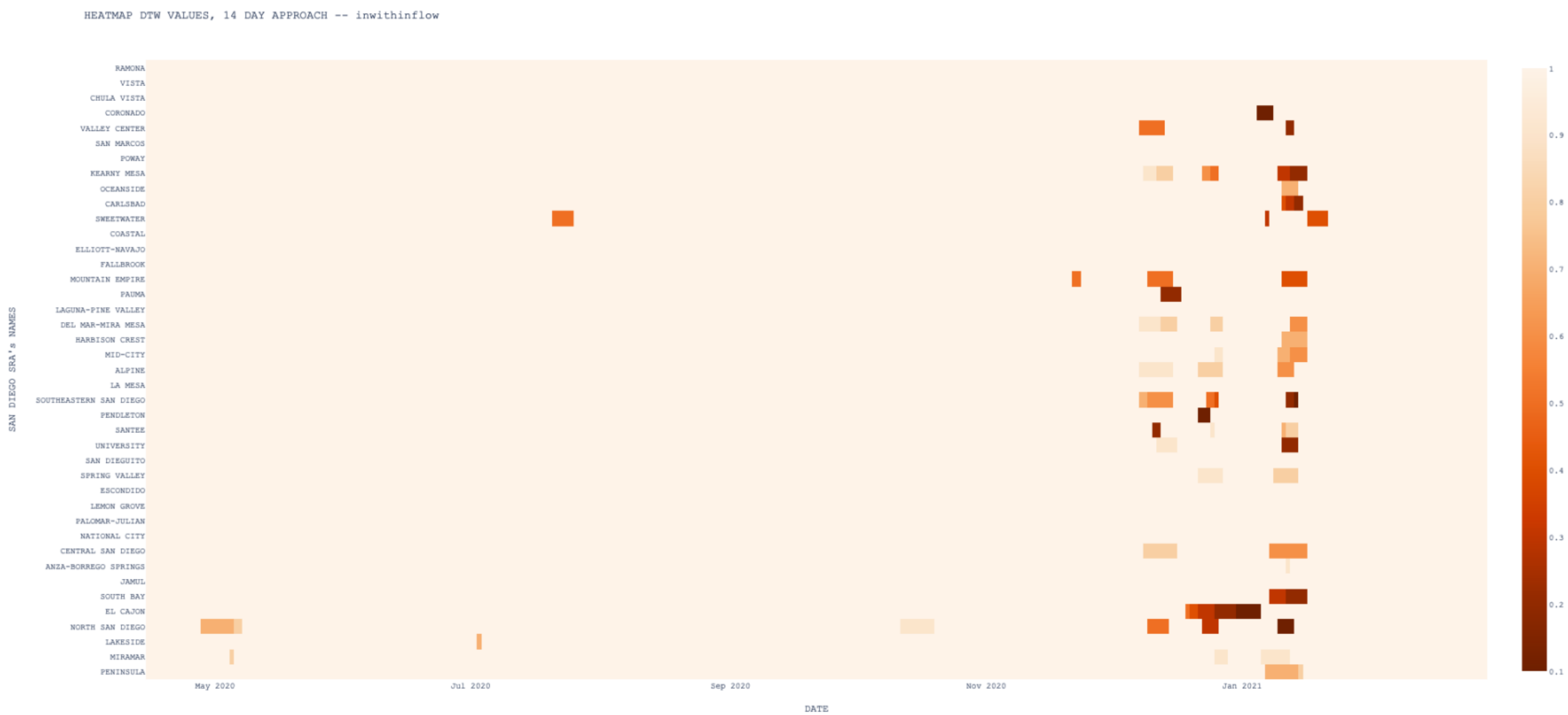


Figure 6.5. Heatmap of DTW values between human mobility inwithinflow and COVID-19.

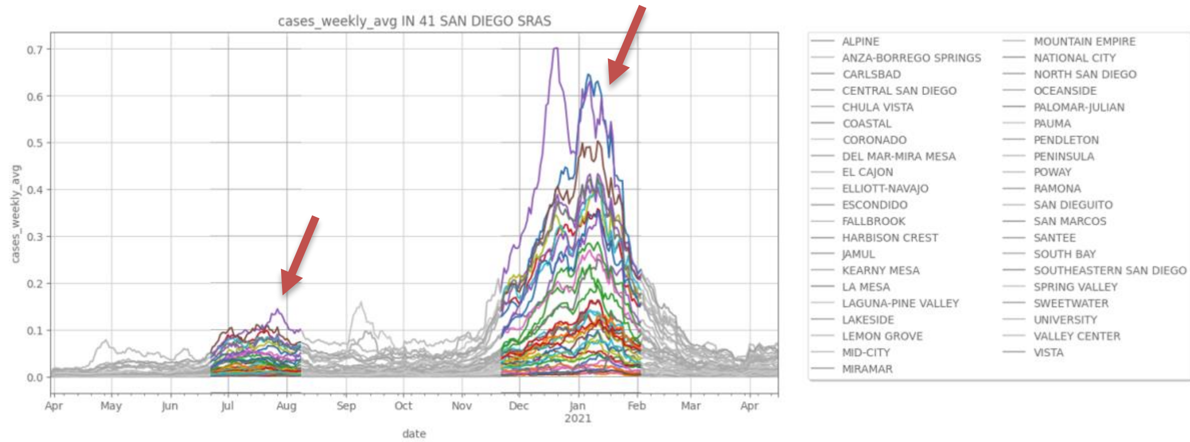


Figure 6.6. High peak of COVID-19 case number in the winter/holiday season 2020.

The withinflow and inwithin flow time series in Figures 6.7 and 6.8 increase slightly during the holiday season.

The inflow in Figure 6.9 does show many up slopes in June-July and late November – December 2020.

The outflow (Figure 6.10) has many lines increasing in the middle of December 2020.

Netflow (Figure 6.11) has many increasing slopes throughout time: June-July, October, and December 2020 and January 2021.

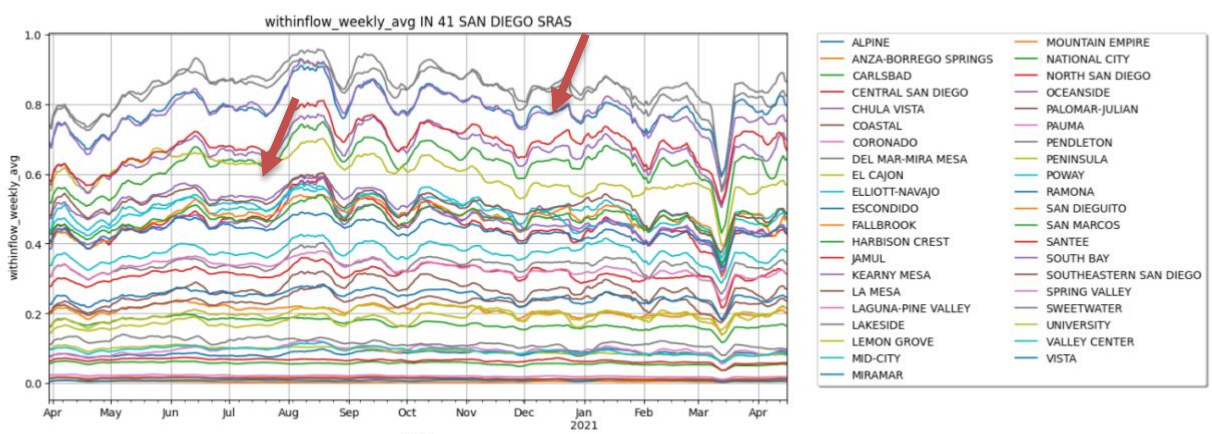


Figure 6.7. Human mobility withinflow in the winter/holiday season 2020.

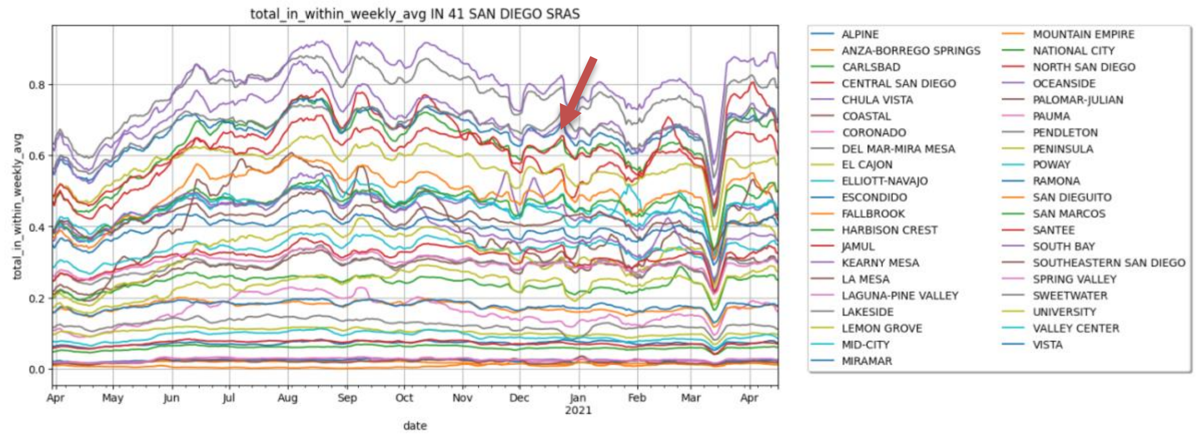


Figure 6.8. Human mobility total_in_within in the winter/holiday season 2020

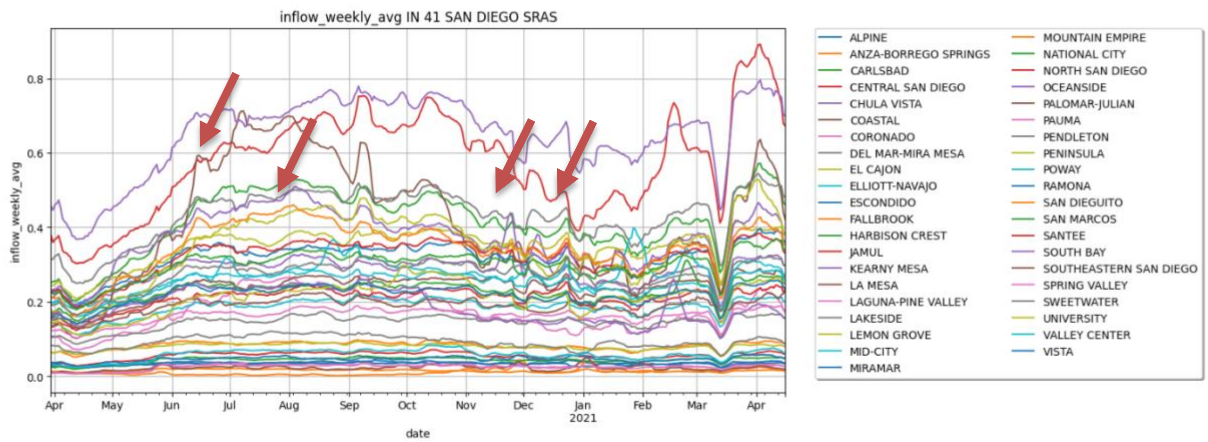


Figure 6.9. Human mobility inflow in the winter/holiday season 2020

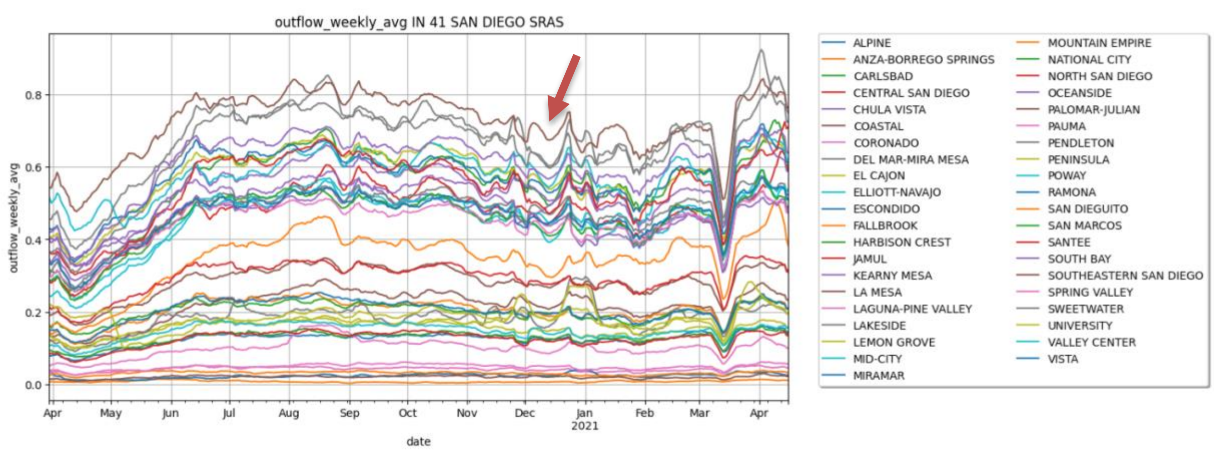


Figure 6.10. Human mobility outflow in the winter/holiday season 2020

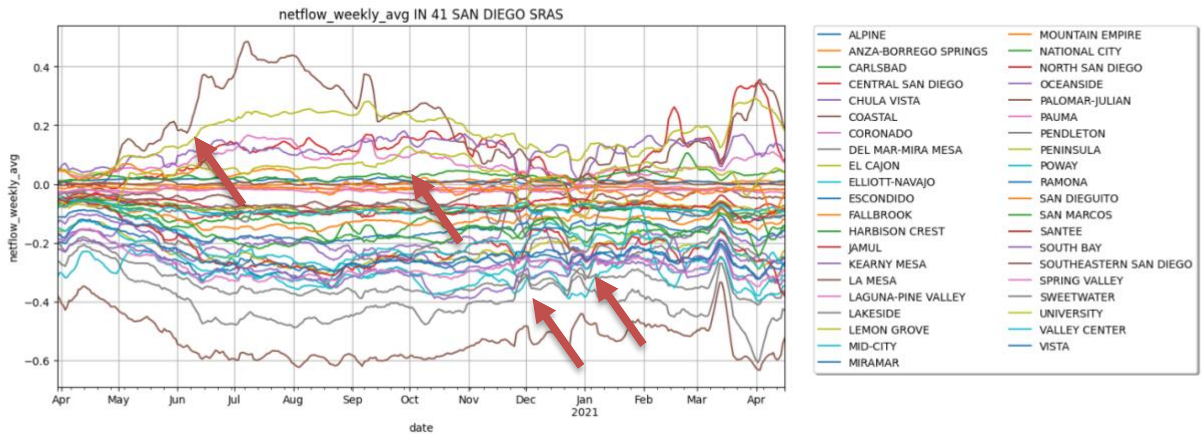


Figure 6.11. Human mobility netflow throughout the time.

6.2 Which SRAs Have More Positive Correlation between Mobility Flow and COVID-19

Each SRA has a count of days that have a positive correlation between human mobility flow and COVID-19 cases. The rows in Figure 6.12 are in the ascending order of the SRA numbers. The SRAs having close SRA numbers are often located next to each other on the map.

The counts are then sorted to reveal which SRAs have the most positive correlation with which flow of human mobility in San Diego County. Below are the SRA names having COVID-19 affected the most by the mobility type (having correlation count > 15 and highlighted in the table).

- Inflow (Figure 6.13): Chula Vista, Southeastern San Diego, Mid-City, South Bay, Elliott-Navajo, National City, Sweetwater, El Cajon, Spring Valley, Santee, Escondido, La Mesa, Poway, University, Vista, San Marcos
- Netflow (Figure 6.14): Central San Diego, Kearny Mesa, San Dieguito, Carlsbad, Fallbrook, La Mesa, Peninsula, Coastal, Chula Vista, Lemon Grove, Santee, National City, Poway, University, San Marcos, Pendleton, Valley Center
- Outflow (Figure 6.15): Coastal, Peninsula, National City, Chula Vista, La Mesa
- Inwithinflow (Figure 6.16): National City, Southeastern San Diego, La Mesa, Vista, University, Lakeside,
- Withinflow (Figure 6.17): La Mesa, Southeastern San Diego, National City, Coastal, Central San Diego, University, Ramona, Peninsula
- SRAs affected by all the flows: La Mesa, National City

SRA Numbers	SRA Names	inflow	netflow	outflow	inwithinflow	withinflow
1	CENTRAL SAN DIEGO	8	60	12	9	21
2	PENINSULA	12	29	25	11	16
3	CORONADO	1	2	3	1	0
4	NATIONAL CITY	30	24	23	31	29
5	SOUTHEASTERN SAN DIEGO	46	0	5	18	31
6	MID-CITY	45	0	0	9	11
10	KEARNY MESA	0	54	0	0	0
11	COASTAL	8	28	29	1	22
12	UNIVERSITY	19	22	15	17	20
13	DEL MAR-MIRA MESA	8	7	0	0	0
14	NORTH SAN DIEGO	8	12	0	0	0
15	POWAY	21	23	0	0	0
16	MIRAMAR	0	0	0	0	0
17	ELLIOTT-NAVAJO	31	0	0	12	0
20	SWEETWATER	30	0	7	0	0
21	CHULA VISTA	47	27	22	9	14
22	SOUTH BAY	32	6	9	7	13
30	JAMUL	3	3	11	3	6
31	SPRING VALLEY	25	0	7	14	15
32	LEMON GROVE	0	26	7	0	0
33	LA MESA	23	31	16	18	35
34	EL CAJON	30	0	9	9	9
35	SANTEE	25	25	6	6	7
36	LAKESIDE	12	0	0	16	5
37	HARBISON CREST	0	7	14	0	5
38	ALPINE	0	11	11	5	12
39	RAMONA	8	7	12	14	20
40	SAN DIEGUITO	3	43	10	0	0
41	CARLSBAD	1	36	0	0	0
42	OCEANSIDE	8	7	0	0	0
43	PENDLETON	14	20	14	11	8
50	ESCONDIDO	25	0	7	5	8
51	SAN MARCOS	16	21	9	4	7
52	VISTA	19	0	8	18	11
53	VALLEY CENTER	6	17	12	0	15
54	PAUMA	0	3	3	0	0
55	FALLBROOK	9	32	12	8	15
60	PALOMAR-JULIAN	5	4	0	4	0
61	LAGUNA-PINE VALLEY	0	0	3	0	0
62	MOUNTAIN EMPIRE	0	0	0	0	0
63	ANZA-BORREGO SPRINGS	0	0	0	0	0
	Total	578	587	321	260	355

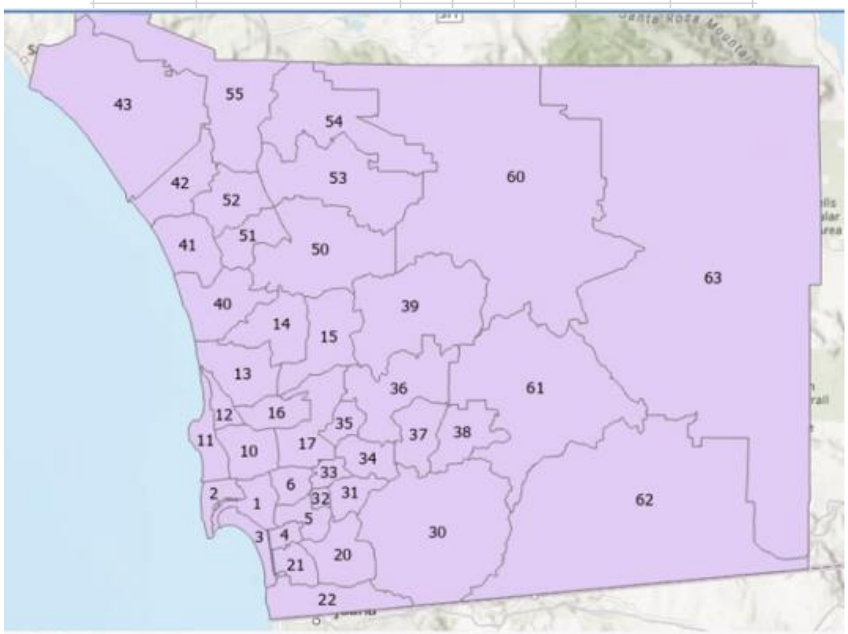


Figure 6.12. Count of days having positive correlation between human mobility and COVID-19 in 41 San Diego SRAs.

SRA names	inflow	netflow	outflow	inwithinflow	withinflow
CHULA VISTA	47	27	22	9	14
SOUTHEASTERN SAN DIEGO	46	0	5	18	31
MID-CITY	45	0	0	9	11
SOUTH BAY	32	6	9	7	13
ELLIOTT-NAVAJO	31	0	0	12	0
NATIONAL CITY	30	24	23	31	29
EL CAJON	30	0	9	9	9
SWEETWATER	30	0	7	0	0
SANTEE	25	25	6	6	7
SPRING VALLEY	25	0	7	14	15
ESCONDIDO	25	0	7	5	8
LA MESA	23	31	16	18	35
POWAY	21	23	0	0	0
UNIVERSITY	19	22	15	17	20
VISTA	19	0	8	18	11
SAN MARCOS	16	21	9	4	7
PENDLETON	14	20	14	11	8
PENINSULA	12	29	25	11	16
LAKESIDE	12	0	0	16	5
FALLBROOK	9	32	12	8	15
CENTRAL SAN DIEGO	8	60	12	9	21
COASTAL	8	28	29	1	22
NORTH SAN DIEGO	8	12	0	0	0
RAMONA	8	7	12	14	20
DEL MAR-MIRA MESA	8	7	0	0	0
OCEANSIDE	8	7	0	0	0
VALLEY CENTER	6	17	12	0	15
PALOMAR-JULIAN	5	4	0	4	0
SAN DIEGUITO	3	43	10	0	0
JAMUL	3	3	11	3	6
CARLSBAD	1	36	0	0	0
CORONADO	1	2	3	1	0
KEARNY MESA	0	54	0	0	0
LEMON GROVE	0	26	7	0	0
ALPINE	0	11	11	5	12
HARBISON CREST	0	7	14	0	5
PAUMA	0	3	3	0	0
LAGUNA-PINE VALLEY	0	0	3	0	0
MIRAMAR	0	0	0	0	0
MOUNTAIN EMPIRE	0	0	0	0	0
ANZA-BORREGO SPRINGS	0	0	0	0	0

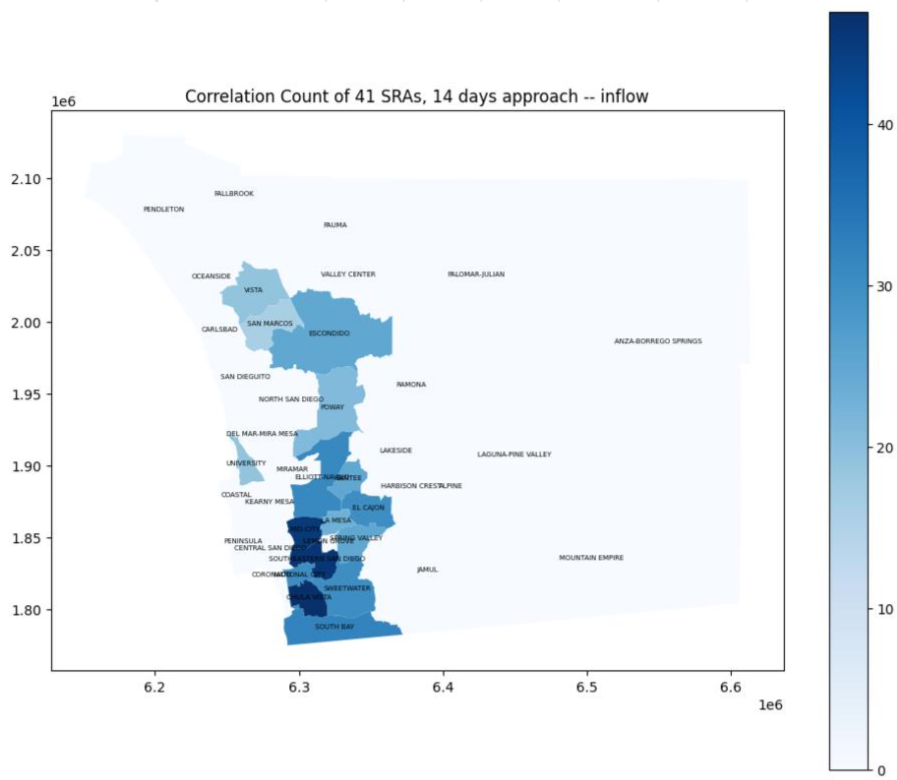


Figure 6.13. San Diego SRAs having the highest correlation day count between human mobility inflow and COVID-19 case number.

SRA names	inflow	netflow	outflow	inwithinflow	withinflow
CENTRAL SAN DIEGO	8	60	12	9	21
KEARNY MESA	0	54	0	0	0
SAN DIEGUITO	3	43	10	0	0
CARLSBAD	1	36	0	0	0
FALLBROOK	9	32	12	8	15
LA MESA	23	31	16	18	35
PENINSULA	12	29	25	11	16
COASTAL	8	28	29	1	22
CHULA VISTA	47	27	22	9	14
LEMON GROVE	0	26	7	0	0
SANTEE	25	25	6	6	7
NATIONAL CITY	30	24	23	31	29
POWAY	21	23	0	0	0
UNIVERSITY	19	22	15	17	20
SAN MARCOS	16	21	9	4	7
PENDLETON	14	20	14	11	8
VALLEY CENTER	6	17	12	0	15
NORTH SAN DIEGO	8	12	0	0	0
ALPINE	0	11	11	5	12
RAMONA	8	7	12	14	20
DEL MAR-MIRA MESA	8	7	0	0	0
OCEANSIDE	8	7	0	0	0
HARBISON CREST	0	7	14	0	5
SOUTH BAY	32	6	9	7	13
PALOMAR-JULIAN	5	4	0	4	0
JAMUL	3	3	11	3	6
PAUMA	0	3	3	0	0
CORONADO	1	2	3	1	0
SOUTHEASTERN SAN DIEGO	46	0	5	18	31
MID-CITY	45	0	0	9	11
ELLIOTT-NAVAJO	31	0	0	12	0
EL CAJON	30	0	9	9	9
SWEETWATER	30	0	7	0	0
SPRING VALLEY	25	0	7	14	15
ESCONDIDO	25	0	7	5	8
VISTA	19	0	8	18	11
LAKESIDE	12	0	0	16	5
LAGUNA-PINE VALLEY	0	0	3	0	0
MIRAMAR	0	0	0	0	0
MOUNTAIN EMPIRE	0	0	0	0	0
ANZA-BORREGO SPRINGS	0	0	0	0	0

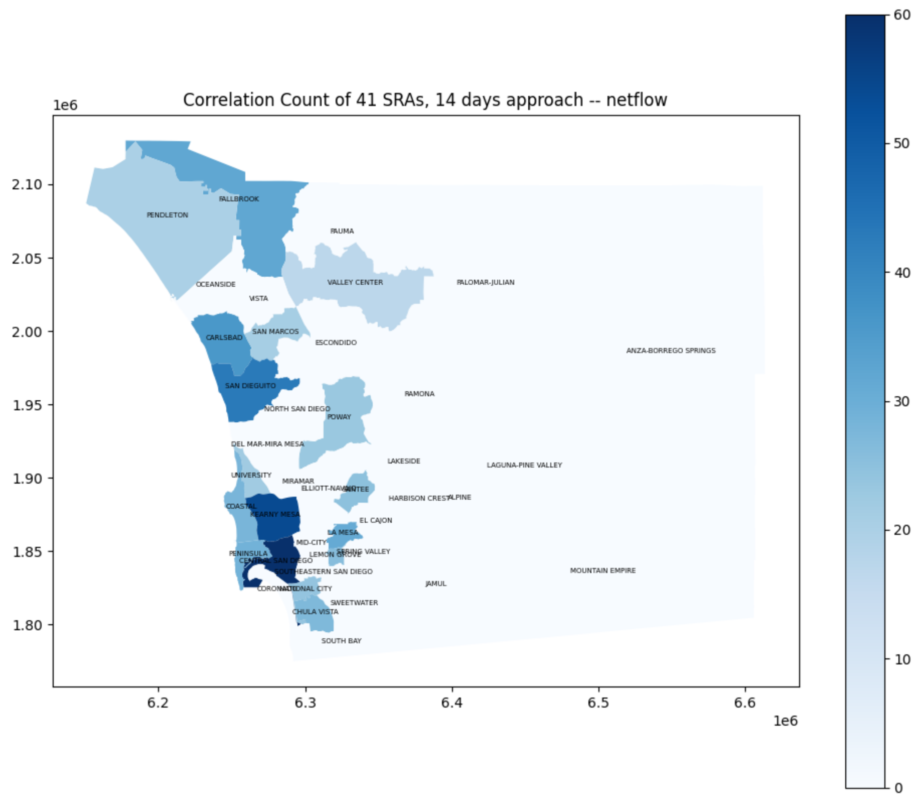


Figure 6.14. San Diego SRAs having the highest correlation day count between human mobility netflow and COVID-19 case number.

SRA names	inflow	netflow	outflow	inwithinflow	withinflow
COASTAL	8	28	29	1	22
PENINSULA	12	29	25	11	16
NATIONAL CITY	30	24	23	31	29
CHULA VISTA	47	27	22	9	14
LA MESA	23	31	16	18	35
UNIVERSITY	19	22	15	17	20
PENDLETON	14	20	14	11	8
HARBISON CREST	0	7	14	0	5
CENTRAL SAN DIEGO	8	60	12	9	21
FALLBROOK	9	32	12	8	15
VALLEY CENTER	6	17	12	0	15
RAMONA	8	7	12	14	20
ALPINE	0	11	11	5	12
JAMUL	3	3	11	3	6
SAN DIEGUITO	3	43	10	0	0
SAN MARCOS	16	21	9	4	7
SOUTH BAY	32	6	9	7	13
EL CAJON	30	0	9	9	9
VISTA	19	0	8	18	11
LEMON GROVE	0	26	7	0	0
SWEETWATER	30	0	7	0	0
SPRING VALLEY	25	0	7	14	15
ESCONDIDO	25	0	7	5	8
SANTEE	25	25	6	6	7
SOUTHEASTERN SAN DIEGO	46	0	5	18	31
PAUMA	0	3	3	0	0
CORONADO	1	2	3	1	0
LAGUNA-PINE VALLEY	0	0	3	0	0
KEARNY MESA	0	54	0	0	0
CARLSBAD	1	36	0	0	0
POWAY	21	23	0	0	0
NORTH SAN DIEGO	8	12	0	0	0
DEL MAR-MIRA MESA	8	7	0	0	0
OCEANSIDE	8	7	0	0	0
PALOMAR-JULIAN	5	4	0	4	0
MID-CITY	45	0	0	9	11
ELLIOTT-NAVAJO	31	0	0	12	0
LAKESIDE	12	0	0	16	5
MIRAMAR	0	0	0	0	0
MOUNTAIN EMPIRE	0	0	0	0	0
ANZA-BORREGO SPRINGS	0	0	0	0	0

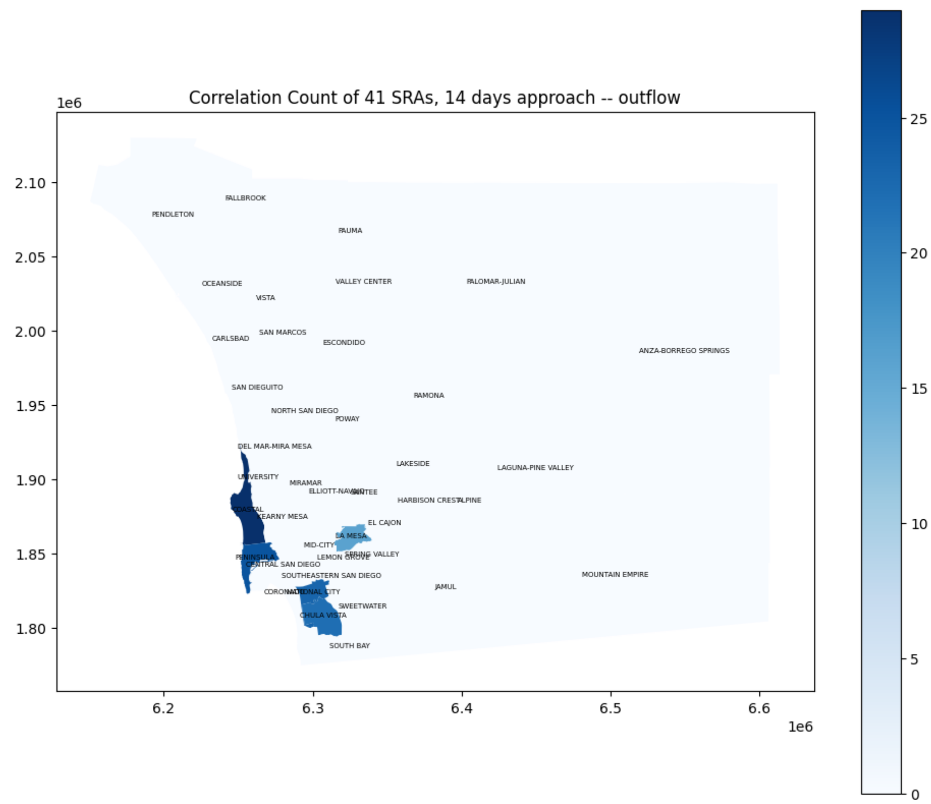


Figure 6.15. San Diego SRAs having the highest correlation day count between human mobility outflow and COVID-19 case number.

SRA names	inflow	netflow	outflow	inwithinflow	withinflow
NATIONAL CITY	30	24	23	31	29
LA MESA	23	31	16	18	35
VISTA	19	0	8	18	11
SOUTHEASTERN SAN DIEGO	46	0	5	18	31
UNIVERSITY	19	22	15	17	20
LAKESIDE	12	0	0	16	5
RAMONA	8	7	12	14	20
SPRING VALLEY	25	0	7	14	15
ELLIOTT-NAVAJO	31	0	0	12	0
PENINSULA	12	29	25	11	16
PENDLETON	14	20	14	11	8
CHULA VISTA	47	27	22	9	14
CENTRAL SAN DIEGO	8	60	12	9	21
EL CAJON	30	0	9	9	9
MID-CITY	45	0	0	9	11
FALLBROOK	9	32	12	8	15
SOUTH BAY	32	6	9	7	13
SANTEE	25	25	6	6	7
ALPINE	0	11	11	5	12
ESCONDIDO	25	0	7	5	8
SAN MARCOS	16	21	9	4	7
PALOMAR-JULIAN	5	4	0	4	0
JAMUL	3	3	11	3	6
COASTAL	8	28	29	1	22
CORONADO	1	2	3	1	0
HARBISON CREST	0	7	14	0	5
VALLEY CENTER	6	17	12	0	15
SAN DIEGUITO	3	43	10	0	0
LEMON GROVE	0	26	7	0	0
SWEETWATER	30	0	7	0	0
PAUMA	0	3	3	0	0
LAGUNA-PINE VALLEY	0	0	3	0	0
KEARNY MESA	0	54	0	0	0
CARLSBAD	1	36	0	0	0
POWAY	21	23	0	0	0
NORTH SAN DIEGO	8	12	0	0	0
DEL MAR-MIRA MESA	8	7	0	0	0
OCEANSIDE	8	7	0	0	0
MIRAMAR	0	0	0	0	0
MOUNTAIN EMPIRE	0	0	0	0	0
ANZA-BORREGO SPRINGS	0	0	0	0	0

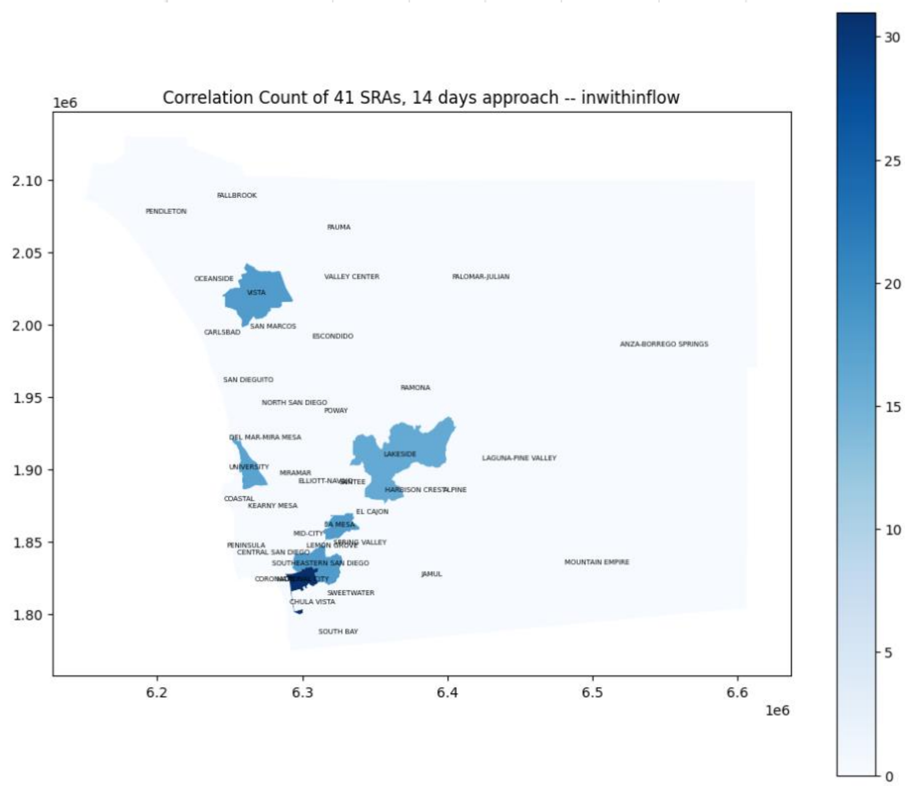


Figure 6.16. San Diego SRAs having the highest correlation day count between human mobility inwithinflow and COVID-19 case number.

SRA names	inflow	netflow	outflow	inwithinflow	withinflow
LA MESA	23	31	16	18	35
SOUTHEASTERN SAN DIEGO	46	0	5	18	31
NATIONAL CITY	30	24	23	31	29
COASTAL	8	28	29	1	22
CENTRAL SAN DIEGO	8	60	12	9	21
UNIVERSITY	19	22	15	17	20
RAMONA	8	7	12	14	20
PENINSULA	12	29	25	11	16
SPRING VALLEY	25	0	7	14	15
FALLBROOK	9	32	12	8	15
VALLEY CENTER	6	17	12	0	15
CHULA VISTA	47	27	22	9	14
SOUTH BAY	32	6	9	7	13
ALPINE	0	11	11	5	12
VISTA	19	0	8	18	11
MID-CITY	45	0	0	9	11
EL CAJON	30	0	9	9	9
PENDLETON	14	20	14	11	8
ESCONDIDO	25	0	7	5	8
SANTEE	25	25	6	6	7
SAN MARCOS	16	21	9	4	7
JAMUL	3	3	11	3	6
LAKESIDE	12	0	0	16	5
HARBISON CREST	0	7	14	0	5
ELLIOTT-NAVAJO	31	0	0	12	0
PALOMAR-JULIAN	5	4	0	4	0
CORONADO	1	2	3	1	0
SAN DIEGUITO	3	43	10	0	0
LEMON GROVE	0	26	7	0	0
SWEETWATER	30	0	7	0	0
PAUMA	0	3	3	0	0
LAGUNA-PINE VALLEY	0	0	3	0	0
KEARNY MESA	0	54	0	0	0
CARLSBAD	1	36	0	0	0
POWAY	21	23	0	0	0
NORTH SAN DIEGO	8	12	0	0	0
DEL MAR-MIRA MESA	8	7	0	0	0
OCEANSIDE	8	7	0	0	0
MIRAMAR	0	0	0	0	0
MOUNTAIN EMPIRE	0	0	0	0	0
ANZA-BORREGO SPRINGS	0	0	0	0	0

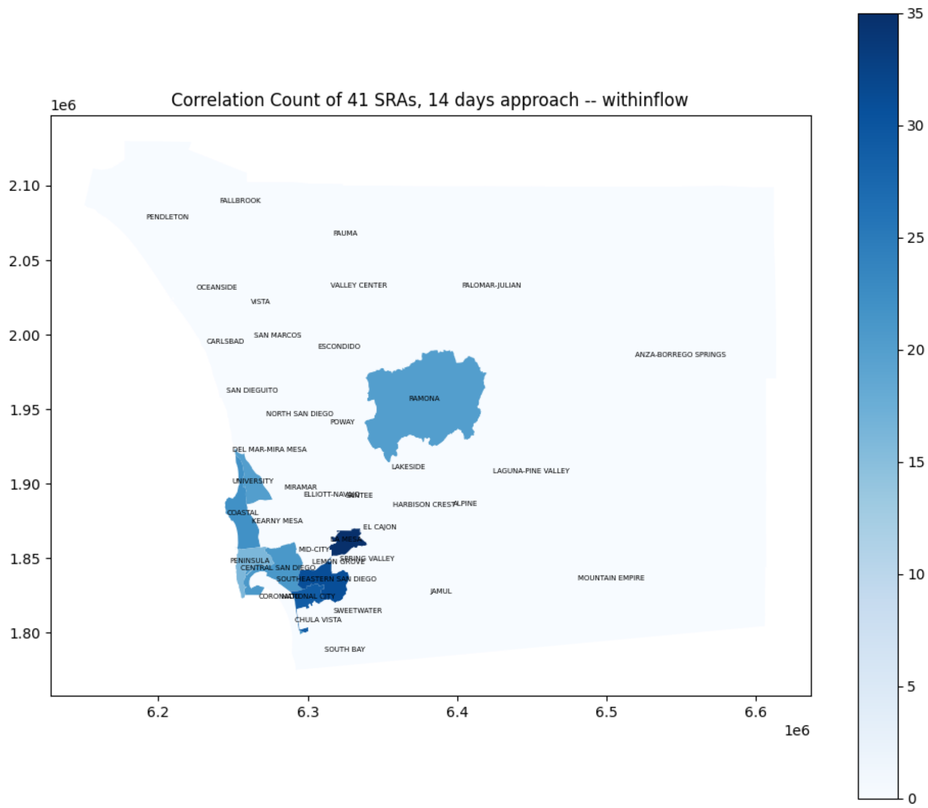


Figure 6.17. San Diego SRAs having the highest correlation day count between human mobility withinflow and COVID-19 case number.

On the choropleth maps, many neighboring SRAs have similar correlation counts (similar color shade), suggesting that the spillover effects exist in some areas of San Diego. For example, the South region across part of the North Center to the east of the North Inland in the inflow choropleth map (Figure 6.13), or the North Center and the lower region of the North Coastal in the netflow choropleth map (Figure 6.14). These areas often have higher population densities (Statistical Atlas, n.d.), which tend to transmit the disease easier.

Examples of the time series of some SRAs that show the positive correlation between COVID-19 and human mobility flows (Figure 6.18).

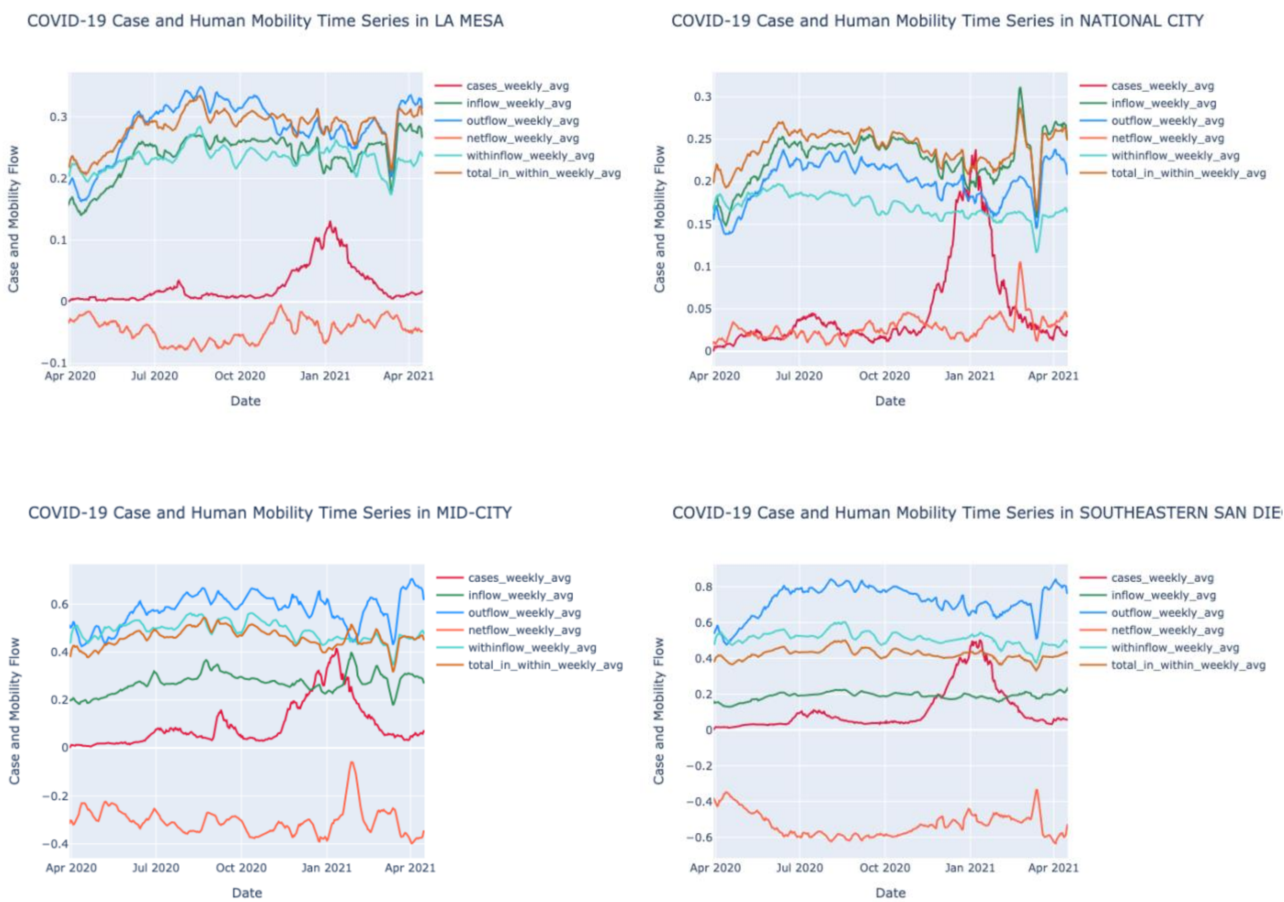


Figure 6.18. Time series of San Diego SRAs having the highest correlation between human mobility netflow and COVID-19 case number.

6.3 Which Human Mobility Flow Has the Most Impact on COVID-19 Case

The human mobility inflow positively correlates with the COVID-19 case in 578 days in San Diego, netflow impacts 587 days, while outflow 321, inwithinflow 260, and withinflow 355 (Supplemental Material 1). On average, inflow and netflow affects COVID-19 case in 14 days per SRA, higher than the corresponding index of outflow, within flow, and inwithin flow, which means that netflow and inflow positively correlate more with the COVID-19 situation in San Diego County. Note that the days that are counted are those that meet the two conditions (DTW values < 1.0 and case and flow both have positive slopes), so inflow and withinflow combined do not equal the inwithinflow as in the original dataset.

Outflow, inwithinflow and withinflow do not have a high impact on COVID-19 since they impact 7, 6, and 8 days of each SRA on average.

Table 6.1. Statistics of Day Counts by Human Mobility Flows

Human mobility flow	total_count	max_count	min_count	mean_count	standard deviation_count
inflow	578	47	0	14.1	13.91
netflow	587	60	0	14.32	15.93
outflow	321	29	0	7.83	7.66
inwithinflow	260	31	0	6.34	7.37
withinflow	355	35	0	8.66	9.69

CHAPTER 7

DISCUSSION & CONCLUSION

The results prove that the positive correlation between human mobility and COVID-19 case numbers depends on time and place, which is more in line with the conclusion that the relationship is diverse in the studies of Nouvellet et al. (2021), Elarde et al. (2021), Li et al. (2021), Praharaj and Han (2021), S. Wang et al. (2020), and Yuan et al. (2020). The correlation existed in May and June 2020, especially between the inflow and netflow and COVID-19 cases. However, the outflow, withinflow, and inwithin flow showed little association. Therefore, the connection between human mobility and COVID-19 does show in the early time of the pandemic, but not as strong as in the prior studies by Badr et al. (2020), Cot et al. (2021), Kraemer et al. (2020), and Linka et al. (n.d.).

With a diverse racial and ethnic composition: 42.8% non-Hispanic white, 30.3% Hispanic, 16.7% Asian, 6.4% Black or African American, and 26.1% foreign-born persons (U.S. Census Bureau, n.d.), San Diego tends to have local people going out to visit friends and family in Mexico and Asian and European countries (outflow), or become a place of meetups for family members and friends coming from other countries and states (inflow). A meeting in close contact is an excellent condition for the virus to spread and increases the COVID-19 case numbers in San Diego.

This study finds that San Diego's human mobility and COVID-19 move in the same direction mainly in the winter/holiday season of 2020 and early 2021. This trend is consistent with all five human mobility flow types. California was the first state to order a lockdown to combat COVID-19 in March 2020. People tend to travel more after months of dealing with the stress from the strict COVID-19 policies and the changes COVID-19 brought to their work and daily lives. Human movements into, out of, and within San Diego neighborhoods for tourism purposes and to visit relatives and friends in late 2020 is one primary reason for the high positive correlation between human mobility and COVID-19 in the winter and holiday season of 2020.

The flows of people coming to San Diego County (inflow and netflow) have more impact on the Covid-19 case increase in San Diego than the flow of people going out of the County. During the holiday seasons, people come to San Diego for tourism purposes to enjoy the warm weather and beautiful beaches – the advantage San Diego has for its tourism attraction. On the other hand, San Diegans can travel outside their neighborhood to get a different winter experience, but usually they do not go out as much as the people coming in, resulting in the positive netflow in San Diego, especially in tourism attracting neighborhoods in the Central Region and North Central Region of the County. An area with positive netflow has a higher chance of being exposed to the virus from outside. The visitors might not be aware that they have the virus in their bodies and can accidentally bring it to a San Diego neighborhood. The chance for travelers to get the virus is higher as they need to communicate (risk of close contact) to use gasoline, food, and airline services and might touch objects having the virus during their travel to San Diego.

Unlike the inflow and netflow, withinflow refers to the movements inside the local neighborhood. For example, it could be the flow of local people going to work or school, shopping for groceries and daily necessaries. The study finds much correlation between withinflow/inwithinflow and COVID-19 at the end of 2020 and early 2021. This pattern could be explained by the more vibrant movements of the local people having more activities around the year-end and holiday seasons.

In terms of geographic locations, La Mesa and National City, with complex demographics, are the two SRAs that have the COVID-19 case number increase affected by at least one human mobility type (inflow, outflow, netflow, withinflow, inwithinflow). The withinflow has the most significant impact on tourist and college places such as Peninsula, Coastal, University, Coronado, Central San Diego, and Southeastern San Diego, compared to its effects on other areas in San Diego. On the other hand, the North Central, including Coastal, University, Kearny Mesa, Peninsula, is affected the most by the mobility netflow. The South region – South Bay, Chula Vista, Southeastern San Diego, National City, and Sweetwater – is most affected by human mobility inflow. The South Region has a high concentration of immigrants and first-generation Americans. Encanto, which belongs to

Southeastern San Diego, is one of the most diverse communities in San Diego (Lovejoy, 2021). The connection between immigrants and COVID-19 spread can be examined in a future study.

This study shows that the spatial association among San Diego neighborhoods does exist, depending on specific time and place (Figures 7.1-7.5). The Figures 7.1-7.5 are the animations that visualize the time and place that both human mobility and COVID-19 increase, based on the DTW values. The animation plays when the play button is clicked. The spatial temporal autocorrelation exists when the color shade shifts from one SRA to its neighboring SRAs on the same or consecutive days.

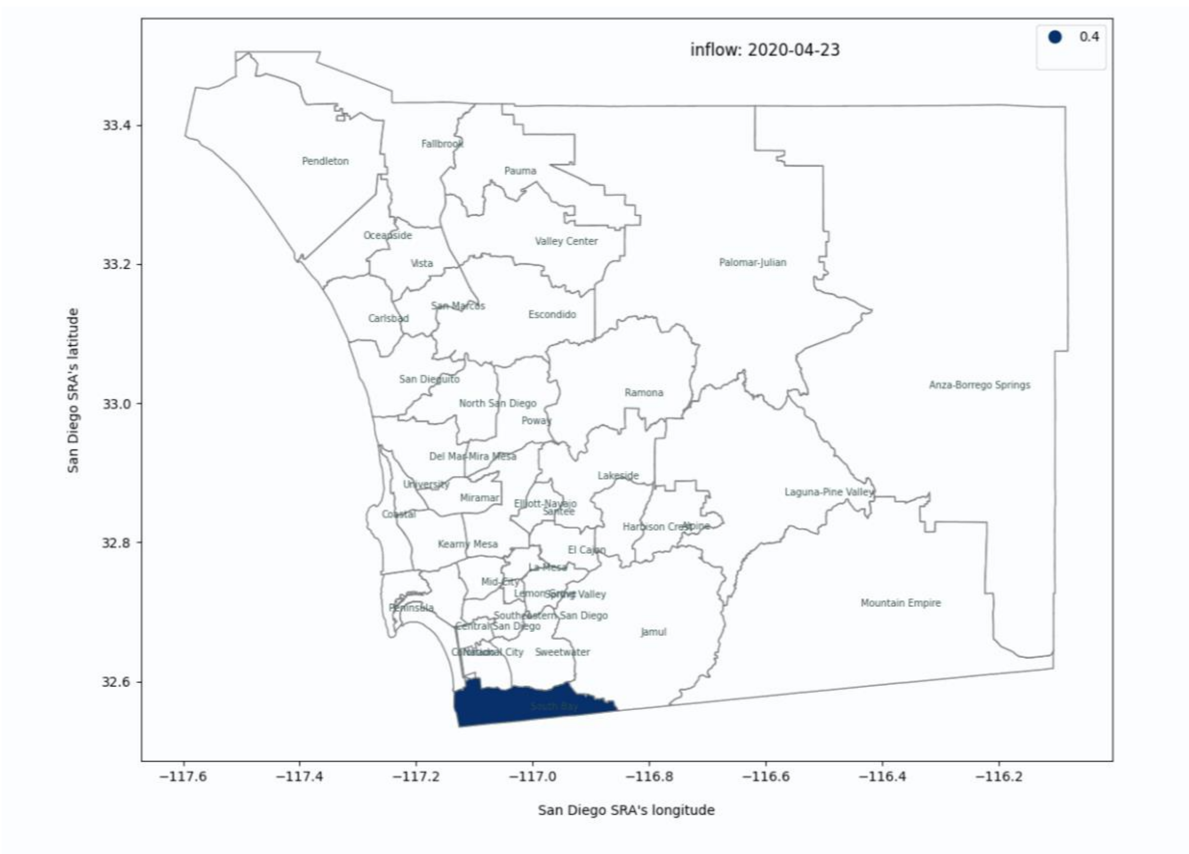


Figure 7.1. The development of DTW values over time in SRAs that have positive correlation between human mobility inflow and COVID-19 cases.

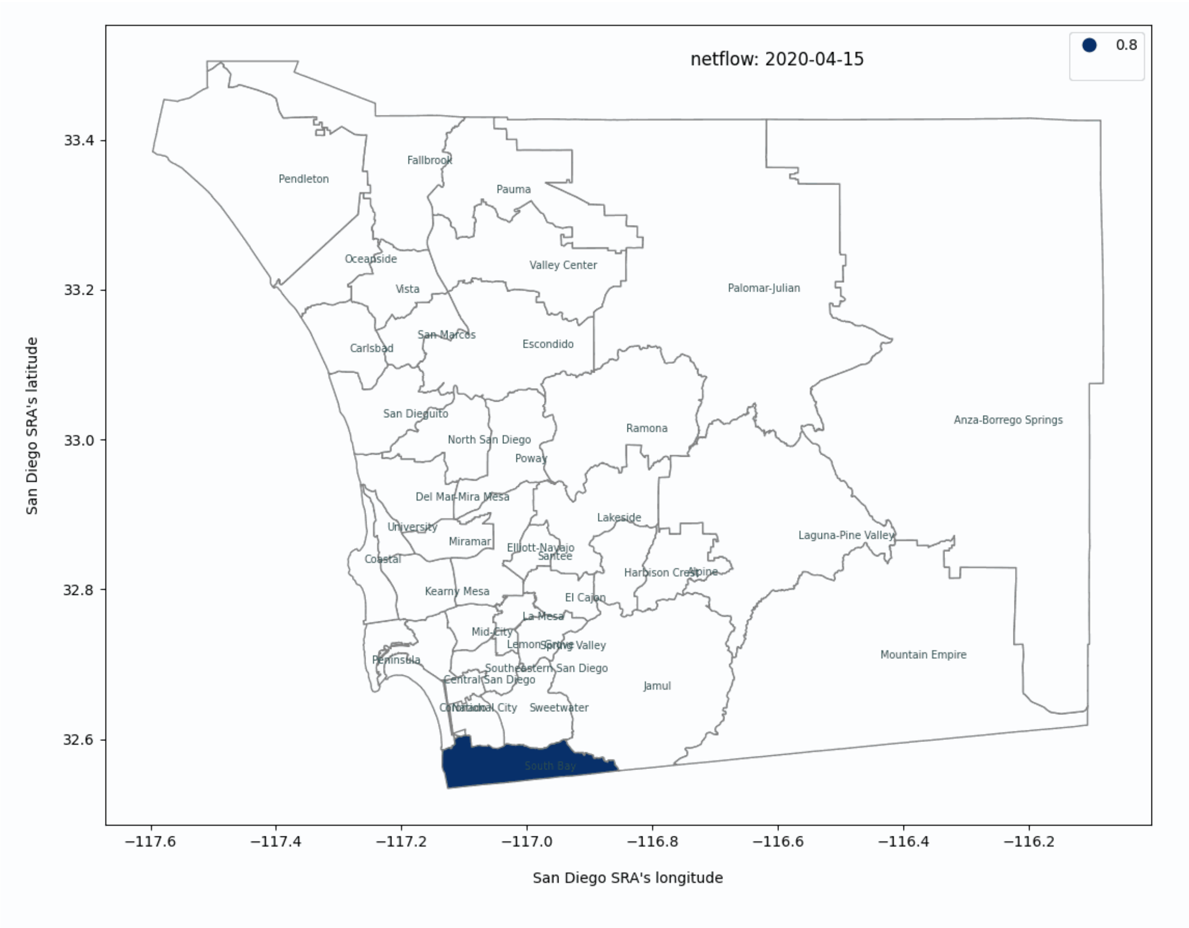


Figure 7.2. The development of DTW values over time in SRAs that have positive correlation between human mobility netflow and COVID-19 cases.

In conclusion, this study analyzes the increase-increase relationship between human mobility and COVID-19 cases in 41 sub-region areas of San Diego County, giving insights into when and where that relationship exists and what types of human mobility have the most positive correlation with the COVID-19 cases.

The revealed patterns are meaningful to the local public health interventions. The insight about the timing helps raise awareness for the community about COVID-19 when it transmits rapidly. Local governments can implement a temporary lock-down or stay-at-home order for communities with similar human mobility and COVID-19 case patterns. Those SRAs that have spatial association among SRAs can set up vaccination sites, recommend mask-wearing or reinforce social gathering restrictions to control the COVID-19 spread in that area group. The findings in this study can assist public health agencies with additional

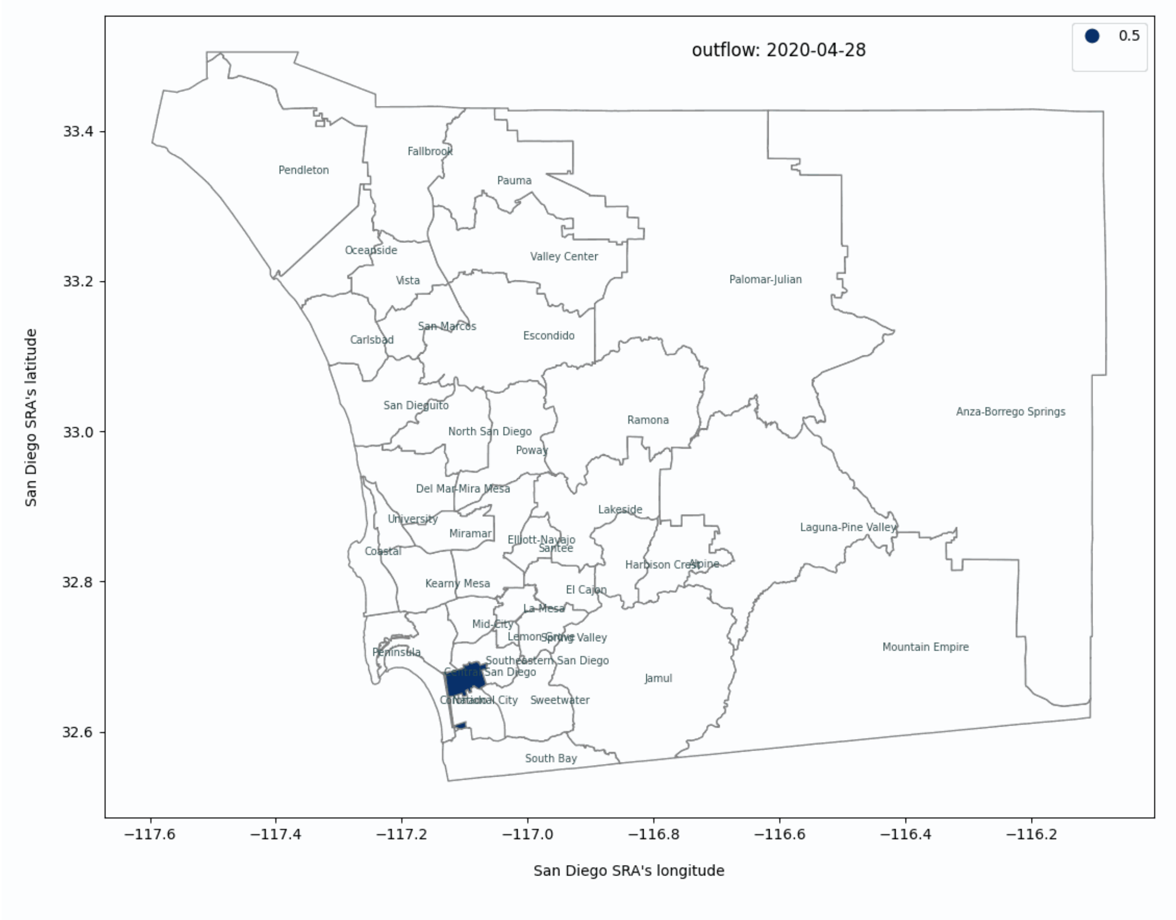


Figure 7.3. The development of DTW values over time in SRAs that have positive correlation between human mobility outflow and COVID-19 cases.

information to evaluate the effectiveness of their COVID-19 or similar airborne disease control policies.

Some limitations that can improve in future studies:

- If the COVID-19 data used in this study covered March 2020, the COVID-19 – mobility relationship would be examined for the earliest stage of the pandemic in San Diego.
- The study focuses only on the positive correlation (increase-increase) between human mobility and COVID-19. Some SRAs have low DTW values, but the correlation between the case and flow does not show in the choropleth map. The reason is that the correlation is not increase-increase, and it could be the decrease-decrease or stable-stable relationship, which is not the focus of this study. Future studies can extend to the decrease-decrease and stable-stable relationship between the two time series.

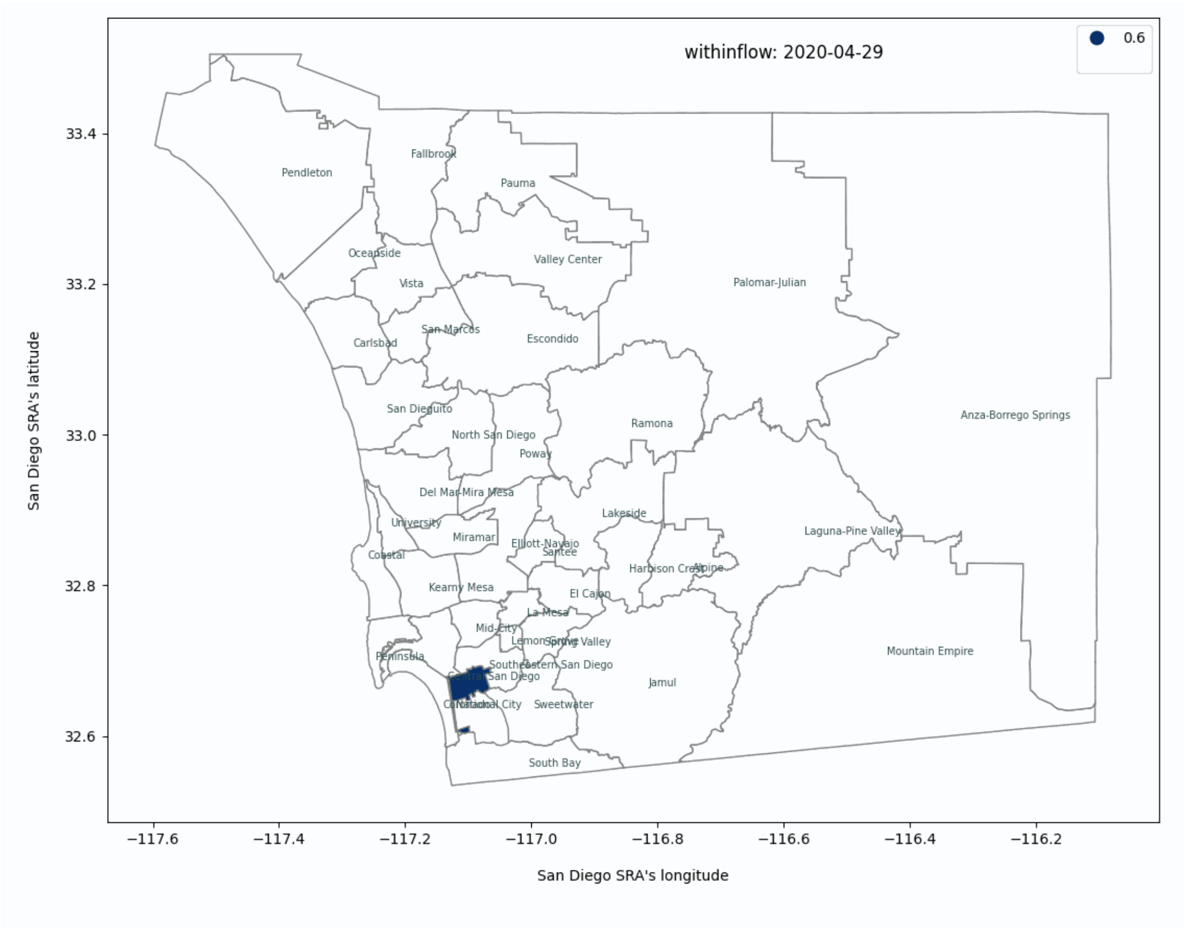


Figure 7.4. The development of DTW values over time in SRAs that have positive correlation between human mobility withinflow and COVID-19 cases.

- The DTW method is applied on the 15-day time series, which should cover the 7-day lag between human mobility and COVID-19 cases. However, a future study can try shifting the human mobility time series seven days backward to test if it can show a stronger correlation with the COVID-19 data.
- The conclusion is based on the slopes being rounded up to three decimal places. The correlation count table will show more values if the slope values have more than three decimal places.
- There are cases when the increase in movements is noticed, but it does not have an association with the COVID-19 case, which is due to other factors, including vaccination efforts and non-pharmaceutical interventions such as hand washing, mask wearing, and social distancing among people. More data about vaccination and human behaviors are needed to extend the scope of future study.

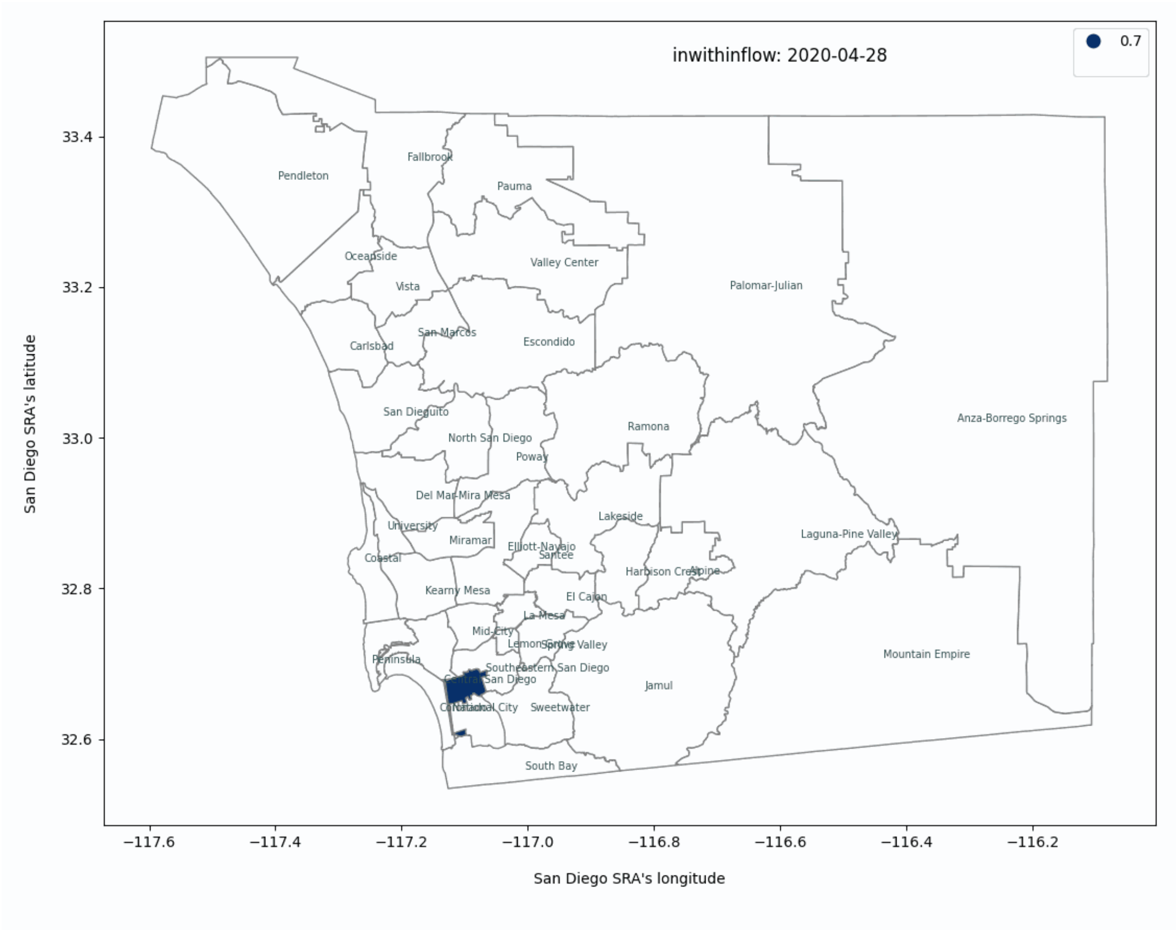


Figure 7.5. The development of DTW values over time in SRAs that have positive correlation between human mobility inwithinflow and COVID-19 cases.

REFERENCES

Alessandretti, L. (2022). What human mobility data tell us about COVID-19 spread. *Nature Reviews Physics*, *4*(1), 12–13. https://doi.org/10.1038/s42254-021-00407-1

Alizadeh, E. (2020, October 13). *An illustrative explanation to dynamic time warping*. Towards Data Science. https://towardsdatascience.com/an-illustrative-introduction-to-dynamic-time-warping-36aa98513b98

Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *The Lancet Infectious Diseases*, *20*(11), 1247–1254. https://doi.org/10.1016/S1473-3099(20)30553-3

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, *589*(7840), 82–87. https://doi.org/10.1038/s41586-020-2923-3

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini, I. M., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395–400. https://doi.org/10.1126/science.aba9757

Cot, C., Cacciapaglia, G., & Sannino, F. (2021). Mining Google and Apple mobility data: Temporal anatomy for COVID-19 social distancing. *Scientific Reports*, *11*(1), Article 4150. https://doi.org/10.1038/s41598-021-83441-4

County of San Diego. (2022). *Health & human services agency*. https://www.sandiegocounty.gov/hhsa/

covid19.ca.gov. (2022). *Current safety measures*. https://covid19.ca.gov/safely-reopening/

Elarde, J., Kim, J.-S., Kavak, H., Züfle, A., & Anderson, T. (2021). Change of human mobility during COVID-19: A United States case study. *PloS One*, *16*(11), Article e0259031. https://doi.org/10.1371/journal.pone.0259031

Eurostat. (2020). *Your key to European statistics. Air transport of passengers*. Retrieved April 13, 2020, from https://ec. europa.eu/eurostat

Halverstadt, L. (2013, January 29). *Fact check: San Diego county's rural landscape*. Voice of San Diego. https://www.voiceofsandiego.org/topics/news/fact-check-san-diego-countys-rural-landscape/

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Open Covid-19 Data Working Group, Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, *368*(6490), 493-497.

Lewis, B. (2021, March 8). *COVID-19 in San Diego: A look back at the timeline of the past year*. CBS 8. https://www.cbs8.com/article/news/health/coronavirus/covid-one-year-san-diego/509-8008290f-7c8f-4ed9-9bf2-6f57de82474e

Li, Q., Bessell, L., Xiao, X., Fan, C., Gao, X., & Mostafavi, A. (2021). Disparate patterns of movements and visits to points of interest located in urban hotspots across US metropolitan cities during COVID-19. *Royal Society Open Science*, *8*(1), Article 201209. https://doi.org/10.1098/rsos.201209

Linka, K., Goriely, A., & Kuhl, E. (n.d.). *Global and local mobility as a barometer for COVID-19 dynamics*. MedRxiv. https://doi.org/10.1101%2F2020.06.13.20130658

Linka, K., Peirlinck, M., Sahli Costabal, F., & Kuhl, E. (2020). Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. *Computer Methods in Biomechanics and Biomedical Engineering*, *23*(11), 710–717. https://doi.org/10.1080/10255842.2020.1759560

Live Well San Diego. (2018). *HHSA-service-region-dashboard-2018*. https://www.livewellsd.org/indicators/Dashboards/HHSA-Service-Region-Dashboard-2018.pdf

Lovejoy, B. (2021, February 25). *What's living in San Diego like? | ULTIMATE moving to San Diego guide*. Republic Moving & Storage. https://republicmoving.com/blog/living-in-moving-to-san-diego/

McKee, K. L., Crandell, I. C., & Hanlon, A. L. (2020). County-level social distancing and policy impact in the United States: A dynamical systems model. *JMIR Public Health and Surveillance*, *6*(4), Article e23902. https://doi.org/10.2196/23902

Menkir, T. F., Chin, T., Hay, J. A., Surface, E. D., De Salazar, P. M., Buckee, C. O., Watts, A., Khan, K., Sherbo, R., Yan, A. W. C., Mina, M. J., Lipsitch, M., & Niehus, R. (2021). Estimating internationally imported cases during the early COVID-19 pandemic. *Nature Communications*, *12*(1), 311. https://doi.org/10.1038/s41467-020-20219-8

Nazario, B. (2021, December 22). *Coronavirus (COVID-19) overview*. WebMD. https://www.webmd.com/lung/coronavirus

Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K. E. C., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N. F., Cattarino, L., Cooper, L. V., Coupland, H., Cucunuba, Z. M., Cuomo-Dannenburg, G., Dighe, A., Djaafara, B. A., Dorigatti, I., Eales, O. D., van Elsland, S. L., Nascimento, F. F., . . . Donnelly, C. A. (2021). Reduction in mobility and COVID-19 transmission. *Nature Communications*, *12*(1), Article 1090. https://doi.org/10.1038/s41467-021-21358-2

Portilla, R., & Heintz, B. (2019, April 30). *Understanding dynamic time warping*. Databricks. https://databricks.com/blog/2019/04/30/understanding-dynamic-time-warping.html

Praharaj, S., & Han, H. (2021). Human mobility impacts on the surging incidence of COVID-19 in India. *Geographical Research*, 60(1), 18-28. https://doi.org/10.1111/1745-5871.12502

Statista. (2021). *Global Twitter user age distribution 2021*. https://www.statista.com/statistics/283119/age-distribution-of-global-twitter-users/

Statistical Atlas. (n.d.). *The demographic statistical atlas of the United States*. Retrieved January 18, 2022, from https://statisticalatlas.com/county/California/San-Diego-County/Population

U.S. Census Bureau. (n.d.). *QuickFacts: San Diego city, California*. Retrieved January 19, 2022, from https://www.census.gov/quickfacts/fact/dashboard/sandiegocitycalifornia/PST045221

U.S. Food and Drug Administration. (2021, August 23). *FDA approves first COVID-19 vaccine*. https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine

Wang, B., Liu, J., Li, Y., Fu, S., Xu, X., Li, L., Zhou, J., Liu, X., He, X., Yan, J., Shi, Y., Niu, J., Yang, Y., Li, Y., Luo, B., & Zhang, K. (2020). Airborne particulate matter, population mobility and COVID-19: A multi-city study in China. *BMC Public Health*, *20*(1), Article 1585. https://doi.org/10.1186/s12889-020-09669-3

Wang, S., Liu, Y., & Hu, T. (2020). Examining the change of human mobility adherent to social restriction policies and its effect on COVID-19 cases in Australia. *International Journal of Environmental Research and Public Health*, *17*(21), Article 7930. https://doi.org/10.3390/ijerph17217930

Wei, Y., Wang, J., Song, W., Xiu, C., Ma, L., & Pei, T. (2021). Spread of COVID-19 in China: Analysis from a city-based epidemic and mobility model. *Cities*, *110*, Article 103010. https://doi.org/10.1016/j.cities.2020.103010

Wong, D. W. S., & Li, Y. (2020). Spreading of COVID-19: Density matters. *PLOS ONE*, *15*(12), Article e0242398. https://doi.org/10.1371/journal.pone.0242398

Worldometer. (n.d.). *Coronavirus death toll and trends*. Retrieved January 12, 2022, from https://www.worldometers.info/coronavirus/coronavirus-death-toll/

Yechezkel, M., Weiss, A., Rejwan, I., Shahmoon, E., Ben-Gal, S., & Yamin, D. (2021). Human mobility and poverty as key drivers of COVID-19 transmission and control. *BMC Public Health*, *21*(1), Article 596. https://doi.org/10.1186/s12889-021-10561-x

Yuan, X., Hu, K., Xu, J., Zhang, X., Lynch, C. F., & Zhang, L. (2020). State heterogeneity in the associations of human mobility with COVID-19 epidemics in the European Union. *American Journal of Translational Research, 12*(11), 7430-7438.

Zeng, C., Zhang, J., Li, Z., Sun, X., Olatosi, B., Weissman, S., & Li, X. (2021). Spatial-temporal relationship between population mobility and COVID-19 outbreaks in South Carolina: Time series forecasting analysis. *Journal of Medical Internet Research*, *23*(4), Article e27045. https://doi.org/10.2196/27045