

UC Irvine

UC GIS Week 2020

Title

Public Health

Permalink

<https://escholarship.org/uc/item/8dt8c7r6>

Authors

Naughton, Colleen

Yang, Jiue-An

Oh, Debby

Publication Date

2020-11-18

Supplemental Material

<https://escholarship.org/uc/item/8dt8c7r6#supplemental>

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at <https://creativecommons.org/licenses/by-nc-sa/4.0/>

UC GIS Week
Wednesday, November 18, 2020
1 PM – 2 PM
Public Health

Sessions:

- Integrated Risk Assessment and Geospatial Analysis for SARS-CoV-2 in Wastewater for Vulnerable Populations | Speaker: Colleen Naughton – UC Merced | Pg. 2 | Video Timestamp: 00:02:16.530
- Visualizing Preterm Birth in the Bay Area | Speaker: Jiue-An (Jay) Yang – UC San Diego | Pg. 11 | Video Timestamp: 00:24:12.750
- Mapping Health Inequities: UCSF Health Atlas and California Health Maps | Speaker: Debby Oh – UC San Francisco | Pg. 18 | Video Timestamp: 00:36:24.120

I'm the moderator for this session.

We have

We're supposed to have three like a more speakers here to talk about to give us three talks, but kinda we have to we have speakers for the two talks.

For now we will have two talks for today's session. So it partitions around 15 minutes and we'll give time for questions at the end of this public transportation

There's a q&a window. If you look at the bottom. As of the screen, feel free to ask Questions.

Window and we'll ask the secrets to answer the questions and welcome the speakers from

UC San Diego, the day going to talk about a topic. I'm going to talk about utilizing a preterm birth DB area. So the speakers are Dr. Yang his that's your special research scientist and developer as Qualcomm Institute UC San Diego and also Marta Jankowska.

She is a system resource centers and geospatial just and just specialist as Qualcomm Institute UC San Diego. So let's welcome our speakers, I will you have time to YouTube.

And you can share your screen.

Okay.

Yes, feel free to share your screen, it will replace my slide with yours.

Are our present presenters from UC Merced that are here. I'm currently working on the technical side to get them ready, but let's

Let's start with

Here we go.

Excellent.

Okay.

And you guys can see my main presentation right now. It's a Presenter view.

That's correct.

Visualizing Preterm Birth in the Bay Area

Speaker: Jiue-An (Jay) Yang – UC San Diego

Abstract:

Preterm birth is the leading cause of death under the age of 5, and can lead to lifelong health issues. Its complex causes are associated with genetic, biological, environmental, and sociodemographic factors. We present a geographic data visualization platform to explore rates of PTB and related risk factors in understanding PTB in the Bay Area. This is an essential step for pattern exploration, hypothesis generation, and presents opportunities to increase public and stakeholder involvement.

Transcript: Video Timestamp: 00:02:16.530

Okay, great. All right. Hi, my name is Martin Costco. I am here today with Jay being

One of our partners in this project disco block couldn't make it today, but we're going to be presenting some of the work we've been doing with the

UCSF preterm birth initiative and discuss how we have tackled visualizing their data and associated risk factors. Today we're gonna be focusing on the Bay Area, but this project has spanned a couple other geographical regions as well. So the California preterm birth initiative is

A project that started a couple years ago and the mission is to eliminate racial disparities and preterm birth and to improve health outcomes for babies born to suit.

And a lot of people are not aware that preterm birth is actually quite a big public health issue.

Specifically in California, but throughout the United States. So one in 12 babies in California are born prematurely. And this can lead to all kinds of different complications of health, it can lead to

Everything from delayed learning and delayed milestones to come respiratory issues and problems along development and in some cases can even even lead to death. So it is a really big problem that has long term health impacts and the really

Tough part swallow about this problem is that there's a huge racial and ethnic disparities approved from birth. So black women are 47% more likely to give birth prematurely.

And this is a disparity that the California pushovers initiative is really looking to try to solve.

So I'm where our team comes in from UC San Diego and the health data at scale collaboratory we were approached by UCSF to help them take a bunch of the data that they had

On and link it back to various environmental democrats a graphic and socioeconomic factors. So

We wanted to provide some kind of public platform that would make it easy for people to visualize the data and navigate the data and something that's accessible to both non researchers and researchers and interested Members of the public.

Work is a really complex issue. There's several factors that come into play when we talked about preterm birth and it's it's not just genetic. It's not just the health of the mother. It's not just the environmental exposures.

All of these things work in tandem to result in potential prematurity so while we know a lot about these individual factors on their own. We really don't know much about how these factors interact and how various environmental and contextual exposures come together to

Compete in a way compound or risk, the risk that a mother may experience from having several these factors at once.

So, and the other thing is that there could be environmental contextual factors that we don't even know about that are related to preterm birth.

And particularly as a spatial problem because what might be affecting preterm birth in one area may not be affecting it and another so certain pesticides or certain exposures may only be occurring in certain places.

So why, why did exploration and why data visualization. So exploratory spatial data analysis really

Has been a field that's been around for a long time, and the whole goal here is to utilize geographically colleges to allow people to explore special patterns.

To compare breaks in one area with rates in a nearby area, area to really look at different clustering patterns.

And also to really help generate hypotheses. So questions such as, where and when to healthy and unhealthy exposures occurred that link down into preacher versus

How much of a certain exposures before actually impacts preterm birth and what are the different patterns we might be observing

And again, with such a complex issues preterm birth, having a lot of different environmental factors to look at once could really help researchers better understand and generate hypotheses about this problem.

So,

Sorry about that.

ended up developing and Jay, my co presenter here is going to show you guys around that in a minute. The tool that we developed was a really iterative process and it really

It took a lot of meetings, a lot of trials, a lot of errors. And it turns out it's it's really challenging to generate a data visualization tool that

Actually gives you good informative and deep data to dig your teeth into, but it's also accessible to public members and stakeholders.

So we went through several meetings with stakeholders in Fresno in San Francisco, Oakland. Those are the two sides for our tools are focused

On because these are two areas that have the highest rate of freedom birth in California, and we went through several meetings. We did a bunch of iterations on the website.

We also had a bunch of focus groups on what topics to include in on the website. So you'll see soon. We have a number of different data topics and these weren't just chosen arbitrarily, they really were.

Chosen through community member and stakeholder feedback. And when I say stakeholder. I mean, people that were hospitals and NGOs.

People who are actively working with mothers kind of colleges, etc. And then once the site was up and running, we had a several workshops that we did with members of the community, both in Fresno and Oakland, both to teach people how to use the site, but also to get feedback on the site.

Some of the results of this research has been for kind of key products. The first two are the Fresno, and the Bay Area data visualization portals.

The third is actually a open and available source code for the websites. So it's not going to be specifically the website, but it will be pieces of code that you can use to generate their similar

Visualizations with your own data. If you want to build your own site. On a different topic or even the same topic. That's all available on GitHub.

And then we wrote a paper that was published in the CDC journal preventing chronic disease about this whole process about how the stakeholder work went and then about the website itself and what it includes

So the data that we use for the site and it gets a little more in depth into this, that kind of crown jewel of the data is this OSH pad first cohort and hospitalization data.

So this is a California agency that basically every time of birth happens in a hospital, but even sometimes not in a hospital. Every time right happens, it gets reported. So there's this mass dataset births going years back in 2011 is not that really as possible year

Of birth burst and then information about that birth, whether the child was born prematurely any issues that the child has and then those

Those data are links to mother hospitalization records. So I knew as a mother was hospitalized previously with issues such as potential and

We have that information as well. And then we pulled in a ton of other data sources. So things like the column bio screen from the EPA

Census and American Community Survey data several local and national data sets, things like walkability crime access to fresh foods. These are all different types of indicators that we've tried to include in the tool.

Alright, so

I'm going to now turn it over to Jay

And he's going to give us a tour of the of the tool.

So jump right into the data exploration platform that we developed for the Bay Area in San Francisco area.

And you, you can start from the candidate introduction video that talks about because there's a collaboration between the UCSF return verse initiative and us as you at UC San Diego. So you can look at the introduction video to understand the mission and the goal to pretend person initiative.

Looking into more details connects to their website as well and as Mark mentioned, our goal is to really create a data into

On topics that we believe and throughout discussion and discussion with community members and domain expertise to create separate different topics that helps us to start by exploring data and then

Maybe cook up a hypothesis for NGOs or for researcher to looking to return burst in different domains. So you will see

The export type experience you tonight, different topics ranging from first data demographics environmental data pollutions all the way to index is created by our research as well.

And on the website, you will always be able to look at the frequently asked question that we collected throughout the workshops and

Empire feedbacks sending for people that this might be useful information that to help you navigate to the websites. There's always the feedback option if you see anything.

That triggers your, your eureka moment you want to use the data you want to understand or even collaborate with us against industry back and we'll get right back to that.

So let's look into some of the topics as the example that's into the first data every time you click on a topic.

The data source that you will be seeing prompts to you in this pop up box so you know that you'll be looking at data from the ash Park cohort. And on the other data set that is within this topic tab.

And for each different topics we develop different visualization component that is better supporting this exploration on this topics.

For example, we want to look at percent of preterm birth. Each of the regions and want to compare to subtype.

Let's say percent of spontaneous between various each of the regions. So you have a map and chart side by side comparisons you hover over the South San Francisco area.

This is by zip code. See the night for once before zip code has a point seven 6% preterm birth under 37 weeks. You can also see how where this

Call rings among us by the percentage of spontaneous preterm birth in the region.

As you move over to a different region explore different zip codes, the Barroso changes location and works the other way to if you want to look at the highest

Area was the percent of spontaneous free to embarrass you can, you'll be able to see the Maskell kind of pen to the area of your interests.

And preach members. This all has different stages. My weeks on characterizations of disappointment advocate the first count either a week's work percent of

Pre tempers in each of the weeks intervals and every time you see this little boxes that visualization component can be future of want to look at only underneath 32 weeks that's in the highest region.

I 4621 we have no idea where is that, so just click on that in the map will pen to cells Oakland brown. The Oakland airport area. This is

We think that using these visualization component better supports exploring preterm birth on the first data science.

And if you want to look back again to the data stores, you can click on data stores prompt you again.

For interested in download the data tab, you can click on Download PDF data, it'll bring you to our shared drive and you will be able to tell that data and use it as you want.

On the select topic Bart jump to different topics we mentioned, there are some topics we're still working on it as new data coming in.

Every year we receive new purchase core data and what we constantly get new data from collaborators that use the site and see that. Well, I do have that data set. I want to on you guys to put it on there might be interesting. So

We continue constantly updating this data.

For example, on the demographic tap on your promise the data source again to the sidebar disappears, by itself, not to distract you.

And for different topics, please time different type of visualization to support better understanding in that data sets. Once again, we select percent of three turnovers hundred 37 weeks and you will be able to compare

Women giving preterm birth and women. Women given turnovers different subcategories based on ethnicity age group, education and place of birth.

If you again look at the South San Francisco area of the data will update on the right hand side so you can compare these things, side by side. So

You'll be able to the future, if you want to what he compare that same education under 2012 years in over two years and you'll see the same

Style side by side comparison we use that throughout for different visualization comparing the preterm birth rate and environmental indicator

That say walk ability as Martin mentioned so you'll be able to see this side is on the lower right hand side information parks, you can navigate through different regions and income here but explore on the indicator that interest you in a Western creature inverse data.

Let's go back to the homepage. One example of art cooperators interested in house insecurity as bar admission just spray chamber is

Really a racial disparity behind that in certain regions. So one example of using the data is to create a story map.

So we take some subset of the data and it create a story map of free time versus housing insecurity to really dive into this narrative and understand whether

The driven force behind that word details in Houston security and preach numbers. So, this

Story map. It's created purely using the data that we have. But you will be able to use the data and then create your own visualization as you want to tell the story about

Certain dimension or aspect of return. First, because this is really still very puzzling to the researcher. There are so many factors in play.

Think Different groups and user will find their weight of how to use our data to create visualizations and tell the story that they're interested in.

So, not only for the for the San Francisco and for us now. And this is for service code and Bay Area.

And the same size similar style. We also feel similar aside for the Fresno County on using slightly different design, but the same idea behind on me was, you will be able to explore different topics and also download the data.

So was this were showing that this visualization platforms really scalable, you can quickly replicated different regions, as we move on to different study region to understand the impact of

Environmental factors and demographic factors and how does that fear was preached embrace different regions. So we just see a potential of different

Researchers in different area and policymakers that we can engage with them utilize our concept to help them to five pretend verse in different regions.

And as we share the Cortana came up for a visualization components. So you to build your own similar style data visualization platform was teach

examples on how we actually do it in this way, so you'll be able to do the same for you and

As the FAQ mention that if you want to know more about how to use the site on our youtube channel of the health data a skillful collaborator.

There are videos that were recorded to help us navigate through the websites, then make the max out of using the websites.

Want to end with on the dissemination on the website. And this whole operations as Martin mentioned, we already host couple of workshops and events to

Talk with company member and engage with them and get feedback from the user and how we can improve the site and how they can use the data and decide to

Hello, sorry, I'm sorry. So interrupt. Now we see your

Your petition slide is is a is a no Simon can see all note here.

Walker, yes.

Maybe this

Is Martin mentioned that we already created.

Where the host couples workshop

A Fresno County.

At Oakland.

In with San Francisco.

Group.

To actually get the feedbacks from the community and the stakeholders and how engage with them actually use the site and get feedback know where we can improve

And very interested in the whole process on how we built the this collaborations to community. The paper is publishing the CDC, Journal of any disease and also it's been being reported in the press on being um it's pretty much our talk.

back tomorrow.

Yes, thank you, about talk. I think we have a question from the chat window. No.

phone calling she couldn't tell me the answers to the Q AMP. A. So I want to mention the question quickly in case the

We can see it as we have more chess here. So, are there plans to expand beyond to be area. Have you found Open Street Map preferable to as many projects have you found associations between environment nasty and other social economic factors.

So as you guys, so we did do the site for Fresno, and we may be doing a version for San Diego for now. Those are the three main local areas where we're planning on I think it is important with something as complicated as preterm birth. I mean, yes, you can look at California wide.

But I think you get a lot more information and context when you really focus in on a specific region. So for now, those are our plans to stick locally. Who knows, maybe someone throw money at us. We'd love to build one for all California, but

That's always an issue. And in terms of associations, yes, we are actually working on pulling in these different data sources to try to create some more complicated or just more complex modeling frameworks. So, ones that actually go at both environmental individual and genetic levels.

And there's been quite a few papers. If you just go to the UCSF preterm birth site, they have a list of their publications. Some of them are by us as well and vicious the wealth of information there. And I like to answer the question about Open Street Map places as three

Yeah, I think we started with the mind of using open source and development because we do want to

The site to be available for a lot of user to view their own realization. I think we started with beef that and open source libraries mainly focus the story map is just one example that how user can use the data to

Storytelling

Application, so yes.

Open Street Map as we map being map whatever map you have access to it doesn't really won't change what the visualizations that you're trying to achieve. So I think comes next night to find open source library, very easy to use.

And for the questions about GitHub link and story men, we will study in the chair.

Okay. Okay. Thank you. Thank you, Jay. And awesome. So we'll see if other questions at the end of the session. So feel free to type the question to the Q AMP a window. So next we will welcome the speakers for the next talk

So Ariana and Fernando

Fernando comes from UC Mossad, and his a undergrads majoring in

The chat. Sorry.

majoring in a human engineering and they will talk about the risk assessment and geo spatial analysis for stock for source code to invest water. So let's welcome Ariana and the financial

You've come to you guys.

Okay, I hope that's working.

Good.

Looks good.

Integrated Risk Assessment and Geospatial Analysis for SARS-CoV-2 in Wastewater for Vulnerable Populations

Speaker: Colleen Naughton – UC Merced

Abstract:

Wastewater monitoring for SARS-CoV-2 RNA has been shown cost effective to predict outbreaks of COVID-19 by 2-10 days. The virus recovered in wastewater can sample whole city populations with results in 1-2 days. Monitoring efforts are also being deployed in at least 27 countries and 88 universities but there is lack of a global repository of these efforts. Thus, we created a global dashboard of SARS-CoV-2 wastewater testing so the public can easily see where testing is happening in their area.

Transcript: Video Timestamp: 00:24:12.750

Good afternoon, everyone.

My name is Fernando Roman and I'd like to thank everyone for attending our

Presentation, I would also like to thank

The organizers and other presenters for making UCS as possible.

My research partner every week and I are both forth your environment engineering majors at the University of California or said

Throughout the semester we have been studying source code to behavior and wastewater.

With water based epidemiology and with washing up plants as a part of a UC center awarded coding 19 average CPC crack research team led by assistant professor calling out at night, using her said

I research was funded by citrus. The Senate for information technology research interest of society. Next slide please.

Our presentation will go over a brief history of wastewater based epidemiology and how it has been used to help combat. So I was cooked to

Will also showcase the senior center Global's are as close to his partner monitoring efforts dashboard and research involving wastewater treatment plants and disadvantaged communities within California. Next slide please.

This presentation will not have been possible without the help of our research team.

Our research is a collaborative effort coordinated by UC Merced UC Davis, our team consists of ASSISTANT PROFESSOR DR mountain from you. Similar said

Assistant professor, Dr. Can you up from UC Davis graduate student jihad from UC Davis graduate student on a grace for the summer said

On with arigato wiki Matt Damon and I pictured on the bottom right, that he was being environmental engineering undergraduates senior said. Next slide please.

What's, what are based epidemiology, also known as soon as it relates monitors levels of targeted targeted biomarkers and chemicals within wastewater to analyze threats within the population.

It has historically been used as a tool to monitor illicit drugs, along with pathogens, such as poliovirus

Recently with water based epidemiology and gather a lot of attention since it can be utilized to monitor levels of source code to within municipalities.

Up to over 50% of people both symptomatic and asymptomatic sheds just cooked you under feces scientists and engineers can detect the RNA Periscope to the effect of people have shared by sampling wastewater and those samples can show transit source code to levels within a population.

Since affected individuals should the viruses early symptoms as an early symptom.

Groups analyzing wastewater can help predict outbreaks. The czar's co to do to 14 days prior to them actually happening along local authorities to help stop spreading the virus.

Well we swear to base epidemiology is not meant to replace individual testing it works well alongside clinical trials and conceptual a large pool.

For relatively low cost. Next slide please.

Many universities and public authorities around the world are using service surveillance, the guy. Look at policies are coded protective measures.

YOU TELL US Department of Environmental Quality has a very in depth to a certain system and monitor as many ways question implants consumers sample sites around the state.

Another good example is the University of Arizona who through testing the dorms wastewater were able to find an isolated incident asymptomatic students who have the virus.

Inspired by the Johns Hopkins coronavirus dashboard inspired by whisper based epidemiologists our research team or user sets out to create a dashboard showcasing Global's our scope to waste water monitoring efforts next language.

Utilizing as accurate as online dashboard tool, we were able to create and share dashboard, which visually summarizes global gyroscope to a spotter monitoring efforts.

Anyone can view the dashboard and see universities countries and sample points vocations using this quarter based epidemiology.

But zoning in on an area in the dashboard, California. For example, what can see that many universities are not testing the wastewater universities, you know, like UC San Diego UC Irvine UC Santa Barbara consumer said UC Berkeley and UC Davis. We also added links to news articles dashboards with water data and publications from any of our points. Next slide please.

We're card a continuance of continuously update our map to find your points we are constantly monitoring slacks. The coven 19 with water based epidemiology collaborative

Attending webinars searching keywords on Twitter conducting Internet searches really publications interviewing request forms from institutions, I would like to be added.

a Twitter account at covert groups 19 is also tagged on a lot of Twitter posts and with the growing number of followers, we have been able to share our map to up again academics what enthusiasts and the general public, like

Now, like to pass it on to area and I will talk about another aspect of a research associate plants in California.

Hello. So I'm really excited to be presenting to all of you today at UC GIS week

As Fernando already mentioned, my name is Arianna Tariqi, and I'm a fourth year undergraduate I use humor said. And today, like to present to you our research on wastewater treatment plants in California.

And their proximity to residential and commercial areas using jazz principles.

So while the potential of stars Kofi to spreading through fecal oral and aerosol born transmission is still being debated. We've been researching wastewater treatment plants in California.

And their proximity to nearby neighborhoods to see if there's a potential risk for aerosol born transmission during the early stages of wastewater treatment. This research is all

During the early stages of wastewater treatment. This research is also beneficial for understanding potential risk of other pathogens and wastewater treatment plant proximity to our workplaces in homes.

So in order to conduct our research, we first need to know the number of wastewater treatment plants in California and their locations.

We started by combining two public data sets. The first thing the Environmental Protection Agency's facility registry service pictures.

pictured as the small enough on the bottom left. The second source was Kelly pH state water resource control boards facility at a glance web page.

And we were able to look through most of their list of locations and find wastewater treatment plants in California.

From here, we looked up location addresses on Google Maps and notice that latitude and longitude points weren't right on top of treatment plans and for researchers imperative that we had the locations.

In the right areas. So the 300 wastewater treatment plants were split among researchers in our group and searched in Google Maps and given longitude and latitude points that corresponded to the middle of the plants.

We also conducted measurements within Google Maps to find the wastewater treatment plants that are 250 meters away from residential or commercial areas.

turned a 50 meters is used as a base measurement, since you found that in California. There are no regulations for the distances that wastewater plants may must be in relation to communities.

This is unlike some other countries, and even some other states in the US who do have published regulations.

Studies have found that communities nearby wastewater treatment plants may be prone to adverse health effects such as headaches respiratory issues, tiredness difficulty concentrating and skin diseases.

Even if starts moving to cannot be spread efficiently through by aerosols, it's still important research to understand

For the potential other pathogens spread from wastewater implants during early stage treatment and how they can potentially affect nearby can be communities.

Many ways for two implants in California are less than 250 meters away.

/And from residential areas and we're built their prior to modern environmental regulation laws, and this is why we decided to research.

A 250 meter proximity for wastewater treatment plants. So we upgraded our total Excel onto our job is to get our visual map of wastewater treatment plants in California, which is what you see on the right of the screen and we managed to find a total of 305 wastewater treatment plants.

Okay, so the next part of our research is not being wastewater treatment plants in vulnerable communities. To do this, we use the calendar screen version 3.0

The map is displayed here with a darker orange and red areas being disadvantaged communities. So what is a disadvantage community.

Well, brief definition is a California defined area that suffers a combination of economic, environmental and health burdens also pictured is from the report on the converse screen are the different attributes listed within this data set.

These can be pollution burden population characteristics environmental effects and socio economic factors and all the characteristics within them that you see in the picture.

scores are kick calculated and areas with high amounts of burden or the top 25% of the calculated column virus score those days and report can be found on

Website, on behalf of Kelly EPA

Yeah. So in our chairs pro within calculated the intersections of wastewater treatment plants in disadvantaged communities.

Of the 37 wastewater treatment plants in the Bay Area pictured here on the slide. So from these statistics. We found that 84% of wastewater treatment plants are less than 250 meters away from residential or business areas.

We found that of these wastewater treatment plants 38% were in disadvantaged communities and of the 14 that are in disadvantaged communities 12 of them were

Less than 250 meters away from residential or commercial areas and these results are preliminary so we're still working on expanding it to all of California.

So some next steps for ongoing and future work. We're going to use ArcGIS Pro to buffer 250 meters away from each wastewater treatment plant in California.

By using either land use data or only data, we can see where buffers intersect with commercial or residential areas and you can kind of see that in the picture here in the middle.

And this will give us a better automated system to find proximity of wastewater treatment plants to communities.

Once we can automate these results will do the same analysis, but with buffering 500 meters away from so other literature, other countries regulations.

Other steps we're currently working towards is mapping areas where sanitary sewer overflow incidents occur.

This can be. This can give us an idea of areas that are more vulnerable because sanitary sewer flows can be overwhelmed who are flow. Sorry, can be caused by different things such as blockages or defects.

Powerful failures or improper sewer designs. These can potentially threaten public or may throw and communities public health because of potential contamination from center sewage.

Another step to our plan is to follow the similar methodology explained before, to find wastewater treatment percentages and urban and rural areas to see if there's any statistical significance of these different locations.

Yeah, so that kind of concludes our presentation. We would like to acknowledge you all as well for putting on this amazing event. And of course, researchers at UC Merced said and UC Davis.

This work will not be possible without them. Thank you to citrus for funding our research and to wastewater workers. Thank you to frontline workers and everyone who's been sampling to make our research possible and thank you to the coven 19 wastewater based epidemiology collaborative

So these are references. And yeah, so if you have any questions we'd love to answer them. And please feel free to follow our Twitter at Kovac poops and friendly reminder to only flush the three P's and thank you.

Thank you.

It's

It's really

Interesting name Kobe.

Oh, it's been comments that

We received the questions at

You so much

Yes, so we will save the questions as an as an end of the

Free to tapping the question to a Q AMP a window.

So next we welcome our last speaker, Debbie.

Debbie is a data scientist from UC San Francisco's department of epidemiology and final sentence about statistics and she will talk about maybe healthy evidence. So that's welcome, Debbie.

Mapping Health Inequities: UCSF Health Atlas and California Health Maps

Speaker: Debby Oh – UC San Francisco

Abstract:

Where we live, including the physical and social structures that surround us, influences health in both positive and negative ways. Combining neighborhood-level data with individual-level health data can give researchers a fuller picture of what causes health inequities, and what can be done to address them. Health Atlas aims to provide researchers with an online platform with visual tools to help researchers apply population health principles to achieve health equity.

Transcript: Video Timestamp: 00:36:24.120

We can hear you. Yes, sorry I my there's too many windows here. So now, now you can hear me. Right. Okay.

Let me see my screen again.

Alright, how does that look

If it looks good.

Okay.

So hi, everyone.

It's great to be able to share with you some of the work that we've been doing in mapping health inequities. So today I'm going to be sharing two of the projects that we've been working on over the past year or so.

My name is Debbie. Oh, and I'm from UCSF, the Department of epidemiology and biostatistics. My background is in epidemiology my title is data scientist but

I think I kind of fall somewhere in between. I also have a background in graphic design. So I'm kind of able to integrate that all into the work that we do in mapping

Okay. So we all know now more acutely than ever that when it comes to health place matters. So the codependent because me that really absolutely clear.

Some populations are clearly disproportionately affected the recent wildfires in California has also made this really painfully obvious. You know, if you were

Constantly refreshing purple air during the wildfires, like me, you know, you know that where you live affects your health and it affects the way that you can live in the way that you can keep yourself healthy

So about a year ago, we decided we wanted to make a tool that researchers could use to better understand factors influencing health at a population level. So we wanted something for researchers to be able to use for like grant writing or study planning hypothesis generation.

So well, the focus was supposed to be for researchers we we understood or we still understand that research alone cannot make public health progress.

So we hope to create something that could also be used by policymakers community advocates and public health professionals, so that they could, you know, take this tool and make data informed decisions in their work.

So what we created was an interactive map of over 100 variables to explore neighborhood level characteristics.

We spent most of last summer kind of thinking about the vision and we officially kicked off project in October 2019. So about a year ago, we were planning to wrap up development in March and then launch in early April, but as you all know, the coven pandemic kind of up ended things

And so we are planning to just kind of release initially released this map with these neighborhood level variables. But at the very end.

We were asked to add Cobra data onto the map. And so we connected with the Los Angeles Times, who has been collecting Cobra data on a daily basis from the 61 different health departments in California.

And so they have graciously shared their data actually not just with us, but it's kind of open source, everybody can access it on their GitHub. So we have

coded case and death data from LA Times hospitalization data is from the California Department of Public Health and a new addition, as of just like less than a week ago. We also have coded in nursing facilities and Kobe in prisons and that's also from the LA Times.

So we collected data. In addition to, you know, that corporate stuff we collected data from many different data sources. I think it's like close to 20

Loosely based on a social determinants of health framework. So we categorized into demographic socio economic community neighborhood and health and healthcare.

Inner map we kind of kept the cobra data separate just to make it easier to find. And you'll kind of see the maps and a little bit

So here's the list of the different places that we grab data from and you may recognize them since it seems like some of the other people presenting day also use kind of similar sources and these are really rich you know resources that anybody can and should use to inform their work.

I would say that of the list, the ones that are in Boulder were the ones that we used most so American Community Survey Los Angeles Times CD pH and 500 cities.

So in our maps. We have three different levels of geography. We have census tract place and county level.

So County, you know, county is something that most people can kind of wrap their minds around census tract.

I would assume or my guess would be that most people in the audience would understand census tracts, but just in case you've never heard of it is census tract is a geography that the Census Bureau.

Has kind of created that supposed to contain around 4000 people around 4000 people live in that area.

The places data was put together, specifically to match the way various departments of public health.

Have been reporting covert data. So it's a little bit of a patchwork so in San Francisco and zip code, but in Los Angeles, it's neighborhood and you know it's we kind of custom built it so that we could show CO, but at a more granular level.

Some of the other features are and I'll walk you through these one by one.

You can compare two variables. You can see database distributions. There's a custom area selection. You can export the data and we have detailed data descriptions and we also have data stories.

So here's an example of what a comparison of two variables looks like in health Atlas.

Here we're looking at unemployed and uninsured so darker blue indicates higher levels of unemployment darker yellow indicates higher levels of an insurance.

darker green indicates higher levels of both. So I should note that for this map, it's two variables laid on top of each other. It's not filtered data.

So when it's dark green. It doesn't mean those people, people in there are both unemployment uninsured and it's just those two variables laid on top of each other.

Okay, so I just popping up the scatter plot a little bit. I just wanted to highlight the other visualizations, besides the map. So on the left hand side, you see a histogram and that is for like the California State distribution.

And then in the scatter plot. You can see also again California State distribution. This one in particular is for census tracts.

And so if you see a black on the map that you're interested in. If there's like an outlier. You can click on it and then the map will kind of take you to that place.

The scatter plot also serves as like a legend for the map.

As well as the history

OK, so we also have a custom selection feature, which I think is pretty useful, especially if you're interested in specific regions or neighborhoods.

So say for example you're interested in looking at coven data for the South East part of San Francisco.

So you can select kind of the two zip codes for that area and get custom data estimates for almost all variables. So this is just a screenshot. If there's extra time, I can do a demo, but I, I suspect there might not be

But if you, you can see that you we have here coven cases per 100,000 people. And then we have extremely low income households.

You can see for that neighborhood that the case rate is actually higher than the state average. So the orange is for that neighborhood.

And the blue is for California. And then below in the histogram. You see that little orange line and that indicates where that neighborhood falls in the California distribution.

Also in the scatter plot. You see that little orange dot. And you can see that yeah really you know that neighborhood has relatively higher levels of extremely low income households, as well as relatively higher

relatively higher Coby cases per 100,000 people

So it's a lot of information kind of packed into one place

If you're interested in exploring more or a pending the data to your own data, then you can export it. And I just wanted to highlight here that you can hire you can download

Data for a custom area. So while the map shows just the two, it actually will give you data for any other variable, almost. Some of them because they're indices are the denominators are kind of more challenging.

Not all variables will come at the neighborhood level or the custom area level, but almost all of them are available and they are estimate. So their population we waited estimates are household weighted estimates.

Now, if you're interested in learning more. This is for maybe the researchers in the audience, you can click on the little i and then

There's information about it, significant sources and some references. This is maybe particularly helpful if you're like writing a grant or just like trying to understand how you can maybe use this data for your own work.

And lastly, if you're interested in exploring a one of our data stories, you can click on explore data stories, the top right hand corner.

Currently we have six stories. There's two encoded one on food insecurity one on housing insecurity one a lot. Next ethnic enclaves and then one that's like an introduction. So it's kind of like a tutorial on how to use our site.

And just wanted to recognize that health Atlas has been a collaborative effort. It was developed with support from the UCSF School of Medicine's dean's office of population health and health equity we got expert guidance from some faculty at UCSF for do stories.

And we're also fortunate enough to have the creative and technical genius of statement design as well as writing as well as writing expertise from plain language health

Okay, so on to the second project. The second project. I want to share with you is our California Health maps tool just as a little background for this project.

UCSF houses the greater Bay Area cancer registry and that i mean it's it's an incredible public resource.

It's like the closest thing that we have to like Scandinavian health records in the US, like every single cancer case is recorded, and there's information.

And there's a lot of rules about how you can use it but it's incredible resource. Currently, however, the data is presented in a style that is

In a format that like researchers and epidemiologist are very familiar with, but might not be as accessible to the general public.

Also, generally speaking, the data is available at the county level as a smallest level of geography, however, we've we've felt an increased demand for subcounty data just to make it more meaningful for public health and policy efforts.

So the goal of this project was to develop an interactive mapping tool of health data for geographies beyond the county level in California.

Right. HEY, DEBBIE. I'm sorry, I interrupted, but we have limited time we have 10 minutes left, so

Sure, I'll just try to get the I have like seven more slides.

Sure, sure.

Yeah.

We have seven geographies, I'm

So zones were let me just say that zones were built specifically for for this type of visual visualization specifically for cancer data and we went through a whole process to make those. But there's also an SSH, which is a California

Geography unique to California, but there's also all these different legislative district as well as county

You can see cancer incidence for 12 of the most common invasive cancer sites and you can filter it so this

Is different than health Atlas in that you can filter by sex and race, ethnicity, and there's also a lot of other data on demographics, you know, other things that might be relevant to cancer outcomes.

So let me just walk you through what it looks like. So here you can select your area type, you can search where you want like the look for your location. You can choose either five year or 10 year cancer incidence

You can choose your cancer type

Filter by sex or race, ethnicity,

So here I have chosen County and lung cancer. And if you look at Los Angeles. It looks like long haul sandals has pretty low rates.

However, if you switch to zone. You can see that it is not uniform

And in a place like Los Angeles, which is really diverse and actually kind of really, really populous in geographically large you would expect to see differences.

Then I zoomed in here on Long Beach and kind of selected one of the zones.

And loops and you notice and excuse me that the rates are higher. So it's actually more purple. You know, so it's like a darker color you know as Long Beach like something there are higher rates, there is higher rates of

LUNG CANCER IN LONG BEACH AND YOU CAN under the map. There's a little bit more detail. You can see that for that particular zone. It has a higher incidence rate of cancer compared to the state you know 5055 compared to 41

You can also download the data if you want

You can also download an app in case you want to put it into a grant or something.

You can also print the page. So this might be useful if you're going to talk to, like, a community partner, you just have a print out and you know

Go talk to people in Long Beach and say hey what's going on in terms of lung cancer and, you know, you know, what are some things that are happening in community, you know, what can we do to change things.

So this is just a little close. I just zoomed in a little bit on the additional data that you can get. So there's demographics healthcare access and use health and well being.

And then lastly, at the very bottom is the Learn section where you can get more detailed information about the project methodology FAQ and definitions.

So this work has been a collaboration with the National Cancer Institute West dad.

And our key player in this project was green info network who helped us build out the website.

And it was so well received that they were actually asked to make it open source and available to other cancer registries, so that is available to others to use.

However,

The cancer registry data is usually owned by the by the individual registry. So it'd be more useful for people within working within the cancer. Cancer world. Um, and I think that's it there's maybe seven minutes for Q and. A for everybody. So thank you.

Q&A

Danny and I think we have a question from the chat window so Connie asks just export inclusive. So when you mentioned you the the layout. The two on online to you can export data.

So the question.

Hello, Danny. I couldn't understand what you said you're a little

Cut off. I'm sorry. So you show us

That online to the first the first project, we can export data from the online Mac.
Right, right.

Export Data. So does the export enclosure file or is only the attributes.

Oh, it's only the. Yeah. So it sounds like a CSV file, it's, it's not the the shape file, but
if you're interested. I mean, we're happy to share. It's

The sheep files. The only one that might be of interest might be places, but that's
maybe very, it's very covert specific

OK. OK, I see. So we have another one. So do you have any improvements you like
to make going forward or no. Do you want to put it on GitHub.

So are we. So we're talking about

The

Both pride or

The project I think both the data or the the online application that the two
probably for my understanding.

Okay, so, so both projects we are wrapping up actually around a development for
this year. I know that moving forward. We will continue to update the cancer
registry data.

Health Atlas. We've been talking

To various parties.

Considering using it, I think, well, we are piloting it actually in electronic health
record. Right now, we're also considering using it, maybe for other applications,
either for community partners or for patient navigators

Right now we're just wrapping up some things with like mobile improvements and, you know, smaller things. We added some new data layers and things like that so

There's a lot of ideas for the future, but we are currently focused on just like sort of finishing this round of development, but we are very open to new ideas and collaborations and I would love to hear from you.

Is the first project. The house has less map on can hop now.

That is not on GitHub. But if the the covert data is available through LA Times on GitHub.

Sure. I think we have another question for me. Anna and financial

So the question was, how, how did you choose which are public pollution data from call embarked screen to a three to us and do use all of it, or choose world feel pollution factors.

THE QUESTIONS FOR INDIANA Indiana and now

Well as efforts in the chat. We

Initially,

Did it with all the

Factors.

But we also mapped out

Lots of individual factors as well.

And even combinations of some

Trying to see any trends.

One interesting thing we did was for disadvantaged communities which represents 25% of the top 25% most you know at risk places we did it. We calculated that specifically for the Bay Area. By doing all the other counties and recalculating just the top 20 top 25 percentile for the bay.

Okay.

Okay, thank you.

I had one more question.

Where hold on one central question.

Yes, real

Do we have any

Gian. I have a question for Debbie in the chat.

And it was I'm

Sorry if I missed it. But how are the zones.

Determined in the

California Health maps.

Right, so I could probably do another 15 minutes just on how the zones were developed. So what we did. We use something called an easy tool and we put in a percent minority percent

urban, rural and poverty and tried to create these zones that were minimum population 50,000 maximum population 100,000, you know, we couldn't all these factors and we came up with a few versions of it and that we tested it with the community.

And so it was this whole process that we did in conjunction with

The National Cancer Institute and West that

I'm happy to connect more on that. I can't see where that question is from. But if you email me, I am happy to give you more information on how we created those zones.

Can ask question our Slack channel. I think all of us, our section right getting close the chat window of our Slack channel joined the channel and we can continue our conversation.

Yes, I'll repost that in just a moment. Join us on Slack to continue the conversation, if you like, that includes our presenters as well. Yes.

So personally, I have a question for G and Marta.

So we are displayed the pre the pre term birth.

You. I want to know the the premise data is about the host host like hospital locations or the, the actual like of the patients locations. So when you realize that map. So you get the data from our hospitals or it's about the actual location of the person.

If someone location of the mother.

See

I see.

Okay, awesome. And also I want to know the right bus you will realize a new horizons, like in aggregated into the phenomena. Since the tracker neighborhood level.

Think you're cut off a little bit. Can you

Repeat that question. Thank you. So the data you have these individual data right

Yes.

Data is individually came in as individual level, but

To to able to

Analyze and showed it

Those aggregate to

Geographic areas.

And also the IRB doesn't let us

Share individual data. So we have to activate it to you think, and have 10 just the minimum amount of birth.

Okay. Okay. Thank you so much.

Thank you for asking the question, thank you. Okay.

I think this is the end of work.

I think there's probably cut off of my voice.

Right, I

POS job seekers, so I'm although the sessions, but we can watch you to continue the conversation sectional is 2020 public house and also thank you so much for attending and thank you so much of our speakers so last meeting for

You. Bye.

Thank you very much. All

Thank you.