

# A DATASCIENCE WORKFLOW: Nonprofit Edition

Dr. Brian Spiering

# Hi, I'm Brian.

Faculty @GalvanizeU  
[brian.spiering@galvanize.com](mailto:brian.spiering@galvanize.com)  
@BrianSpiering



galvanize

# A Data Science Workflow

- ① Ask
- ② Acquire
- ③ Process
- ④ Model
- ⑤ Deliver

---

1

---

# Ask

Start by asking the right questions

# How can Data Science be a force multiplier for good?



---

2

---

## Acquire

Find the right team & obtain the proper data



Adam Crawley



Eric Williams



# Citizen Science for Tracking Flu

**REPORT IN**  
Share how you're feeling

The image shows a smartphone screen with a pink header. The app interface includes a logo with a virus icon and the text "flu near you". Below this is a section titled "Select Symptoms" with the question "What symptoms have you experienced in the past week?". A list of symptoms with checkboxes follows, with "Cough" checked. Other symptoms listed include Fever, Sore throat, Short breath, Chills/night sweats, Fatigue, Nausea or vomiting, Diarrhea, Body aches, and Headache. The next section asks "What day did you start feeling ill?" with a field containing "Monday, January 26". The final section asks "Did you receive a flu vaccination since you last took the survey?" with three radio button options: "Yes", "No", and "Don't know", where "Don't know" is selected. A blue "Next" button is at the bottom.

flu near you

Select Symptoms

What symptoms have you experienced in the past week.

Fever       Fatigue  
 Cough       Nausea or vomiting  
 Sore throat       Diarrhea  
 Short breath       Body aches  
 Chills/night sweats       Headache

What day did you start feeling ill?

Monday, January 26

Did you receive a flu vaccination since you last took the survey?

Yes     No     Don't know

Next

---

3

---

## Process

Create a structure to empower the team  
(and clean the data)



---

4

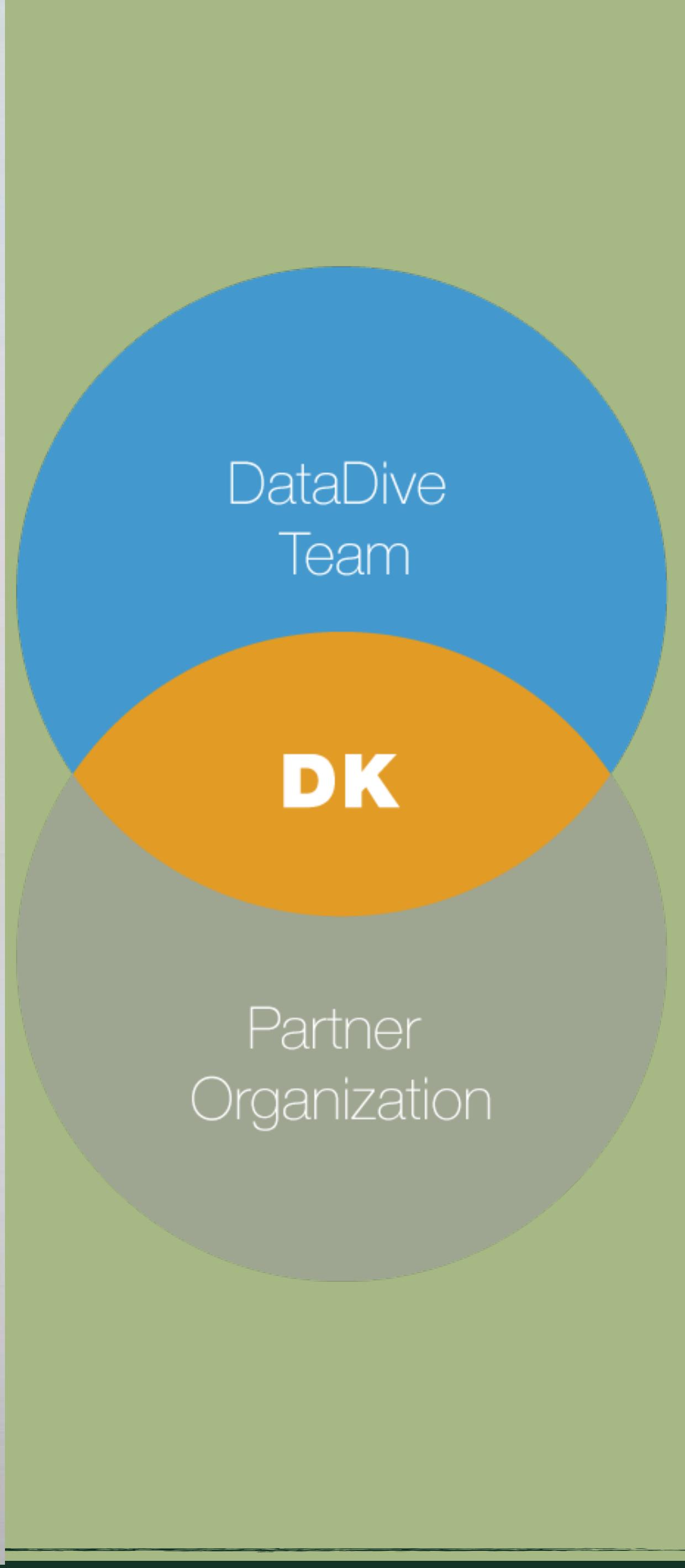
---

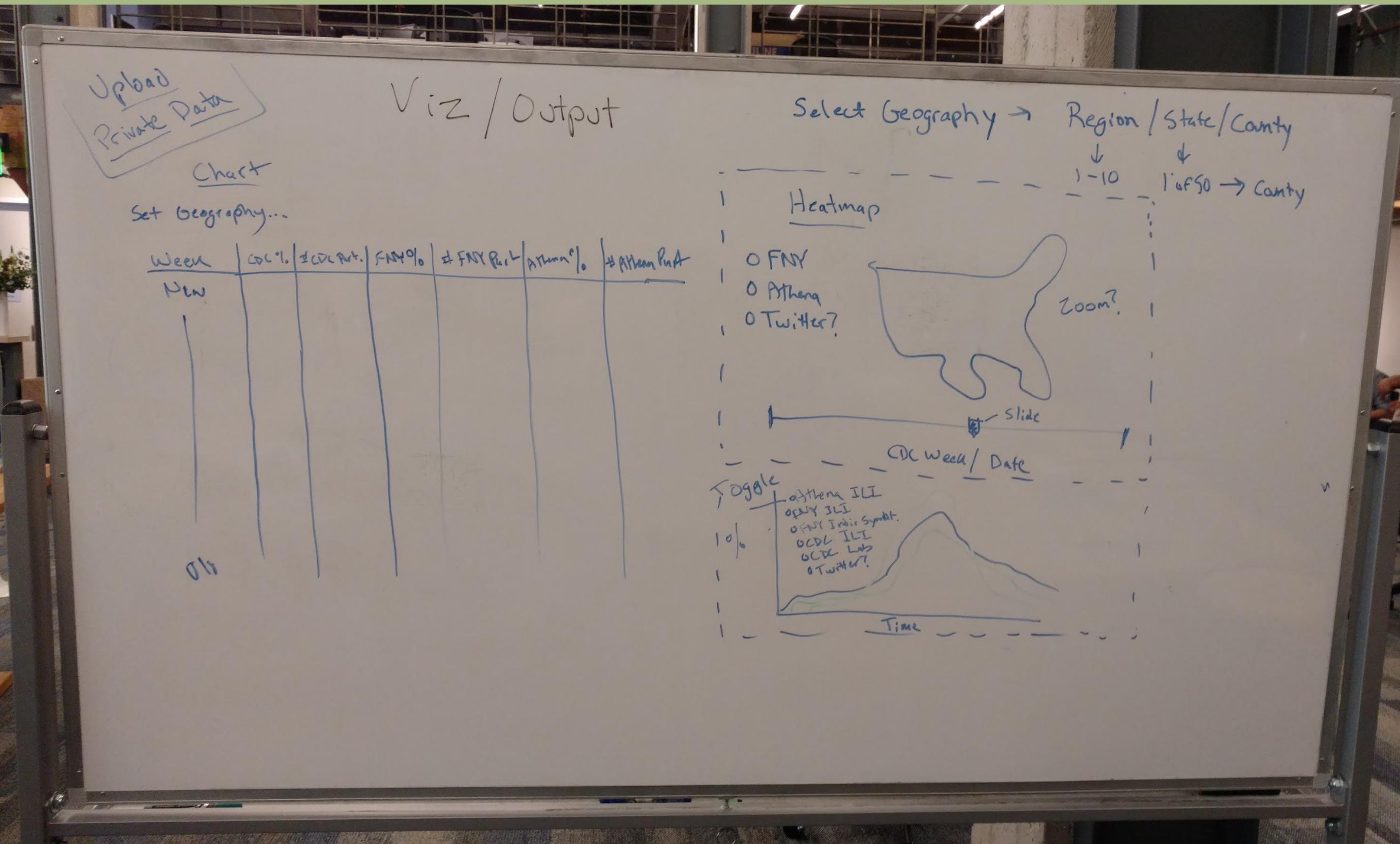
## Model

Leverage your team to find meaning in the  
data

#DATADIVE







```

names_to_r <- function(names){
  names <- str_replace_all(names, "^ ", "")
  names <- str_replace_all(names, " ", ".")
  names <- str_replace_all(names, "__", ".")
  names <- str_replace_all(names, "_", ".")
  names <- str_replace_all(names, "\\\.\\"., ".")
  names <- tolower(names)
}

#####
## Transformed data
load("./data/transformed/fny_ilinet_cdclab_athena_2013_to_2015.RData")

fips.codes <- read_csv("./data/fips_code_by_county.csv")
names(fips.codes) <- names_to_r(names(fips.codes))
#fips.codes$fips.state <- sprintf("%02d", strtoi(fips.codes$fips.state))
#fips.codes$fips.county <- sprintf("%03d", strtoi(fips.codes$fips.county))

tmp.states <- unique(fips.codes[,c("state","fips.state")])
fips.states <- as.list(tmp.states$fips.state)
names(fips.states) <- tmp.states$state

#fips.states <- unique(fips.codes[,c("state","fips.state")])
#fips.states <- as.list(unique(fips.codes$fips.state))
#the_51_states <- unique(fips.codes$state)
#names(fips.states) <- c(the_51_states[1:8], the_51_states[10:51])

# county.summary
load("./data/transformed/county_summary.RData")

county.summary.names <-
  names(county.summary)[4:length(county.summary)]

# table.summary
load("./data/transformed/table_summary.RData")

# region.summary
load("./data/transformed/region_summary.RData")
region.summary[is.na(region.summary)] <- 0
region.summary.names <-
  names(region.summary)[4:length(region.summary)]

#####
## County
county.shapes <- readOGR("./data/cb_2014_us_county_20m/cb_2014_us_county_20m.shp",
                           layer="cb_2014_us_county_20m")

county.shapes$county <- paste0(county.shapes$STATEFP, county.shapes$COUNTYFP )

```

---

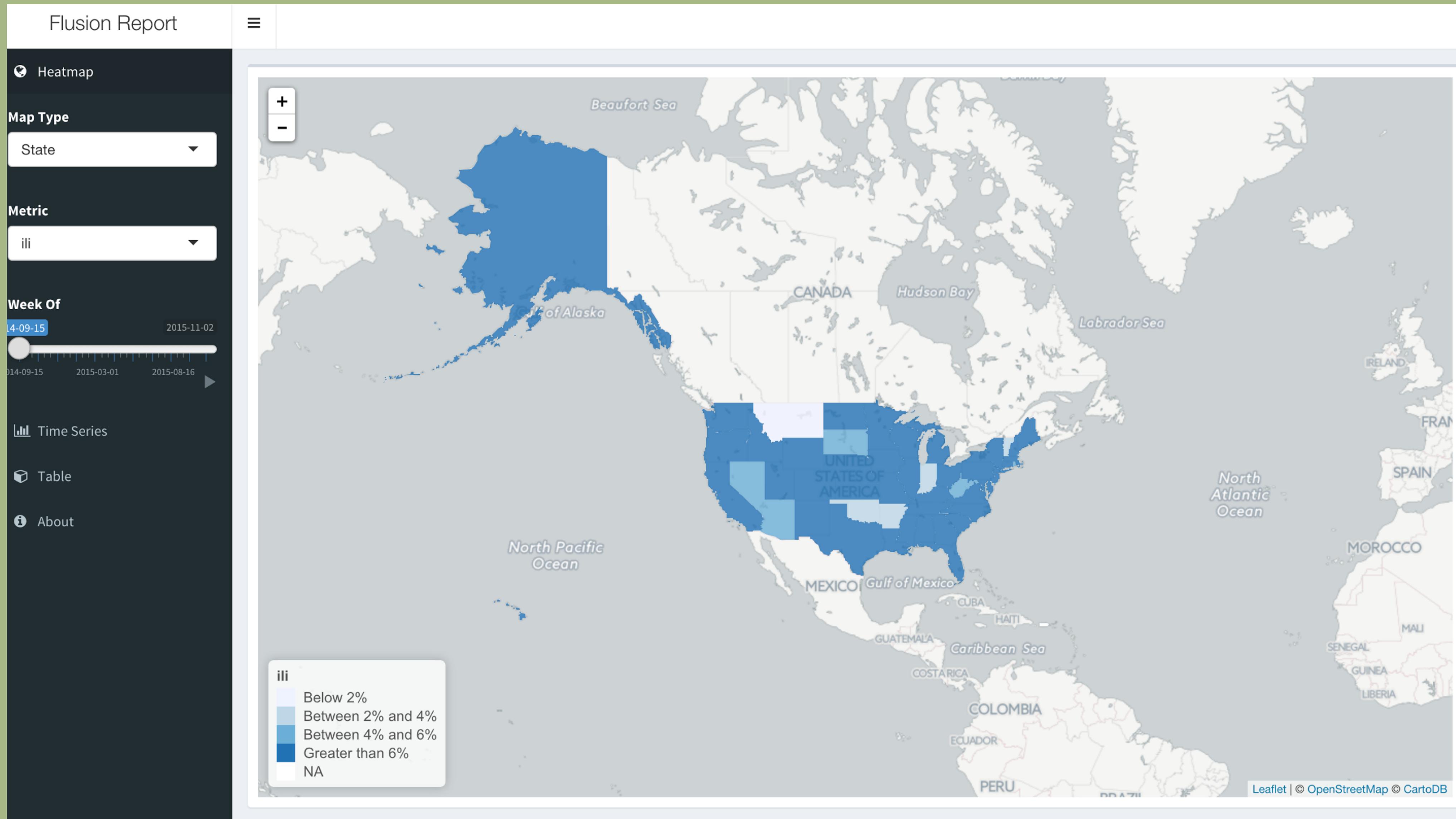
5

---

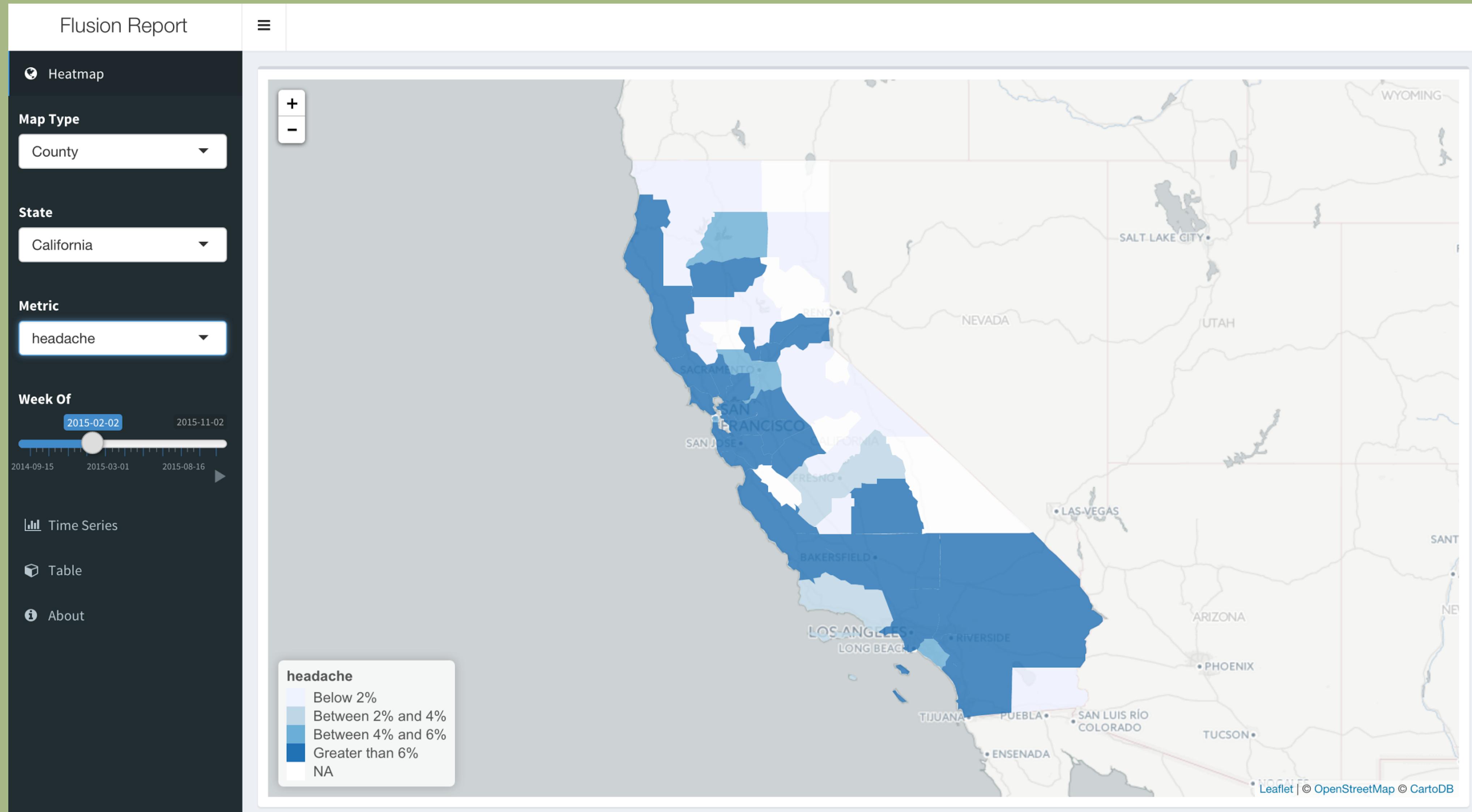
## Deliver

Make sure what you make has value

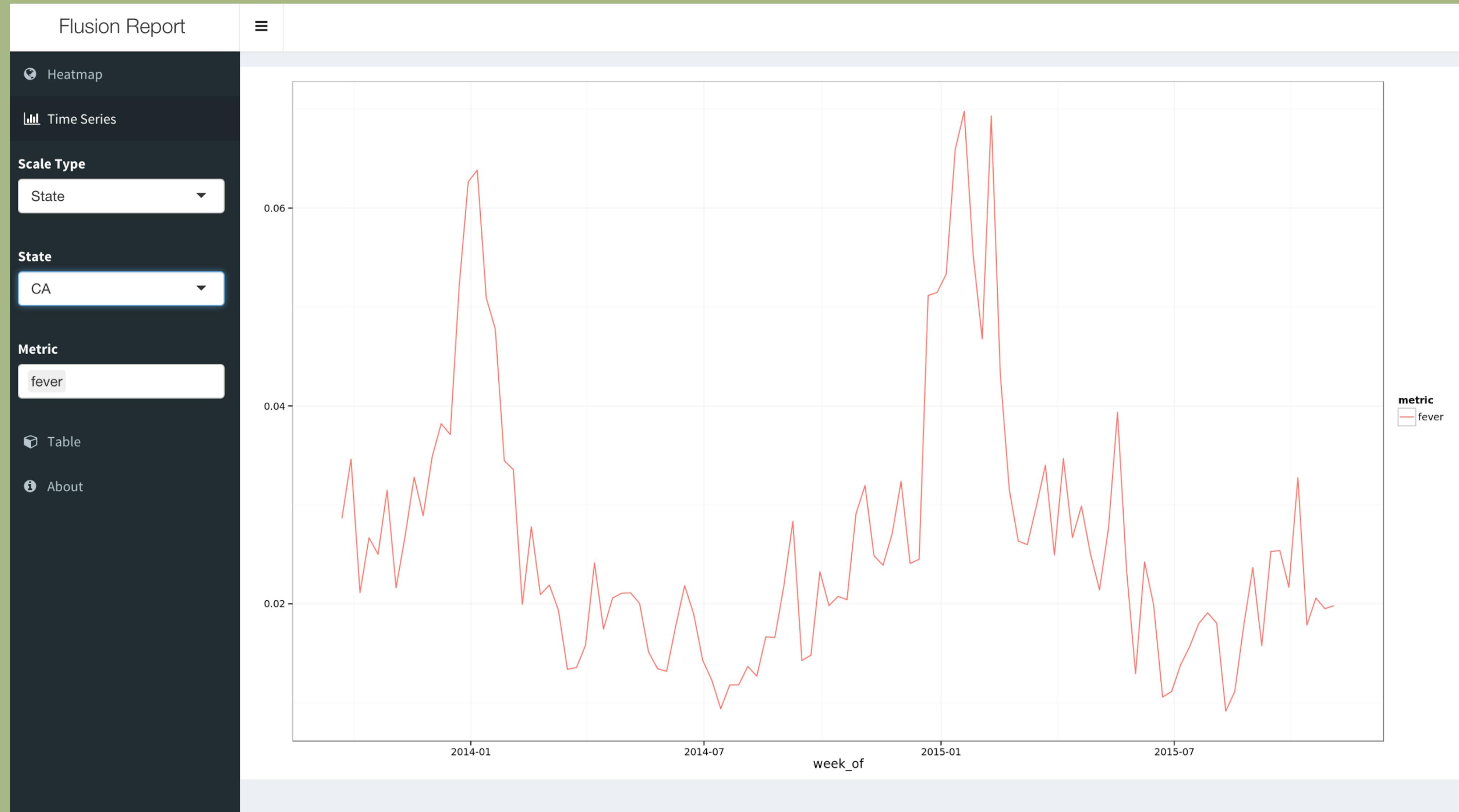
# “Flusion” Report for Health Authorities



# “Flusion” Report for Health Authorities

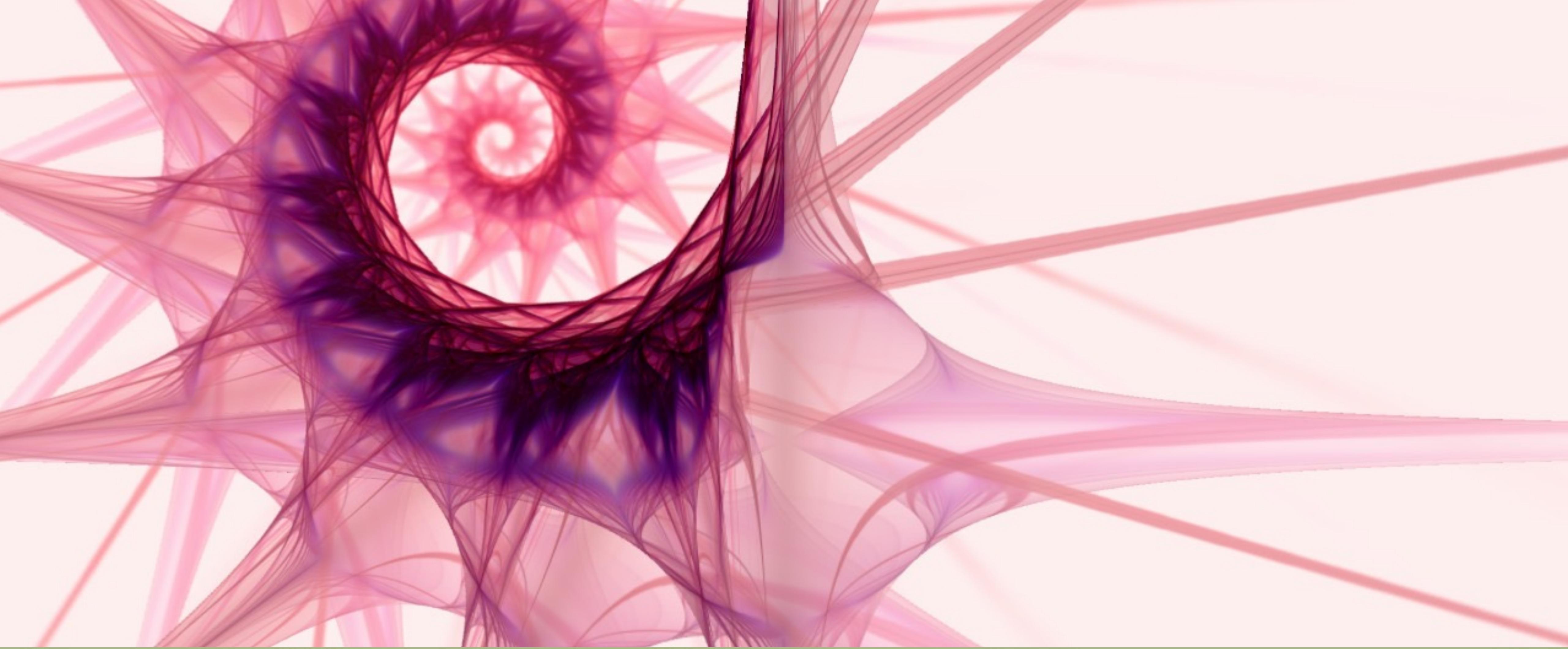


# “Flusion” Report for Health Authorities



# A Data Science Workflow

- ① Ask
- ② Acquire
- ③ Process
- ④ Model
- ⑤ Deliver



**Life is nonlinear**  
**(your process should be, too)**

---

# Thank You!

---