Visualizing cumulative cancer risk from air toxins

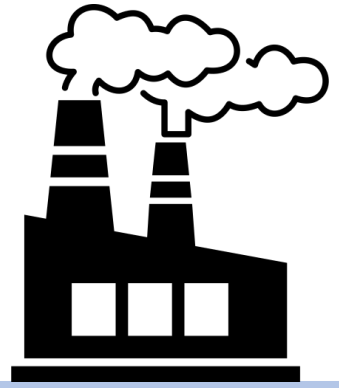
Using R and Leaflet to create an interactive map to explore cancer risk from air toxins in Oregon

Meenakshi Rao
Oregon DEQ
meenakshi.rao@deq.state.or.us

Presentation to NACAA Air Toxics Committee 5th Aug 2021

BACKGROUND

WHAT IS THE NEED FOR THE VISUALIZATION?



Division 245
Cleaner Air Oregon Program

Regulates industrial sources of toxic air contaminants

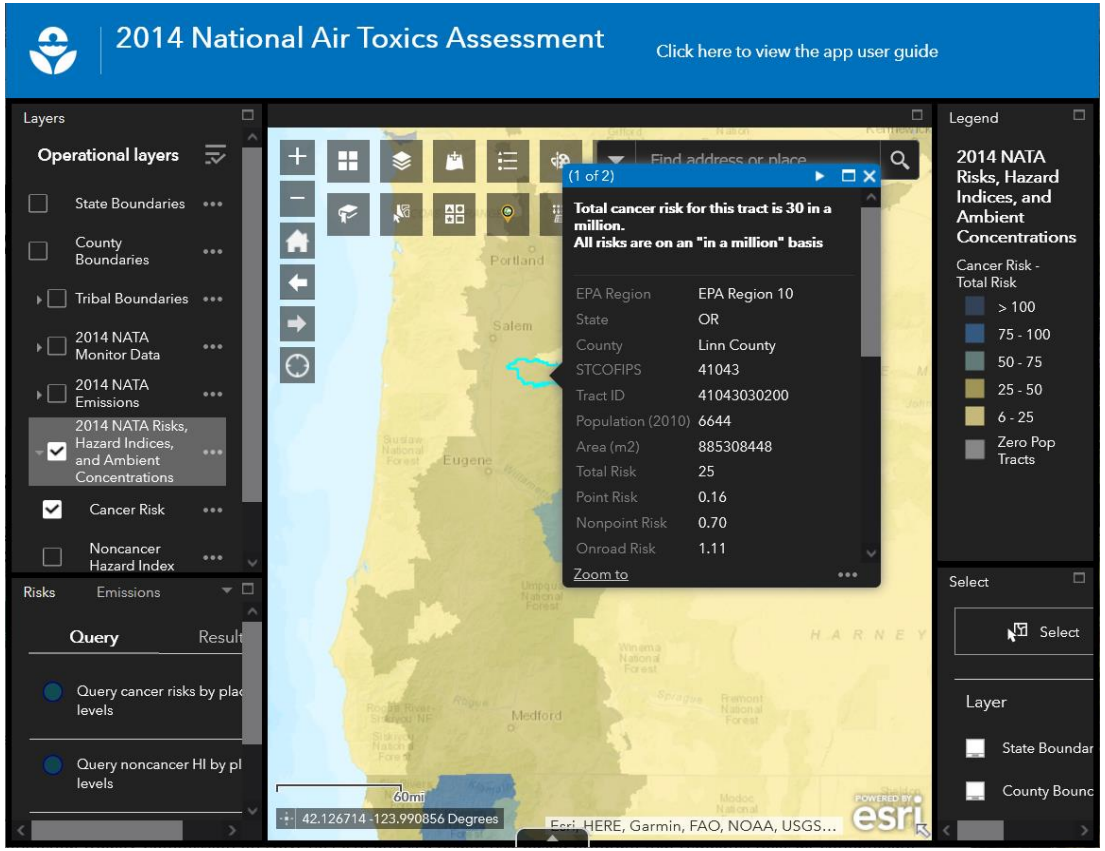
Division 246
Oregon State Air Toxics Program

Assesses impact of toxic air contaminants from all sources

CHALLENGE

MAKING NATA 2014 DATA ACCESSIBLE

- Data/analysis challenges
  - OR-specific risk
    (ambient concentrations, OR TRVs)
- Communication challenges
  - Accessible
  - Local/relevant
  - Interesting



Data Challenge

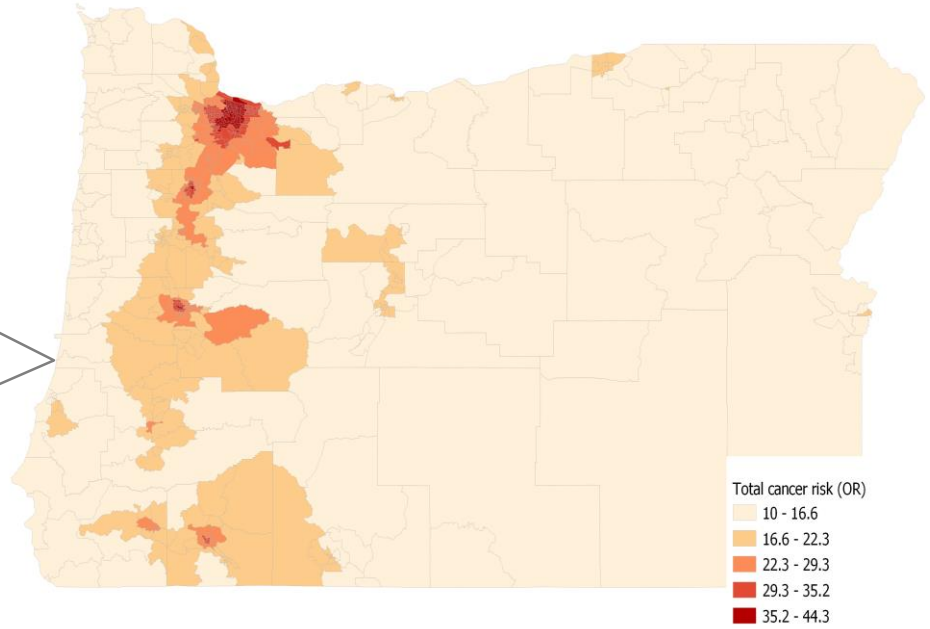
Calculating OR-specific risk

- Ambient concentrations vs. exposure concentrations
- OR specific standards

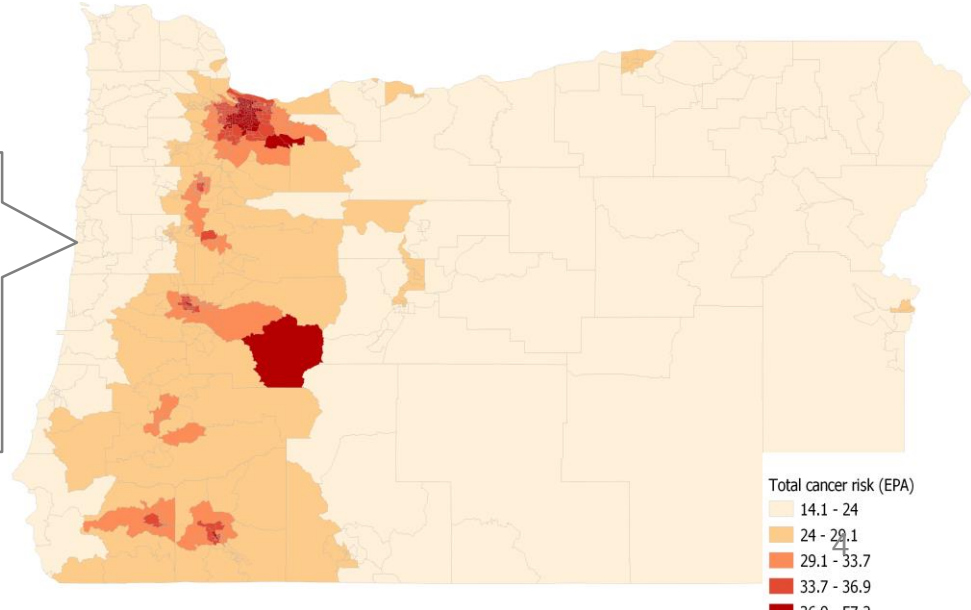
- Fairly straightforward in R:
- iterate through list of air toxins
- divide the ambient concentration by reference values (TRV) for that air toxin
- Bind the rows together

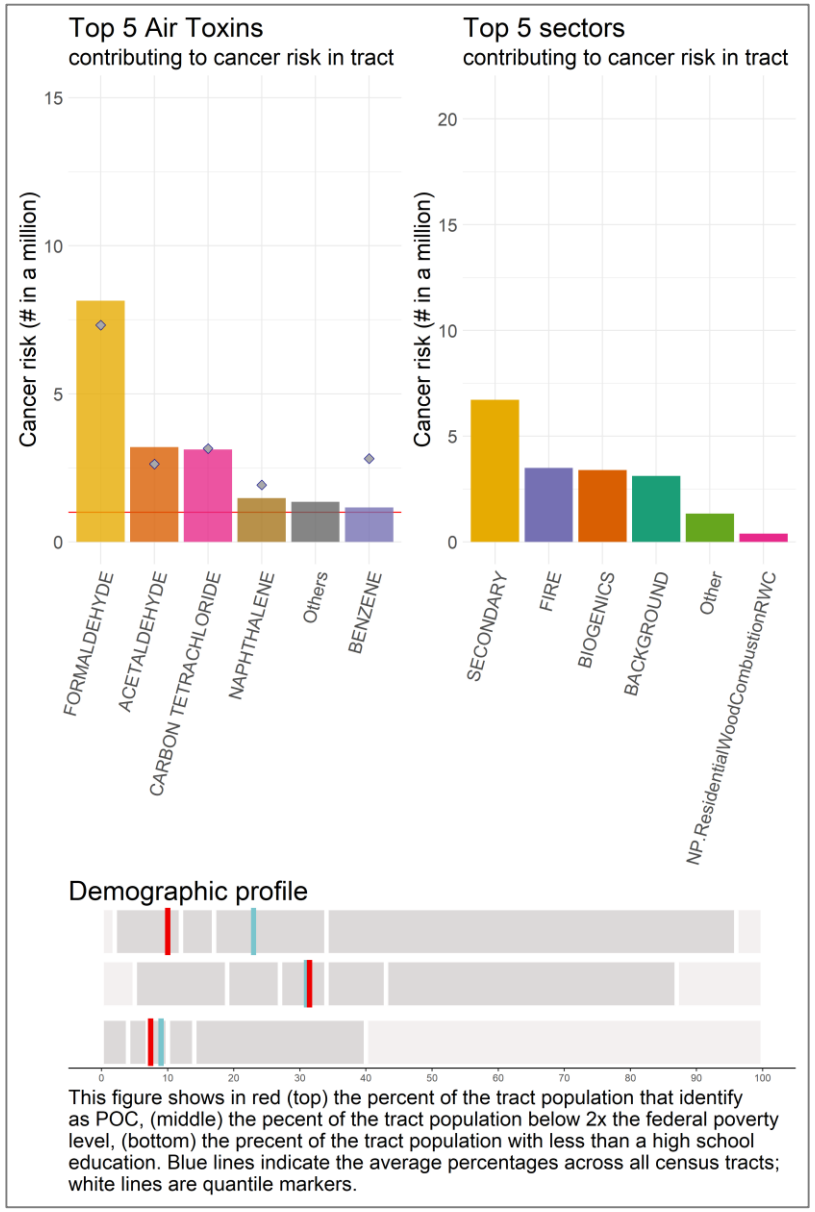
(for R code, see slide at end of deck)

Cancer risk based on ambient conc. & OR TRVs



Cancer risk based on exp. Conc. & EPA reference values





COMMUNICATION CHALLENGE

IDENTIFYING RELEVANT DATA & MAKING IT ACCESSIBLE

Making information accessible

- ❑ focus on top five air toxins
- ❑ focus on top five sectors

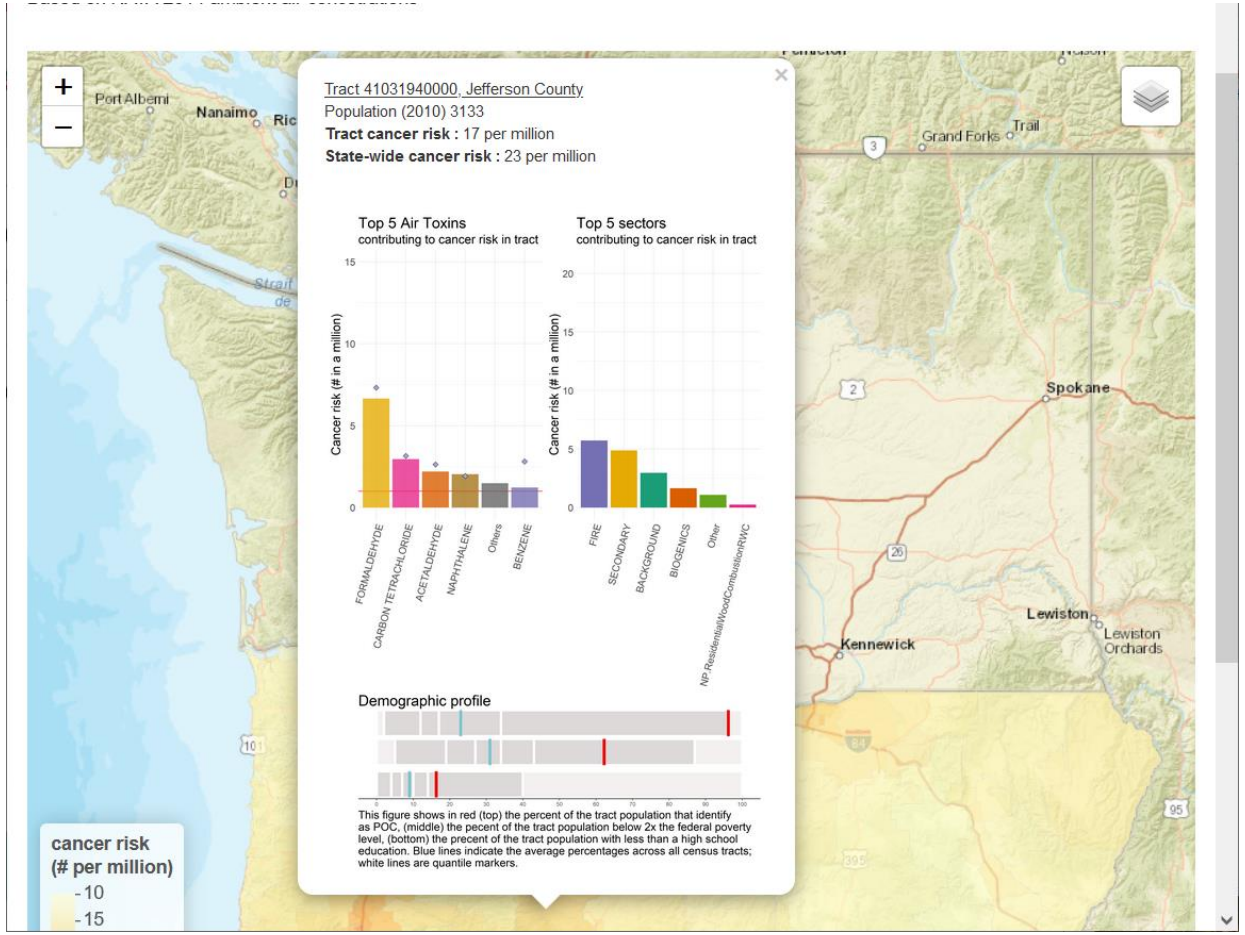
Making information local & relevant

- ❑ create census tract specific information
- ❑ add socio-demographic variables
- ❑ add comparison to state-wide averages

For R code to extract the top 5 air toxins contributing to cancer risk in the tract, see slide at end of deck

Putting it All Together

Interactive Map with Tract-specific Data Visualization

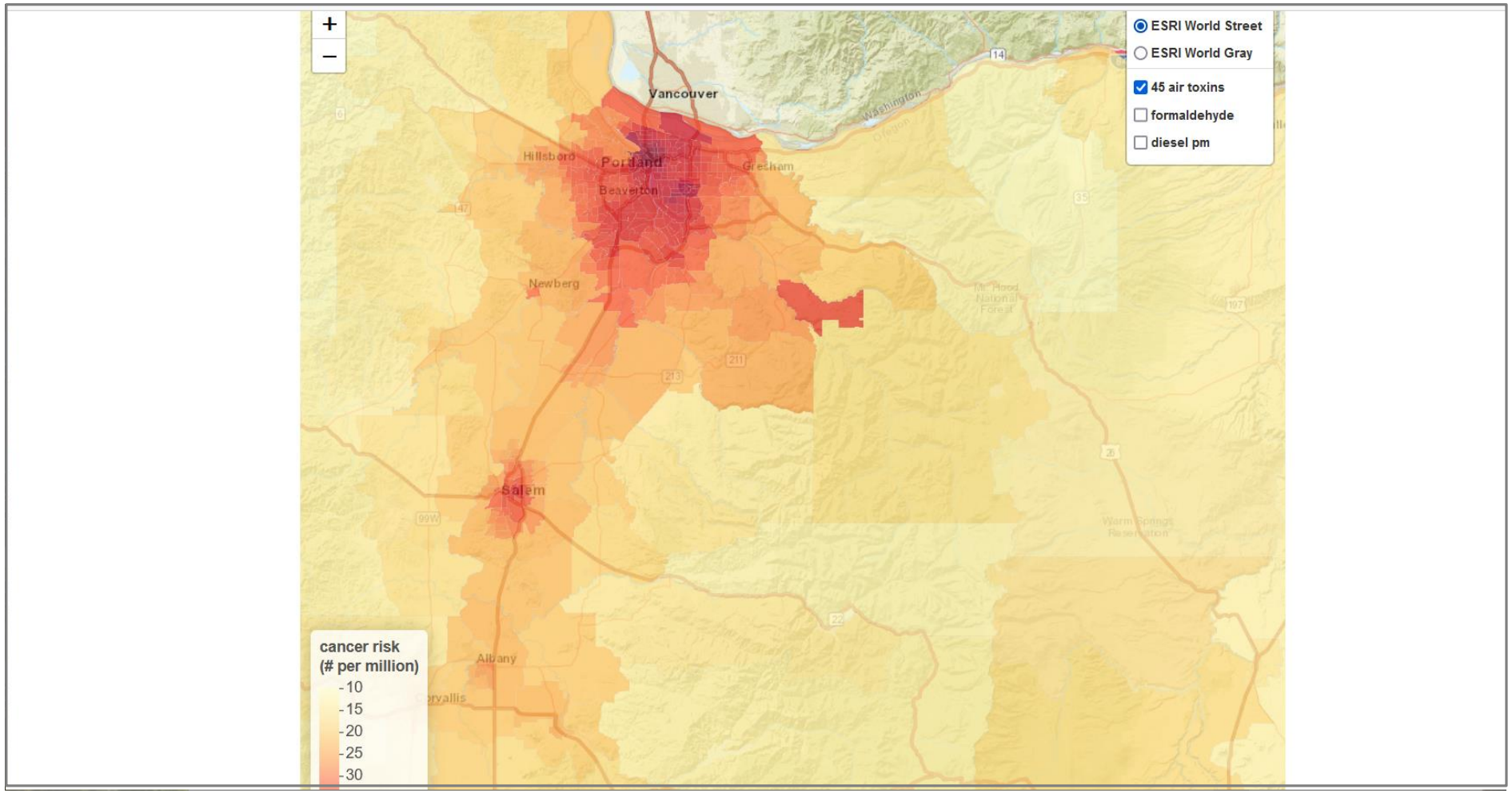


R packages used:

- ❑ rgdal
- ❑ leaflet
- ❑ leafpop
- ❑ tidycensus
- ❑ patchwork

Map created using RMarkdown, and “published” to DEQ’s Connect Server

Demo of Interactive Map

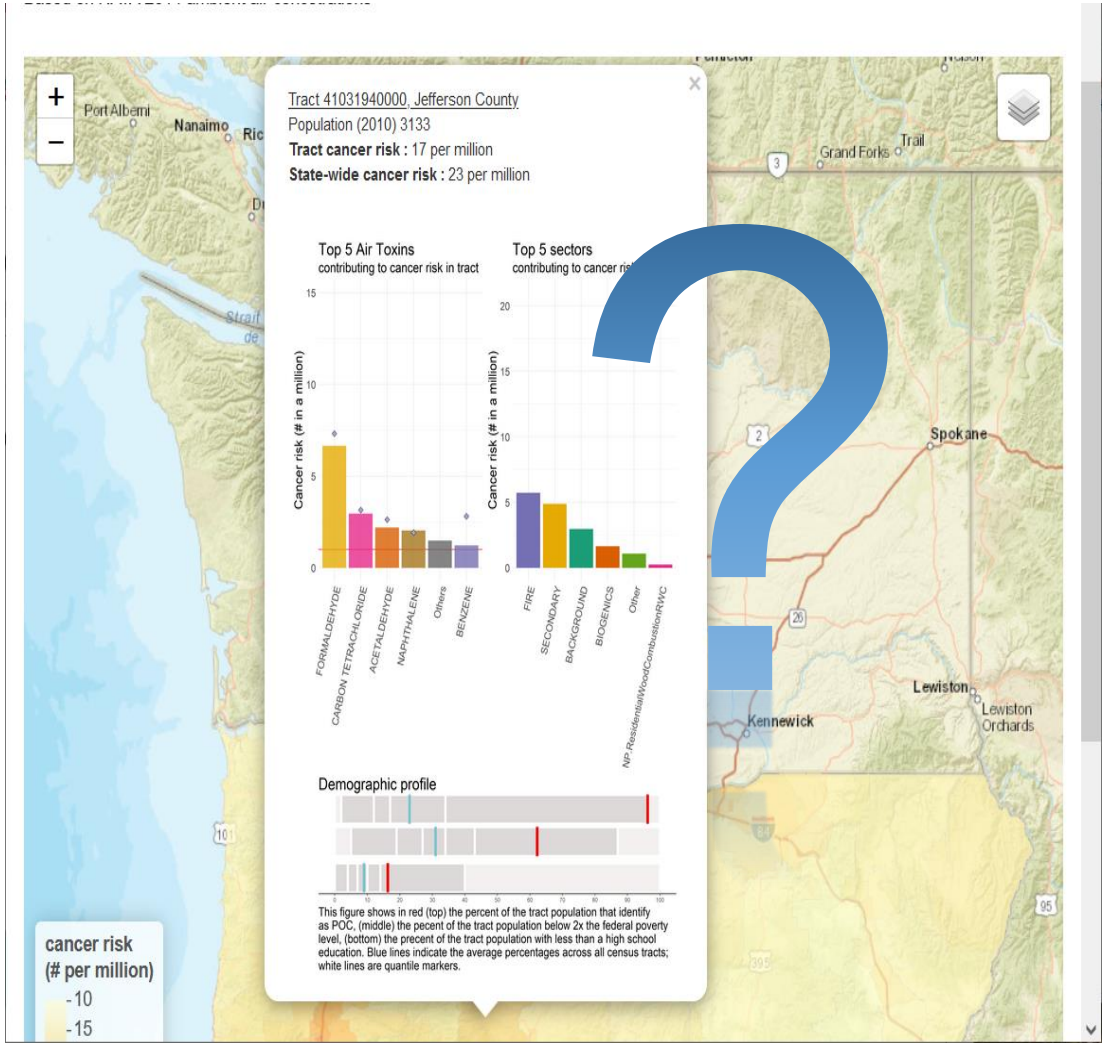


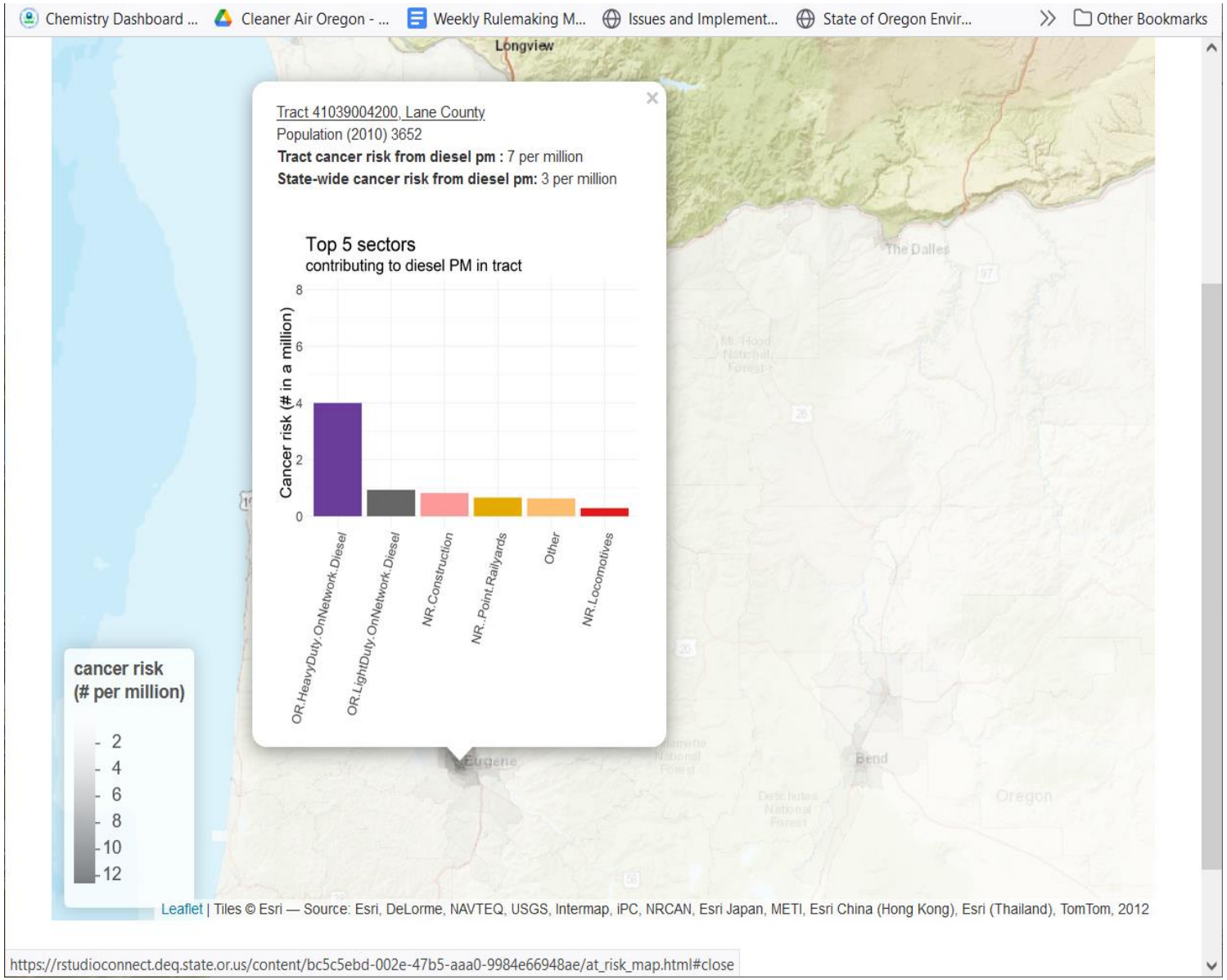
Next Steps

But…Is This What Users Want?

Feedback from community members & advocacy groups

- ❑ accessible?
- ❑ relevant?
- ❑ interesting?
- ❑ useful background for providing input on upcoming rule updates?





DISCUSSION

QUESTIONS…
FEEDBACK…
SUGGESTIONS…

meenakshi.rao@deq.state.or.us

DATA CHALLENGE

CALCULATING OR-SPECIFIC RISK

A “mere” matter of R programming!

```
# calculate cancer risk based on Oregon TRV cancer risk
len <- length(ure_trv$nata.name)                                    ## ure_trv holds OR TRVs
or_ris <- ris[FALSE, 1:46]                                          ## create an empty dataframe with same columns as NATA risk
for ( i in 1:len) {
  if (!is.na(ure_trv$trv_can1_ugm.3[i])) {
    at_exp <- exp %>% filter(Pollutant.Name == ure_trv$nata.name[i]) %>%
                  filter(FIPS*10^6 != Tract)                        ## filter out state and county summary rows from ambient conc dataframe
    at_ris <- ris[, 1:46] %>% filter(Pollutant.Name == ure_trv$nata.name[i]) %>%
                        filter(FIPS*10^6 != Tract)                  ## filter out state and county summary rows from risk dataframe
    at_ris[, 8:46] <- 0  ## set risk from sectors to 0 – ready to be filled in by OR-specific values
    at_ris[, 8:46] <- at_exp[, 8:46]/ure_trv$trv_can1_ugm.3[i]      ## calculate the risk
    or_ris <- rbind(or_ris, at_ris)                                 ## bind rows for each air toxin to the or_ris dataframe
    print(paste(as.character(i), ure_trv$nata.name[i]))             ## print air toxin just completed
   }
}
## add 'pahpom' to the or_ris table (special case)
pah_ris <- ris[, 1:46] %>% filter(Pollutant.Name == "PAHPOM") %>%
  filter(FIPS*10^6 != Tract)
or_ris <- rbind(or_ris, pah_ris)
```

ACCESSIBILITY

IDENTIFYING TOP FIVE TOXINS

```
# extract top five air toxins contributing to cancer risk for each census tract
yat <- at_all %>% group_by(Tract) %>%
      arrange(desc(Total.Cancer.Risk..per.million.)) %>%
      slice(1:5) %>%
      select("Tract", "County", "Population", "Pollutant.Name", "Total.Cancer.Risk..per.million.")
colnames(yat) <- c("Tract", "County", "Population", "Air.Toxin", "Cancer.Risk")


# for each tract, create a "Others" category, summing up cancer risk from remaining air toxins
yato <- at_all %>% group_by(Tract) %>%
                arrange(desc(Total.Cancer.Risk..per.million.)) %>%
                slice(6:n()) %>%
                summarize(County = first(County),
                        Population = first(Population),
                        Air.Toxin = "Others",
                        Cancer.Risk = sum(Total.Cancer.Risk..per.million., na.rm = TRUE))
# combine the two data frames
yat <- bind_rows(yat, yato)
yat <- merge(yat, st_avg, by.x = "Air.Toxin", by.y = "Pollutant.Name", all.x = TRUE)
```