Probabilistic answers to data science questions: using Bayesian hierarchical modeling in analysis

Eric Lind, Ph.D

Strategic Initiatives

Metro Transit
a service of the Metropolitan Council

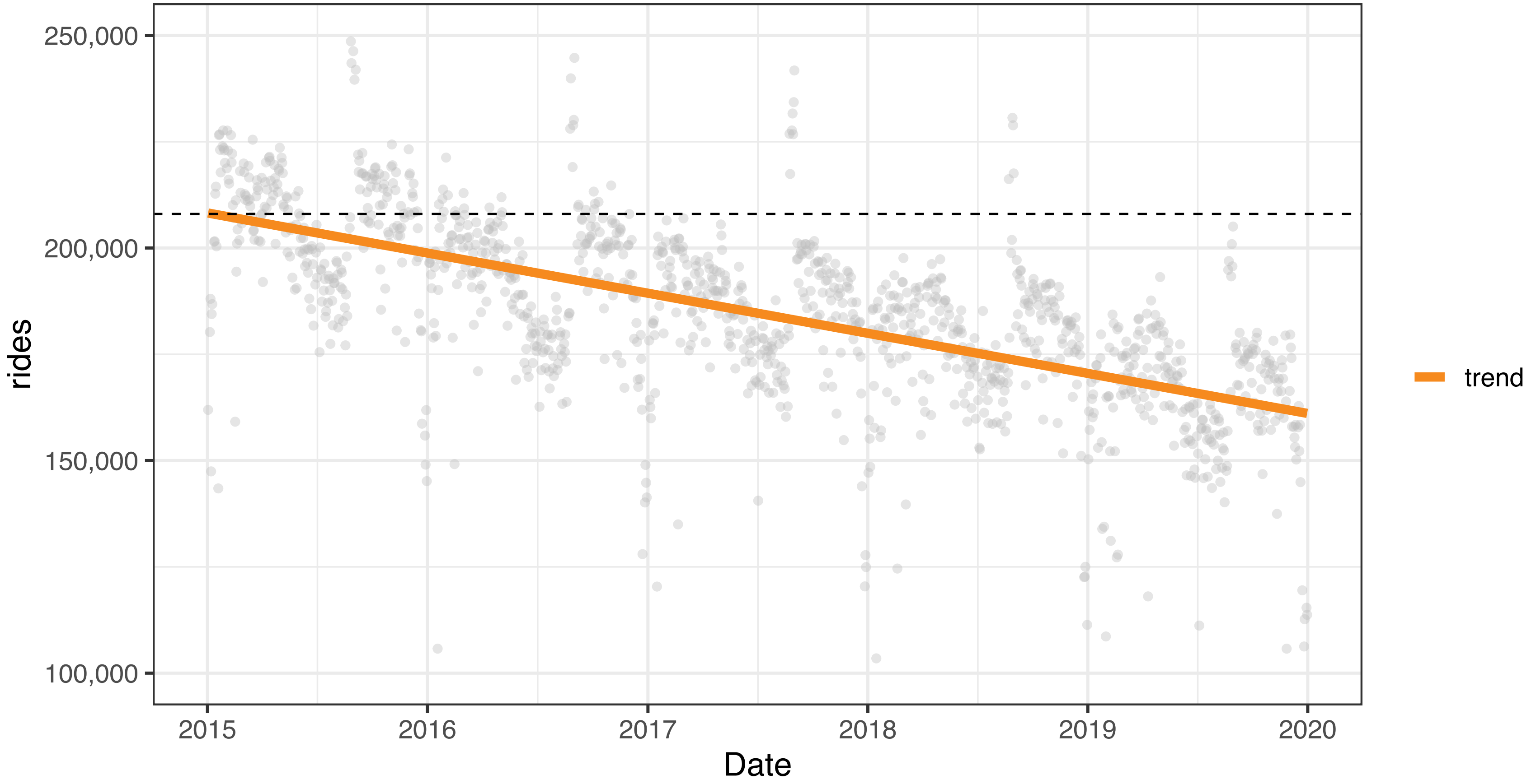
Metro Transit at a Glance

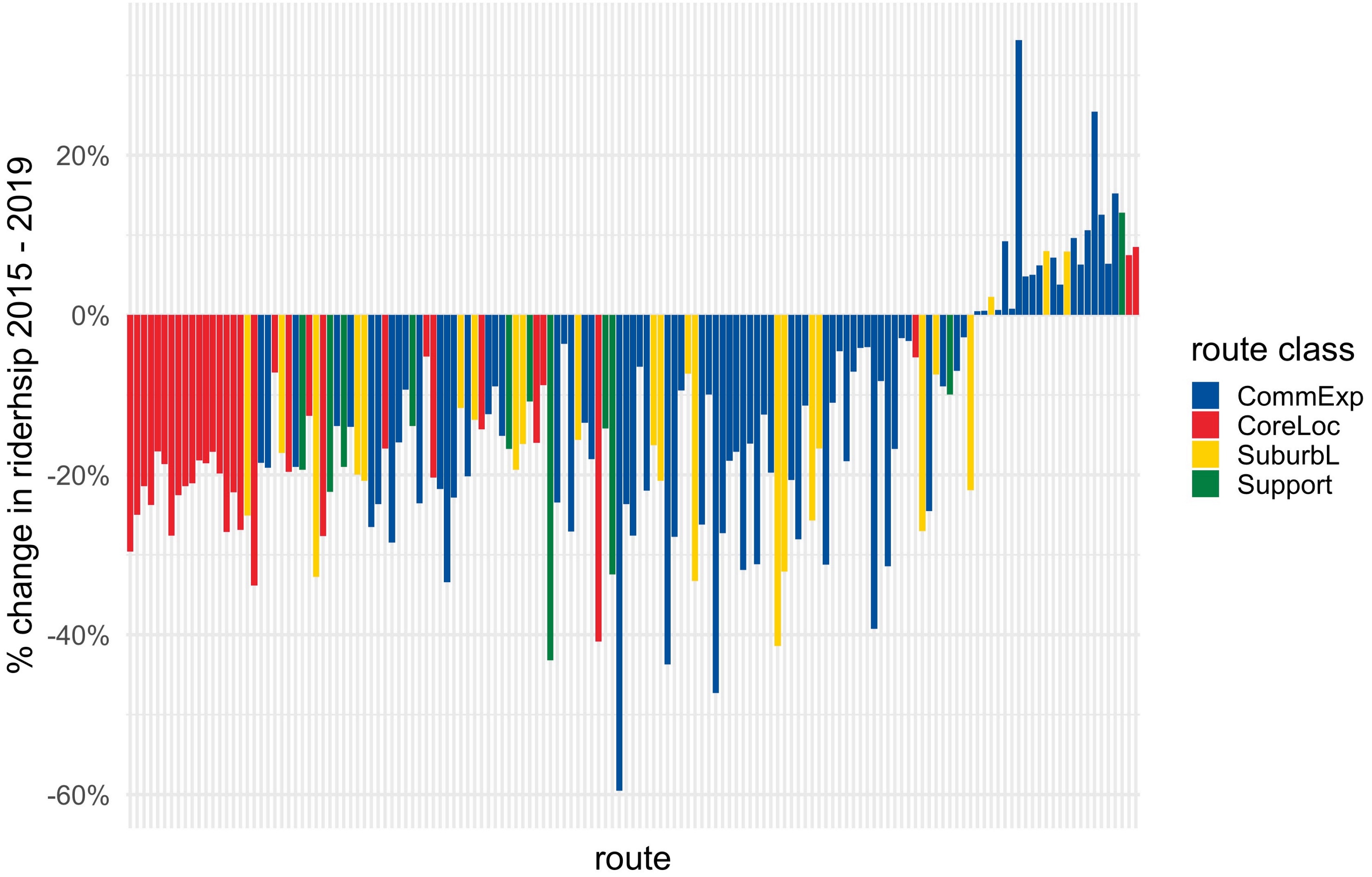
- 15th largest in U.S.
- 7 counties, 90 cities
- 122 bus routes
  - 2 arterial BRT
- 2 light rail lines
- 1 commuter rail line
- 900 buses, 86 LRVs
- 3,200+ employees
- 250,000 daily rides

"We at Metro Transit deliver environmentally sustainable transportation choices that link people, jobs and community conveniently, consistently and safely."

5
BUS STOP
1 HR PARKING
Adopt-A-Shelter
Cystic Fibrosis Pen Pal Club and Support Group
Not everything that is faced can be changed But nothing can be changed until it is faced
Justice Now! Arrest the Cops!
ONLY COWARDS KILL THE UNARMED
STOP KILLIN' BLACK PEOPLE
SAME CRIME
MY LIFE MATTERS
Dear Black
My Crys went un Answered
My Shouts was Ignored
Now every one knows
Stand and Speak for me
Rest shall come when Justice is served. The line
George Floyd

Weekday total Regular Bus rides by day

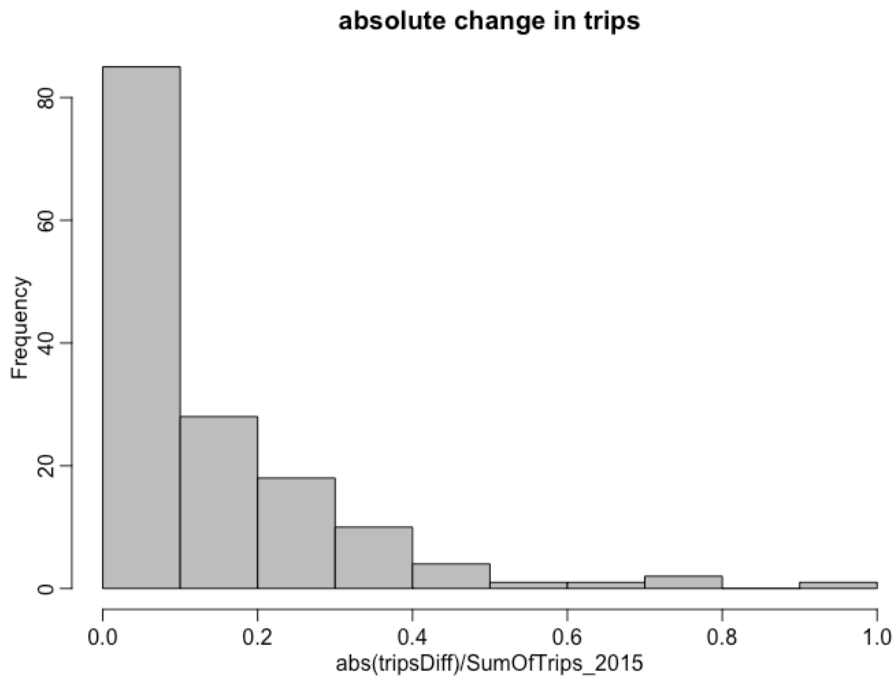




% change in riderhsip 2015 - 2019
20%
0%
-20%
-40%
-60%
route
route class
CommExp
CoreLoc
SuburbL
Support

predicting change

- route class
  - Core Local
  - Suburban Local
  - Supporting Local
  - Commuter & Express
- change in service
  - trips or hours
  - proposed change cutoff 30%
    - 10 routes increased
    - 13 routes decreased
    - 127 routes not changed



goal: characterize ridership change by bus route

- capture:
  - network-wide ridership change
    - most routes share trend

hierarchical grouping

  - differences by route class
    - classes capture type of route, ridership patterns

hierarchical grouping

  - differences by change in amount of service
    - trips, in-service hours

predictor

Rev. Thomas Bayes (1701-1761)

Bayes' Theorem:

The probability of variable ***y*** given ***x*** is determined by the joint probability of ***x*** and ***y***, and the probability of ***x***.

$$\Pr(\text{model}|\text{data}) = \frac{\Pr(\text{data}|\text{model}) * \Pr(\text{model})}{\Pr(\text{data})}$$

$$p_Y(y \mid x) = \frac{p_X(x \mid y)\, p_Y(y)}{p_X(x)}$$

quantifying change

y_i = log(2019 rides / 2015 rides)

translation: the relative rate of change of ridership for route *I*

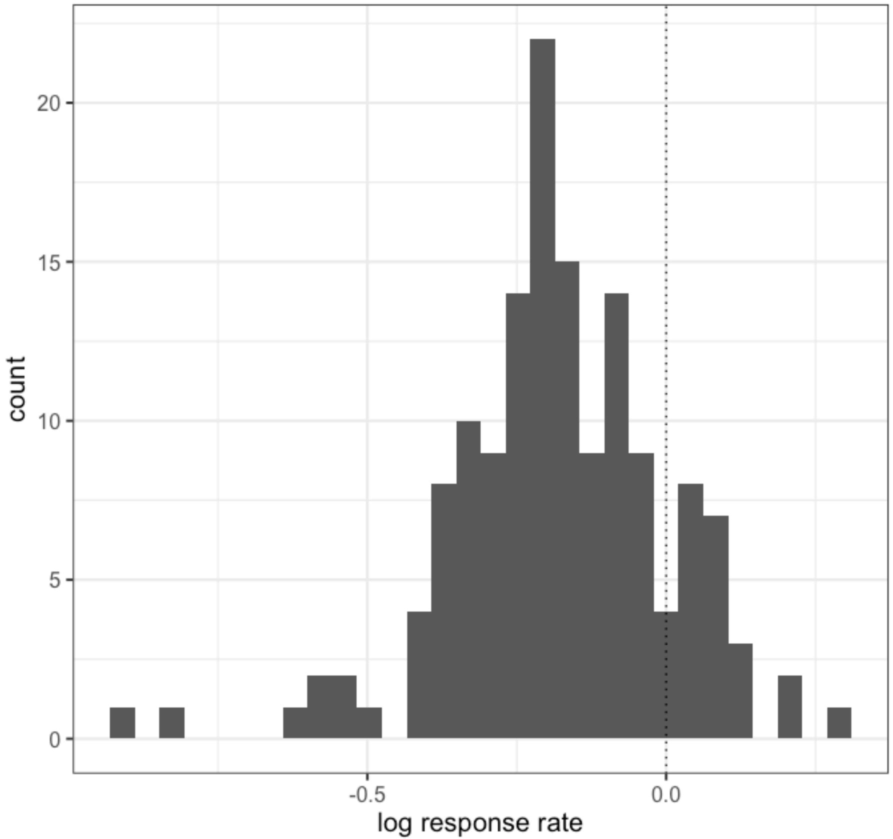
example: Route 21

4,117,347 rides in 2015

3,080,527 rides in 2019

y = -0.290

e^y = 74.8% of 2015 ridership



2015 to 2019 change by route

modeling change

- hierarchical model

$$y_i \sim \mathrm{N}(\mu, \sigma^2)$$

$$\mu = \beta_{0_{class}} + \beta_{1_{class}} * \Delta hrs_i$$

$$\beta_{0_{class}} \sim \mathrm{N}(\overline{\beta_0}, \sigma_{int}^2)$$
$$\beta_{1_{class}} \sim \mathrm{N}(\overline{\beta_1}, \sigma_{hrs}^2)$$

route change is drawn from distribution of probable changes…

…which is determined by the expected route type change, and the change in service for that route…

…the coefficients of which can change by route type

modeling change

- hierarchical model

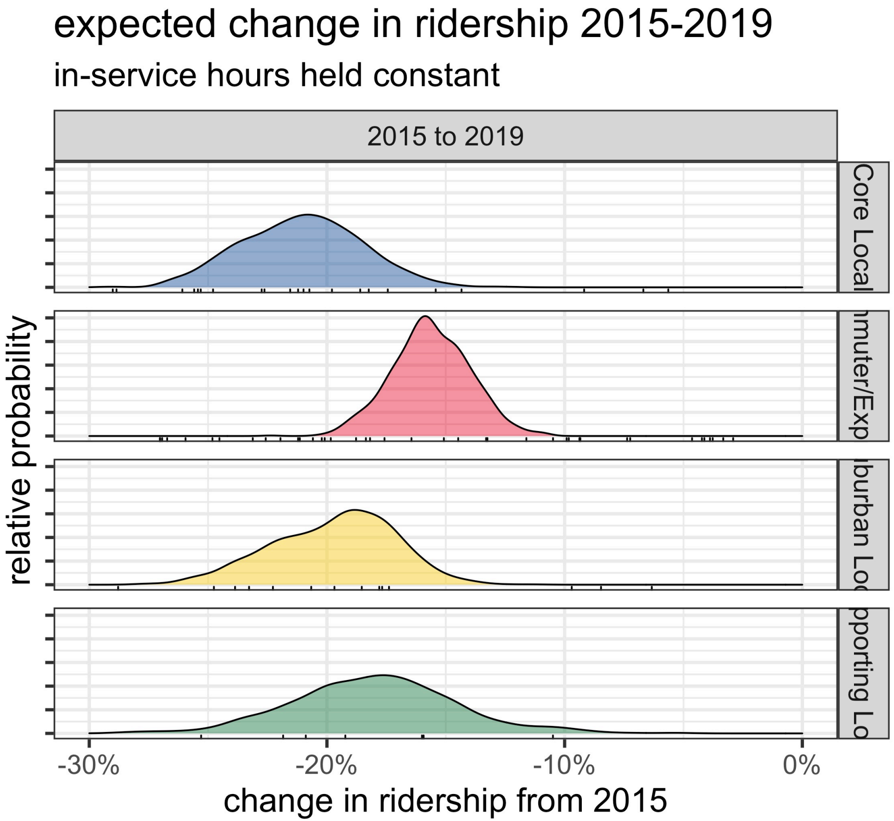
$$y_i \sim \mathrm{N}(\mu, \sigma^2)$$

$$\mu = \beta_{0_{class}} + \beta_{1_{class}} * \Delta hrs_i$$

$$\beta_{0_{class}} \sim \mathrm{N}(\overline{\beta_0}, \sigma_{int}^2)$$
$$\beta_{1_{class}} \sim \mathrm{N}(\overline{\beta_1}, \sigma_{hrs}^2)$$

```
m4 <- map2stan(
  alist(
    y ~ dnorm(muhat, sigma),
    muhat <- a_C[cls] + b_H[cls]*hrs_chg,
    a_C[cls] ~ dnorm(A, sigma_class),
    b_H[cls] ~ dnorm(B_H, sigma_hours),
    c(A, B_H) ~ dnorm(0, 2),
    c(sigma, sigma_class, sigma_hours) ~ dcauchy(0, 1)
  ),
  data = tomod3, chains = 4, cores = 4, debug = T,
  control = list(max_treedepth = 15, adapt_delta = 0.98)
)
```

results



- Overall decline: 18% [12 - 24]

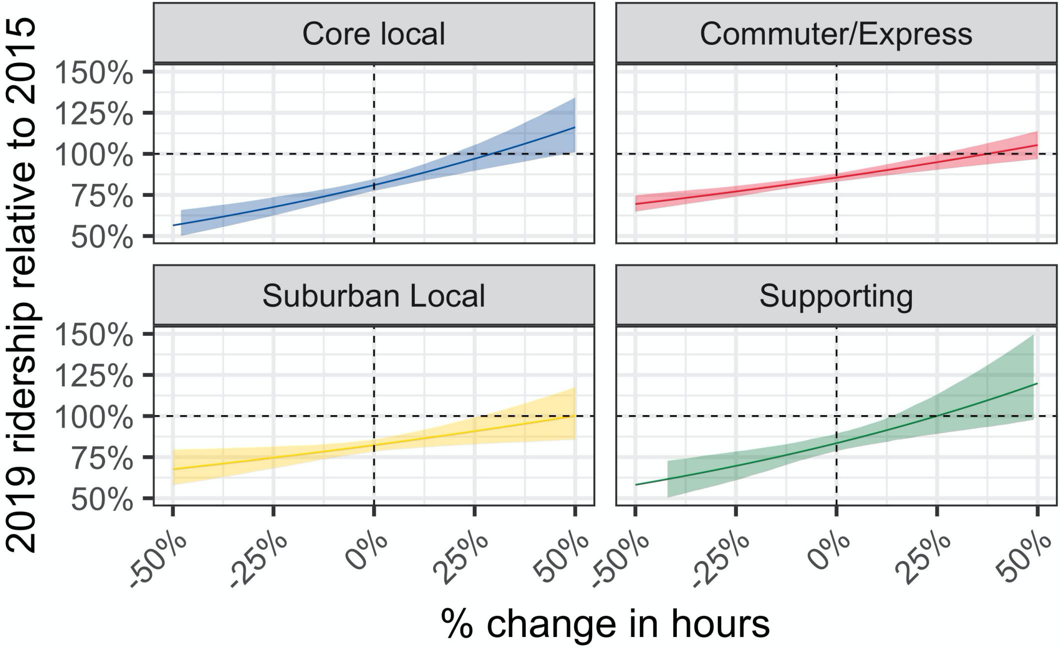
Core local	-23%
Commuter	-17%
Suburban	-22%
Supporting	-19%

results

- hours change varies by class
  - core local steeper
  - commuter flatter
- *vertical dashed line:* expected ridership @ no service change
- *horizontal dashed line:* service change needed to maintain 2015 ridership level

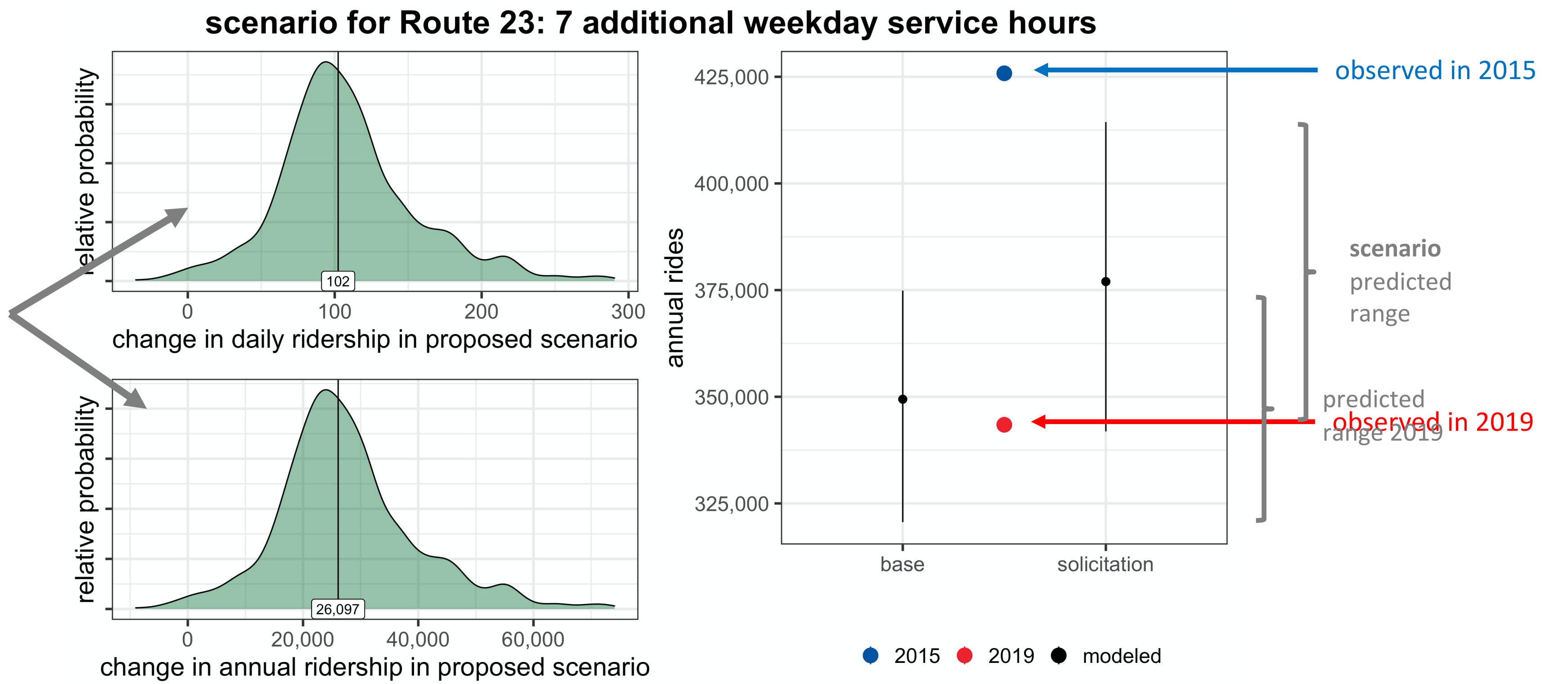
expected change in ridership 2015-2019

response to change in in-service hours



use in scenario forecasting

change in annual from base



probabilistic solutions in data science

- reflect *uncertainty* and range of knowledge
- Bayesian: true nested structure
- straightforward interpretation / extrapolation
- domain knowledge + stats knowledge