

THE BENEFITS OF BEING BAYESIAN

HOW MULTI-LEVEL MODELS ENABLE MORE GRANULAR INSIGHTS

June 6, 2018



Elicit

CURT BERGMANN

SENIOR DATA SCIENTIST

curt.bergmann@elicitinsights.com





WE LITERALLY
WROTE THE BOOK
ON THE SUBJECT



www.geeknerdsuit.com

"With today's business being driven by data, the Geek Nerd Suit philosophy is a must to be successful in the shifting customer landscape."

CARRIE THARP - CMO, NEIMAN MARCUS

GEEK NERD SUIT

BREAKING DOWN WALLS, UNIFYING TEAMS, AND
CREATING CUTTING-EDGE CUSTOMER CENTRICITY

CHUCK
DENINGER

BROOKE
NIEMIEC

MASON
THELEN

THE TAKEAWAY FOR TODAY



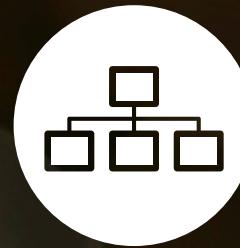
Regression
might perform
poorly



Bayesian
Methods
aren't magic



Prior
information
can improve
regression



Hierarchy can
be a part of
that prior

HOW MULTI-LEVEL MODELS ENABLE MORE GRANULAR INSIGHTS

WHAT DOES THIS MEAN?

PAY PER CLICK

UNDERSTANDING CONVERSION RATES



BOWER'S BIKES

PAID SEARCH

Search:

BOWER'S BIKES

PAID SEARCH

Search: Trek

BOWER'S BIKES

PAID SEARCH

Search: Trek

Joe's Bike Shop: \$2.00

Bower's Bikes: \$2.50

Pro Bike Shop: \$2.25

BOWER'S BIKES

PAID SEARCH

Search: Trek

Joe's Bike Shop: \$2.00

Bower's Bikes: \$2.50

Pro Bike Shop: \$2.25

BOWER'S BIKES

PAID SEARCH

Search: Trek

Trek Bicycles at Bower's Bikes

[ad] <https://www.bowersbikes.com>

Get \$30 off your next Trek bicycle

Trek Bikes - The world's best bikes and cycling gear | Trek Bikes

https://www.trekbikes.com/us/en_US/

Driven by adventure, guided by our history, ...



HOW MUCH TO BID?

MAXIMUM BID

Part of the optimization strategy is to determine your maximum bid

Max Bid = Average Conversion Rate × Margin

CONVERSIONS HAVE TWO POSSIBILITIES

NO CONVERSION

1. An ad appears on a search results page
2. A customer clicks on the ad
3. We count **one click**
4. The customer leaves the page
5. We count **zero conversions**

CONVERSION

1. An ad appears on a search results page
2. A customer clicks on the ad
3. We count **one click**
4. Customer prints coupon
5. Customer buys bike with coupon
6. We count **one conversion**



HOW MUCH TO BID?

MAXIMUM BID

Part of the optimization strategy is to determine your maximum bid

Max Bid = Average Conversion Rate × Margin

$$\$3.00 = 3.00\% \times \$100.00$$

HOW MUCH TO BID?

WHY THE MAXIMUM BID?

Revenue = **3.00%** x 100 clicks x \$100 = \$300

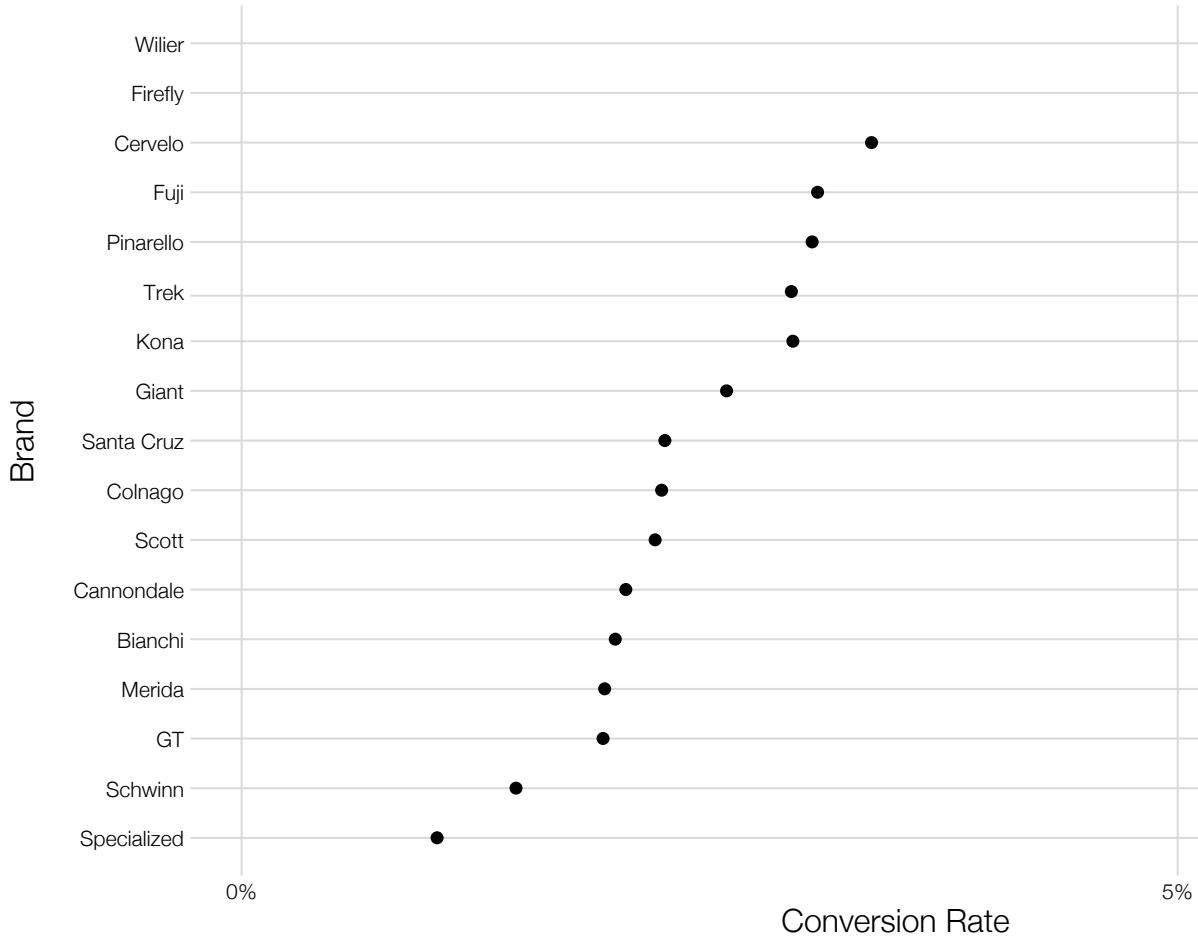
Cost = **\$-4.00** x 100 clicks x \$100 = \$-400

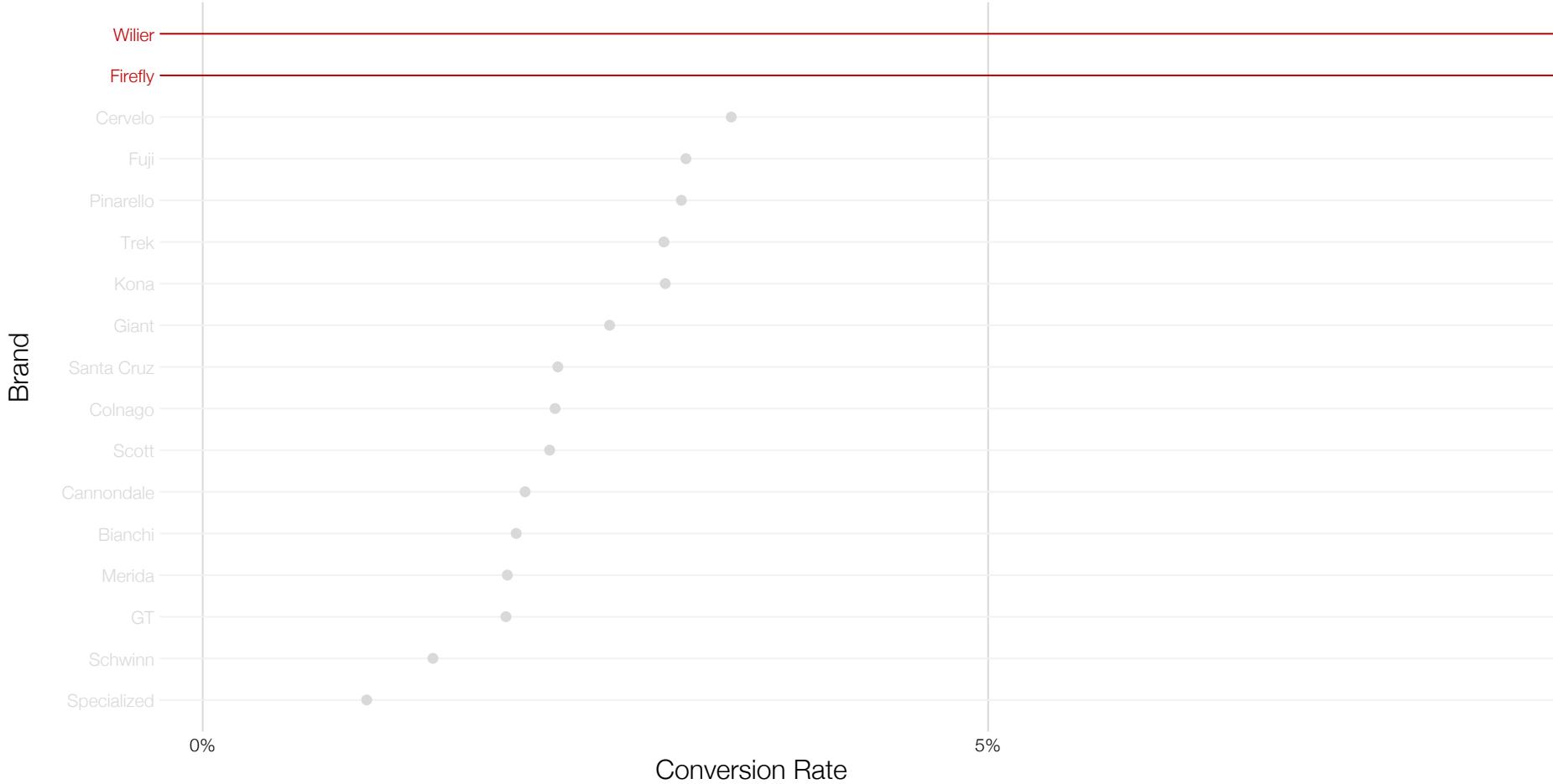
\$300 - \$400 = **\$-100**

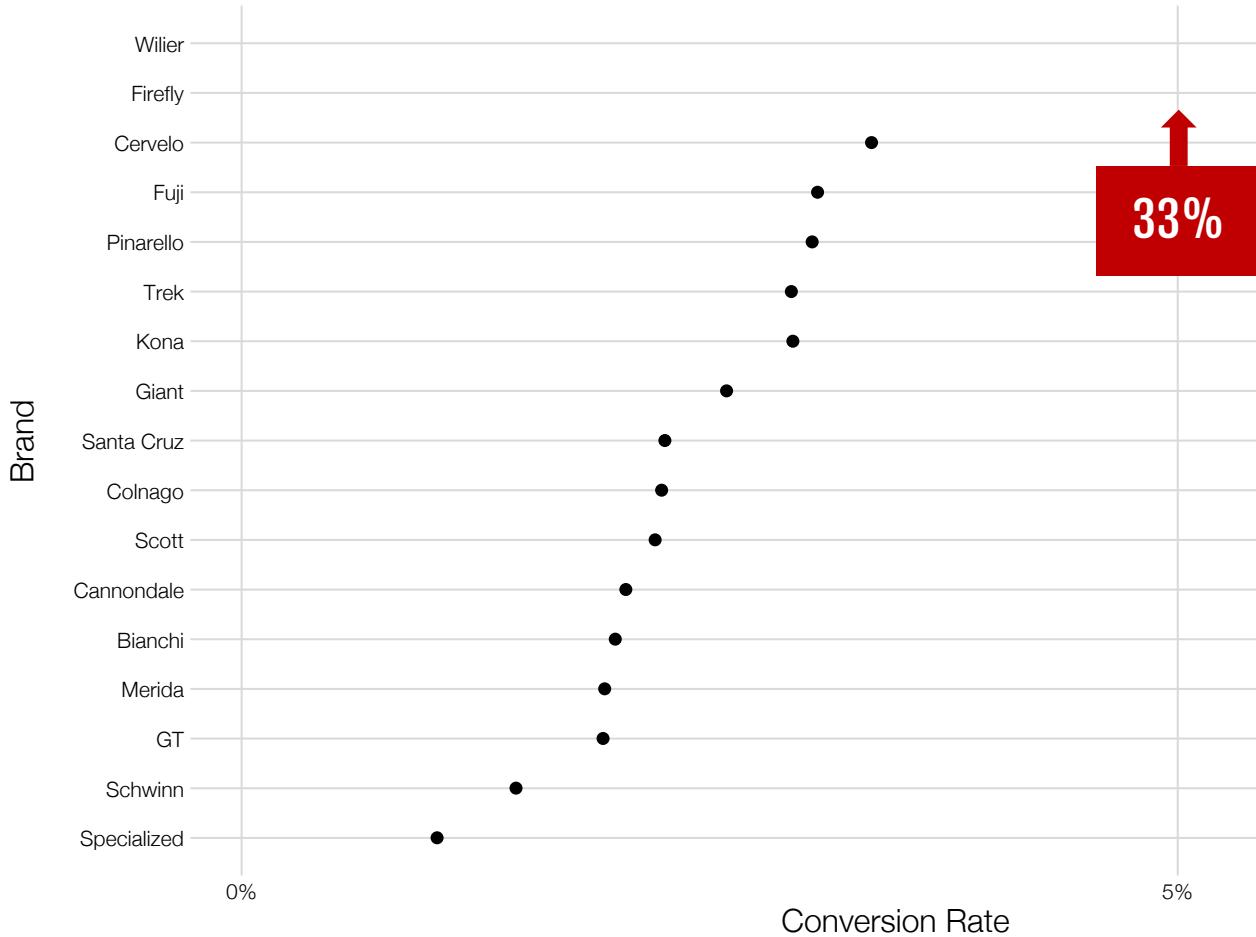
PAID SEARCH

WHAT WE WOULD LIKE TO KNOW

What is the average conversion rate?





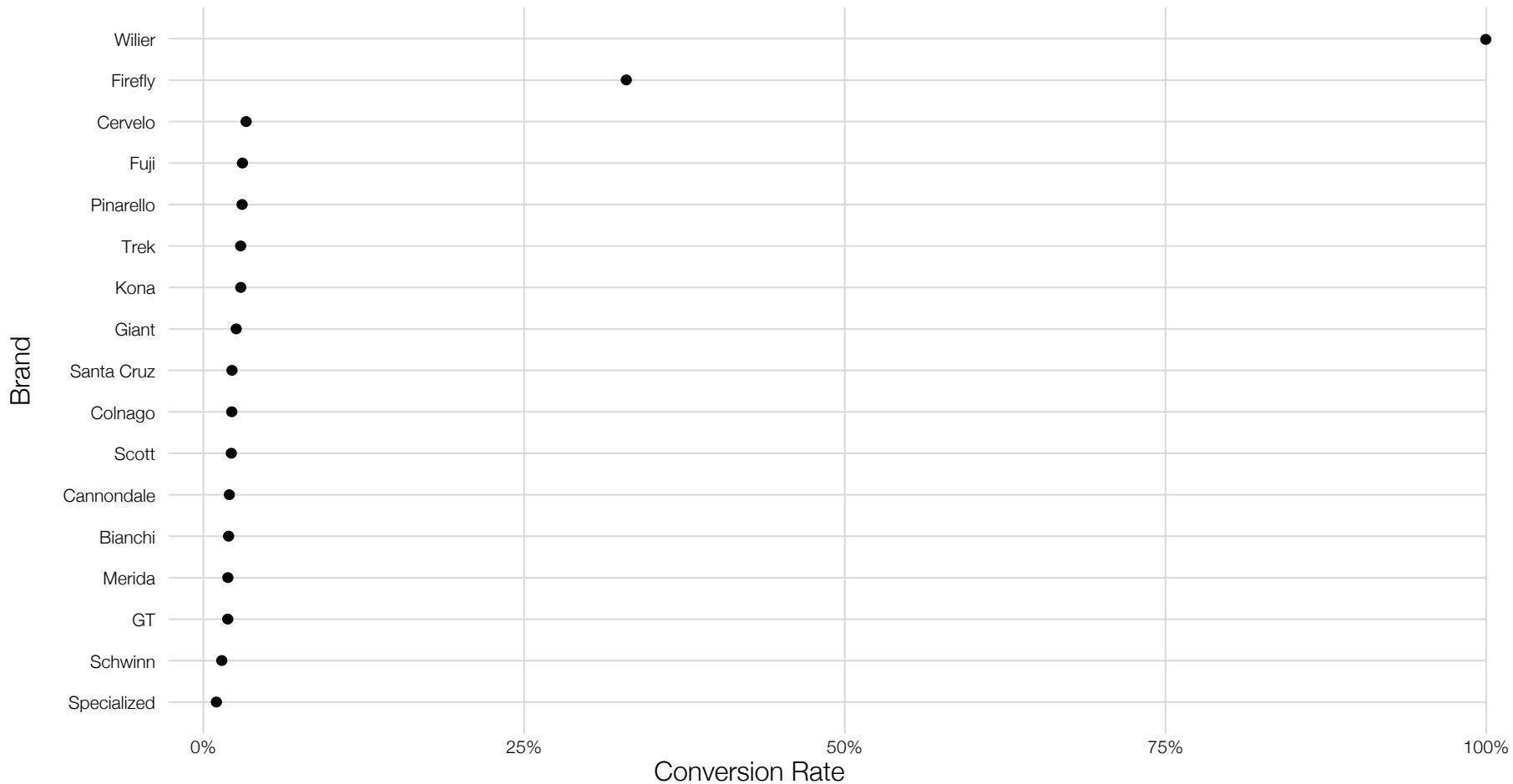




100%

30%

35%



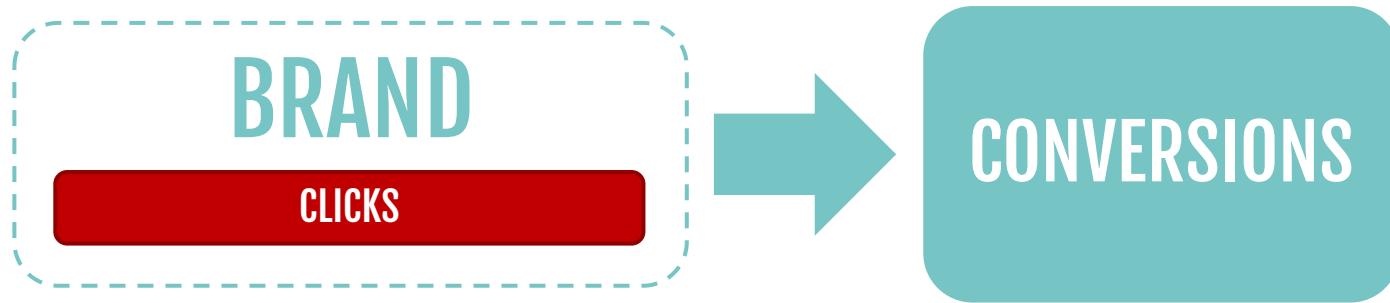
BICYCLE SHOP

PAID SEARCH

BRAND	CLICKS	CONVERSIONS	CONVERSION RATE
Wilier	3	3	100.00%
Firefly	3	1	33.33%
Cervelo	208	7	3.37%
Fuji	195	6	3.08%
Pinarello	197	6	3.05%
Trek	204	6	2.94%



INTUITIVE MODEL



REGRESSION ALGEBRA

$$\text{CONVERSIONS} = \text{BRAND RATE} \times \text{CLICKS}$$

BINOMIAL DISTRIBUTION

$$\left(\text{CONVERSIONS}, \text{CLICKS} - \text{CONVERSIONS} \right) \sim \text{BRAND RATE}$$

BINOMIAL REGRESSION IN R

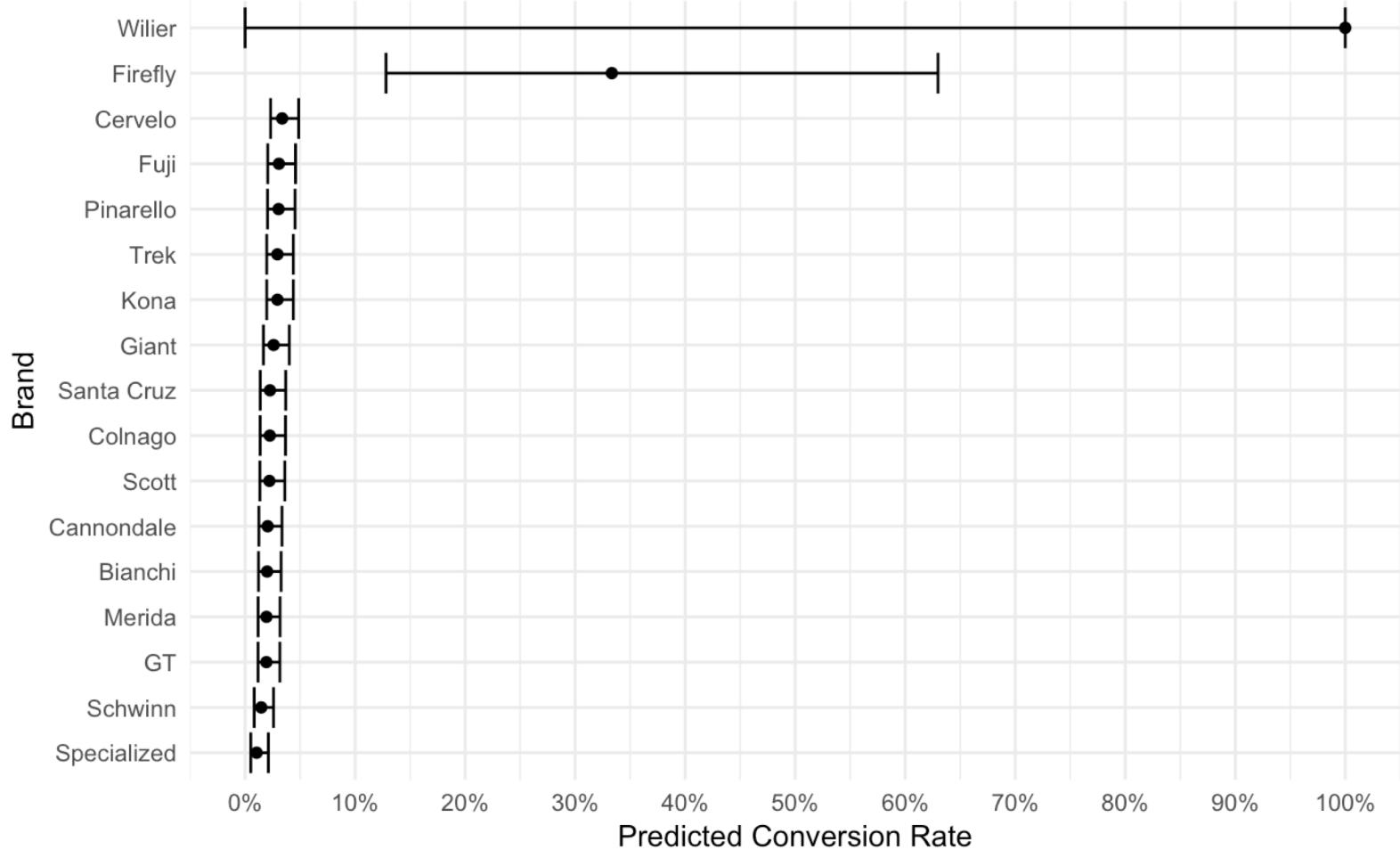
```
glm(  
  cbind( CONVERSIONS , CLICKS - CONVERSIONS ) ~ BRAND RATE ,  
  family = binomial(link = "logit")  
)
```

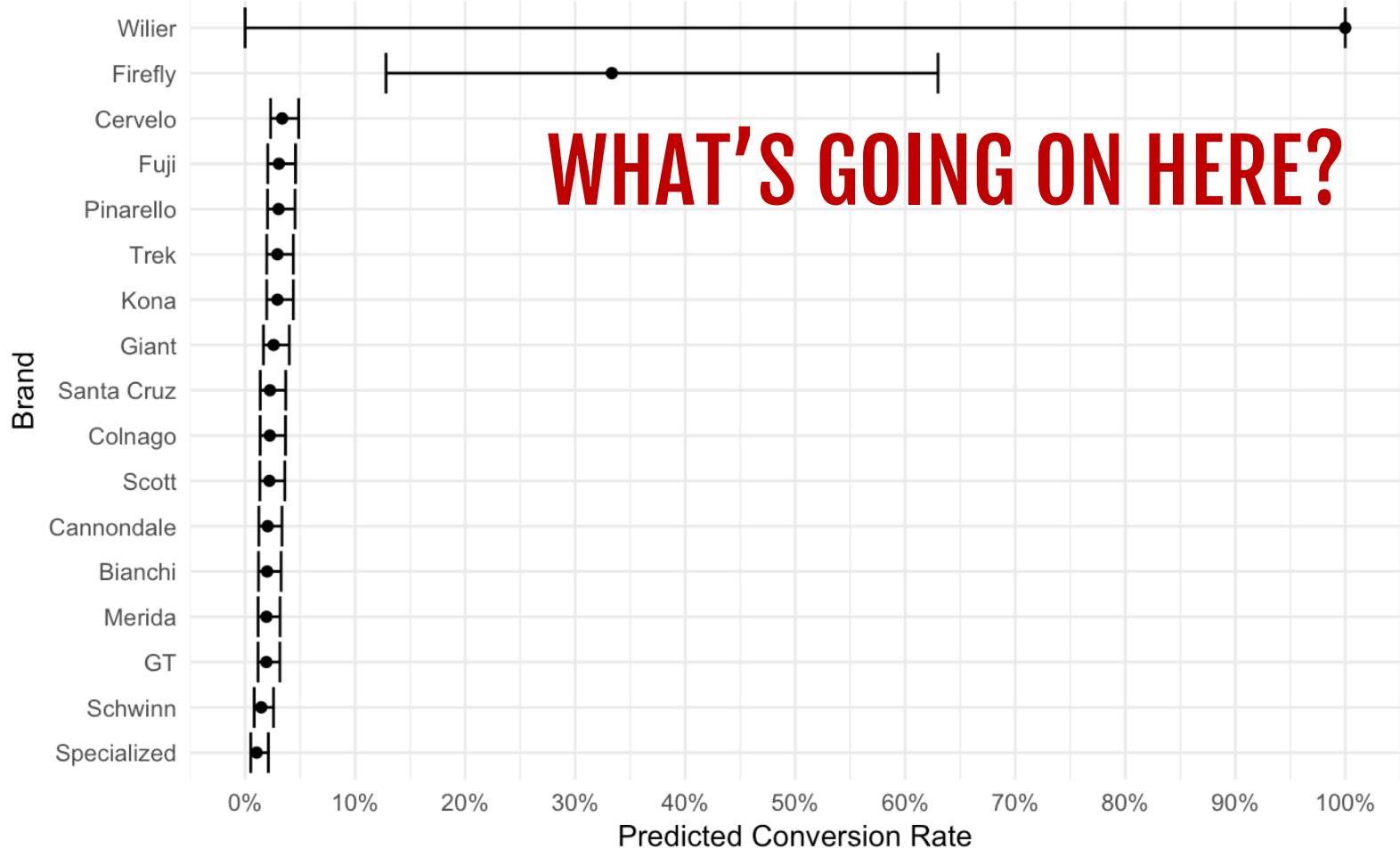


BINOMIAL REGRESSION IN R

```
glm(  
  cbind(conversions, clicks - conversions) ~ brand,  
  family = binomial(link = "logit")  
)
```

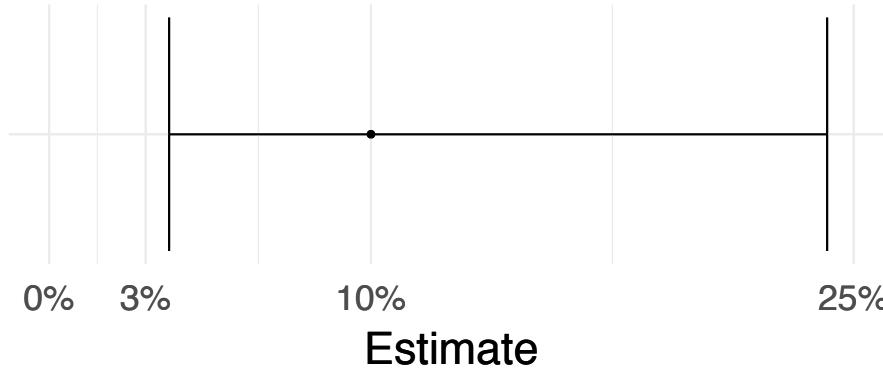






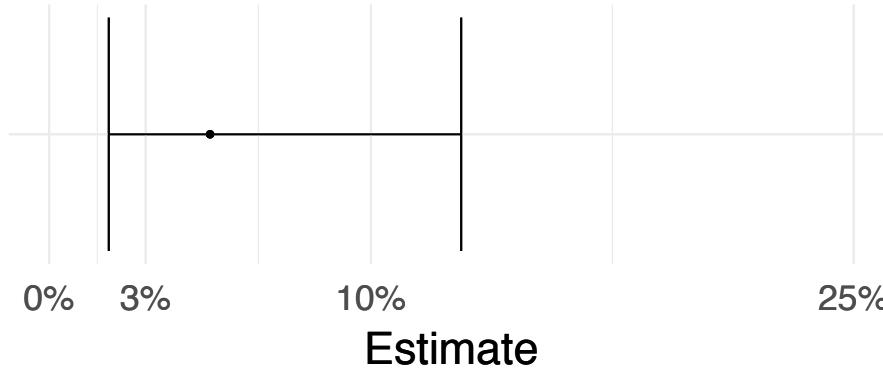
PAID SEARCH

1 Conversion 10 Clicks



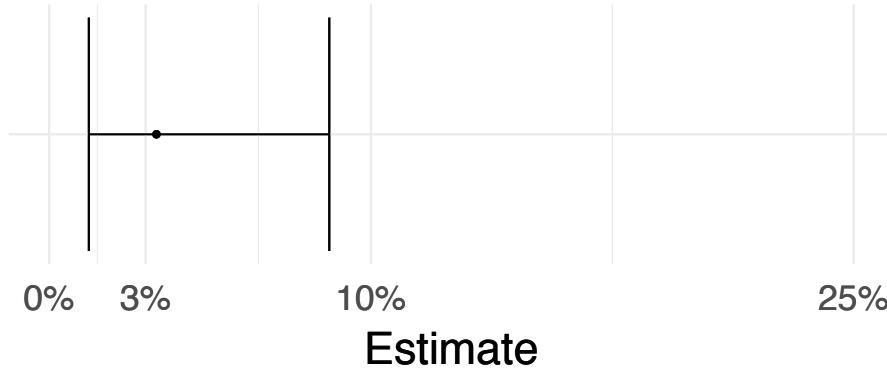
PAID SEARCH

1 Conversion 20 Clicks



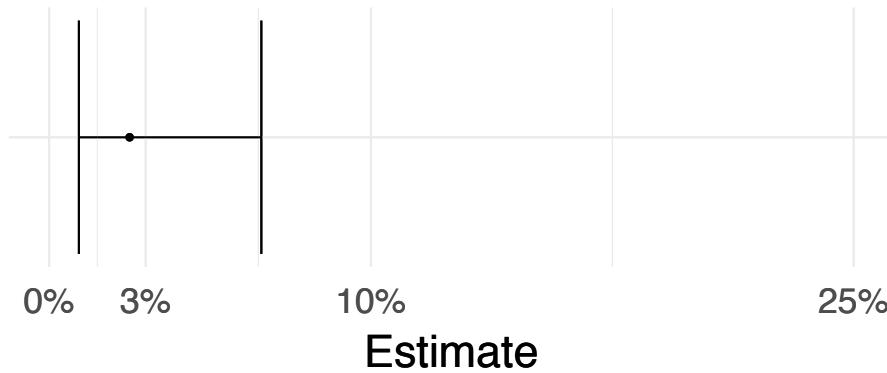
PAID SEARCH

1 Conversion 30 Clicks



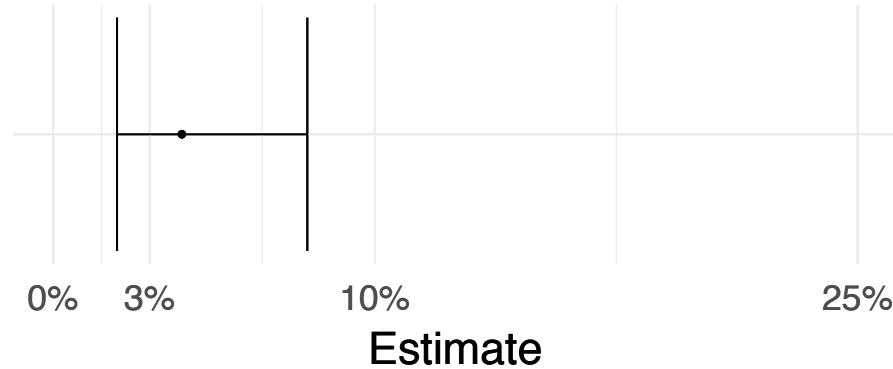
PAID SEARCH

1 Conversion 40 Clicks



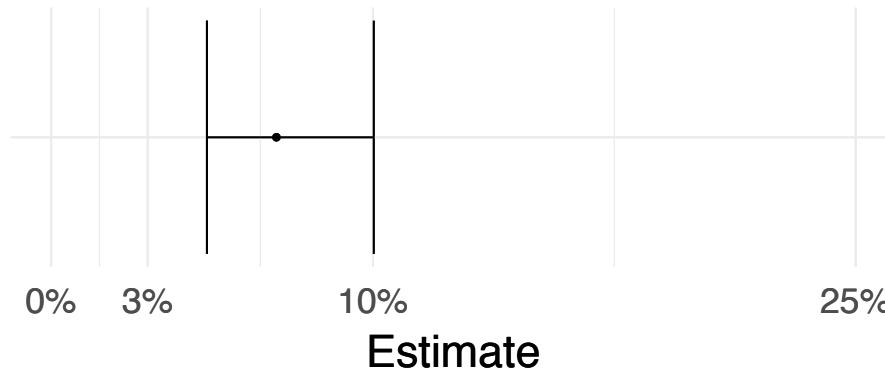
PAID SEARCH

2 Conversions 50 Clicks



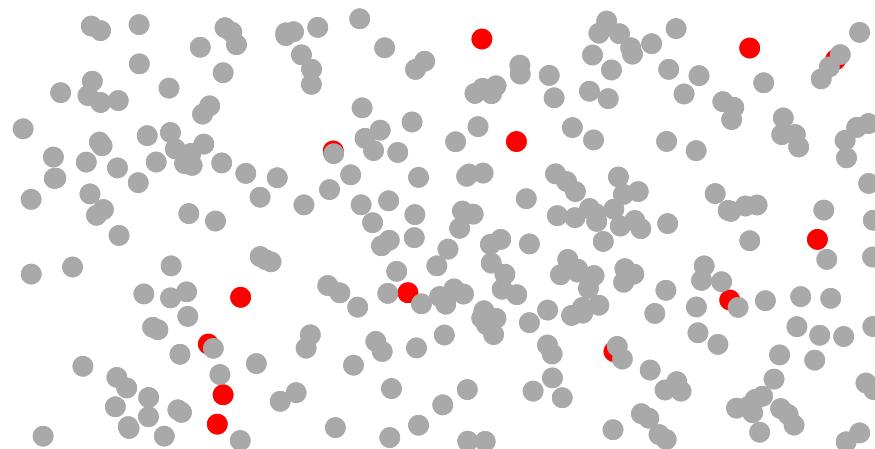
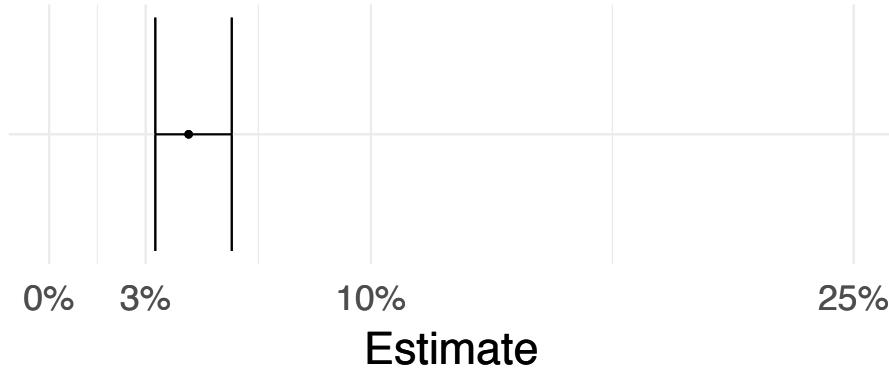
PAID SEARCH

7 Conversions 100 Clicks



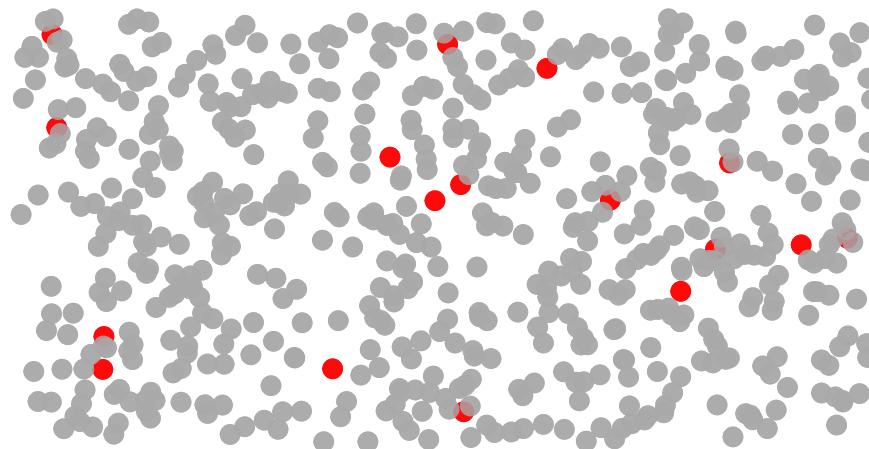
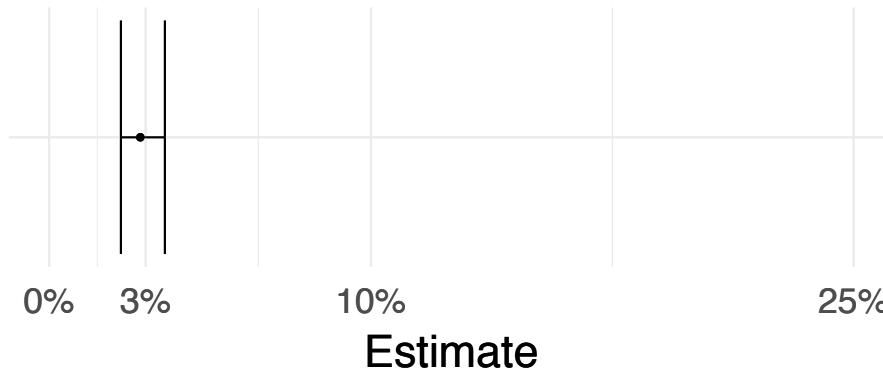
PAID SEARCH

13 Conversions 300 Clicks

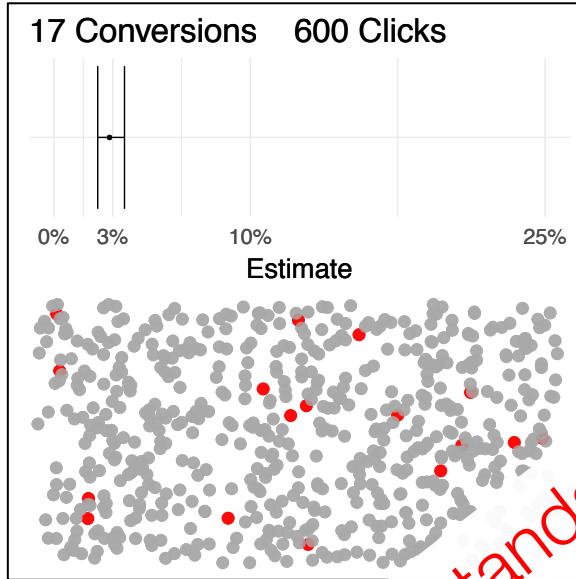


PAID SEARCH

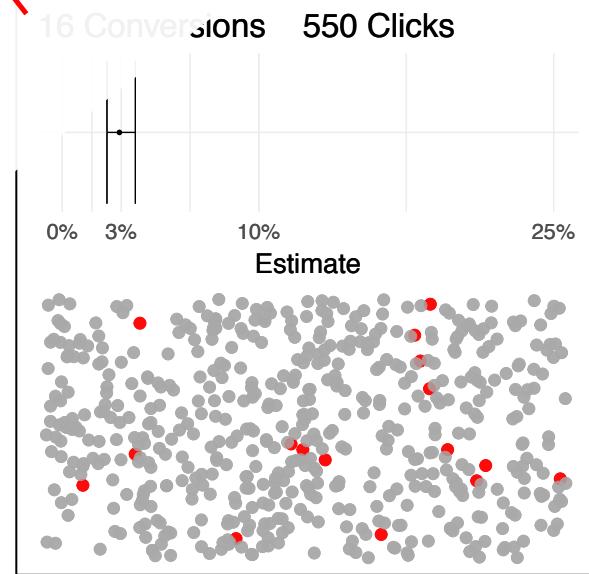
17 Conversions 600 Clicks



PAID SEARCH



Standard Regression Performs Poorly
in this Scenario



HOW DO YOU CHOOSE A CONVERSION RATE?

NO PRESCRIBED METHOD

1. Current Actual Value?
2. Mean value of prior brands?
3. Weighted value of all brands?

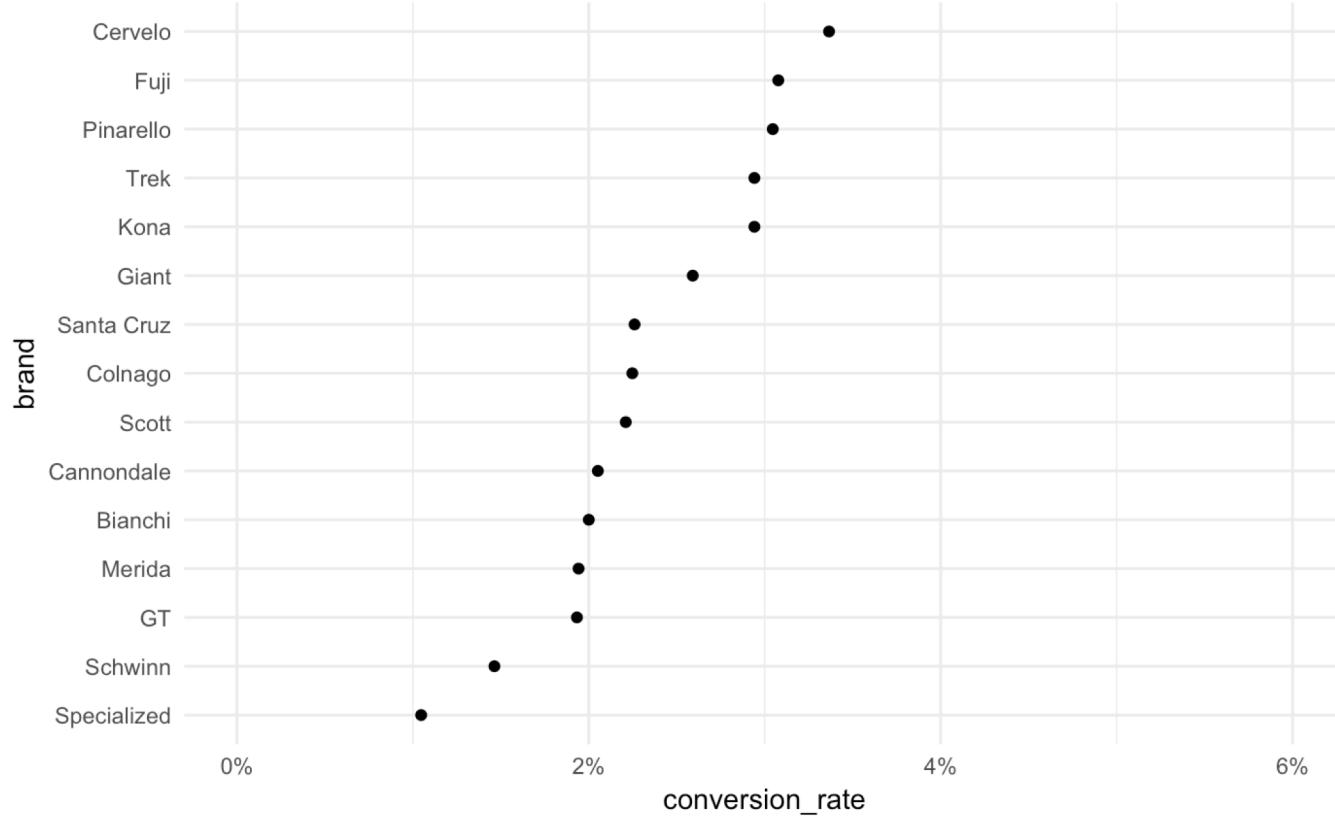
HOW DO YOU CHOOSE A CONVERSION RATE?

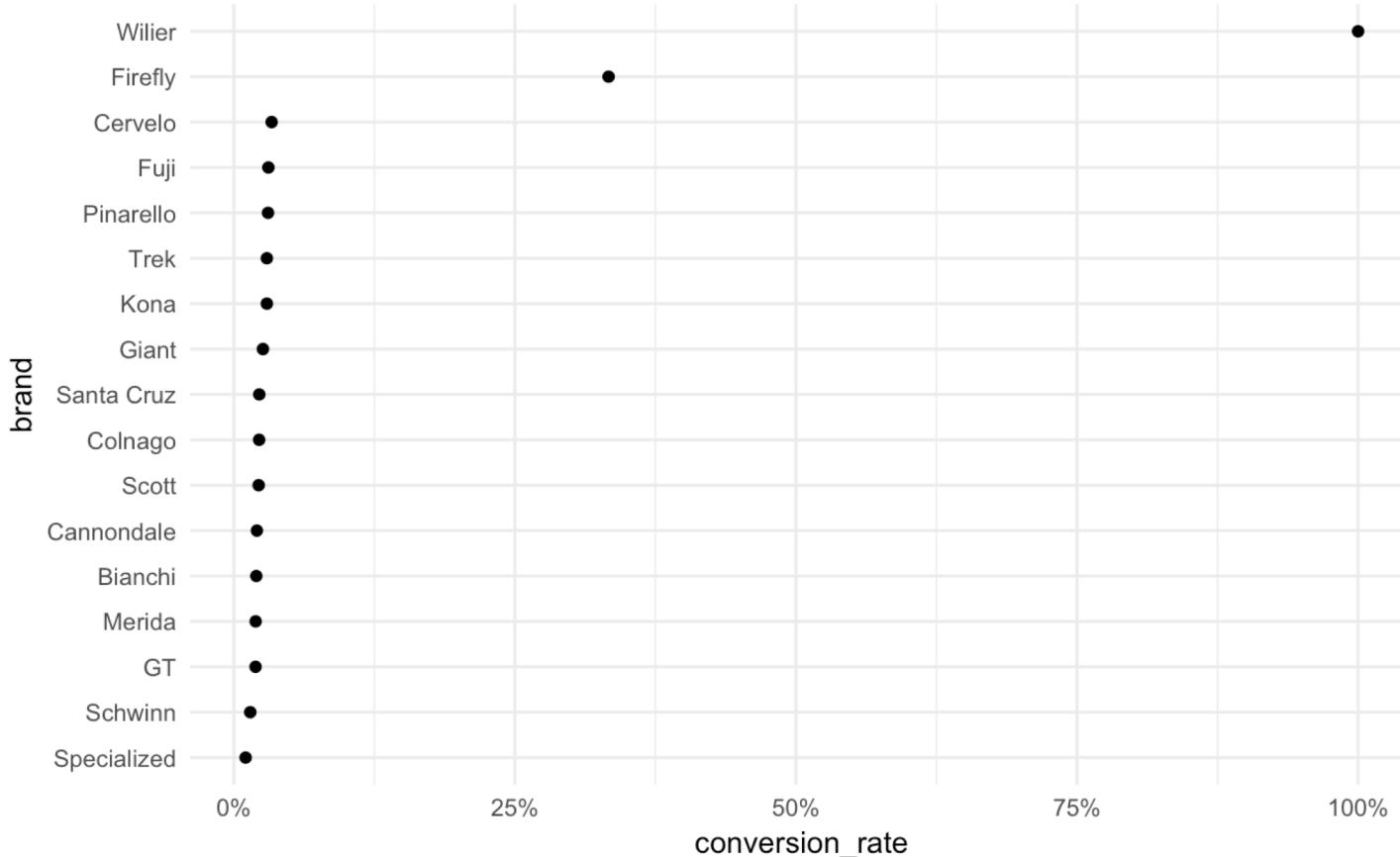
NO PRESCRIBED METHOD

1. Current Actual Conversion Rate?
2. Mean value of all brands?
3. Weighted average value of all brands?

2 and 3 do not allow modeling at the most granular brand level

WHAT DO WE KNOW?





PAID SEARCH

WHAT WE WOULD LIKE TO KNOW

We've Heard of Bayesian.

Let's Try a Bayesian Regression.

BINOMIAL DISTRIBUTION

$$\left(\text{CONVERSIONS}, \text{CLICKS} - \text{CONVERSIONS} \right) \sim \text{BRAND RATE}$$

REGRESSION ALGEBRA

$$\text{CONVERSIONS} \sim \left(\text{CLICKS}, \text{BRAND RATE} \right)$$

STAN CODE FOR REGRESSION WITHOUT PRIORS

PAID SEARCH

```
model {  
    CONVERSIONS ~  
}
```

STAN CODE FOR REGRESSION WITHOUT PRIORS

PAID SEARCH

```
model {  
    CONVERSIONS ~  
}  
    CLICKS
```

STAN CODE FOR REGRESSION WITHOUT PRIORS

PAID SEARCH

```
model {  
    CONVERSIONS ~ ( CLICKS X BRAND RATE )  
}
```



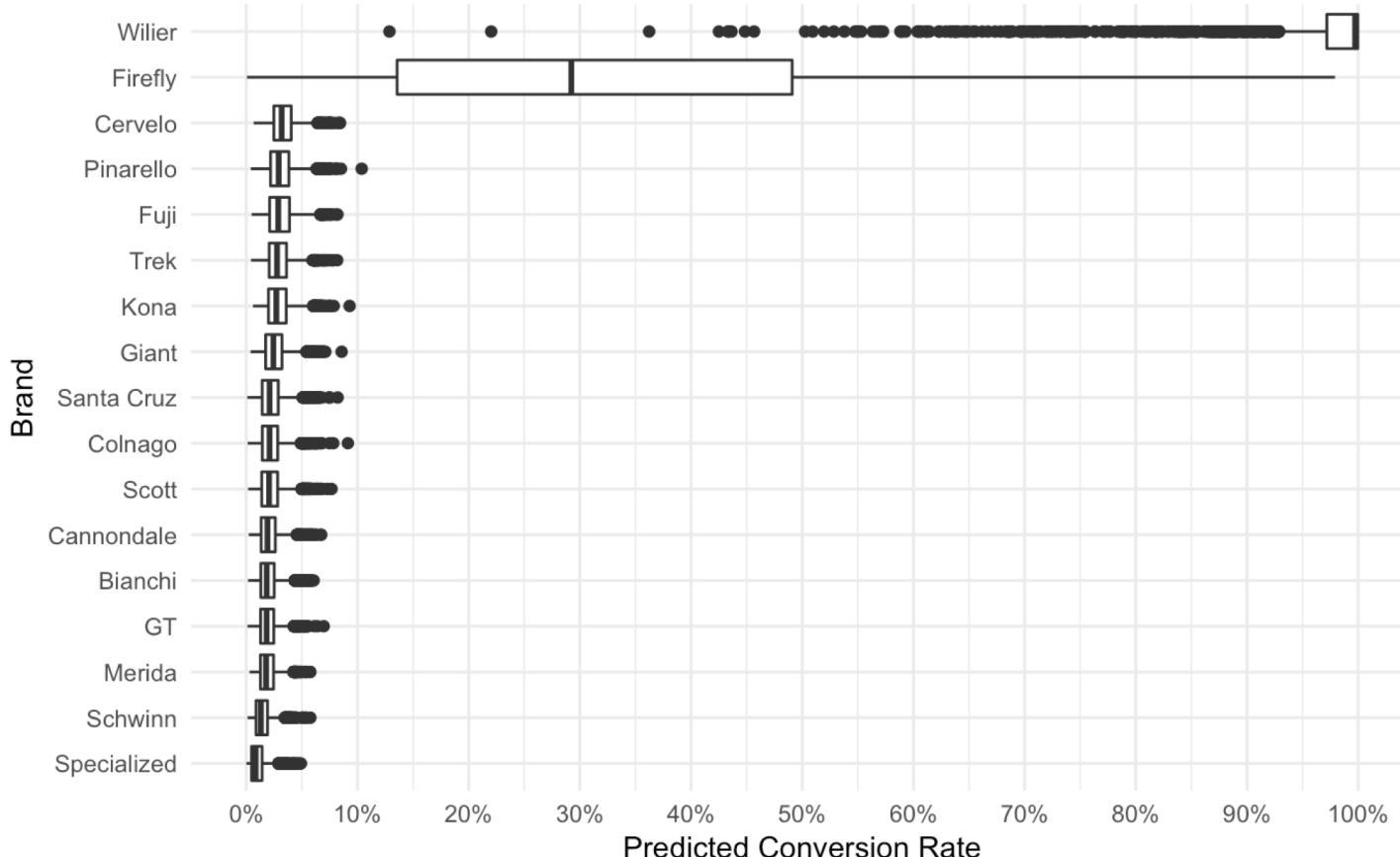
STAN CODE FOR REGRESSION WITHOUT PRIORS

PAID SEARCH

```
model {  
    CONVERSIONS ~ binomial_logit ( CLICKS X BRAND RATE )  
}
```

STAN CODE FOR REGRESSION WITH HIERARCHICAL PRIORS

```
model {  
    conversions ~ binomial_logit(clicks, brands * rate);  
}
```



PRIOR INFORMATION

Bayesian Analysis Can Also Use

Prior Information

BAYESIAN INFERENCE

$$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$$

BAYESIAN IS NOT MAGIC

$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$

Could be any number

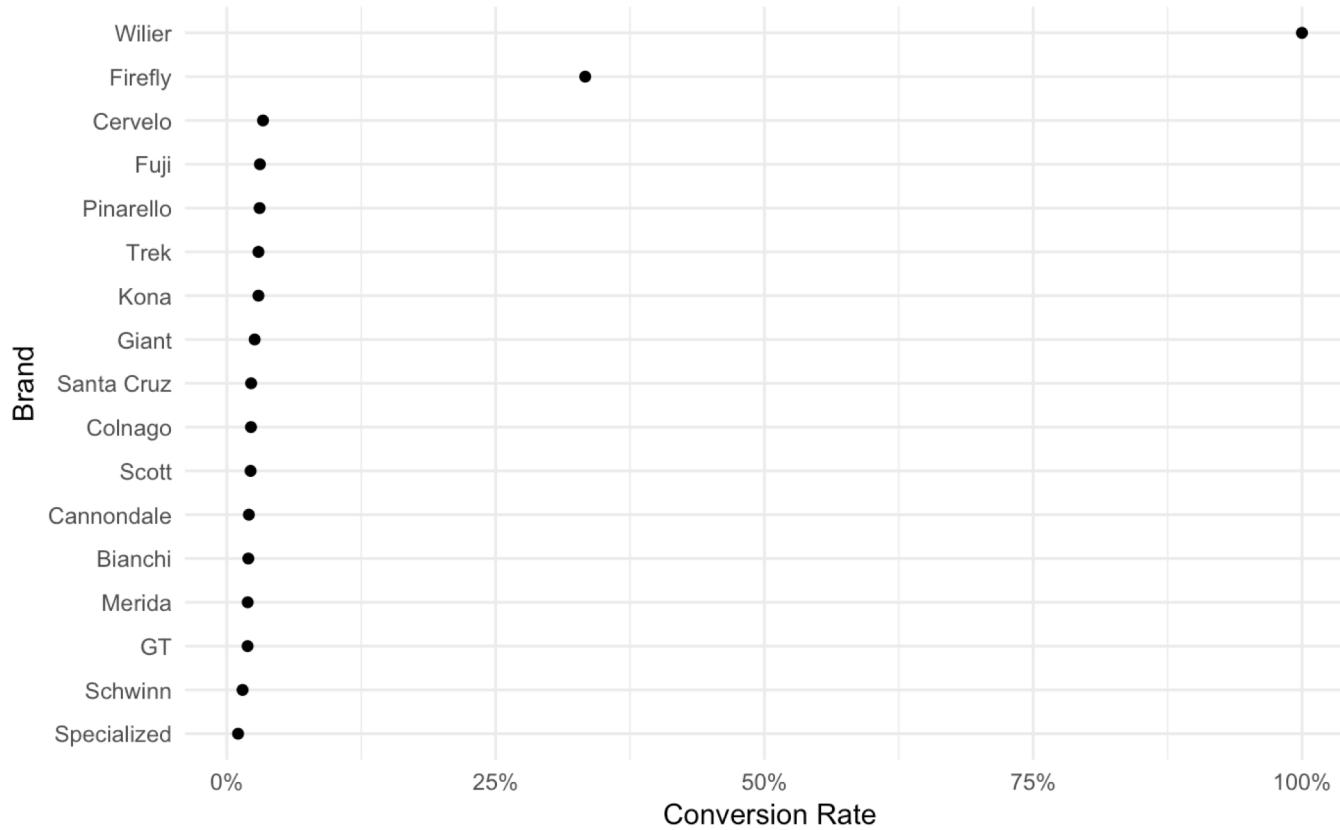
BAYESIAN INFERENCE

$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$

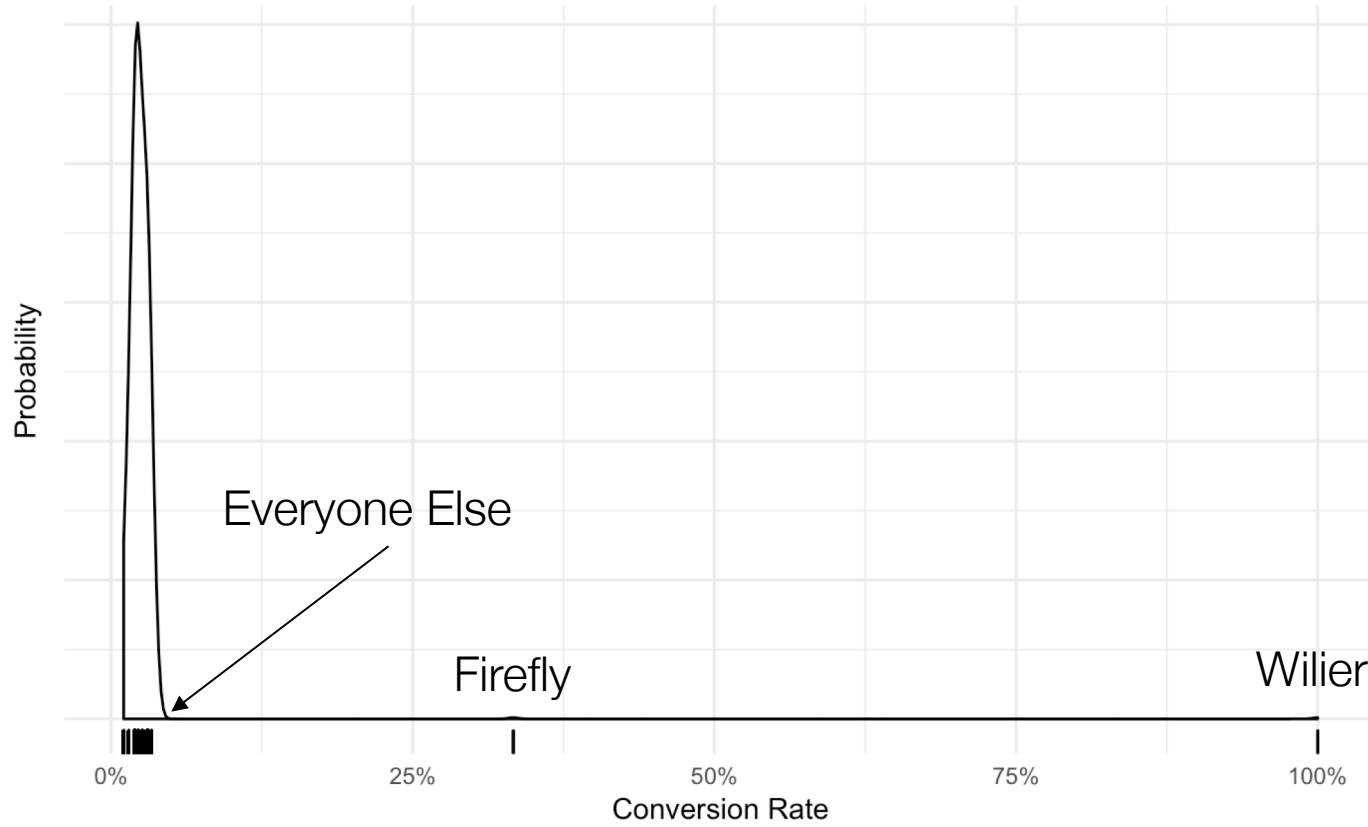
About 1% to 3%



BRAND HIERARCHY



THE PRIOR



BAYESIAN INFERENCE

What are the likely
rate values?

$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$

About 1% to 3%

BAYESIAN INFERENCE

WHAT WE WOULD LIKE TO KNOW

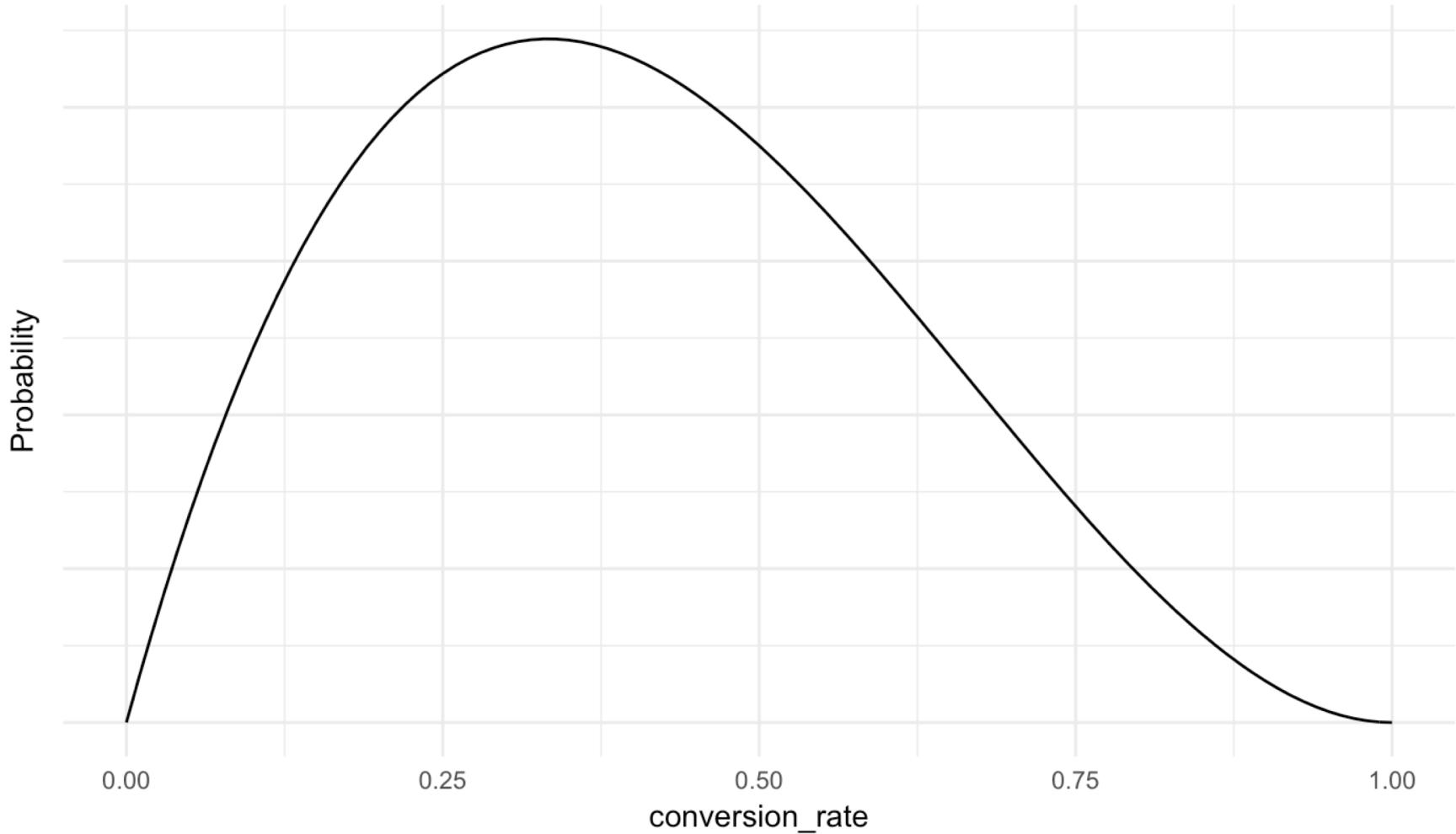
Binomial

$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$

About 1% to 3%



Firefly



BAYESIAN INFERENCE

WHAT WE WOULD LIKE TO KNOW

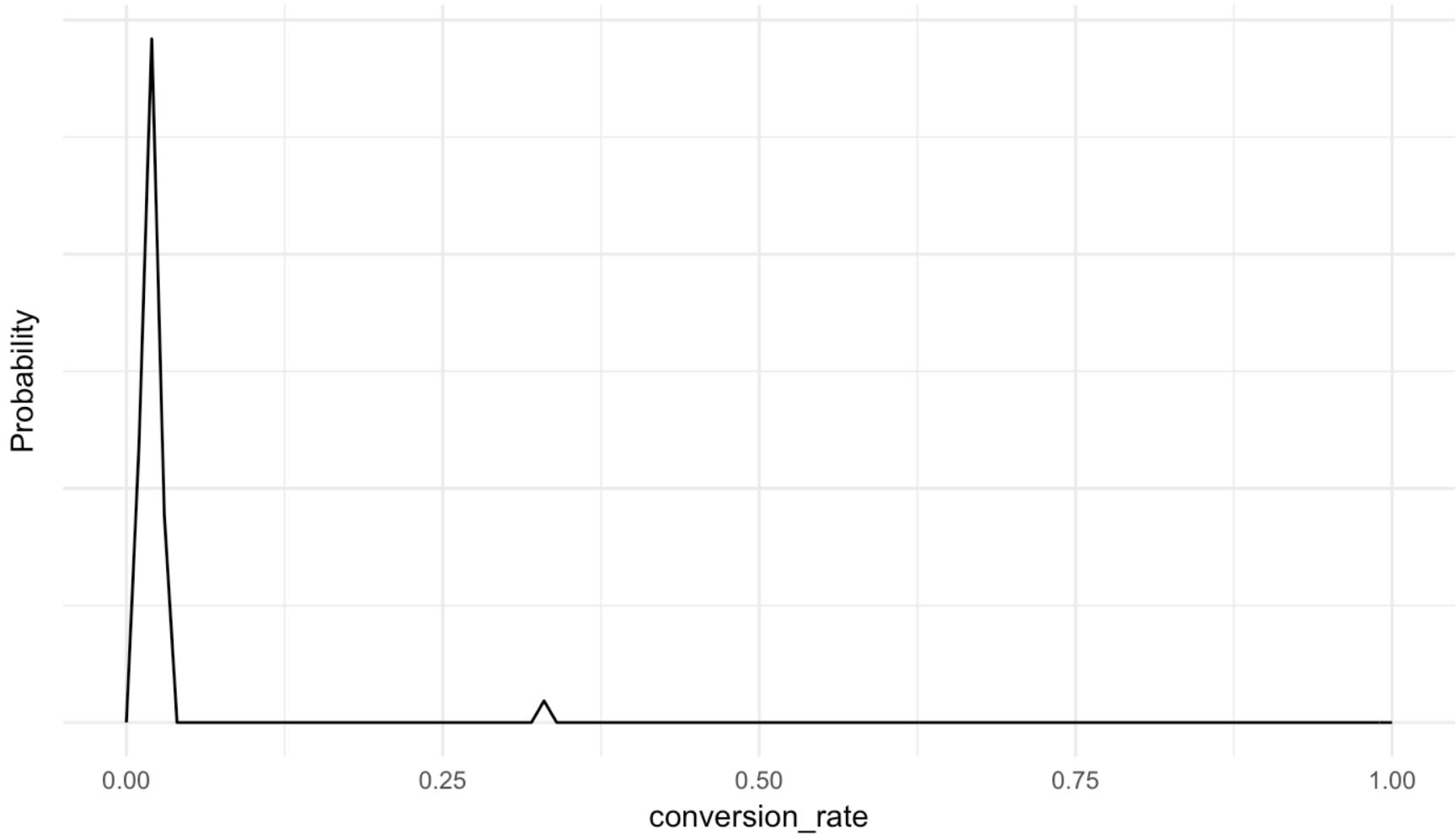
Binomial

$P(\text{Conversion Rate} \mid \text{Data}) \sim \text{Prior} * \text{Likelihood}$

About 1% to 3%



Bayesian Inference



HOW DO WE USE THIS?



STAN CODE FOR REGRESSION WITH HIERARCHICAL PRIORS

```
model {  
    conversions ~ binomial_logit(clicks, brands * rate); // Likelihood  
}  
                                About 1% to 3%
```

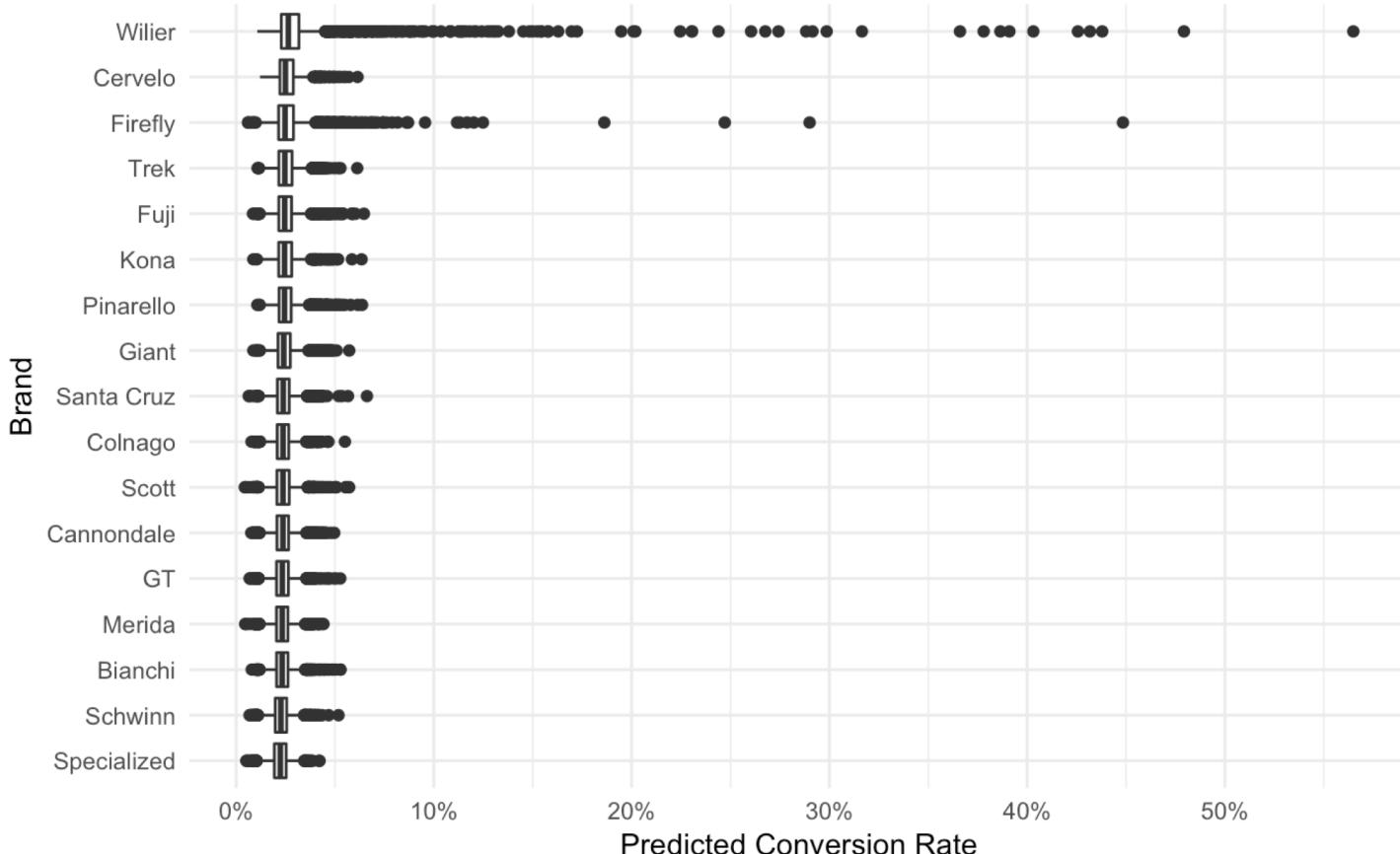
STAN CODE FOR REGRESSION WITHOUT PRIORS

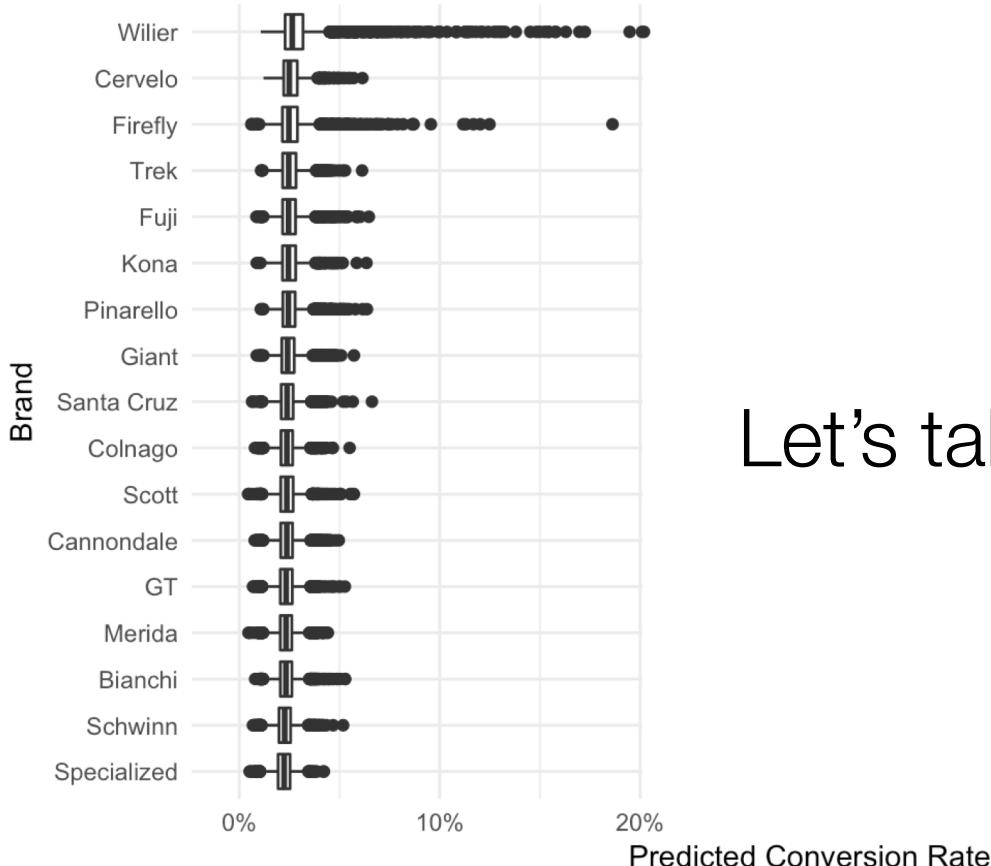
PAID SEARCH

```
model {  
    conversions ~ binomial_logit(clicks, brands * rate); // Likelihood  
    rate ~ normal(rate_mean, rate_var); // Prior  
}
```

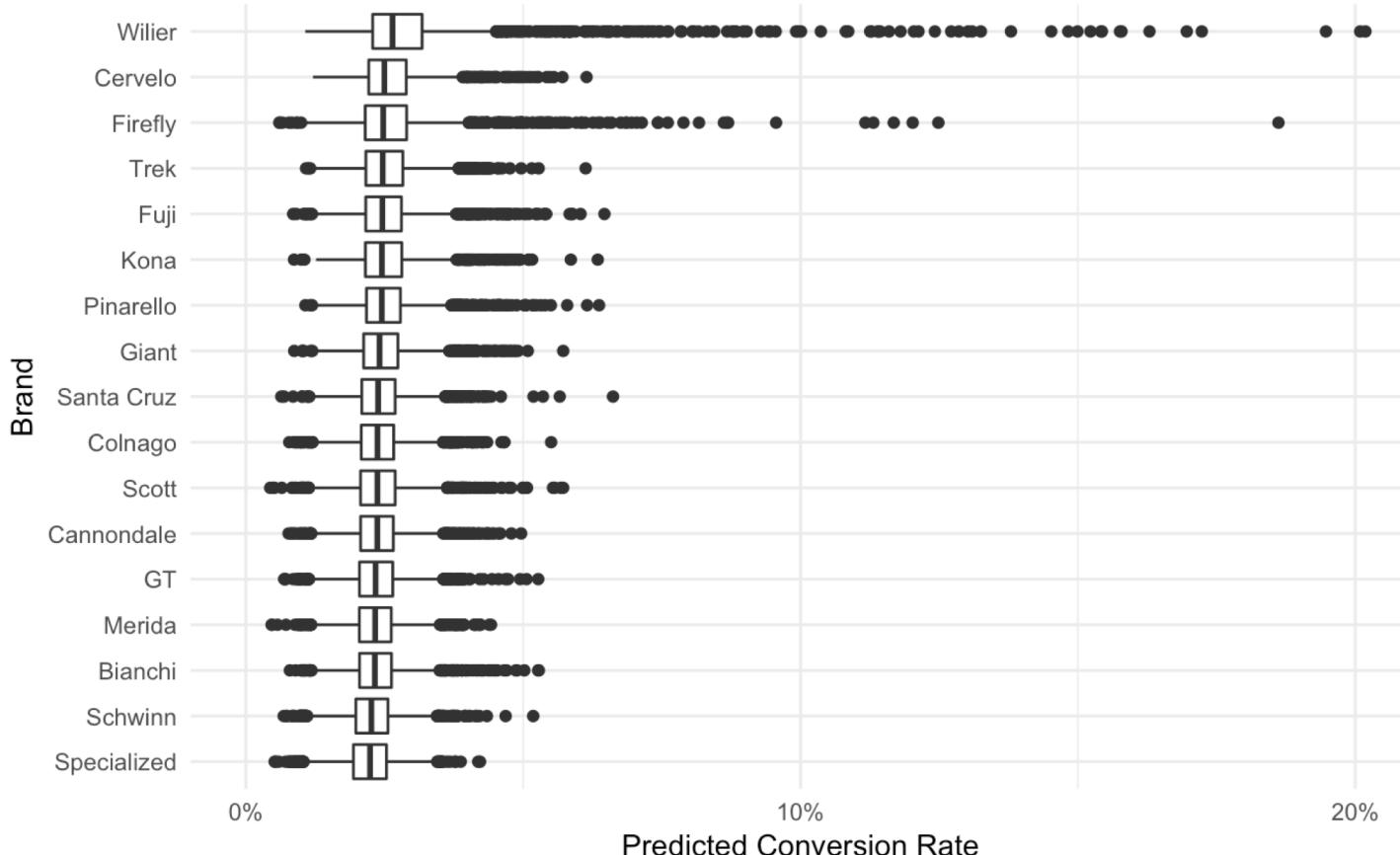


About 1% to 3%





Let's take a closer look



MAXIMUM BIDS

WILIER

Max Bid = Average Conversion Rate × Margin

$$\$2.72 = 2.72\% \times \$100.00$$

FIREFLY

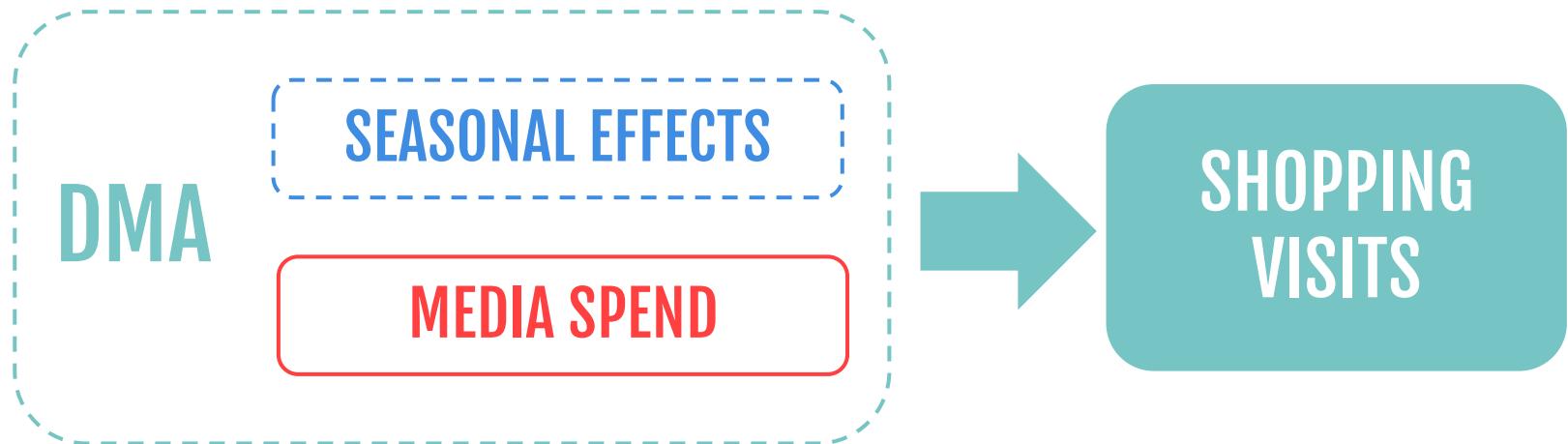
Max Bid = Average Conversion Rate × Margin

$$\$2.53 = 2.53\% \times \$100.00$$

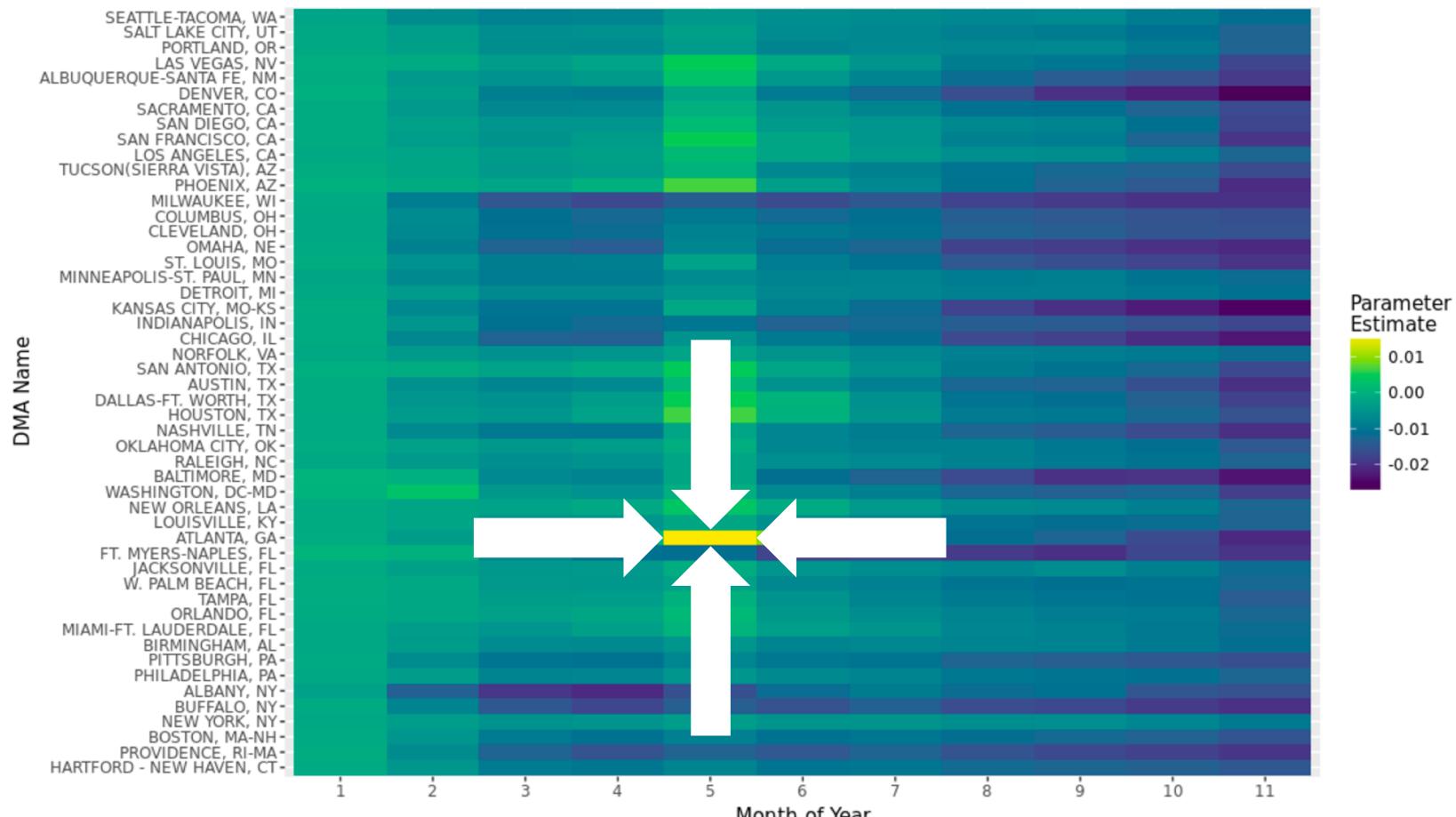
MEDIA MIX MODELS



NESTED MODEL STRUCTURE



Seasonal Effects



STAN CODE FOR MMO MODEL

```
model {  
    visits ~ normal(month * month_rate + dma_spend * dma_rate, var);  
    month_rate ~ normal(month_mean, month_var);  
    dma_rate ~ normal(dma_mean, dma_var);  
}
```



HOW DOES A MULTI-LEVEL MODEL COMPARE WITH THE CURRENT MODEL?

IN TERMS OF PREDICTION AND MODEL FIT, THE MULTI-LEVEL MODEL IS A SIGNIFICANT IMPROVEMENT

STANDARD MODEL

R²: 81%

MAPE: 15%

MULTI-LEVEL MODEL

R²: 90%

MAPE: 9%

THE TAKEAWAY FOR TODAY



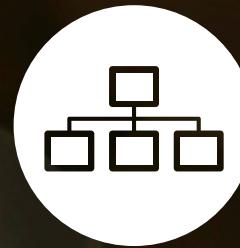
Regression
might perform
poorly



Bayesian
Methods
aren't magic



Prior
information
can improve
regression



Hierarchy can
be a part of
that prior

THE BENEFITS OF BEING BAYESIAN

HOW MULTI-LEVEL MODELS ENABLE MORE GRANULAR INSIGHTS

REFERENCES

- Statistical Rethinking, A Bayesian Course with Examples in R and Stan
 - Richard McElreath
 - <http://xcelab.net/rm/statistical-rethinking/>
- Data Analysis Using Regression and Multilevel/Hierarchical Models
 - Andrew Gelman, Jennifer Hill
 - <https://www.amazon.com/Analysis-Regression-Multilevel-Hierarchical-Models/dp/052168689X>
- Gist containing code used for the Bike Shop example
 - <https://bit.ly/pawbayes>

