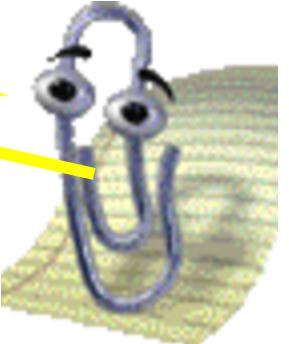




# IDS702 Team Project

October 21<sup>st</sup>, 2021

Anna, Dauren, aTHENA, Himangshu, mOritz



(just kidding)

# **Methadone Prices on the Black Market Across the United States**

IDS702 Team Project Group Yellow

Duke

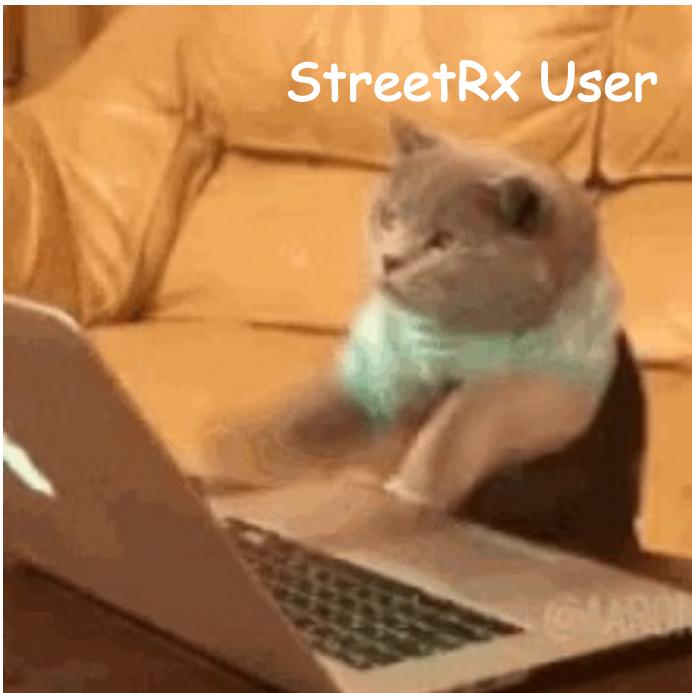
# Data & Modeling Process

---

## Data Set: Methadone

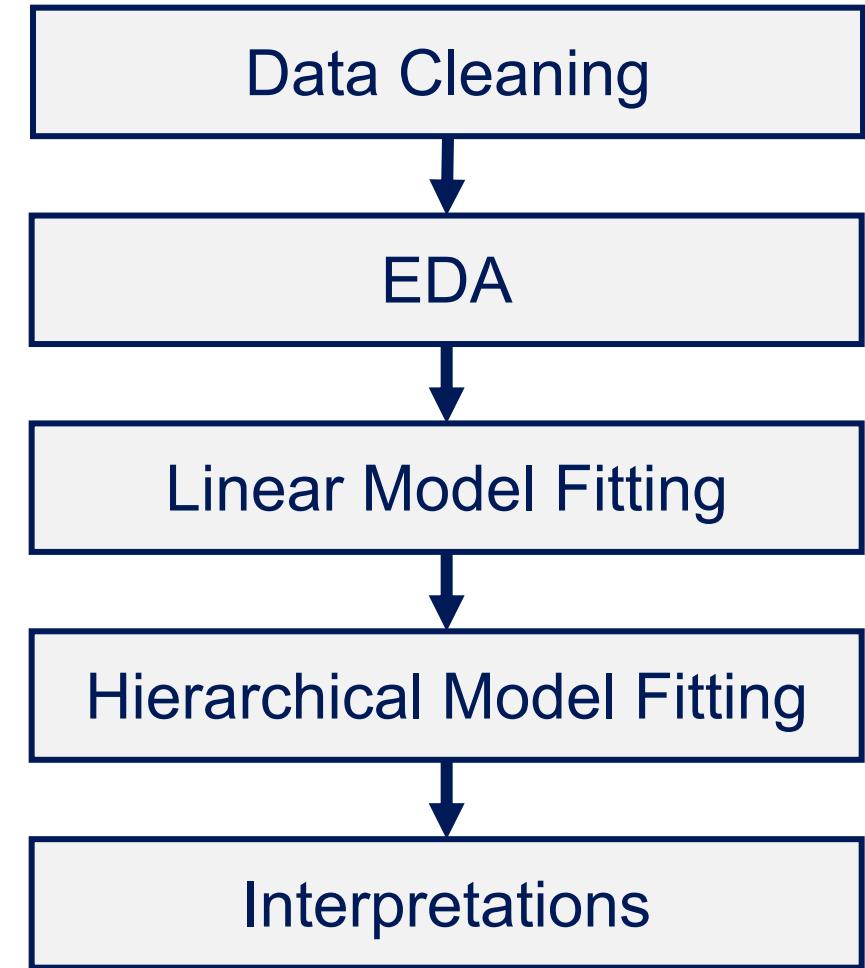
---

- **StreetRx:** Crowdsourced data on diverted prescription drugs
- Unreliable data inputs



## Process

---



# Data Cleaning

---

## Consolidation of source

---



## Missing Values

---

- Removed N/A values in ppm  
⇒ also eliminated N/A values in mgstr

## Levels of mgstr

---

mgstr	1	2.5	5	10	15	40
count	2	1	781	3047	1	351

fac_mgstr	ppm	state	USA_region	source	bulk_purchase
15	0.01	USA	Other/Unknown	Internet	Bulk
1	10.00	Pennsylvania	Northeast	Internet	Not Bulk
1	40.00	Washington	West	No Input	Not Bulk
2.5	40.00	Nevada	West	No Input	Not Bulk

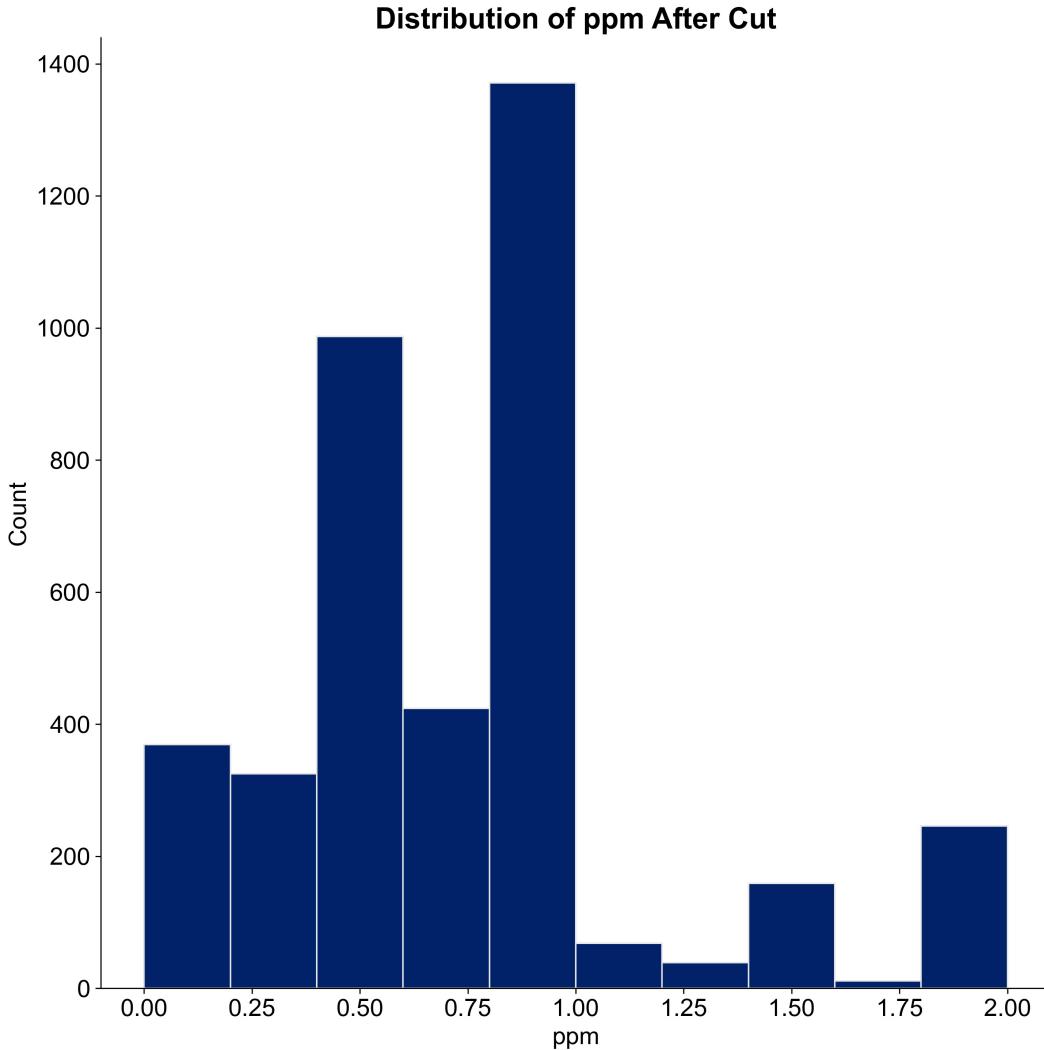
Removed uncommon mgstr & factorized it

# EDA

---

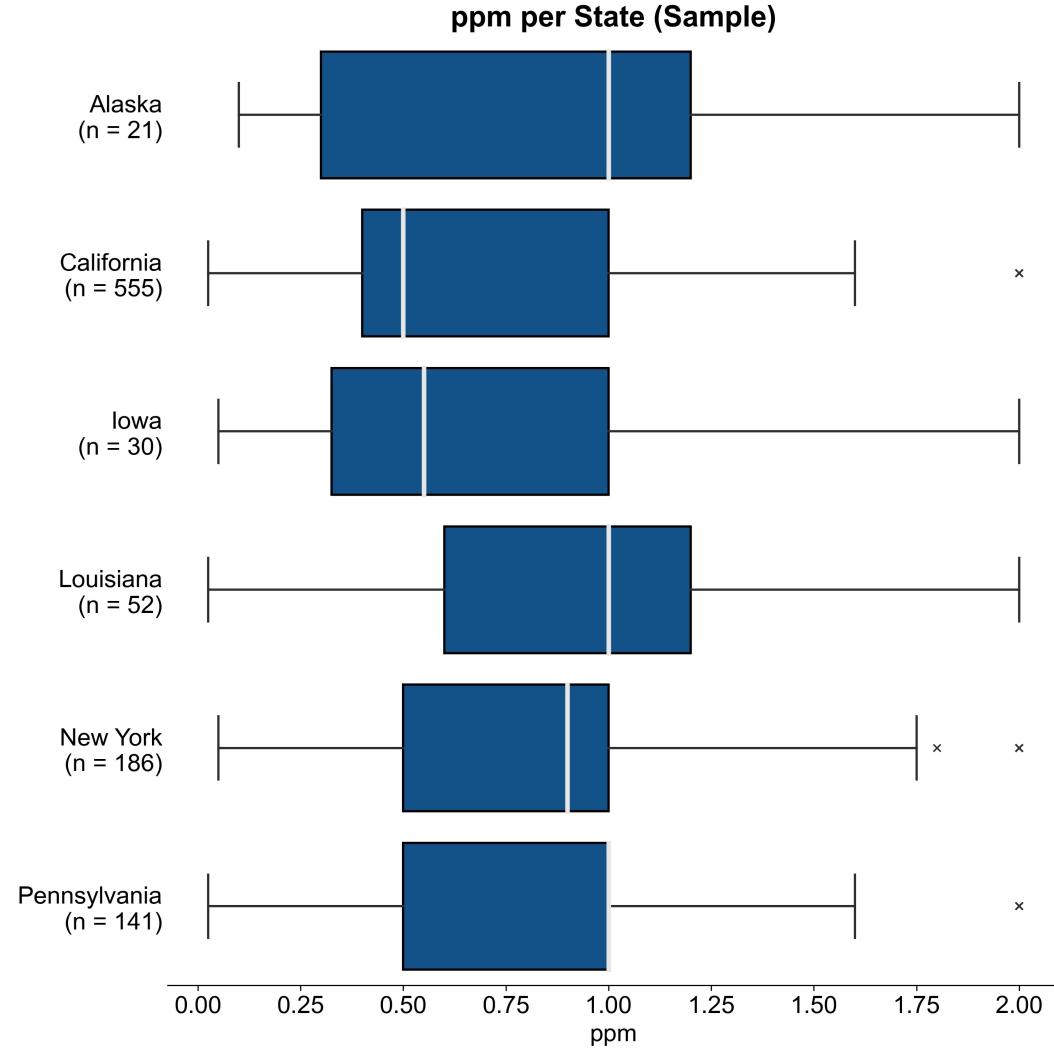
## Distribution of Price per mg

---



## Distribution by Location

---



# Linear Model Fitting

**Null Model**  
Intercept only

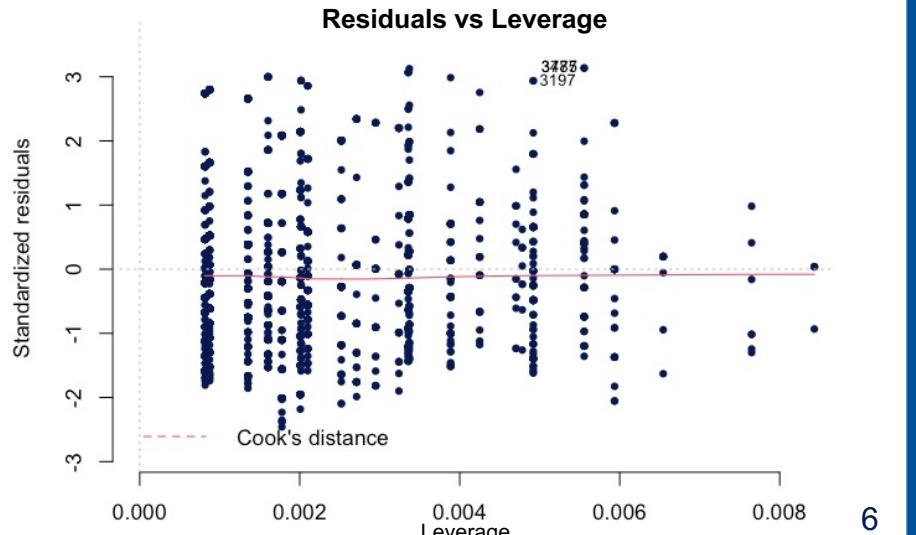
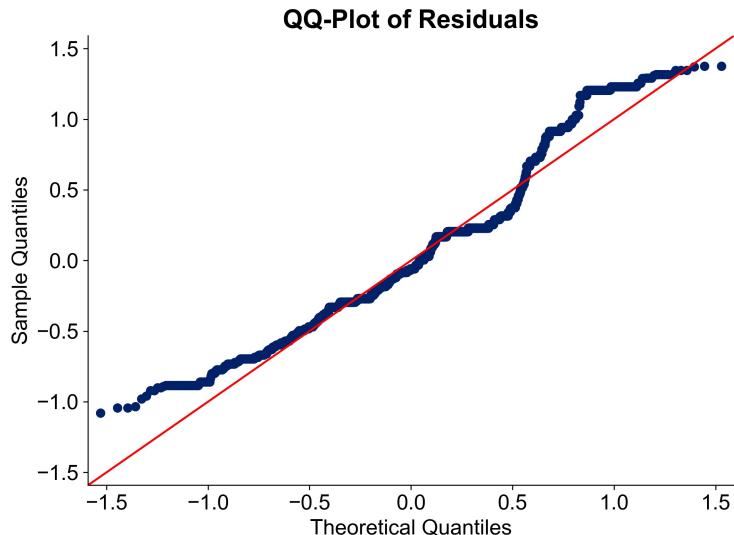
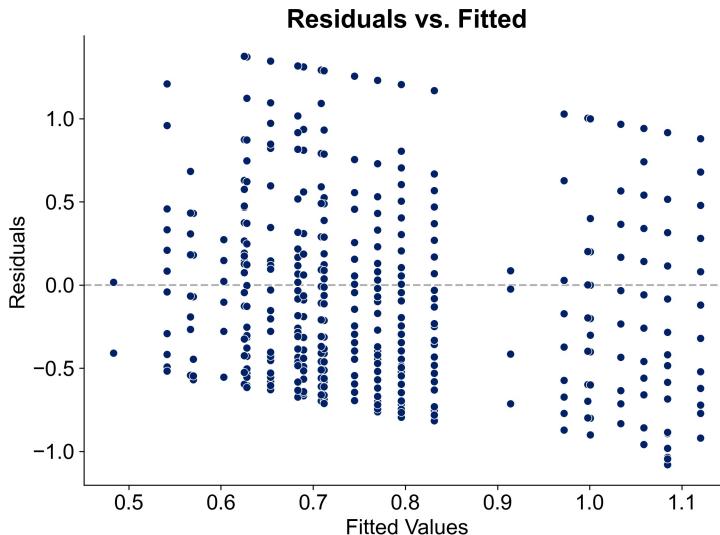
Parameters  
AIC  
Direction = both

**Full Model**  
All non-location predictors

## Resulting Model

$$\widehat{ppm}_i = \widehat{\beta}_0 + \sum_{j=2}^3 \widehat{\beta}_{1j} \mathbb{I}[fac\_mgstr_i = j] + \widehat{\beta}_{2j} \mathbb{I}[bulk\_purchase_i = j] + \sum_{j=2}^4 \widehat{\beta}_{3j} \mathbb{I}[source_i = j]$$

## Model Assessment



# Hierarchical Model Fitting

## Choice of Levels



## State-Only vs. State & Region

	npar	AIC*	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
state only	9.00	4717.62	4774.26	-2349.81	4699.62			
state and region	10.00	4715.35	4778.29	-2347.68	4695.35	4.27	1	0.04

\* Difference from above due to ML vs. REML refitting

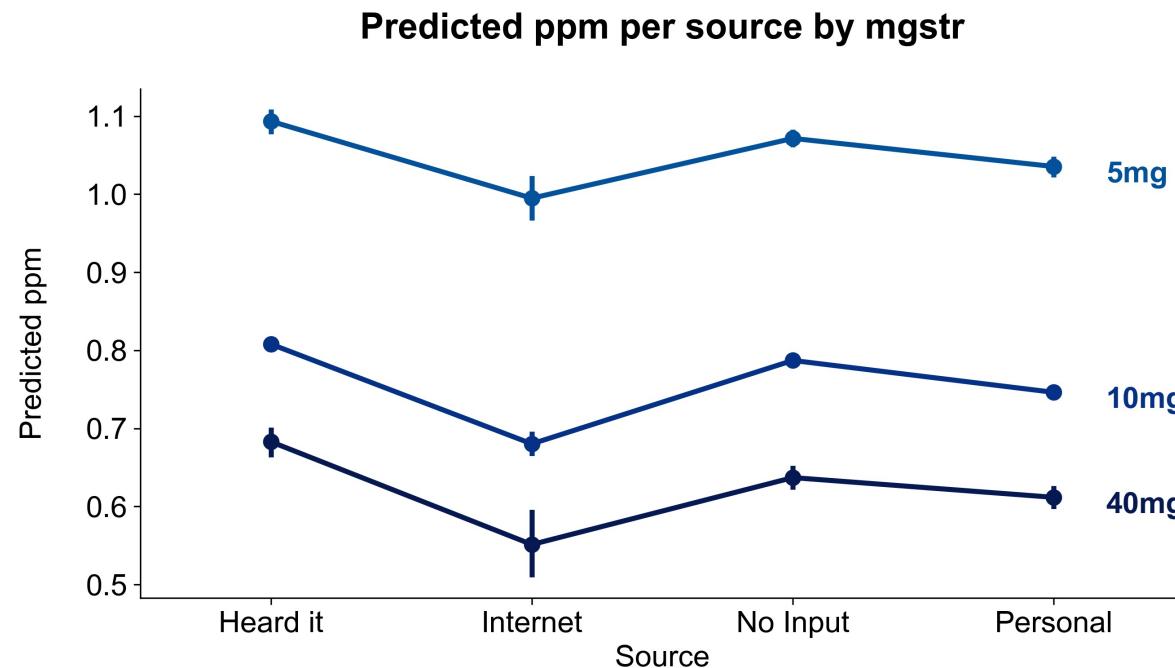


```
ppm ~ fac_mgstr + bulk_purchase + source + (1 | USA_region) + (1 | state)
```

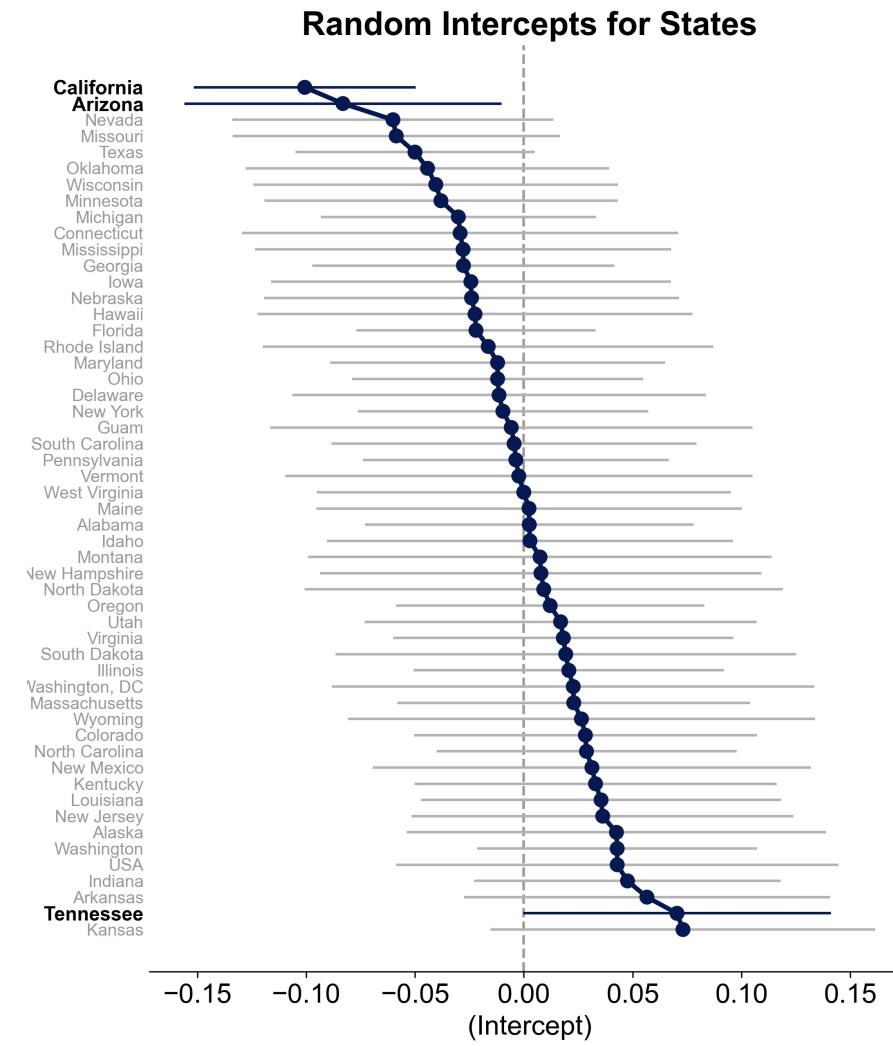
# Interpretation

## Fixed Effects

	Estimate	Std. Error	t value	Lower Bound	Upper Bound
(Intercept)	1.14	0.03	34.65	1.0789	1.2131
fac_mgstr10	-0.29	0.02	-16.06	-0.3283	-0.2569
fac_mgstr40	-0.46	0.03	-15.87	-0.5191	-0.4049
bulk_purchaseBulk	-0.08	0.02	-4.47	-0.1110	-0.0435
sourceInternet	-0.12	0.03	-3.55	-0.1843	-0.0531
sourceNo Input	-0.03	0.02	-1.84	-0.0721	0.0022
sourcePersonal	-0.06	0.02	-3.09	-0.0947	-0.0213

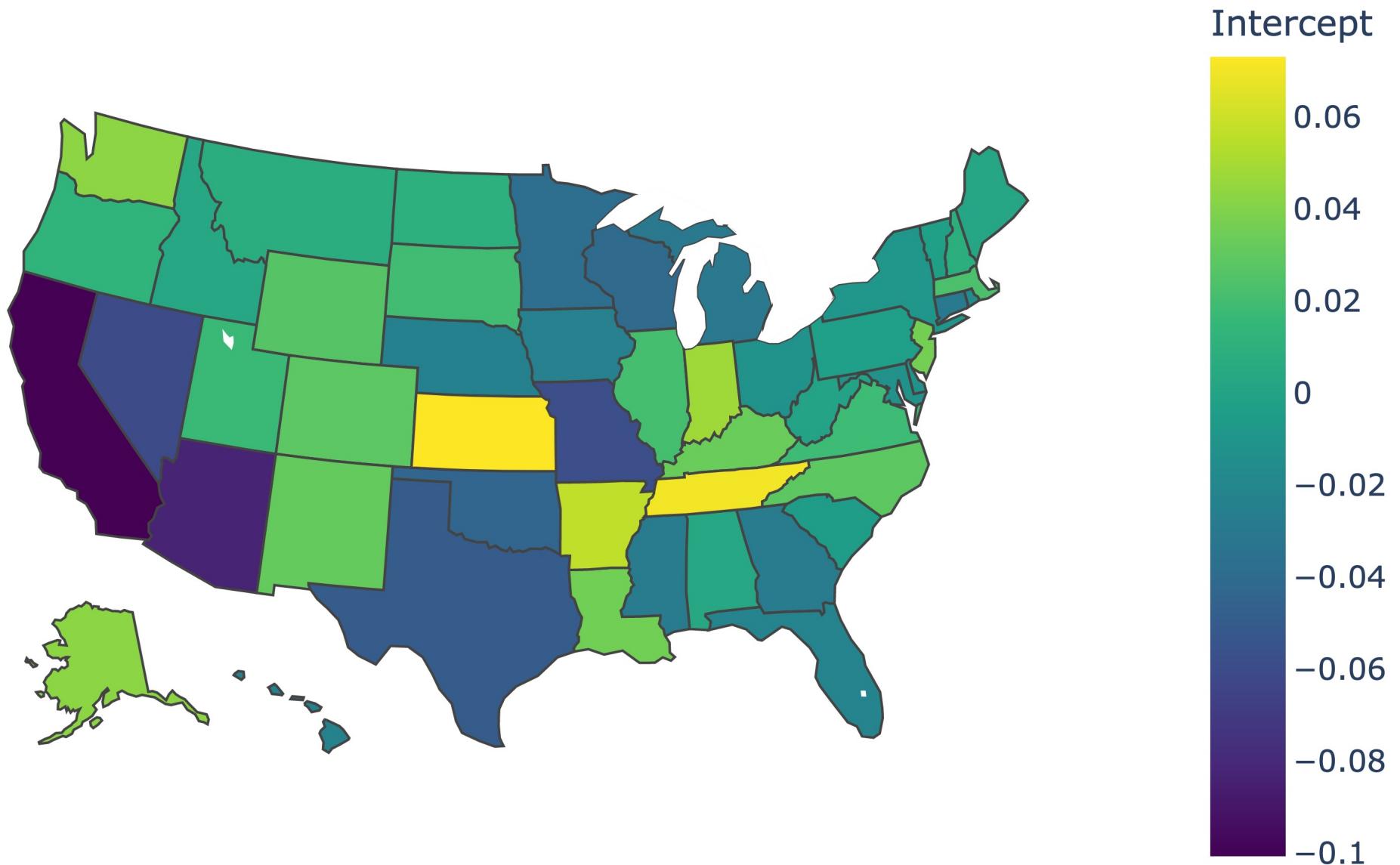


## Random Effects: State Level



# Geographical Visualization

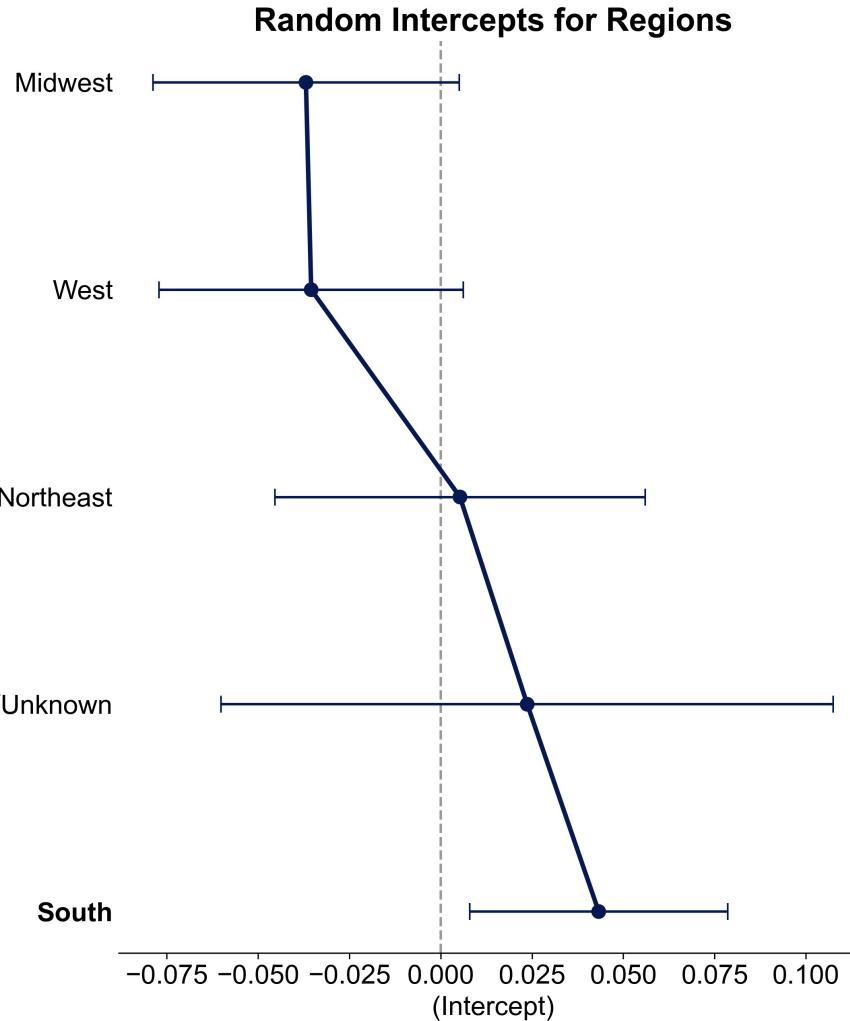
---



# Interpretation

---

## Random Effects: Region Level



## Random Effects: Summary

Groups	Name	Variance	Std.Dev.
state	(Intercept)	0.0033	0.0570
USA_region	(Intercept)	0.0021	0.0456
Residual		0.1878	0.4334

