US HONEY PRODUCTION PROJECT PRESENTATION

Presentation by Group 8 [Harshini Karthikeyan, Adam Kong, and Alisa Krasilnikov]

INTRODUCTION

RESEARCH QUESTION: What factors work to predict the value of honey production? Does this vary by state? Does this vary by year?

DATA SOURCE: US Honey Production 1995 - 2021

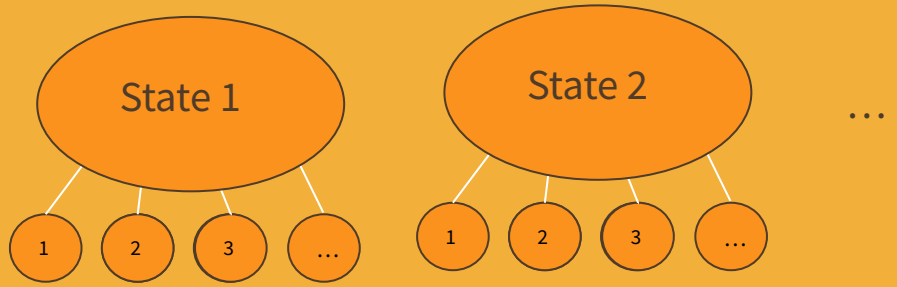
Kaggle, https://www.kaggle.com/datasets/mohitpoudel/us-honey-production-19952021/data

Overview of Dataset

Response: Value of Production

Level 1 Units: Occasions

Level 2 Units: States

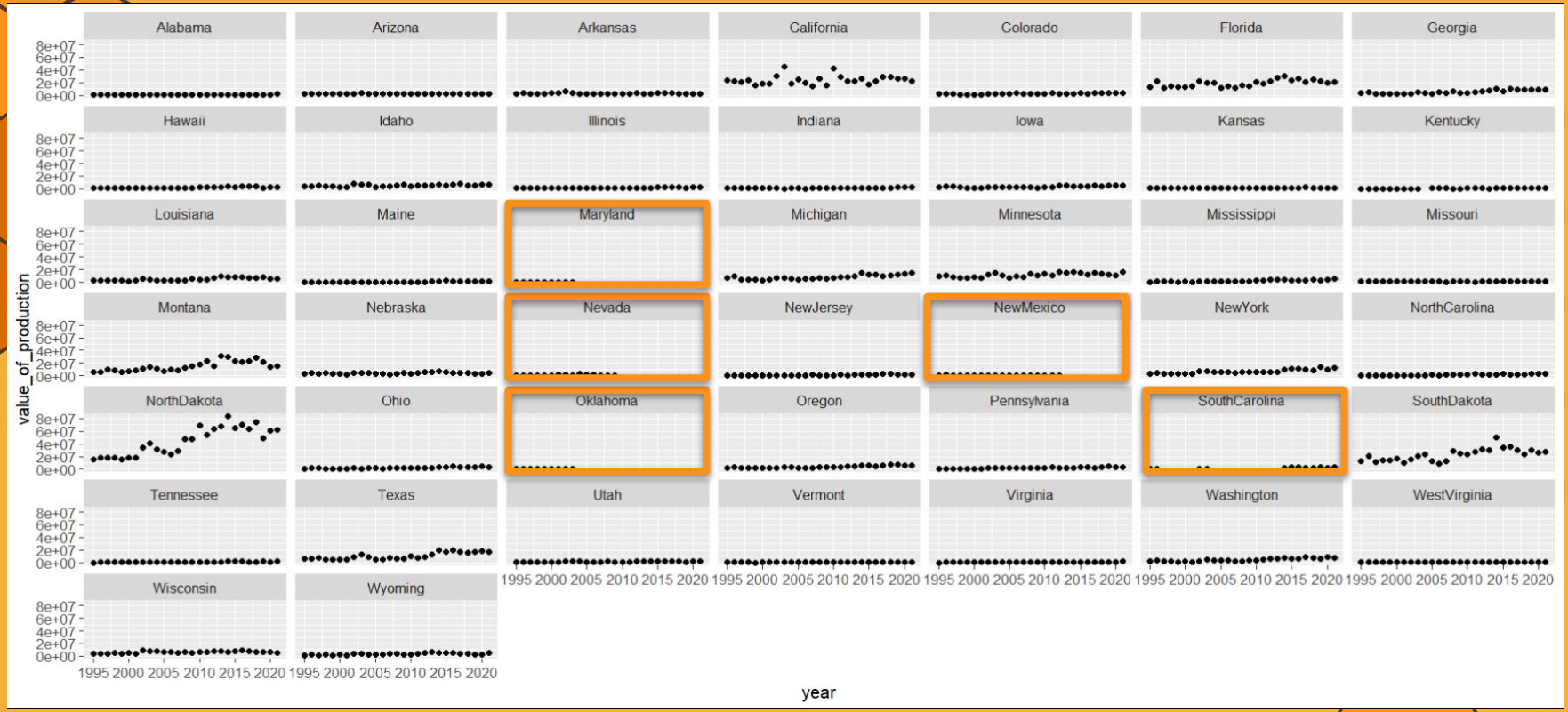


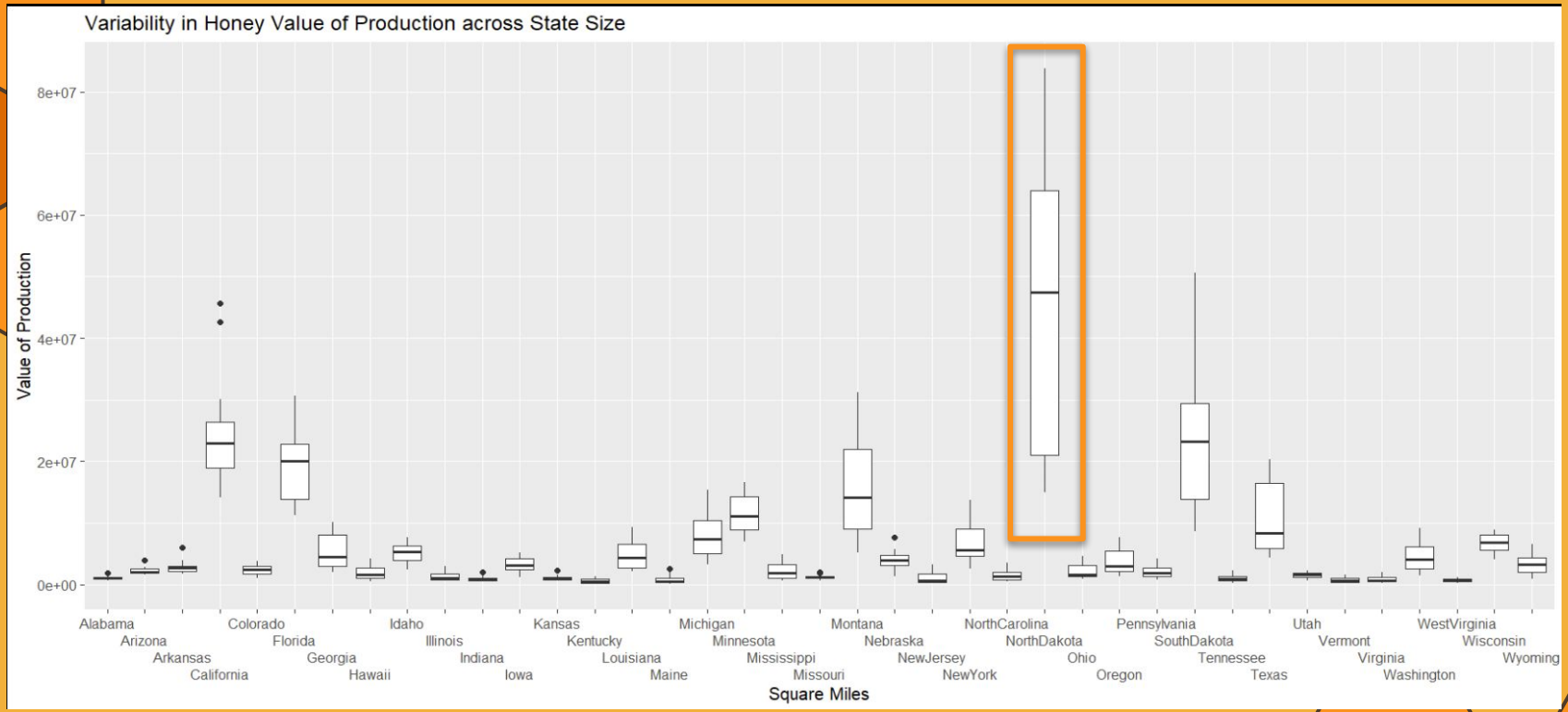
Level 1 Variables: Colonies Number, Yield Per Colony, Production, Stocks, Average Price

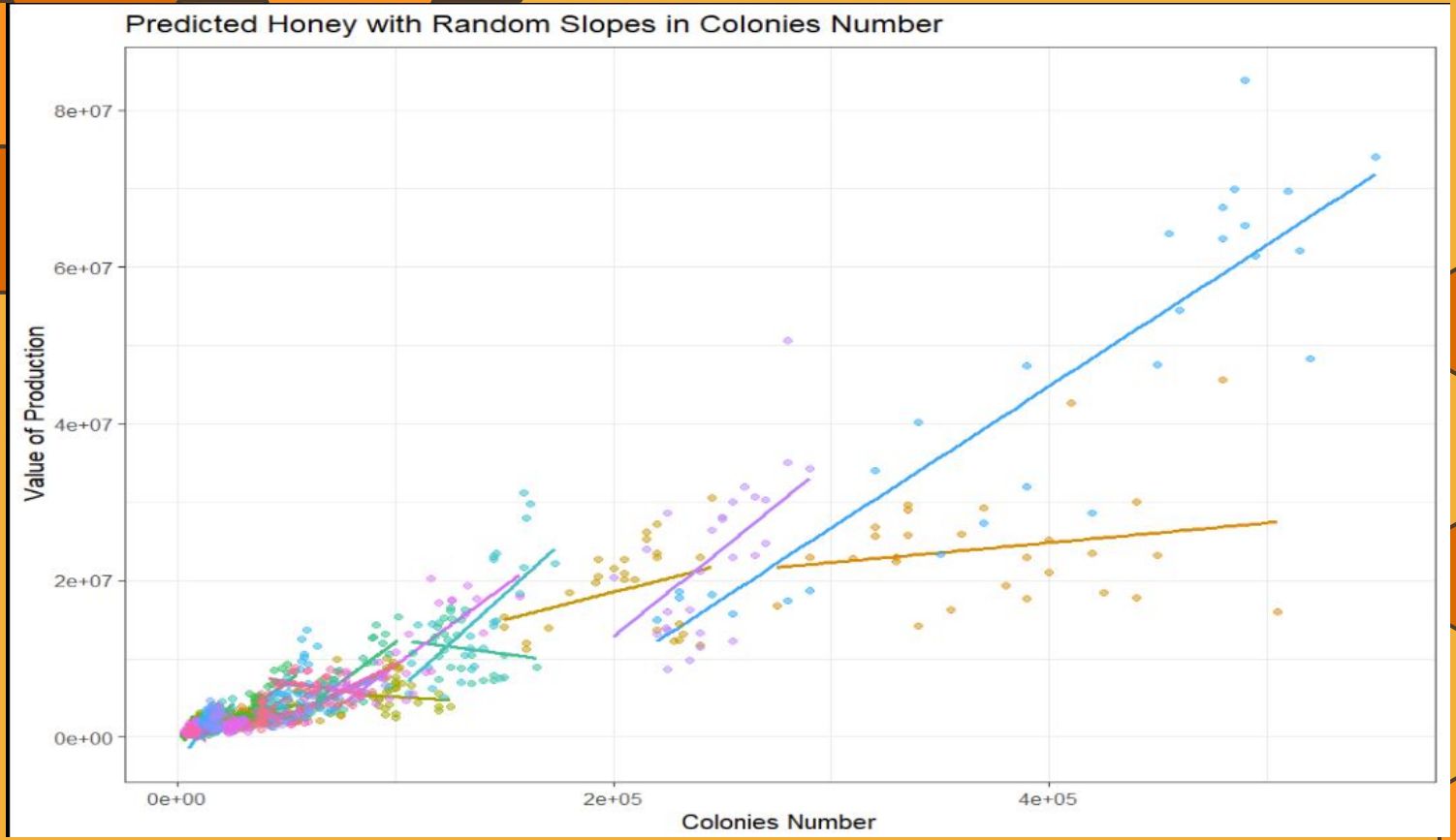
Level 2 Variables: Land mass

ADDITIONAL QUESTIONS FOR RESEARCH

- Observations we needed to get rid of?
- Any other variables not included in this initial dataset?
  - Land mass per state? (merged in for initial analysis)
  - Size of agriculture fields?
  - Grasslands/pastures size?
  - Temperature?


Alabama
Arizona
Arkansas
California
Colorado
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
NewJersey
NewMexico
NewYork
NorthCarolina
NorthDakota
Ohio
Oklahoma
Oregon
Pennsylvania
SouthCarolina
SouthDakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
WestVirginia
Wisconsin
Wyoming
value_of_production
0e+00
2e+07
4e+07
6e+07
8e+07
1995 2000 2005 2010 2015 2020
year


Variability in Honey Value of Production across State Size
Value of Production
8e+07
6e+07
4e+07
2e+07
0e+00
Alabama
Arizona
Arkansas
California
Colorado
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
NewJersey
NewYork
NorthCarolina
NorthDakota
Ohio
Oregon
Pennsylvania
SouthDakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
WestVirginia
Wisconsin
Wyoming
Square Miles


Predicted Honey with Random Slopes in Colonies Number
Value of Production
8e+07
6e+07
4e+07
2e+07
0e+00
0e+00
2e+05
4e+05
Colonies Number

FIRST MODEL

```
Random effects:
 Groups   Name        Variance  Std.Dev.
 state    (Intercept) 7.206e+13 8488898
 Residual             2.117e+13 4600820
Number of obs: 1052, groups:  state, 39

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5947462    1366693   4.352
```

Null Model

Model with States as a random effect

CURRENT MODEL (?)

```
Random effects:
 Groups   Name                Variance  Std.Dev.  Corr
 state    (Intercept)         7.661e+12 2.768e+06
          cen_colonies_number 3.880e+03 6.229e+01 1.00
 Residual                     5.715e+12 2.391e+06
Number of obs: 1052, groups:  state, 39

Fixed effects:
                     Estimate Std. Error t value
(Intercept)          3.075e+06  5.505e+05   5.586
cen_colonies_number  8.545e+01  1.141e+01   7.492
cen_stocks          -1.424e+00  8.224e-02 -17.319
year_95              2.015e+05  1.037e+04  19.436
cen_stocks:year_95   1.052e-01  5.913e-03  17.783
```

model5 <- lmer(value_of_production ~ 1 + cen_colonies_number + cen_stocks + year_95 + cen_stocks:year_95 + (1 + cen_colonies_number | state), data = honey_final, REML = FALSE)

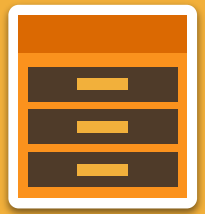
Final Model

WARNING TO CONSIDER

```
fit warnings:
Some predictor variables are on very different scales: consider rescaling
```

```
Warning: Some predictor variables are on very different scales: consider rescalingboundary
(singular) fit: see help('isSingular')
```

UNUSED VARIABLES



YIELD PER COLONY

Repetitive*

PRODUCTION

Repetitive

AVERAGE PRICE

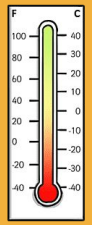
Difficult Inclusion

SQUARE MILES

Insignificant*

GRASSLAND

Insignificant*



TEMPERATURE

Not enough variability between states

* Variables that are not ruled from the final model

INTERPRETATIONS (RANDOM EFFECTS)

```
Random effects:
 Groups   Name                 Variance  Std.Dev.  Corr
 state    (Intercept)          7.698e+12 2.775e+06
          cen_colonies_number  3.664e+03 6.053e+01 1.00
 Residual                      5.591e+12 2.365e+06
Number of obs: 1052, groups:  state, 39
```

Total Variation: 1.3271e+13

7.68e+12
7,680,000,000,00 0

State to state variability

% of Total Variation: 57%

3.664e+03
3,664

Variability in the slope of centered colonies number across states

(2.76e−8) %

5.591e+12
5,591,000,000,00 0

Unexplained variability within states

42%

INTERPRETATIONS (FIXED EFFECTS)

(Intercept)	3,075,000
cen_colonies_number	85.45
cen_stocks	-1.424
year_95	201,500
cen_stocks:year_95	0.1052

```
Fixed effects:
                    Estimate Std. Error t value
(Intercept)        3.075e+06  5.505e+05   5.586
cen_colonies_number 8.545e+01  1.141e+01   7.492
cen_stocks         -1.424e+00  8.224e-02 -17.319
year_95            2.015e+05  1.037e+04  19.436
cen_stocks:year_95 1.052e-01  5.913e-03  17.783
```

CONCLUSION

THANK YOU!

Any Questions?