Honey Production in the USA (1998-2012)

By way of **US Honey Production 1995 - 2021**

Margaret Bowers, NYCDSA
Data Analysis Project with Python
September 2024

Introduction

Purpose of This Project

Questions I Addressed

Photo by Alexander Grey, Unsplash

Exploratory Data Analysis

US Honey Production 1995 – 2021

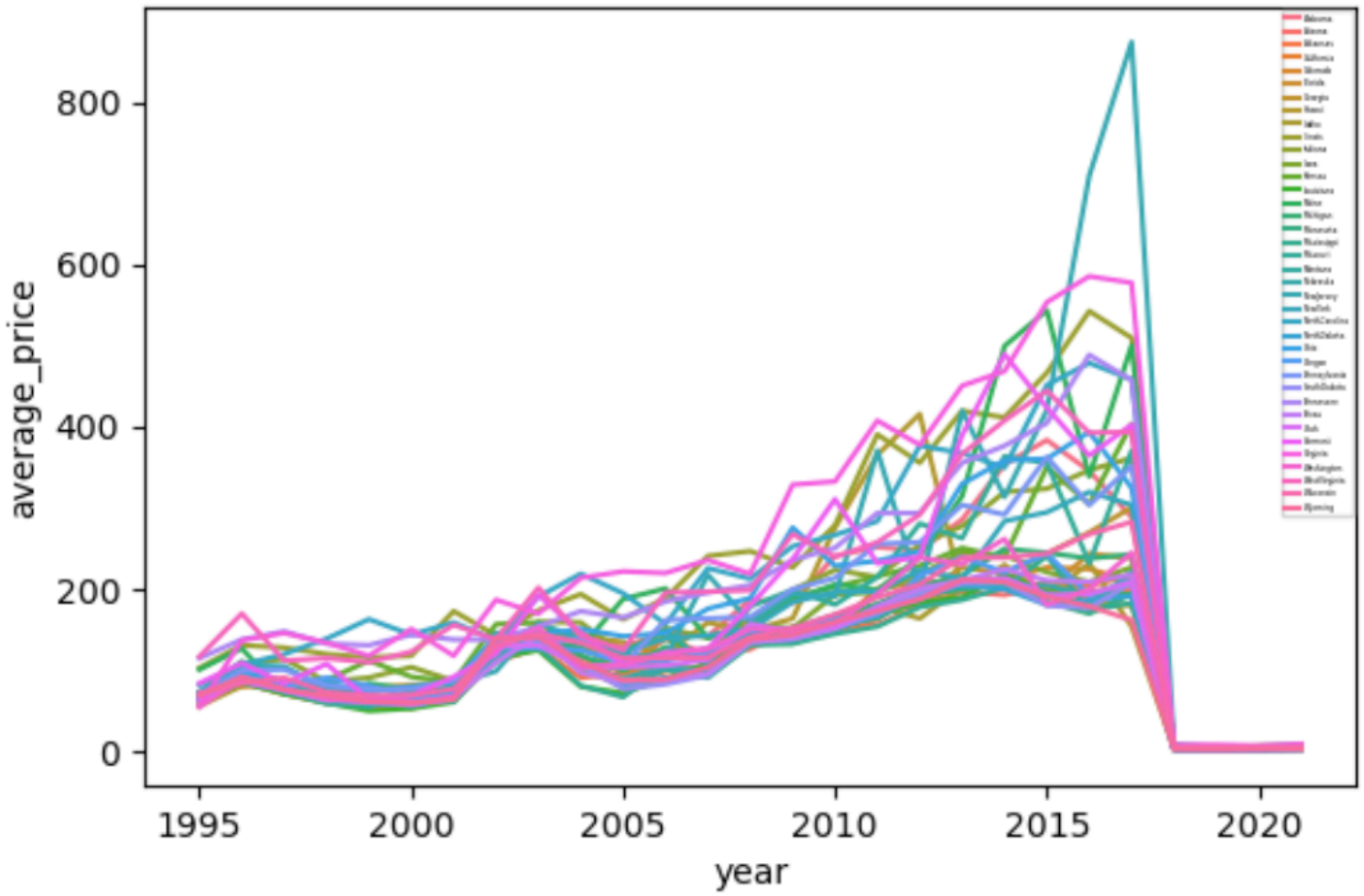
Features:

- state
- number of colonies
- yield per colony
- total production
- stocks
- average price
- value of production
- year

EDA

- No null values
- Missing data:
  - Missing 6 states
  - 6 states did not have data for every year
- Average Price of Honey
  - Unbelievable range in price: $1.30 to $874.00
  - Dollars to cents

First look at average price in 1995-2021 dataset



Exploratory Data Analysis

US Honey Production 1995 – 2021

Features:

- **state**
- **number of colonies**
- **yield per colony**
- **total production**
- **stocks**
- **average price**
- **value of production**
- **year**

EDA

- **No null values**
- **Missing data:**
  - **Missing 6 states**
  - **6 states did not have data for every year**
- **Average Price of Honey**
  - **Dollars to cents**
- **Derived columns:**
  - **Value of production != price*production**
  - **Total production != colonies*yield**

Honey Production in the USA (1998 - 2012)

Photo by Eric Ward on Unsplash

Exploratory Data Analysis

Honey Production in the USA 1998 – 2012

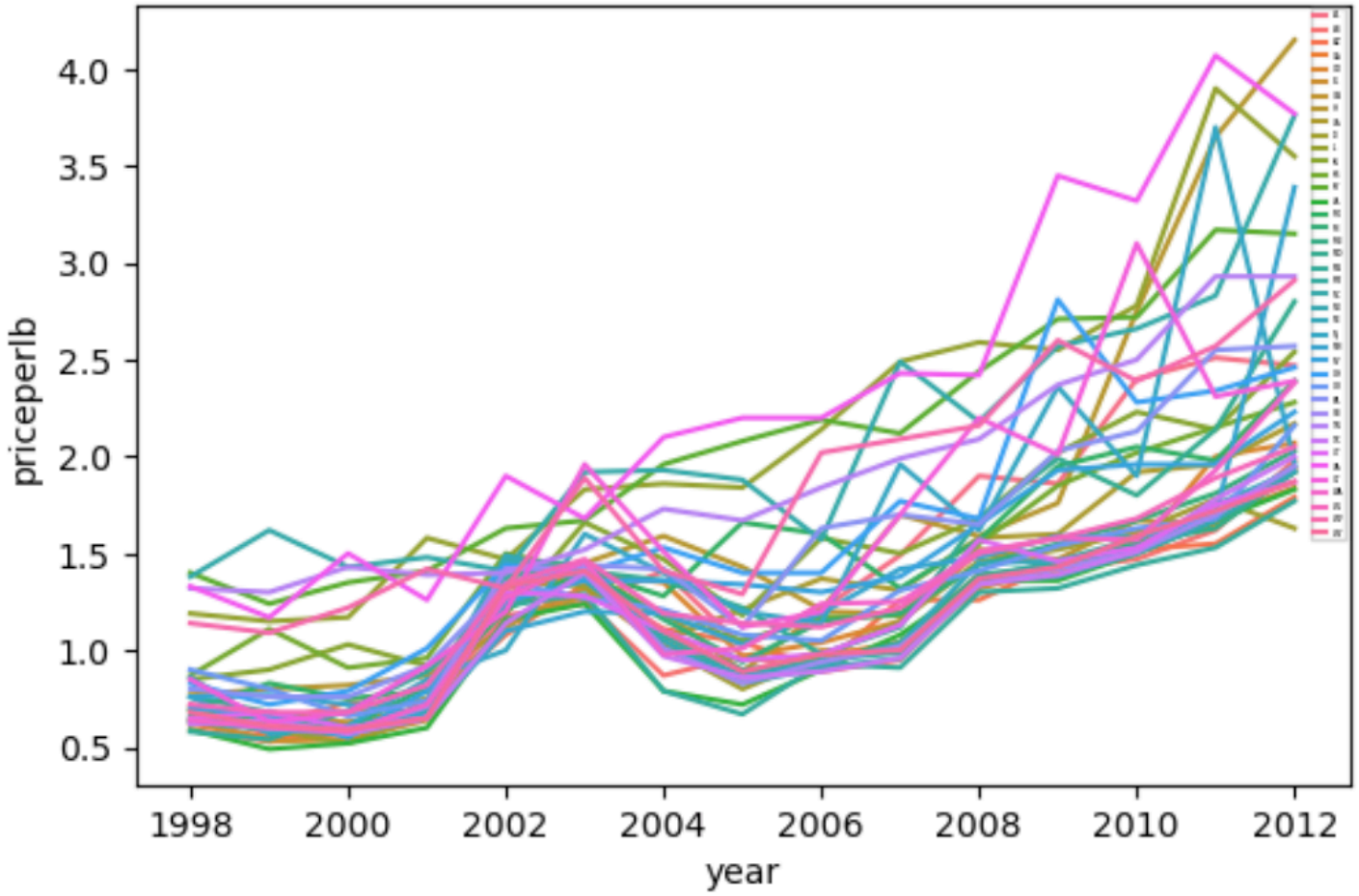
Features:

- state
- number of colonies
- yield per colony
- total production
- stocks
- average price
- value of production
- year

EDA

- No null values
- Missing data:
  - Missing 6 states
  - 4 states did not have data for every year
- Average Price of Honey
  - Consistent throughout
  - Believable price range:
    $0.49 to $4.15

Average price per pound from 1998 - 2012 dataset



Exploratory Data Analysis

Honey Production in the USA 1998 – 2012

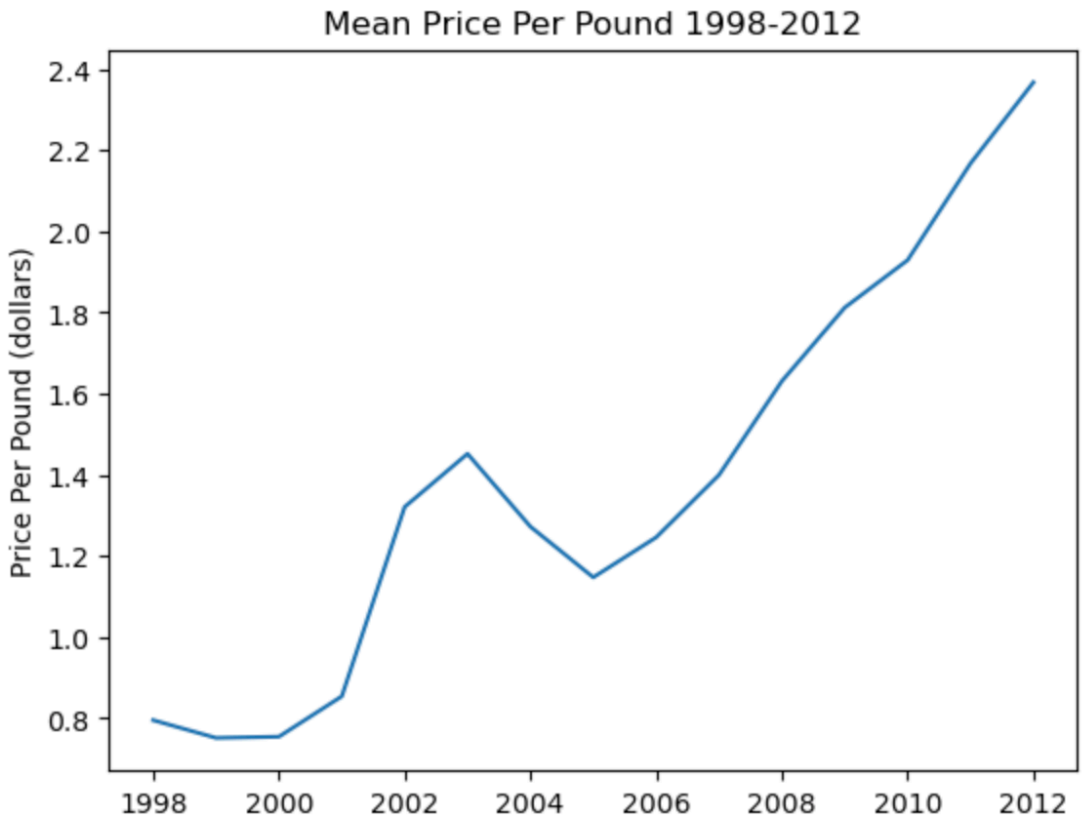
Features:

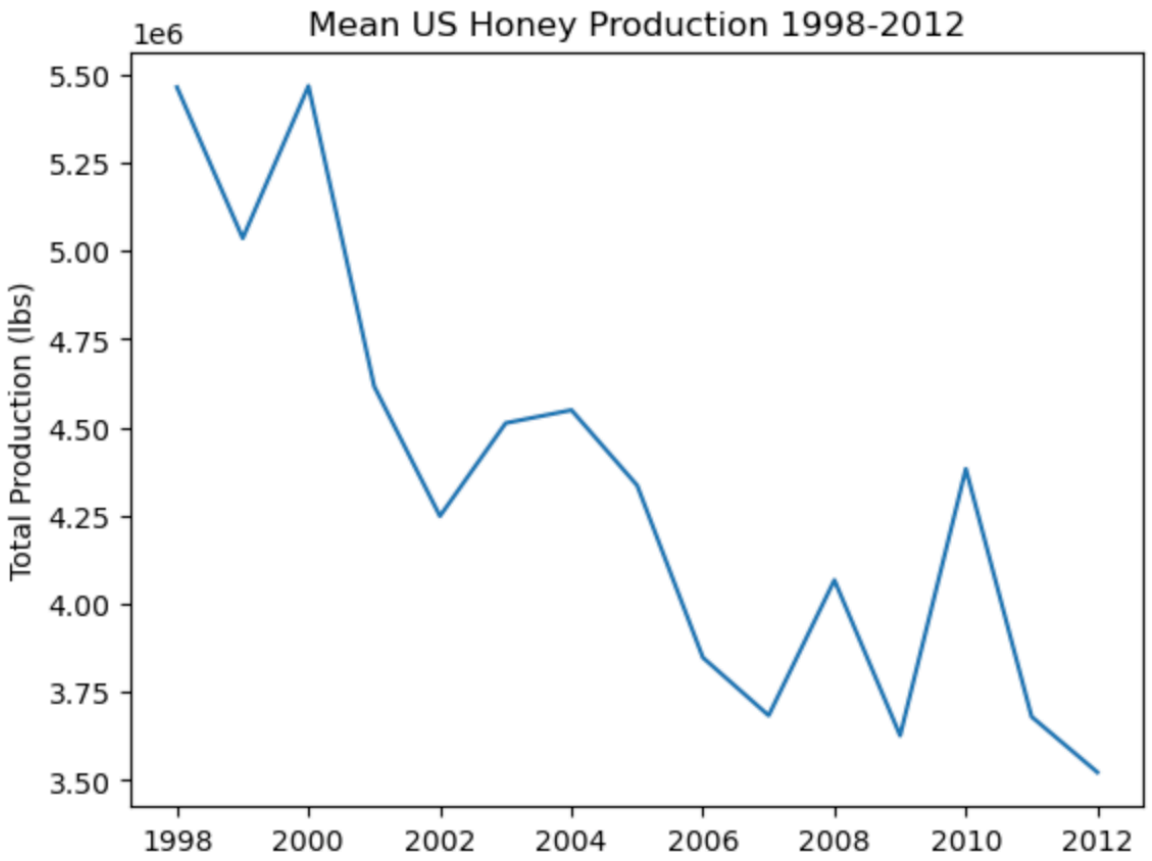
- **state**
- **number of colonies**
- **yield per colony**
- **total production**
- **stocks**
- **average price**
- **value of production**
- **year**

EDA

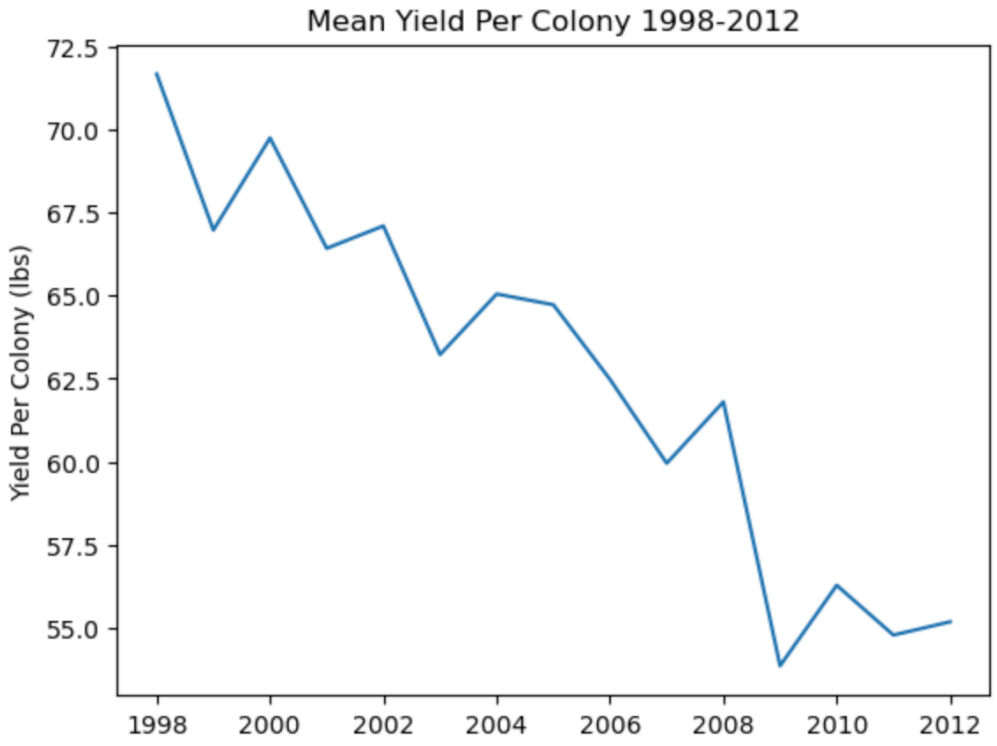
- **No null values**
- **Missing data:**
  - **Missing 6 states**
  - **4 states did not have data for every year**
- **Average Price of Honey**
  - **Consistent throughout**
- **Derived columns:**
  - **Total production = colonies*yield**
  - **Value of production != price*production**

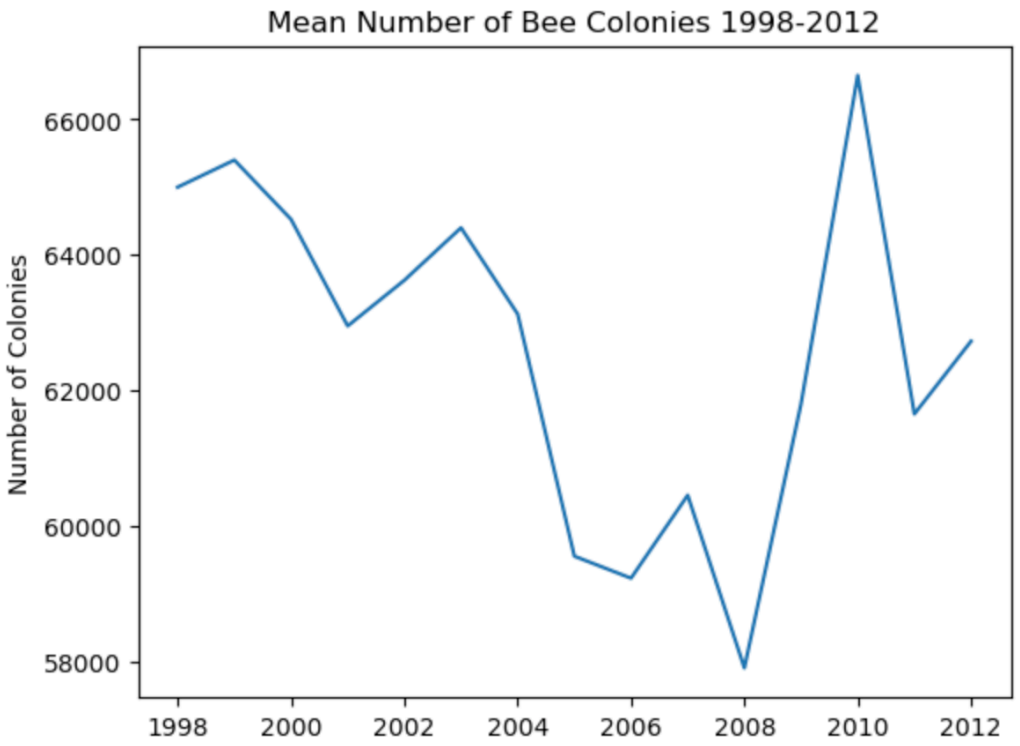
First look at the data:

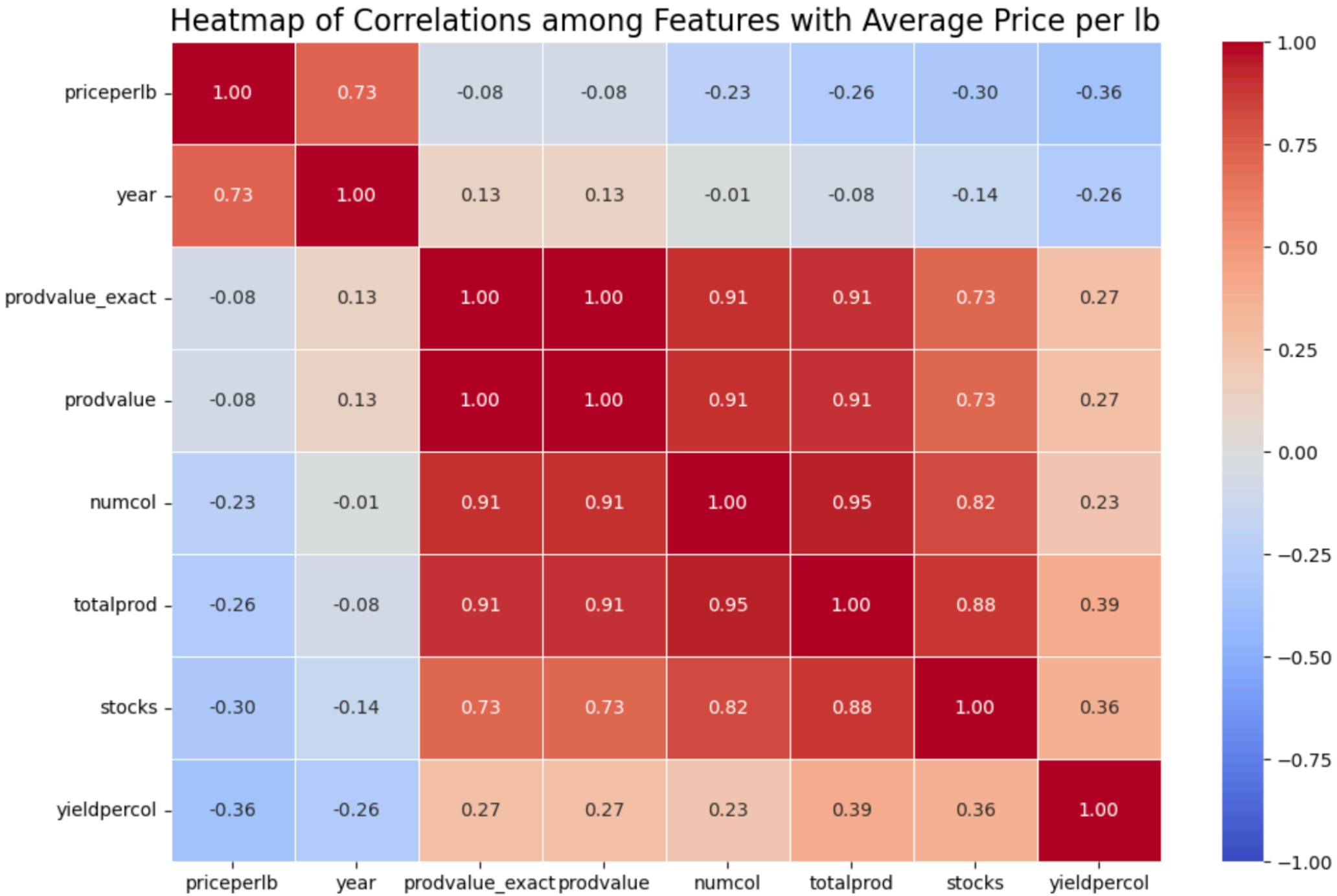




First look at the data:






Heatmap of Correlations among Features with Average Price per lb
priceperlb
year
prodvalue_exact
prodvalue
numcol
totalprod
stocks
yieldpercol
1.00 0.73 -0.08 -0.08 -0.23 -0.26 -0.30 -0.36
0.73 1.00 0.13 0.13 -0.01 -0.08 -0.14 -0.26
-0.08 0.13 1.00 1.00 0.91 0.91 0.73 0.27
-0.08 0.13 1.00 1.00 0.91 0.91 0.73 0.27
-0.23 -0.01 0.91 0.91 1.00 0.95 0.82 0.23
-0.26 -0.08 0.91 0.91 0.95 1.00 0.88 0.39
-0.30 -0.14 0.73 0.73 0.82 0.88 1.00 0.36
-0.36 -0.26 0.27 0.27 0.23 0.39 0.36 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Price compared to:

yield per colony (blue), number of colonies (red), total production (green), stocks (purple)

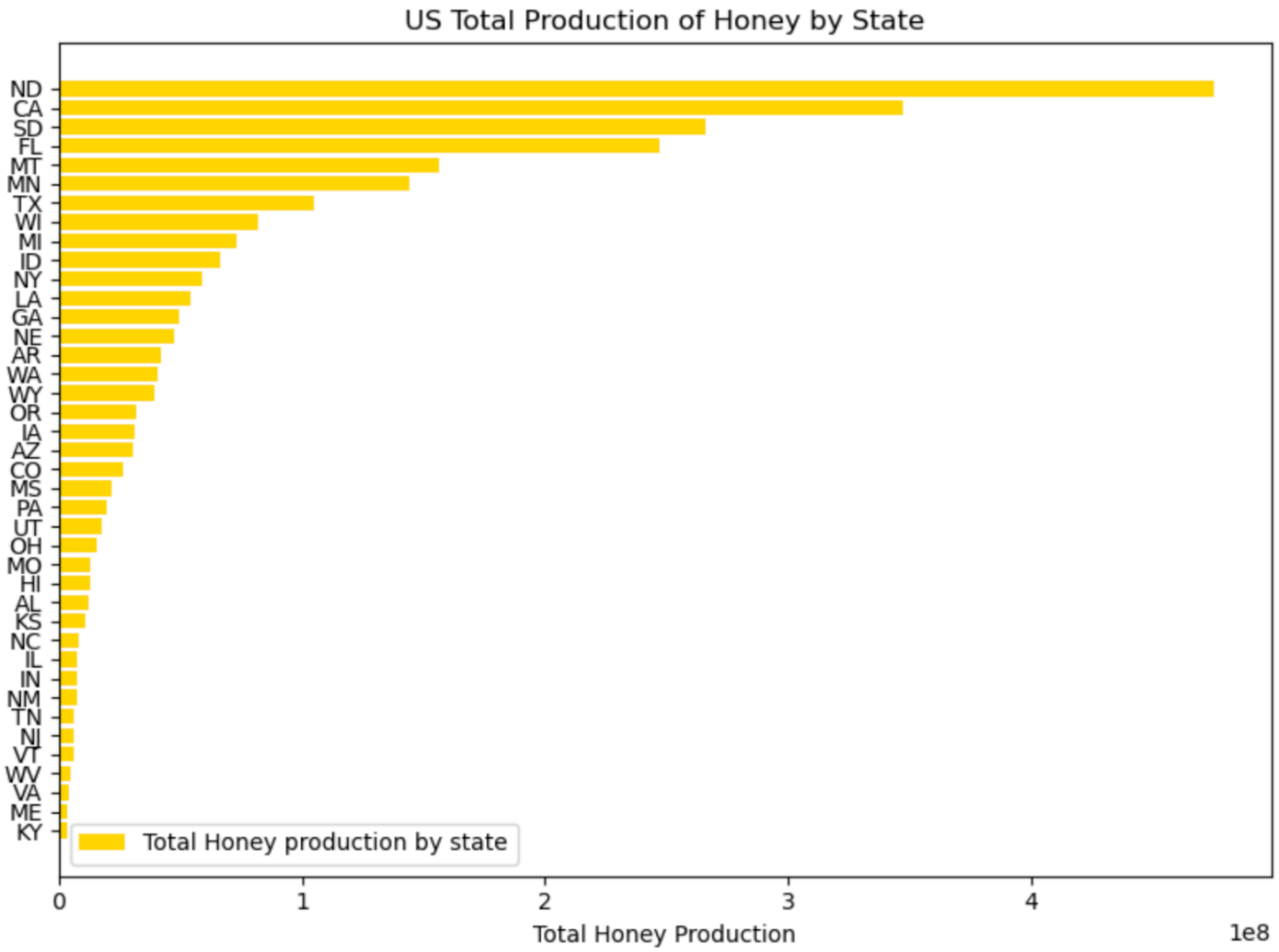
Individual States

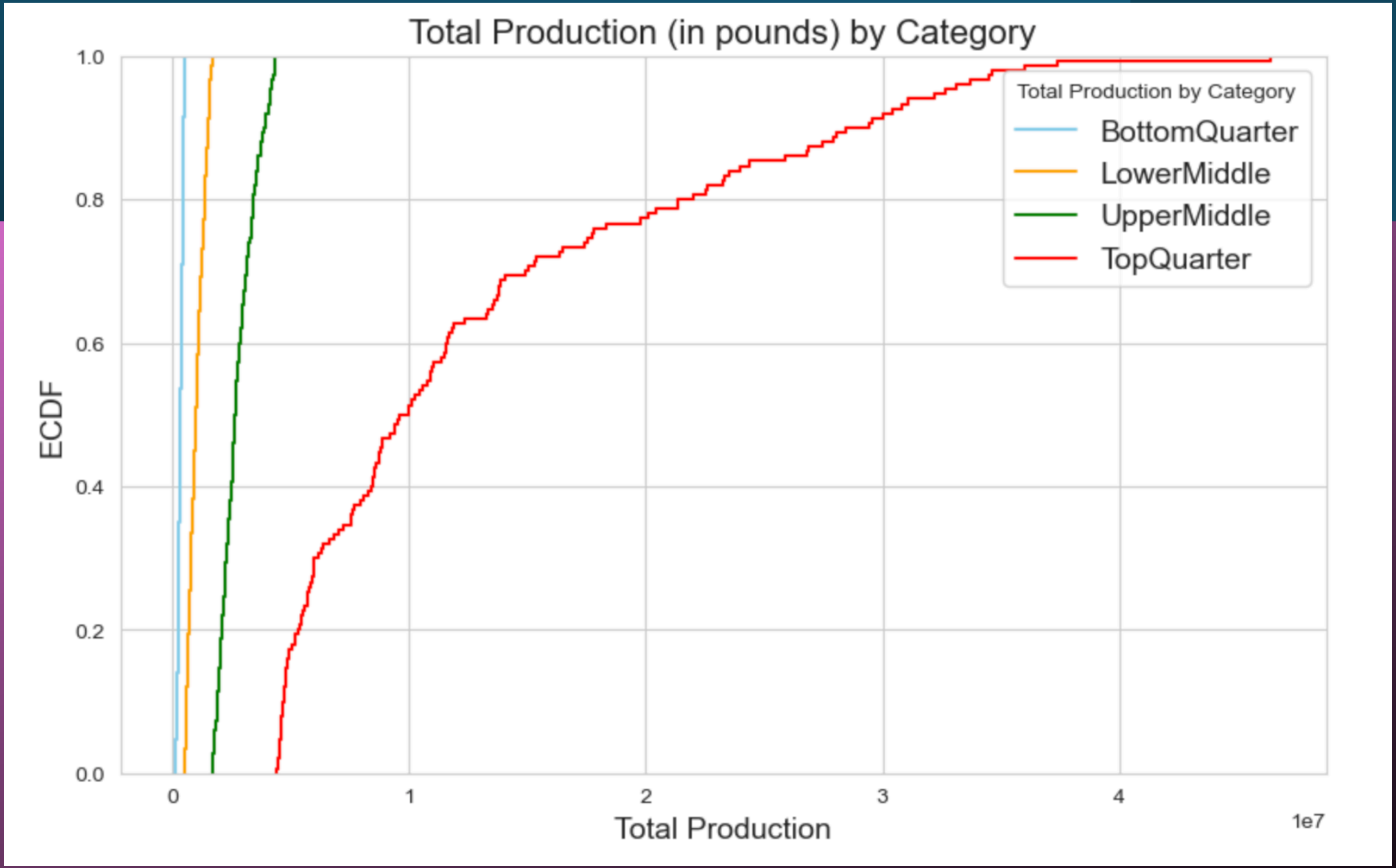
Grouping the data by top performers

Quantiles of total production

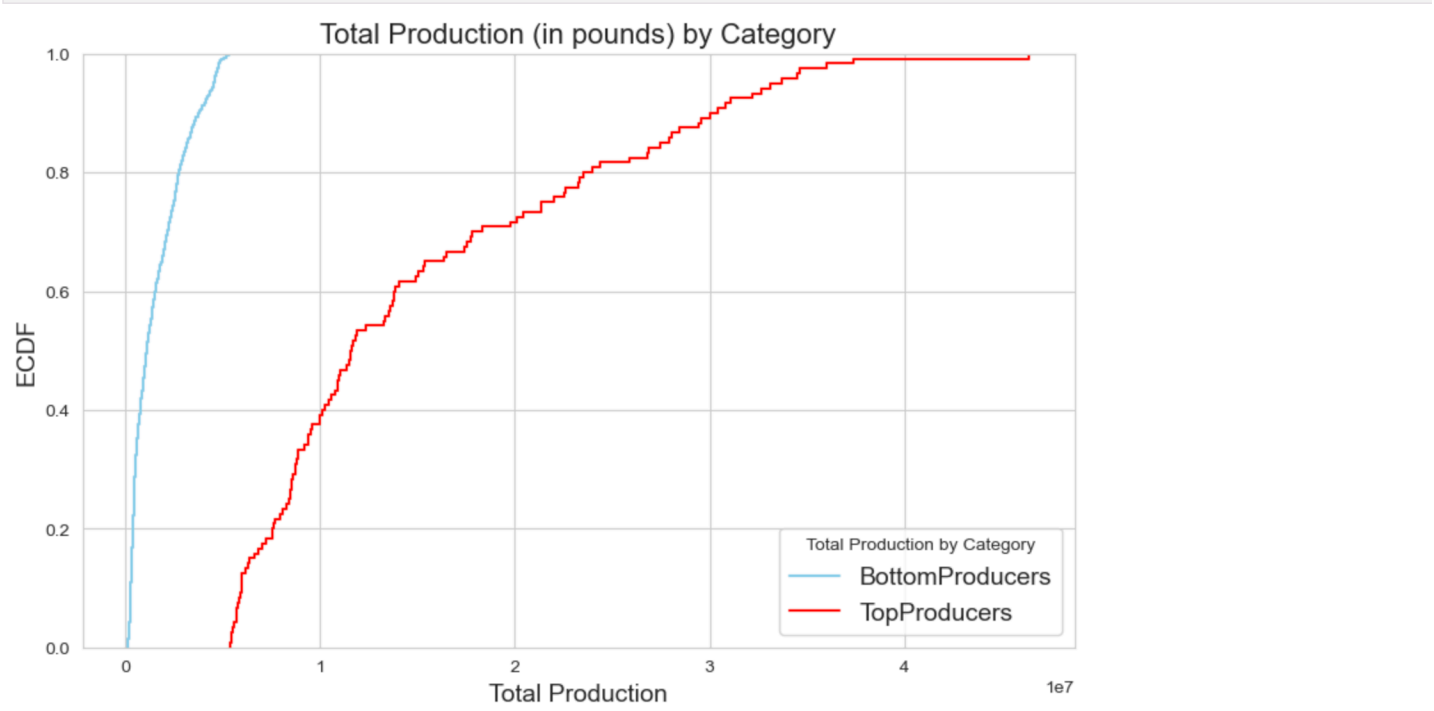
state	numcol	priceperlb	prodvalue	totalprod	yieldpercol
ND	370333.333333	1.125333	3.637613e+07	3.167233e+07	88.066667
CA	406666.666667	1.133333	2.435933e+07	2.316900e+07	55.800000
SD	236266.666667	1.140000	1.937673e+07	1.774273e+07	75.800000
FL	195200.000000	1.100667	1.661527e+07	1.646987e+07	83.066667
MT	136000.000000	1.150667	1.179353e+07	1.043747e+07	77.333333
MN	128933.333333	1.152000	1.052333e+07	9.624000e+06	74.266667
TX	98533.333333	1.114667	7.429200e+06	6.993600e+06	70.533333
WI	70727.272727	1.080000	6.325364e+06	6.020545e+06	85.000000
ID	105375.000000	0.958750	5.047375e+06	5.299750e+06	50.375000
MI	72000.000000	1.092000	5.573600e+06	5.226600e+06	72.500000
NY	63800.000000	0.904000	4.658200e+06	5.074800e+06	80.000000
LA	44666.666667	0.560000	2.742667e+06	4.931000e+06	110.333333
AR	51333.333333	0.583333	2.798333e+06	4.788000e+06	93.333333
GA	64000.000000	1.360000	6.304500e+06	4.638500e+06	72.500000
NE	57666.666667	0.826667	3.720333e+06	4.495000e+06	78.666667

Total Honey Production by State




Total Production (in pounds) by Category
Total Production by Category
BottomQuarter
LowerMiddle
UpperMiddle
TopQuarter
ECDF
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
1e7
Total Production

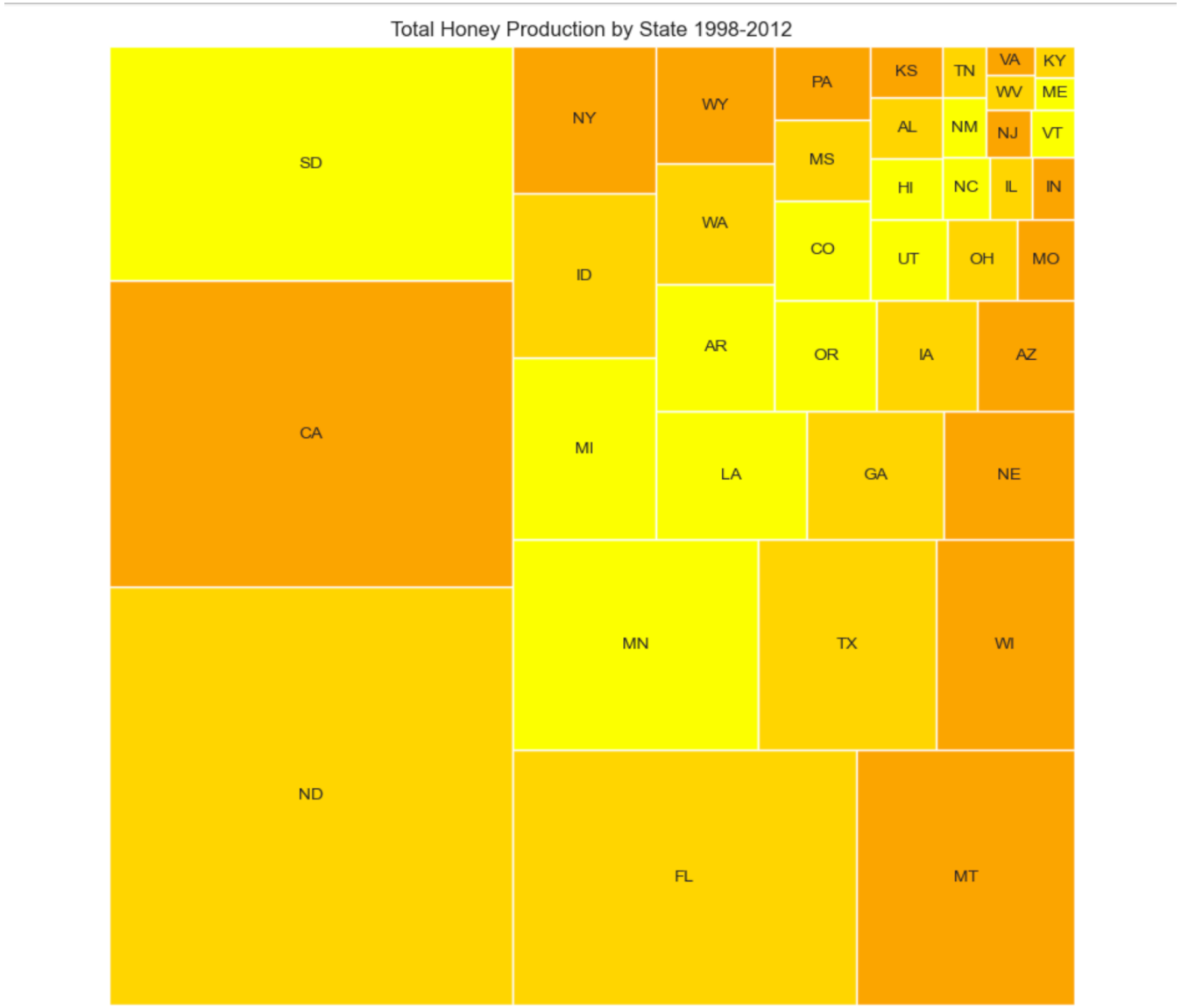
Producers: bottom 80% and top 20%



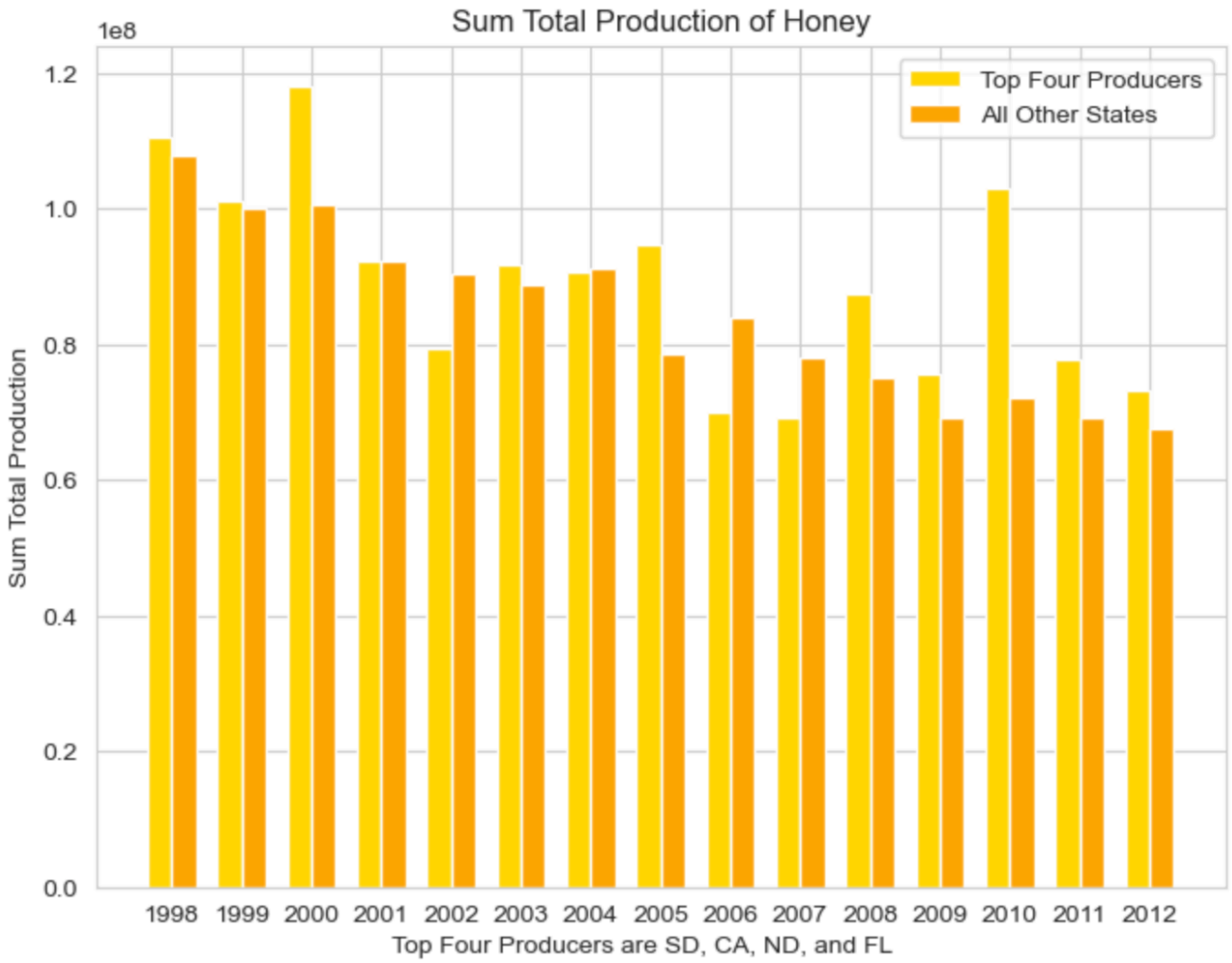
Honey Production by state:

4 states were responsible for over 50% of the total US production of honey:

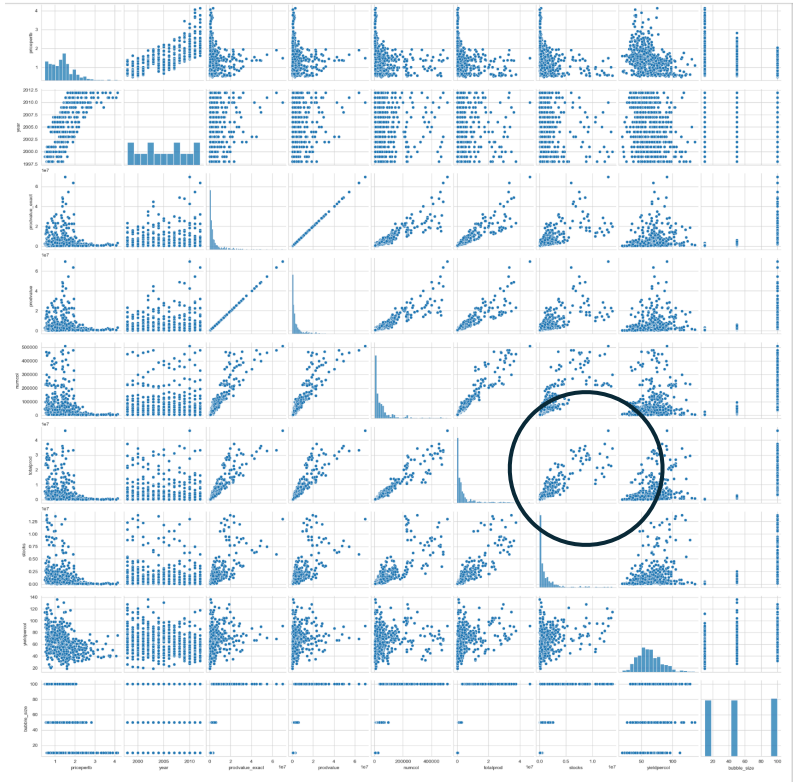
North Dakota 18%
California 13%
South Dakota 10%
Florida 9.5%

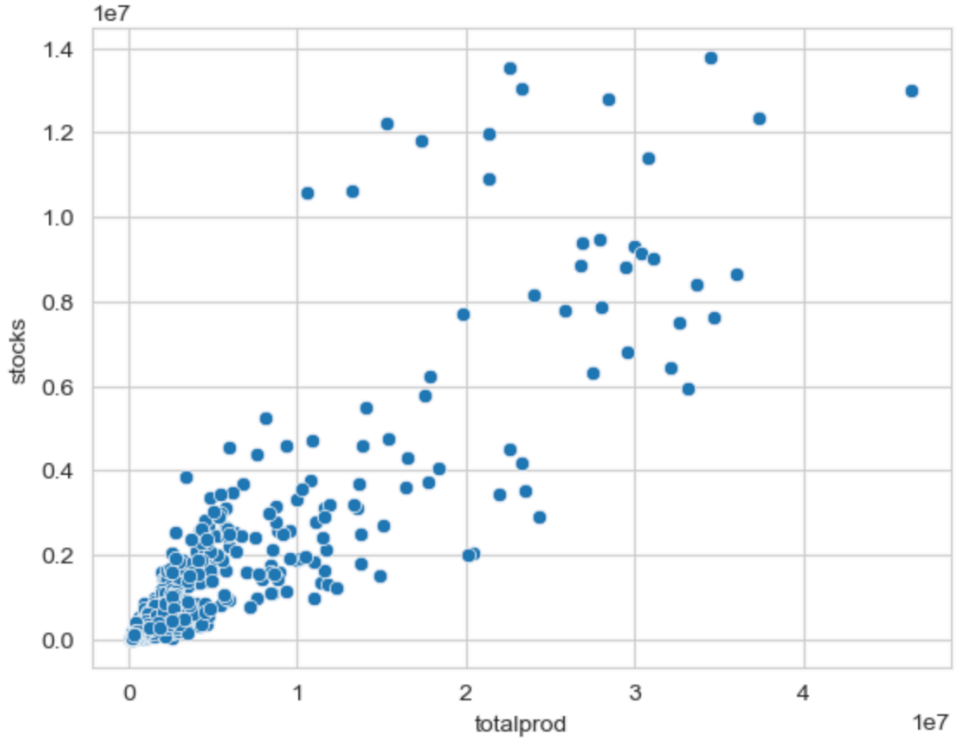


Top 4 producers compared to 36 states:

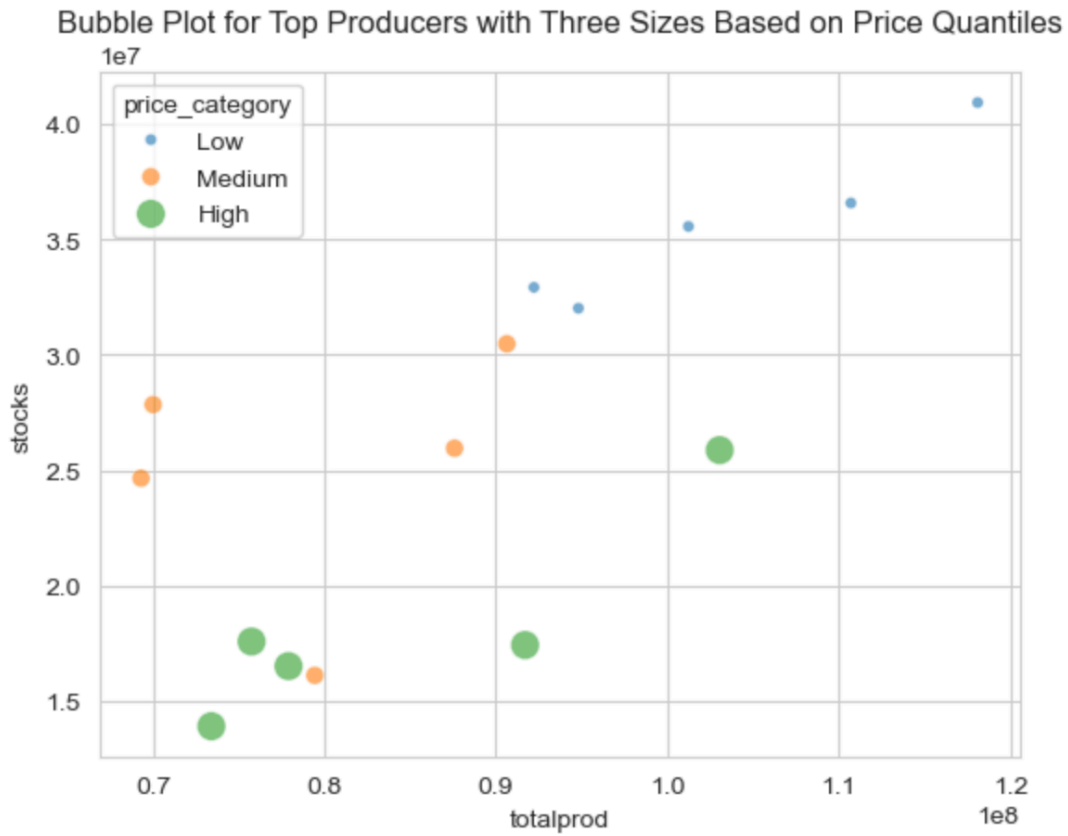


Stocks and production in the top and bottom producers

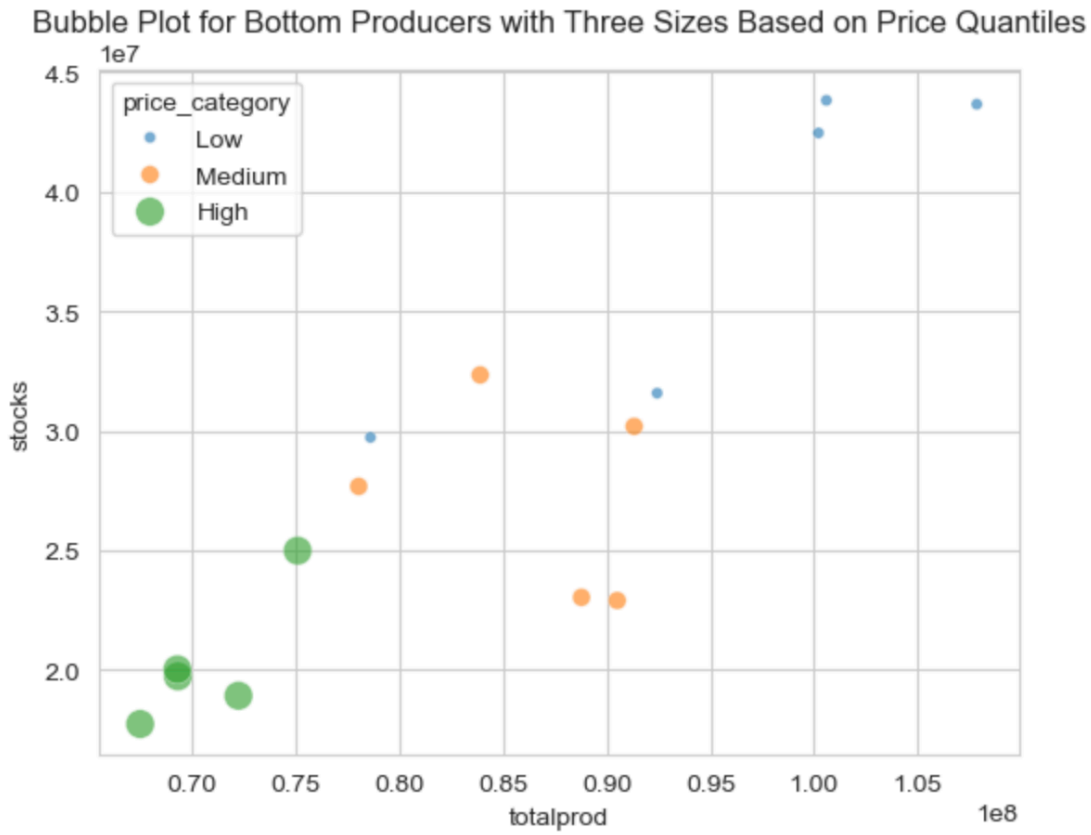




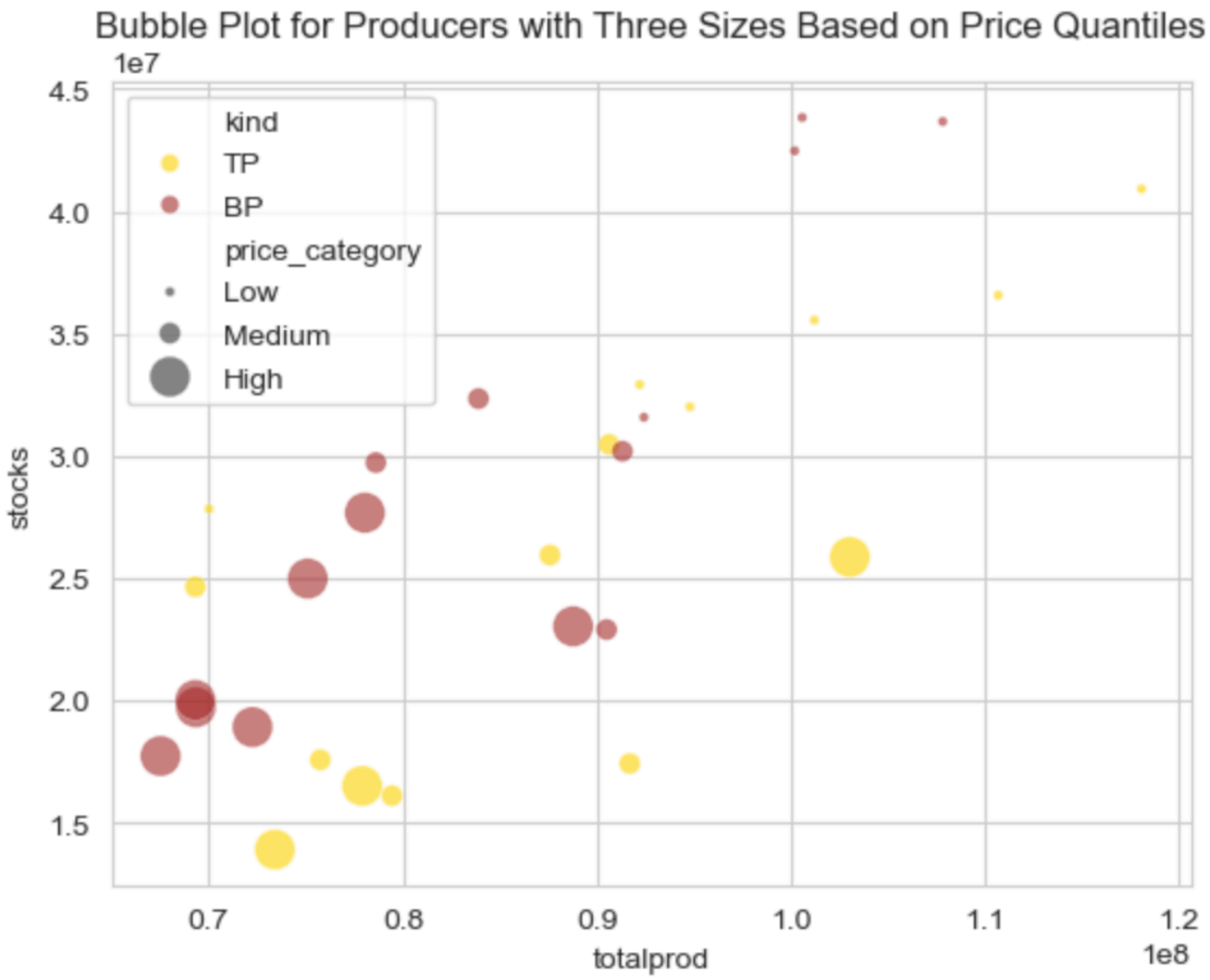
Top Producers



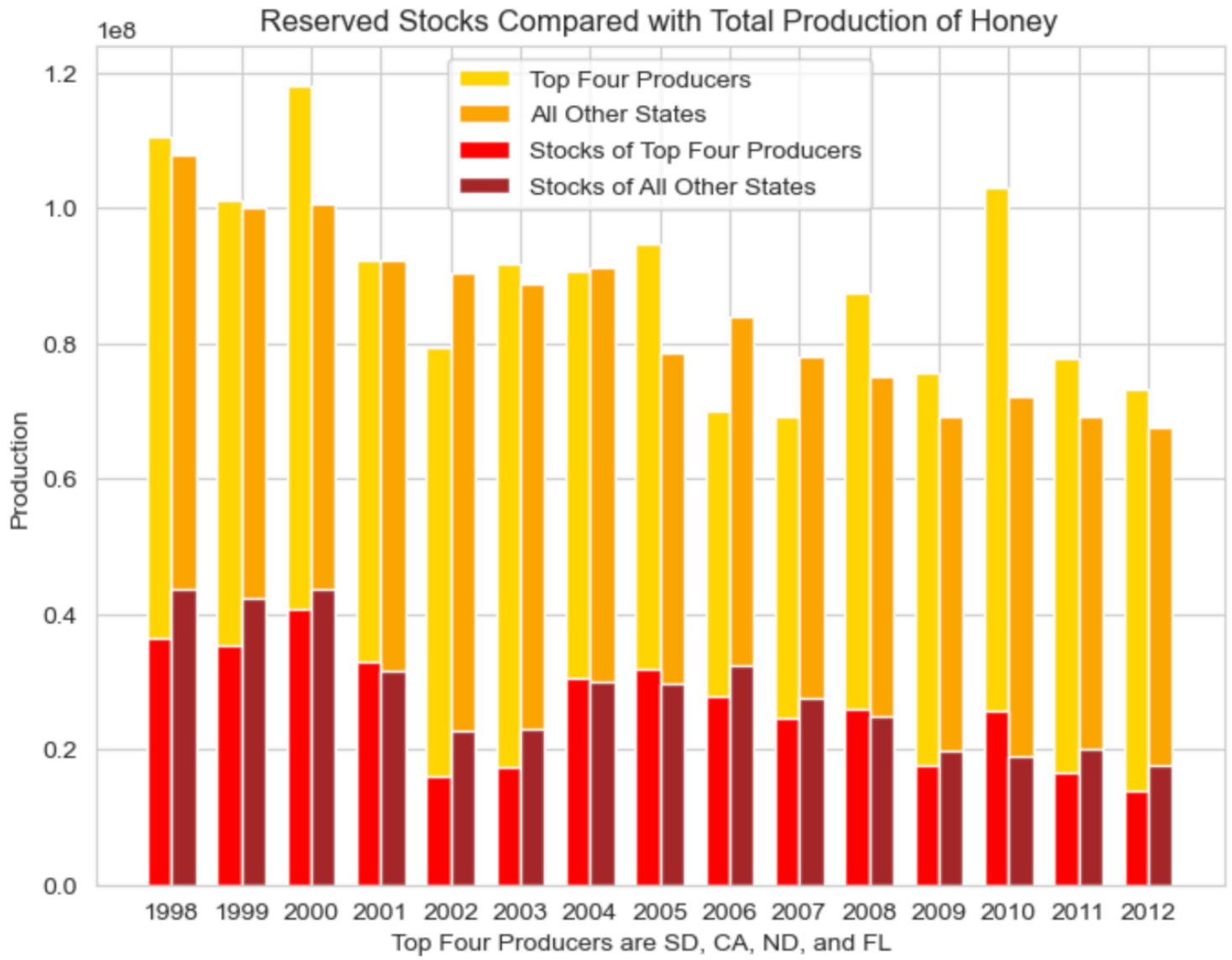
Bottom Producers



Top and Bottom Producers



Stocks and Total Production for Top and Bottom Producers



Conclusions:

Four states are responsible for over 50% of US honey production:

ND, CA, SD, FL

States with lowest levels of production:

ME, KY, VA

Highest average prices of honey:

VA, IL, KY

Lowest average prices of honey:

LA, MS, AR

Further Research:

Relationships between crops and honey production

Photo by Pixabay: https://www.pexels.com/photo/brown-and-black-bee-on-yellow-flower-nectar-460961/

Special Thanks To My Mentors:

Vinod Chugani
Philippe Heitzmann

Photo by ROMAN ODINTSOV: https://www.pexels.com/photo/spoon-of-honey-6422029/

Questions

Photo by Alexander Grey, Unsplash