

ECON 0150 | Economic Data Analysis

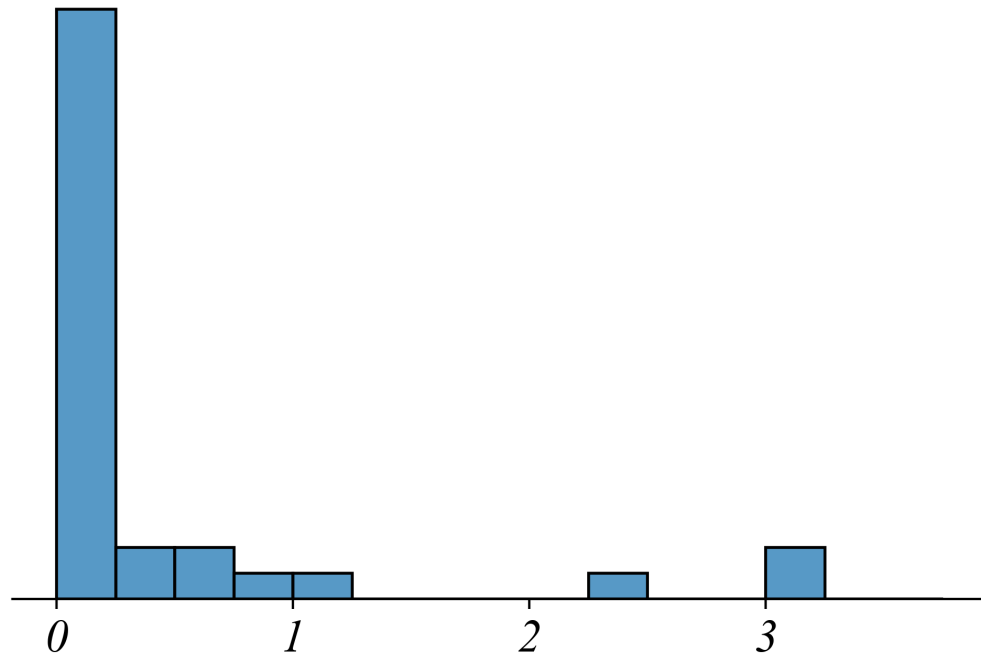
The economist's data analysis skillset.

Part 1.5 | Bivariate Relationships

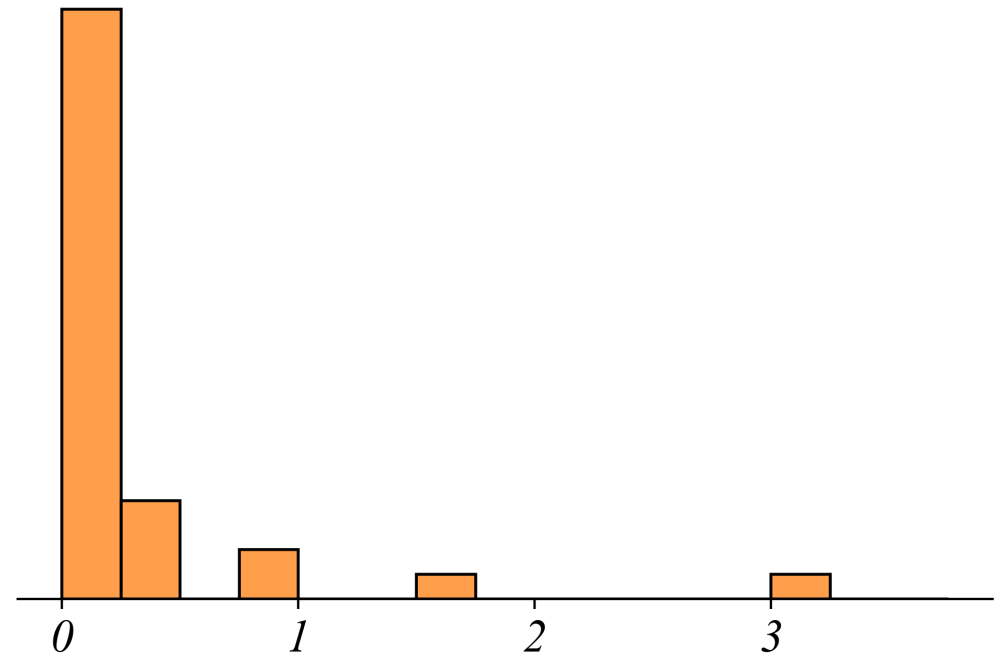
Cross-Sectional Relationships

Q. Is there a relationship between GDP and coffee production?

Gross Domestic Product (GDP)



Coffee Production (tonnes)

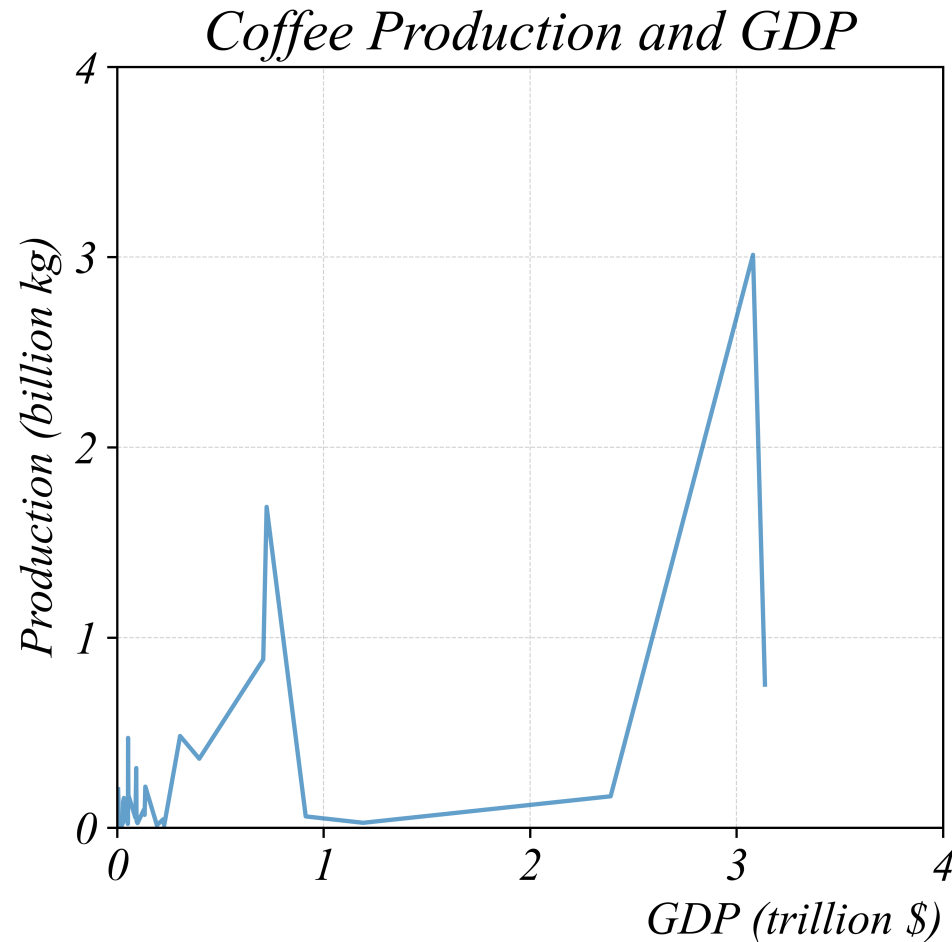


> *maybe, but it's hard to see*

> *lets use a two dimensional graph*

Cross-Sectional Relationships

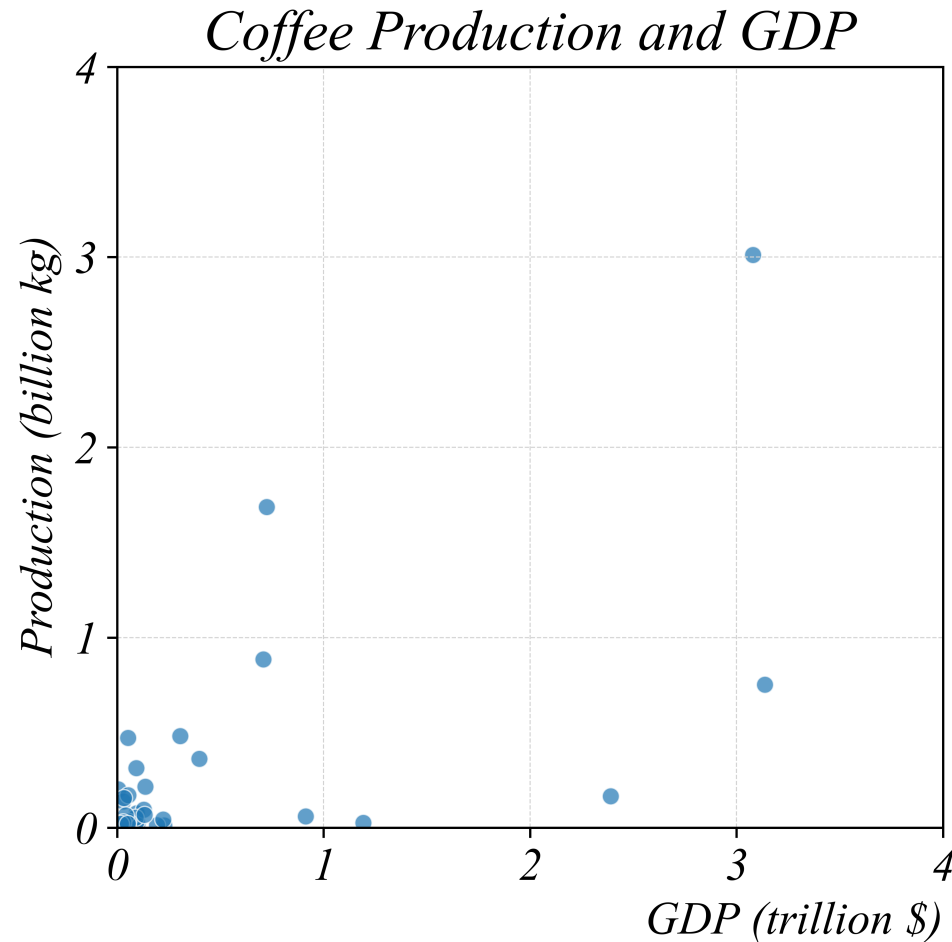
Q. Is there a relationship between GDP and coffee production?



> two dimensions is nice, but the points have no meaningful relationships

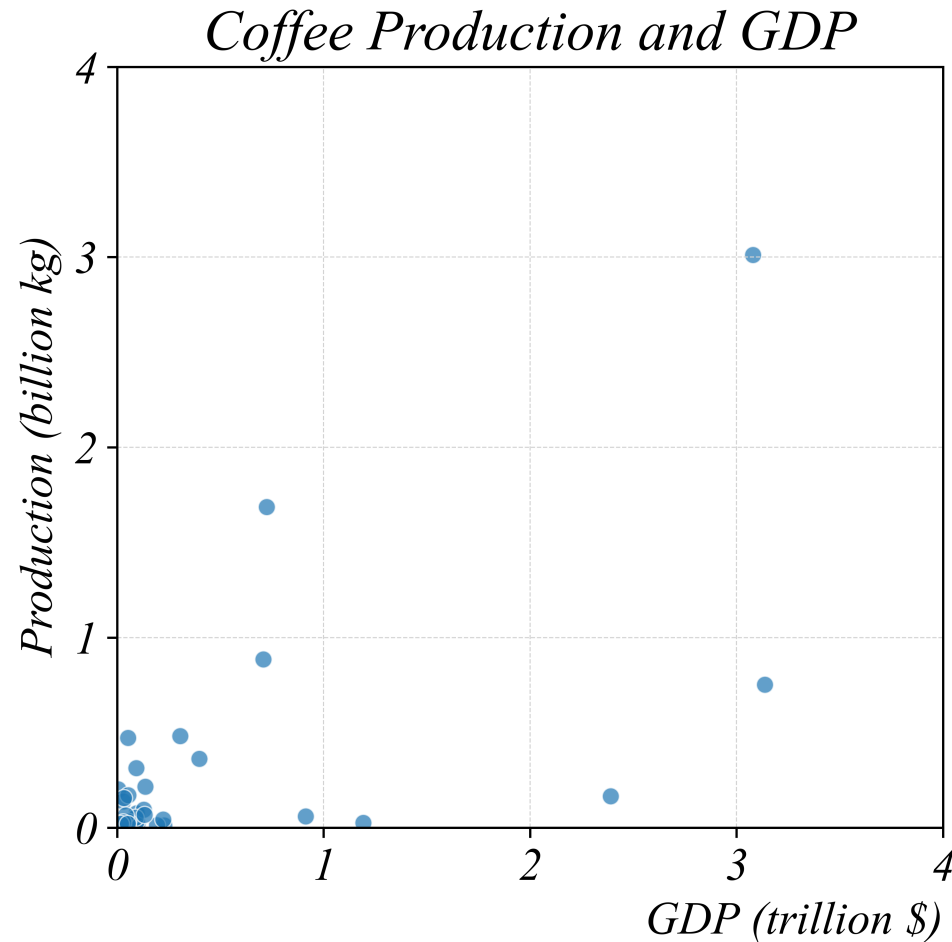
Cross-Sectional Relationships

Q. Is there a relationship between GDP and coffee production?



Cross-Sectional Relationships

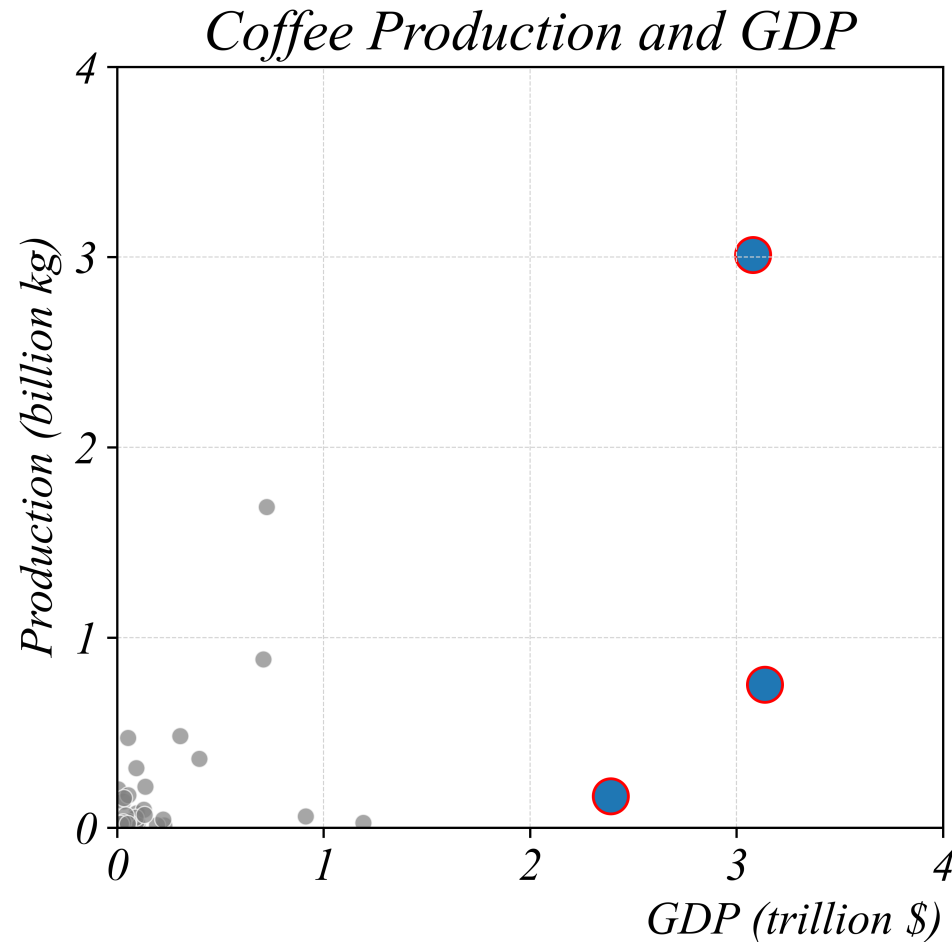
Which countries have a GDP above \$2 trillion?



> *look at the horizontal axis and select all that are greater than 2*

Cross-Sectional Relationships

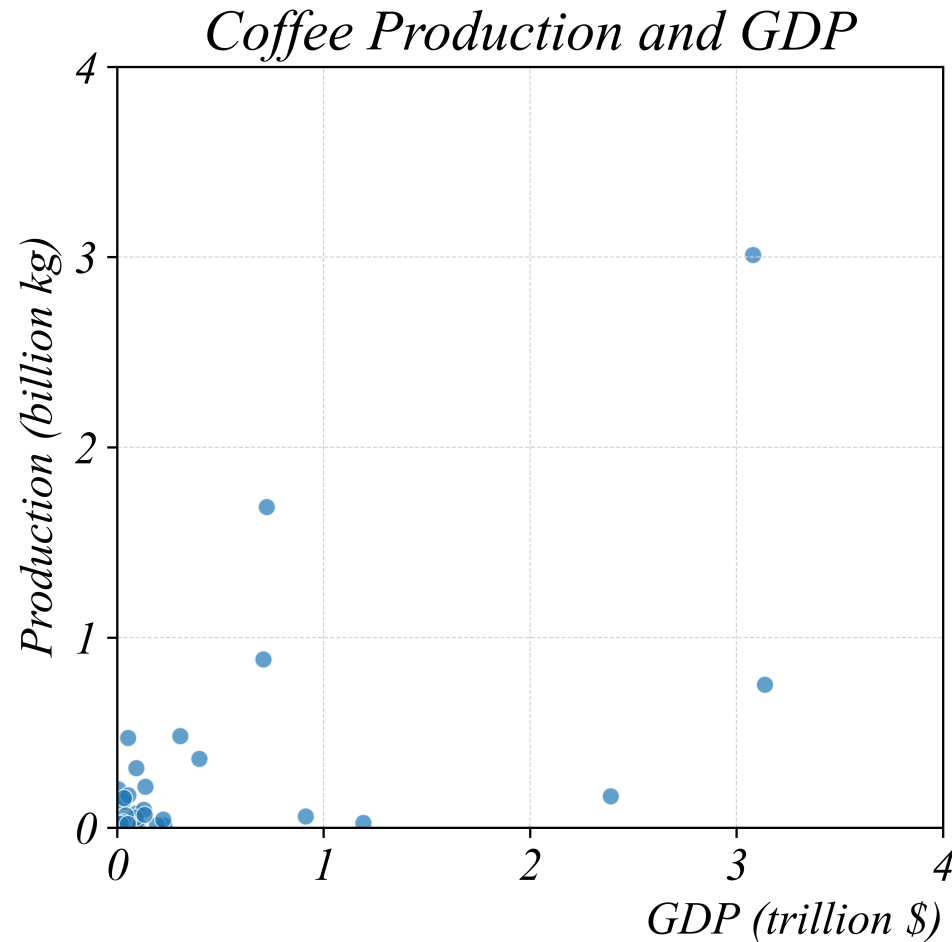
Which countries have a GDP above \$2 trillion?



> *look at the horizontal axis and select all that are greater than 2*

Cross-Sectional Relationships

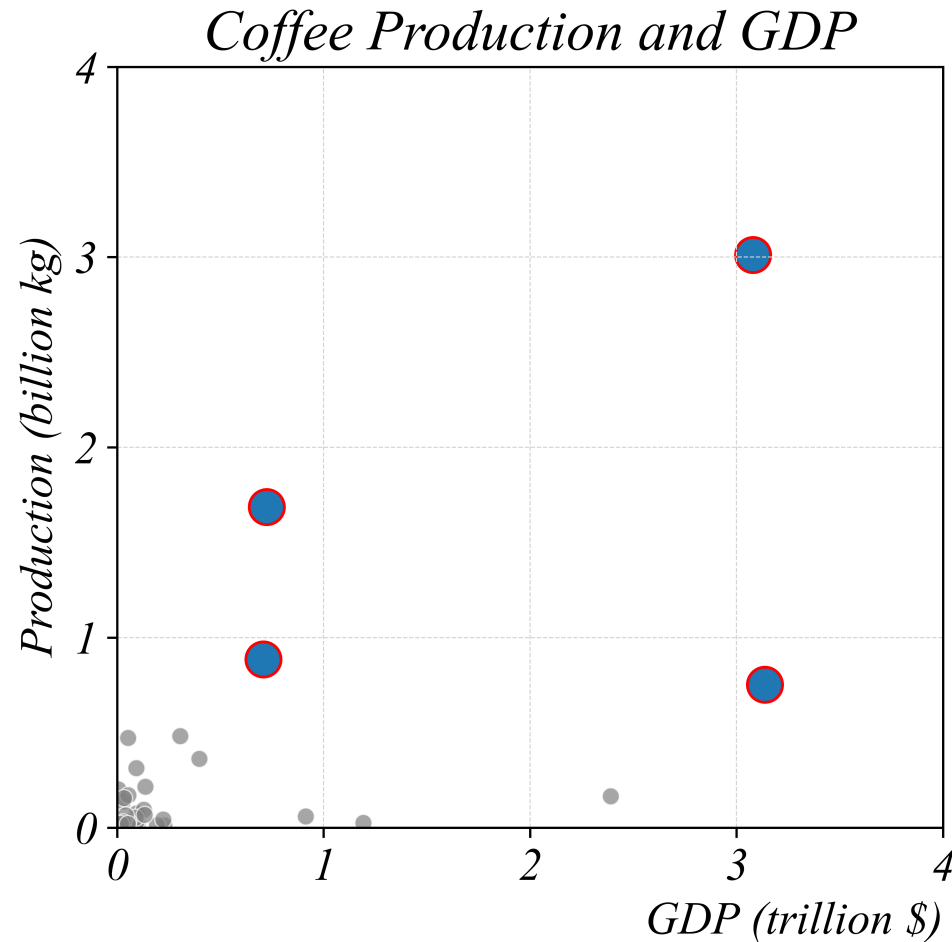
Which countries have a production above $\frac{1}{2}$ billion kg?



> *and we can use either axis*

Cross-Sectional Relationships

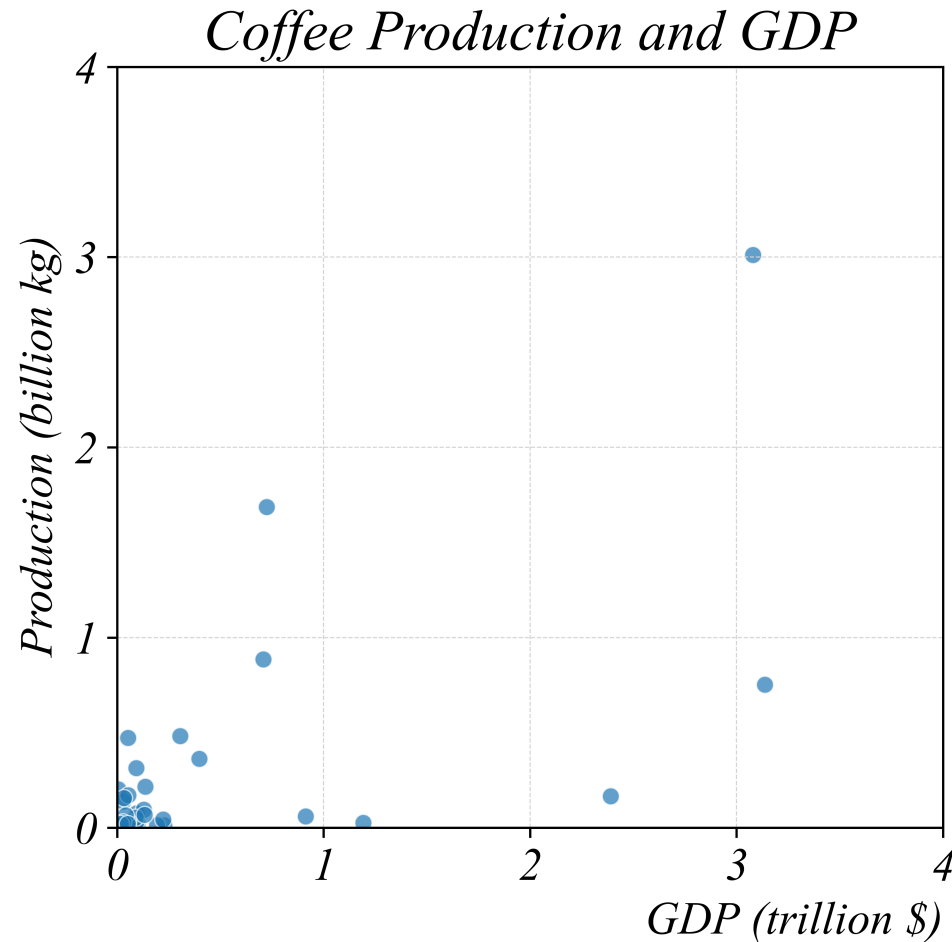
Which countries have a production above ½ billion kg?



> *and we can use either axis*

Cross-Sectional Relationships

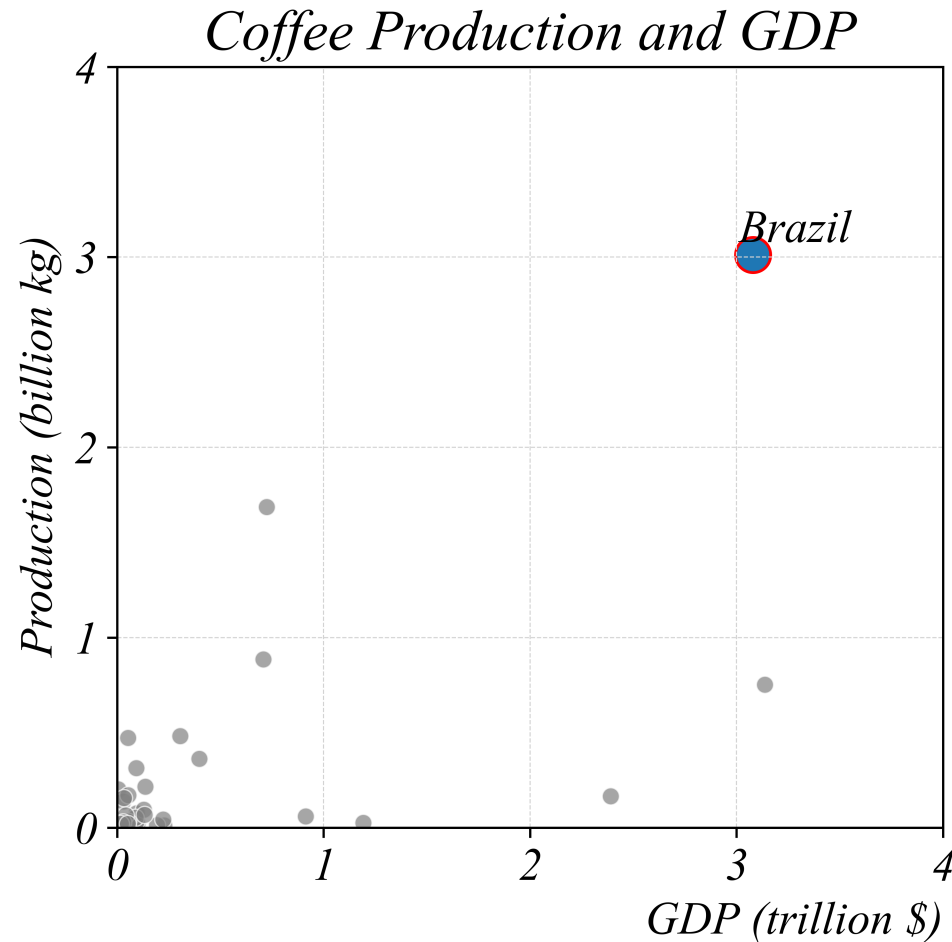
Which countries produce less coffee per dollar than Brazil?



> we can also compare *BETWEEN* data points

Cross-Sectional Relationships

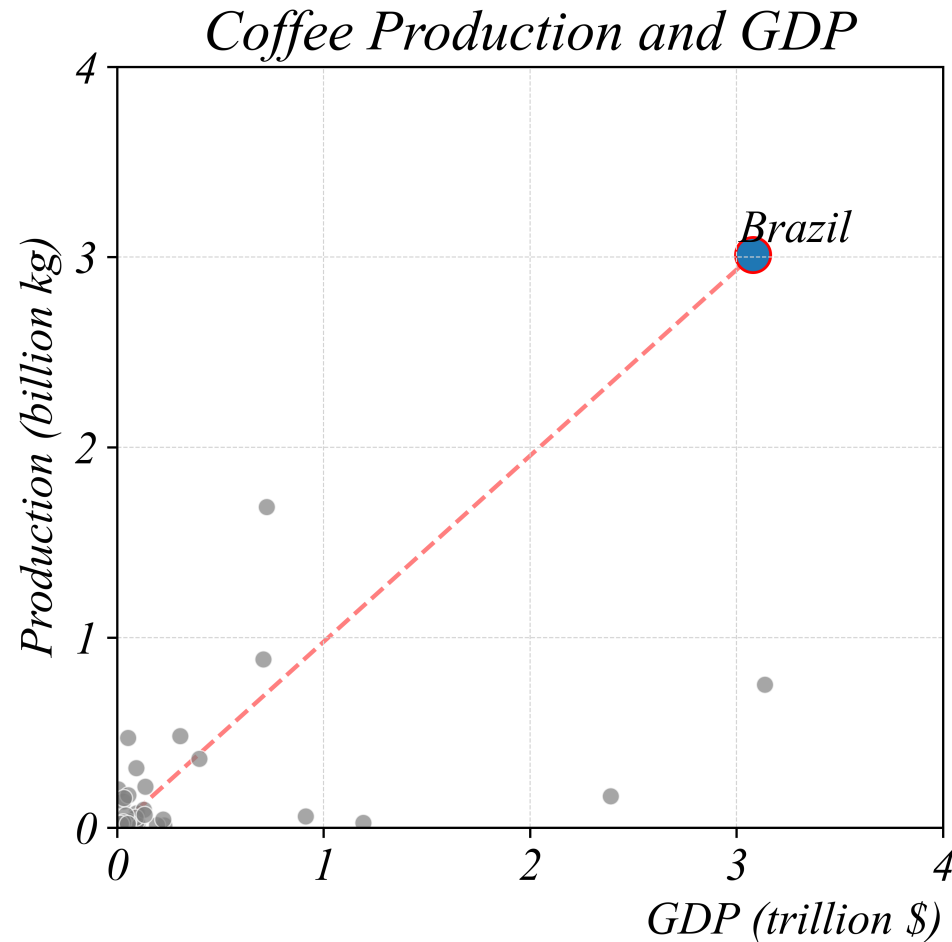
Which countries produce less coffee per dollar than Brazil?



> we can also compare *BETWEEN* data points

Cross-Sectional Relationships

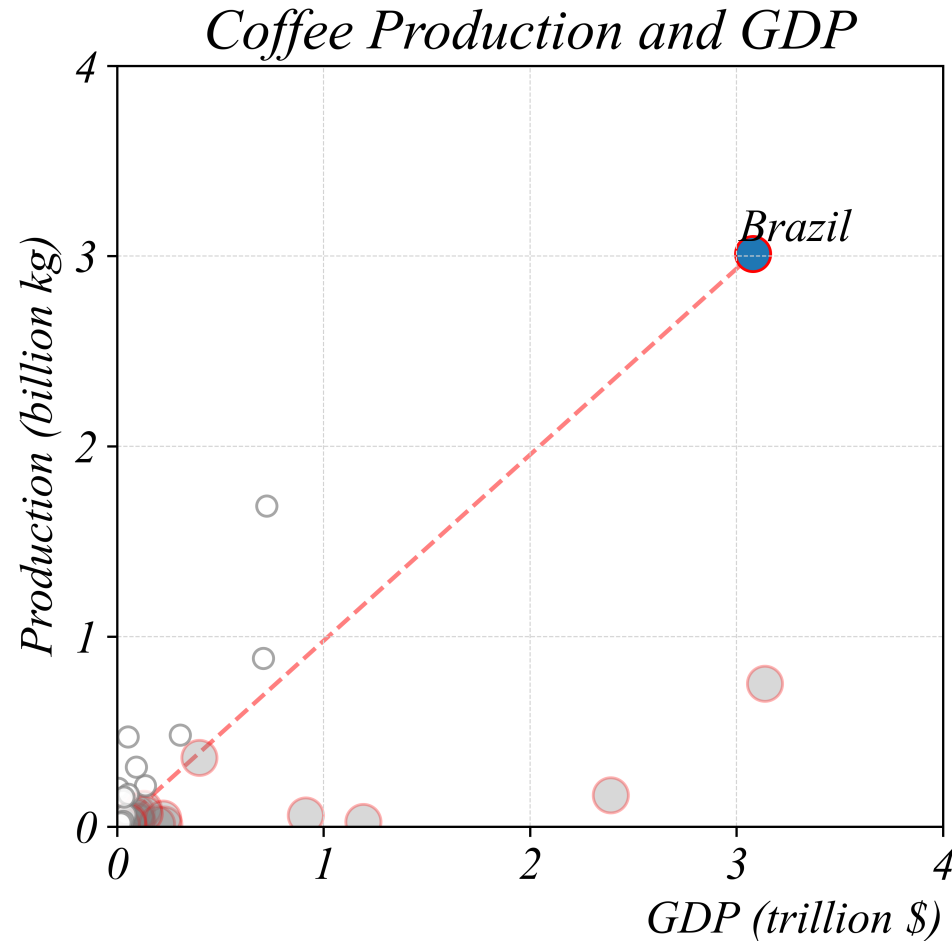
Which countries produce less coffee per dollar than Brazil?



> *separating lines can help make comparisons between ratios*

Cross-Sectional Relationships

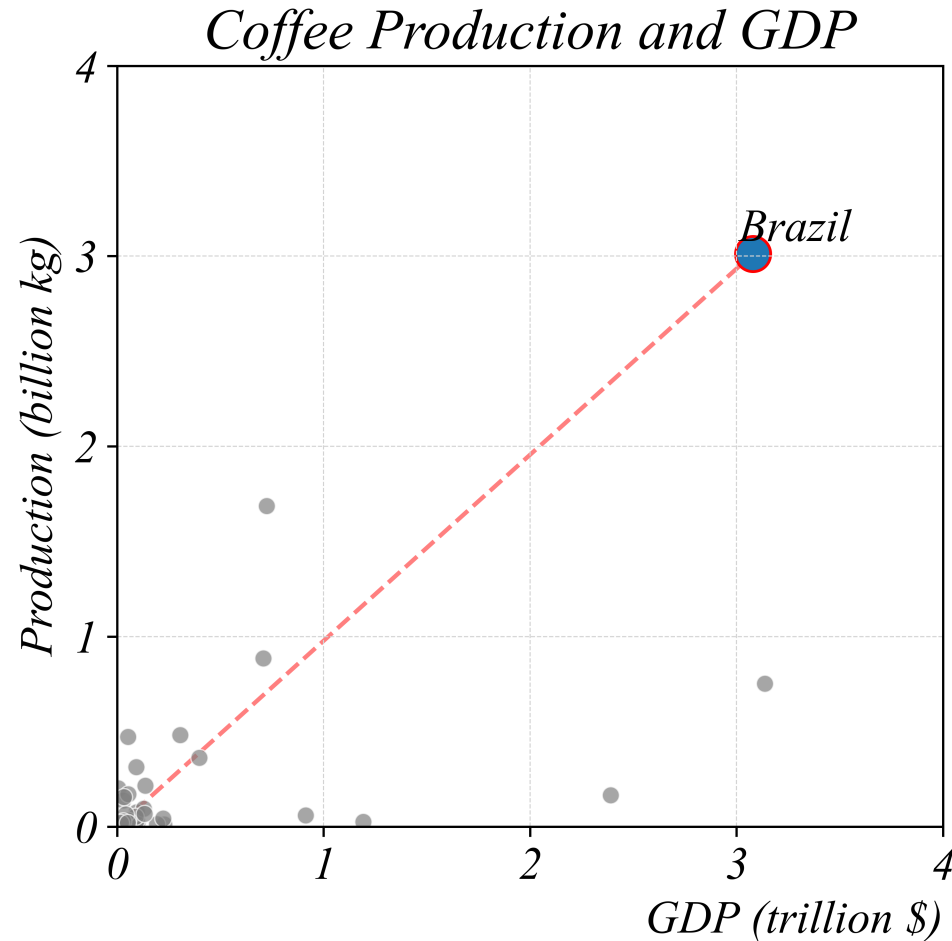
Which countries produce less coffee per dollar than Brazil?



> *separating lines can help make comparisons between ratios*

Cross-Sectional Relationships

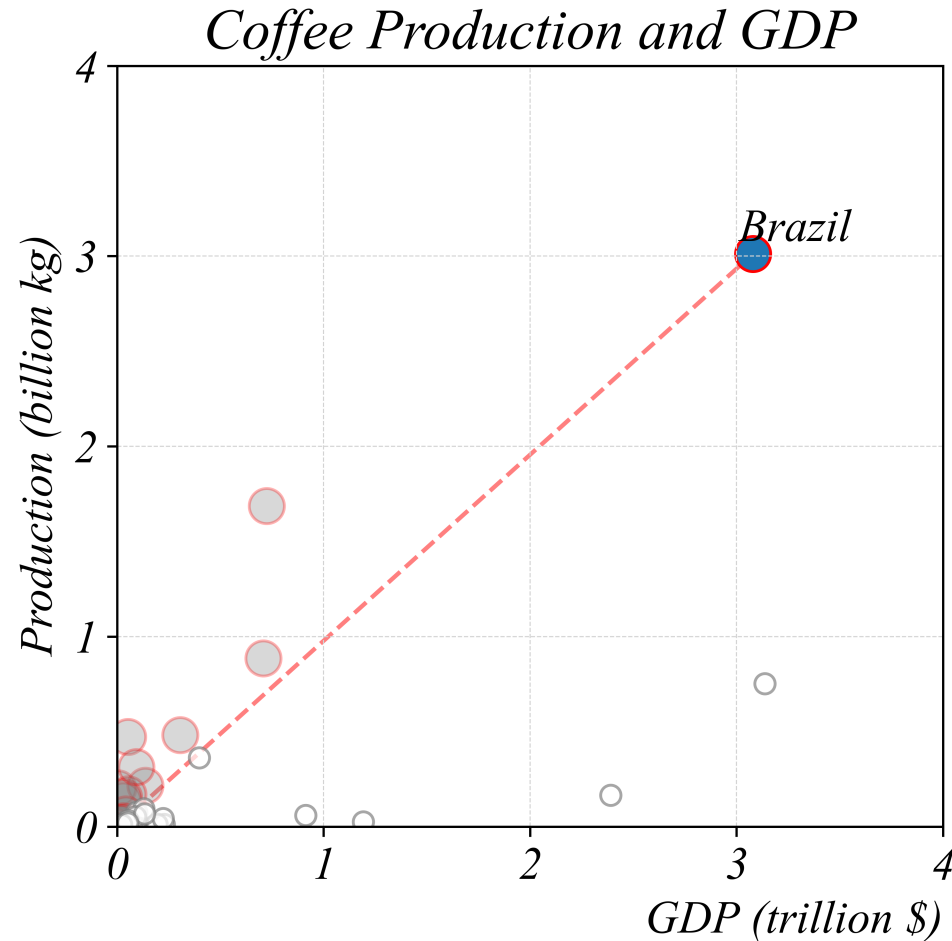
Which countries produce more coffee per dollar than Brazil?



> *separating lines can help make comparisons between ratios*

Cross-Sectional Relationships

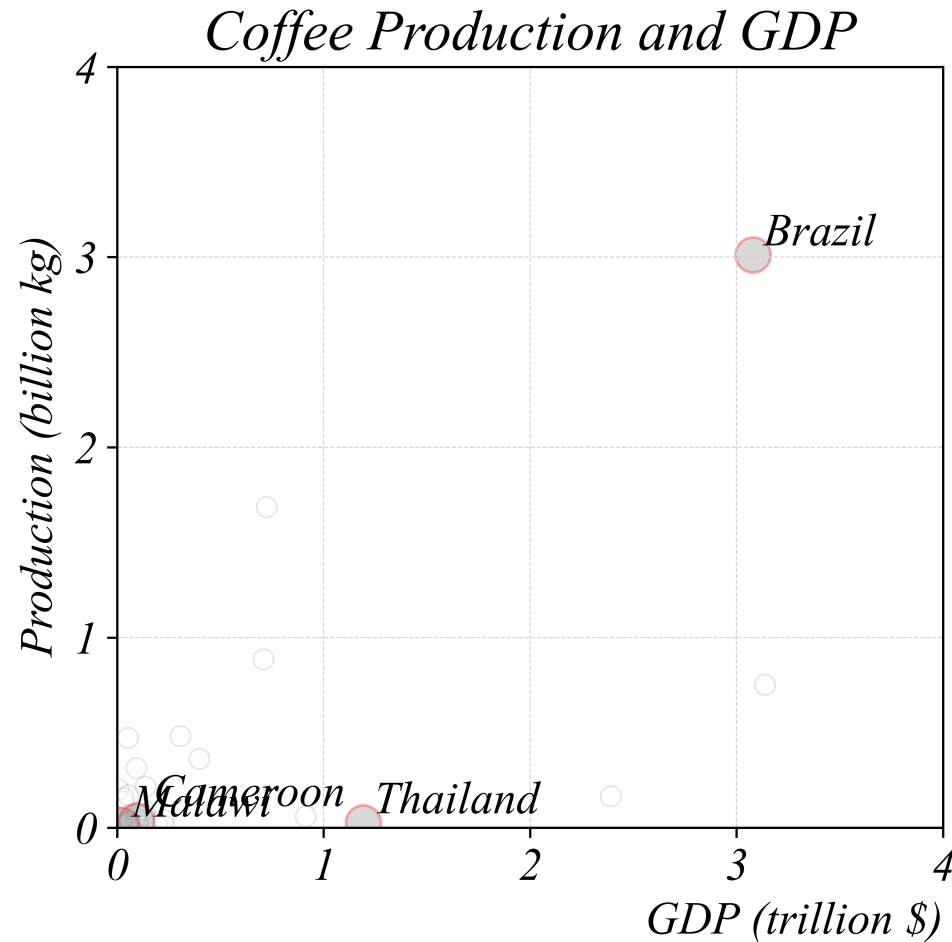
Which countries produce more coffee per dollar than Brazil?



> *separating lines can help make comparisons between ratios*

Cross-Sectional Relationships

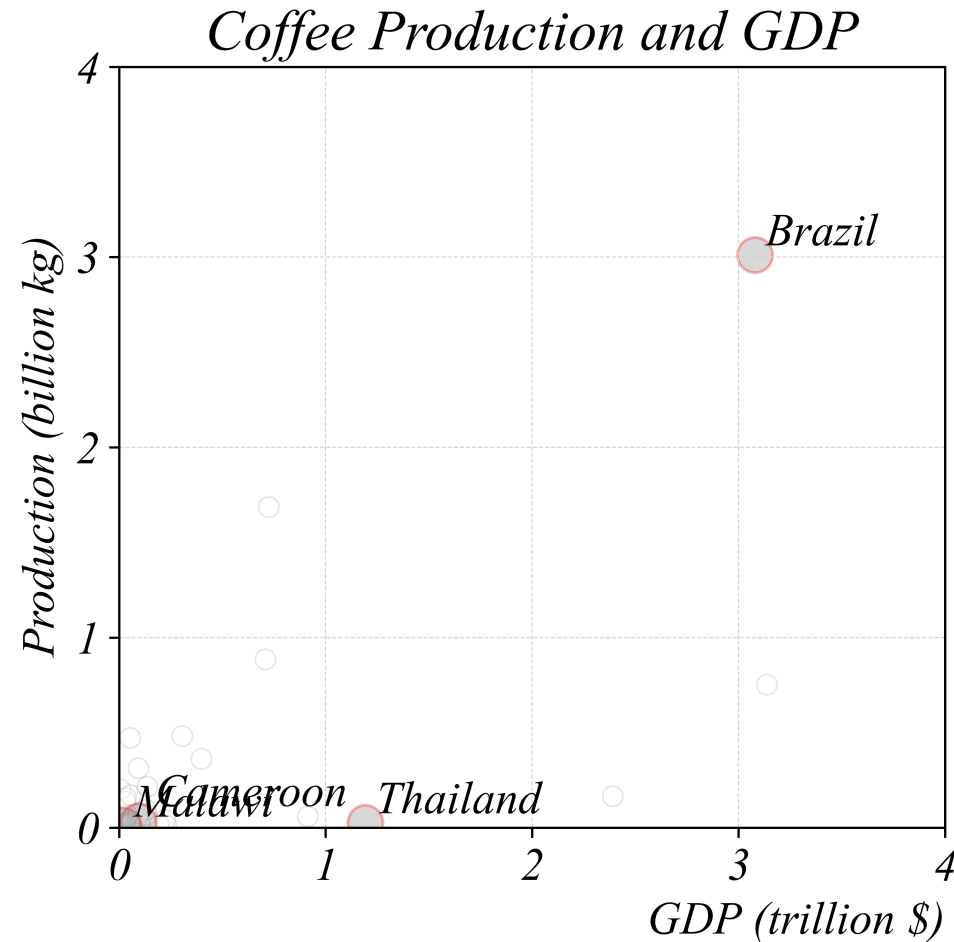
Do the GDPs of the upper or lower pair differ by a larger amount?



> use the differences on the horizontal axis to measure differences

Cross-Sectional Relationships

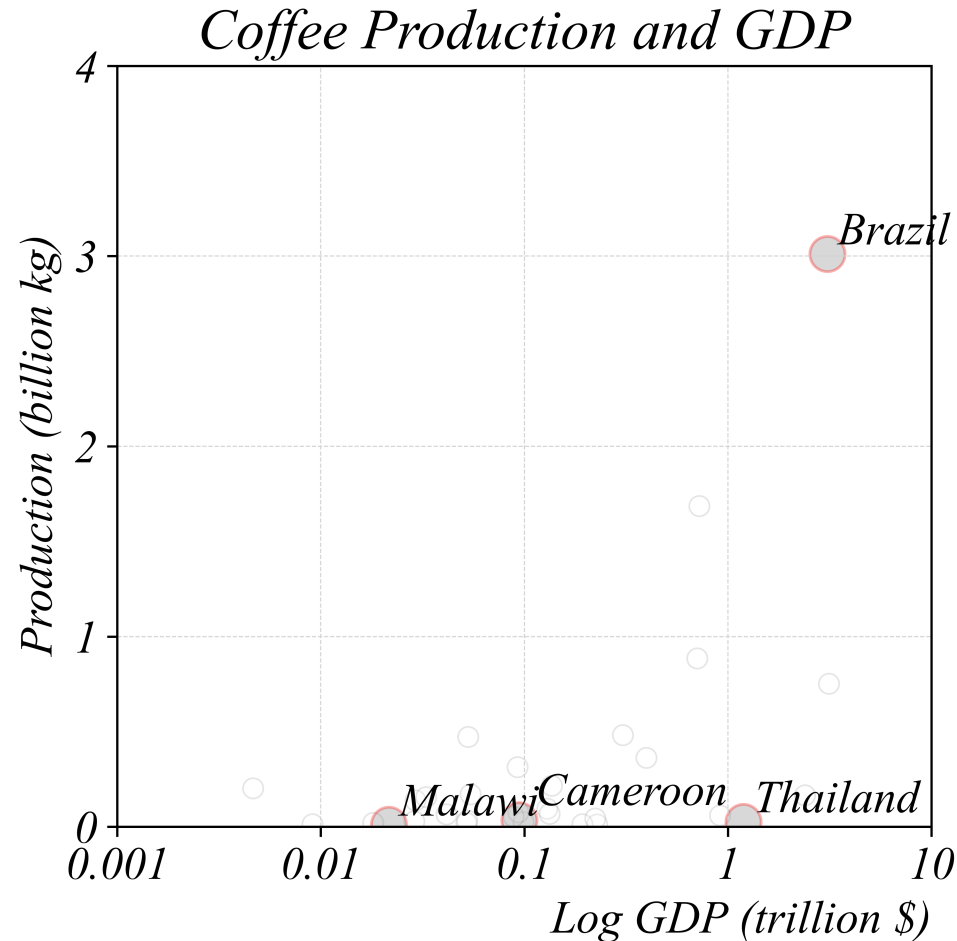
Which is larger: the ratio of GDPs of the upper or lower pair?



> this question is difficult to answer with this scale

Cross-Sectional Relationships

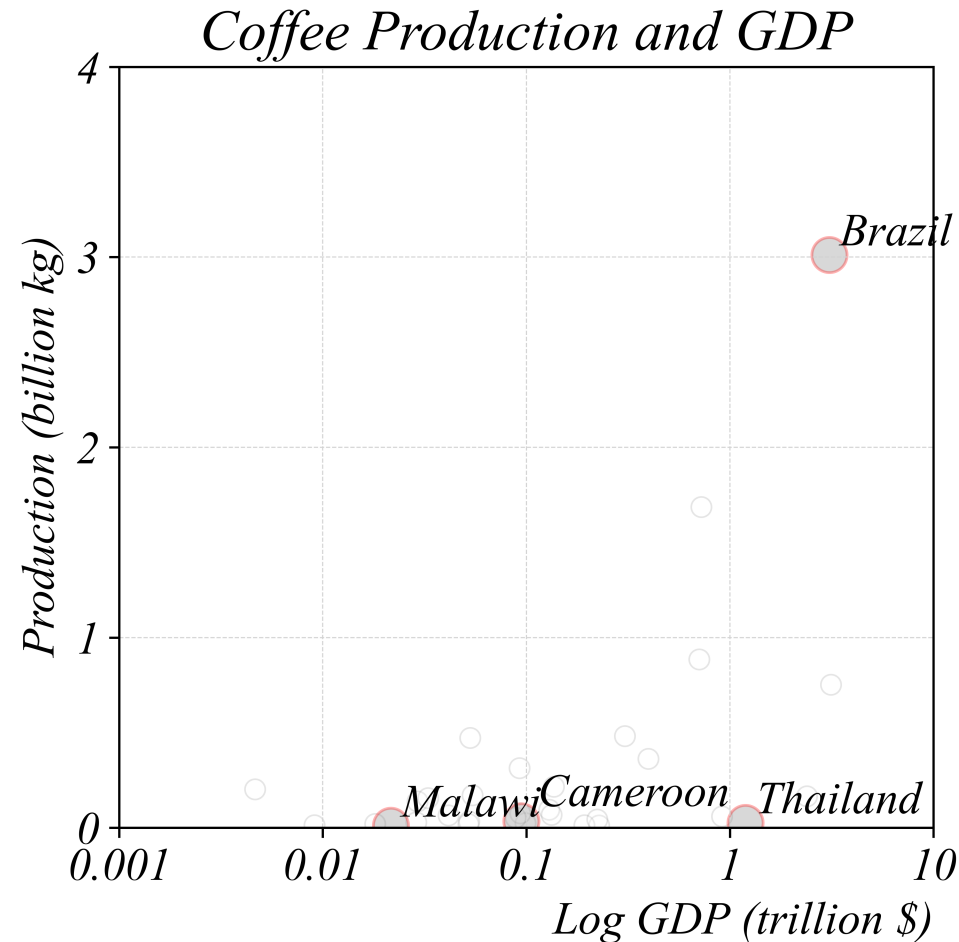
Which is larger: the ratio of GDPs of the upper or lower pair?



> a log scale makes RATIOS easier to visualize: each tick is 10x larger

Cross-Sectional Relationships

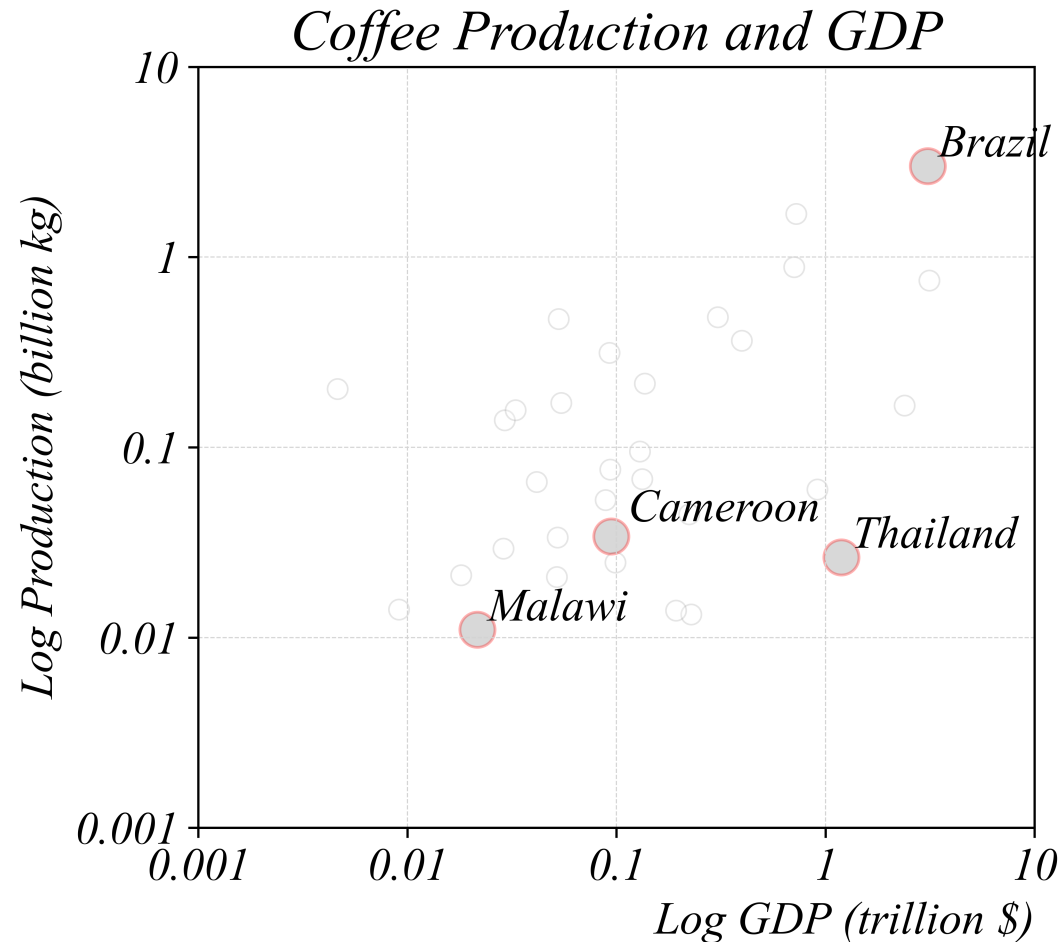
Which country produces the second highest output of coffee?



> *a log scale also makes it easier to see SCALING*

Cross-Sectional Relationships

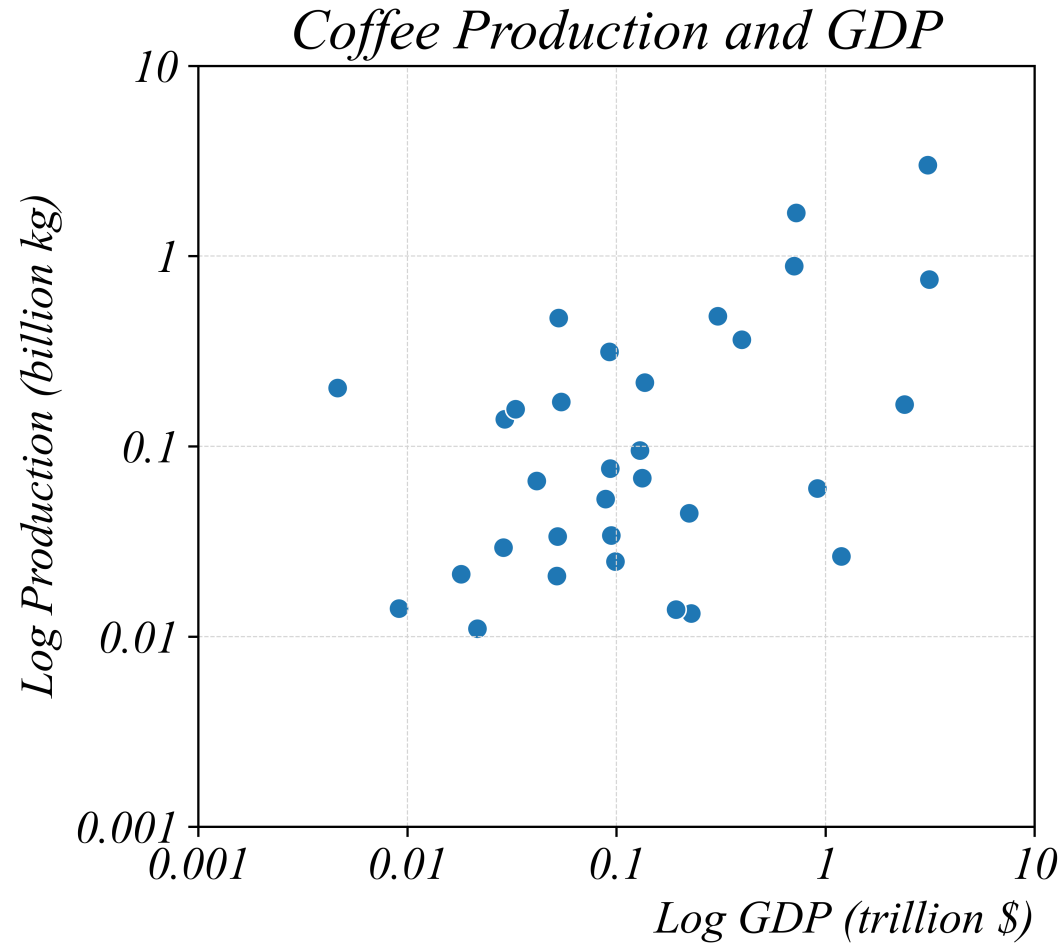
Which country produces the second highest output of coffee?



> *scaling the vertical axis in logs clarifies both small and large variation*

Cross-Sectional Relationships

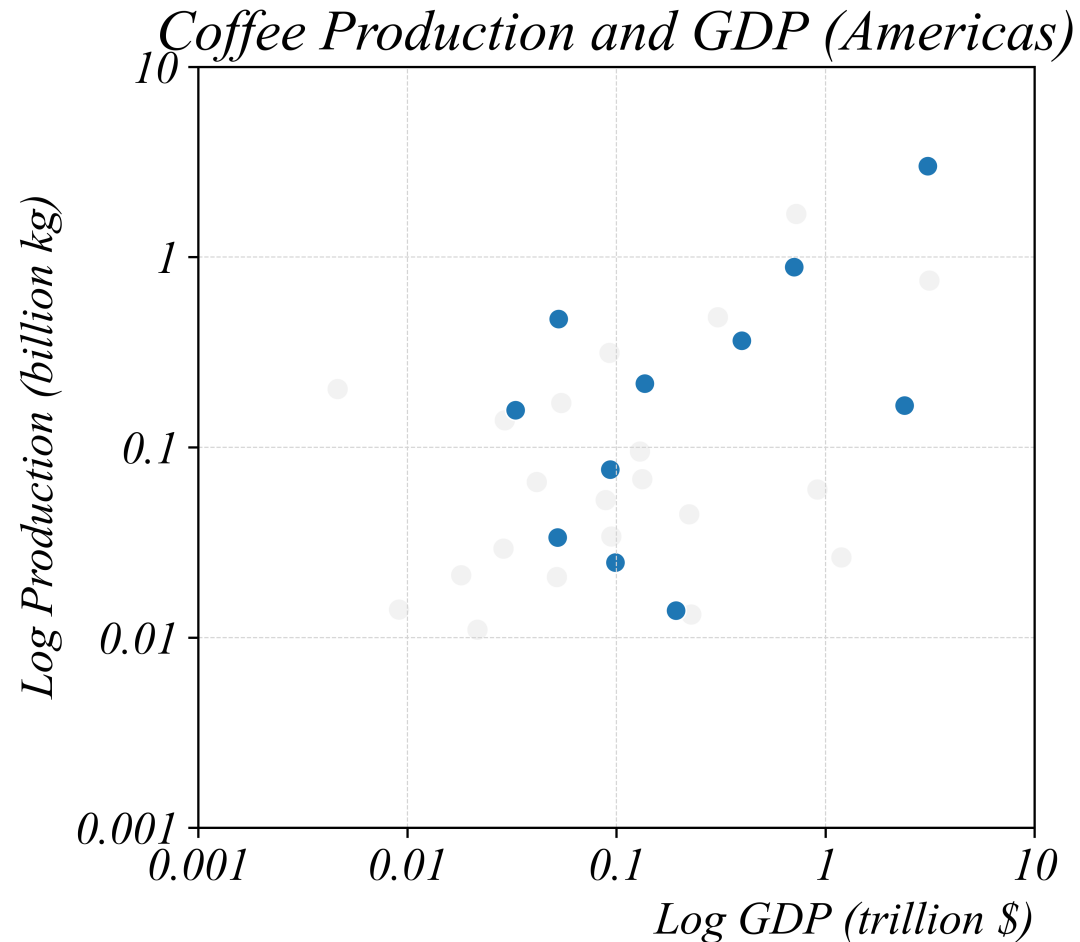
How does GDP relate to coffee production in the Americas?



> lets use a filter with this data

Cross-Sectional Relationships

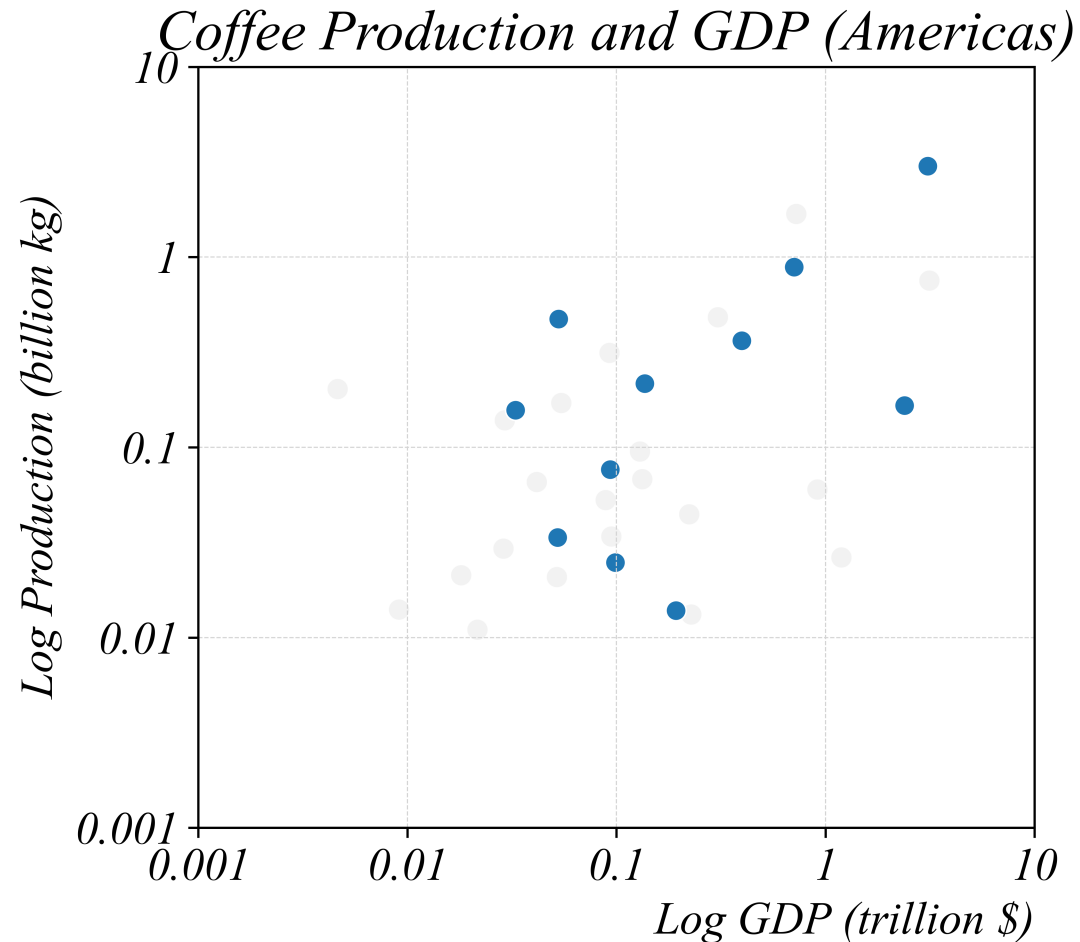
How does GDP relate to coffee production in the Americas?



> looks positive, but we'll formally test this in Part 4

Cross-Sectional Relationships

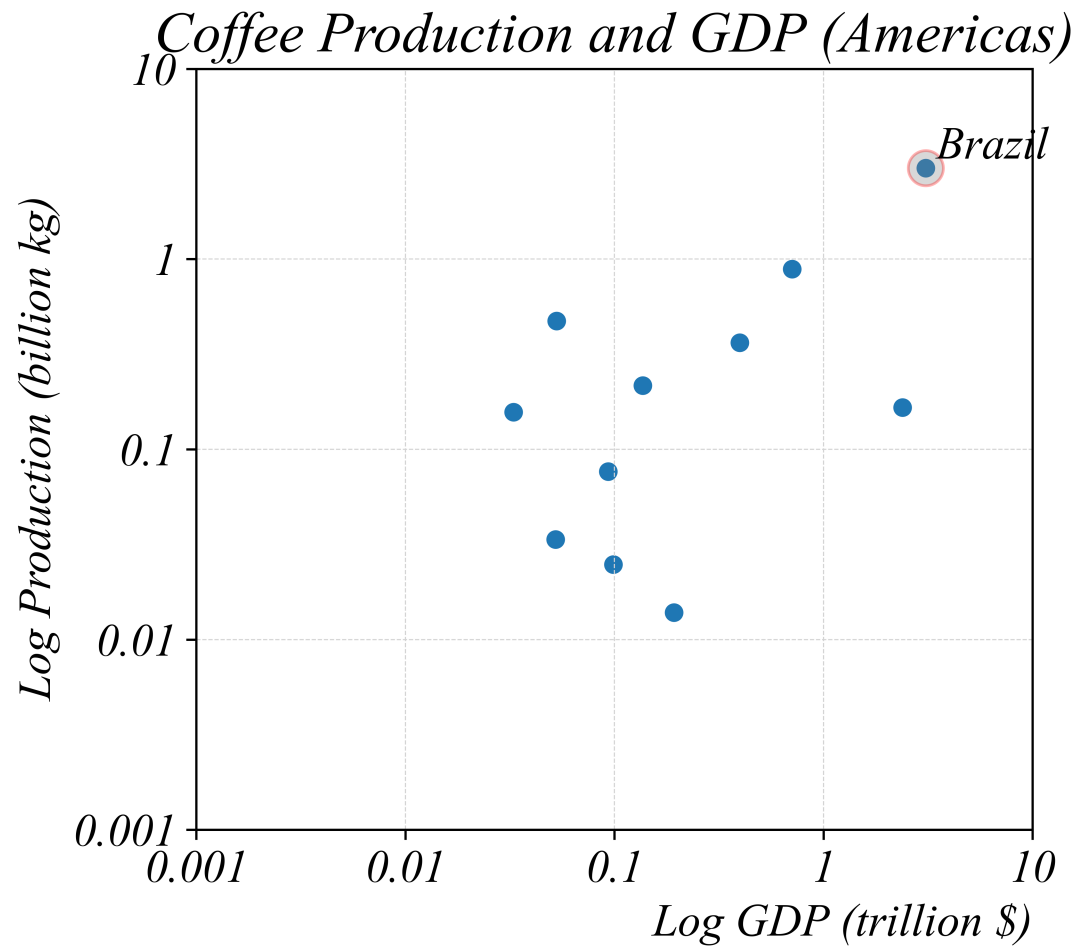
Which country in the Americas produces the most coffee?



> *looks positive, but we'll formally test this in Part 4*

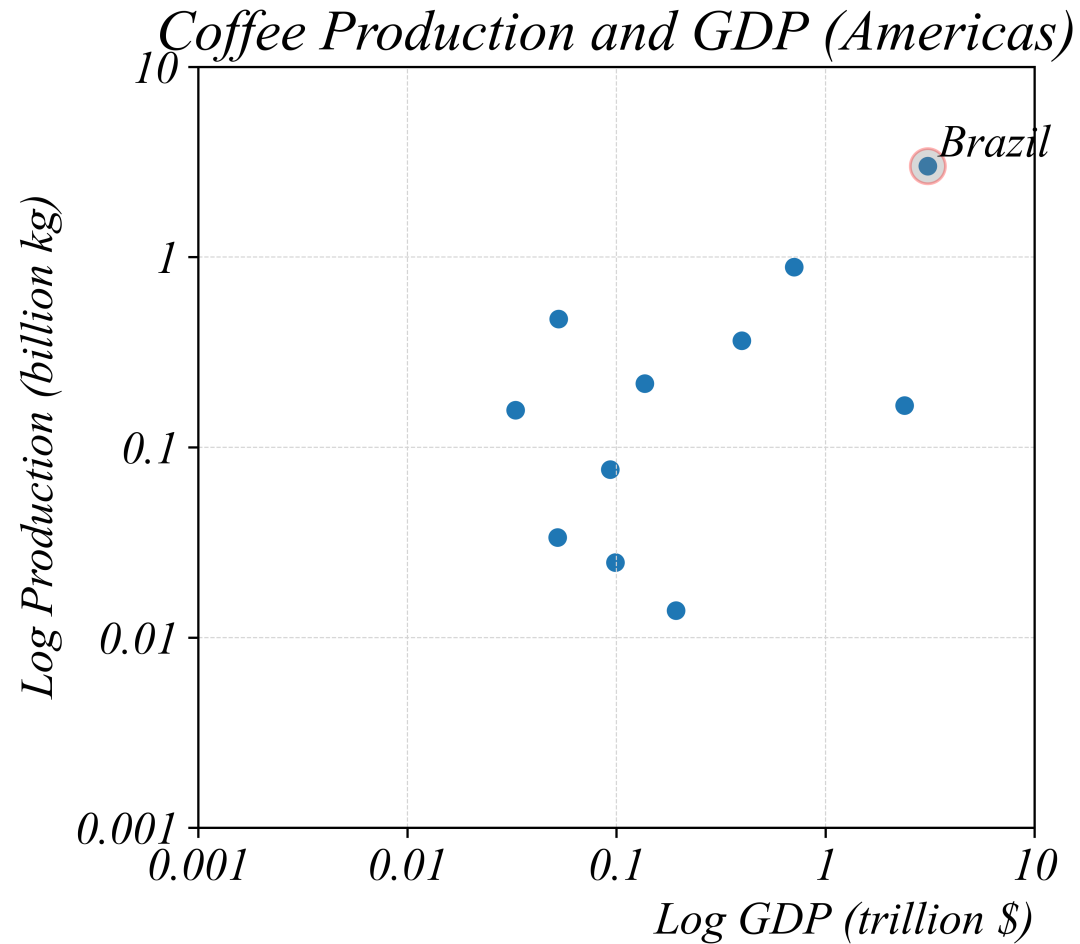
Cross-Sectional Relationships

Which country in the Americas produces the most coffee?



Cross-Sectional Relationships

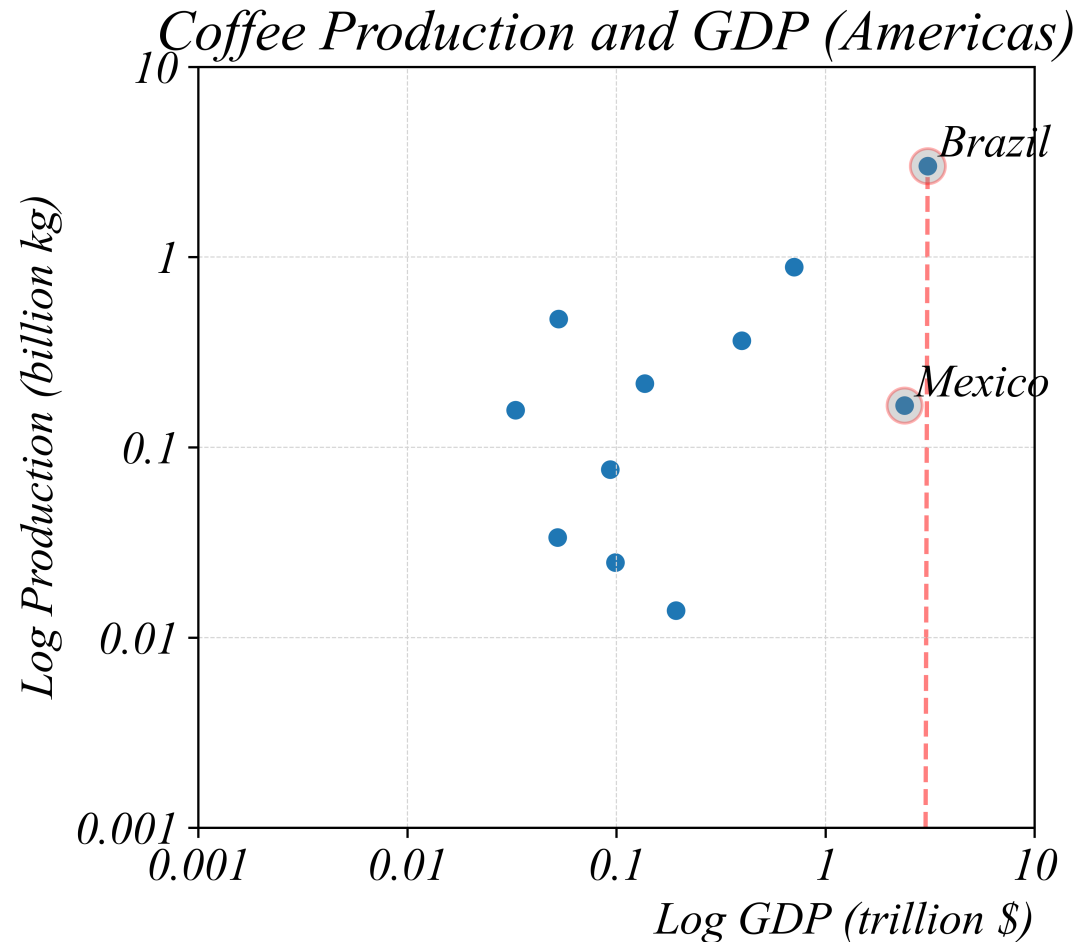
Which country's GDP is closest to Brazil's GDP?



> we can use the horizontal axis

Cross-Sectional Relationships

Which country's GDP is closest to Brazil's GDP?



> we can use a vertical line here to find the closest on the horizontal axis

Summary

- *Relationships between two variables can be easily summarized in scatterplots.*
- *Scatterplots make it easy to visually filter either axis.*
- *Scatterplots make it easy to compare absolute differences by axis.*
- *A log-scale transformation makes it easy to compare ratios.*
- *A log-log-scale makes it easy to see variation in both large and small values.*

Exercise: Cross-Sectional Scatterplots

Visualizing GDP and Coffee Production Relationships

We're going to use a scatterplot to visually examine the relationship between coffee production and GDP.

- ***Data:*** *Beans_GDP_2019.csv*

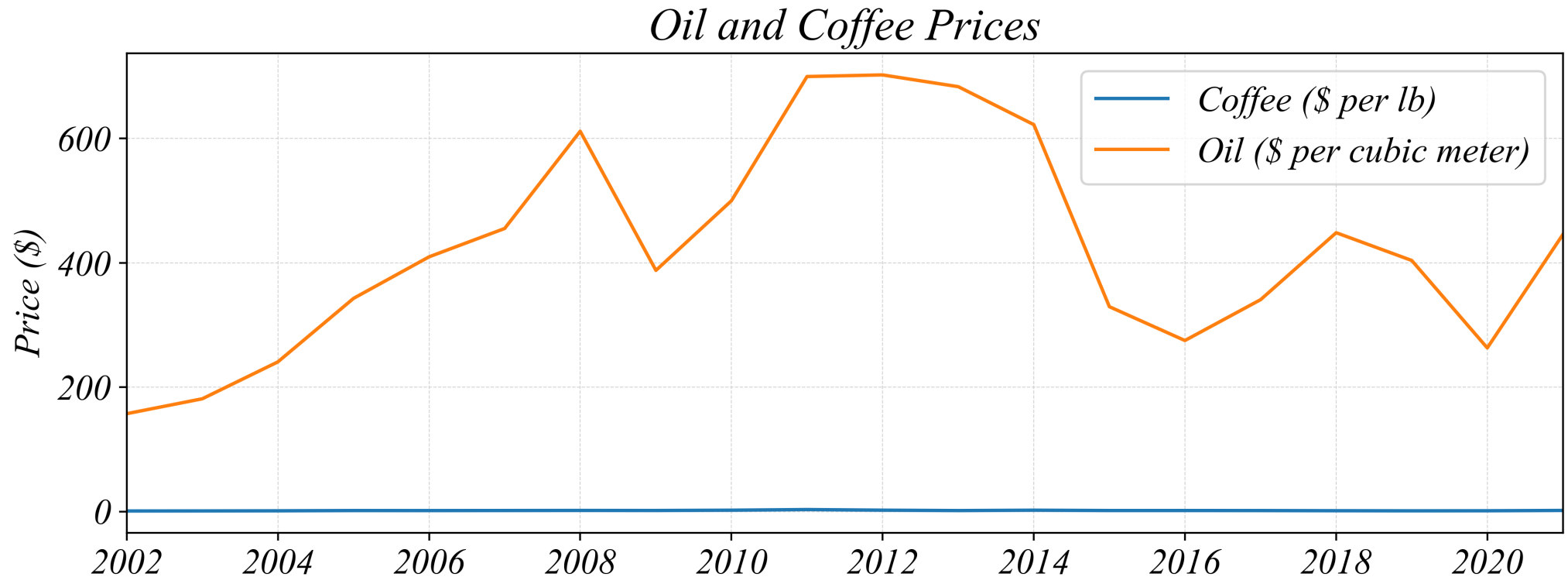
Exercise: Cross-Sectional Scatterplots

Visualizing GDP and Coffee Production Relationships

```
1 # Scatterplot  
2 sns.scatterplot(gdp, x='GDP', y='coffee_prod')
```

Timeseries Relationships

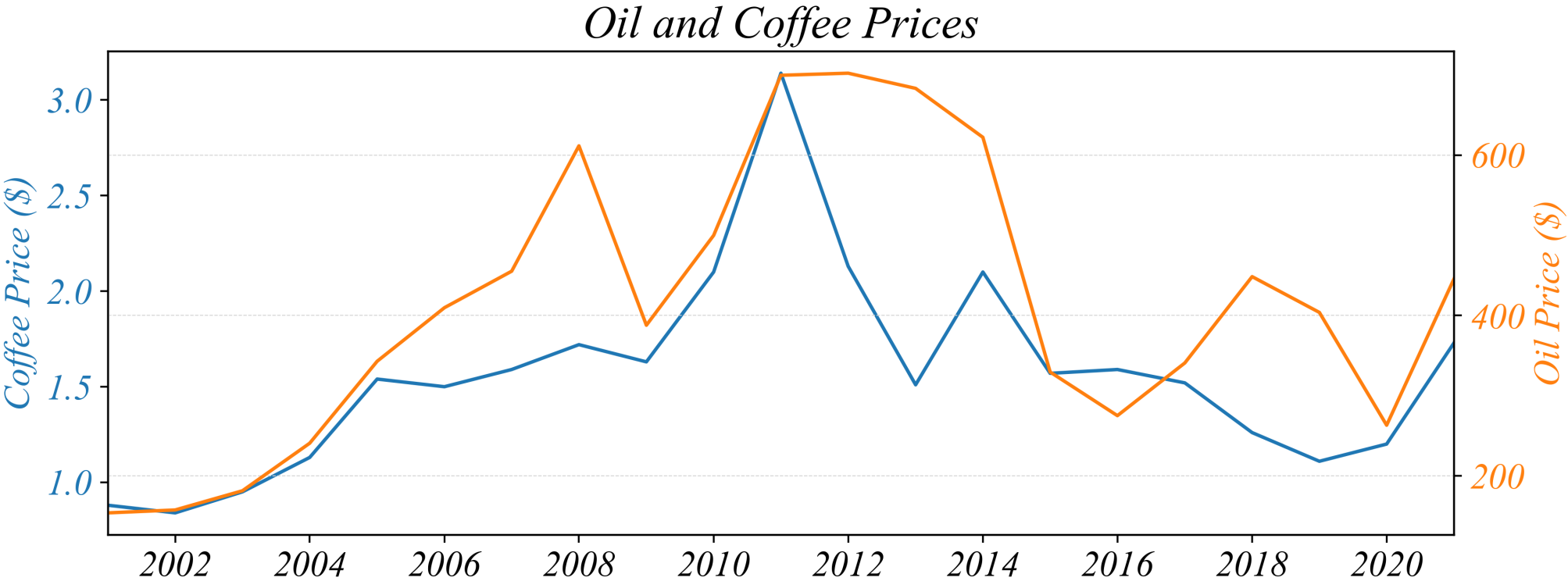
How do the two commodity prices relate to each other?



> *difficult to tell because of the axis scale*

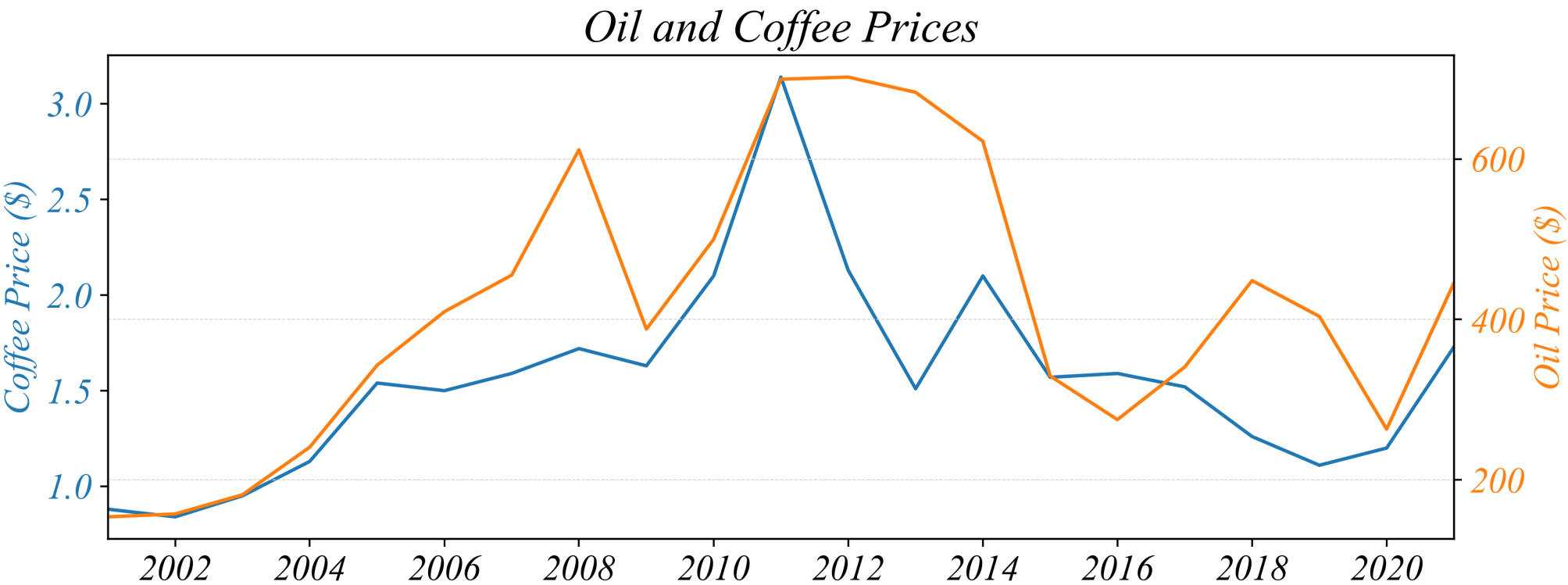
Timeseries Relationships

How do the two commodity prices relate to each other?



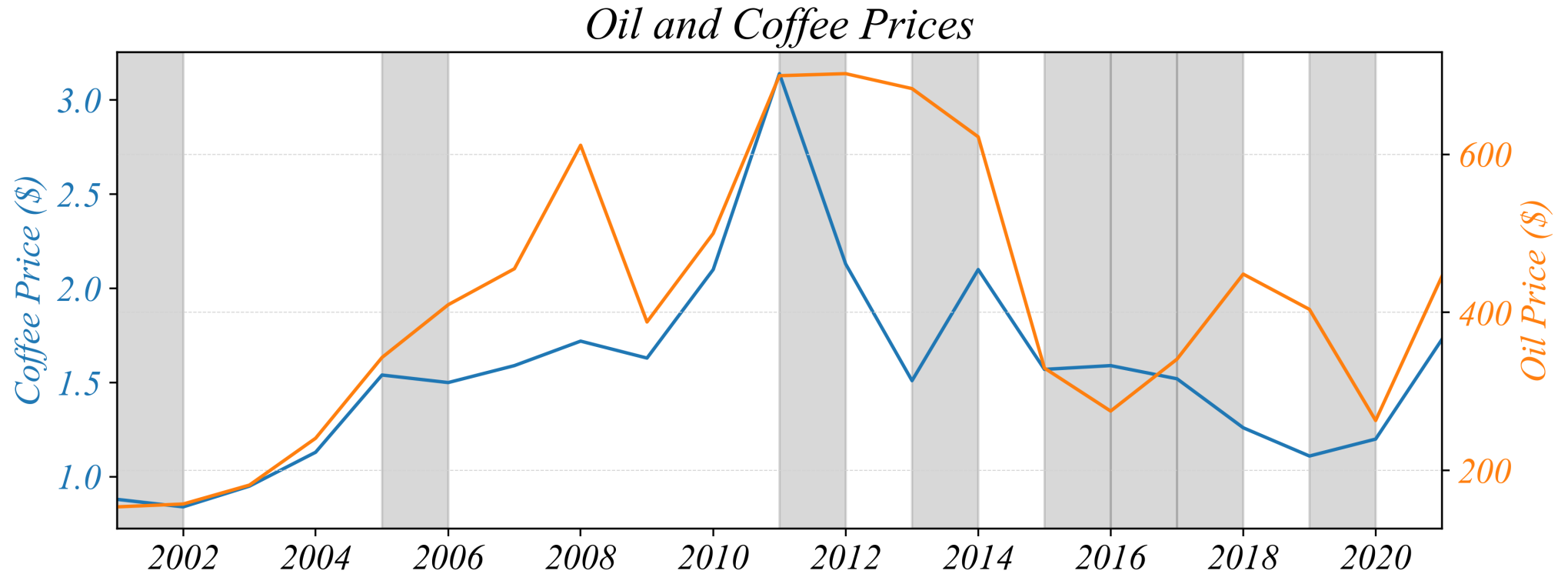
Timeseries Relationships

*In which years did oil and coffee prices move in **opposite** directions?*



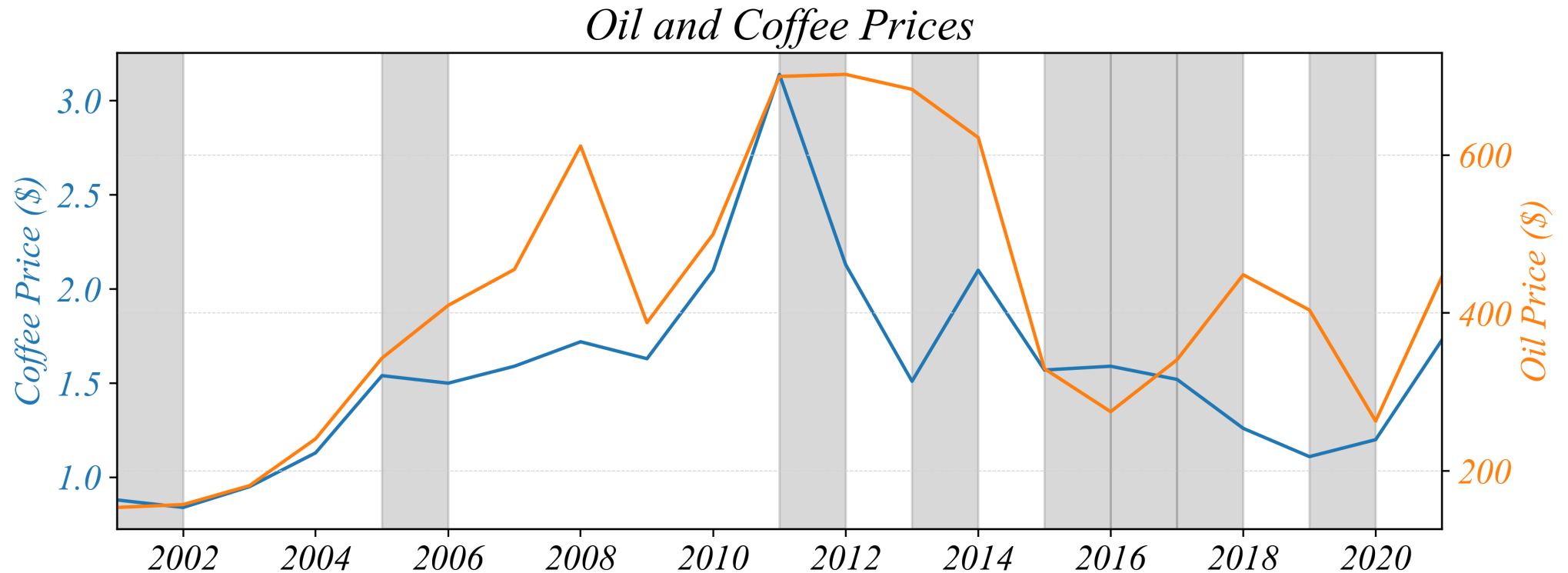
Timeseries Relationships

*In which years did oil and coffee prices move in **opposite** directions?*



Timeseries Relationships

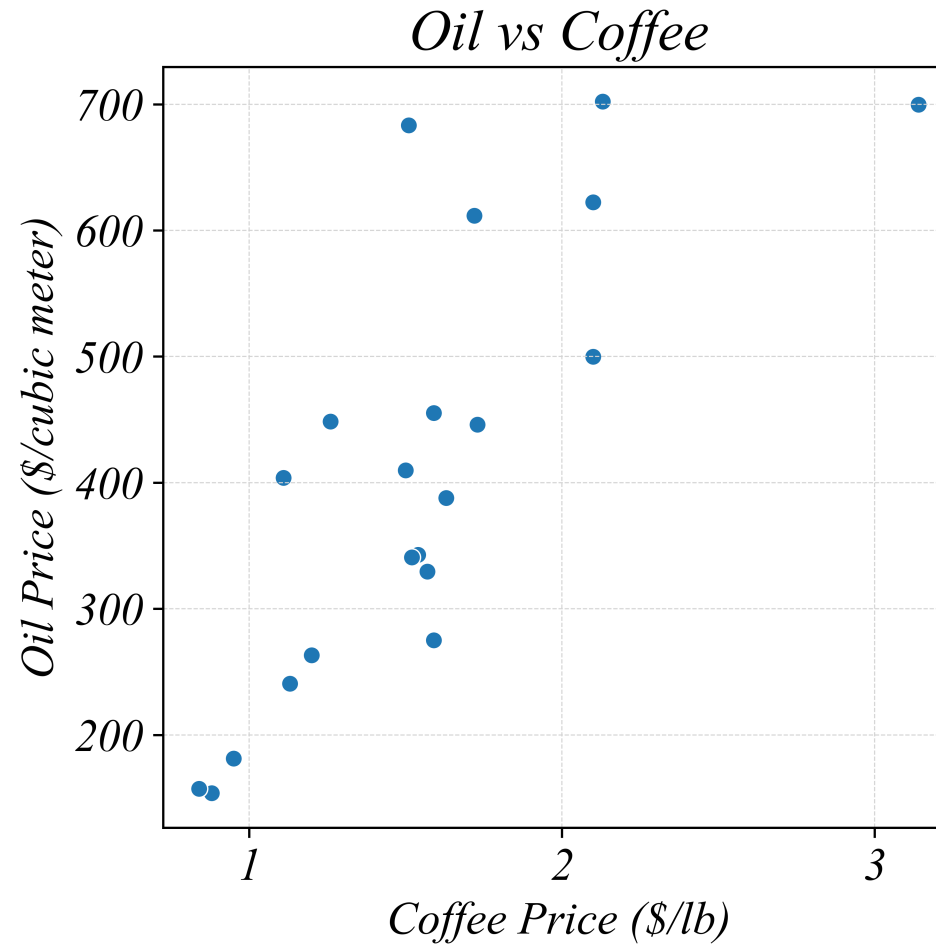
But are the two prices positively or negatively related to each other?



> *this is difficult to see with just a Multi-Lineplot...*

Timeseries Relationships

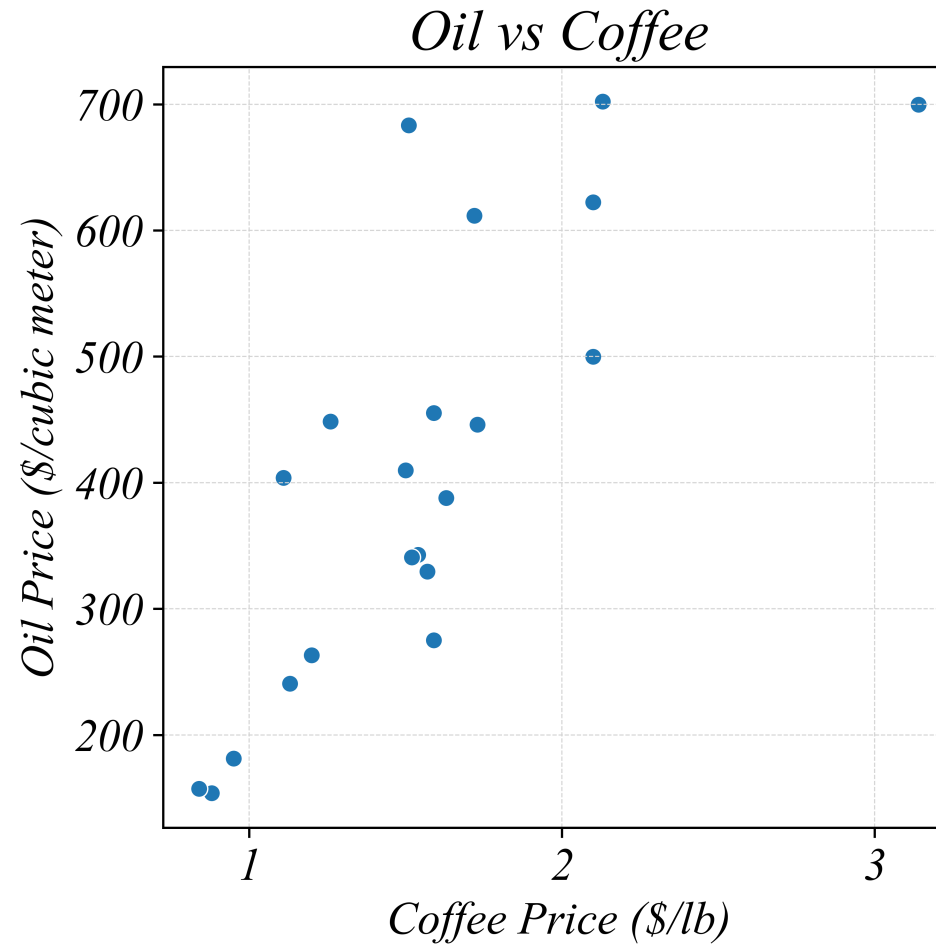
But are the two prices positively or negatively related to each other?



> *a Scatterplot can show the relationship between two variables through time*

Timeseries Relationships

Does the price of oil determine the price of coffee?



> *a Scatterplot can only show associations not causation :(*

Exercise: Timeseries Scatterplots

Visualizing Coffee Prices and Oil Prices

We're going to use a scatterplot to visually examine the relationship between coffee prices and oil prices.

- **Data:** *Coffee_Oil.csv*

Exercise: Timeseries Scatterplots

Visualizing Coffee Prices and Oil Prices

```
1 # Scatterplot  
2 sns.scatterplot(gdp, x='GDP', y='coffee_prod')
```