ECON 0150 | Economic Data Analysis

The economist's data analysis pipeline.

Part 1.3 | Relationships Through Time

Data Structures

...three main relationships between data points.

The most effective summarization tool depends on the relationship between the data points.

	Cross-Sectional	**Time-Series**	**Panel Data**
Focus	Multiple units, one time point	One unit, many times	Multiple units, many time points
Shape	Wide format	Long format	Long format
Ex.	Household income, 2025	US GDP, 10 years	Household income, 10 years

> we've spent the first part of the class on cross-sectional data

> we'll spend a bit of time on panel and geographic data later

Exercise 1.3 | Data Structures

Lets identify the variable type for each dataset.

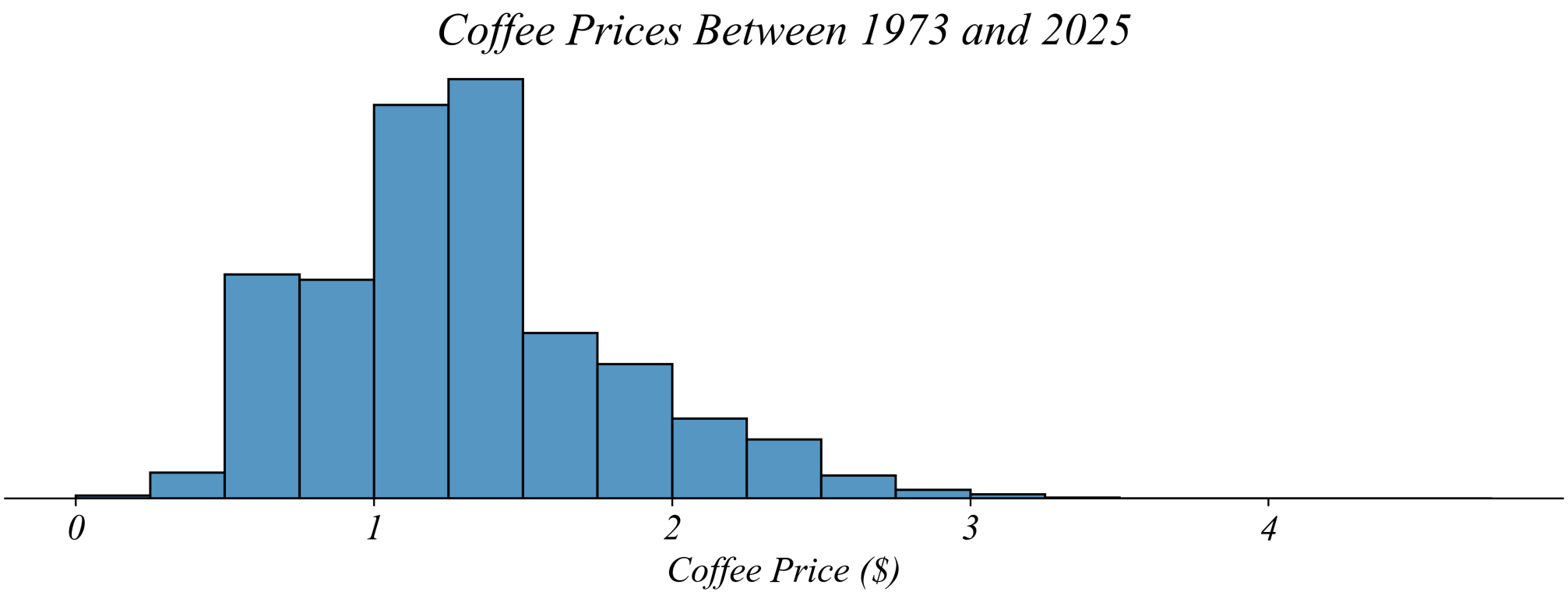
- ***Dataset 1:*** `household_incomes.csv`
- ***Dataset 2:*** `household_savings.csv`
- ***Dataset 3:*** `Monthly_Coffee_Prices.csv`

Timeseries: Coffee Prices

What information should we use to set prices in January 2026?

Timeseries: Coffee Prices

What information should we use to set prices in January 2026?

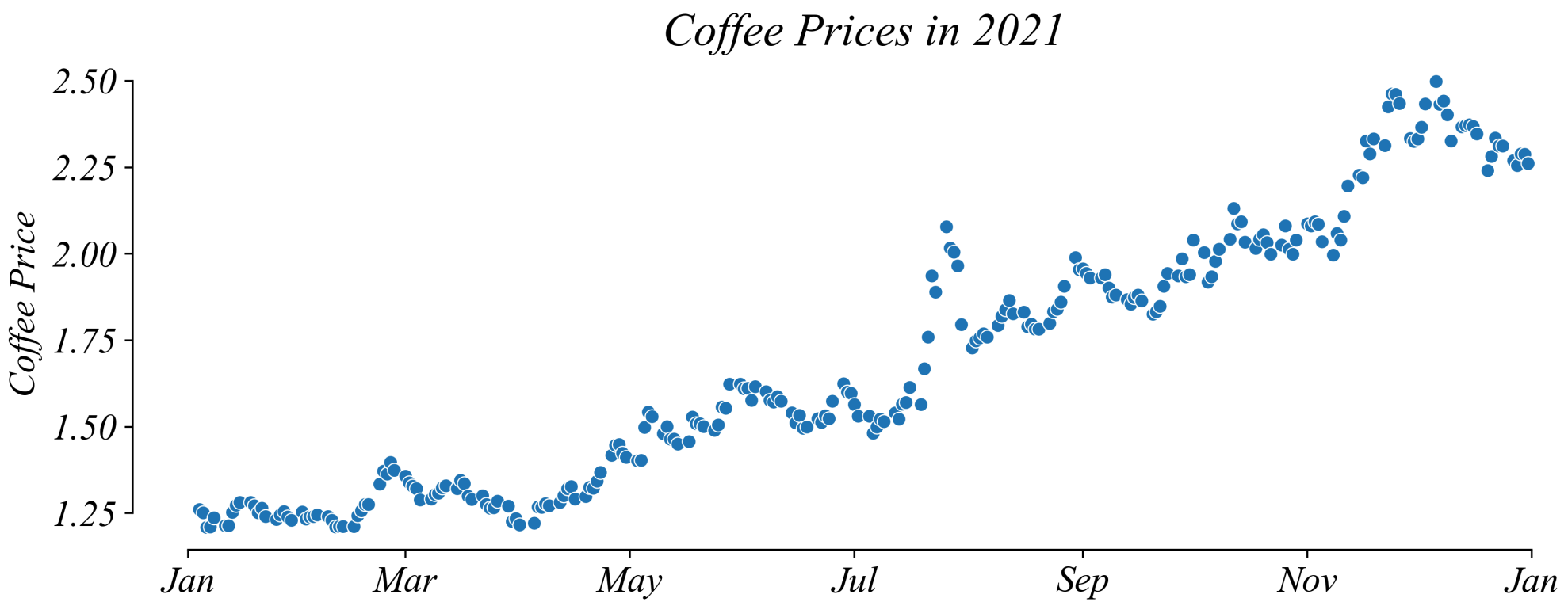


> *it's difficult to know… do we choose the mode?*

> *lets just plot the price against time*

Timeseries: Coffee Prices

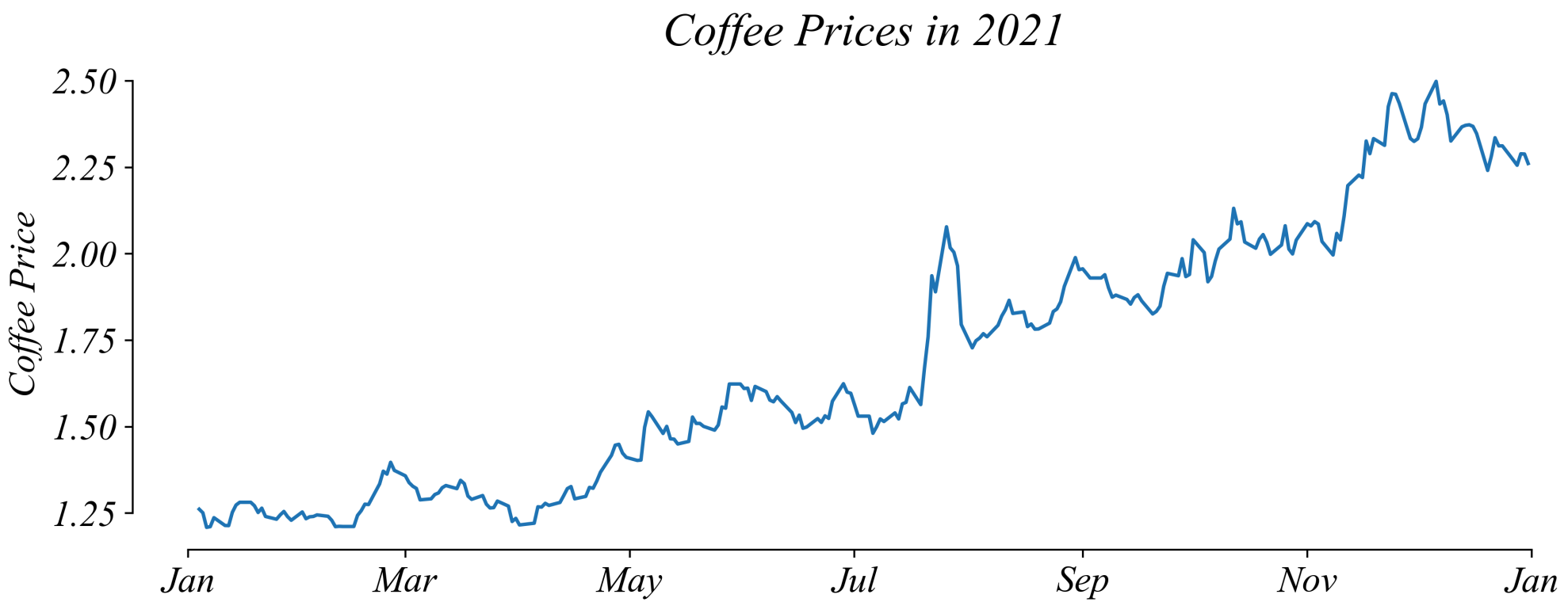
What information should we use to set prices in January 2026?



> *lets indicate with a line that these points are in squence*

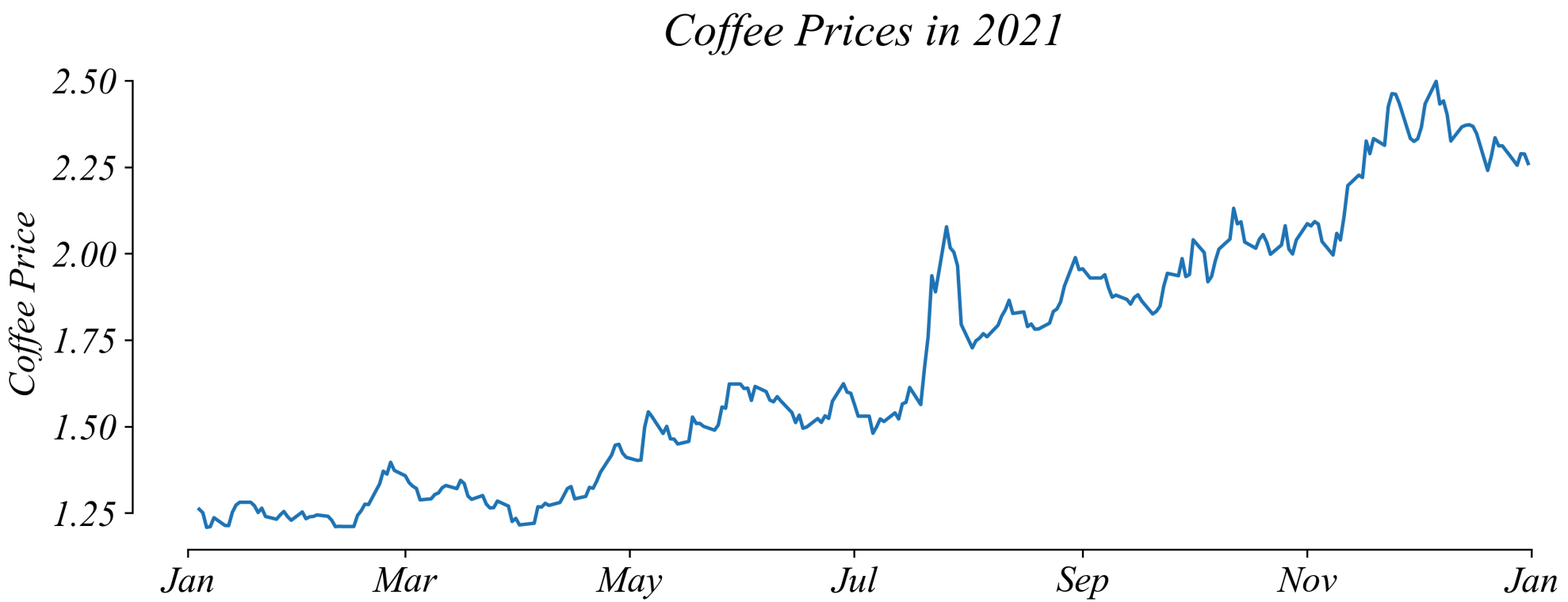
Timeseries: Line Graph

What information should we use to set prices in January 2026?



Timeseries: Trends

*Do you notice a **trend** in price?*

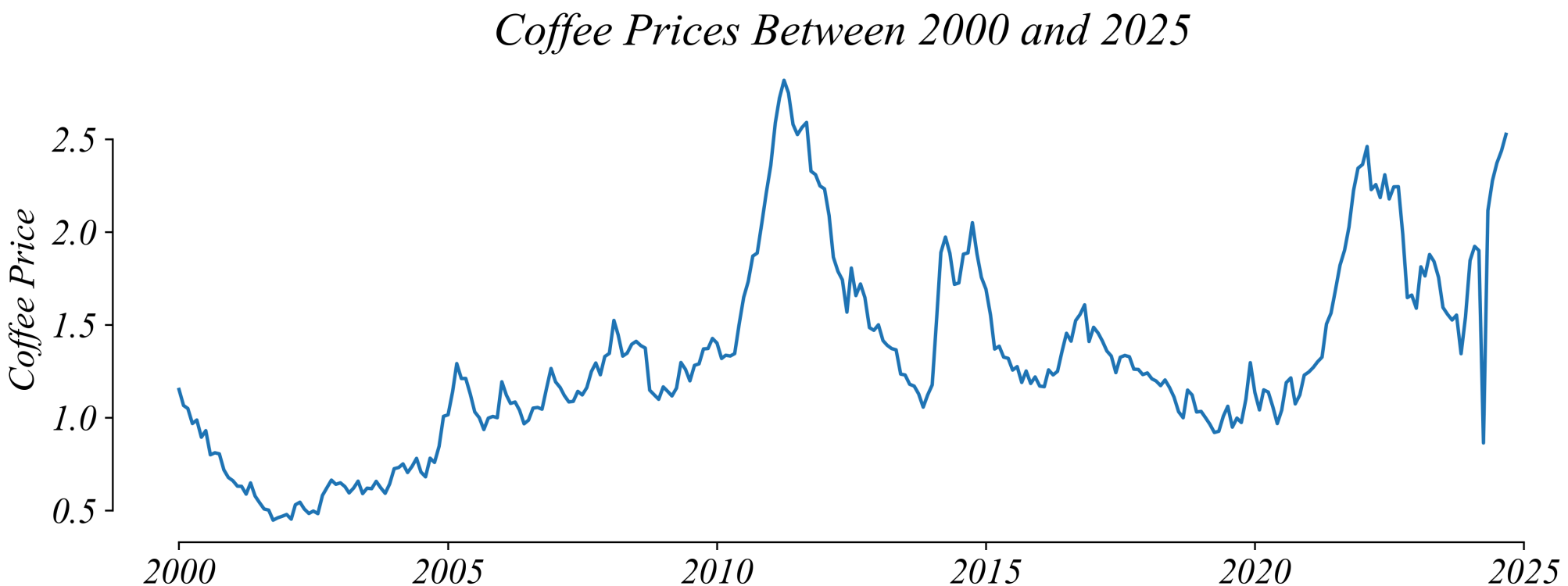


> *there was a **positive trend** in 2021*

> *we can zoom out to get a bigger picture*

Timeseries: Trends + Subtrends

*Do you notice a **trend** in price?*

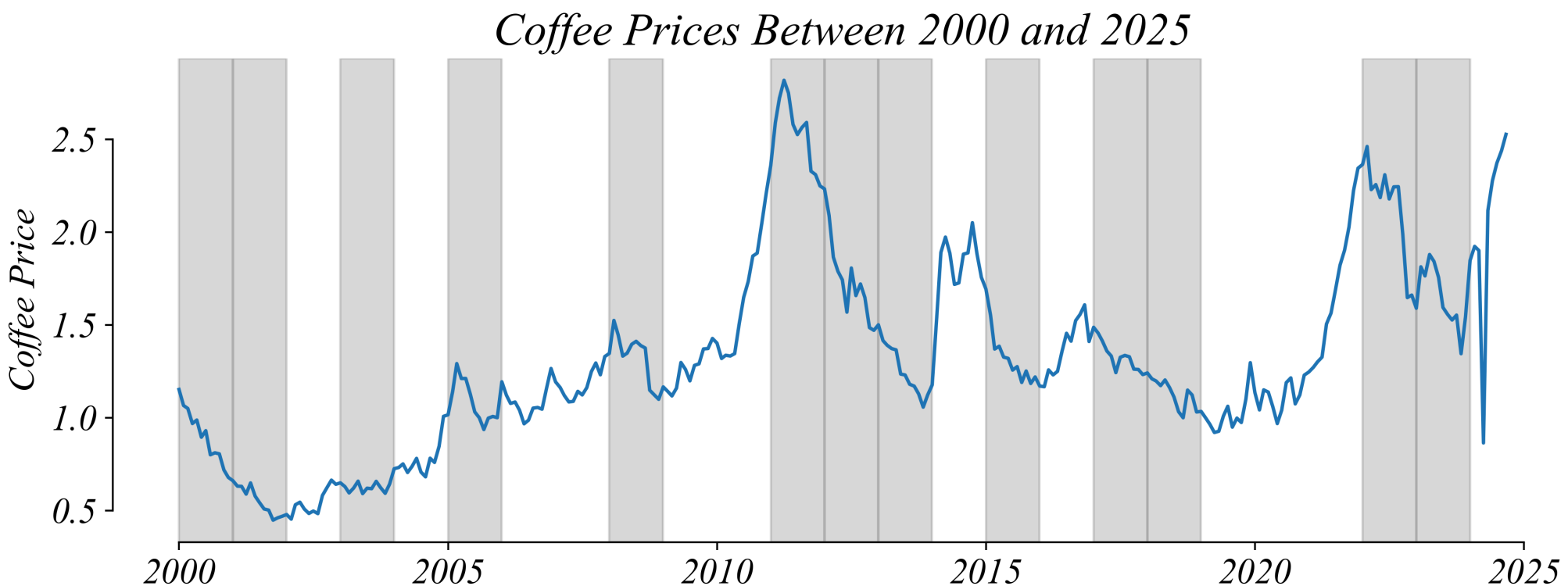


> *how have prices changed since 2000?*

> *prices have increased somewhat, with many periods of decrease*

Timeseries: Background Shading

What information should we use to set prices in January 2026?



> *with background shading its easier to see periods with a negative trend in price*

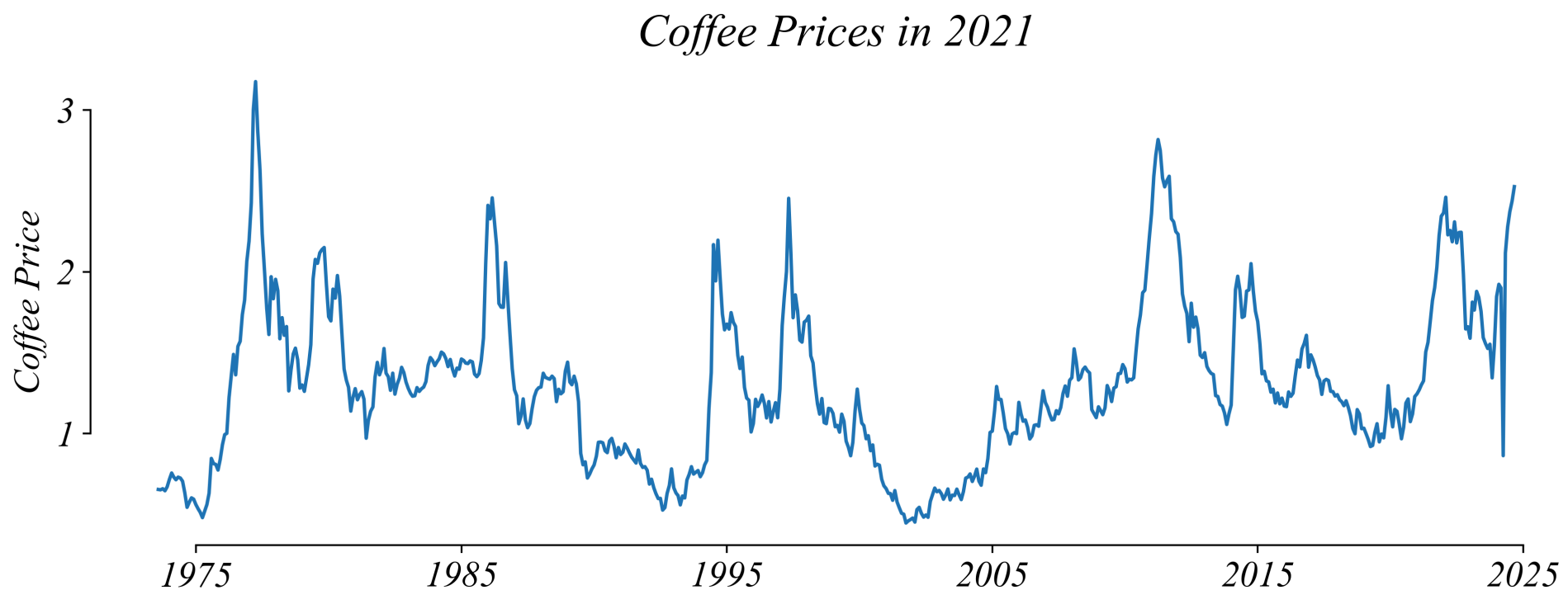
Exercise 1.3: Timeseries

Lets use a linegraph to examine the trends in coffee prices.

- ***Data:*** `Coffee_Prices.csv`

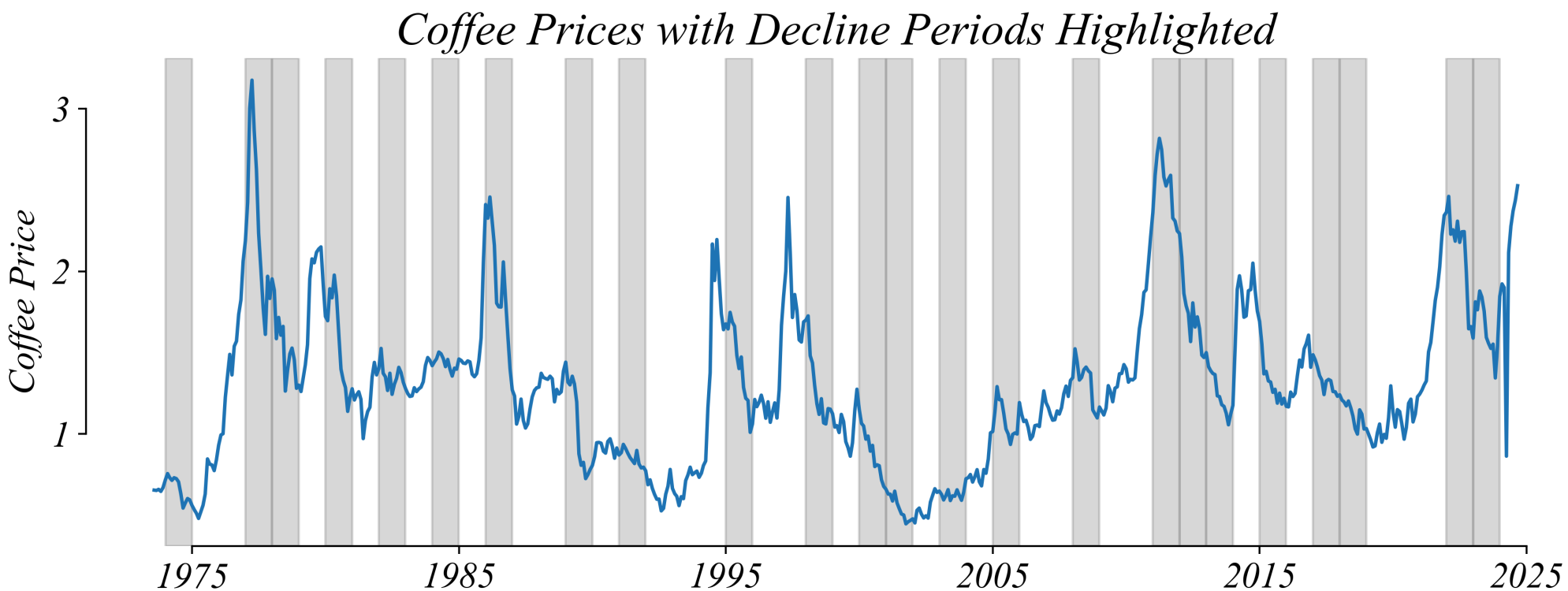
Exercise 1.3: Timeseries

```
1 # Lineplot
2 sns.lineplot(prices, y='price', x='date')
```



Timeseries

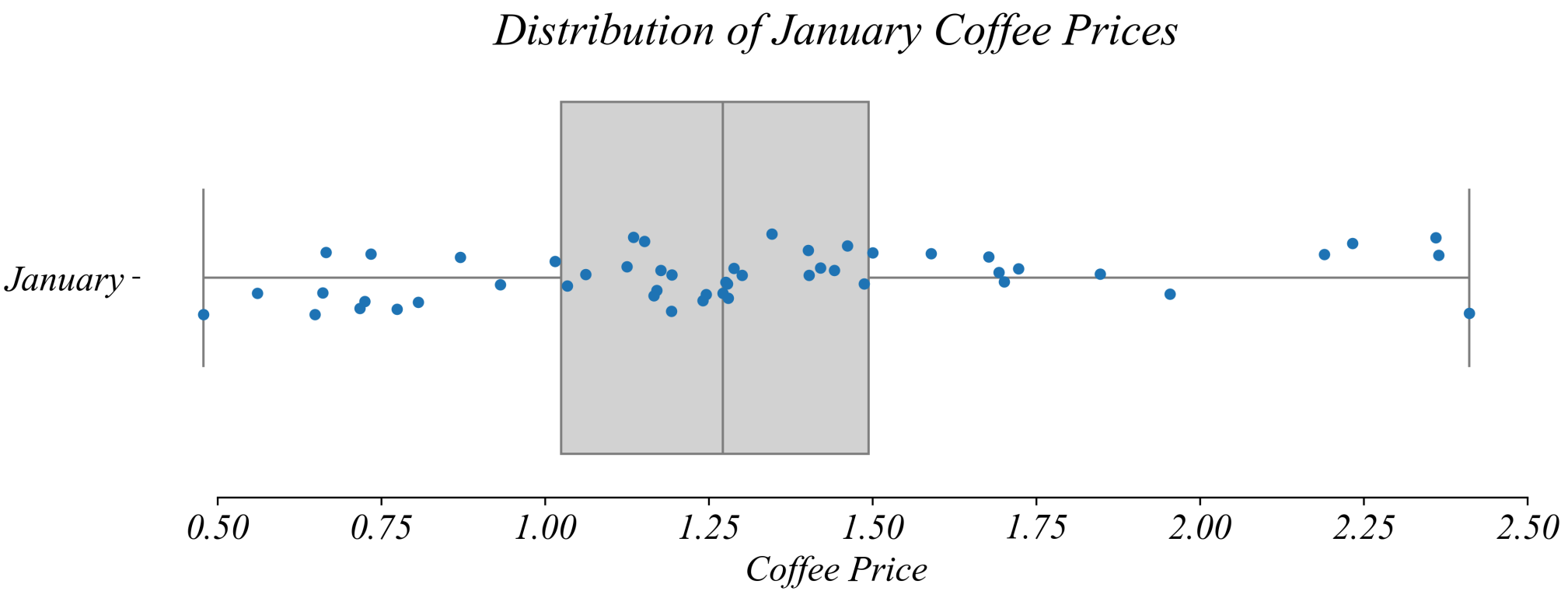
What information should we use to set prices in January 2026?



> *could there be seasonal trends within the larger trend?*

Seasonality: January

What information should we use to set prices in January 2026?

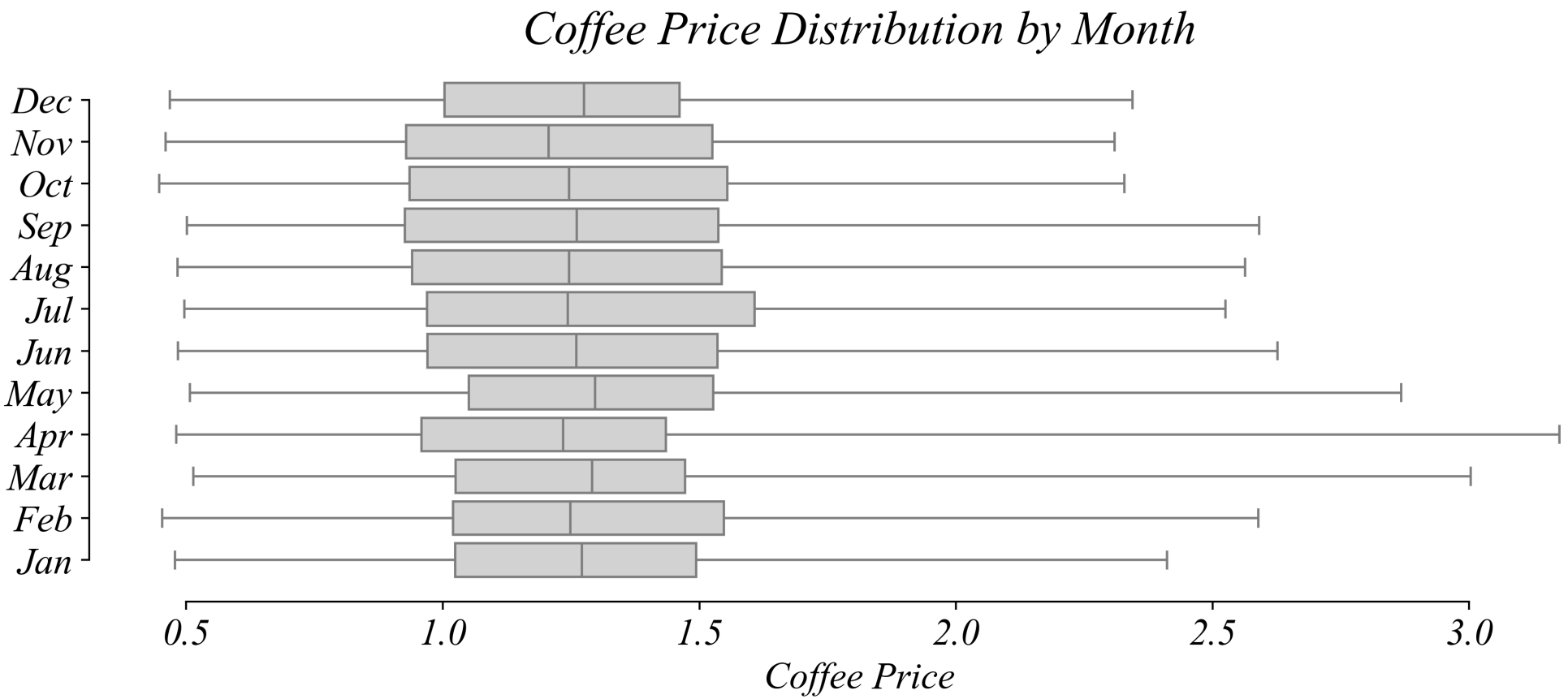


> *a boxplot gives us a picture of the prices just in January*

> *lets compare this to other months*

Seasonality: Monthly Boxplots

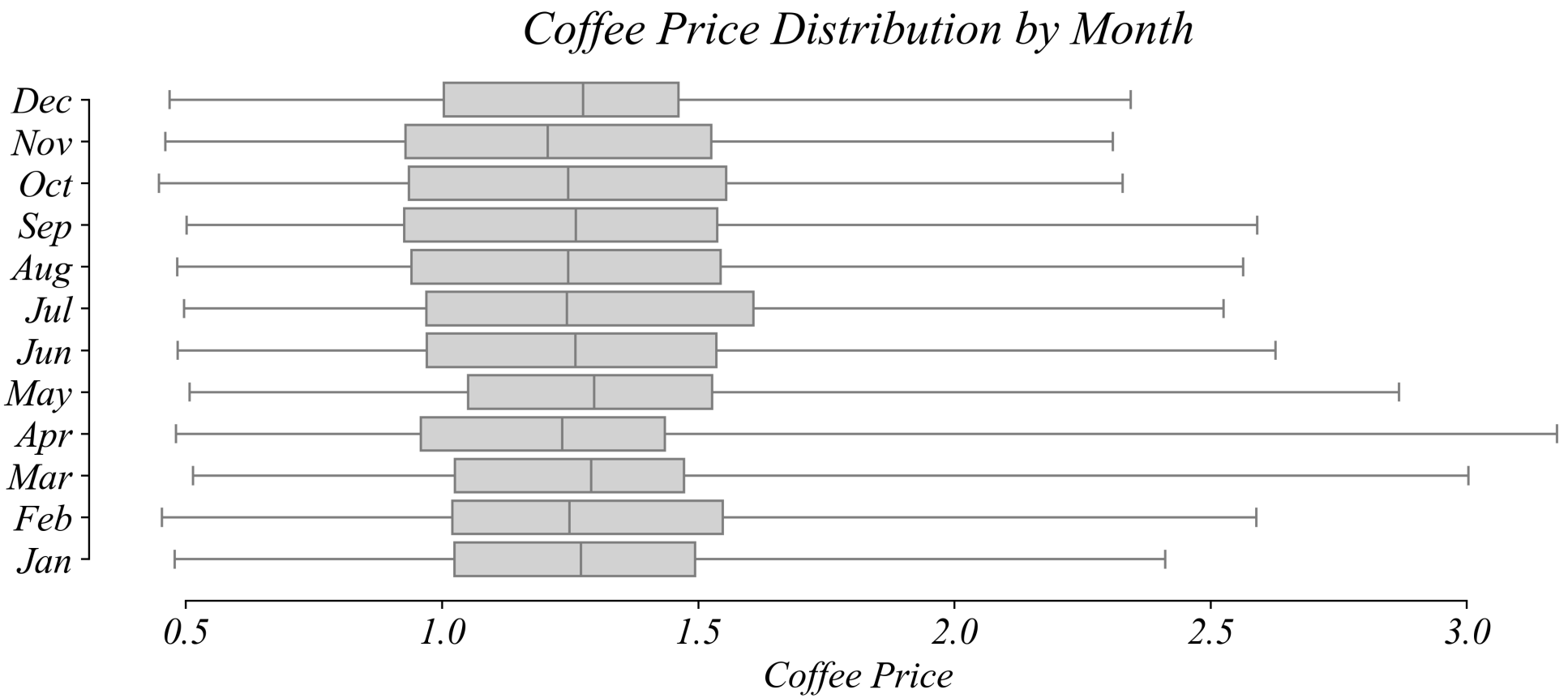
In addition to the overall trend, are there monthly patterns?



> lets be more specific…

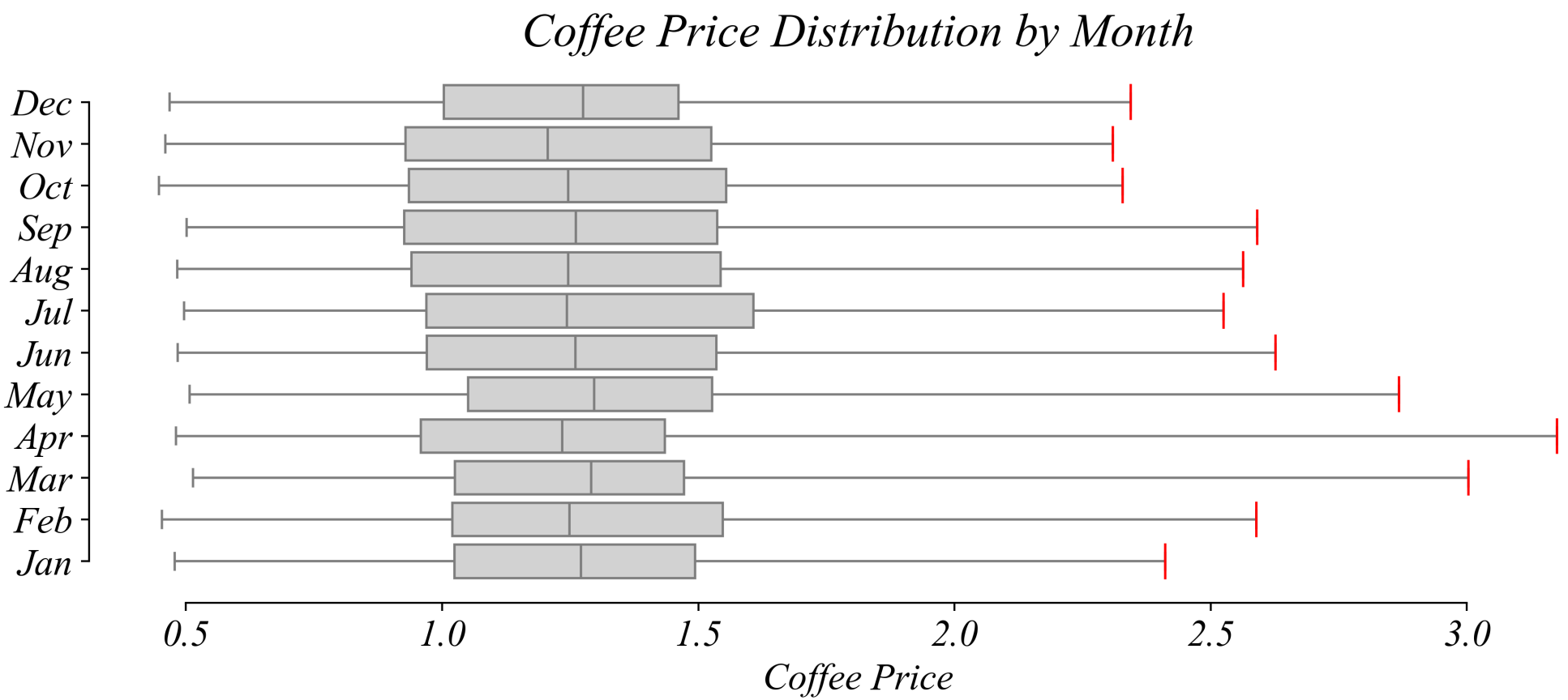
Seasonality: Monthly Boxplots

In which month was the record highest price set?



Seasonality: Monthly Boxplots

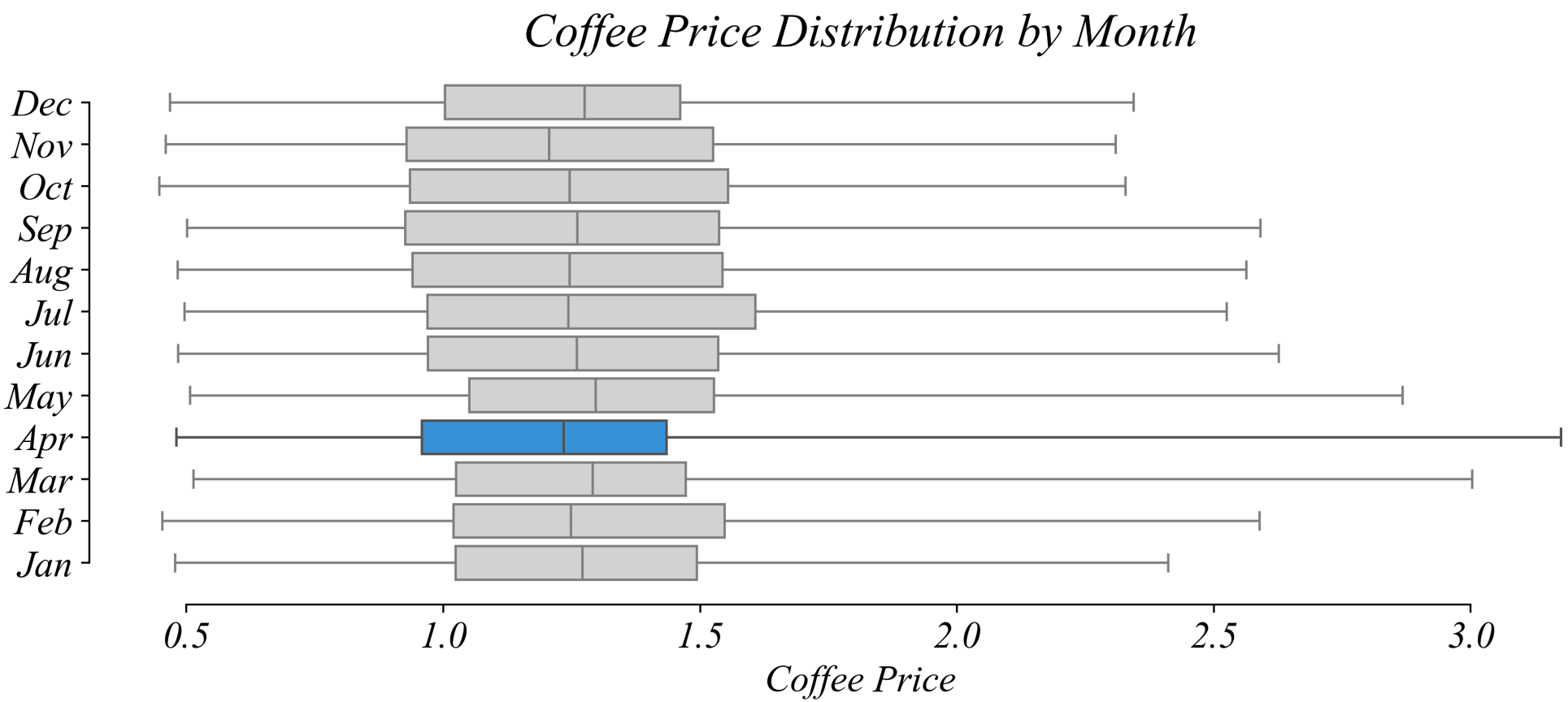
In which month was the record highest price set?



> *look at the maximums*

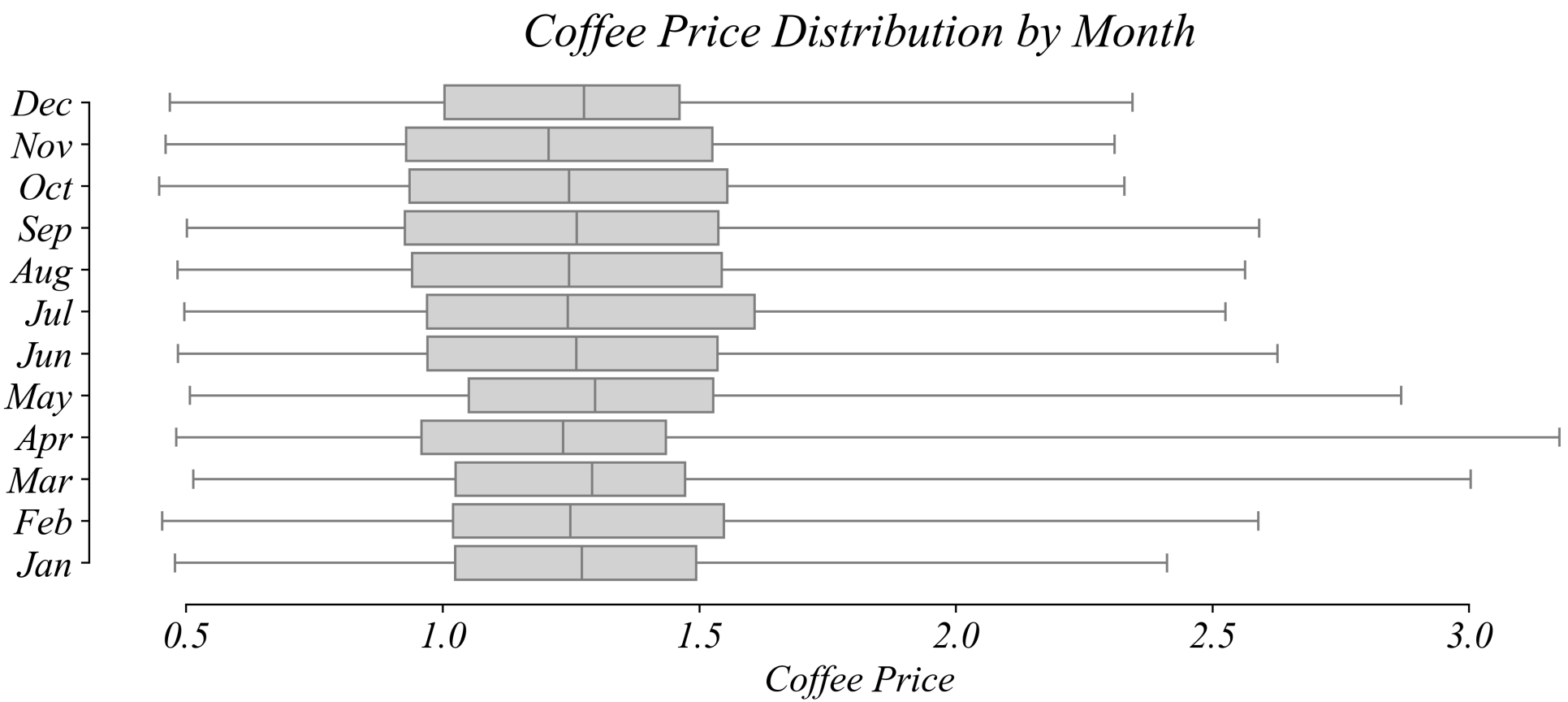
Seasonality: Monthly Boxplots

In which month was the record highest price set?



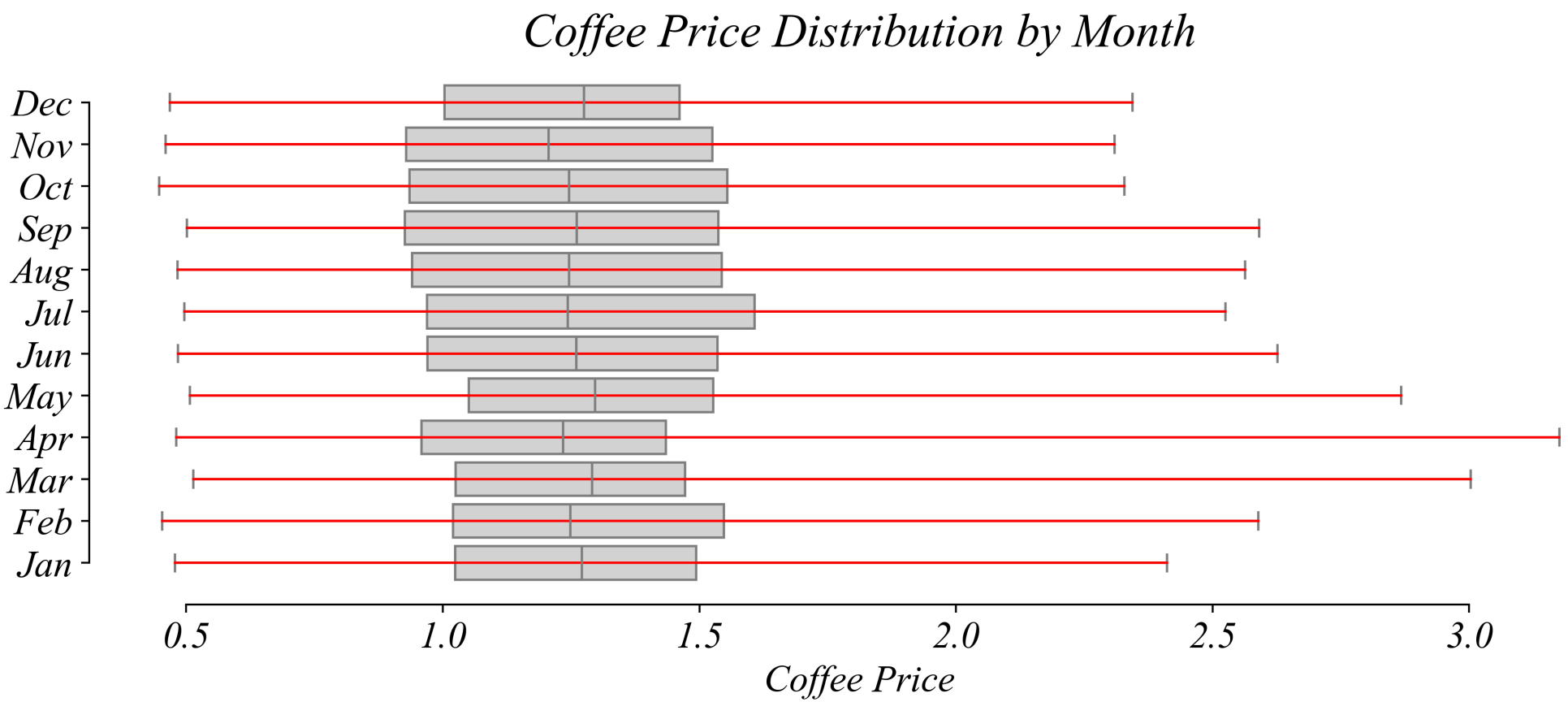
Seasonality: Monthly Boxplots

In which season are prices most spread out?



Seasonality: Monthly Boxplots

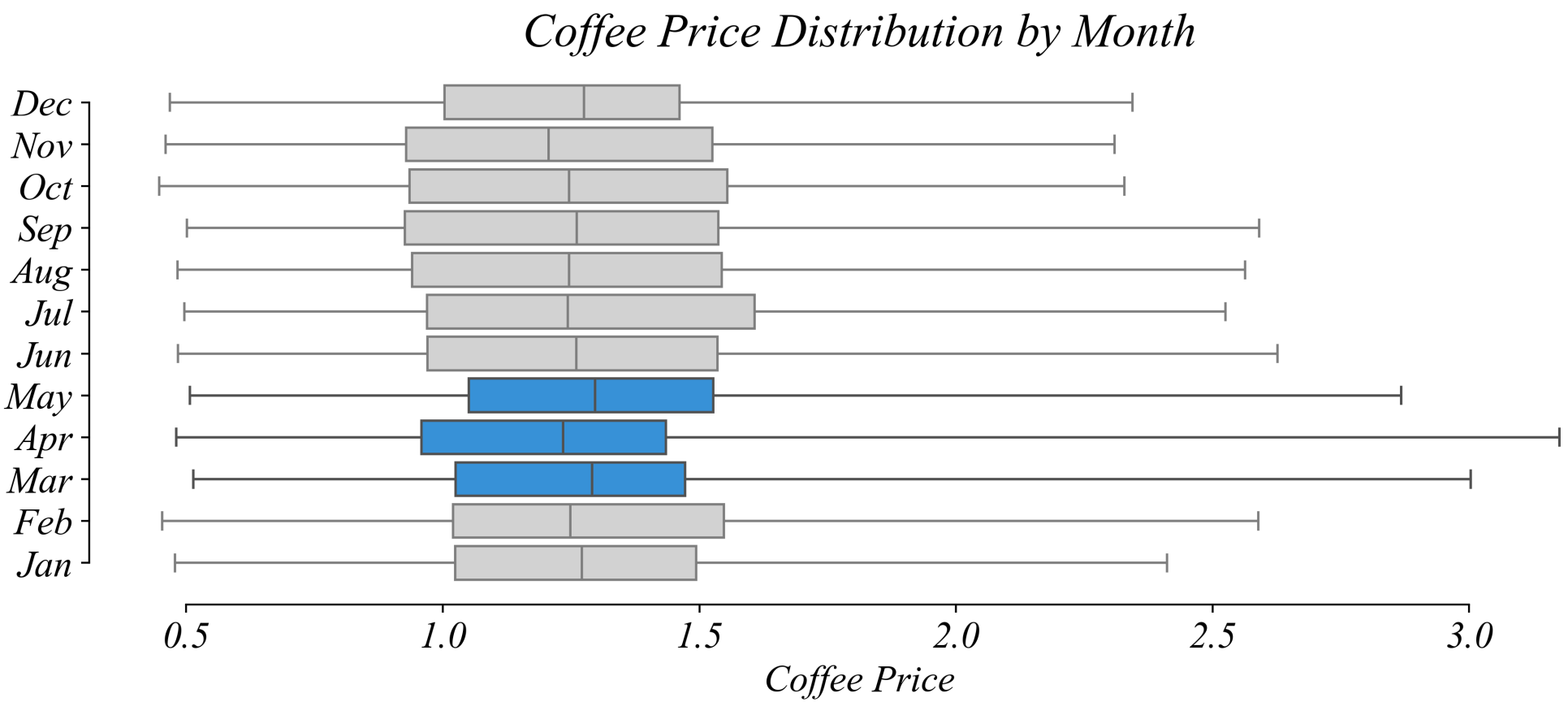
In which season are prices most spread out?



> *look at the ranges*

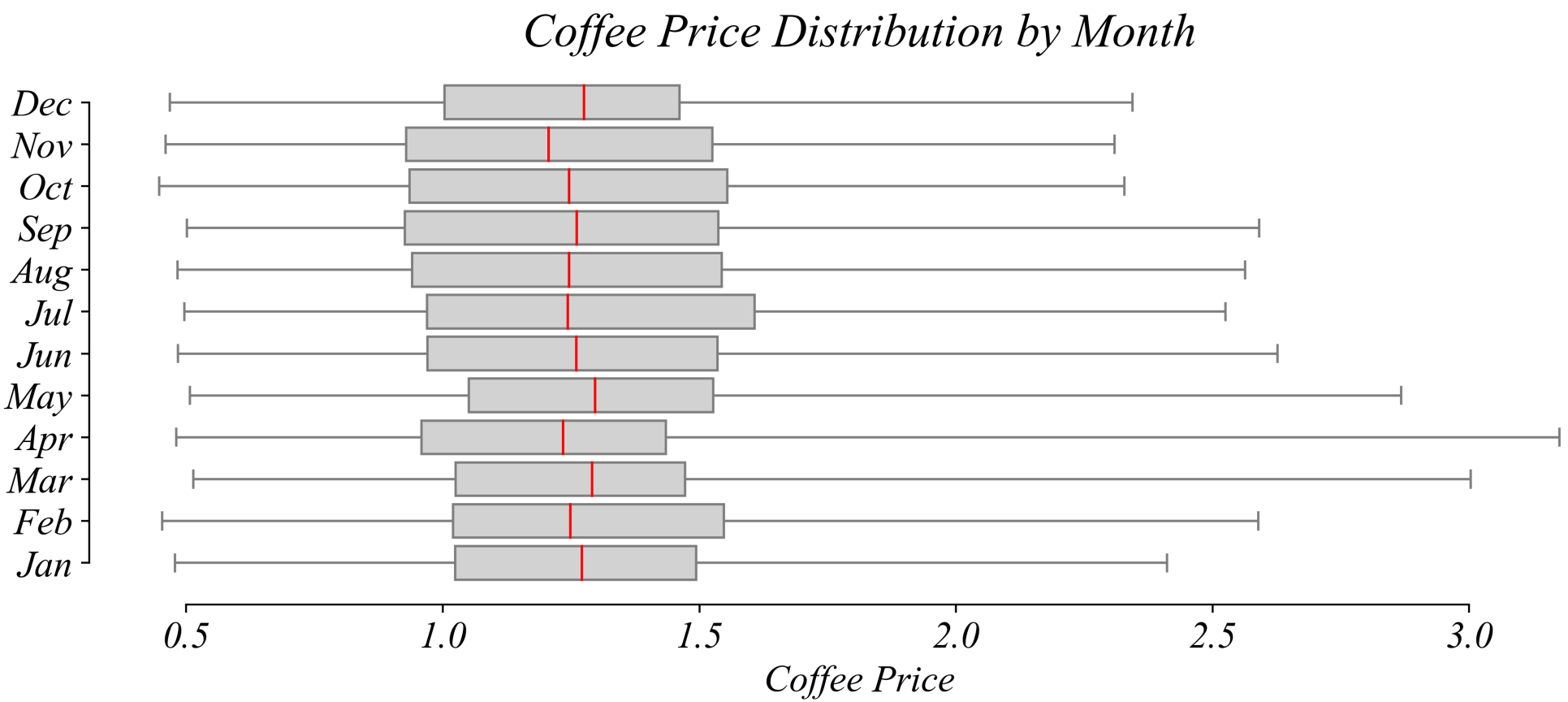
Seasonality: Monthly Boxplots

In which season are prices most spread out?



Seasonality: Multi-Boxplot

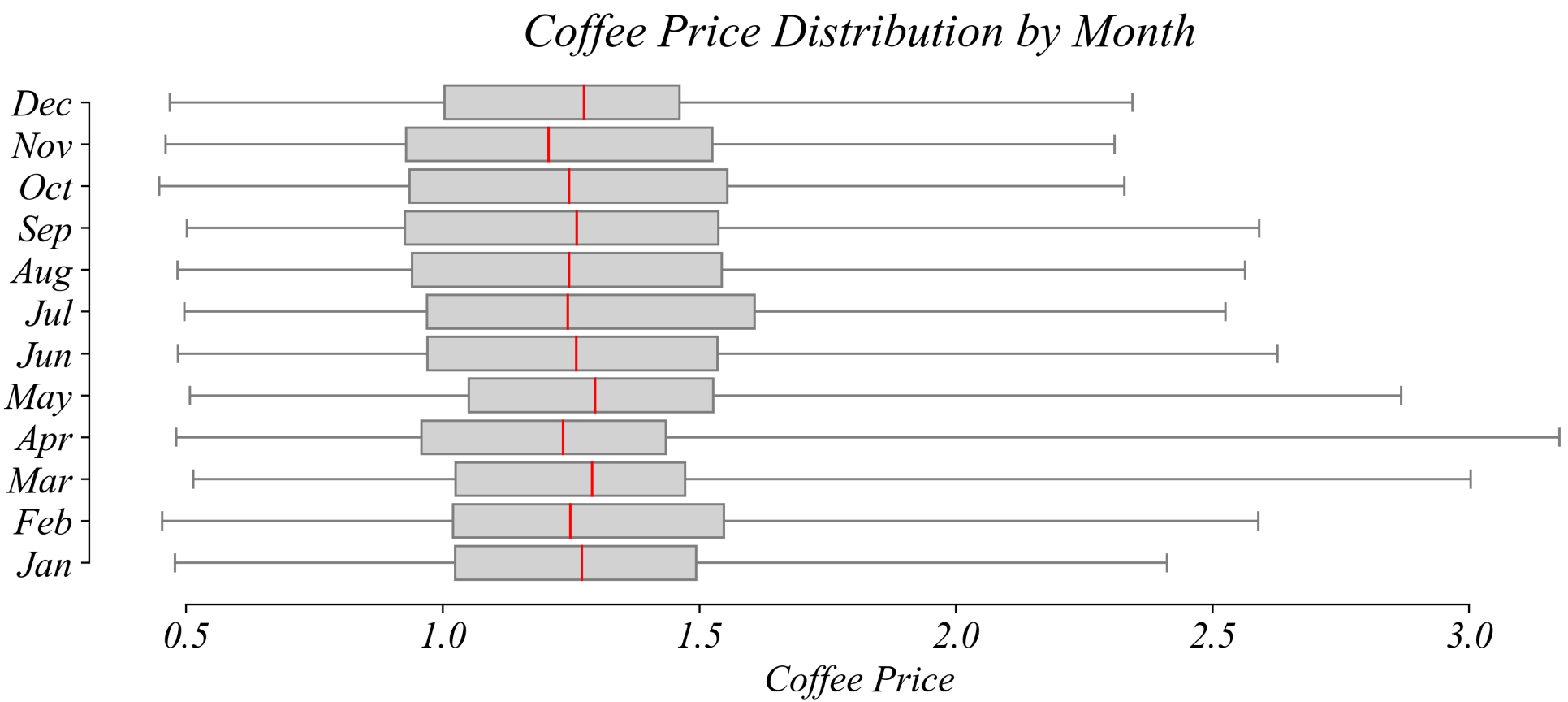
What is the trend in median price?



> *look at the medians…*

Seasonality: Multi-Boxplot

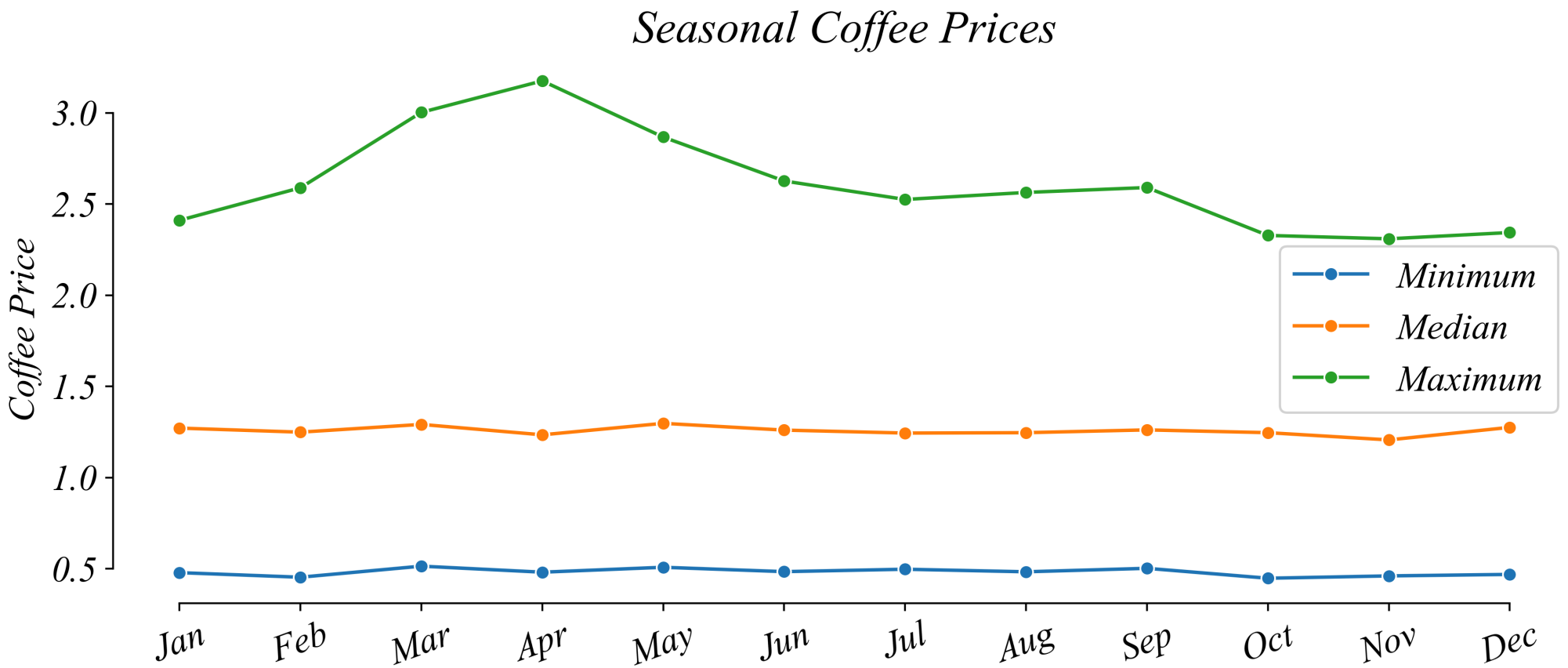
What is the trend in median price?



> *look at the medians… pretty difficult to see*

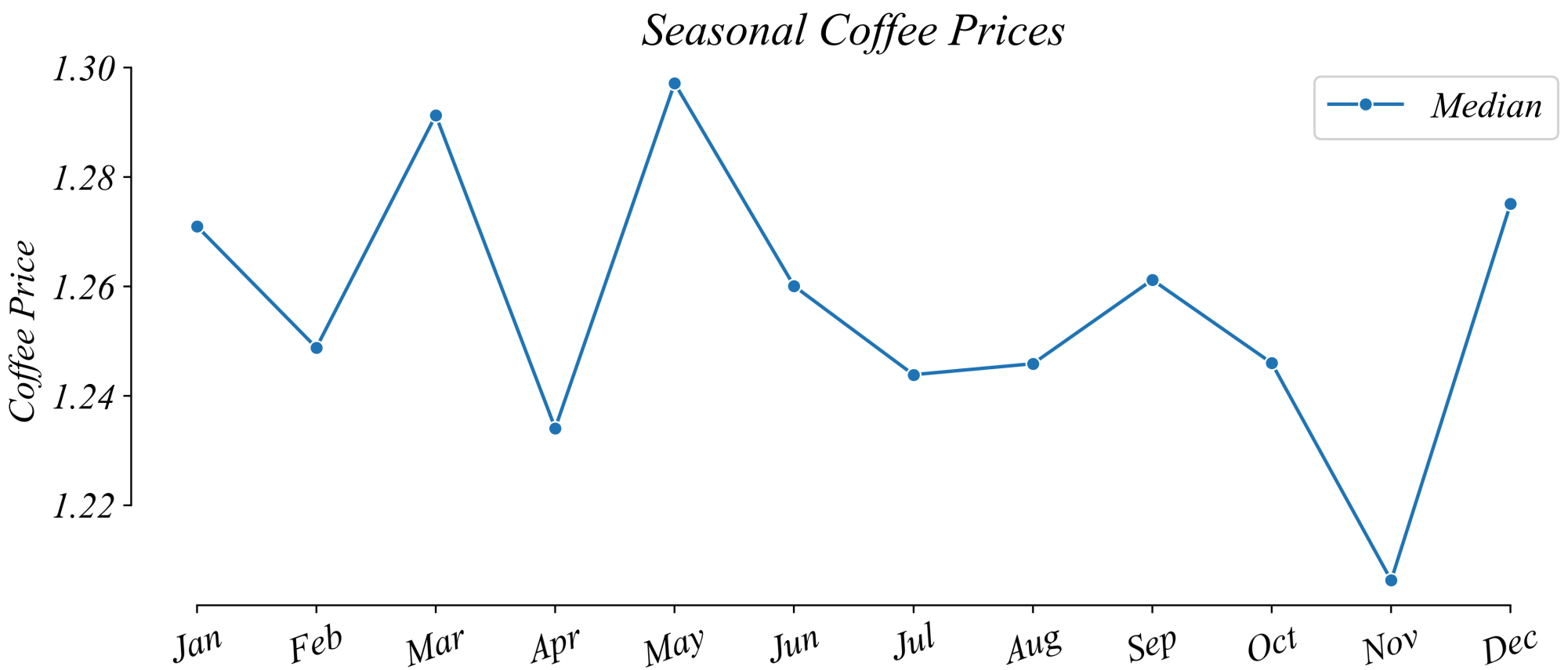
Seasonality: Quartile Lineplot

What is the trend in median price?



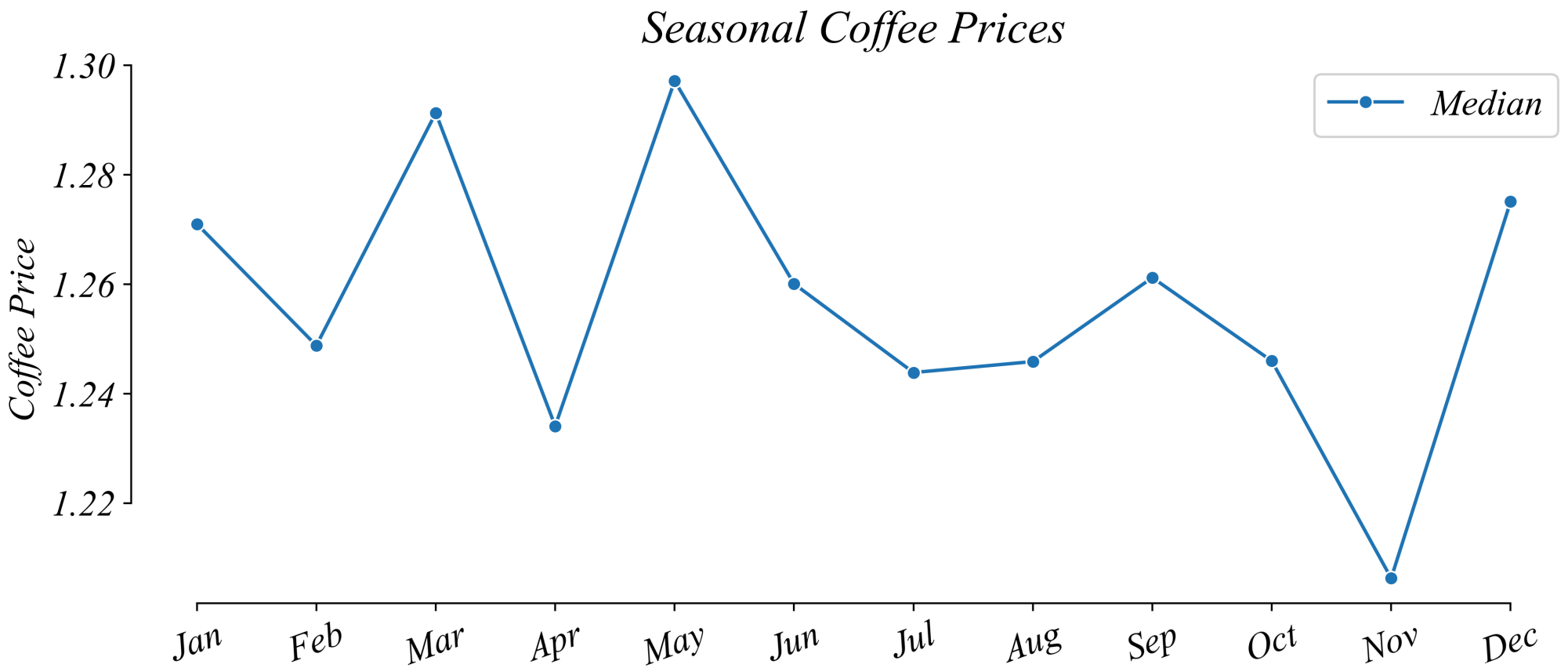
Seasonality: Quartile Lineplot

What is the trend in median price?



Seasonality: Quartile Lineplot

What is the difference between the largest and the smallest median price per pound?



> *something like $1.30 - $1.21 = $0.09*

Timeseries: Summary

Linegraphs show trends; multi-boxplots show between-period patterns.

- *Use a **linegraph** to show a numerical variable through time.*
- *Highlight changes in a linegraph using **shading**.*
- *Use a **multi-boxplot** to show the distribution between multiple periods.*

Exercise 1.3: Seasonality

Lets use a multi-boxplot to examine the seasonal patterns of coffee prices.

- ***Data:*** `Coffee_Prices.csv`

Exercise 1.3: Seasonality

```
1 # Multi-Boxplot
2 sns.boxplot(prices, y='month', x='price', whis=(0,100))
```