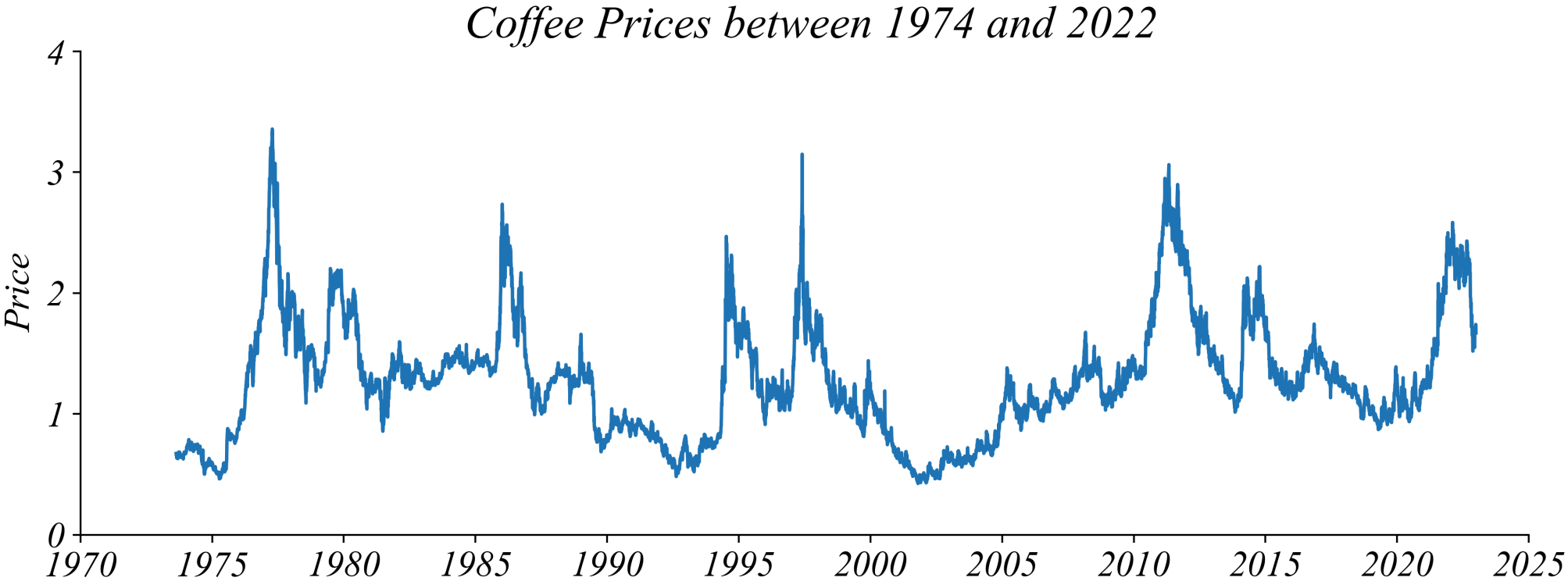# ECON 0150 | Economic Data Analysis

The economist's data analysis pipeline.

Part 2.2 | Transforming Data

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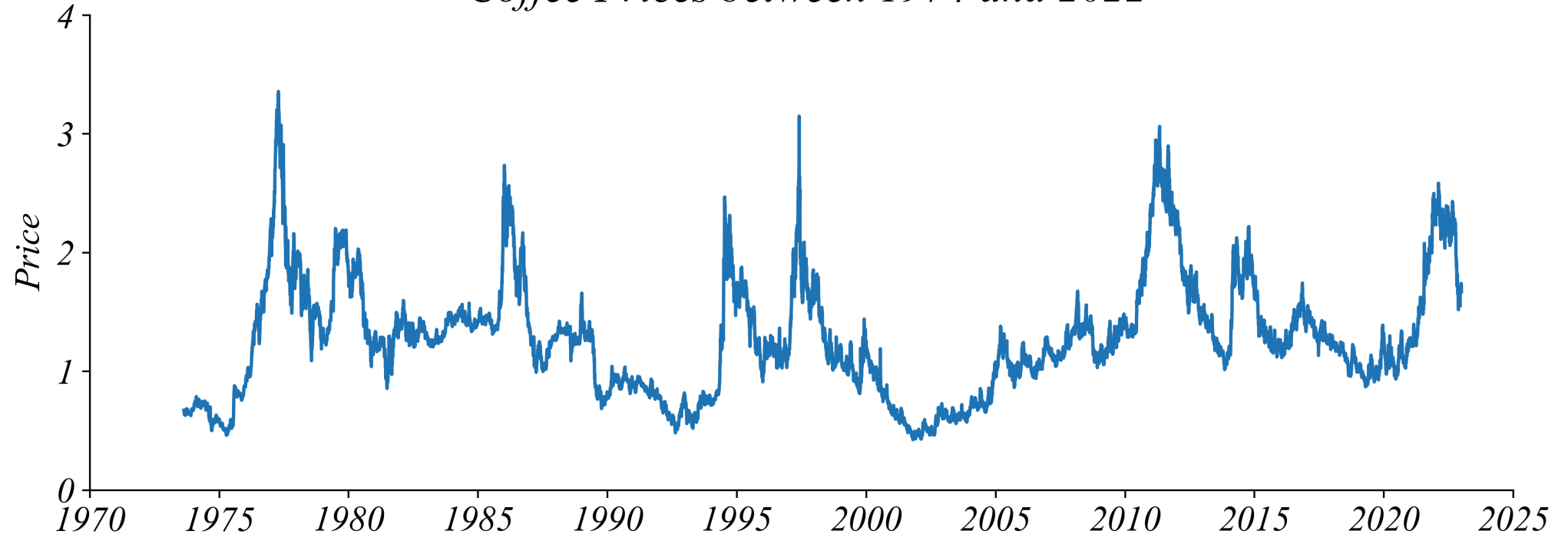
*Do you notice a **trend** in price?*

Coffee Prices between 1974 and 2022

Coffee Prices

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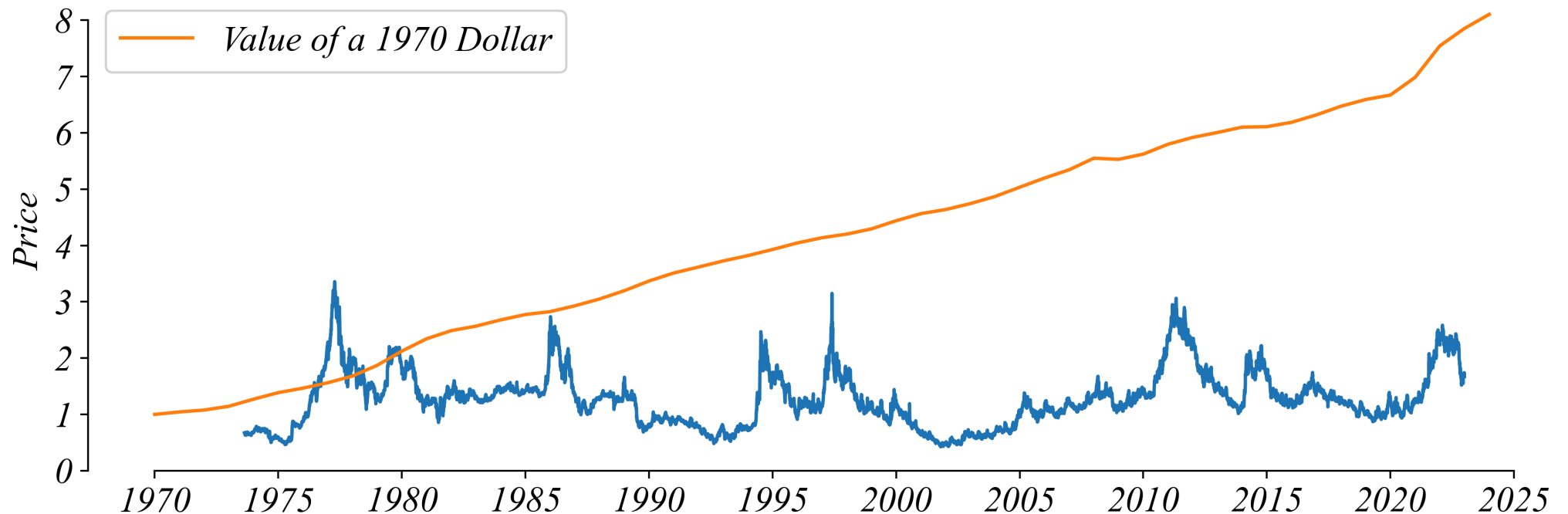

Coffee Prices between 1974 and 2022

> *was coffee about as expensive in 1980 as it is today?*

Coffee Prices: Nominal vs Real Prices

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

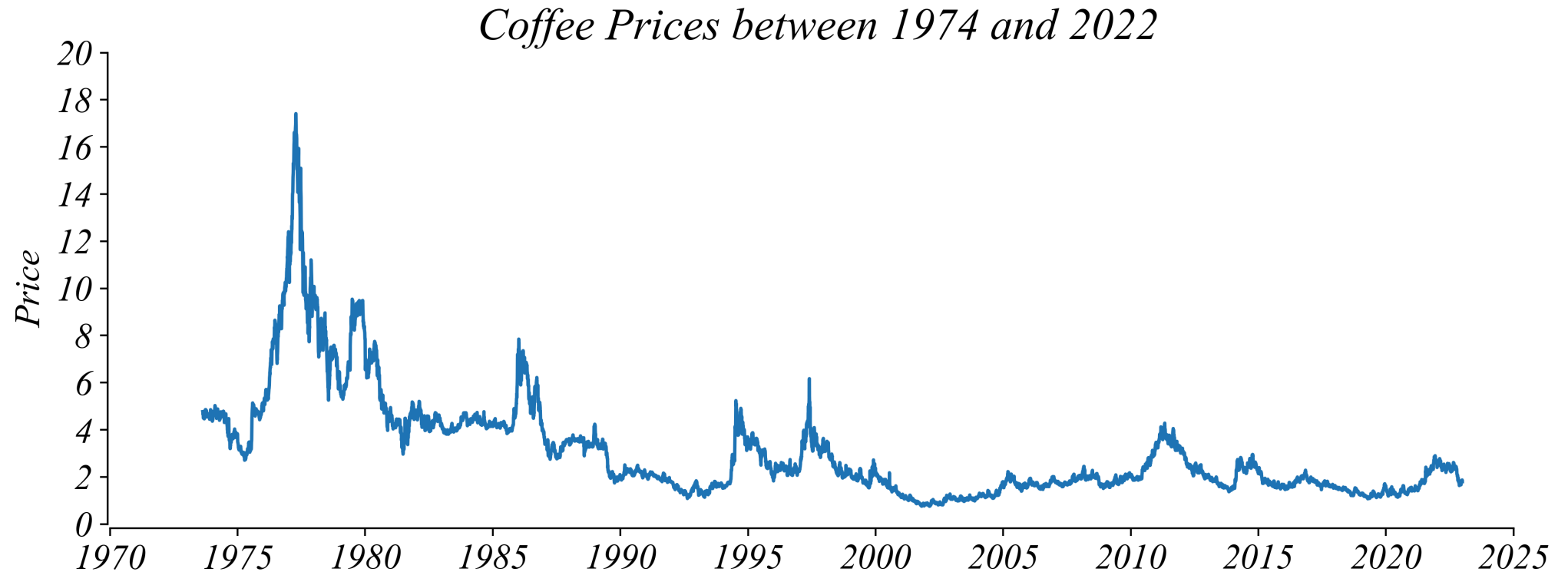

Coffee Prices and CPI between 1974 and 2022

> *no! a dollar today is worth much less than in 1980!*

> *adjusting for inflation makes the picture clearer*

Real Coffee Prices: Adjust For Inflation

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

Coffee Prices between 1974 and 2022

> *prices have dropped a lot since 1970 and have been stable since 2000*

Exercise 2.2 | Real Price Adjustment

Is there a trend in the real price of coffee?

Lets transform coffee prices from nominal dollars to real dollars.

- **Data:** *Coffee_Prices_CPI.csv*

Starbucks' Global Server Capacity

How many shops are opening at any given time?

- *Starbucks manages many shops globally and needs to maintain server capacity for all of them around the clock.*

- *Starbucks has a massive operation to make sure their shops are able to open every morning.*

- *Lets investigate how many coffee shops are opening at any hour of the day.*

Opening Times: Starbucks' Global Capacity

How many shops are opening at any given time?

Looking at the data is a good place to start.

	country	open	close	GMT
0	HK	8	22	8
1	HK	7	22	8
2	HK	8	22	8
3	HK	8	22	8
4	HK	8	20	8

>*as is common, it's difficult to understand the raw data on its own*

Opening Times

Lets start by looking at what times shops open in local time.
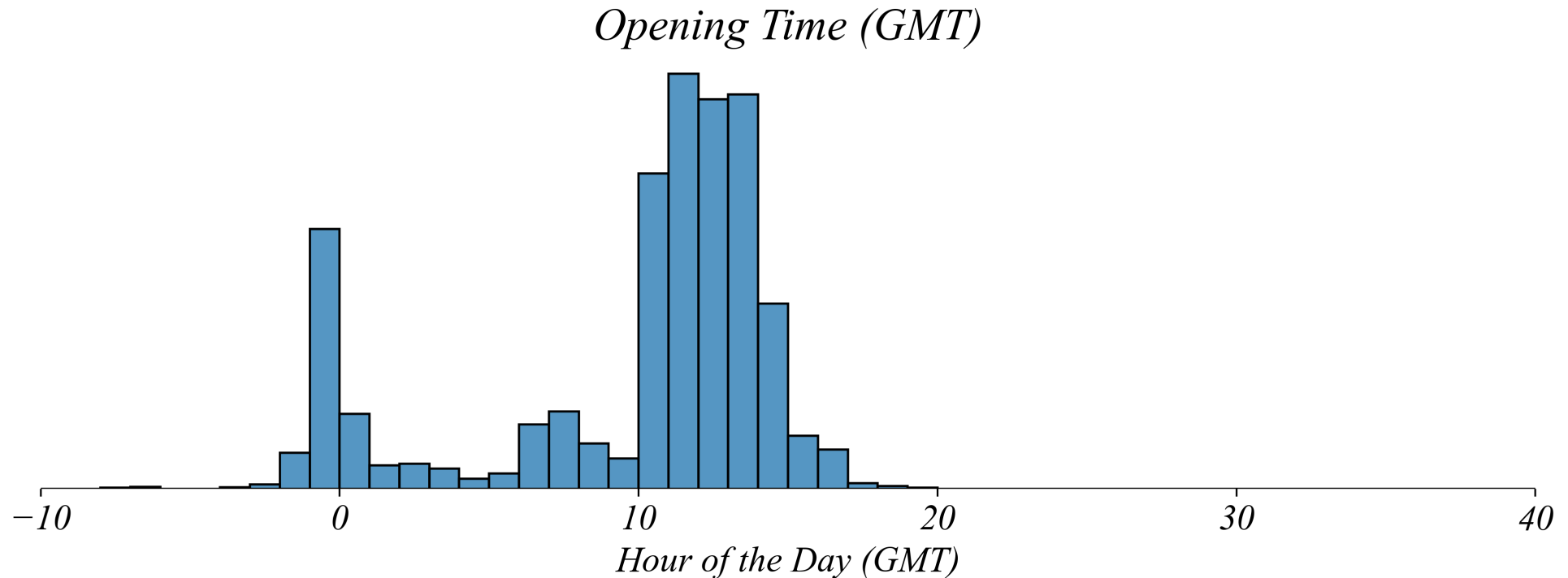
Opening Time

Hour of the Day

> but does this tell us how many shops are opening at one time?

Opening Times: Standardize by GMT

What times do shops open in GMT?

Lets standardize all times in *Greenwich Mean Time* (GMT).

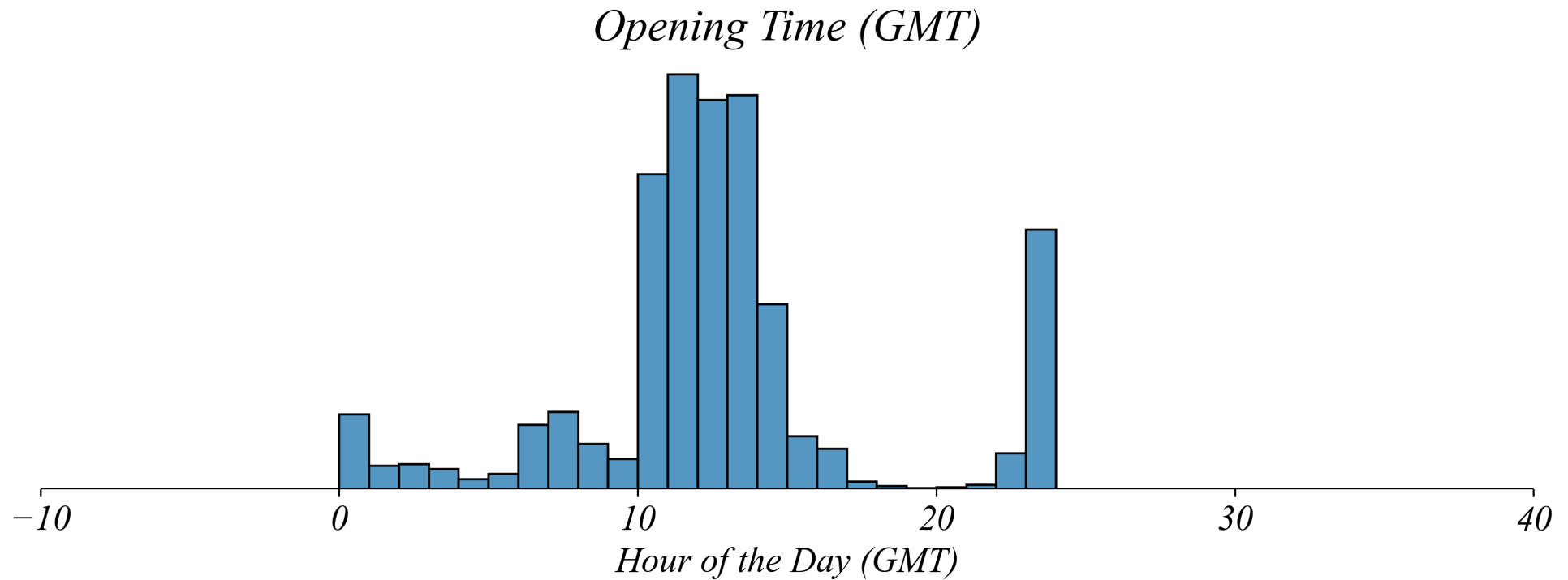*Opening Time (GMT)*

Hour of the Day (GMT)

> *what do the negative values mean?*

> *hour -1 (1 hour before GMT midnight) is the same as opening at hour 23*
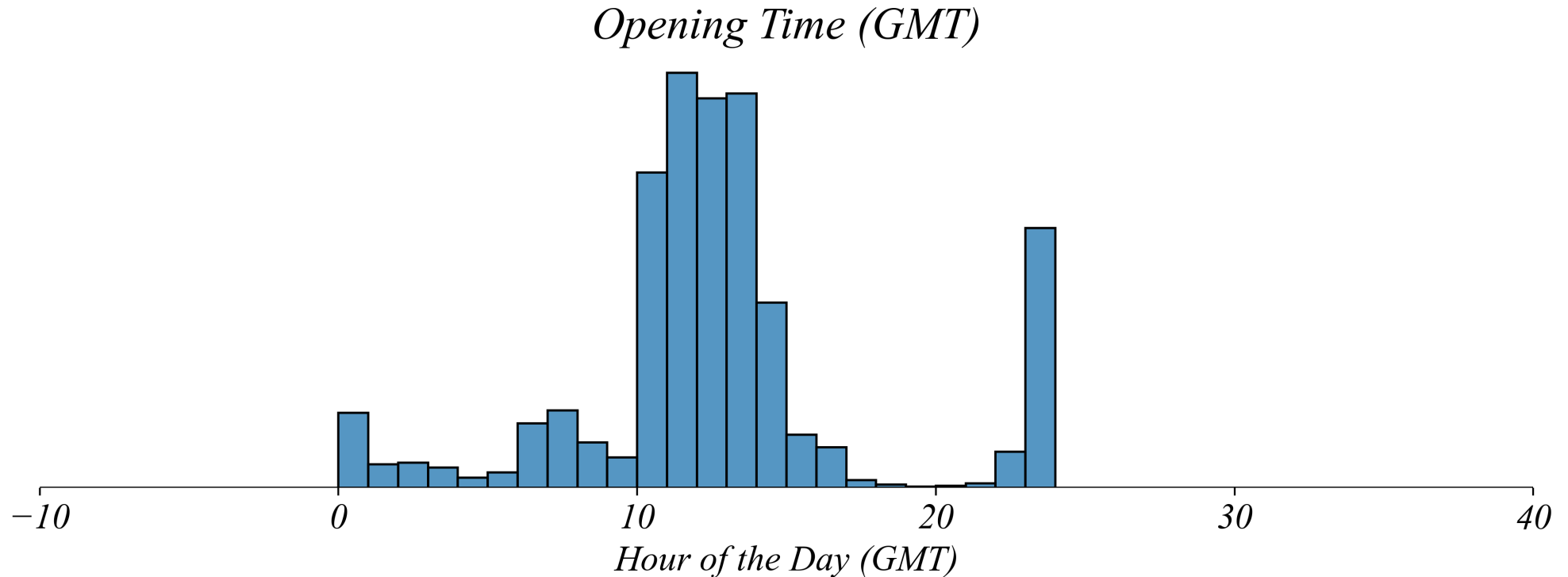
Opening Times: Normalize to 24 Hours

Normalize the negative values to 24 hours.

Lets add 24 if the number is negative.

Opening Time (GMT)

Hour of the Day (GMT)

Opening Times: Standardizing Hours

How many shops are opening at any given time?

Opening Time (GMT)

Hour of the Day (GMT)

> *a small bump during morning in Europe*

> *a huge spike during morning in the Americas*

> *a smaller spike during morning in Asia*

Exercise 2.2 | Starbucks' Global Capacity

How many shops are open at any given time?

- *Starbucks manages many shops globally and needs to maintain server capacity for all of them around the clock.*

- *We want to investigate how many coffee shops are open at any given hour to better understand server loads and Starbucks' global capacity needs.*

- *It's also just pretty interesting.*

Exercise 2.2 | Starbucks' Global Capacity
How many shops are opening at any given time?

Looking at the data is a good place to start.

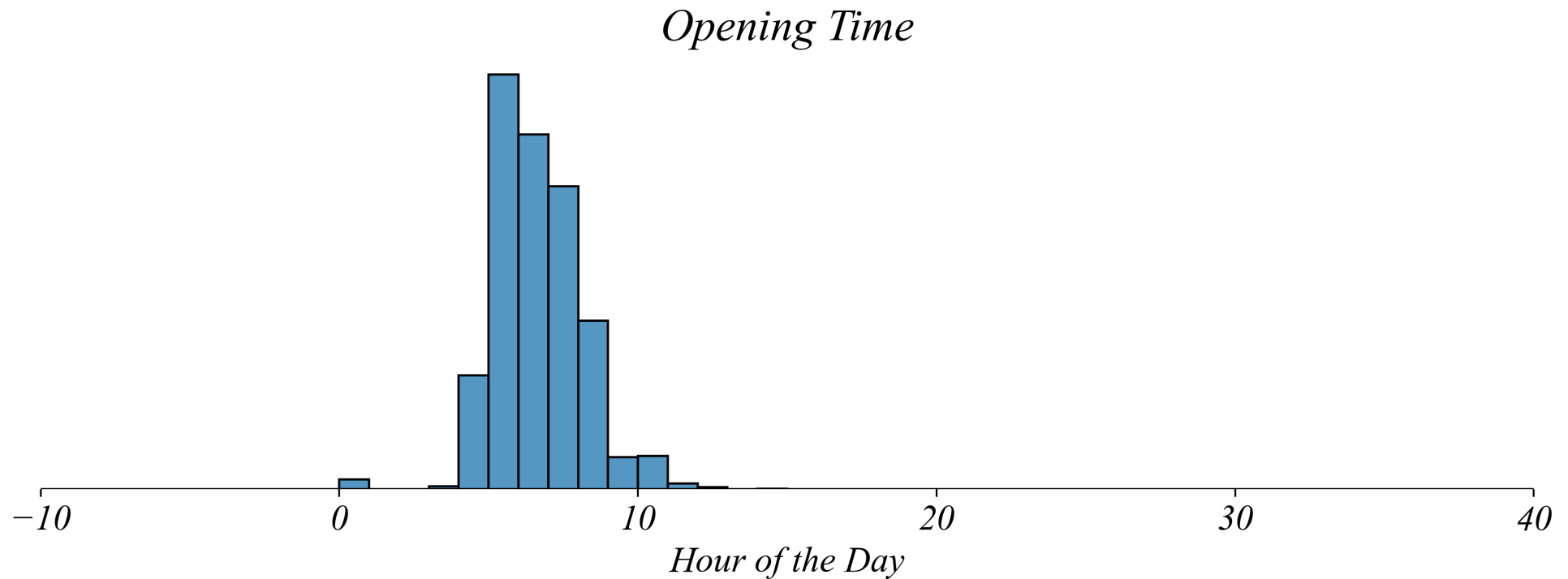|	country	open	close	GMT
0	HK	8	22	8
1	HK	7	22	8
2	HK	8	22	8
3	HK	8	22	8
4	HK	8	20	8

>as is common, it's difficult to understand the raw data on its own

Opening Times

What times do shops open in their local times?

Lets start by looking at what times shops open in local time.

```
1  # Histogram of opening times
2  sns.histplot(hours, x='open')
```
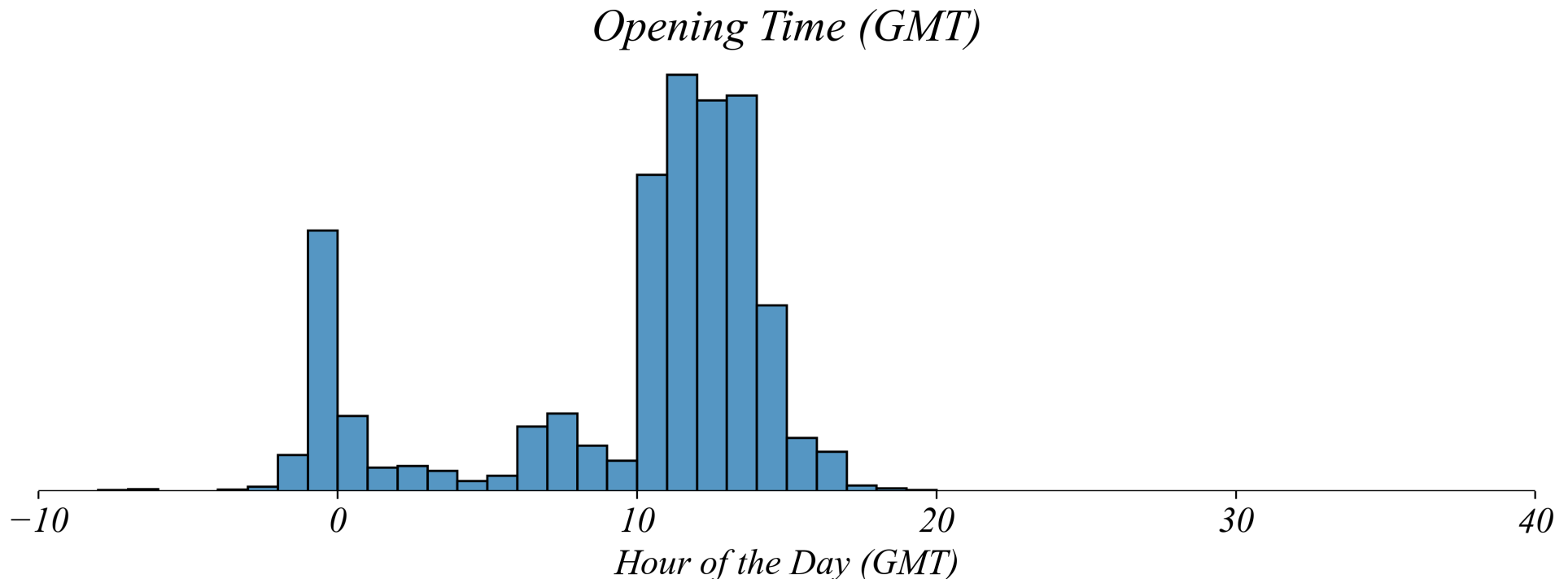

Opening Time

Opening Times: Standardize by GMT

What times do shops open (GMT)?

Lets standardize all times in *Greenwich Mean Time* (GMT).

```
1  # Normalize to GMT
2  hours['OpenGMT'] = hours['open'] - hours['GMT']
3
4  # Histogram of opening times (GMT)
5  sns.histplot(hours, x='OpenGMT')
```

Opening Time (GMT)

Hour of the Day (GMT)

Opening Times: Standardizing Hours

Normalize the negative values to 24 hours.

Lets add 24 if the number is negative.

```python
# Normalize to 24 hours
hours['OpenGMT24'] = hours['OpenGMT'].mod(24)

# Histogram of opening times (GMT, 24)
sns.histplot(hours, x='OpenGMT24')
```

Opening Time (GMT)

Hour of the Day (GMT)