

## Lesson 16 Summary: Correlation

Data Science with Python – Key Concepts

Data Science Program

## Correlation Analysis



### Key Insights:

`corrcoef()` | rolling corr | confounders | causation

*Correlation does not imply causation!*

---

Correlation measures linear relationship between variables

## Linear relationship measure:

- Range:  $[-1, 1]$
- $+1$ : Perfect positive linear relationship
- $-1$ : Perfect negative linear relationship
- $0$ : No linear relationship

## Python:

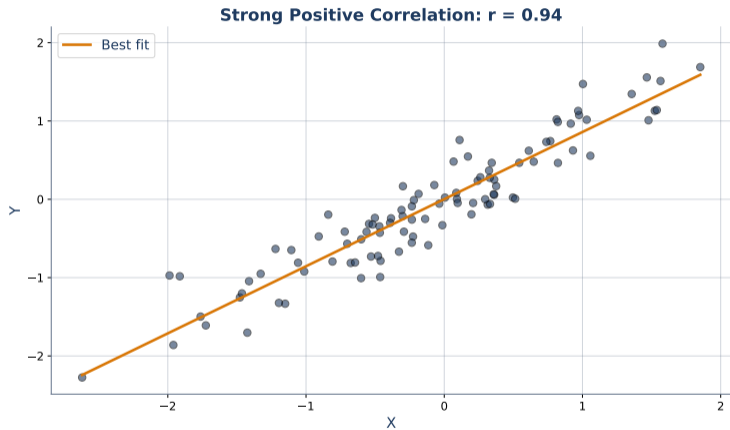
```
df['A'].corr(df['B'])
```

```
np.corrcoef(a, b)[0, 1]
```

---

Pearson assumes linear relationship and normality

# Correlation Strength



Visual inspection helps interpret correlation values

## Rank-based correlation (non-parametric):

- Measures monotonic (not just linear) relationship
- Robust to outliers
- Use when data is non-normal or ordinal

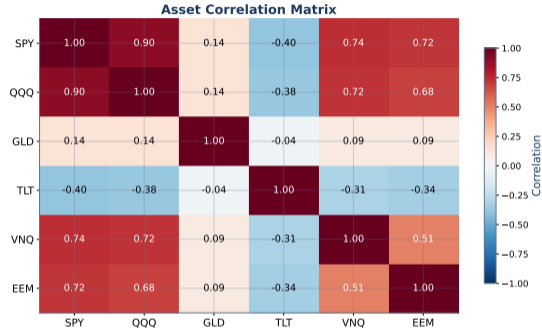
## Python:

```
from scipy import stats  
stats.spearmanr(a, b)
```

---

**Spearman is often more appropriate for financial data**

# Correlation Matrix Heatmap



Heatmaps visualize pairwise correlations efficiently

## Common mistakes:

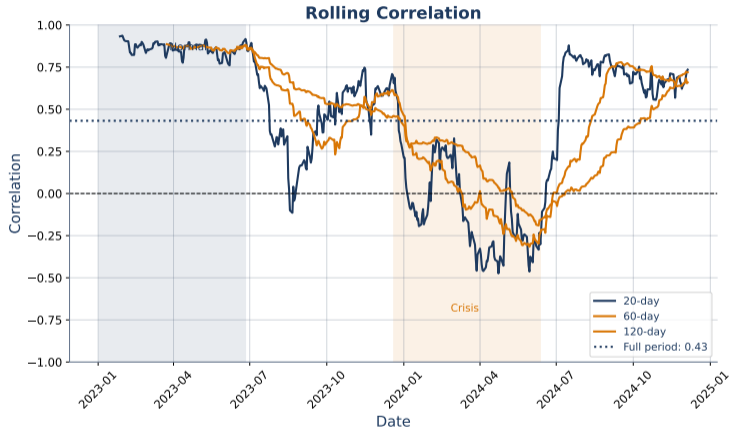
- **Confounders:** Third variable drives both
- **Non-linear:** Pearson misses curved relationships
- **Outliers:** Single points can distort correlation
- **Spurious:** Random coincidence in time series

**Golden rule:** Correlation  $\neq$  Causation

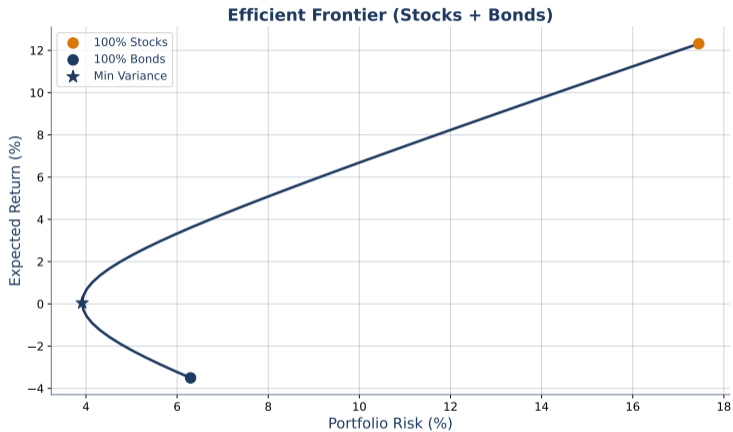
---

Always visualize data; don't rely on correlation alone

# Rolling Correlation



Correlations change over time – monitor dynamically



Low correlation enables portfolio risk reduction

### Essential Correlation Operations:

Operation	Syntax
Pearson (two cols)	<code>df['A'].corr(df['B'])</code>
Full matrix	<code>df.corr()</code>
Spearman	<code>df.corr(method='spearman')</code>
Rolling	<code>df['A'].rolling(60).corr(df['B'])</code>
Heatmap	<code>sns.heatmap(df.corr())</code>
Interpretation	
$ r  > 0.7$	Strong
$0.3 <  r  < 0.7$	Moderate
$ r  < 0.3$	Weak

---

Use correlation for diversification and risk management