

INCASE

แนวคิดหลักของการวิเคราะห์

ในการวิเคราะห์นี้ เราต้องการทำความเข้าใจว่า "ดัชนีชี้วัดความกว้างของตลาด" (**Market Breadth Indicators**) ต่างๆ มีความสัมพันธ์อย่างไรกับ "ดัชนี SET100" และความสัมพันธ์เหล่านั้นมีนัยสำคัญทางสถิติมากน้อยแค่ไหน

ทำไมถึงต้องวิเคราะห์แบบนี้?

- **ความสัมพันธ์ (Correlation):** บอกเราว่าดัชนีชี้วัดความกว้างของตลาดเคลื่อนไหวไปในทิศทางเดียวกัน หรือทิศทางตรงข้ามกับ SET100 มากน้อยเพียงใด ค่า **Correlation** ใกล้ +1 หมายถึงเคลื่อนที่ไปในทิศทางเดียวกันอย่างแข็งแกร่ง, ใกล้ -1 หมายถึงเคลื่อนที่สวนทางกันอย่างแข็งแกร่ง, และใกล้ 0 หมายถึงแทบไม่มีความสัมพันธ์กันเลย
- **นัยสำคัญทางสถิติ (Statistical Significance) หรือ p-value:** บอกเราว่าความสัมพันธ์ที่เราเห็นนั้น "ไม่ใช่เรื่องบังเอิญ" ถ้าค่า **p-value** ยิ่งน้อย (เช่น น้อยกว่า 0.05) นั่นหมายความว่าความสัมพันธ์ที่เราเห็นนั้นมีความน่าเชื่อถือสูง ไม่ใช่แค่เรื่องบังเอิญ

สรุปคือ: เรากำลังมองหาดัชนีชี้วัดความกว้างของตลาดที่มี "ความสัมพันธ์ที่ดี" (เช่น ค่า **Correlation** สูง) และมีความสัมพันธ์นั้น "น่าเชื่อถือทางสถิติ" (เช่น ค่า **p-value** ต่ำ) เพื่อใช้เป็นเครื่องมือในการวิเคราะห์ตลาด SET100

อธิบายวิธีการคำนวณ p-value และ Correlation

ในโค้ดที่เราเขียนขึ้นมา มีการคำนวณค่าเหล่านี้ดังนี้ครับ:

1. การเตรียมข้อมูล (Data Preparation):

- เรามี **daily_scores_df** ซึ่งเป็น **DataFrame** ที่เก็บค่าของดัชนีชี้วัดความกว้างของตลาดแต่ละตัวในแต่ละวัน
- เรามี **set100_returns** ซึ่งเป็นการคำนวณ "ผลตอบแทนรายวัน" ของดัชนี SET100 (ได้มาจากการเปลี่ยนแปลงของราคาปิดในแต่ละวัน)
- สิ่งสำคัญ: ก่อนที่เราจะคำนวณ เรามีการจัดการกับค่าว่าง (**missing values**) โดยการใช้ **.fillna(0)** ซึ่งหมายถึง ถ้าข้อมูลของดัชนีชี้วัดใด หรือผลตอบแทน SET100 ไม่มีค่าในวันนั้น เราจะใส่ค่า 0 เข้าไปแทน ซึ่งเป็นวิธีหนึ่งในการจัดการข้อมูลเพื่อให้สามารถคำนวณได้โดยไม่มีข้อผิดพลาด

2. การวนลูปคำนวณ (Looping and Calculation):

- โค้ดจะทำการ "วนลูป" ไปทีละตัวของ **indicator_names** หรือพูดง่ายๆ คือ วนไปที่ละดัชนีชี้วัดความกว้างของตลาดที่เราสนใจ

- ในแต่ละรอบของการวนลูป:
 - เราจะนำข้อมูลของ "ดัชนีชี้วัดนั้นๆ" (`daily_scores_df[indicator].fillna(0)`)
 - มาเปรียบเทียบกับข้อมูล "ผลตอบแทนของ SET100" (`set100_returns.fillna(0)`)
 - โดยใช้ฟังก์ชัน `stats.pearsonr()` จากไลบรารี `scipy.stats`

3. `stats.pearsonr()` คืออะไร?

- ฟังก์ชันนี้คือหัวใจสำคัญในการคำนวณของเราครับ
- `pearsonr` ย่อมาจาก "**Pearson Correlation Coefficient**" ซึ่งเป็นวิธีมาตรฐานในการวัดความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรสองตัว
- ผลลัพธ์ที่ได้จาก `stats.pearsonr()`:
 - **correlation:** นี่คือการวัดความสัมพันธ์ (Pearson Correlation Coefficient, r) ที่เราได้อธิบายไปแล้ว ยิ่งใกล้ $+1$ หรือ -1 แสดงว่ามีความสัมพันธ์กันมาก
 - **p_value:** นี่คือการวัดการทดสอบสมมติฐาน เราตั้งสมมติฐานว่า "ไม่มีความสัมพันธ์ระหว่างตัวแปรทั้งสอง" (หรือที่เรียกว่า **Null Hypothesis**) ค่า **p-value** ที่ต่ำมากๆ หมายความว่า "โอกาสที่จะเห็นความสัมพันธ์แบบนี้ หากสมมติฐานที่ว่าไม่มีความสัมพันธ์นั้นเป็นจริงนั้นมีน้อยมาก" ดังนั้น เราจึงปฏิเสธสมมติฐานที่ว่าไม่มีความสัมพันธ์ และสรุปว่า "มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ"
 - เกณฑ์มาตรฐานที่เราใช้คือ: ถ้า $p\text{-value} < 0.05$ (หรือ 5%) เราจะถือว่าความสัมพันธ์นั้น "มีนัยสำคัญทางสถิติ" และเชื่อถือได้

4. การจัดเก็บและแสดงผล (Storing and Displaying Results):

- ค่า **correlation** และ **p_value** ที่ได้จากการคำนวณในแต่ละรอบ จะถูกเก็บไว้ในลิสต์ **correlations** และ **p_values** ตามลำดับ
- จากนั้น ข้อมูลทั้งหมดจะถูกนำมาสร้างเป็นตาราง **significance_results** ซึ่งทำให้เราเห็นภาพรวมของ **Indicator** แต่ละตัว พร้อมค่า **Correlation**, **P-value** และสถานะว่า **"Significant"** หรือไม่

ที่เราเลือกใช้เกณฑ์ **p-value** ที่ **0.05** (หรือ 5%) เป็นค่ามาตรฐานในการตัดสินว่าผลลัพธ์มีนัยสำคัญทางสถิติหรือไม่นั้น ไม่ได้มีเหตุผลตายตัวแบบคณิตศาสตร์เป๊ะๆ แต่เป็น ข้อตกลงร่วมกัน (**convention**) ในวงการวิจัยและสถิติมาอย่างยาวนานครับ

จากครั้งก่อนนะครับที่ได้ไปplot กราฟของ **score** ของแต่ละ อินดิเคเตอร์ เทียบกับ **set 100**

ที่ดอลล่าได้ให้ไปหาแต่ละ **indicators** นั้น ดีเลย์แค่ไหนจาก ตัว **set100** และก็หา **correlation** กับ **p value** ครับ

"ในการวิเคราะห์ครั้งนี้ เราต้องการหาว่า**indicators**ที่วัด**market breadth** ที่เราสร้างขึ้นนั้น มีความสามารถในการอธิบายหรือมีความสัมพันธ์กับ **SET100** ได้มากน้อยเพียงใด และความสัมพันธ์เหล่านั้นมีความน่าเชื่อถือทางสถิติหรือไม่"

"ผมได้ใช้วิธีการทางสถิติที่เรียกว่า '**Pearson Correlation**' เพื่อวัด 'ความสัมพันธ์' ครับ โดยเรียกใช้**library Scipy.stats** ฟังก์ชันนี้จะบอกเราว่า เมื่อดัชนีชี้วัดตัวหนึ่งเพิ่มขึ้นหรือลดลง **SET100** มีแนวโน้มที่จะเคลื่อนไหวไปในทิศทางเดียวกัน สวนทางกัน หรือไม่เกี่ยวข้องเลย ยิ่งค่า **Correlation** ใกล้ **+1** หรือ **-1** เท่าไหร่ ก็แปลว่ามีความสัมพันธ์ที่แข็งแกร่งมากเท่านั้น"

นอกจากการวัดความสัมพันธ์แล้ว เรายังได้คำนวณ 'ค่า **p-value**' ซึ่งเป็นตัวชี้วัด 'นัยสำคัญทางสถิติ' ครับ ค่า **p-value** นี้จะบอกเราว่า ความสัมพันธ์ที่เราเห็นนั้น 'เป็นเรื่องบังเอิญ' หรือ 'มีความน่าเชื่อถือ' ถ้าค่า **p-value** ที่คำนวณได้ 'น้อยกว่า **0.05**' นั้นหมายความว่า ความสัมพันธ์ที่เราค้นพบนั้น มีนัยสำคัญทางสถิติสูง และไม่ได้เกิดขึ้นจากความบังเอิญครับ

"ดังนั้น จากกราฟที่เรานำเสนอในส่วนบน จะเห็นว่าแท่งสีเขียวคือ**indicators**ที่มีนัยสำคัญทางสถิติ (**p-value** น้อยกว่า **0.05**) ยิ่งแท่งสูง ยิ่งมีความน่าเชื่อถือสูง ส่วนในแต่ละแท่ง เราได้ระบุค่า **Correlation (r)** ไว้ด้วย เพื่อให้เห็นถึงทิศทางและความแข็งแกร่งของความสัมพันธ์ และในส่วนล่างเป็น **Heatmap** ที่สรุปค่า **Correlation** ของทุกตัวอีกครั้งครับ

การวิเคราะห์นี้จะช่วยให้เราสามารถคัดเลือก**indicators**ที่ใช้ในการคำนวณ **market breadth**ที่มีประสิทธิภาพและน่าเชื่อถือ เพื่อนำไปใช้ในการวิเคราะห์และประกอบการตัดสินใจ ต่อไปครับ

Part correlation and p value

ส่วนที่ 1: นัยสำคัญทางสถิติ (Statistical Significance)

- **แกนตั้ง:** นัยสำคัญทางสถิติ (Statistical Significance) แสดงด้วยค่า $-\log_{10}(p\text{-value})$ โดยค่าที่สูงกว่าแสดงถึงนัยสำคัญที่มากกว่า
- **เส้นประสีดำ:** เส้นประนี้คือเกณฑ์นัยสำคัญที่ $p < 0.05$ หรือ $-\log_{10}(p\text{-value}) > -\log_{10}(0.05) \approx 1.3$
- **แถบสีเขียวอ่อน:** ตัวชี้วัดที่มีแถบสีเขียวอ่อนคือ "Significant" หมายถึงมีนัยสำคัญทางสถิติที่ $p < 0.05$
- **แถบสีแดง:** ตัวชี้วัดที่มีแถบสีแดงคือ "Not Significant" หมายถึงไม่มีนัยสำคัญทางสถิติที่ $p \geq 0.05$

ข้อสังเกตจากส่วนที่ 1:

- **Indicators** ส่วนใหญ่แสดงนัยสำคัญทางสถิติอย่างชัดเจน โดยเฉพาะ "Advance/Decline Volume Ratio" และ "Stochastic Oscillator Trend" ที่มีค่า $-\log_{10}(p\text{-value})$ สูงมาก (ประมาณ 39 และ 36 ตามลำดับ)
- ตัวชี้วัดที่อยู่ทางขวา เช่น "MACD Positive-Negative" และ "ATR Trend" แสดงนัยสำคัญทางสถิติที่ต่ำหรือไม่แสดงนัยสำคัญเลย

ส่วนที่ 2: ความสัมพันธ์กับผลตอบแทน SET100 (Correlation with SET100 Returns)

- **แกนตั้ง:** ความสัมพันธ์ (Correlation)
- **แถบสี:** แสดงค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient, r) โดยสีเข้มแสดงถึงความสัมพันธ์เชิงบวกที่สูง และสีเหลืองแสดงถึงความสัมพันธ์ที่ต่ำหรือติดลบเล็กน้อย
- **ตัวเลขในแถบ:** ค่าสัมประสิทธิ์สหสัมพันธ์ (r) ของแต่ละตัวชี้วัด

ข้อสังเกตจากส่วนที่ 2:

- indicators ที่มีนัยสำคัญทางสถิติสูงในส่วนแรก มักจะมีความสัมพันธ์เชิงบวกที่สูงกับ SET100 ด้วย เช่น "Advance/Decline Volume Ratio" ($r = 0.56$) และ "Stochastic Oscillator Trend" ($r = 0.54$)
- "New High-Low (Month)" และ "McClellan Oscillator" ก็แสดงความสัมพันธ์เชิงบวกที่แข็งแกร่งเช่นกัน ($r = 0.39$ และ $r = 0.34$)
- ตัวชี้วัดที่อยู่ทางขวา เช่น "MACD Positive-Negative" และ "ATR Trend" มีความสัมพันธ์ที่ต่ำมากหรือติดลบเล็กน้อย ($r = -0.03$ สำหรับ "MACD Positive-Negative" และ $r = 0.03$ สำหรับ "ATR Trend") ซึ่งสอดคล้องกับผลลัพธ์ด้านนัยสำคัญทางสถิติด้านบน

สรุปโดยรวม

ภาพนี้แสดงให้เห็นว่า indicators ที่ใช้วัด market breadth โดยเฉพาะอย่างยิ่ง "Advance/Decline Volume Ratio" และ "Stochastic Oscillator Trend" มีทั้งนัยสำคัญทางสถิติที่สูงและความสัมพันธ์เชิงบวกที่แข็งแกร่งกับ SET100 ซึ่งบ่งชี้ว่า indicators เหล่านี้อาจมีประโยชน์ในการวิเคราะห์และคาดการณ์ทิศทางของตลาด SET100 ในอนาคต

หลังจากที่เราทราบถึงนัยสำคัญและความสัมพันธ์ไปแล้ว ภาพนี้จะพาเรามาดูที่ 'ประสิทธิภาพในการให้สัญญาณ' ของ indicators เหล่านี้ครับ/ค่ะ"

ส่วนบน: Accuracy (%)

- "กราฟแท่งแนวนอนด้านบนนี้ แสดงถึง 'ความแม่นยำ (Accuracy)' ของแต่ละ indicators ในการให้สัญญาณการเคลื่อนไหวของตลาดครับ/ค่ะ"

การวัด 'ความแม่นยำ' (Accuracy):

- เมื่อ indicators ส่งสัญญาณออกมา เราจะเปรียบเทียบกับ 'แนวโน้มราคาจริง' ที่เกิดขึ้น ณ จุดเปลี่ยนแนวโน้มที่ใกล้ที่สุดครับ
- ถ้าทิศทางของสัญญาณ indicators ตรงกับ ทิศทางของแนวโน้มราคาจริง = สัญญาณนั้น 'ถูกต้อง' (Correct Signal)
- ถ้าทิศทางของสัญญาณ indicators ไม่ตรงกับ ทิศทางของแนวโน้มราคาจริง = สัญญาณนั้น 'ผิดพลาด' (False Signal)
- 'ความแม่นยำ' ก็คือ ร้อยละของสัญญาณที่ถูกต้อง เทียบกับจำนวนสัญญาณทั้งหมดที่ indicators นั้นส่งออกมาครับ

จากภาพก็จะเห็นได้ว่า

- "'McClellan Volume Oscillator' และ 'McClellan Oscillator' แสดงความแม่นยำสูงสุดที่ 88.0% และ 86.7% ตามลำดับ ซึ่งถือว่าเป็นระดับที่สูงมากครับ"
- นอกจากความแม่นยำแล้ว ตัวเลขด้านล่าง ยังบอกเราถึง 'ลักษณะของสัญญาณ' ด้วยครับว่าสัญญาณนั้นเป็น 'Lagging' หมายถึงสัญญาณที่ indicators ให้นั้น 'เกิดขึ้นตามหลัง' การเคลื่อนไหวของตลาดเล็กน้อย เช่น McClellan Oscillator เป็น Lagging 0.5d (0.5 วัน) หรือ McClellan Volume Oscillator เป็น Lagging 0.6d (0.6 วัน)"
- การวัด 'จังหวะสัญญาณ' (Lead/Lag Analysis):
 - เมื่อ indicators ส่งสัญญาณ และสัญญาณนั้น 'ถูกต้อง' เราจะไปดูว่า 'สัญญาณจาก indicators ' เกิดขึ้น 'ก่อน' หรือ 'หลัง' การเปลี่ยนแปลงแนวโน้มราคาจริงที่วัน
 - ขออธิบายยกตัวอย่างเพิ่มเติมเพื่อความเข้าใจนะครับ โดยตัว
 - 'Lagging Signal' (สัญญาณตาม): ถ้าสัญญาณจาก indicators เกิดขึ้นหลังจาก การเปลี่ยนแปลงแนวโน้มราคาจริง เช่น ราคาจริงเปลี่ยนทิศทางวันนี้ แต่ indicators เพิ่งจะเปลี่ยนทิศทางตามไปอีก 3 วันข้างหน้า แบบนี้เรียกว่าสัญญาณ lag 3 วัน

- **'Leading Signal' (สัญญาณนำ):** ถ้าสัญญาณจาก indicators เกิดขึ้นก่อน การเปลี่ยนแปลงแนวโน้มราคาจริง เช่น indicators เปลี่ยนทิศทางวันนี้ แต่ราคาจริงไปเปลี่ยนทิศทางในอีก 2 วันข้างหน้า แบบนี้เรียกว่าสัญญาณlead 2 วัน

จากกราฟแสดงผลข้างต้นจะครบก็จะแสดงให้เห็นถึงความแม่นยำและความดีเลย์ของindicatorsแต่ละตัวนะครับ

ถัดมาโดยในส่วนล่างนะครับคือ

: Lead/Lag Signal Distribution

- "กราฟแท่งด้านล่างนี้จะเจาะลึกลงไปในเรื่องของ 'ประเภทของสัญญาณ' ครับ
- "สีเขียว แสดงสัดส่วนของสัญญาณที่เป็น 'Leading' และ สีส้ม แสดงสัดส่วนของสัญญาณที่เป็น 'Lagging'

หลังจากที่เราทราบข้อมูลข้างต้นดังกล่าวทั้งหมดที่กล่าวมานะครับ เราจะมาดู

การเสริมกันของทั้งสามมิติ

"ทำไมเราถึงต้องดูทั้งสามอย่างพร้อมกัน?"

"เพราะทั้งสามมิติให้ข้อมูลที่แตกต่างแต่เสริมกันครับ

- **"Correlation และ P-value เป็นพื้นฐาน:** บอกเราว่าindicatorsนั้นๆ มี 'ความสัมพันธ์ที่มั่นคงและน่าเชื่อถือ' กับตลาดโดยรวมหรือไม่ ถ้าindicatorsใดไม่มีความสัมพันธ์ที่น่าเชื่อถือทางสถิติแล้ว แม้ว่าในบางช่วงจะดูเหมือนแม่นยำ ก็อาจเป็นเพียงเรื่องบังเอิญ และไม่ควรมานำมาใช้เป็นเครื่องมือหลักครับ

ทำไมถึงไม่ควรนำมาใช้เป็นเครื่องมือหลัก" หากไม่มีความสัมพันธ์ที่น่าเชื่อถือทางสถิติ (คือ P-value สูง) แม้จะดูเหมือนแม่นยำในบางช่วง:

- **เสี่ยงต่อการเจอ 'ความสัมพันธ์จอมปลอม' (Spurious Correlation / Data Snooping Bias):**
 - ในชุดข้อมูลขนาดใหญ่ หรือเมื่อเราทดสอบดัชนีจำนวนมาก มันเป็นไปได้ที่ **โดยความบังเอิญ** ดัชนีบางตัวจะแสดงความแม่นยำสูงในบางช่วงเวลา หรือมีความสัมพันธ์ที่ดีกับตลาดในช่วงสั้นๆ
 - ถ้าเราไม่ได้ใช้ p-value มากรอง เราอาจหลงเชื่อในความสัมพันธ์ที่เกิดขึ้นเองโดยสุ่มนี้ และคิดว่าดัชนีนั้นมีประสิทธิภาพจริง ทั้งที่มันเป็นแค่ "โชค" ชั่วคราว
 - เมื่อเรานำดัชนีที่ "บังเอิญ" แม่นยำนี้ไปใช้ในอนาคต (ซึ่งสถานการณ์จะเปลี่ยนแปลงไป) ประสิทธิภาพก็มักจะหายไป เพราะไม่มีความสัมพันธ์ที่แท้จริงรองรับ
- **ขาด 'ความมั่นคง' และ 'ความสม่ำเสมอ':**
 - ดัชนีที่ไม่มีนัยสำคัญทางสถิติ แสดงว่าความสัมพันธ์ที่สังเกตเห็นนั้นไม่มีความสม่ำเสมอ หรือไม่คงที่ในระยะยาว
 - การที่มันแม่นยำในบางช่วงเวลา อาจเป็นผลจากปัจจัยเฉพาะกิจที่ยังยั่งยืน เมื่อปัจจัยเหล่านั้นหายไป ความแม่นยำก็หายไปด้วย
 - การสร้างกลยุทธ์จากเครื่องมือที่ขาดความมั่นคงเช่นนี้จะทำให้ผลลัพธ์ไม่น่าเชื่อถือและคาดเดาได้ยาก
-

- **"Accuracy** คือการประเมินประสิทธิภาพการใช้งาน: เมื่อเรามั่นใจแล้วว่า indicators มีความสัมพันธ์ที่น่าเชื่อถือ (จาก **Correlation** และ **P-value**) เราจึงจะมาดูต่อว่า 'ความสัมพันธ์นั้น ถูกแปลงเป็นสัญญาณที่แม่นยำในสถานการณ์จริงได้ดีแค่ไหน' **Accuracy** จึงเป็นเหมือนการทดสอบภาคปฏิบัติว่า indicators นั้นๆ สามารถใช้งานได้จริงตามวัตถุประสงค์หรือไม่"

ดังนั้น การที่เราได้เห็น indicators ที่มีทั้ง **Correlation สูง, P-value ต่ำ** (หมายถึงนัยสำคัญสูง) และยังแสดง **Accuracy** ที่ดี (เช่น **McClellan Oscillator** ที่มีทั้งนัยสำคัญและความแม่นยำสูง) นั้นยังเป็นการตอกย้ำว่า indicators นั้นๆ มีประสิทธิภาพที่ครบถ้วนและเชื่อถือได้ในหลายมิติ ทำให้เรามีความมั่นใจในการนำไปใช้ในการวิเคราะห์และประกอบการตัดสินใจลงทุนมากยิ่งขึ้นครับ

"โดยสรุปแล้ว การวิเคราะห์ทั้งสามมิตินี้ ทำให้เราสามารถคัดเลือก indicators ที่ใช้คำนวณ **market breadth** ที่มีศักยภาพสูงสุด ทั้งในด้าน 'ความสัมพันธ์เชิงโครงสร้าง' และ 'ความสามารถในการให้สัญญาณ' ครับ

โดยในส่วนต่อไปที่จะทำเพิ่มเติมหลังจากที่ได้อัปเดตกับพี่ดอลล่าแล้วก็คือ

นำ indicators ที่มี ประสิทธิภาพสูงๆ มาคำนวณ **score breadth** อีกครั้งครับ และเพิ่มในส่วนของการ **weight** น้ำหนักของแต่ละ indicators เรียงตาม **accuracy** ครับ ในส่วนนี้กำลังศึกษาและทดลองทำเพิ่มเติมต่อไปครับ

ของผมมีเท่านี้ครับ ขอบขอบคุณครับ

INCASE

คำแนะนำ: การวิเคราะห์ความแม่นยำและจังหวะของสัญญาณ (Indicator Accuracy and Lead/Lag Analysis)

Indicators แต่ละตัวให้สัญญาณการเปลี่ยนแปลงแนวโน้มของราคาได้แม่นยำแค่ไหน' และที่สำคัญกว่านั้นคือ 'สัญญาณนั้นเกิดขึ้นก่อน หรือเกิดขึ้นตามหลังการเปลี่ยนแปลงของราคาจริงมากน้อยเพียงใด'

"การวิเคราะห์นี้มีความสำคัญอย่างยิ่ง เพราะจะช่วยให้เราเข้าใจถึงธรรมชาติของแต่ละดัชนี และนำไปใช้ประโยชน์ได้อย่างเหมาะสม ทั้งในการ 'คาดการณ์ล่วงหน้า' หรือการ 'ยืนยันแนวโน้ม' ครับ/ค่ะ"

แนวคิดหลักของการวิเคราะห์ (Core Concepts)

ในการวิเคราะห์นี้ เราจะยึดหลักการดังต่อไปนี้ครับ

1. การนิยาม 'แนวโน้มราคาจริง' (Actual Trend):

- เราเริ่มต้นด้วยการกำหนด 'แนวโน้มราคาจริง' ของหุ้น (ในที่นี้คือ SET100) ในแต่ละวันครับ/ค่ะ
- โดยเราจะดูจากการเปลี่ยนแปลงของราคาปิด:
 - ถ้าวันนี้ราคาปิด สูงกว่า เมื่อวาน = แนวโน้ม ขึ้น (กำหนดให้เป็น +1)
 - ถ้าวันนี้ราคาปิด ต่ำกว่า เมื่อวาน = แนวโน้ม ลง (กำหนดให้เป็น -1)
 - ถ้าวันนี้ราคาปิด เท่าเดิม = แนวโน้ม คงที่ (กำหนดให้เป็น 0)
- นี่คือ 'ความเป็นจริง' ที่เราจะนำมาเปรียบเทียบกับสัญญาณจาก indicators ต่างๆ ครับ/ค่ะ

2. การระบุ 'สัญญาณการเปลี่ยนแปลง' จากดัชนี (Indicator Signals):

- สำหรับแต่ละ indicators เราจะดูว่าเมื่อไหร่ที่ 'ค่าของ indicators นั้นๆ มีการเปลี่ยนแปลงทิศทาง' ครับ (เช่น จากขึ้นเป็นลง หรือจากลงเป็นขึ้น)
- เมื่อ indicators เปลี่ยนทิศทาง เราจะถือว่านั่นคือ 'สัญญาณ' ที่ indicators ส่งออกมา

3. การวัด 'ความแม่นยำ' (Accuracy):

- เมื่อ indicators ส่งสัญญาณออกมา เราจะเปรียบเทียบกับ 'แนวโน้มราคาจริง' ที่เกิดขึ้น ณ จุดเปลี่ยนแนวโน้มที่ใกล้ที่สุดครับ
- ถ้าทิศทางของสัญญาณ indicators ตรงกับ ทิศทางของแนวโน้มราคาจริง = สัญญาณนั้น 'ถูกต้อง' (Correct Signal)

- ถ้าทิศทางของสัญญาณดัชนี **ไม่ตรงกับ** ทิศทางของแนวโน้มราคาจริง = สัญญาณนั้น **'ผิดพลาด' (False Signal)**
- **'ความแม่นยำ'** ก็คือ **ร้อยละของสัญญาณที่ถูกต้อง** เทียบกับจำนวนสัญญาณทั้งหมดที่ดัชนีนั้นส่งออกมาครับ/ค่ะ

4. การวัด 'จังหวะสัญญาณ' (Lead/Lag Analysis):

- เมื่อ indicators ส่งสัญญาณ และสัญญาณนั้น 'ถูกต้อง' เราจะไปดูว่า **'สัญญาณจาก indicators ' เกิดขึ้น 'ก่อน' หรือ 'หลัง' การเปลี่ยนแปลงแนวโน้มราคาจริงที่วัน**
- ขออธิบายยกตัวอย่างเพิ่มเติมเพื่อความเข้าใจนะครับ โดยตัว
- **'Leading Signal' (สัญญาณนำ):** ถ้าสัญญาณจาก indicators เกิดขึ้นก่อน การเปลี่ยนแปลงแนวโน้มราคาจริง เช่น indicators เปลี่ยนทิศทางวันนี้ แต่ราคาจริงไปเปลี่ยนทิศทางในอีก 2 วันข้างหน้า แบบนี้เรียกว่าสัญญาณ **lead 2 วัน**
- **'Lagging Signal' (สัญญาณตาม):** ถ้าสัญญาณจากดัชนี เกิดขึ้นหลังจาก การเปลี่ยนแปลงแนวโน้มราคาจริง เช่น ราคาจริงเปลี่ยนทิศทางวันนี้ แต่ดัชนีเพิ่งจะเปลี่ยนทิศทางตามในอีก 3 วันข้างหน้า แบบนี้เรียกว่าสัญญาณ **lag 3 วัน**
- เราจะคำนวณสัดส่วนของสัญญาณนำและสัญญาณตาม รวมถึงค่าเฉลี่ยจำนวนวันนำ/ตาม เพื่อให้เห็นภาพที่ชัดเจนครับ/ค่ะ (โดยเราจะมองหาจุดเปลี่ยนแนวโน้มราคาที่ใกล้ที่สุดภายในกรอบเวลาที่เรากำหนด เช่น ± 30 วัน เพื่อให้การวิเคราะห์มีความสมเหตุสมผล)

วิธีการคำนวณในโค้ด (Methodology in the Code)

"ในส่วนของวิธีการคำนวณนะครับ/ค่ะ โค้ดของเราจะทำงานเป็นขั้นตอนดังนี้ครับ/ค่ะ:"

1. สร้างคอลัมน์ **'Actual_Trend'**: คำนวณแนวโน้มราคาจริงในแต่ละวันจากราคาปิด
2. วนลูปที่ละดัชนี: โค้ดจะไล่ไปที่ละดัชนีวัดความกว้างของตลาดที่เราสนใจ
3. ตรวจสอบการเปลี่ยนแปลงสัญญาณ: ในแต่ละดัชนี โค้ดจะสแกนหาจุดที่ค่าของดัชนีมีการเปลี่ยนแปลงทิศทาง (นั่นคือ 'สัญญาณ' ที่ดัชนีส่งออกมา)
4. ค้นหาจุดเปลี่ยนแนวโน้มราคาจริงที่ใกล้ที่สุด: เมื่อดัชนีส่งสัญญาณ โค้ดจะมองหาจุดที่ **'Actual_Trend'** ของราคาจริงมีการเปลี่ยนแปลงทิศทางที่ใกล้ที่สุดในรอบประมาณ 30 วัน ทั้งก่อนหน้า และหลังจากวันที่ดัชนีส่งสัญญาณ
5. เปรียบเทียบทิศทาง:

- ถ้าทิศทางของสัญญาณดัชนี (ขึ้น/ลง) ตรงกับทิศทางของแนวโน้มราคาจริง = นับเป็น 'สัญญาณถูกต้อง'
- ถ้าไม่ตรง = นับเป็น 'สัญญาณผิดพลาด'

6. **คำนวณ Lead/Lag:** สำหรับสัญญาณที่ถูกต้อง ใ้คิดจะคำนวณจำนวนวันที่สัญญาณดัชนีเกิดก่อนหรือหลังแนวโน้มราคาจริง และนำไปเก็บไว้เพื่อคำนวณค่าเฉลี่ยและสัดส่วน

7. **สรุปผล:** หลังจากวนลูปครบทุกดัชนี ใ้คิดจะสรุปผลทั้งหมดออกมาในรูปแบบตาราง ซึ่งรวมถึง:

- Accuracy (%):** ร้อยละความแม่นยำของสัญญาณ
- Lead Ratio (%):** สัดส่วนของสัญญาณที่นำตลาด
- Lag Ratio (%):** สัดส่วนของสัญญาณที่ตามตลาด
- Avg. Lead Days:** ค่าเฉลี่ยจำนวนวันที่นำตลาด (ถ้ามีสัญญาณนำ)
- Avg. Lag Days:** ค่าเฉลี่ยจำนวนวันที่ตามตลาด (ถ้ามีสัญญาณตาม)

ความสำคัญของผลลัพธ์ (Importance of Results)

"ผลการวิเคราะห์นี้จะช่วยให้เราสามารถ:"

- ระบุ 'ดัชนีนำ' (Leading Indicators):** ดัชนีที่มีสัดส่วน Lead Ratio สูงและ Avg. Lead Days ที่เหมาะสม จะมีประโยชน์อย่างยิ่งในการช่วยคาดการณ์การเปลี่ยนแปลงของตลาดล่วงหน้า
- ระบุ 'ดัชนียืนยัน' (Confirming/Lagging Indicators):** ดัชนีที่มี Accuracy สูงแต่ Lagging อาจใช้เพื่อยืนยันแนวโน้มที่เกิดขึ้นแล้ว หรือใช้เป็นสัญญาณในการเข้าหรือออกจากสถานะการลงทุน
- ปรับปรุงกลยุทธ์:** เราสามารถนำข้อมูลนี้ไปปรับปรุงกลยุทธ์การลงทุนให้สอดคล้องกับพฤติกรรมของแต่ละดัชนี เพื่อเพิ่มโอกาสในการตัดสินใจที่แม่นยำและตรงจังหวะมากยิ่งขึ้นครับ/ค่ะ"

"จากผลการวิเคราะห์นี้ เราจะเห็นว่าดัชนีบางตัวมีบทบาทในการให้สัญญาณนำ ในขณะที่บางตัวโดดเด่นในเรื่องความแม่นยำในการยืนยันแนวโน้ม ซึ่งจะช่วยให้เราสร้างระบบการวิเคราะห์ที่มีประสิทธิภาพมากขึ้นครับ/ค่ะ"

"ขอบคุณครับ/ค่ะ"

