

Journey

Cleansing Data

1. มีการแก้ไข column ข้อมูลของ nav เพราะว่า ชนิดของข้อมูลตั้งต้นเป็นแบบ object เราต้องแปลงชนิดเพื่อที่จะนำมาทำข้อมูลต่อ
2. มีการ filter และเลือกเฉพาะ column ที่จำเป็นใช้ในการวิเคราะห์
3. Column 'management_style_th' ซึ่งเป็นข้อมูลที่แสดงถึงการจัดการกองทุน โดยข้อมูลจะออกไปทาง category และมีข้อมูลที่ค่อนข้างยาว จึงได้ทำการเปลี่ยนชื่อข้อมูลใน column นั้นเพื่อให้สะดวกต่อการ filter ข้อมูล

Data Analysis

1. ทำการ filter ทั้งสอง dataset โดยเลือก filter ตาม hypothesis ที่เรารู้อยู่ ดังนี้
 - a. Hypothesis 1 : การขึ้นดอกเบี้ยของภาครัฐ ส่งผลให้นักลงทุนโยกย้ายเงินลงทุนไปลงทุนจากกองทุนที่มีความเสี่ยงสูง (ตราสารทุน) ไปลงทุนในกองทุนที่มีความเสี่ยงต่ำกว่า (ตราสารหนี้) เราจะทำการ filter ข้อมูล column 'POLICY_THAI_DESC' โดยเลือกเฉพาะข้อมูลที่เป็น ตราสารหนี้ และตราสารทุน ของทั้ง dataset เดือนธันวาคม และเดือนกรกฎาคม
 - b. Hypothesis 2 : ในช่วงเศรษฐกิจตกต่ำ ส่งผลให้นักลงทุนโยกย้ายเงินลงทุนไปลงทุนจากกองทุน Passive ไปลงทุนในกองทุน Active หรือไม่ เราจะทำการ filter ข้อมูล column 'management_style_th' โดยเลือกเฉพาะข้อมูล ที่เป็นกองทุน Active และกองทุน Passive
2. ทำการหาค่าความเปลี่ยนแปลง (หน่วย : เปอร์เซนต์) จากของเดือนกรกฎาคม และเดือนธันวาคม
3. ทำการ filter outlier ออกด้วยวิธี IQR outlier detection
4. ทำการคำนวณช่วงความเชื่อมั่น ที่นัยสำคัญ 90% และทำการสรุปผลข้อมูลจากช่วงความเชื่อมั่น
 - a. หากช่วงความเชื่อมั่นทั้งสองทาง เป็นลบ => นักลงทุนเลือกที่จะลงทุนกองทุนนั้นๆในเดือน ธันวาคม มากกว่าเดือน กรกฎาคม
 - b. หากช่วงความเชื่อมั่นทั้งสองทาง มีค่าศูนย์อยู่ในช่วง => นักลงทุนเลือกที่จะลงทุนกองทุนนั้นๆในเดือน ธันวาคม และเดือนกรกฎาคม ไม่แตกต่างกัน
 - c. หากช่วงความเชื่อมั่นทั้งสองทาง เป็นบวก => นักลงทุนเลือกที่จะลงทุนกองทุนนั้นๆในเดือน ธันวาคม น้อยกว่าเดือน กรกฎาคม

