**Tự động hóa Google sheet – Phần 2**

**Gộp dữ liệu**

Ở phần trước mình đã giới thiệu cách lọc dữ liệu bằng Google sheet. Từ dữ liệu đầu vào là tổng tài sản của các công ty theo từng quý trong năm, mình đã lọc dữ liệu ra theo tùng năm và đặt dữ liệu vào từng sheet riêng biệt. Ngoài tài sản sẽ có hai chỉ số cần quan tâm trong bảng báo cáo tài chính là lợi nhuận và vốn chủ sở hữu, từ ba chỉ số này sẽ dễ dàng tính được hai chỉ số quan trọng trong việc đánh giá kết quả hoạt động của công ty là ROA và ROE. Cụ thể ROA chính là tỉ số lợi nhuân sau thuế trên tài sản, còn ROE là tỉ số lợi nhuận trên vốn chủ sở hữu. Trong bài viết này mình sẽ hướng dẫn cách gộp dữ liệu từ tài sản, lợi nhuận và vốn chủ sở hữu vào một sheet theo năm từ đó tính toán hai chỉ số ROA & ROE

Về cách làm cụ thể, từ một sheet ban đầu bao gồm tất cả các mã cổ phiếu, mình sẽ lấy dữ liệu tất cả các mã cổ phiếu và thêm 12 cột vào sau mỗi mã, cụ thể 4 cột đầu là lợi nhuận theo từng quý trong năm, tương tự 4 cột sau là vốn chủ sở hữu và 4 cột cuối là tài sản. Khi đã thêm được tất cả mã cổ phiếu vào sheet, mình sẽ cho lặp qua toàn bộ dữ liệu về lợi nhuận, vốn chủ sở hữu và tài sản, nếu trong chúng có dữ liệu thì phần dữ liệu này sẽ được ghi vào sheet, trường hợp không có sẽ được ghi 0

**Javascript vs App script**

Để lặp được qua tất cả các giá trị trong hàng và cột, mình sẽ sử dụng App script , app script sẽ chạy trên ngôn ngữ Javascript. Một điểm lưu ý là nếu không hiểu cách thức vận hành của Spreadsheet, việc chạy lệnh javascript trong App script sẽ rất lâu, và khi quá lâu Google sheet sẽ hủy bỏ quá trình, cùng xem xét một ví dụ dơn giản sau để hiểu rõ vấn đề

Cho cột A có mười dòng dữ liệu là các con số, từ các con số này hãy điền vào cột B giá trị ‘pass’ nếu số ở cột A lớn hơn 500, ngược lại sẽ nhận giá trị ‘fail’ nếu số bên cột A nhỏ hơn hoặc bằng 500

**Hình batch0**

Để giải quyết bài toán này mình sẽ lặp qua giá trị của cột A tổng cộng 10 lần bằng cách đọc các cột từ A1 đến A10, sau đó kiểm tra giá trị và viết vào cột B tương ứng. Như vậy trường hợp này Spreadsheet sẽ đọc giá trị từ range của cột A 10 lần và ghi giá trị vào cột B 10 lần tương ứng

**Hinh batch1**

Phương án thứ 2 mình sẽ tiến hành đọc Range dữ liệu từ cột A một lần duy nhất, các giá trị này sẽ được ghi vào 1 array hai lớp [[a1Value], [a2Value]… [a10Value]]. Tiếp theo mình sẽ kiểm tra dữ liệu trong array này và trả kết quả là fail hay pass, kết quả này sẽ được ghi vào một array khác. Array chứ kết quả Fail hoặc pass sẽ được ghi vào cột B. Với phương án này Spreadsheet sẽ đọc giá trị của cột A 1 lần duy nhất và ghi giá trị lên cột B cũng 1 lần duy nhất

**Hinh batch2**

Từ hai trường hợp trên có thể dễ dàng nhận ra phương án 2 cho xử lý nhanh hơn phương án một khá nhiều (1 giây so với 6 giây), tại sao lại như vậy? Câu trả lời là Spreadsheet mất nhiều thời gian cho việc đọc và ghi dữ liệu vào sheet, do đó một nguyên tắc quan trọng khi làm việc với spreadsheet là phải hạn chế tối đa số lần đọc hoặc ghi giá trị của range, để làm được điều này bạn phải biết đóng gói các dữ liệu và đặt chúng trong các array. Làm sao để có thể đóng gói được dữ liệu vào một array và ghi lên các vùng khác chúng ta cần phải hiểu được giá trị trong sheet được hiển thị trong array như thế nào, hãy cũng xem xét ví dụ dưới đây

Cho dữ liệu ở cột A và B là các con số

**Hình practice1**

Dữ liệu cột A sẽ được lưu vào array aValues qua câu lệnh sau

var aValues = sheet.getRange("A1:A4").getValues();

Làm tương tự cho dữ liệu côt AB

var abValues = sheet.getRange("A1:B4").getValues();

In 2 giá trị aValues & abValues cho ra kết quả lần lượt là

[[1.0], [3.0], [5.0], [7.0]]

[[1.0, 2.0], [3.0, 4.0], [5.0, 6.0], [7.0, 8.0]]

**Hình practice2**

Từ kết quả trên có thể kết luận dữ liệu của range trong sheet sẽ được lưu vào một array 2 lớp, với lớp bên trong hiển thị tất cả các giá trị của cột thuộc một dòng bất kỳ

Khi đã có được array aValues, có thể dễ dàng ghi toàn bộ array này lên cột C qua câu lệnh

sheet.getRange("C1:C4").setValues(aValues)

Lúc này cột C sẽ được ghi giá trị giống cột A

**Hình practice 3**

**Gộp dữ liệu từ các file vào 1 file**

Để tính được 2 chỉ số ROA và ROE cho các mã chứng khoán, bước đầu tiên cần gộp được dữ liệu từ 3 file profit, capital & asset thành 1 file. Tiếp theo tính giá trị trung bình của profit, capital & asset. Với 3 giá trị trung bình này, có thể dễ dàng tính được 2 chỉ số ROE & ROA

Cụ thể, bước đầu tiên cần lấy được toàn bộ dữ liệu đầu vào bao gồm các giá trị của profit, asset & capital. Để thực hiện cần bật các file chứa các giá trị này và ghi hết giá trị của chúng vào một array, chi tiết cách thức ghi và đọc dữ liệu vào array các bạn có thể xem lại ở phần trên

Một thông tin đầu vào cũng quan trọng không kém là các mã cổ phiếu. Để lấy được các mã này cần bật sheet stockcode lên, sau đó lấy hết giá trị mã cổ phiếu và đặt chúng vào 1 array stockcode

Dưới đây mình sẽ hướng dẫn phần code chi tiết chạy trong app script

**Hinh code1**

Tiếp theo tạo một array chứa tiêu đề của sheet

**Hình code2**

Tiến hành cho lặp các giá trị của profit, asset & capital theo từng mã cổ phiếu và ghi các giá trị này vào 3 array tương ứng profitValue, assetValue & capitalValue, nếu không tìm được giá trị sẽ ghi là 0 vào array

**Hình code3**

**Hình code4**

**Hình code5**

Khi đã có đầy đủ các array, tiến hành ghi các array này lên các range trong sheet, sau đó thiết lập công thức tính giá trị trung bình của các giá trị profit, asset & capital

**Hình code6**

Cuối cùng set công thức tính ROE (%) và ROA (%) lần lượt cho cột Q và R

**Hình code7**