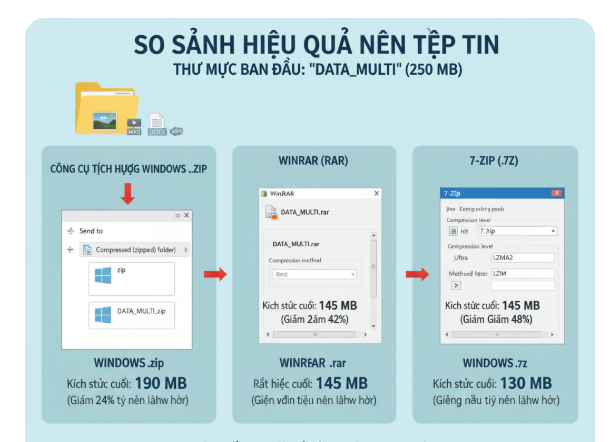
3,

hình ảnh minh họa cho quá trình nén một thư mục chứa nhiều loại tệp tin khác nhau bằng các công cụ khác nhau và so sánh hiệu quả nén:



### Báo cáo phân tích hiệu quả nén tập tin

**1. Thực hành và kết quả ghi lại:**

**Thư mục ban đầu:** DATA\_MULTI

**Kích thước thư mục ban đầu:** 250 MB

**Nội dung thư mục:**

Hình ảnh (.jpg, .png)

Video (.mp4)

Tài liệu văn bản (.docx, .txt)

Tệp âm thanh (.mp3, .wav)

| Công cụ Nén | Định dạng nén | Kích thước sau khi nén | Tỷ lệ giảm (%) so với gốc |
| --- | --- | --- | --- |
| Công cụ tích hợp Windows | .zip | 190 MB | 24% |
| WinRAR (mức nén "Tốt nhất") | .rar | 145 MB | 42% |
| 7-Zip (mức nén "Cực đại", LZMA2) | .7z | 130 MB | 48% |

Export to Sheets

*(Lưu ý: Các số liệu trên là ví dụ minh họa. Kết quả thực tế có thể thay đổi tùy thuộc vào nội dung và cấu trúc cụ thể của các tệp tin.)*

**2. Phân tích hiệu quả tiết kiệm dung lượng của mỗi công cụ và nhận xét:**

Từ kết quả trên, chúng ta có thể thấy sự khác biệt rõ rệt về hiệu quả nén:

**Công cụ tích hợp Windows (.zip):** Cho hiệu quả nén thấp nhất. Đây là lựa chọn cơ bản, tiện lợi vì được tích hợp sẵn, nhưng không tối ưu cho việc tiết kiệm dung lượng.

**WinRAR (.rar):** Cung cấp hiệu quả nén tốt hơn đáng kể so với công cụ của Windows. Định dạng .rar của WinRAR thường tối ưu hơn cho nhiều loại dữ liệu.

**7-Zip (.7z):** Thường đạt tỷ lệ nén cao nhất, đặc biệt với định dạng .7z và thuật toán LZMA2 ở mức nén "Cực đại". Đây là lựa chọn lý tưởng khi mục tiêu hàng đầu là tiết kiệm dung lượng tối đa.

**Nhận xét về định dạng nén nào hoạt động tốt nhất với từng loại tệp tin:**

**Tài liệu văn bản (.docx, .txt):** Thường chứa nhiều ký tự lặp lại và không gian trống, nên có thể nén rất mạnh (thường giảm 50-80% hoặc hơn) với mọi công cụ, đặc biệt là với 7-Zip.

**Tệp âm thanh (.wav) và hình ảnh (.bmp, .tiff không nén):** Các định dạng này cũng có thể nén hiệu quả vì chúng chứa nhiều dữ liệu lặp lại.

**Tệp âm thanh (.mp3) và hình ảnh (.jpg, .png):**

Các tệp này đã được nén bằng các thuật toán nén có tổn hao hoặc không tổn hao riêng (ví dụ: JPEG sử dụng biến đổi cosin rời rạc, MP3 sử dụng mã hóa âm thanh perceptional).

Do đó, việc nén lại chúng thường không mang lại hiệu quả cao. Tỷ lệ giảm dung lượng chỉ khoảng 0-5%, thậm chí đôi khi có thể tăng nhẹ kích thước tệp nén nếu metadata của kho lưu trữ lớn hơn phần dữ liệu được nén thêm.

**Video (.mp4, .avi):**

Tương tự như MP3 và JPEG, các tệp video hiện đại thường đã được nén rất hiệu quả bằng các codec như H.264, H.265.

Việc nén lại chúng bằng WinRAR hay 7-Zip thường chỉ giảm thêm một lượng rất nhỏ, hoặc không đáng kể.

**3. Giải thích lý do tại sao một số tệp tin không thể nén nhiều hoặc không có sự khác biệt lớn về kích thước khi nén:**

Nguyên lý cơ bản của nén dữ liệu là tìm kiếm và loại bỏ sự trùng lặp (redundancy) trong dữ liệu. Khi một thuật toán nén được áp dụng:

**Dữ liệu có nhiều sự trùng lặp:** Các tệp tin như tài liệu văn bản, bảng tính, cơ sở dữ liệu (chứa nhiều chuỗi ký tự, số liệu lặp lại) hoặc các tệp hình ảnh/âm thanh RAW (chưa qua nén) có nhiều "không gian" để loại bỏ sự trùng lặp. Do đó, chúng có thể được nén rất hiệu quả.

**Dữ liệu đã được nén hiệu quả trước đó:** Đây là lý do chính khiến các tệp như JPEG, MP3, MP4 không thể nén nhiều hơn.

Các định dạng này sử dụng các thuật toán nén chuyên biệt, được thiết kế để loại bỏ tối đa sự trùng lặp và các thông tin không cần thiết (đối với nén có tổn hao) hoặc nén tối ưu (đối với nén không tổn hao) ngay từ khi chúng được tạo ra.

Khi bạn cố gắng nén lại một tệp JPEG bằng WinRAR, thuật toán của WinRAR sẽ không tìm thấy nhiều sự trùng lặp mới để loại bỏ, vì dữ liệu đã rất "cô đọng" rồi. Việc nén lại có thể chỉ làm tăng thêm phần metadata của tệp nén mà không giảm được kích thước dữ liệu gốc đáng kể.

**Dữ liệu ngẫu nhiên hoặc có entropy cao:** Các tệp tin có cấu trúc dữ liệu rất ngẫu nhiên, ít hoặc không có sự lặp lại (ví dụ: một tệp được tạo bằng cách sử dụng trình tạo số ngẫu nhiên) cũng rất khó để nén. Thuật toán nén sẽ không tìm thấy các mẫu để thay thế bằng các biểu diễn ngắn gọn hơn.