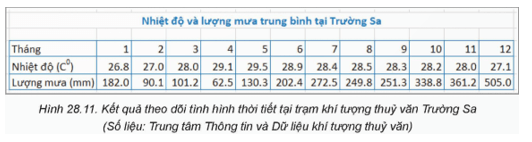
# Bài 28: Thực hành trải nghiệm trích rút thông tin và tri thức

**Giải Tin học 12 Bài 28: Thực hành trải nghiệm trích rút thông tin và tri thức**  
**Khởi động trang 149 Tin học 12**: Có thể hiểu phân tích dữ liệu là việc trích rút thông tin hữu ích giúp tạo ra tri thức mới từ dữ liệu đã thu thập được. Trong thực tế, công việc này thường gắn với việc xử lí để biến đổi dữ liệu về dạng thuận tiện, phù hợp với yêu cầu phân tích. Hãy trao đổi và cho biết, nếu dữ liệu dạng file Excel có 2 cột: Số tuổi và Thu nhập, trong trường hợp muốn tổng hợp kết quả thu nhập theo độ tuổi thì cần bổ sung thêm cột dữ liệu nào? Dữ liệu cột đó có thể lấy từ đâu và bằng cách nào?  
**Lời giải:**  
Để tổng hợp kết quả thu nhập theo độ tuổi từ dữ liệu dạng file Excel có 2 cột: Số tuổi và Thu nhập, bạn cần bổ sung thêm một cột dữ liệu mới là Nhóm độ tuổi. Cách tạo cột này như sau:  
- Thêm cột Nhóm độ tuổi:  
+ Tạo một cột mới trong bảng tính và đặt tên là "Nhóm độ tuổi".  
+ Sử dụng công thức để phân loại tuổi vào các nhóm (ví dụ: 18-24, 25-34, 35-44, v.v.). Một ví dụ công thức có thể là:  
=INT((B2-1)/5)\*5+1&"-"&(INT((B2-1)/5)+1)\*5  
Trong đó, B2 là ô chứa số tuổi của mỗi người. Công thức này sẽ tạo ra các nhóm tuổi như "18-24", "25-34",…  
- Lấy dữ liệu từ nguồn khác:  
+ Nếu em không có dữ liệu về nhóm độ tuổi, bạn có thể tìm kiếm dữ liệu thống kê về thu nhập theo độ tuổi từ các nguồn đáng tin cậy như cơ quan thống kê, báo cáo nghiên cứu, hoặc dữ liệu từ tổ chức chính phủ.  
+ Sau khi có dữ liệu, em có thể nhập vào cột "Nhóm độ tuổi" tương ứng với từng số tuổi.  
- Tính tổng hợp thu nhập:  
+ Sử dụng Pivot Table để tổng hợp thu nhập theo nhóm độ tuổi. Nhóm thu nhập nào có tỉ lệ phần trăm khách hàng có khả năng tín dụng mức Tốt cao nhất?  
**Nhiệm vụ 4 trang 153 Tin học 12**: Hãy cho biết:  
Nhóm thu nhập nào có tỉ lệ phần trăm khách hàng có khả năng tín dụng mức Kém cao nhất?  
Nhóm thu nhập nào có số lượng khách hàng có khả năng tín dụng Tốt gần gấp đôi số khách hàng có khả năng tín dụng Kém?  
Nhóm thu nhập nào có khả năng tín dụng mức Kém cao hơn mức Tốt?Chọn cột "Nhóm độ tuổi" làm hàng dọc và cột "Thu nhập" làm giá trị. Pivot Table sẽ tự động tính tổng hợp thu nhập cho từng nhóm độ tuổi.  
**Lời giải:**  
- Nhóm thu nhập Cao có tỉ lệ phần trăm khách hàng có khả năng tín dụng mức Tốt cao nhất.  
- Nhóm thu nhập Thấp có tỉ lệ phần trăm khách hàng có khả năng tín dụng mức Kém cao nhất.  
- Nhóm thu nhập Cao có số lượng khách hàng có khả năng tín dụng Tốt gần gấp đôi số khách hàng có khả năng tín dụng Kém.  
- Nhóm thu nhập Thấp có tỉ lệ phần trăm khách hàng với khả năng tín dụng mức Kém cao hơn mức Tốt.  
Như vậy, thông qua biểu đồ, chúng ta có cái nhìn tổng quan về khả năng tín dụng của từng nhóm thu nhập.  
**Luyện tập 1 trang 154 Tin học 12**: Có thể sử dụng hàm IF lồng trong nhau kết hợp với thao tác “kéo – thả” công thức trực tiếp trong bảng dữ liệu ban đầu để tạo các cột phân loại Mức thu nhập và Nhóm tuổi. Theo em, cách làm này có khuyết điểm gì so với việc sử dụng Power Query?  
**Lời giải:**  
Sử dụng hàm IF lồng trong nhau và thao tác "kéo – thả" công thức trực tiếp trong bảng dữ liệu ban đầu để tạo các cột phân loại Mức thu nhập và Nhóm tuổi có một số khuyết điểm so với việc sử dụng Power Query:  
- Khó duy trì và hiệu suất kém: Khi sử dụng hàm IF lồng nhau, công thức trở nên phức tạp và khó duy trì. Nếu bạn cần thay đổi logic hoặc thêm điều kiện mới, việc chỉnh sửa sẽ rất khó khăn. Ngoài ra, hiệu suất tính toán có thể bị ảnh hưởng.  
- Khả năng xử lý dữ liệu lớn: Power Query được tối ưu hóa để xử lý dữ liệu lớn một cách hiệu quả. Nó cho phép bạn kết nối, làm sạch và biến đổi dữ liệu từ nhiều nguồn khác nhau. Sử dụng Power Query giúp bạn xử lý dữ liệu lớn một cách nhanh chóng và linh hoạt hơn.  
  
**Luyện tập 2 trang 154 Tin học 12**: Nếu chỉ quan sát trực tiếp bảng dữ liệu ban đầu, em có thể dễ dàng trả lời các câu hỏi nêu trong Nhiệm vụ 4 không? Hãy nêu một vài nhận xét về những trải nghiệm em thu được thông qua việc thực hiện các Nhiệm vụ thực hành trong bài học.  
**Lời giải:**  
Nếu chỉ quan sát trực tiếp bảng dữ liệu ban đầu mà không có bất kỳ phân tích hay biểu đồ hỗ trợ, có thể khó để trả lời các câu hỏi chi tiết trong Nhiệm vụ 4. Việc phân tích dữ liệu và biểu diễn thông qua bảng tổng hợp và biểu đồ giúp chúng ta có cái nhìn tổng quan và dễ dàng nhận ra các mẫu, mối quan hệ và xu hướng trong dữ liệu.  
Thông qua việc thực hiện các Nhiệm vụ thực hành trong bài học, ta có thể thu được một số trải nghiệm như sau:  
- Quá trình phân tích dữ liệu yêu cầu sự tổ chức và phân loại dữ liệu một cách cẩn thận. Việc biểu diễn dữ liệu theo các đồ thị, biểu đồ hoặc bảng tổng hợp giúp ta dễ dàng nhìn thấy các mẫu, xu hướng và mối quan hệ trong dữ liệu.  
- Đối với mục tiêu phân tích dữ liệu, việc đặt câu hỏi cần được xác định rõ ràng từ trước để có thể tìm hiểu và rút ra kết luận từ dữ liệu được tổng hợp.  
- Trong quá trình phân tích, việc sử dụng các công cụ và phần mềm thống kê, hình dung dữ liệu có thể giúp chúng ta xác định các xu hướng, tính chất và mối quan hệ trong dữ liệu một cách hiệu quả.  
- Việc biểu diễn dữ liệu theo các hình thức trực quan như biểu đồ cột, biểu đồ đường, biểu đồ tròn, heatmap và sơ đồ tương quan giúp chúng ta dễ dàng nhìn thấy và hiểu rõ hơn về dữ liệu.  
  
**Luyện tập 3 trang 154 Tin học 12**: Tạo bảng tổng hợp và biểu đồ khả năng tín dụng theo nhóm tuổi. Nêu nhận xét về kết quả thu được.  
**Lời giải:**  
Hướng dẫn làm:  
- Tạo Bảng Tổng Hợp:  
+ Mở Excel và tạo một workbook mới.  
+ Tạo các cột: Nhóm tuổi, Số lượng khách hàng, và Khả năng tín dụng.  
+ Nhóm tuổi có thể là các khoảng như "18-24", "25-34", v.v.  
+ Điền số lượng khách hàng và khả năng tín dụng tương ứng cho từng nhóm.  
- Biểu Đồ Khả Năng Tín Dụng:  
+ Sử dụng biểu đồ cột hoặc biểu đồ hình tròn để trực quan hóa khả năng tín dụng theo nhóm tuổi.  
+ Trục x của biểu đồ là nhóm tuổi, trục y là khả năng tín dụng hoặc số lượng khách hàng.  
- Nhận Xét Kết Quả:  
+ Qua bảng tổng hợp và biểu đồ, bạn có thể nhận xét về mức độ khả năng tín dụng của từng nhóm tuổi.  
+ Có thể phát hiện xu hướng hoặc sự khác biệt giữa các nhóm.  
+ Điều này có thể giúp bạn đưa ra quyết định về chiến lược kinh doanh hoặc tiếp thị.  
  
**Vận dụng trang 154 Tin học 12**: Trong Hình 28.11 là nhiệt độ và lượng mưa đo được tại Trường Sa. Những thông tin hữu ích nào có thể rút ra từ dữ liệu này? Nếu biết mùa mưa là mùa có 3 tháng liên tiếp lượng mưa trung bình trên 100 mm và lớn hơn các tháng còn lại, thì mùa mưa ở Trường Sa là những tháng nào?  
  
**Lời giải:**  
Dữ liệu trong Hình 28.11 cho thấy nhiệt độ và lượng mưa trung bình hàng tháng tại Trường Sa. Dựa trên dữ liệu này, chúng ta có thể rút ra các thông tin hữu ích như sau:  
- Nhiệt độ:  
+ Nhiệt độ trung bình hàng tháng dao động từ 26.8°C đến 29.5°C, cho thấy Trường Sa có một khí hậu ấm.  
+ Trong khoảng nhiệt độ này, môi trường thích hợp cho cuộc sống và hoạt động của người dân và động vật.  
- Lượng mưa:  
+ Lượng mưa trung bình tăng lên từ tháng 4 và đạt điểm cao nhất vào các tháng 9, 10 và 11 với lượng mưa trên 250 mm/tháng.  
+ Theo tiêu chí đã được xác định (3 tháng liên tiếp có lượng mưa trung bình trên 100 mm và cao hơn các tháng còn lại), mùa mưa tại Trường Sa diễn ra vào các tháng 9, 10 và 11.  
Như vậy, mùa mưa ở Trường Sa thường kéo dài từ tháng 9 đến tháng 11.