***Take action with data***

**6 pha phân tích dữ liệu**

* ***Pha hỏi (Ask phase)***

Bạn không thể giải quyết vấn đề nếu không biết nó là cái gì.

Các vấn đề cần xem xét:

* Định nghĩa vấn đề cần giải quyết
* Đảm bảo thấu hiểu kỳ vọng của các bên liên quan
* Tập trung vào vấn đề thật sự, tránh các yếu tố xao lãng
* Cộng tác với các bên liên quan và duy trì giao tiếp cởi mở
* Lùi một bước để xem xét tổng thể tình huống trong ngữ cảnh

Câu hỏi đặt ra cho chính mình:

* Các bên liên quan nói gì về vấn đề của họ?
* Sau khi hiểu vấn đề, tôi cần làm gì để giúp các bên liên quan giải đáp câu hỏi của họ?
* ***Pha chuẩn bị (Prepare phase)***

Dữ liệu nào cần thu thập để trả lời câu hỏi và tổ chức dữ liệu như thế nào để dữ liệu có ích.

Bạn dựa vào tác vụ kinh doanh của mình để quyết định

* Việc đo lường được thực hiện trên độ đo nào
* Định vị dữ liệu trong cơ sở dữ liệu
* Tạo các thước đo bảo mật để bảo vệ dữ liệu trên

Câu hỏi đặt ra cho chính mình:

* Tôi cần tìm ra cách giải quyết vấn đề này là gì?
* Tôi cần phải nghiên cứu điều gì?
* ***Pha xử lý (Process Phase)***

Dữ liệu sạch là dữ liệu tốt nhất → cần làm sạch dữ liệu để loại bỏ mọi vấn đề tiềm ẩn về lỗi, sai lệch, hoặc không nhất quán

*Điều này có nghĩa là bạn cần phải thực hiện những việc sau*:

* Dùng hàm trên bảng tính để tìm dữ liệu bị nhập sai
* Dùng hàm SQL để kiểm tra các vấn đề khác của dữ liệu
* Loại bỏ các hạng mục dữ liệu trùng
* Kiểm tra vấn đề dữ liệu lệch càng nhiều càng tốt

*Câu hỏi đặt ra cho chính mình:*

* Những lỗi dữ liệu hoặc sự không chính xác nào có thể mắc phải trong quá trình tìm câu trả lời tốt nhất cho vấn đề đang giải quyết?
* Làm sạch dữ liệu như thế nào để thông tin có được nhất quán hơn?
* ***Pha phân tích (Analyze Phase)***

Bạn sẽ muốn suy nghĩ phân tích về dữ liệu của mình.

Sắp xếp và định dạng dữ liệu sẽ giúp dễ dàng hơn khi:

* Thực hiện các phép tính toán
* Kết hợp dữ liệu từ nhiều nguồn
* Tạo bảng với kết quả hiện có

***Câu hỏi đặt ra cho chính mình:***

* Dữ liệu đang kể với tôi câu chuyện gì?
* Dữ liệu giúp giải quyết vấn đề như thế nào?
* Ai cần sản phẩm hay dịch vụ của công ty tôi? Đối tượng nào có khả năng sử dụng nó nhất?
* ***Pha chia sẻ (Share Phase)***

Đây là cơ hội để các bên liên quan thấy được con đường giải quyết vấn đề của bạn.

Tóm tắt kết quả phân tích bằng hình ảnh trực quan rõ ràng và hấp dẫn thông qua các công cụ (biểu đồ, trang tổng quan, ...)

Việc chia sẻ dĩ nhiên sẽ giúp đội ngũ của bạn

* Tạo ra các quyết định tốt hơn và sáng suốt hơn
* Dẫn đến kết quả mạnh mẽ hơn
* Truyền đạt thành công những khám phá

***Câu hỏi đặt ra cho chính mình:***

* Cần làm gì để những phần trình bày cho các bên liên quan trở nên hấp dẫn và dễ hiểu?
* Điều gì sẽ giúp tôi hiểu điều này nếu tôi là người nghe?
* ***Pha hành động (Act Phase)***
* Đưa vào sử dụng mọi thứ học được từ việc phân tích dữ liệu
* Cung cấp cho các bên liên quan lời gợi ý dựa trên khám phá dữ liệu để họ có thể đưa ra quyết định dựa trên dữ liệu

***Câu hỏi đặt ra cho chính mình:***

* Làm thế nào để sử dụng phản hồi nhận được từ pha Chia sẻ để thật sự thỏa mãn nhu cầu và kỳ vọng của các bên liên quan?

***Tư duy cấu trúc (structured thinking)***

Sáu bước thực hiện giúp phân rã quy trình phân tích dữ liệu thành những phần nhỏ hơn và có thể quản lý được.

Việc phân rã này được gọi là tư duy cấu trúc, bao gồm bốn hoạt động cơ bản

* Nhận biết vấn đề hay tình huống hiện tại
* Tổ chức thông tin hiện có
* Bộc lộ những khoảng trống và cơ hội
* Xác định các lựa chọn

**Terms and definitions for Course 2, Module 1**

**Action-oriented question:** A question whose answers lead to change

**Cloud:** A place to keep data online, rather than a computer hard drive

**Data analysis process:** The six phases of ask, prepare, process, analyze, share, and act whose purpose is to gain insights that drive informed decision-making

**Data life cycle:** The sequence of stages that data experiences, which include plan, capture, manage, analyze, archive, and destroy

**Leading question:** A question that steers people toward a certain response

**Measurable question:** A question whose answers can be quantified and assessed

**Problem types:** The various problems that data analysts encounter, including categorizing things, discovering connections, finding patterns, identifying themes, making predictions, and spotting something unusual

**Relevant question:** A question that has significance to the problem to be solved

**SMART methodology:** A tool for determining a question’s effectiveness based on whether it is specific, measurable, action-oriented, relevant, and time-bound

**Specific question:** A question that is simple, significant, and focused on a single topic or a few closely related ideas

**Structured thinking:** The process of recognizing the current problem or situation, organizing available information, revealing gaps and opportunities, and identifying options

**Time-bound question:** A question that specifies a timeframe to be studied

**Unfair question:** A question that makes assumptions or is difficult to answer honestly