Tuần 2:

* Thêm link data, chỉnh lại phần đầu, cân nhắc bài toán thứ 2
* Bổ sung Đặc trưng bộ dữ liệu:
  + Thời gian thu thập
  + Thống kê: số lượng khóa học, số loại khóa học, số trường đại học, lĩnh vực,
  + …

| **Tên tài nguyên** | **Số lượng** |
| --- | --- |
| Tài nguyên khóa học (Course Resource) | **4,216** khóa học |
| Video | **2,798,892** video |
| Bài tập (Exercise) | **358,265** bài tập |
| Vấn đề (Problem) | **2,454,397** vấn đề |
| Hồ sơ học sinh (Student Profile) | **3,330,294** hồ sơ |
| Hành vi xem video (Video watching behavior) | **154,332,174** dữ liệu gốc về nhật ký xem video |
| Hành vi bài tập (Exercise behavior) | **133,384,333** bản ghi |
| Trả lời và bình luận (Comment and Reply) | **8,422,134** bản ghi phản hồi bình luận |

→ cho thấy được độ đa dạng của dữ liệu

→ Nhận xét bộ dữ liệu có mới, có tương thích, có phức tạp

⇒ Bài toán: Gợi ý giảng viên,...

* Bảng Problem → check kiểu câu hỏi, option key, value
* Bài tập: dùng một file để dùng các kĩ thuật à quay video; vận dụng vào bài tập của nhóm

Tuần 3: