**CHƯƠNG 1: GIỚI THIỆU VỀ ĐỀ TÀI**

**1.1. Đề tài**

Tối ưu hóa dịch vụ nhà hàng thông qua Business Intelligence: Nghiên cứu tình huống tại chuỗi nhà hàng Pizza B&P.

**1.2. Mục tiêu**

Tối ưu hóa hoạt động kinh doanh của chuỗi nhà hàng Pizza B&P thông qua các công cụ BI bằng cách:

* **Dự đoán doanh số bán hàng** theo món ăn.
* **Tối ưu hóa thực đơn** dựa trên phân tích dữ liệu bán hàng và phản hồi của khách hàng.
* **Cải thiện hiệu suất dịch vụ**, giảm thời gian phục vụ.

**1.3. Lý do chọn đề tài**

Nhiều chuỗi nhà hàng gặp khó khăn trong việc:

* **Dự đoán nhu cầu khách hàng** và tối ưu hóa nguồn lực.
* **Hiểu rõ phản hồi khách hàng** để cải thiện dịch vụ.
* **Tăng doanh số** của các món ăn có hiệu suất kém.  
  ⟹ Nhóm chúng tôi quyết định sử dụng dữ liệu thực tế từ chuỗi nhà hàng Pizza B&P để trực tiếp hỗ trợ các mục tiêu trên, đồng thời cung cấp thông tin phong phú để triển khai các giải pháp BI.

**1.4. Phạm vi nghiên cứu**

Phân tích dữ liệu từ chuỗi nhà hàng Pizza B&P trong vòng **12 tháng**, bao gồm:

* Dữ liệu doanh số theo từng món ăn.
* Phản hồi của khách hàng.

**1.5. Kết quả dự kiến**

1. Báo cáo chi tiết với các phân tích và trực quan hóa dữ liệu.
2. Dashboard BI hiển thị doanh số và phản hồi của khách hàng.
3. Mô hình dự đoán doanh số và phân khúc khách hàng.
4. Đề xuất chiến lược cải thiện hiệu suất hoạt động của nhà hàng.

**CHƯƠNG 2: GIỚI THIỆU VỀ DỮ LIỆU**

**2.1. Mô tả dữ liệu**

Tệp **Expanded\_Pizza\_Sales.xlsx** bao gồm 48,620 dòng và 22 cột, cụ thể như sau:

1. **Thông tin đơn hàng**:
   * order\_details\_id, order\_id, pizza\_id, quantity.
2. **Thời gian đơn hàng**:
   * order\_date, order\_time.
3. **Thông tin giá và loại pizza**:
   * unit\_price, total\_price, pizza\_size, pizza\_category.
4. **Thành phần và tên món ăn**:
   * pizza\_ingredients, pizza\_name.
5. **Thông tin khách hàng**:
   * customer\_name, customer\_gender, customer\_dob.
6. **Phản hồi khách hàng**:
   * customer\_feedback, feedback\_platform, feedback\_date.
7. **Hoạt động nhân viên**:
   * employee\_id, employee\_activity.

**2.2. Mối quan hệ giữa các đặc trưng dữ liệu**

1. **Dữ liệu đơn hàng**:
   * order\_date: Từ ngày 01/01/2015 đến 31/12/2015.
   * order\_time: Có 16,382 giá trị khác nhau, ghi nhận rất chi tiết về thời gian.
   * quantity (trung bình): ~1.02, hầu hết khách hàng đặt 1 pizza/lần.
2. **Thông tin giá**:
   * unit\_price: Từ $9.75 đến $35.95, trung bình $16.49.
   * total\_price: Giá trị lớn nhất là $83, có thể do đơn hàng lớn.
3. **Chi tiết pizza**:
   * Có 32 loại pizza, chia thành 4 danh mục chính và 5 kích thước khác nhau.
   * Món bán chạy nhất: The Classic Deluxe Pizza.
4. **Phản hồi khách hàng**:
   * Có tổng cộng 64 khách hàng với các phản hồi đa dạng, ví dụ: “Rất thích không gian quán.”
   * customer\_gender: Gồm 3 nhóm giá trị: Nam, Nữ và nhóm “khác” hoặc không xác định.
5. **Hoạt động nhân viên**:
   * Bao gồm các nhiệm vụ như: “Đặt hàng”, “Giao hàng”, “Chuẩn bị món ăn” và “Chăm sóc khách hàng”.

**2.3. Lý do chọn dataset này thay vì dataset khác**

1. **Phù hợp với mục tiêu nghiên cứu**
   * Dataset này cung cấp dữ liệu toàn diện, liên quan trực tiếp đến mục tiêu nghiên cứu:
     + **Dự đoán doanh số**: Thông qua các trường dữ liệu như order\_date, order\_time, unit\_price, total\_price, và quantity.
     + **Phân tích hành vi khách hàng**: Với các trường customer\_feedback, customer\_gender, customer\_dob, và pizza\_preferences.
     + **Tối ưu hóa thực đơn**: Dựa vào thông tin doanh số (quantity) và thành phần món ăn (pizza\_ingredients).
2. **Mức độ chi tiết và tính thực tế cao**
   * Dataset ghi nhận **48,620 giao dịch thực tế** trong vòng 1 năm, đảm bảo:
     + **Tính toàn diện**: Bao quát mọi thời điểm trong năm (mùa cao điểm và thấp điểm).
     + **Phân tích xu hướng**: Xác định biến động doanh thu theo thời gian (ngày/giờ) và theo chi nhánh.
   * Dữ liệu phản ánh hoạt động kinh doanh thực tế, không phải dữ liệu giả lập.
3. **Đa dạng thông tin**
   * Dataset chứa không chỉ dữ liệu doanh số mà còn bao gồm:
     + **Thông tin khách hàng**: Phân tích nhóm khách hàng mục tiêu dựa trên độ tuổi, giới tính, hành vi.
     + **Phản hồi khách hàng**: Đánh giá sự hài lòng và cải thiện dịch vụ.
     + **Hoạt động nhân viên**: Đánh giá hiệu suất làm việc và tối ưu hóa thời gian phục vụ.
4. **Khả năng khai thác bằng các kỹ thuật BI**
   * Dataset phù hợp để áp dụng các kỹ thuật phân tích tiên tiến:
     + **Trực quan hóa dữ liệu**: Tableau có thể hiển thị xu hướng bán hàng, nhóm khách hàng, và hiệu suất dịch vụ.
     + **Phân tích nâng cao**:  
       ♣ **Dự báo doanh số**: Sử dụng các mô hình hồi quy.  
       ♣ **Phân cụm khách hàng**: Phân loại khách hàng theo hành vi.  
       ♣ **Phân tích cảm xúc**: Xác định mức độ hài lòng dựa trên phản hồi.

**CHƯƠNG 3: TIỀN XỬ LÝ VÀ TRỰC QUAN HÓA DỮ LIỆU**

**3.1. Tiền xử lý**

Các bước tiền xử lý dữ liệu:

1. **Kiểm tra và xử lý dữ liệu thiếu**
   * Xác định các cột có giá trị thiếu và quyết định:
     + Xóa cột (nếu thiếu quá nhiều).
     + Điền giá trị thay thế (trung bình, trung vị, hoặc mode).
2. **Xử lý dữ liệu không hợp lệ**
   * Kiểm tra các giá trị bất thường, ví dụ:
     + Giá trị âm trong các cột doanh thu hoặc số lượng.
     + Ngày tháng không hợp lệ hoặc nằm ngoài phạm vi mong muốn.
3. **Chuẩn hóa định dạng**
   * Chuyển các cột ngày tháng (order\_date và feedback\_date) sang định dạng chuẩn datetime.
   * Chuẩn hóa dữ liệu danh mục, ví dụ: phân loại kích thước pizza (pizza\_size) hoặc danh mục món ăn (pizza\_category).
4. **Tạo thêm các cột giá trị**
   * **Doanh thu bình quân trên mỗi pizza**: revenue\_per\_pizza = total\_price / quantity.
   * **Thời gian trong ngày**: Tạo cột order\_hour để phân tích giờ đặt hàng.