

Hiệu trưởng duyệt

**Mẫu đề thi đáp ứng chuẩn đầu ra học phần (CLO) và phục vụ đo lường đánh giá mức đạt PLO/PI**

*(Phần dành cho giảng viên khi thiết kế đề thi và các cán bộ quản lý đào tạo)*

TRƯỜNG ĐẠI HỌC VĂN LANG  
**KHOA THƯƠNG MẠI**

**ĐỀ THI ..... HỌC PHẦN**  
**Học kỳ 1, Năm học 2023-2024**

**I. Thông tin chung**

Học phần: KHAI THÁC VÀ PHÂN TÍCH DỮ LIỆU (trong Marketing) Số tín chỉ: 3

Mã học phần: 231\_7TM0410\_01 Mã nhóm lớp học phần: 231\_7TM0410\_01

Thời gian làm bài: **2 tuần** Hình thức thi: Báo cáo nhóm

SV được tham khảo tài liệu: Có  Không

Giảng viên nộp đề thi, đáp án Lần 1  Lần 2

**II. Các yêu cầu của đề thi nhằm đáp ứng CLO**

*(phần này phải phối hợp với thông tin từ đề cương chi tiết của học phần)*

Ký hiệu CLO	Nội dung CLO	Hình thức đánh giá	Trọng số CLO trong thành phần đánh giá (%)	Câu hỏi thi số	Điểm số tối đa	Lấy dữ liệu đo lường mức đạt PLO/PI
(1)	(2)	(3)	(4)	(5)	(6)	(7)
CLO1	Phân tích các chiến lược marketing kỹ thuật số thông qua các kiến thức chuyên sâu về khách hàng, dữ liệu và các công cụ marketing kỹ thuật số nhằm đưa ra các giải pháp nâng cao hiệu quả hoạt động marketing cho doanh nghiệp.	Báo cáo nhóm	20%	Task 1, Task 2, Task 3	Task 1: 25/100 Task 2: 15/100 Task 3: 60/100	M – PI 4.1 R – PI 5.1 R – PI 8.3 R – PI 9.2 R – PI 10.1
CLO2	Áp dụng các phương pháp, công cụ định tính và định lượng để phục vụ việc phân tích, tổng hợp, đánh giá dữ liệu và thông tin về hoạt động kinh doanh và marketing của doanh nghiệp.	Báo cáo nhóm	50%	Task 1, Task 2, Task 3	Task 1: 25/100 Task 2: 15/100 Task 3: 60/100	M – PI 4.1 R – PI 5.1 R – PI 8.3 R – PI 9.2 R – PI 10.1
CLO3	Vận dụng hiệu quả kỹ năng làm việc nhóm, kỹ năng làm việc độc lập nhằm phát triển bản thân và thực hiện công việc một cách hiệu quả	Báo cáo nhóm	15%	Task 1, Task 2, Task 3	Task 1: 25/100 Task 2: 15/100 Task 3: 60/100	M – PI 4.1 R – PI 5.1 R – PI 8.3 R – PI 9.2 R – PI 10.1
CLO4	Hình thành ý thức học tập suốt đời để làm việc hiệu quả và phát triển sự nghiệp	Báo cáo nhóm	15%	Task 1, Task 2, Task 3	Task 1: 25/100 Task 2: 15/100 Task 3: 60/100	M – PI 4.1 R – PI 5.1 R – PI 8.3 R – PI 9.2 R – PI 10.1

### Chú thích các cột:

(1) Chỉ liệt kê các CLO được đánh giá bởi đề thi kết thúc học phần (tương ứng như đã mô tả trong đề cương chi tiết học phần). Lưu ý không đưa vào bảng này các CLO không dùng bài thi kết thúc học phần để đánh giá (có một số CLO được bố trí đánh giá bằng bài kiểm tra giữa kỳ, đánh giá qua dự án, đồ án trong quá trình học hay các hình thức đánh giá quá trình khác chứ không bố trí đánh giá bằng bài thi kết thúc học phần). Trường hợp một số CLO vừa được bố trí đánh giá quá trình hay giữa kỳ vừa được bố trí đánh giá kết thúc học phần thì vẫn đưa vào cột (1).

(2) Nêu nội dung của CLO tương ứng.

(3) Hình thức kiểm tra đánh giá có thể là: trắc nghiệm, tự luận, dự án, đồ án, vấn đáp, thực hành trên máy tính, thực hành phòng thí nghiệm, báo cáo, thuyết trình, ..., phù hợp với nội dung của CLO và mô tả trong đề cương chi tiết học phần.

(4) Trọng số mức độ quan trọng của từng CLO trong đề thi kết thúc học phần do giảng viên ra đề thi quy định (mang tính tương đối) trên cơ sở mức độ quan trọng của từng CLO. Đây là cơ sở để phân phối tỷ lệ % số điểm tối đa cho các câu hỏi thi dùng để đánh giá các CLO tương ứng, bảo đảm CLO quan trọng hơn thì được đánh giá với điểm số tối đa lớn hơn. Cột (4) dùng để hỗ trợ cho cột (6).

(5) Liệt kê các câu hỏi thi số (câu hỏi số ... hoặc từ câu hỏi số... đến câu hỏi số...) dùng để kiểm tra người học đạt các CLO tương ứng.

(6) Ghi điểm số tối đa cho mỗi câu hỏi hoặc phần thi.

(7) Trong trường hợp đây là học phần cốt lõi - sử dụng kết quả đánh giá CLO của hàng tương ứng trong bảng để đo lường đánh giá mức độ người học đạt được PLO/PI - cần liệt kê ký hiệu PLO/PI có liên quan vào hàng tương ứng. Trong đề cương chi tiết học phần cũng cần mô tả rõ CLO tương ứng của học phần này sẽ được sử dụng làm dữ liệu để đo lường đánh giá các PLO/PI. Trường hợp học phần không có CLO nào phục vụ việc đo lường đánh giá mức đạt PLO/PI thì để trống cột này.

(Phần công bố cho sinh viên)

## I. Thông tin chung

Học phần: KHAI THÁC VÀ PHÂN TÍCH DỮ LIỆU (trong Marketing) Số tín chỉ: 3

Mã học phần: 231\_7TM0410\_01

Mã nhóm lớp học phần: 231\_7TM0410\_01

Thời gian làm bài: **2 tuần**

Hình thức thi: Báo cáo nhóm

SV được tham khảo tài liệu:

Có

Không

Giảng viên nộp đề thi, đáp án

Lần 1

Lần 2

## II. Nội dung câu hỏi thi

- **HƯỚNG DẪN SINH VIÊN TRÌNH BÀY:**

- ✓ **Trang bìa:** Tên trường, Tên môn học, Tiêu đề, Tên Giảng viên, Tên nhóm
- ✓ **Trang 2:** Tên các thành viên + Đóng góp của từng thành viên trong bài report cuối kỳ.
- ✓ **Table of contents**
- ✓ **Nội dung bài báo cáo**

- **ĐỀ BÀI:**

❖ **TASK 1: REGRESSION PROBLEM**

Dựa vào tập dữ liệu **IceCream\_Sales.csv** hãy thực hiện những yêu cầu phân tích sau đây (có thể sử dụng Excel, SPSS, hoặc bất kỳ phần mềm phân tích nào):

**1) Descriptive Statistical Analysis**

- 1.1. Calculate mean, median, mode, range, standard deviation.
- 1.2. Frequency for "Public Holidays".

**2) Skewness Analysis, Visualization, and Time Series**

- 2.1. Determine skewness (threshold:  $|\text{skewness}| > 0.5$  indicates significant skew).
- 2.2. Create histograms for distribution visualization.
- 2.3. Monthly averages visualization over two years for continuous variables.

### 3) **Boxplot Analysis for Outliers**

3.1. Generate boxplots for each continuous variable.

### 4) **Correlation Analysis**

4.1. Analyze correlations between independent variables and Ice Cream Sales.

### 5) **Regression Analysis**

5.1. Conduct linear regression to identify significant variables impacting sales.

5.2. Interpret coefficients, p-values, and R-squared.

### 6) **Interpretation and Insights**

6.1. Provide analytical insights based on analyses.

6.2. Discuss brief implications in a marketing context.

## ❖ **TASK 2: CLASSIFICATION PROBLEM**

Dựa vào tập dữ liệu `Customer_transaction_record.csv`, hãy thực hiện yêu cầu sau đây:

**\*\*Sử dụng kỹ thuật **K-Means** để phân nhóm khách hàng.**

**\*\*Các bước tính toán cần được làm thủ công và thể hiện trong bài report kết quả của từng bước.**

**\*\*Task Instruction (hướng dẫn các bước):**

1. Create a SCATTER PLOT.
2. Based on a provided scatter plot, decide the value of K.
3. Select Initial Centroids: Choose k data points from the dataset as the initial centroids.
4. Assign Points to Clusters:

For each point in the dataset, calculate the distance to each centroid and assign the point to the nearest cluster. You may use the Euclidean distance formula for this purpose.

5. Recalculate Centroids:

After all points have been assigned to clusters, recalculate the centroids of each cluster. The new centroid is the mean position of all points in the cluster.

## 6. Iterate the Process:

Repeat the process of assigning points and recalculating centroids. Continue this iterative process until the cluster assignments no longer change. Document each iteration step, including the positions of centroids and the points assigned to each cluster.

### ❖ TASK 3: DATA VISUALIZATION

Hãy chọn tập dữ liệu **Walmart\_retail\_data.csv** được đính kèm. Vận dụng kỹ năng trực quan hóa dữ liệu để phân tích tập dataset đó theo dạng tạo biểu đồ.

**\*\*Tiêu chí tối thiểu mỗi nhóm sử dụng 10 biểu đồ để biểu đạt các khía cạnh ẩn sau tập dữ liệu:**

- Nhóm nào có thêm nhiều hình ảnh thì số điểm càng cao.
- Nhóm nào có độ sáng tạo trong cách trực quan hóa dữ liệu (kết hợp đa thông tin, nhiều cột dữ liệu trong cùng một biểu đồ; hoặc cấu trúc biểu đồ có độ “lạ” và hiệu quả).
- Những giải thích lập luận cho các biểu đồ có đi kèm trích dẫn nguồn, trình bày khoa học, độ sâu và nhạy bén của phân luận giải của từng biểu đồ.

#### 1) Giới thiệu về tập dữ liệu

- 1.1. Dataset này thể hiện dữ liệu của ngành công nghiệp nào?
- 1.2. Giới thiệu sơ lược về ngành công nghiệp đó, ở nơi mà tập dữ liệu này được tạo ra (1 trang).
- 1.3. Cấu trúc dataset này như thế nào? Ví dụ: liệt kê các cột dữ liệu số (numeric: continuous, discrete), các cột dữ liệu dạng định danh (categorical: nominal, ordinal), các cột dữ liệu tọa độ, etc.
- 1.4. Nêu ra các cột dữ liệu chứa missing values? Chỉ rõ bao nhiêu dòng, và bao nhiêu % số dòng thuộc cột đó gặp tình trạng missing values?
- 1.5. Có xử lý missing values không? Nêu phương pháp Imputation (trám dữ liệu) cho từng cột dữ liệu có chứa missing values.

#### 2) Các bước chuẩn bị

- 2.1. Bao nhiêu cột được sử dụng trong bài phân tích? Liệt kê các cột.
- 2.2. Nêu sơ lược các nội dung muốn truyền tải đến người đọc thông qua bài phân tích.

Ví dụ:

*Bài phân tích nêu lên các nội dung chính bao gồm: (1) thống kê doanh thu tổng của chuỗi siêu thị theo (a) năm, (b) quý, (c) tháng; (2) doanh thu của từng category theo (a) năm, (b) quý, (c)*

tháng; (3) phân tích doanh thu/lợi nhuận của từng category hàng hóa, (4) lượng tiêu thụ từng loại hàng hóa theo khu vực, (n) etc.

2.3. Nêu ra các cột/biến (variables) tham gia trong từng mục (theo 2.2)

Ví dụ:

*Visual chart 1 = cột A + cột B + cột C*

*Visual chart 2 = cột C + cột D + cột F*

2.4. Nếu mục nào có tạo biến mới thì kê khai biến đó ra. Và nếu biến mới được tạo ra bởi hàm/syntax (theo 2.2) thì kê khai hàm/câu lệnh ra, hoặc nêu cách tạo biến mới.

Ví dụ:

*Visual chart 1 có tạo ra biến mới, tên là Profit ratio (vốn không có trong dataset gốc). Vậy biến mới Profit ratio được tạo ra bằng cách:*

*Profit ratio = Profit (cột A trong dataset) / Revenue (Cột B trong dataset)*

...

### **3) Data visualization**

3.1. Trình bày các Figures (hình ảnh) ứng với thứ tự đã kê khai trong 2.2

Ví dụ:

*(Visual chart 1a) thống kê doanh thu tổng của chuỗi siêu thị theo năm*

*(Visual chart 1b) thống kê doanh thu tổng của chuỗi siêu thị theo quý*

*(Visual chart 1c) thống kê doanh thu tổng của chuỗi siêu thị theo tháng*

3.2. Giải thích insight ứng với các Figures trong 3.1

### **4) Kết luận và nhận định sau khi phân tích**

4.1. Nhận định chung về tình hình (kinh doanh, dịch bệnh,...) đã quan sát được từ việc trực quan hóa tập dữ liệu

4.2. Đưa ra suggestions để cải thiện issues mà Walmart Retail gặp phải.

Ví dụ:

Theo tập dữ liệu về tình hình kinh doanh của siêu thị, nhóm chúng tôi phát hiện sự chậm trễ thường xảy ra trong việc giao những món hàng abcd, vậy cần cải thiện bằng cách xyz. Hoặc,

Vào tháng mưa, thì các mặt hàng nên được nhập chủ đạo bao gồm: a,b,c,d,... Vào tháng nắng thì các mặt hàng nên được nhập nhiều bao gồm; x,y,z,...

• **TIÊU CHÍ CHẤM ĐIỂM:**

- a) Đây là báo cáo nhóm, điểm số sẽ được cho dựa trên độ hoàn thiện của bài báo cáo. Nhóm trưởng đánh mức độ hoàn thiện công việc của từng thành viên (theo thang 0-100%).
- b) Điểm số các cá nhân = điểm tổng chung của bài nhóm \* % mức độ hoàn thiện công việc của từng thành viên.

TP. Hồ Chí Minh, ngày 09 tháng 12 năm 2023

**NGƯỜI DUYỆT ĐỀ**



**ThS Nguyễn Văn Tâm**

**GIẢNG VIÊN RA ĐỀ**



**TS LƯƠNG THÁI HÀ**