



## BÁO CÁO

### Lab 02: Làm việc và trực quan hóa dữ liệu chuỗi thời gian

**Bộ môn: Trực quan hóa dữ liệu**

#### NHÓM 22

<b>Nhóm trưởng</b>	20120128 – Nguyễn Thị Cẩm Lai
<b>Thành viên</b>	20120037 – Trần Thị Minh Anh
<b>Thành Viên</b>	20120125 – Bùi Anh Kiệt
<b>Thành viên</b>	20120232 – Võ Duy Trường
<b>Thành viên</b>	20120547 – Võ Thành Phong

**Giáo viên hướng dẫn:** Nguyễn Thị Thu Hằng  
Nguyễn Bảo Long  
Lê Nhựt Nam

## MỤC LỤC

<b>I.</b>	<b>Thông tin chung .....</b>	<b>2</b>
1.	Thông tin thành viên.....	2
2.	Phân công công việc .....	2
3.	Đánh giá mức độ hoàn thành.....	3
<b>II.</b>	<b>Tài liệu tham khảo .....</b>	<b>3</b>

## I. Thông tin chung

### 1. Thông tin thành viên

Họ và tên	MSSV	Email
Nguyễn Thị Cẩm Lai (nhóm trưởng)	20120128	<a href="mailto:20120128@student.hcmus.edu.vn">20120128@student.hcmus.edu.vn</a>
Trần Thị Minh Anh	20120037	<a href="mailto:20120073@student.hcmus.edu.vn">20120073@student.hcmus.edu.vn</a>
Bùi Anh Kiệt	20120125	<a href="mailto:20120125@student.hcmus.edu.vn">20120125@student.hcmus.edu.vn</a>
Võ Duy Trường	20120232	<a href="mailto:20120232@student.hcmus.edu.vn">20120232@student.hcmus.edu.vn</a>
Võ Thành Phong	20120547	<a href="mailto:20120547@student.hcmus.edu.vn">20120547@student.hcmus.edu.vn</a>

- **Github nhóm:**

[https://github.com/DataVisualizationGroup22/DataVisualization\\_Lab02](https://github.com/DataVisualizationGroup22/DataVisualization_Lab02)

### 2. Phân công công việc

Phần	Yêu cầu	Người thực hiện	Mức độ hoàn thành
A. Thu thập dữ liệu	Lựa chọn dataset	Cả nhóm	100%
	Trình bày ngữ cảnh và động cơ lựa chọn	Cẩm Lai	100%
B. Khám phá dữ liệu (thường đan xen với pha tiền xử lý dữ liệu)	1. Đọc dữ liệu và tính số dòng và cột	Duy Trường	100%
	2. Mỗi dòng có ý nghĩa gì? Có vấn đề các dòng có ý nghĩa khác nhau không?		
	3. Dữ liệu có các dòng bị lặp không?		
	4. Tỷ lệ giá trị thiếu và thống kê mô tả của từng cột		
	5. Kiểu dữ liệu của mỗi cột		
	6. Xem xét tập giá trị của các thuộc tính phân loại	Anh Kiệt	100%
	7. Xem xét sự phân bố giá trị của các cột dữ liệu dạng số		
	8. Xem xét sự phân bố giá trị của các cột dữ liệu không phải dạng số		
C. Khám phá mối quan hệ trong dữ liệu	Khám phá mối quan hệ đơn biến (độc lập)	Yêu cầu tối thiểu: <ul style="list-style-type: none"><li>• Minh Anh (4 khám phá)</li><li>• Cẩm Lai (4 khám phá)</li><li>• Anh Kiệt (2 khám phá)</li><li>• Duy Trường (2 khám phá)</li><li>• Thành Phong (4 khám phá)</li></ul>	100%
	Khám phá mối quan hệ đa biến/nhân quả (tích hợp)		
D. Tổng hợp		Cẩm Lai	100%

### 3. Đánh giá mức độ hoàn thành

Phần	Yêu cầu	Mức độ hoàn thành
A. Thu thập dữ liệu	Lựa chọn dataset	100%
	Trình bày ngữ cảnh và động cơ lựa chọn	100%
B. Khám phá dữ liệu	Khám phá dữ liệu	100%
	Tiền xử lý dữ liệu	100%
C. Khám phá mối quan hệ trong dữ liệu	Khám phá mối quan hệ đơn biến/đa biến	100%
	Sử dụng màu sắc để thể hiện dữ liệu	100%
	Thể hiện trực quan một số dữ liệu biến đổi theo thời gian	100%
	Sử dụng các kỹ thuật trực quan hóa	100%
	Trực quan hóa trên nhiều loại biểu đồ	100%

## II. Tài liệu tham khảo

**Tham khảo một số ý tưởng vẽ biểu đồ tại:**

<https://www.kaggle.com/code/atrisaxena/animated-way-of-visualization-ipyvizu>

<https://www.kaggle.com/code/bhuvanchennoju/data-science-salaries-plotly-tableau>