BỘ GIÁO DỤC VÀ ĐÀO TẠO



TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

KHOA CÔNG NGHỆ THÔNG TIN

NHẬP MÔN DỮ LIỆU LỚN

**BÁO CÁO BÀI TẬP MÔN HỌC**

**NỘI DUNG: LAB 02 - ADVANCED HADOOP & SPARK STRUCTURED APIs**

**Giảng viên hướng dẫn** : Huỳnh Lâm Hải Đăng

**Lớp** : CQ2022/21

**Sinh viên thực hiện** : Trương Tiến Anh- 22120017

Nguyễn Minh Cường- 22120043

Đinh Viết Lợi- 22120188

Nguyễn Trần Lợi-22120190

*Hồ Chí Minh, ngày 8 tháng 4 năm 2025*

# MỤC LỤC

[MỤC LỤC 1](#_Toc195009574)

[PHẦN 1: BÁO CÁO BÀI TẬP 2](#_Toc195009575)

[I. Exercise 2.1- Calculate revenue in the last 3 days for each country 2](#_Toc195009576)

[1. Ý tưởng thực hiện 2](#_Toc195009577)

[2. Mô tả chi tiết 2](#_Toc195009578)

[II. Exercise 2.2- Calculate the number of products sold in the last 7 days of each SKU, report every Monday 2](#_Toc195009579)

[1. Ý tưởng thực hiện 2](#_Toc195009580)

[2. Mô tả chi tiết 2](#_Toc195009581)

[III. Exercise 2.3- Find pairs of overlapping shapes 2](#_Toc195009582)

[1. Ý tưởng thực hiện 2](#_Toc195009583)

[2. Mô tả chi tiết 2](#_Toc195009584)

[IV. References 2](#_Toc195009585)

[PHẦN 2: BÁO CÁO NHÓM 3](#_Toc195009586)

# PHẦN 1: BÁO CÁO BÀI TẬP

## Exercise 2.1- Calculate revenue in the last 3 days for each country

### **Ý tưởng thực hiện**

### **Mô tả chi tiết**

## Exercise 2.2- Calculate the number of products sold in the last 7 days of each SKU, report every Monday

### **Ý tưởng thực hiện**

### **Mô tả chi tiết**

## Exercise 2.3- Find pairs of overlapping shapes

### **Ý tưởng thực hiện**

### **Mô tả chi tiết**

## References

1. Apache Spark, ‘Quickstart: DataFrame’ [Trực tuyến]. Đường dẫn: <https://spark.apache.org/docs/latest/api/python/getting_started/quickstart_df.html>

# PHẦN 2: BÁO CÁO NHÓM

* + Phương thức trao đổi và làm việc:
    - Nhóm yêu cầu tất cả thành viên đều tham gia giải tất cả bài tập nhằm mục đích bao quát trường hợp, nắm bắt lỗi và học tập hiệu quả.
    - Nhóm tổ chức họp nhóm trao đổi sau 4 ngày làm việc cá nhân kết hợp trao đổi liên tục hằng ngày.
    - Phiên họp nhằm kiểm tra kết quả, ý tưởng giải quyết cho từng vấn đề.
    - Giao công việc cho từng thành viên trên workspace của nhóm sau khi thống nhất kết quả.
  + Bảng phân công công việc:

|  |  |  |
| --- | --- | --- |
| Thành viên | Công việc | Lưu ý |
| Trương Tiến Anh |  |  |
| Nguyễn Minh Cường |  |  |
| Đinh Viết Lợi |  |  |
| Nguyễn Trần Lợi |  |  |