

BÁO CÁO NHÓM 7 - CASE STUDY 04 – PHÂN TÍCH DOANH THU / LỢI NHUẬN QUÁN CÀ PHÊ

Mục tiêu chính của dự án là phân tích doanh thu và lợi nhuận của quán cà phê từ đó rút ra các nhận xét hỗ trợ cho quyết định kinh doanh

Mục 1: Phương pháp xử lý và phân tích dữ liệu

- Sử dụng thành thạo pandas trong việc xử lý và phân tích dữ liệu

Cụ thể: Ứng dụng các hàm và kỹ thuật trong từng nhiệm vụ:

Nhập xuất file: `read_csv()` , `to_csv()`

Làm sạch dữ liệu để phân tích các bài toán truy xuất: `dropna()`, `fillna()`, `astype()`, `to_datetime()`, `str.strip()`, `str.title()`,...

Truy vấn để lọc những dữ liệu theo điều kiện: `loc`, `iloc`, `isin`

Sử dụng `Groupby()`, `agg()` để gom dữ liệu và `Merge()` để nhóm các dữ liệu lại với nhau

`Pivot_table()` tổng hợp và xoay dữ liệu, `stack()`, `unstack()` chuyển đổi giữa `DataFrame` với `Series`,....

- Ngoài ra cần hiểu biết công việc kinh doanh để rút ra kết luận từ kết quả phân tích.

Mục 2: Các bảng kết quả chính

order_id	product_id	order_time	qty
OD2000	CF05	2024-01-01 05:58	2x
OD2001	CF04	2024-01-01 11:45	2x
OD2002	CF03	2024-01-01 14:12	2
OD2003	CF03	2024-01-01 06:16	2
OD2004	CF02	2024-01-01 04:55	1
OD2005	CF05	14:05	-1
OD2006	CF02	2024-01-01 06:56	3
OD2007	CF01	15:21	2x
OD2008	CF06	2024-01-01 01:45	-1
OD2009	CF06	20:38	1
OD2010	CF02	20:03	-1
OD2011	CF02	11:21	2
OD2012	CF02	04:00	3
OD2013	CF05	06:58	2x
OD2014	CF03	2024-01-01 12:59	2x
OD2015	CF06	2024-01-01 10:31	3
OD2016	CF05	16:58	2

Order_log chưa qua làm sạch dữ liệu

	order_id	product_id	order_time	qty
1	OD2000	CF05	2024-01-01 05:58:00	2.0
2	OD2001	CF04	2024-01-01 11:45:00	2.0
3	OD2002	CF03	2024-01-01 14:12:00	2.0
4	OD2003	CF03	2024-01-01 06:16:00	2.0
5	OD2004	CF02	2024-01-01 04:55:00	1.0
6	OD2005	CF05		1.0
7	OD2006	CF02	2024-01-01 06:56:00	3.0
8	OD2007	CF01		2.0
9	OD2008	CF06	2024-01-01 01:45:00	1.0
10	OD2009	CF06		1.0
11	OD2010	CF02		1.0
12	OD2011	CF02		2.0
13	OD2012	CF02		3.0
14	OD2013	CF05		2.0
15	OD2014	CF03	2024-01-01 12:59:00	2.0
16	OD2015	CF06	2024-01-01 10:31:00	3.0
17	OD2016	CF05	16:58:00	2.0

Order_log đã qua làm sạch dữ liệu

Dữ liệu sau khi làm sạch đã đảm bảo tính nhất quán, sẵn sàng cho phân tích truy vấn dữ liệu.

```
product_id
CF01      27.0
CF02      37.0
CF03      32.0
CF04      44.0
CF05      31.0
CF06      26.0
Name: qty, dtype: float64
product_id
CF04       22
CF02       19
CF01       16
CF05       15
CF03       14
CF06       14
Name: count, dtype: int64
product_id
CF01      1.687500
CF02      1.947368
CF03      2.285714
CF04      2.000000
CF05      2.066667
CF06      1.857143
dtype: float64
Món có hiệu suất thấp nhất: CF01
Giá trị hiệu suất: 1.6875
```

Tính tổng số lượng đã bán theo từng món, số đơn hàng chứa từng món và giá trị hiệu suất của sản phẩm

Thông qua truy vấn dữ liệu các món có mã sản phẩm bán CF04, CF02, CF03 và CF05 là những món có tổng số lượng đã bán cao nhất, trong đó món có lượt bán thấp nhất CF06.

	hour	total_sold
0	0	1
1	1	4
2	2	7
3	3	1
4	4	7
5	5	6
6	6	14
7	7	3
8	8	3
9	9	2
10	10	7
11	11	5
12	12	3
13	13	4
14	14	5
15	15	4
16	16	3
17	17	8
18	18	2
19	19	1
20	20	2
21	21	4
22	22	4
23	23	4

Giờ có doanh thu cao nhất: 6.0:00

Doanh thu: 435,090

Món có doanh thu cao nhất: Trà Dao

Doanh thu: 1,080,000

Combo giờ-món có doanh thu cao nhất:

Giờ: 1.0:00

Món: Matcha Latte

Doanh thu: 140,000

product_name	Bac Xiu	Ca Phe Den	Ca Phe Sua	Matcha Latte	Sinh To Xoai \
order_date					
2024-01-01	629784.0	481000.0	500000.0	312000.0	-179640.0

product_name	Trà Dao
order_date	
2024-01-01	864000.0

Kết quả pivot lợi nhuận theo ngày của các món

Sử dụng Pivot Table giúp nhìn dữ liệu theo nhiều chiều một cách trực quan nhất hỗ trợ rút ra được các insight quan trọng cho hoạt động quán cà phê cụ thể:

Khung giờ có doanh thu cũng như số lượng bán cao nhất là khung giờ số 6, với doanh thu 435.090 và số lượng 14 cốc.

Món có doanh thu cao nhất là món Trà Đào với doanh thu 1.080.000, cũng là món có lợi nhuận cao nhất 864000.0, món có lợi nhuận thấp nhất là Sinh Tố Xoài với lợi nhuận âm

Tổng hợp phân tích và nhận xét: Dựa trên kết quả truy vấn và tổng hợp ở các nhiệm vụ giúp xác định cho thấy những thông tin hữu ích mà quán cà phê cần thu thập để biết được tiến độ phát triển của quán. Các món bán chạy: CF04-Trà Đào-45000, CF02-Cà phê sữa-35000, CF03- Bạc xỉu-35000 món có lợi nhuận cao như : CF04,CF02,CF03 trong đó món có mã CF04 là món bán chạy nhất. Khung giờ kinh doanh hiệu quả: khung giờ buổi sáng khoảng khung giờ số 6,từ đó quán có thể có kế hoạch chiến dịch hợp lý để duy trì đẩy mạnh những món đã có lợi nhuận, doanh số bán cao. Đồng thời cũng lưu ý những món đang có hiệu suất thấp và số lượng bán, lợi nhuận thấp: CF01- Cà phê đen, CF06-Matcha latte, doanh thu thấp: CF05- Sinh tố Xoài ít, cần có chiến lược sửa đổi phù hợp để thu hút vị giác người mua hơn.

Mục 3: Kết luận

Thông qua Case 04 nhóm đã vận dụng kiến thức lập trình python và thư viện pandas để xử lý phân tích dữ liệu thực tế cụ thể là phân tích doanh thu lợi nhuận của quán cà phê. Dự án giúp củng cố kỹ năng lập trình, tư duy phân tích và hiểu rõ hơn về vai trò của dữ liệu trong việc hỗ trợ kinh doanh

Mục 4: Nhật ký nhóm

- Nhiệm vụ 1: Phạm Thu Trang - Đọc và làm sạch dữ liệu
- Nhiệm vụ 2 : Đồng Thị Yến Nhi - Truy vấn và thống kê mô tả
- Nhiệm vụ 3: Nguyễn Thị Lan Anh 086- GroupBy và Tổng hợp
- Nhiệm vụ 4 : Ngô Đức Quang - Merge dữ liệu
- Nhiệm vụ 5: Phùng Quang Thanh- Pivot Table + Stack/Unstack