Day 21: Excel Final Project

FINAL PROJECT

A group of people in circle frames

Description automatically generated

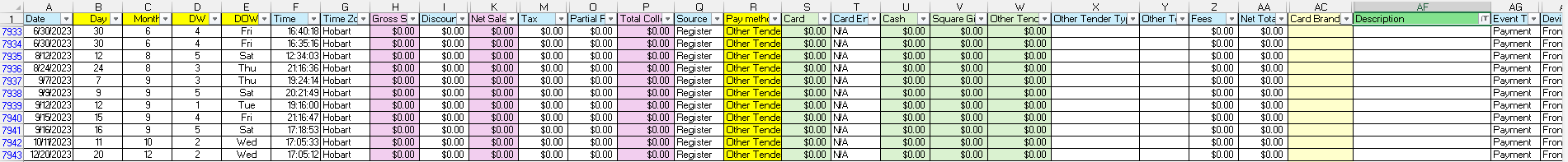
# 

# TÓM TẮT

Team C bước đầu sẽ đọc hiểu dữ liệu bên trong Dataset, tìm hiểu mối liên kết giữa các bảng, các trường dữ liệu với nhau. Team sẽ dựa vào từng Task, tìm hiểu xem với Task đó sẽ cần những thông tin gì. Dataset được làm sạch bằng cách dùng hàm tạo thêm trường dữ liệu mới, từ dữ liệu cũ. Đối với những dữ liệu cũ, dữ liệu ko dùng tới sẽ được ẩn đi (Hide). Sau khi dữ liệu được làm sạch, sẽ thực hiện tổng hợp, vẽ chart, phân tích và đưa ra góp ý nhằm giúp nhà hàng phát triển hơn trong tương lai.

**Link đề:** <https://discord.com/channels/1124195842820161637/1211551059005145138/1221257092674097283>

11 row hiển thị toàn số 0 ở cột tiền, ko có ý nghĩa, ko có mô tả

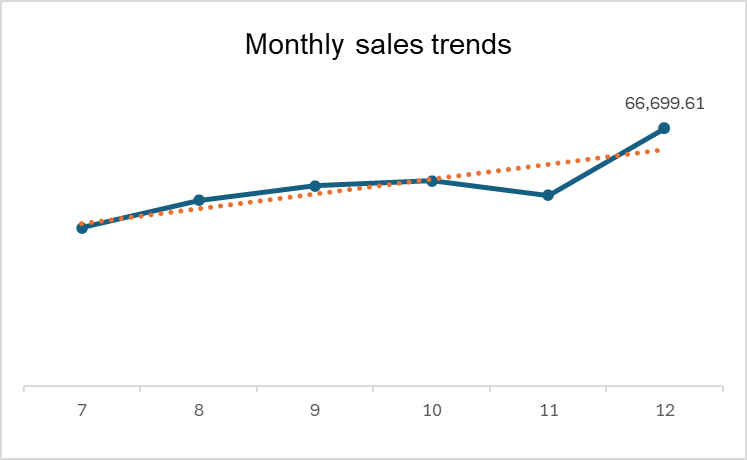


**Task 1**:

Ko có tên trục Y trục X, trendline là thừa. 7 8 9 10 11 12 ko hiểu là số gì

trục Y ko phân cấp, ko biết chênh lệch như thế nào

Nhà phân tích chỉ nhìn liếc qua 1 phút để nắm thông tin.



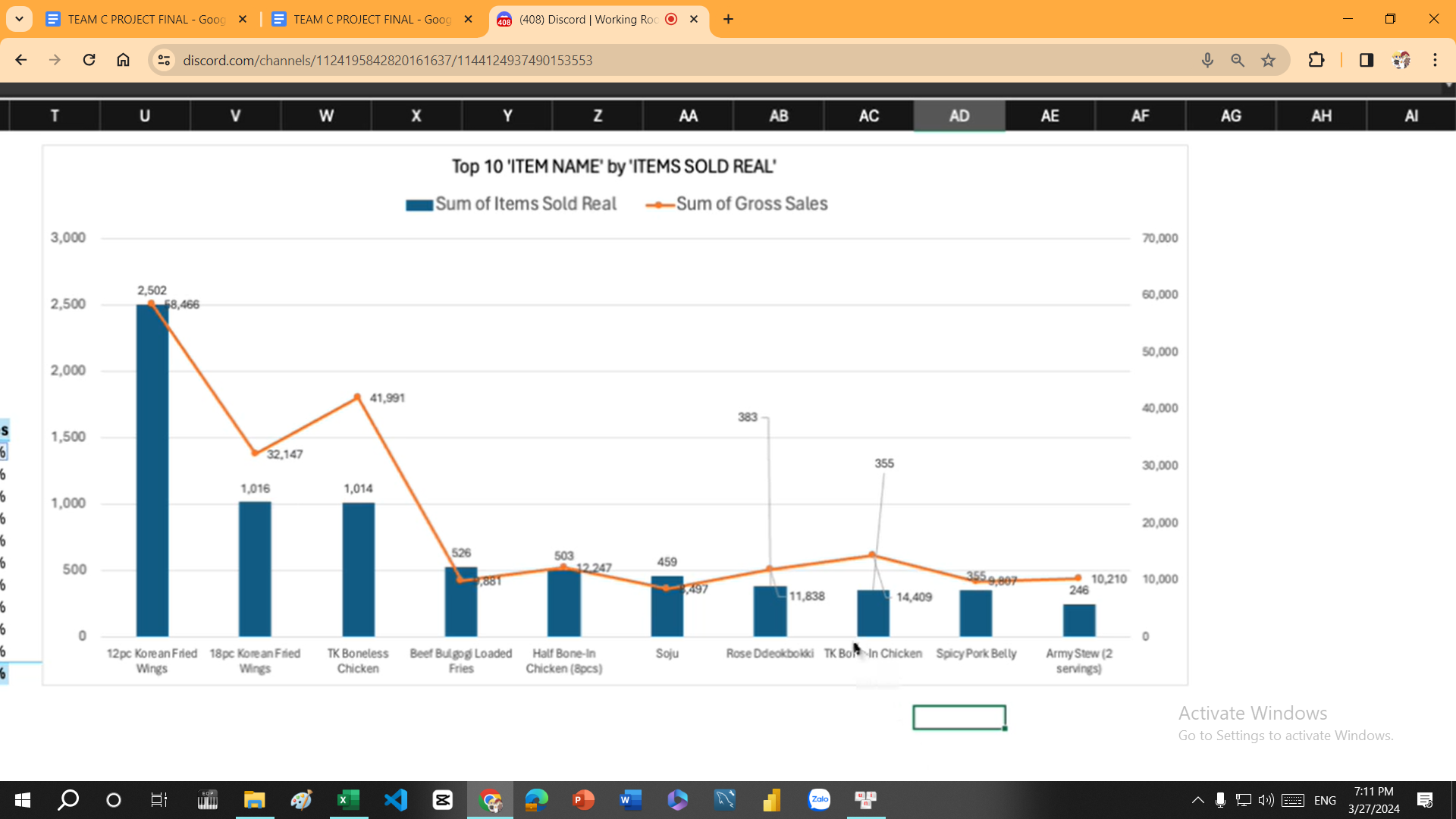
**Task 2:** Nên dùng Net Sales, hạn chế dùng Gross Sales

Họ toàn quan tâm doanh thu sau cùng

Task 2 ghi nhận xét sai

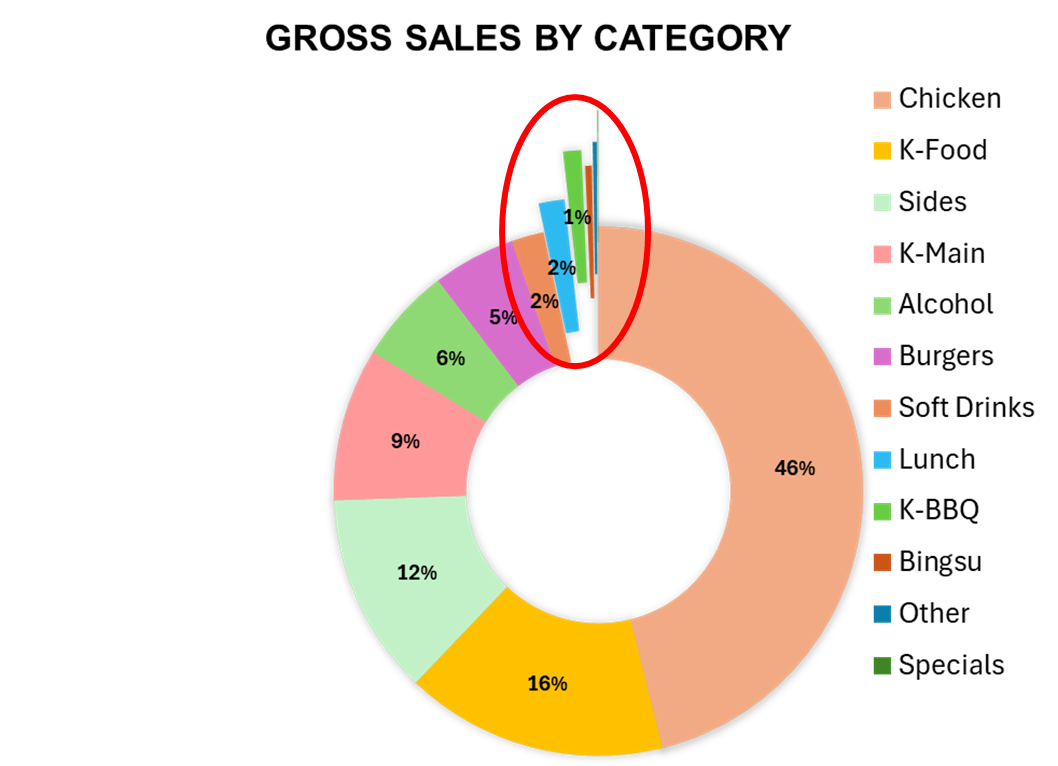
Dùng 2 trục này nhưng lại ko ghi đơn vị, ko biết dữ liệu nào dùng bên cột 2.

Có mấy cái cột nhìn bằng nhau => Nên có Data Label cho từng điểm.



**Task 4**: Màu gần giống nhau. Nước ngoài nhiều người mù màu. Đối với 1 số nhà phân tích, họ ko thích để nhiều râu ria.

12% 6% 1% là giảm dần nhưng màu lại đậm dần



[**TÓM TẮT 1**](#_i1iu14tdlqrj)

[**Intermediate Tasks 4**](#_4k99037d0eo)

[Task 1: Monthly Sales Patterns 4](#_bpnwnua66uo5)

[1.1. Khảo sát 4](#_1cz06w15be21)

[1.2. Làm sạch dữ liệu 5](#_nlcfwcs5k32o)

[1.3. Phân tích 5](#_pnuc337fn88)

[Task 2: Top 10 Product Performance 7](#_bpnwnua66uo5)

[2.1. Khảo sát 7](#_4sdwj61xofs5)

[2.2. Làm sạch dữ liệu 7](#_tumvpcg7696s)

[2.3. Phân tích 8](#_19q5leqm937t)

[Task 3: Sales by Day Analysis 10](#_hw3umy4azqyu)

[3.1. Khảo sát 10](#_1cumvr594ui6)

[3.2. Làm sạch dữ liệu 10](#_1307gyq3efbd)

[3.3. Phân tích 10](#_7utykzp9ormn)

[Task 4: Category Sales Contribution 12](#_tokpbrf1t247)

[4.1. Khảo sát 12](#_8efml2mjd6of)

[4.2. Phân tích 12](#_ui8weal0vux6)

[Task 5: Peak Hours Identification 14](#_82oish3vebx5)

[5.1. Khảo sát 14](#_cj3kbcpukqfa)

[5.2. Phân tích 15](#_2je681mi04lj)

[Task 6: Payment Method Preferences 17](#_wwwfjiyxxj1m)

[6.1. Khảo sát 17](#_869z8wbvfqye)

[6.2. Làm sạch dữ liệu 17](#_przprxtxlec9)

[6.3. Phân tích 19](#_5flwgwba37q)

[Task 7: Source of Payment 21](#_eakz90gjj5zi)

[7.1. Khảo sát 21](#_msnormjmdhft)

[7.2. Phân tích 21](#_8slbd1kgccfm)

[**Advanced Tasks 23**](#_g0n3fjfbedlu)

[Task 8: Correlation Between Website Traffic and Sales 23](#_7kop36wvv6x5)

[8.1. Khảo sát 23](#_q2m5bl35hgrr)

[8.2. Làm sạch dữ liệu 23](#_jagy2cgydupu)

[8.3. Phân tích 24](#_ve2e0zh20hh6)

[Task 9: Predictive Sales Analysis: 26](#_i41vtw9cps65)

# 

# 

# Intermediate Tasks

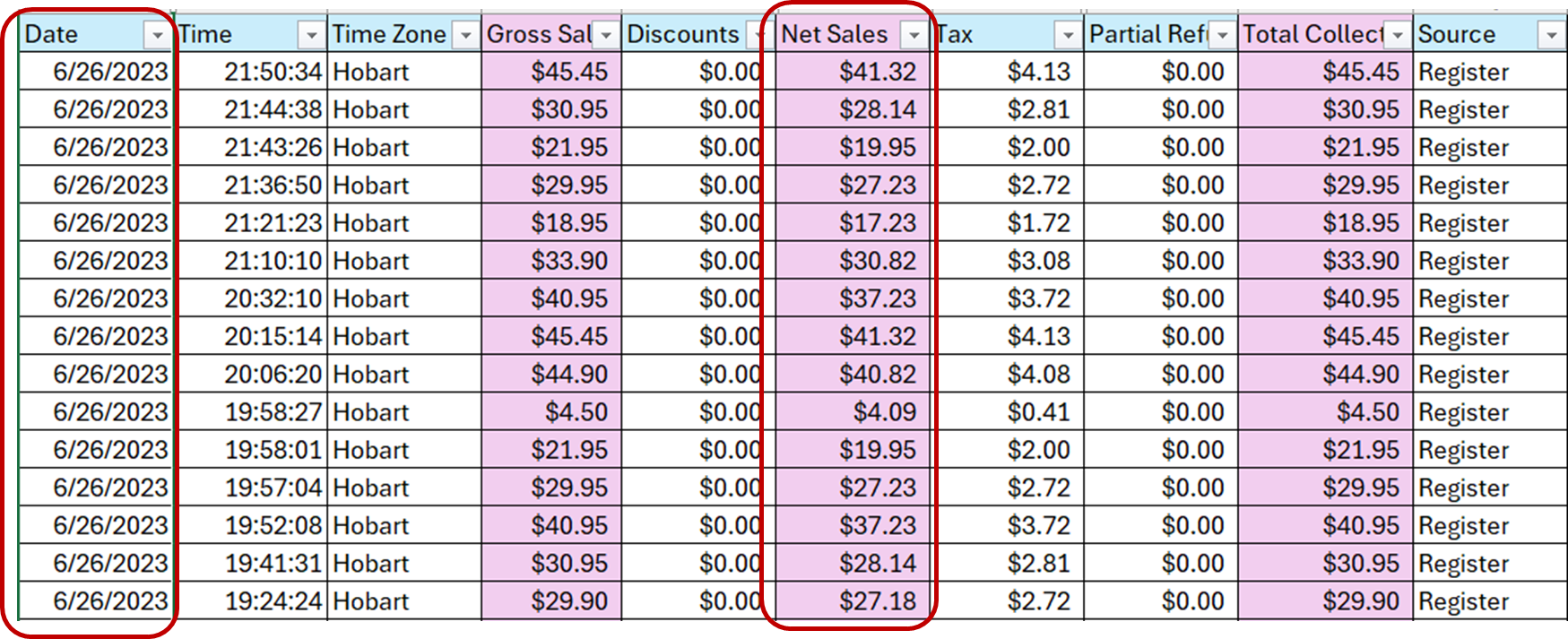
## Task 1: Monthly Sales Patterns

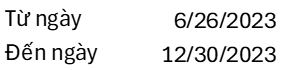
“**Monthly Sales Patterns:** Dive into the last 6 months of sales data to identify patterns, including peak and trough periods. Understanding these trends is crucial for our inventory and marketing strategies.”

### 1.1. Khảo sát

**“Mô hình bán hàng hàng tháng**: Đi sâu vào dữ liệu bán hàng trong 6 tháng qua để xác định mô hình, bao gồm cả giai đoạn cao điểm và đáy. Hiểu được những xu hướng này là rất quan trọng đối với chiến lược tiếp thị và hàng tồn kho của chúng tôi.”

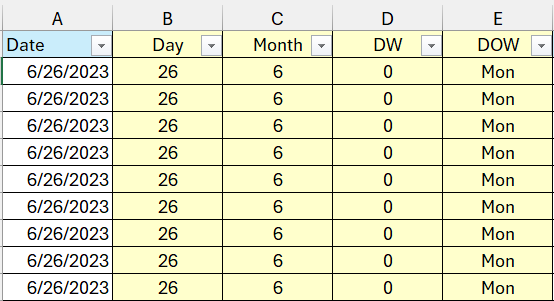
Để phân tích, ta quan tâm đến dữ liệu thời gian (Date) và Doanh thu thuần (Net Sales) trong bảng “**Transactions Data”**

****



Dùng hàm Min, Max biết được thời gian trong dataset là từ 26-06 đến 30-12 năm 2023. Vì tháng 6 ko đủ dữ liệu nên với task này, nhóm chỉ phân tích 6 tháng cuối năm (tháng 7 -> tháng 12).

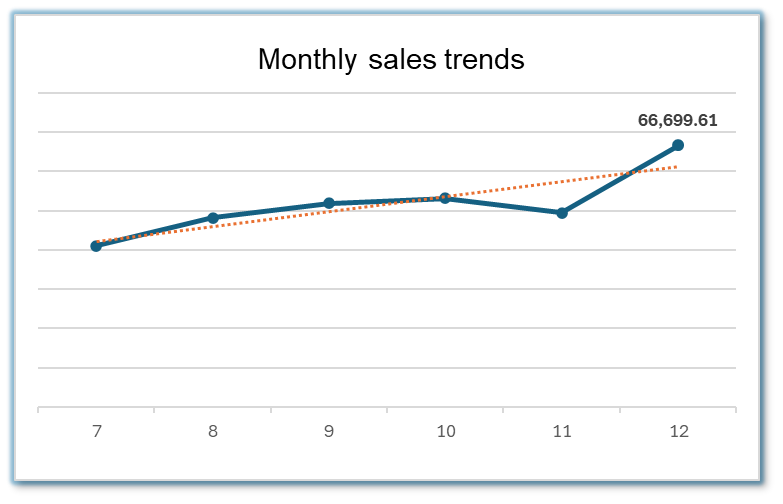
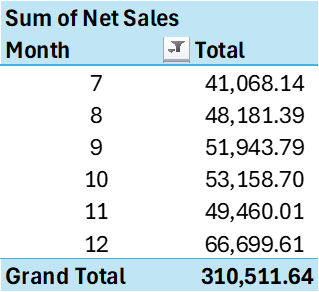
### 1.2. Làm sạch dữ liệu



Ở đây, team C tạo 2 cột mới “Day” và “Month” để tách riêng từng thành phần trong “Date”

### 1.3. Phân tích

Dùng công cụ Pivot Table, tạo bảng tính tổng doanh thu thuần theo từng tháng. Dùng bộ lọc, bỏ qua dữ liệu tháng 6. Từ đó vẽ thêm hình minh họa, biểu diễn bằng biểu đồ đường.





**Báo cáo:**

* Tổng doanh thu bán hàng ròng trong 6 tháng qua là **310,611.64 $.**
* Doanh thu trung bình mỗi tháng là **51,771.94 $.**
* Tháng có doanh thu cao nhất là tháng **12** (66,699.61 $), thấp nhất là tháng **7** (41,068.14 $).
* Xu hướng kinh doanh của nhà hàng theo 6 tháng gần nhất: doanh thu ổn định, có xu hướng phát triển tăng dần theo thời gian.

**Phân tích:**

* **Xu hướng tăng dần:** Doanh thu bán hàng có xu hướng tăng dần từ tháng 7 đến tháng 10, cho thấy nhu cầu của khách hàng tăng cao trong giai đoạn này.
* **Tháng cao điểm:** Tháng 12 là tháng có doanh thu cao nhất, có thể do ảnh hưởng của các dịp lễ cuối năm như lễ Giáng Sinh, Tết dương lịch. Các hội nhóm có xu hướng ăn mừng tổng kết năm, thường chọn các địa điểm ăn uống để tụ tập. Trẻ em ở Úc còn được nghỉ hè từ tháng 12, nên các gia đình có thể cùng con tới trải nghiệm ẩm thực ở các nhà hàng.
* **Tháng thấp điểm:** Tháng 7 là tháng có doanh thu thấp nhất, có thể do đây là do thời tiết rất lạnh, nên khách hàng ngại ra ngoài hàng ăn. Tháng 7 ở Úc là đông, mọi người có thể có xu hướng tích trữ đồ, và ngồi trong ngôi nhà ấm cúng.

**Đánh giá:**

* Nhu cầu của khách hàng có xu hướng tăng cao vào cuối năm.
* Nhà hàng cần chuẩn bị đủ lượng nguyên liệu để đáp ứng nhu cầu tăng cao vào cuối năm.
* Nhà hàng nên triển khai các chương trình marketing tập trung vào các dịp lễ cuối năm để thu hút khách hàng.

**Đề xuất:**

* Tháng 12, nhà hàng khá đông khách, có thể thuê thêm nhân viên part time để đáp ứng đủ nhu cầu thực khách vào dịp cao điểm này.
* Thường xuyên đánh giá hiệu suất của từng chiến dịch Marketing mang lại.

## Task 2: Top 10 Product Performance

**“Top 10 Product Performance:** Analyze which products have been our top performers in terms of revenue. This insight will guide our product development and promotional efforts.”

### 2.1. Khảo sát

**“Hiệu suất sản phẩm của 10 sản phẩm hàng đầu**: Phân tích những sản phẩm nào mang lại hiệu quả cao nhất về mặt doanh thu. Cái nhìn sâu sắc này sẽ hướng dẫn các nỗ lực quảng cáo và phát triển sản phẩm của chúng tôi.”

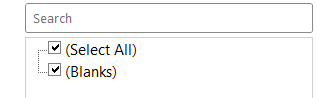
10 sản phẩm hàng đầu sẽ là 10 sản phẩm bán chạy nhất, ta cần dữ liệu về:

* Các sản phẩm **Item Name**, loại sản phẩm **Category**
* Số lượng hàng đã bán ra (**Items Sold**)
* Doanh thu thu được từ các sản phẩm đó (**Gross Sales**).

### 2**.2. Làm sạch dữ liệu**

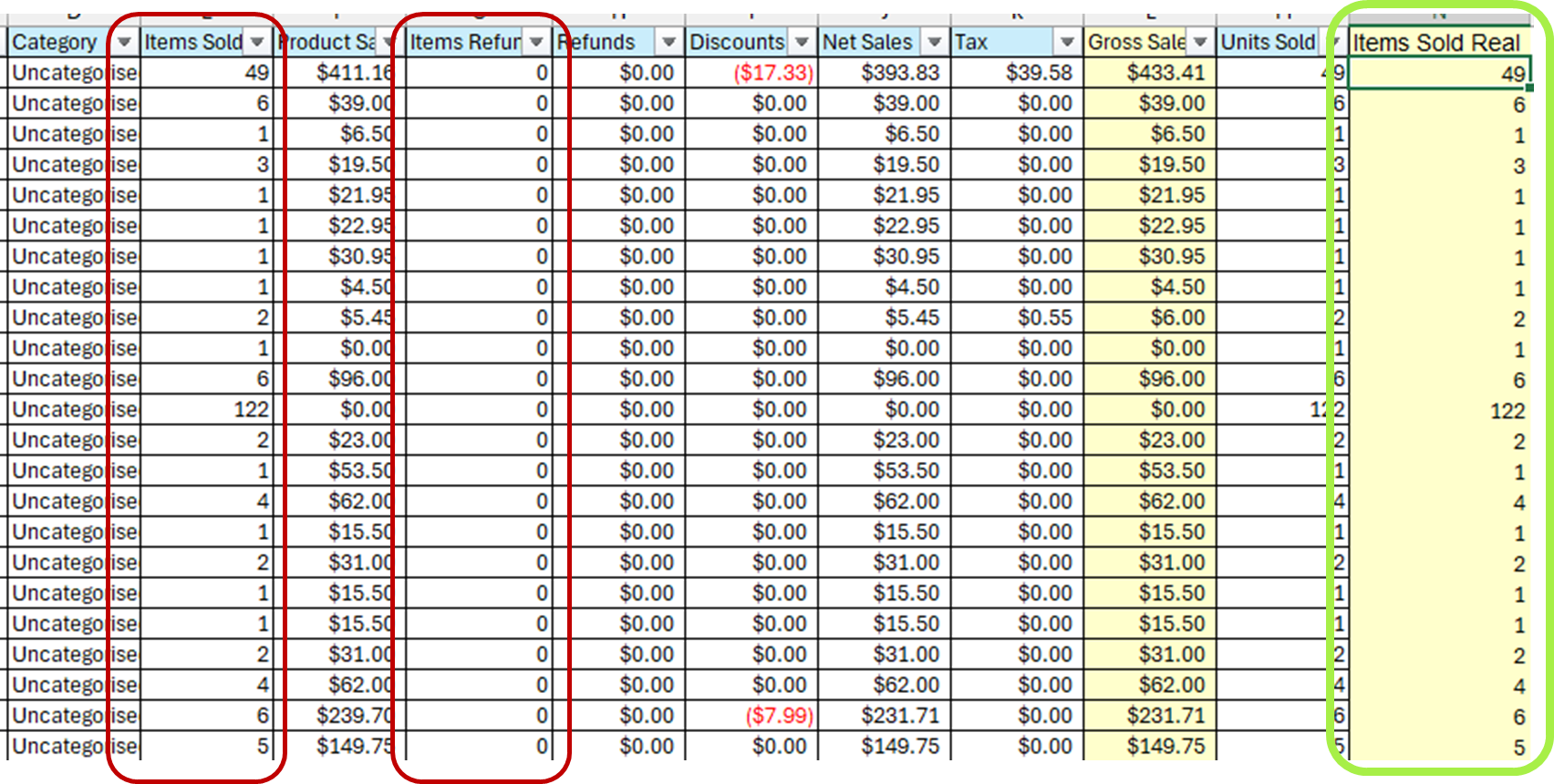
Xét data trong bảng **“Items Sale Data”**

Trường dữ liệu “SKU” là dữ liệu trống, nên ẩn trường này đi.



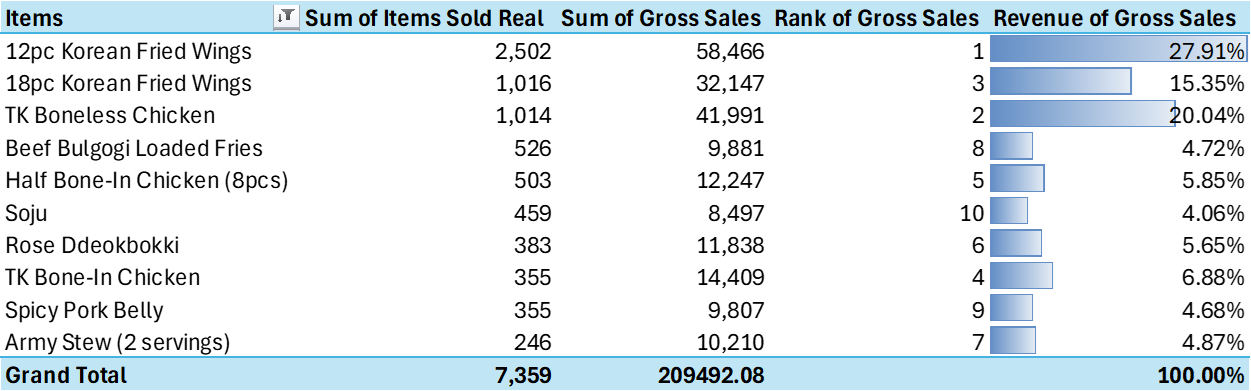
Team thêm một trường mới là “**Items Sold Real**”, tính số lượng hàng thực sự đã bán ra ngoài thị trường, bằng công thức số lượng hàng đã bán + số lượng hàng trả về:

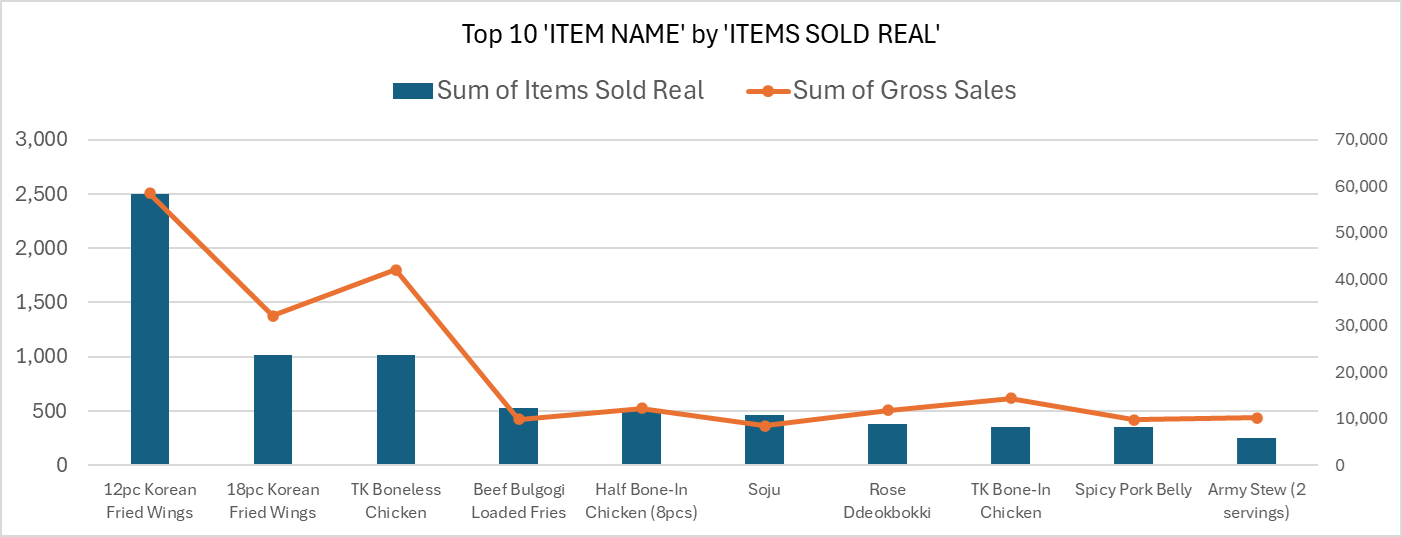
**Items Sold Real** = **Items Sold** + **Items Refunded**



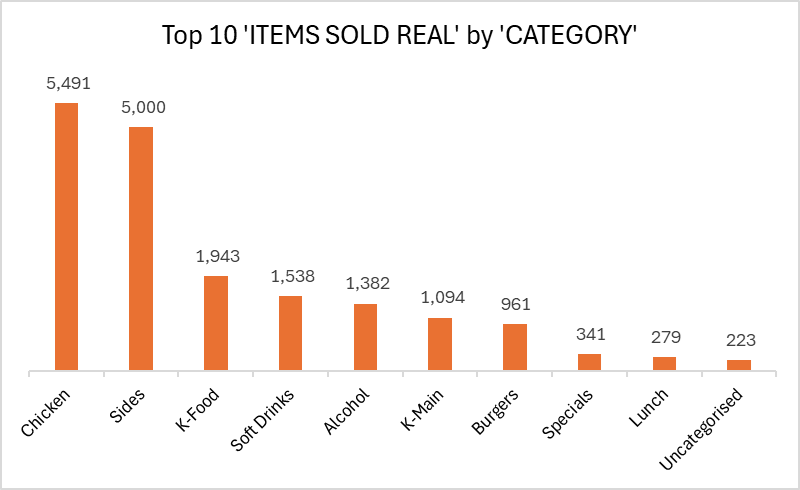
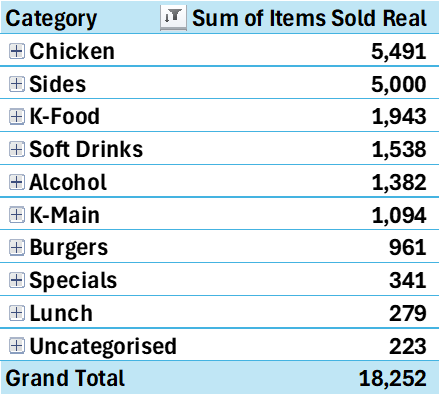
### 2.3. Phân tích

Team C, tạo Pivot table thể hiện tổng doanh số và doanh thu của từng các món ăn. Sử dụng bộ lọc để lấy Top 10 sản phẩm có doanh số cao nhất. Tính thêm “**Rank of Goss Sales**” để show thứ hạng doanh thu của từng món ăn. Tính thêm “**Revenue** **of Goss Sales**” tỷ lệ đóng góp doanh thu của Top 10 món trên.





Team C, tiếp tục dùng Pivot table thống kê 10 dòng sản phẩm bán chạy nhất.



**Báo cáo:**

1. **Top danh mục về số lượng sản phẩm bán được:**

Chicken (Gà) và Sides (Món phụ) là hai danh mục hàng đầu với số lượng sản phẩm bán được cao nhất, đóng góp lớn vào tổng doanh số bán hàng.

1. **Danh mục có đóng góp cao vào doanh thu:**

Để đưa ra kết luận về đóng góp vào doanh thu, chúng ta cần dữ liệu về doanh thu tương ứng với từng danh mục. Trong trường hợp này, dữ liệu chỉ cho biết số lượng sản phẩm bán được (Items Sold), không có thông tin về doanh thu.

1. **Danh mục cần chú ý đến:**

Lunch (Bữa trưa), Uncategorised (Chưa phân loại), K-BBQ, và Bingsu có số lượng sản phẩm bán được thấp so với các danh mục khác. Điều này có thể yêu cầu xem xét lại chiến lược quảng cáo, phát triển sản phẩm, hoặc các chiến dịch khuyến mãi đối với những danh mục này.

1. **Chiến lược phân phối sản phẩm:**

Dựa vào số lượng sản phẩm bán được, có thể cần xem xét lại chiến lược phân phối sản phẩm để đảm bảo sự cân đối và hiệu quả cho mỗi danh mục.

1. **Tiềm năng phát triển sản phẩm:**

Có thể tập trung vào việc phát triển sản phẩm trong các danh mục có số lượng sản phẩm bán được thấp như K-BBQ và Bingsu để tăng cường doanh số và đa dạng hóa sản phẩm.

## Task 3: Sales by Day Analysis

**Sales by Day Analysis**: Determine the most lucrative days of the week, enabling us to optimize our staffing and promotional activities.

### 3.1. Khảo sát

**Phân tích doanh số theo ngày**: Xác định những ngày sinh lợi nhất trong tuần, cho phép chúng tôi tối ưu hóa các hoạt động nhân sự và quảng cáo.

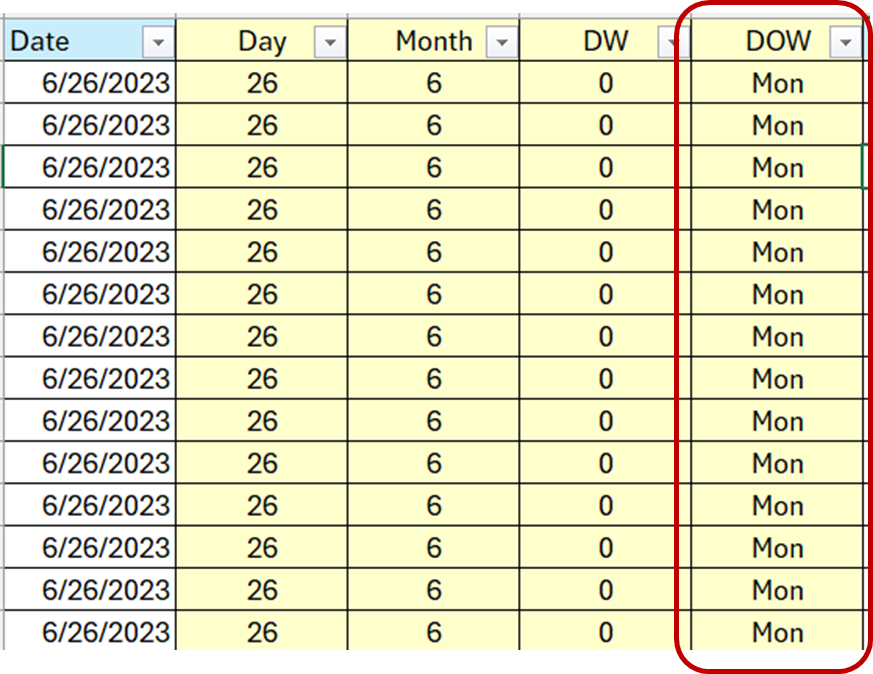
Ta cần dữ liệu về:

* Thời gian => các ngày trong tuần.
* Doanh thu

### 3.2. Làm sạch dữ liệu

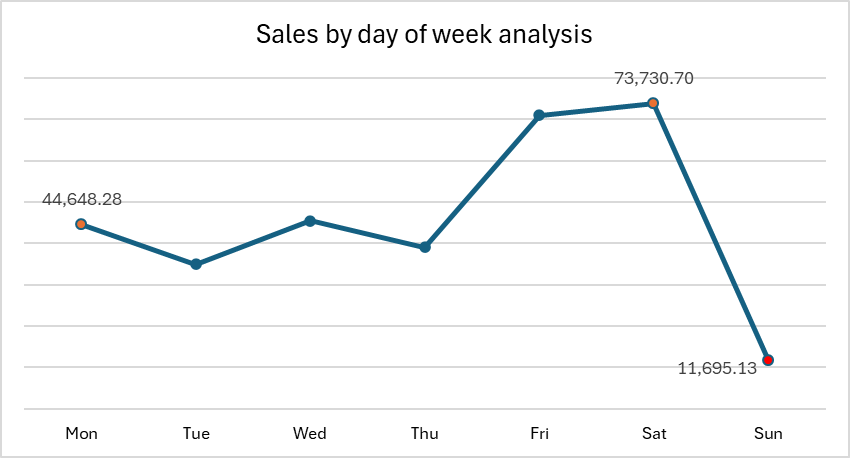
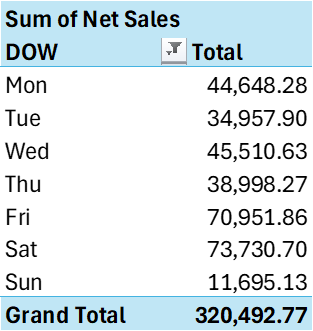
Trong bảng “**Transactions Data”**

Tạo thêm dữ liệu về các ngày trong tuần.



### 3.3. Phân tích

Dùng công cụ Pivot Table, tạo bảng tính tổng doanh thu thuần theo từng ngày trong tuần. Từ đó vẽ thêm hình minh họa, biểu diễn bằng biểu đồ đường.



**Phân tích:**

* **Thứ Bảy** là ngày có doanh thu cao nhất, **gần gấp đôi** so với các ngày còn lại.
* **Thứ Sáu** có doanh thu cao thứ hai, **cao hơn 50%** so với mức trung bình.
* Doanh thu các ngày **Thứ Hai, Ba, Tư, Năm** tương đối thấp và **bằng nhau**.
* **Chủ Nhật** có doanh thu thấp nhất.

**Nhận định:**

* Nhu cầu mua sắm của khách hàng cao nhất vào **cuối tuần** (Thứ Sáu và Thứ Bảy).
* Doanh nghiệp cần tập trung nguồn lực nhân sự vào **cuối tuần** để đảm bảo phục vụ khách hàng tốt nhất.
* Có thể triển khai các chương trình khuyến mãi vào **Thứ Hai, Ba, Tư, Năm** để kích thích nhu cầu mua sắm của khách hàng.

**Đề xuất:**

* Phân tích doanh thu theo giờ trong ngày để xác định thời điểm khách hàng đến mua sắm nhiều nhất.
* Phân tích doanh thu theo từng khu vực, kênh bán hàng để có chiến lược phù hợp.
* Nghiên cứu hành vi mua hàng của khách hàng để hiểu rõ hơn về nhu cầu của họ.

## 

## Task 4**:** Category Sales Contribution

**Category Sales Contribution**: Assess how different product categories contribute to our overall sales, aiding in inventory and category management decisions.

### 4.1. Khảo sát

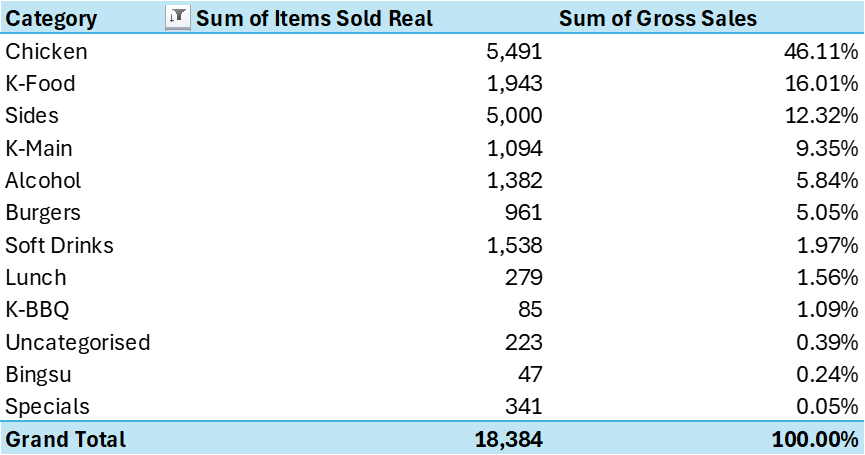
**Đóng góp bán hàng theo danh mục**: Đánh giá mức độ đóng góp của các danh mục sản phẩm khác nhau vào doanh số bán hàng tổng thể của chúng tôi, hỗ trợ các quyết định quản lý danh mục và hàng tồn kho.

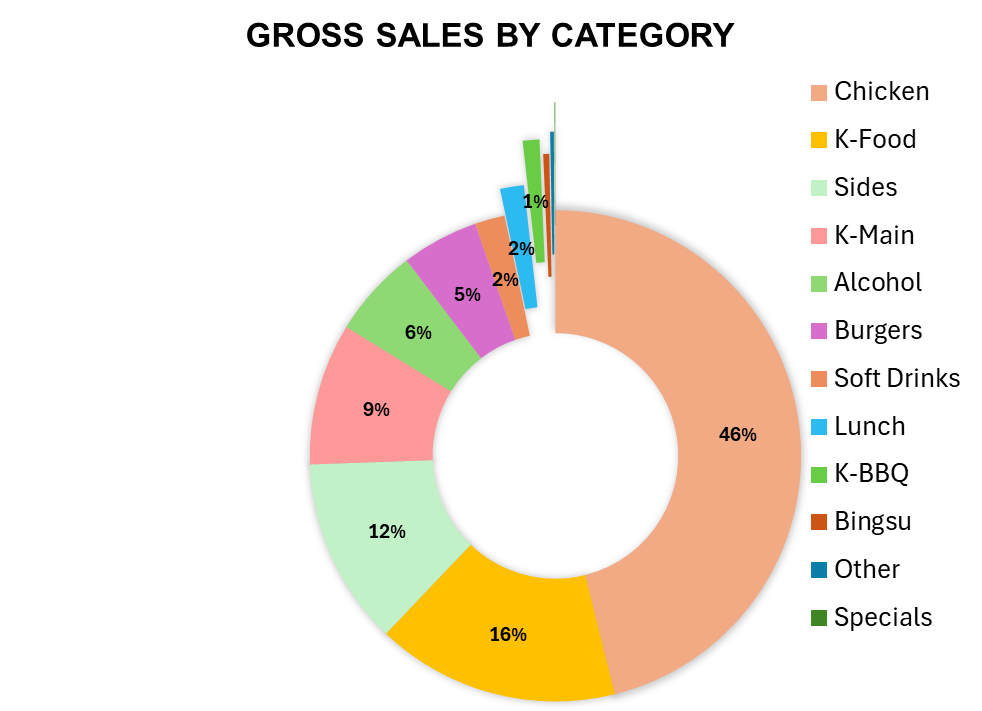
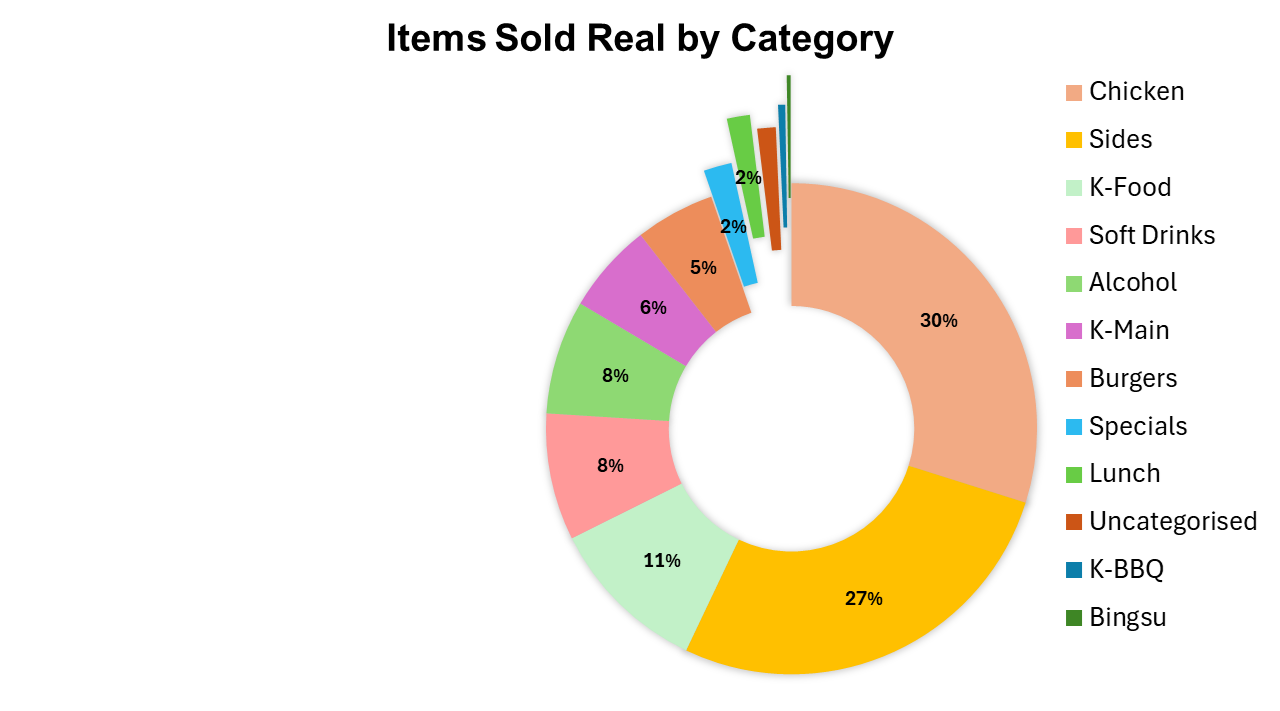
Ta cần dữ liệu về:

* Thông tin sản phẩm
* Thông tin danh mục sản phẩm
* Doanh thu theo từng sản phẩm

### 4.2. Phân tích

Tương tự Task 2, team C tạo bảng Pivot table lấy top 12 dòng sản phẩm bán chạy nhất, tính doanh thu tương ứng với từng sản phẩm. Để hiểu rõ hơn về tỷ lệ đóng góp của từng loại, chọn hiển thị doanh thu theo phần trăm.





**Báo cáo:**

* Dựa trên biểu đồ "**Gross Sales by Category**", các sản phẩm từ thịt gà được tính toán chiếm tỷ trọng cao nhất (**46%**) trong tổng doanh thu trong 6 tháng cuối năm. Điều này cho thấy các món ăn chế biến từ thịt gà rất được thực khách yêu thích khi đến nhà hàng. Do đó, cần sử dụng nguồn thịt gà với số lượng lớn.
* Bên cạnh thịt gà, danh mục Các món ăn phụ (Sides) cũng có doanh số với tỷ trọng (**27%**) gần bằng với tỷ trọng của thịt gà (**30%**) được thể hiện trên biểu đồ "**Items Sold Real by Category**". Mặc dù đây không phải là mặt hàng có tác động lớn nhất đến tổng doanh thu, nhưng lại là mặt hàng được khách hàng gọi nhiều.

**Đề xuất:**

Bên cạnh việc bổ sung thêm các món ăn từ thịt gà, nhà hàng cũng nên bổ sung thêm nhiều sản phẩm vào danh mục Các món ăn kèm.

## 

## Task 5: Peak Hours Identification

**Peak Hours Identification**: Identify our peak sales hours to better align our staffing and inventory management with customer demand.

### 5.1. Khảo sát

**Xác định giờ cao điểm**: Xác định giờ bán hàng cao điểm của chúng tôi để điều chỉnh tốt hơn việc quản lý nhân sự và hàng tồn kho của chúng tôi cho phù hợp với nhu cầu của khách hàng.

Ta cần dữ liệu về:

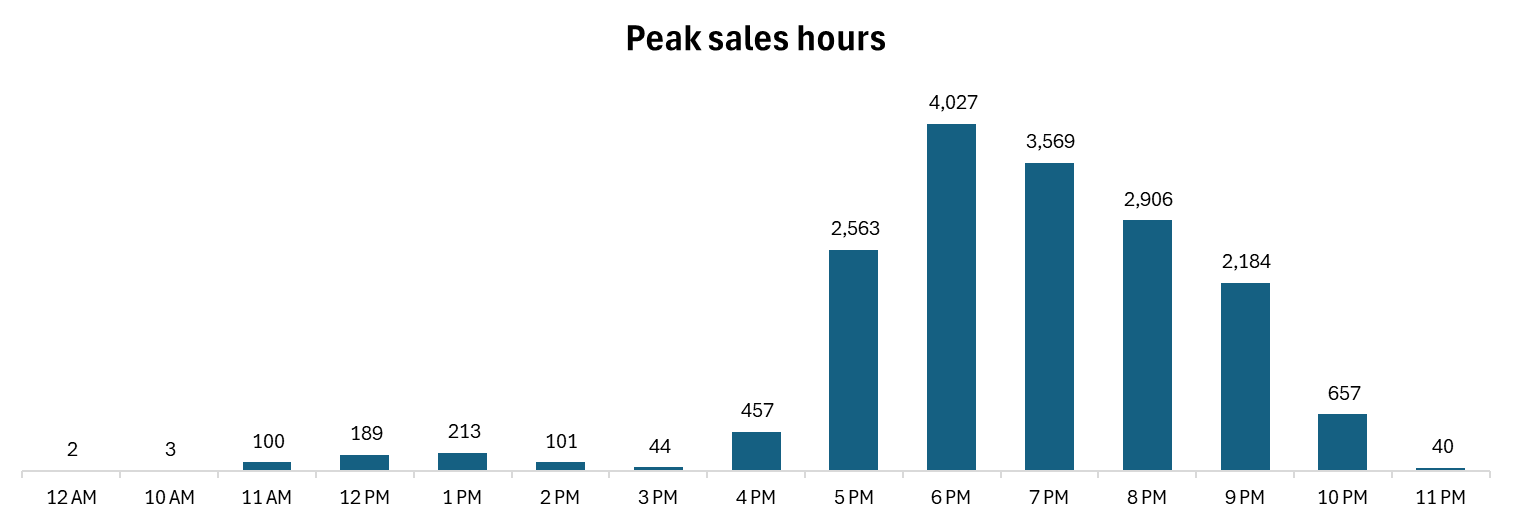
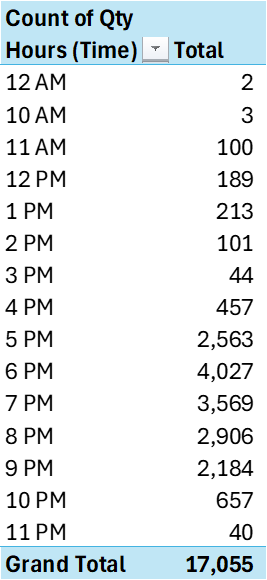
* Thời gian trong ngày (**Time**).
* Số lượng các món khách gọi (**Qty** trong bảng “**Items Data”**).



### 

### 5.2. Phân tích

Từ trường dữ liệu **Time** và **Qty** trong bảng “**Items Data”,** tạo bảng Pivot Table, thể hiện doanh số bán hàng theo từng giờ. Sủ dụng biểu đồ cột để minh họa.



**Phân tích:**

* Giờ cao điểm bán hàng là từ **4 giờ chiều đến 10 giờ tối**, với lượng sản phẩm bán ra cao nhất **từ 5 giờ chiều đến 9 giờ tối**.
* Lượng sản phẩm bán ra thấp nhất vào **ban ngày**, từ **12 giờ sáng đến 3 giờ chiều**

**Nhận định:**

* Khách hàng có xu hướng ăn hàng nhiều nhất vào **buổi tối** và **buổi chiều.**
* Nhà hàng cần tập trung nguồn lực đầu bếp, phục vụ vào **giờ cao điểm** để đảm bảo phục vụ khách hàng tốt nhất.
* Cần dự trữ lượng hàng hóa **đầy đủ**, chế biến sẵn, sơ qua, cho **giờ cao điểm** để tránh tình trạng không kịp chuẩn bị.

**Đề xuất:**

* Phân tích doanh thu theo giờ trong ngày để xác định thời điểm khách hàng đến mua sắm nhiều nhất.
* Phân tích doanh thu theo từng khu vực, kênh bán hàng để có chiến lược phù hợp.
* Nghiên cứu hành vi mua hàng của khách hàng để hiểu rõ hơn về nhu cầu của họ.

## 

## Task 6: Payment Method Preferences

**Payment Method Preferences**: Investigate customer preferences in payment methods to enhance the checkout experience.

### 6.1. Khảo sát

**Tùy chọn phương thức thanh toán**: Điều tra tùy chọn của khách hàng trong các phương thức thanh toán để nâng cao trải nghiệm thanh toán.

Ta cần dữ liệu về:

* Các hình thức thanh toán:
  1. Card (Thẻ thanh toán)
  2. Card Entry Methods (Thẻ quẹt)
  3. Cash (Tiền mặt)
  4. Square Gift Card (Thẻ quà tặng)
  5. Other Tender
* Tổng số tiền thực tế được khách hàng thanh toán (**Total Collected**)

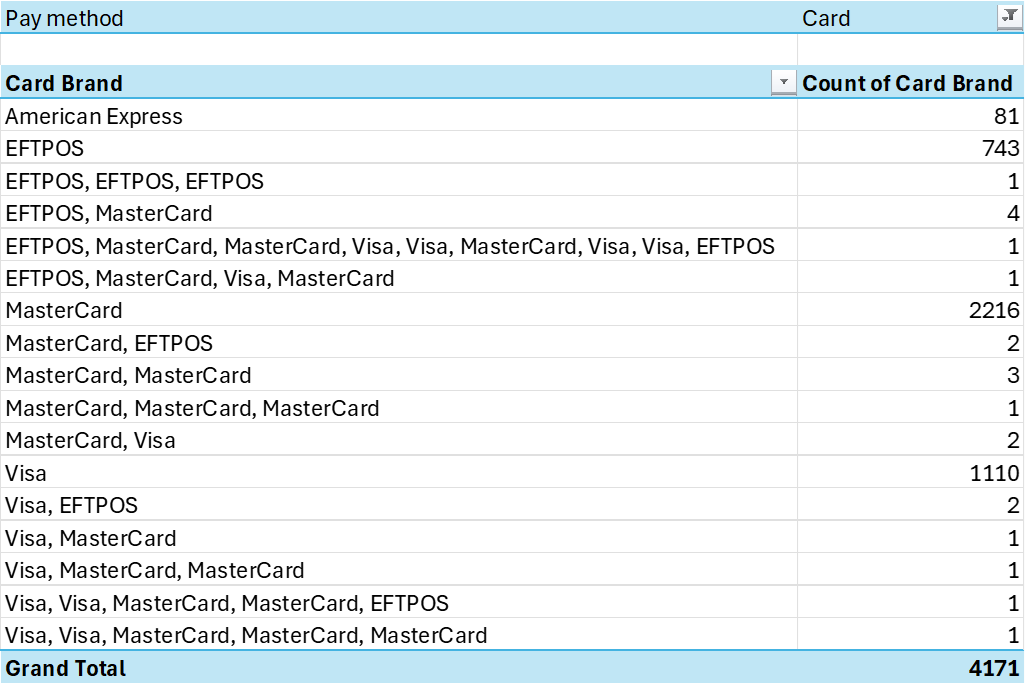
### 6.2. Làm sạch dữ liệu

5 trường **Card, Card Entry Methods, Cash, Square Gift Card, Other Tender**  trong bảng “**Transactions Data”** đều đang được hiện thị dưới số tiền thanh toán theo từng loại. Chưa có trường nào phân loại giao dịch theo 5 hình thức thanh toán trên. Vậy nên nhóm tạo thêm trường dữ liệu “**Pay method**”. Loại hình thanh toán sẽ tương ứng với giá trị tiền lớn hơn 0 ở 5 trường trên.

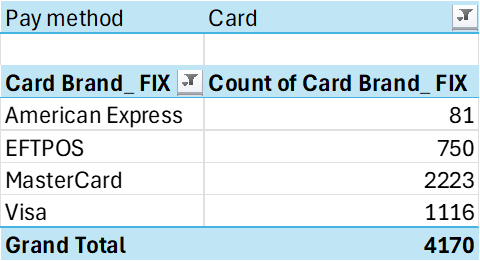


Nếu muốn phân tích sâu vào loại thẻ thanh toán, ta cần tìm hiểu các brand của thẻ đó. Tuy nhiên trường “**Card Brand**” bị lỗi hiển thị quá nhiều brand thẻ trên 1 giao dịch. Team làm sạch dữ liệu bằng cách lấy brand đầu tiên mà khách liệt kê.

**After**



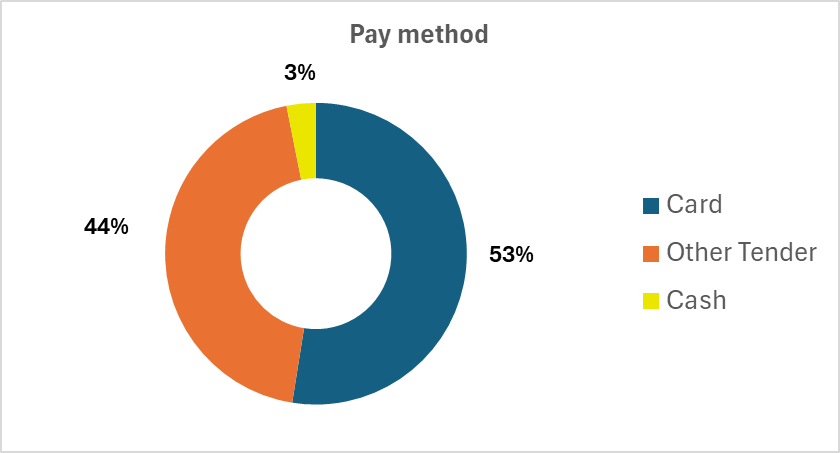
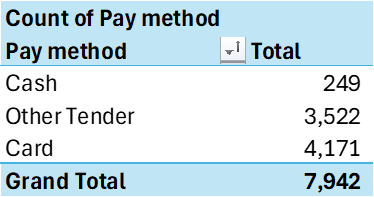
**Before**

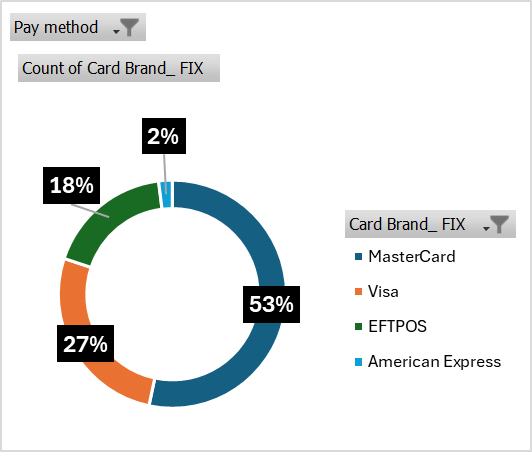
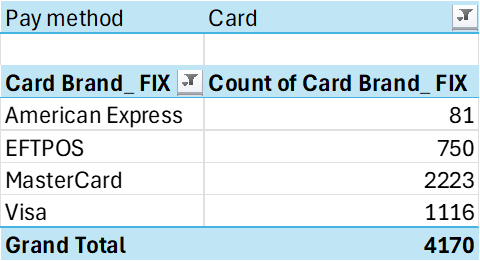
****

### 

### 6.3. Phân tích

Team C dùng Pivot Table để tính số lượng giao dịch trên mỗi phương thức thanh toán





**Phân tích Task 6:**

**Phương thức thanh toán yêu thích:**

* **Thẻ** là phương thức thanh toán được sử dụng phổ biến nhất, chiếm **52,63%** tổng doanh thu. Trong đó loại dùng nhiều nhất là MasterCard (**28%**) và Visa (**12.5%**)
* **Thanh toán khác** (ví điện tử, chuyển khoản ngân hàng) là phương thức thanh toán được sử dụng thứ hai, chiếm **44%** tổng doanh thu.
* **Tiền mặt** chỉ chiếm **3%** tổng doanh thu.

**Nhận định:**

* Khách hàng có xu hướng sử dụng thẻ và ví điện tử để thanh toán, vì tính tiện lợi, ko phải mang theo nhiều tiền ra ngoài. Việc share bill trở nên dễ dàng hơn.
* Đa số khách đến ăn nhà hàng là người nước ngoài, nên việc họ thường dùng MasterCard và Visa. Khách hàng ít dùng tiền mặt vì đổi tiền ngoại tệ có thể chịu một khoản chi phí ko mong muốn.

**Đề xuất:**

Cung cấp nhiều phương thức thanh toán đa dạng để đáp ứng nhu cầu của khách hàng.

Tích hợp các dịch vụ thanh toán trực tuyến phổ biến, uy tín.

Đưa ra các chương trình khuyến mãi cho các phương thức thanh toán ít được sử dụng.

Phân tích hành vi mua hàng của khách hàng theo từng phương thức thanh toán để có chiến lược phù hợp.

## 

## Task 7: Source of Payment

**Source of Payment**: Delve deeper into the source of payments to understand our revenue streams better.

### 7.1. Khảo sát

**Nguồn thanh toán:** Tìm hiểu sâu hơn về nguồn thanh toán để hiểu rõ hơn về nguồn doanh thu của chúng tôi.

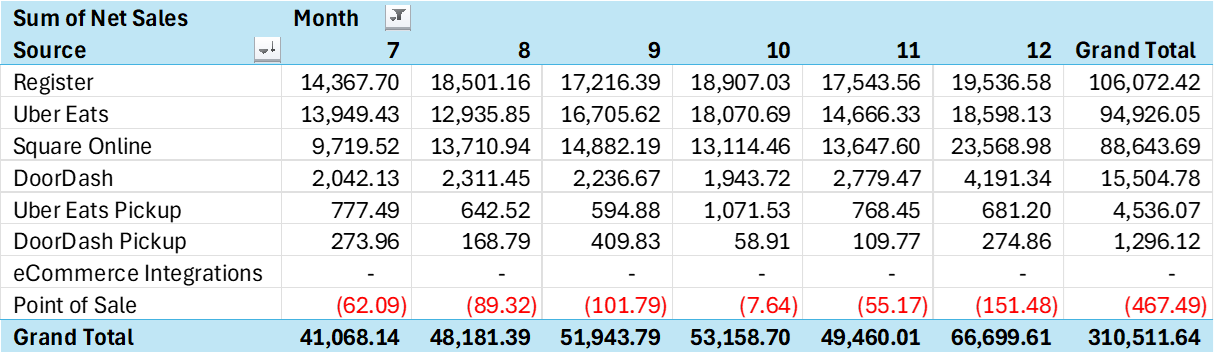
Team sẽ tính xem nguồn doanh thu cao nhất là qua hình thức bán hàng nào.

1. Đơn đặt tại quầy
   1. Register
   2. Point of Sale
2. Square Online (trang web bán hàng trực tuyến do Square cung cấp).
3. Các ứng dụng giao hàng bên thứ ba:
   1. Uber Eats
   2. Uber Eats Pickup
   3. DoorDash
   4. DoorDash Pickup
4. ECommerce Integrations (các nền tảng thương mại điện tử khác được tích hợp với hệ thống thanh toán của cửa hàng).

### 7.2. Phân tích

Xét bảng “**Transactions Data”**

Team C dùng Pivot Table để tạo bảng tính tổng doanh thu thuần (**Net Sales**) theo từng nguồn khách hàng mua hàng (**Source**).



Team C vẽ hình minh họa áp dụng cho những nguồn mang lại nguồn doanh thu dương.



**Phân tích:**

* **Register** là nguồn doanh thu phổ biến nhất, chiếm **34%** tổng doanh thu.
* Theo sau đó là **Uber Eats** chiếm **31%** và **Square Online** với **29%** tổng doanh thu.
* Tuy chỉ chiếm 1 phần nhỏ, nhưng **Uber Eats Pickup** và **DoorDash** đã mang lại 6% doanh thu cho nhà hàng.
* Kinh doanh qua sàn thương mại điện tử chưa tốt lắm, ko mang lại doanh thu cho nhà hàng.
* **Point of Sale**: doanh thu từ việc bán hàng trực tiếp, có xu hướng âm.

**Nhận định:**

* Ứng dụng **Uber** “phủ sóng” tại nhiều thành phố lớn của nước Úc.
* Ở các cửa hàng, doanh thu từ việc bán hàng trực tiếp, ko bù nổi tiền thuê mặt bằng, nên có xu hướng âm.

**Đề xuất:**

* Tăng nhận diện thương hiệu, marketing tên tuổi các trang mạng xã hội, sàn điện tử.
* Làm cho mặt tiền các cửa hàng trở nên nổi bật, ko bị bão hòa với các nhà hàng khác. Biển hiệu màu nổi sẽ thu hút người qua đường, kích thích sự thèm ăn.

# Advanced Tasks

## Task 8: Correlation Between Website Traffic and Sales

**Correlation Between Website Traffic and Sales:** Utilize Excel’s advanced correlation functions to explore the relationship between our website traffic and sales figures, providing insights into the effectiveness of our online presence.

### 8.1. Khảo sát

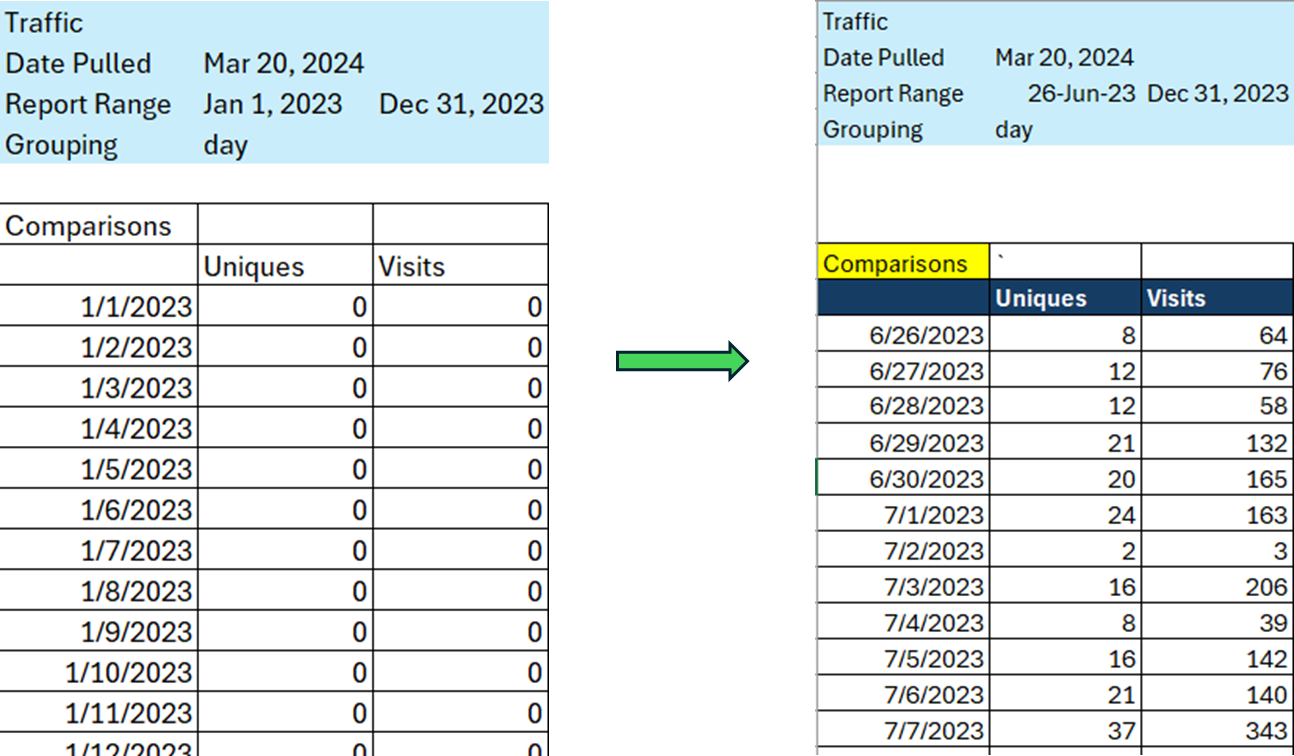
**Mối tương quan giữa lượng truy cập trang web và doanh số bán hàng:** Sử dụng các hàm tương quan nâng cao của Excel để khám phá mối quan hệ giữa lưu lượng truy cập trang web và số liệu bán hàng của chúng tôi, cung cấp thông tin chi tiết về hiệu quả của sự hiện diện trực tuyến của chúng tôi.

Cần những dữ liệu sau:

* Số lượng truy cập = số lượt truy cập lần đầu + số lượt quay lại truy cập lần 2+.
* Doanh thu theo từng ngày.

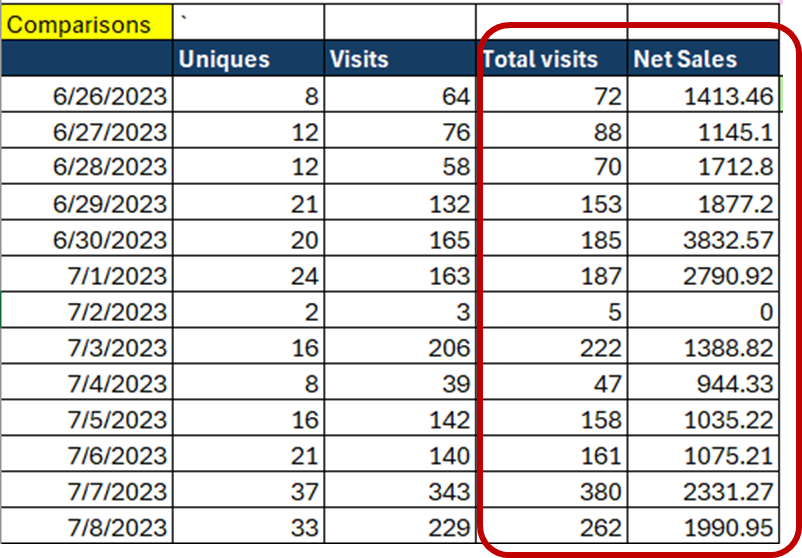
### 8.2. Làm sạch dữ liệu

Dữ liệu bảng “Transactions Data” cho từ ngày 26 tháng 6 đến 30 tháng 12. Nên để tính doanh thu theo thời gian, ta sẽ chỉ lấy dữ liệu từ ngày ngày 26 tháng 6 đến 30 tháng 12 trong bảng “Traffic Summary Data”



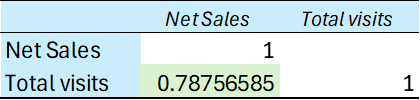
Tạo thêm 2 cột mới là

1. “**Total visits” = “Unique” + “Visits”**
2. **“Net Sales”** bằng cách dùng SUMIFS, tổng hợp bên bảng “Transactions Data”



### 8.3. Phân tích

Sử dụng các hàm tương quan CORREL(**Net Sales**, **Total visit**) cho ra giá trị tương quan là **0.79.**



Căn cứ vào kết quả phân tích tương quan ta thấy: doanh số nhà hàng có mối tương quan dương/ cùng chiều với Total visits, với hệ số tương quan ~0.74

=> Có mối tương quan mạnh mẽ

**Đề Xuất:**

Website đặt tên miền ngắn gọn dễ nhớ, thường xuyên truy cập web trên các trình duyệt khác nhau, đồng thời gắn link web trên nhiều nền tảng mạng xã hội, báo điện tử. Điều này giúp Google thực hiện thuật toán, giúp từ khóa website của nhà hàng dễ tìm hơn, một số trường hợp kết quả xuất hiện ngay đầu tiên khi search.

## Task 9: Predictive Sales Analysis:

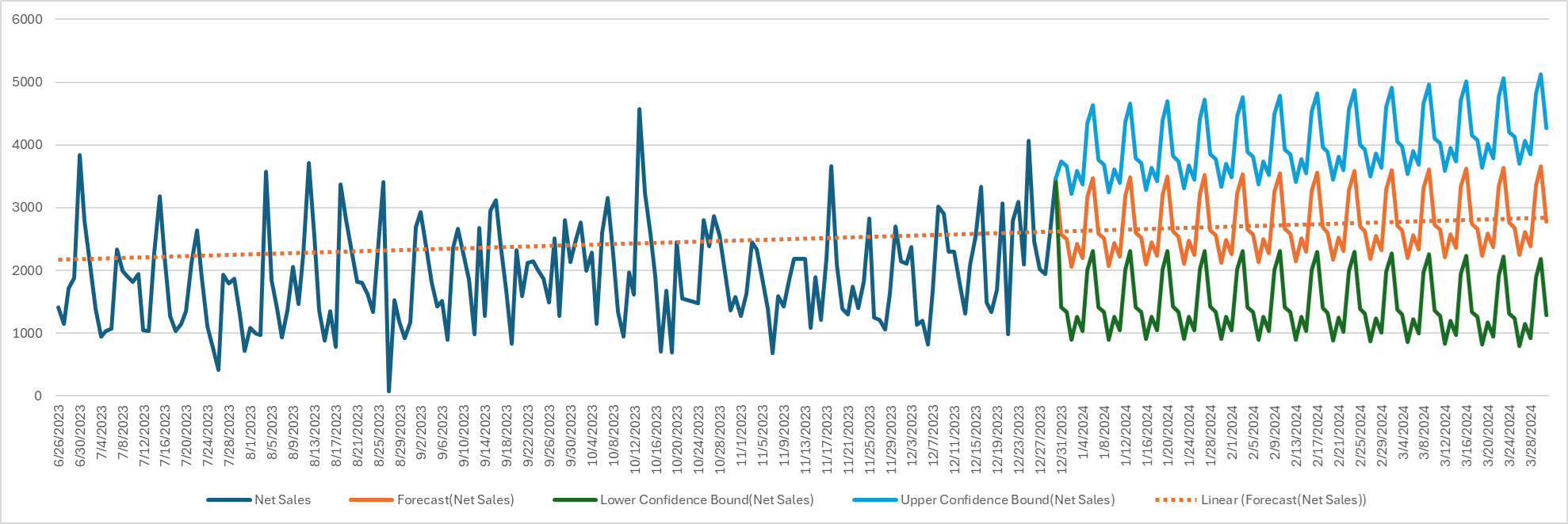
**Predictive Sales Analysis**: Employ linear regression analysis to predict our sales for the upcoming quarter, using historical sales data. This predictive insight will be invaluable for our strategic planning.

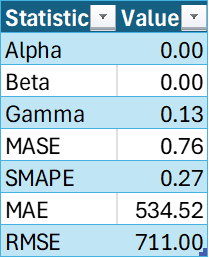
**Phân tích dự đoán doanh số bán hàng:** Sử dụng phân tích hồi quy tuyến tính để dự đoán doanh số bán hàng của chúng tôi trong quý sắp tới, sử dụng dữ liệu bán hàng lịch sử. Cái nhìn sâu sắc mang tính dự đoán này sẽ rất có giá trị cho việc hoạch định chiến lược của chúng tôi.

**Liên quan đến hồi quy:**

* Hàm FORECAST.ETS sử dụng một phiên bản nâng cao của hồi quy tuyến tính, gọi là **Hồi quy tuyến tính phi tuyến (Nonlinear Linear Regression)**, để ước tính các tham số của mô hình ETS.
* Hàm FORECAST.LINEAR sử dụng hồi quy tuyến tính đơn giản để dự đoán giá trị tương lai.

Sử dụng data làm sạch ở **Task 8**, dùng công cụ Forecast Sheet trong Tab Data để vẽ biểu diễn dữ liệu “**Net Sales**” theo dòng thời gian, kèm dự báo tương lai với độ tin cậy **95%**, ta sẽ được đồ thị như hình bên dưới, cùng 1 số chỉ số thống kê liên quan Forecast.





**Phân tích:**

* **Alpha** và **Beta** đều có giá trị là **0**, điều này có nghĩa là không có sự làm trơn theo xu hướng và theo mùa, tương ứng.
* **Gamma** có giá trị là **0.13**, cho thấy sự ảnh hưởng của yếu tố mùa.
* **MASE** có giá trị là **0.74 < 1**, cho thấy mô hình có hiệu suất tốt hơn so với dự đoán ngẫu nhiên hoặc đơn giản chỉ dựa trên trung bình của dữ liệu lịch sử.
* Với giá trị **SMAPE** là **0.27**, điều này cho biết mức độ lỗi của mô hình dự đoán là **27%.**
* Với giá trị **MAE** là **534.52**, điều này chỉ ra rằng trung bình, mô hình dự đoán sai số khoảng 534,52 đơn vị so với giá trị thực tế trên tất cả các điểm dữ liệu.
* Với giá trị **RMSE** là **711**, điều này chỉ ra rằng trung bình, mô hình dự đoán sai số khoảng **711** đơn vị so với giá trị thực tế trên tất cả các điểm dữ liệu.
* Dựa vào dự báo cho thấy, Net sales của hàng có xu hướng tăng trong 3 tháng đầu năm 2024 ( đường trendline cho thấy).