CAFÉ
Final Project

By: Kevin and Angel

Here is where your presentation begins

TABLE OF CONTENTS

Introduction

Scenario:

 You and your partner have been hired by the owner of a popular café chain to help clean and explore their messy transaction dataset.

The Business Problems are the following:
■ What are the peak sales months?

■ Are weekends busier than weekdays?

■ Any other findings that stand out and relevant to the stakeholder?

■ How reliable is the dataset? What cleaning was needed?

Data Cleaning

In order to clean this data, we had to remove NA values for the total spent. This is because the total spent is what allows us to find most of the other missing values based on the item and quantity of said item.

Not only that but every value that is classified as a NA value such as "Unknown" or "Error" was changed into an NA to make tables look less crowded than it should.

Lastly, the months and days of the week were made as columns and were added to the dataset for future use.

Total Spent	Months	Days of the Week
NA	2023-09	Friday
NA	2023-05	Tuesday
NA	2023-07	Wednesday
NA	2023-04	Thursday
NA	2023-06	Sunday
NA	2023-03	Friday
NA	2023-10	Friday
NA	2023-10	Saturday
NA	2023-07	Friday
	2023-12	Sunday
	2023-11	Tuesday

```
=TEXT(H2, "yyyy-mm")
=TEXT(H2, "dddd")
```

```python
#Remove blanks from transaction date
cafe_sales = cafe_sales[cafe_sales['Transaction Date'] != 'BLANK']
```

```python
# Group by month and sum of total spent
monthly_totals = cafe_sales.groupby('Month')['Total Spent'].sum()

# Define Months
months_order = ['January', 'February', 'March', 'April', 'May', 'June',
                'July', 'August', 'September', 'October', 'November', 'December']
monthly_totals = monthly_totals.reindex(months_order)
```

Key Findings

- Peak sales month is June, followed by October and January.

- The weekends are busier than the weekdays. Thursdays experience the most sales.

- Salads made up the majority of their sales monthly while cookies were the opposite.

Our Recommendations

- Offer weekday promotions to drive sales
  - 10% off Coffee or Tea to incentivise sales on slower weekdays.
- Leverage salad sales to improve other type of item sales
  - Bundle salads with juice or sandwich to improve those item's sales
- Create bundles such as with the purchase of a coffee get a cookie for 15% off
  - This should improve overall sales for cookies
  - Or buy 2 get 1 for free on cookies

Our Reflections

Challenges we faced:
- I would say a challenge we faced was the decision making needed to choose whether or not a chunk of data was needed to be dropped or a lot more work needed for a few.
- Initial Data Cleaning

What would we do differently next time:
- More research on cleaner ways to approach a dataset without spending too much time messing with it. For example, forward and backward fill.
-

THANK YOU