Methodology: Bright Coffee Shop Sales Analysis

1. Objective

The objective of this analysis was to uncover actionable business insights from Bright Coffee historical sales data. This included identifying peak sales periods, best-selling products categories, monthly revenue performance, and store level performance to support strategic decision making.

2. Tools utilised

-Snowflake: Used for data cleaning, transformation and SQL based aggregation.
- Microsoft excel: Used for generating pivot tables, data analysis and visualisations(charts).
- Canva: Used to prepare slides to present the key insights to the CEO
: Used for planning the layout and visuals.

3. ETL (Extract, Transform, Load) process.

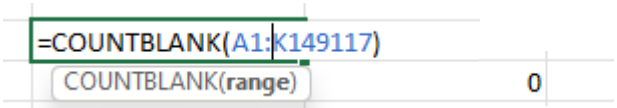
3.1 Extract process

The original dataset was provided into an excel format and was converted into a CSV file to enable smooth ingestion into snowflake for processing.

3.2 Transform

3.2.1 Checked for blank or missing values on the dataset using excel.

Based on the excel COUNTBLANK) our dataset has zero blanks.



3.2.2 Converted the transaction date into proper format in Excel before loading into snowflake.

3.2.3 Cleaned the unit _price column and fixed the error format from '3.1' to '3.1'.

The syntax that was used was:

```
UPDATE coffee_Shop_sales
SET unit_price=REPLACE(UNIT_PRICE,',','.')
```

3.2.4 Updated the table with a new calculated column ' total -amount' as unit _price* transaction quantity.

The syntax used was:

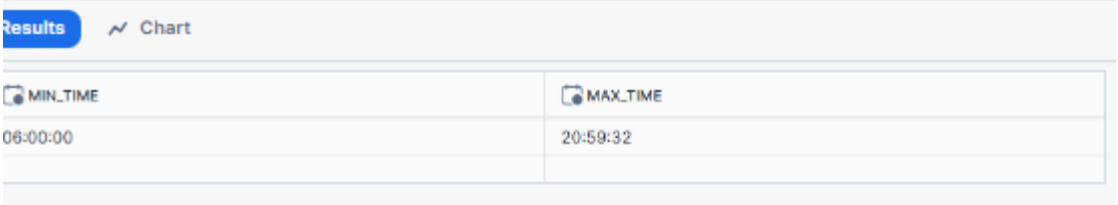
```
ALTER table coffee_shop_sales
ADD COLUMN total_amount decimal(10,2)AS (UNIT_PRICE*TRANSACTION_QTY);
```

3.2.5 Created transaction_ time buckets to group transactions into hours intervals.

In order to create the time buckets I first identified the earliest and latest time using the MIN() and MAX() time functions in Snowflake which helped define the time range for creating the time bucket.

The syntax used was:

```
SELECT min(transaction_time) AS min_time,
max(transaction_time) AS max_time
FROM coffeeshop_sales.public.coffee_shop_sales;
```



After determining the minimum and maximum time values, I proceeded to create the time buckets using CASE statement in Snowflake. The time intervals were defined as follows from 06:00 to 11:59 as 'Morning', 12:00 to 17:59 as 'Afternoon' , 18:00 to 20:59 as 'Evening'

and any other time was classified as ‘Night’. This allowed for effective categorisation of data based on the time of the day.

The syntax used was:

```
SELECT Transaction_time,
CASE
        WHEN transaction_time BETWEEN '06:00:00' AND '11:59:59' THEN 'Morning'
        WHEN transaction_time BETWEEN '12:00:00' AND '16:59:59' THEN 'Afternoon'
        WHEN transaction_time BETWEEN '17:00:00' AND '20:59:59' THEN 'Evening'
        ELSE 'Night'
END AS Time_Buckets
FROM coffeeshop_sales.public.coffee_shop_sales;
```

ults | Chart

TRANSACTION_TIME	TIME_BUCKETS
07:06:11	Morning
07:08:56	Morning
07:14:04	Morning
07:20:24	Morning
07:22:41	Morning
07:22:41	Morning
07:25:49	Morning
07:33:34	Morning
07:39:13	Morning
07:39:34	Morning
07:43:05	Morning
07:44:35	Morning
07:45:51	Morning
07:48:19	Morning

3.2.6 Date Transformation

- I renamed the Transaction_date column to Purchase date for more clarity.
- Using the TO_CHAR () function , I extracted the month name from the date.
- I also extracted the day of the week from the date using the DAYNAME()

Aggregation and Calculation

I used SQL aggregate function to summarize the data.

- Counted the transaction_id ti determine the number of transactions.
- Calculated Total revenue using the SUM(transaction_qty*unit price)
- Calculated total units sold using SUM(transaction_qty).

Grouping

- I grouped the data by relevant dimensions such as time_bucket,purchase_date, month_name,store location,product detail, product_category, day_name,product_type.

Syntax used was:

COFFEE_SHOP_SALES.PUBLIC Settings

```
SELECT COUNT (DISTINCT TRANSACTION_ID) AS NUMBER_OF_SALES,
SUM(TRANSACTION_QTY*TO_NUMBER(REPLACE(UNIT_PRICE,',','.'))) AS TOTAL_REVENUE,
SUM(TRANSACTION_QTY)AS NUMBER_OF_UNITS_SOLD,
TO_DATE(TRANSACTION_DATE) AS PURCHASE_DATE,
TO_CHAR(TO_DATE(TRANSACTION_DATE),'mon') AS MONTH_NAME,
DAYNAME(TO_DATE(TRANSACTION_DATE)) AS DAY_OF_WEEK,
CASE
  WHEN TRANSACTION_TIME BETWEEN '06:00:00' AND '11:59:59' THEN 'MORNING'
  WHEN TRANSACTION_TIME BETWEEN '12:00:00' AND '16:59:59' THEN 'AFTERNOON'
  WHEN TRANSACTION_TIME BETWEEN '17:00:00' AND '20:59:59'  THEN 'EVENING'
  ELSE'NIGHT'
  END AS TIME_BUCKETS,
STORE_LOCATION,
PRODUCT_CATEGORY,
PRODUCT_DETAIL,
PRODUCT_TYPE,
FROM COFFEE_SHOP_SALES.PUBLIC.COFFEE_SHOP_SALES
GROUP BY
PURCHASE_DATE,
STORE_LOCATION,
PRODUCT_CATEGORY,
PRODUCT_DETAIL,
PRODUCT_TYPE,
MONTH_NAME,
DAY_OF_WEEK,
TIME_BUCKETS;
```

With the output table:


```
SELECT COUNT (DISTINCT TRANSACTION_ID) AS NUMBER_OF_SALES,
SUM(TRANSACTION_QTY*TO_NUMBER(REPLACE(UNIT_PRICE,',','.'))) AS TOTAL_REVENUE,
SUM(TRANSACTION_QTY)AS NUMBER_OF_UNITS_SOLD,
TO_DATE(TRANSACTION_DATE) AS PURCHASE_DATE,
TO_CHAR(TO_DATE(TRANSACTION_DATE),'mon') AS MONTH_NAME,
DAYNAME(TO_DATE(TRANSACTION_DATE)) AS DAY_OF_WEEK,
CASE
  WHEN TRANSACTION_TIME BETWEEN '06:00:00' AND '11:59:59' THEN 'MORNING'
  WHEN TRANSACTION_TIME BETWEEN '12:00:00' AND '16:59:59' THEN 'AFTERNOON'
  WHEN TRANSACTION_TIME BETWEEN '17:00:00' AND '20:59:59'  THEN 'EVENING'
  ELSE'NIGHT'
  END AS TIME_BUCKETS,
STORE_LOCATION,
PRODUCT_CATEGORY,
PRODUCT_DETAIL,
PRODUCT_TYPE,
FROM COFFEE_SHOP_SALES.PUBLIC.COFFEE_SHOP_SALES
GROUP BY
PURCHASE_DATE,
STORE_LOCATION,
PRODUCT_CATEGORY,
```


Results Chart

NUMBER_OF_SALES	TOTAL_REVENUE	NUMBER_OF_UNITS_SOLD	PURCHASE_DATE	MONTH_NAME	DAY_OF_WEEK	TIME_BUCKETS	STORE_LOCATION	PRC
1	5	1	2023-01-01	Jan	Sun	MORNING	Hell's Kitchen	Drinkin
3	12	4	2023-01-01	Jan	Sun	MORNING	Hell's Kitchen	Tea
1	4	1	2023-01-01	Jan	Sun	MORNING	Hell's Kitchen	Drinkin

4. Data Analysis

Used Microsoft Excel to generate pivot tables and Charts such as charts, line graphs, pie charts and heatmaps.

- Identified total revenue by product category.
- Analysed peak our sales (morning, afternoon, and evening)
- Compared weekday and weekend performance.
- Assessed store level performance across different time buckets.
- Generated percentage growth per month.

5. Insights Generated

- Morning hours generate the highest sales across all the stores.
- Monday is peak sales days, with weekends (Saturday and Sunday) showing the lowest revenue.
- Coffee, tea, bakery and drinking chocolate are the top revenue drivers, with coffee being top performing product category.
- The product categories where segmented into Core drivers, Balanced performers, and Premium products.
- Monthly revenue growth observed growth observed post February, with June more than 100% increase.

6. Strategic Recommendations.

- Maximise operations during morning hours by having sufficient stocks, having in place top performing staff and early hours promotions.
- Focus marketing and efficiency efforts on weekdays, especially on Monday and Friday.
- To increase revenue generation highlight premium product categories with exclusive offers and targeted positioning.
- Realign weekend operations to reduce cost or repurpose for clearance and brand engagement.