Personal Data Storytelling Through Exploratory Data Analysis(EDA)

CIA I – Component 1

Name: Lavanya C

Register Number: 2582431

Course: 1st MSc Data Analytics

Tool Used: Excel and Power BI

Theme: Step Count, Yoga, and Menstrual Cycle Phase

SECTION 1: DATA DESCRIPTION

This dataset represents my personal daily step count tracked throughout the month of May 2025. Alongside the step count, I also recorded whether I did yoga that day, the day of the week, and which phase of the menstrual cycle I was in.

The purpose of collecting this data was to understand how my physical activity, especially the number of steps I walked, varied based on my energy levels, daily routine, and physiological changes.

The data was manually logged each day using a pedometer app and a personal calendar. In total, the dataset includes 31 days of observations with the following attributes:

- Date
- Steps Walked
- Day of the Week
- Menstrual Cycle Phase
- Mood (based on phase)
- Yoga (Yes/No)

SECTION 2: TYPES OF VARIABLES

Variable	Type	Level of Measurement	Description
Date	Qualitative	Ordinal	The column represents the chronological order of Data collection
Day of the Week	Qualitative	Nominal	This column indicates the specific day in which the Data was recorded
Steps Walked	Quantitative	Ratio	The column indicating the numerical values of the daily step count
Phase	Qualitative	Ordinal	Represents the different phases of the Menstrual cycle(Menstrual, Follicular, Ovulation and Luteal)

Mood	Qualitative	Ordinal	Inferred from general phase-based energy levels, not self-recorded
Yoga	Qualitative	Nominal	The column indicates whether yoga was done on the particular day or not

SECTION 3: DATA QUALITY

The dataset consists of 31 daily records for the month May 2025. Overall, the data is clean and reliable with the following observations.

- **Missing Values-** There is only one missing value in the "Phase" and "Mood" columns for May 19th. This day was labelled as "NA" using Power BI since it did not clearly fall under any specific menstrual phase. No other missing values were found**.**
- **Outliers-** One record (May 11th) has a very low step count of 693, which is significantly lower than the rest.
- **Duplicates-** There is absence of duplicates.
- **Data Formatting-** An extra space in the Day of the week-Friday, was replaced using MS Excel's, Find and Replace option to maintain clean categories.

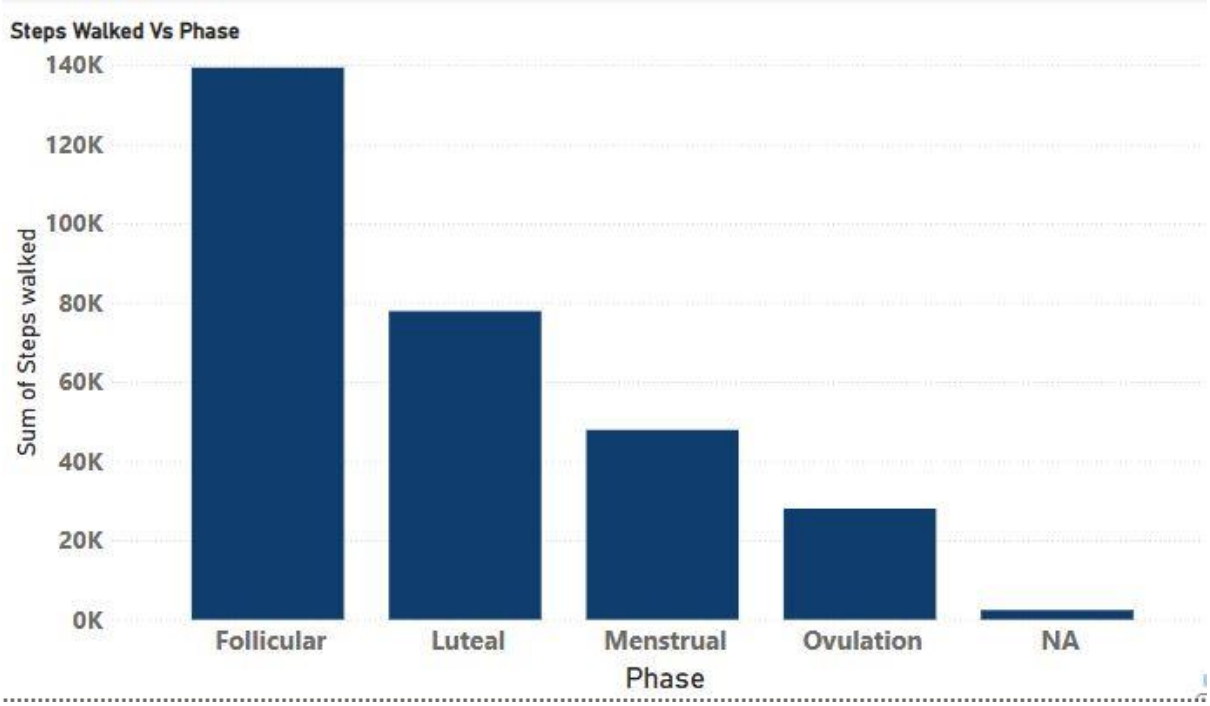
Overall the data provided for EDA with minimum data quality issues which have been addressed using pre-processing in MS Excel and Power BI.

SECTION 4: SUMMARY STATISTICS

To understand the overall distribution of my physical activity, I used Power BI cards to display key statistics. These indicators give a quick overview of my step count trends and the Yoga done across the month of May.

SECTION 5: DATA VISUALIZATIONS

1) **Average Steps By Phase**- This bar chart shows the average number of steps walked during each menstrual cycle phase. Step count was highest in the follicular phase and lowest in the ovulation phase, ignoring the NA value.
   Chart type: Bar Chart
   Chart Title: Steps Walked Vs Phase



Observations

- The average number of steps is highest during the follicular phase.
- The lowest average step count is during the ovulation phase.
- The luteal phase and menstrual phase have moderate average step counts.
- One entry (May 19) is marked as NA and excluded from this comparison.

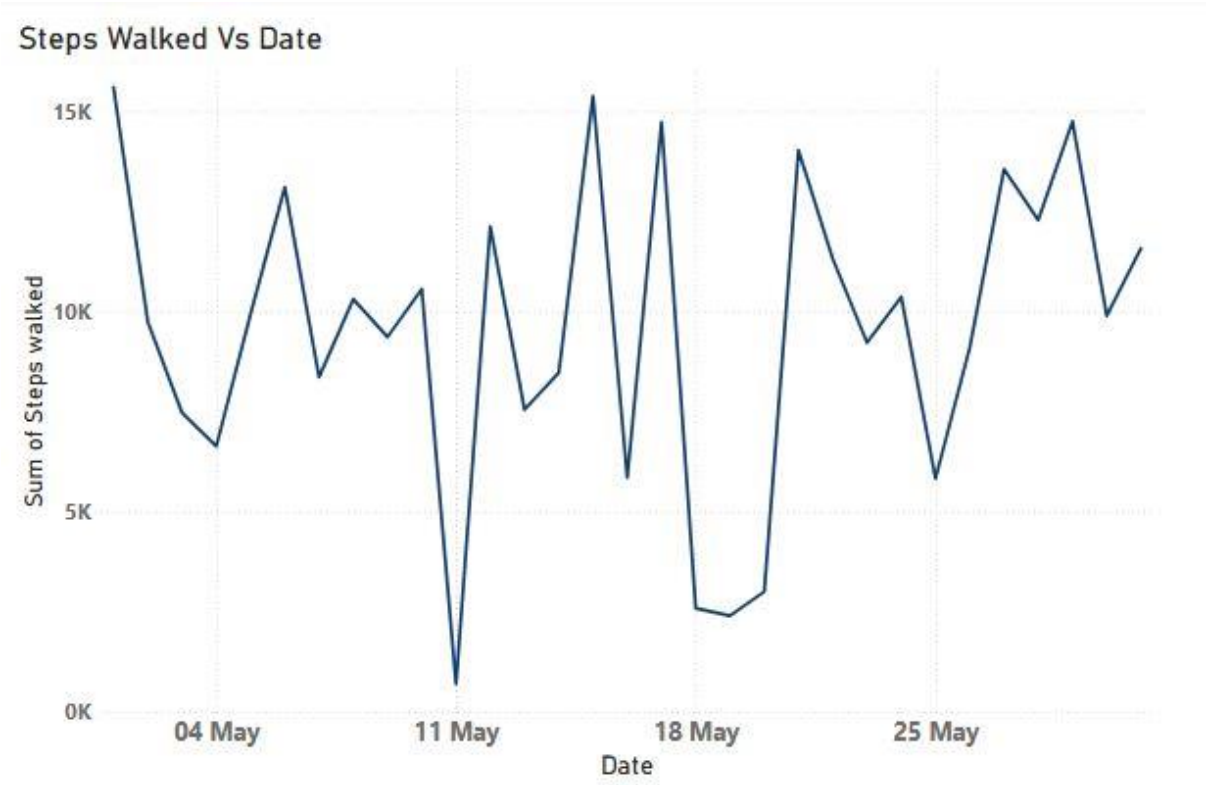
Insights

- The follicular phase shows the highest average step count, indicating that energy and physical activity levels are at their peak during this phase. This aligns with scientific understanding that estrogen levels rise, leading to increased motivation and stamina.
- The lowest steps during ovulation may suggest fatigue, discomfort, or schedule conflicts, showing that physical activity is not always predictable based on biology alone. This reveals the importance of planning high-movement days during the follicular phase and being flexible during ovulation.

2) **Daily Step Trend Over the Month-**This line chart illustrates the fluctuation in daily step count throughout May 2025. While physical activity peaked during the early follicular phase, noticeable dips were observed during the ovulatory phase, reflecting deviations from typical energy expectations**.**

Chart type: Line Chart

Chart title: Steps Walked Vs Date



Observations-

- Step counts fluctuate significantly throughout the month.
- High step counts (15,000+) are observed on May 1, 14, 17, 21, and 28.
- Very low step counts are observed on May 11 (693 steps) and May 19 (~2,000 steps).
- There is no consistent increasing or decreasing trend

Insights

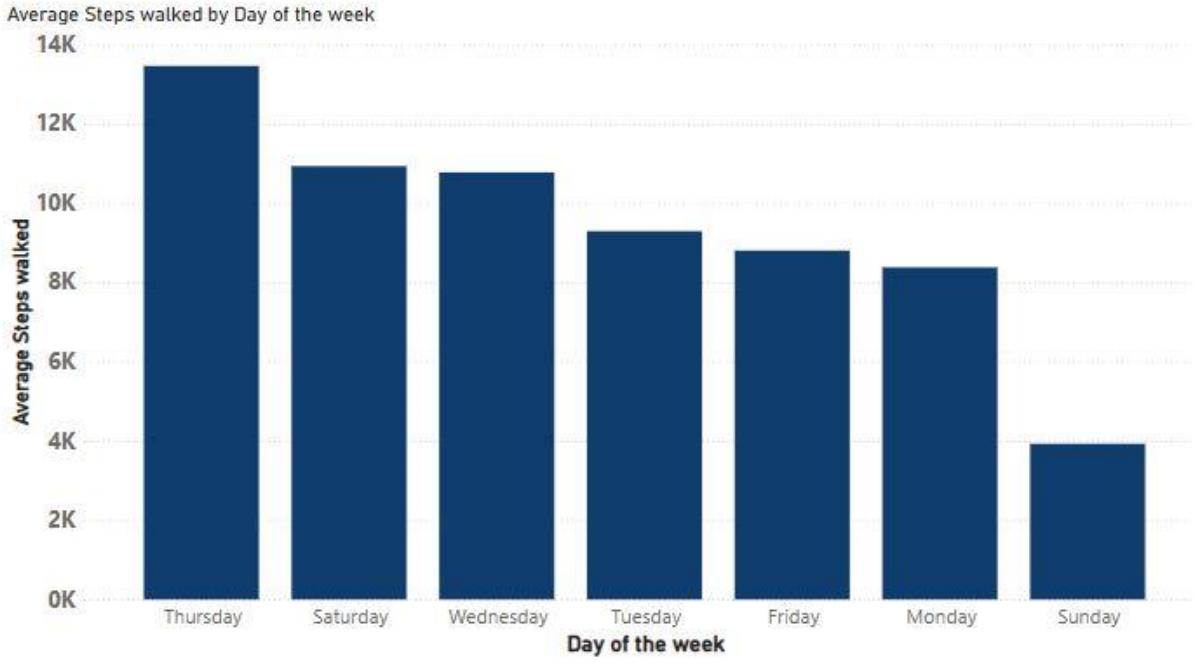
There is a highly fluctuating pattern in step count across the month. Peaks on days like May 1, 14, 17, 21, and 28 suggest intentional effort or active schedules, while dips on May 11 and 19 show days of very low movement, possibly due to rest, mood changes, or lack of motivation. This inconsistency highlights a need for a more balanced

and sustainable routine rather than relying on bursts of activity.

3) Total Steps by Day of the Week

Chart Type: Bar Chart

Chart Title: Average Steps walked by Day of the Week



Observations

- Thursday has the highest average step count (~13,500 steps).
- Saturday and Wednesday also have relatively high step averages.
- Sunday has the lowest average step count (~4,000 steps).

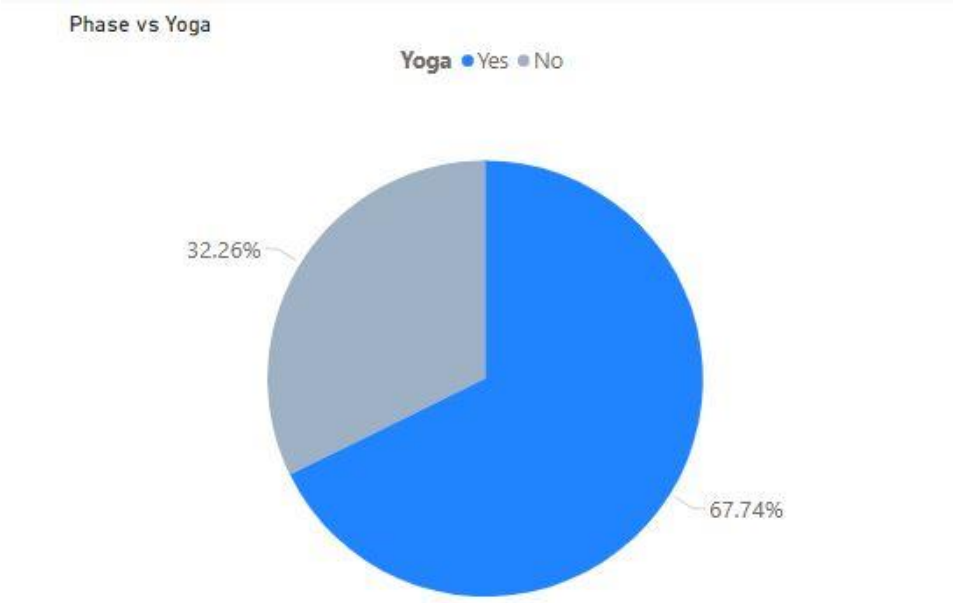
Insights

The highest average steps occur on Thursdays, possibly due me settling in the weekly hustle. Saturday and Wednesday also show high activity, which could be linked to errands or weekly tasks. Sunday has the lowest average steps, because it as a rest day. This pattern confirms that external schedule structures (like classes) significantly influence my physical movement**.**

4) Yoga Participitaion Distribution

Chart Type: Pie Chart

Chart Title: Phase Vs Yoga



Observations

- Yoga was done on approximately 67.74% of the days.
- Yoga was not done on 32.26% of the days.
- The pie chart shows a majority of “Yes” days compared to “No” days.

Insights

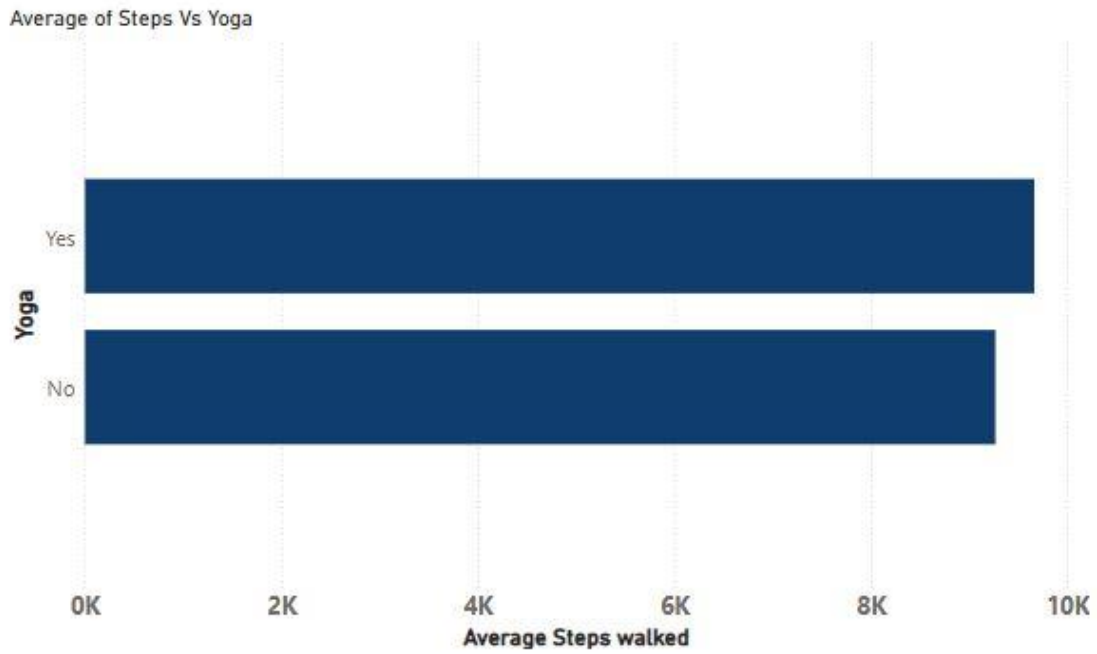
- Yoga was practiced on 67.74% of the days, showing a strong commitment to physical and mental wellness.

- However, 32.26% of the days without yoga point to room for improvement in consistency.
- The data reveals that yoga is already a significant part of my lifestyle, but making it a daily habit could lead to better physical and emotional balance across all phases of the cycle.

5) Average Steps Walked – Yoga vs No Yoga Days

Chart Type**:** Horizontal Bar Chart

Title: Average Steps vs Yoga



Observations

- The average number of steps is higher on yoga days compared to non-yoga days.
- There is a noticeable difference in step count between the two categories.
- The bar for "Yes" (Yoga done) is longer than the "No" (Yoga not done).

Insights

- Days on which yoga was performed had noticeably higher average steps, suggesting that doing yoga may boost energy, discipline, and overall motivation to move more throughout the day.
- This correlation shows that yoga positively influences your broader physical activity levels, and incorporating it consistently could be key to sustaining an active lifestyle.

SECTION 7: REFLECTION.

This Exploratory Data Analysis has helped me draw insights from my daily physical activity and habits like:

- I tend to be more physically activity during the follicular phase of my menstrual cycle.
- Yoga positively influences my overall step count, showing that starting my day with yoga encourages more movement throughout the day.
- Sunday tends to be a rest day, but it would be better if I do light physical movements that would not lead it to complete inactivity.

Analysis from this report has made me want to plan my days making it more efficient and consistent by:

- Making yoga a consistent daily habit
- Planning my most active tasks during high energy phases like follicular phase
- Maintain minimum activity on holidays to prevent complete inactivity.

APPENDIX

Data Source

The dataset used in this report was self-collected over the month of May 2025.

- Step count and menstrual cycle phases was recorded daily using a Google Fit on my phone.
- Yoga activity was manually logged in my personal journal
- Data was organized and cleaned using MS Excel and Power BI.
- The data was recorded in an Excel file and I have copied it to this file

Date	Steps walked	Day of the week	Phase	Mood	Yoga
01-05-2025	15590	Thursday	Follicular	Energy increasing	Yes
02-05-2025	9733	Friday	Follicular	Energy increasing	No
03-05-2025	7463	Saturday	Follicular	Energy increasing	Yes
04-05-2025	6630	Sunday	Follicular	Energy increasing	Yes
05-05-2025	9931	Monday	Follicular	Energy increasing	No
06-05-2025	13099	Tuesday	Follicular	Energy increasing	Yes
07-05-2025	8351	Wednesday	Ovulation	Highest energy	Yes
08-05-2025	10305	Thursday	Ovulation	Highest energy	No
09-05-2025	9359	Friday	Ovulation	Highest energy	Yes
10-05-2025	10553	Saturday	Luteal	Tired/moody	Yes
11-05-2025	693	Sunday	Luteal	Tired/moody	Yes
12-05-2025	12113	Monday	Luteal	Tired/moody	Yes
13-05-2025	7540	Tuesday	Luteal	Tired/moody	No

14-05-2025	8464	Wednesday	Luteal	Tired/moody	Yes
15-05-2025	15379	Thursday	Luteal	Tired/moody	No
16-05-2025	5841	Friday	Luteal	Tired/moody	No
17-05-2025	14730	Saturday	Luteal	Tired/moody	Yes
18-05-2025	2577	Sunday	Luteal	Tired/moody	Yes
19-05-2025	2393	Monday	NA	NA	Yes
20-05-2025	2995	Tuesday	Menstrual	Low energy	No
21-05-2025	14023	Wednesday	Menstrual	Low energy	Yes
22-05-2025	11313	Thursday	Menstrual	Low energy	Yes
23-05-2025	9217	Friday	Menstrual	Low energy	Yes
24-05-2025	10360	Saturday	Menstrual	Low energy	Yes
25-05-2025	5820	Sunday	Follicular	Energy increasing	No
26-05-2025	9091	Monday	Follicular	Energy increasing	Yes
27-05-2025	13549	Tuesday	Follicular	Energy increasing	No
28-05-2025	12279	Wednesday	Follicular	Energy increasing	Yes
29-05-2025	14745	Thursday	Follicular	Energy increasing	Yes
30-05-2025	9884	Friday	Follicular	Energy increasing	Yes
31-05-2025	11575	Saturday	Follicular	Energy increasing	No