Project Title: Instagram Performance Analytics for Data Analyst Duo

Problem Statement:

As part of the Data Analyst Duo team, you are tasked with analysing Instagram performance data to extract key insights on content engagement, follower growth, and posts performance. The goal is to provide actionable recommendations to improve reach, engagement, and content strategy.

Data Description:

Topic	Description
Reach	The number of unique accounts that saw your post or story. Measures audience size.
Profile Visits	How many users visited your profile after viewing the post. Indicates interest in the page.
New Followers	The number of new people who followed your account on the day the content was posted.
Views	Total number of times the video content was viewed. May include multiple views by the same user.
Interaction	Total engagement actions (likes, comments, shares, saves) on a post.
Description	The caption or text content you added with the post. Can contain hashtags and tags.
Duration (secs)	Length of the video post in seconds. Relevant for Reels or video posts.
Publish time	The date when the post was published. Helps with trend analysis over time.
Post type	The type of post — e.g., Image, Video, Reel, Carousel.
Impressions	Total number of times the post was seen (includes repeat views).
Shares	Number of times users shared the post via DMs or stories.
Follows	Number of follows attributed to this specific post. Slightly different from overall "New Followers".
3s views	Number of users who watched the video for at least 3 seconds (a key metric for Reel engagement).
Likes	Number of likes on the post. A basic but strong indicator of engagement.
Comments	Number of user comments on the post. Helps track conversation and sentiment.
Saves	Number of users who saved the post to their collections. Indicates strong content value.

Objectives:

- Analyze trends in content performance over time
- Identify top-performing posts and their characteristics
- Determine optimal days and times for posting
- Assess follower growth patterns and engagement strategies
- Recommend strategies for improving reach and engagement

Skills Required:

- Excel formulas (SUMIFS, AVERAGEIFS, IF, VLOOKUP/XLOOKUP, TEXT, DATE functions)
- Pivot tables for summarization
- Charts: Line, Column, Bar, Pie, Combo Charts
- Insight storytelling in Excel
- Clean formatting and professional visualization

Analysis Plan:

Goal 1: Instagram Daily Performance Analytics

Primary Objective: Create a clean daily dataset using VLOOKUP by joining multiple metrics and uncover trends, correlations, and conversion insights to support growth strategy.

Dataset Challenges & Instructions

1. **Date Join Using VLOOKUP:**
    - Start by creating a **master date column** from 9th Jan 2022 to 2nd May 2025.
    - Use VLOOKUP or XLOOKUP to pull Reach, Profile Visits, Followers, Views, and Interaction from their respective sheets.
    - Missing values (due to different start dates) should be left **blank**, not 0. Use IFERROR logic smartly.
2. **Final Output:** A single sheet called Daily_Stats with:
    - Date, Reach, Profile Visits, New Followers, Views, Interaction

Analysis Sections & Guidelines

1. Daily Trend Analysis

Chart Type: Line or Area Charts

- Plot daily Reach, Profile Visits, New Followers, Views, and Interaction over time
- Observe peak days, drop periods, and long-term momentum
- Use chart slicers or filters for zooming into specific time periods

2. Day of Week Analysis

Steps:

- Add a new column: *=TEXT(Date, "dddd")* to extract weekday
- Create a pivot table to calculate **average reach, views, followers by day**
  Goal: Identify which days perform best (e.g., “Saturdays drive highest new followers”)
  Chart: Column chart with days on X-axis

3. Weekend vs Weekday Comparison

Steps:

- Create a new column: *=IF(WEEKDAY(Date, 2)>5, "Weekend", "Weekday")*
- Use pivot table to compare average metrics
  Insights Expected:
    - Do weekends generate more views?
    - Is engagement rate higher during weekdays? *(Interaction ÷ Reach)*

4. Correlation Analysis (Scatterplot + Logic)

Pairs to explore:

- Reach vs Profile Visits
- Profile Visits vs New Followers
- Views vs Interaction
- Reach vs New Followers
  Chart Type: Scatter Plot
  Expected Outcome:
  Visually interpret relationships, not just guess. Add trendlines where possible.

5. Follower Conversion Rate

Metric: *New Followers ÷ Reach*

- Create a new column to calculate daily conversion rate
- Plot this as a line chart
- Highlight dates with highest conversion — what happened on those days?

6. Growth Rate in Followers

Metric: *= (Today's Followers – Yesterday's Followers) ÷ Yesterday's Followers*
Chart: Line + Highlight spikes

7. Rolling Averages

- Create new column *=AVERAGE(OFFSET(...)) or =AVERAGE(range)* for **7-day moving average**
- Smoother trends help reduce noise. Especially useful for Views & Followers

8. Content Effectiveness (Using Description)

- Flag top 5 days by: Highest Reach, Highest New Followers, Highest Engagement *(Interaction ÷ Reach)*
- Look up content posted on those days — any patterns? Reels? Hooks? Carousels?

Expected Learning Outcomes

- Join datasets using lookup functions with inconsistent date ranges
- Handle missing data (blank vs 0) thoughtfully
- Build interactive dashboards using pivot tables and slicers
- Analyze and interpret Excel visuals for business impact
- Derive data-driven content strategy recommendations

Goal 2: Post-Level Content Performance Analysis

Metrics to Compute:

1. Duration Buckets

Create a new column Duration Bucket with:

- <10 secs
- 10–30 secs
- 31–45 secs
- >45 secs

Purpose: Analyze average engagement/conversion per duration group using a pivot table.

2. Retention Rate

Formula: *Retention Rate = 3s Views ÷ Impressions*

*Why **Impressions**? Because every view is an impression, but not every impression results in 3s watch. This tells us how many people "stuck around."*

3. Follower Conversion Rate

Formula: *Follows ÷ Reach*

This tells you **how persuasive the content is** in converting casual viewers into followers. Use scatterplots vs reach, impressions, etc.

4. Engagement Rate

Option 1 (Simple Average):

Engagement Rate = AVERAGE (Likes ÷ Reach, Shares ÷ Reach, Saves ÷ Reach, Comments ÷ Reach)

Option 2 (Total Interactions ÷ Reach):

Engagement Rate = (Likes + Shares + Saves + Comments) ÷ Reach

5. Post Timing Analysis

- Extract **Day of Week** from Publish time
- Find patterns: Which days drive more reach, engagement, or follows?

6. Post Type Analysis

- Use pivot tables to compare: Avg Reach, Followers, Likes by Post Type (e.g., IG Reels, Carousel, Static)
- Combine with Duration Buckets to find **best format-duration pair**

7. Content Hook Effectiveness (Qualitative + Quantitative)

- Pick top 5 posts by conversion rate or engagement
- Read the Description for pattern (e.g., use of #, hooks like "how to", emojis)
- Was it carousel, storytelling, tutorial, motivational?

8. Engagement Composition

- For each post, calculate share of total engagement:
    - *% Likes, % Shares, % Comments, % Saves*
    - *Total Engagement = Likes + Shares + Comments + Saves*
    - *% Likes = (Likes ÷ Total Engagement), etc.*
- Helps identify content type:
    - Informative = high saves
    - Viral = high shares
    - Community = high comments

9. Virality Score

Custom metric: *Virality Score = (Shares + Saves) ÷ Reach*

Indicates likelihood of the content being reshared or saved — signs of high-perceived value.

10. Outlier Detection using Box-Whisker plot

Identify which posts had unusually high or low Reach, and then analyze why those posts stood out — what kind of content, timing, or structure they had.

For each point beyond the whiskers:

- Look up the corresponding post using INDEX or manual match
- Create a table or annotation like this:

Post	Reach	Why It May Be an Outlier
"Top 5 SQL Tips"	48,000	Posted on weekend, short & relatable, carousel
"Funny Reel on Data"	52,000	Trending audio, lots of shares

Ask yourself:

- What was the **hook** in these posts?
- Was the post type different? IG Reel vs Carousel?
- What was the **posting time/day**?
- Did they use a trending sound or format?

Expected Learning Outcomes

- Understand and interpret key post-level Instagram metrics (Reach, Engagement, Follows, etc.)
- Calculate derived KPIs like Engagement Rate, Retention Rate, and Follower Conversion Rate
- Use pivot tables and charts to compare post types, durations, and performance patterns
- Identify outlier posts using Box & Whisker plots and explain why they stood out
- Derive content strategy insights backed by data and visual storytelling

Goal 3: Audience Demographics Analysis

Analysis Plan & Tasks

1. Country-wise Audience Distribution

- Use a **Bar Chart** or **Map Chart** (if Excel version supports it) to show followers by country
- Sort descending to show top countries clearly
- Highlight the % share of India vs Rest of the World
- Calculate *% Share = Followers from Country ÷ Total Followers*

2. City-level Breakdown (Zoom into India)

- Use **Horizontal Bar Chart** for Top Cities
- Group **Other cities** into a single bar (already done in your data)
- Highlight insights like: "Bangalore alone contributes to X% of our total Indian audience"
- Optional: Create a **Treemap** to show relative city sizes

3. Age & Gender Analysis

- Split the dataset into Male and Female using Excel filters
- Create:
  - **Stacked Bar Chart** or **Clustered Bar Chart**: Age on X-axis, follower count on Y-axis, color-coded by gender
  - **Pie Chart**: Total Male vs Female audience

4. Create Summary KPIs

Create a small metrics block using formulas to calculate:

Metric	Formula
% Female Audience	Total Female ÷ Total Followers
% 25–34 Age Group	Total 25–34 ÷ Total Followers
Top Age Group	Use =MAX() across all age groups
Top Country	Use =INDEX + MATCH or just filter

5. Audience Profile Summary Slide

Summarize key findings in a one-pager:

- Most of our audience is...
- Gender skew is...
- We're popular in these cities...
- Our 25–34 Male audience dominates x%...

Expected Learning Outcomes

- Build audience personas using Excel analysis
- Practice charting techniques like bar, pie, treemap, and stacked charts

- Use basic Excel functions and pivot tables for demographic segmentation
- Derive business insights like “where to run ads” or “what content to create more of”
- Build professional slides summarizing target audience profiles

Deliverables:

As part of this project, you will work in **groups** to create a comprehensive **Project Report** and a compelling **Presentation Deck** based on the Instagram performance dataset provided. This format simulates how data analysts work in real-world teams — conducting individual analyses and coming together to build impactful business narratives.

Step 1: Individual Analysis (Solo Work)

Each student is expected to:

- Complete **all assigned analyses** independently
- Document insights and observations
- Be ready to contribute to the group with their findings

Step 2: Group Collaboration

You will be assigned to a group. As a team:

- Discuss individual insights and consolidate them
- Divide tasks: Who writes the report, who designs the slides, who presents what
- Collaborate to build **two main deliverables**:

Deliverable 1: Project Report (Detailed, In-Depth)

A document that captures the entire analysis journey. Use the following format:

1. **Problem Statement** – Briefly explain what you are trying to solve or uncover
2. **Objective** – Define the business goals (e.g., improve engagement, understand audience, optimize content)
3. **Data Cleaning Summary** – Talk about how you handled missing data, blank dates, merging, formatting issues, etc.
4. **Analysis & Insights** – Deep dive into all three goals (daily performance, post-level metrics, audience) with charts, numbers, and interpretations
5. **Recommendations** – What should be done next? Suggest content strategies, audience targeting ideas, or post optimizations

Deliverable 2: PPT Presentation (Concise & Visual)

This will be presented live in class. Focus on:

- **Visual storytelling** – Less text, more impactful charts and visuals
- Use 1–2 slides per section (max 10 slides total)
- Each team member **must present** a portion of the deck
- Use the **same structure** as the report:

Tips for a Great Report & Presentation

For the Report:

- Include charts with captions
- Explain calculations briefly
- Keep tone professional and analytical
- Highlight unique findings from each member's work

For the PPT Deck:

- Use large, clean visuals — pivot charts, bar graphs, trend lines
- Avoid clutter – 3–4 key points per slide
- Stick to a consistent color scheme and font
- Tell a story: "What did we find?" → "Why does it matter?" → "What should we do?"

Presentation Expectations:

- Each student must **explain at least one insight or slide**
- Be ready for follow-up questions from the instructor or peers
- Treat it like you're presenting to a business team!

Submission Timeline:

Project Launch – 4th May 2025

Final Report + Deck Submission – 15th May 2025

Presentations – 17th May 2025

Note:

As we begin our first project, I wanted to clarify how presentations will be managed across the bootcamp.

Since we have multiple teams and limited time in each session, **not every team will present during every project**. Instead, for each project, **a few teams will be selected to present their PPT live in class**. This way, we can ensure everyone gets a fair chance to present at least once across the entire bootcamp.

That said, **every team must submit both the Project Report and the PPT Deck** for every project — regardless of whether they are presenting or not. The teams chosen to present will be informed **2–3 days in advance**, so please be prepared.

This approach helps us balance learning with time management, especially considering your work schedules. Looking forward to seeing your creativity, insights, and storytelling skills in the submissions!