Retain IQ

BY Raghav Vijayvergia
BTECH From IIT Jodhpur

Q1 Solution Path Overview (which i have followed)-
MOST IMPORTANT FOR EVALUATION

1. **Data Preparation & Metric Calculation**
   - Loaded and explored campaign data.
   - Computed CTOR (Click-to-Open Rate) as `Click Rate / Open Rate` to assess email body effectiveness.
   - Converted Open Rate (OR) from percentage to decimal by dividing by 100 to ensure consistency in calculations.
2. **Day-of-Week Performance Analysis**
   - Identified which days of the week had the highest concentration of top 25th percentile performers for both CTOR and OR.
   - Using heatmaps and scatter plots, determined that Monday, Wednesday, and Friday consistently yielded superior performance.
3. **Best Time-of-Day Determination**
   - Analyzed CTOR and OR distributions by hour using heatmaps.
   - Found that the optimal send window is between 2 PM and 5 PM, with 3 PM (15:00) as the peak.
4. **Campaign Content Effectiveness Comparison**
   - Compared top 25% vs bottom 25% performing campaigns.
   - Evaluated differences in subject line tone, structure, urgency, personalization, and clarity.
   - Assessed email body patterns, including layout, CTA visibility, clarity of offers, and reinforcement techniques.
5. **Top vs Bottom 5% Analysis (OR & CTOR)**
   - Studied subject lines and email bodies from the top and bottom 5% performing campaigns.

- ○ Extracted specific language, formatting, and engagement tactics that distinguish high-performing emails from low-performing ones.

6. **Campaign Date Selection for January 2025**
   - ○ Selected all Mondays, Wednesdays, and Fridays in January.
   - ○ Supplemented with specific historical high-performing dates from previous Januarys, guided by top-performing trends in OR and CTOR.
7. **Final Calendar Development**
   - ○ Using all insights from content analysis, timing, and audience behavior, built a data-backed January 2025 campaign calendar.
   - ○ Ensured each scheduled campaign aligned with strategic themes, high-performing subject line patterns, and optimal send times.
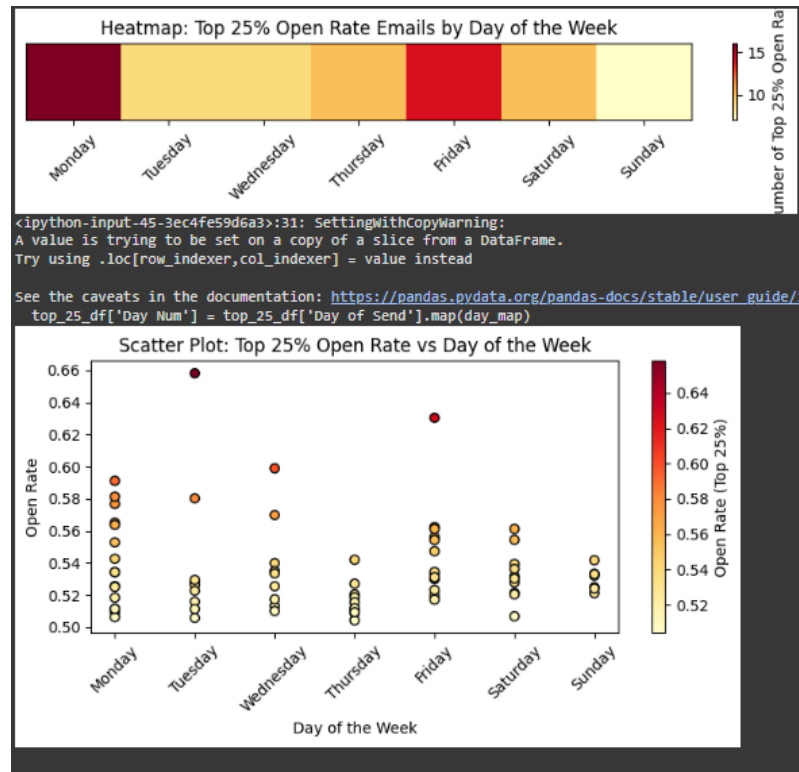
Collab Link - 🔗 RetainIQ1.ipynb Chatgpt link - Chatgpt

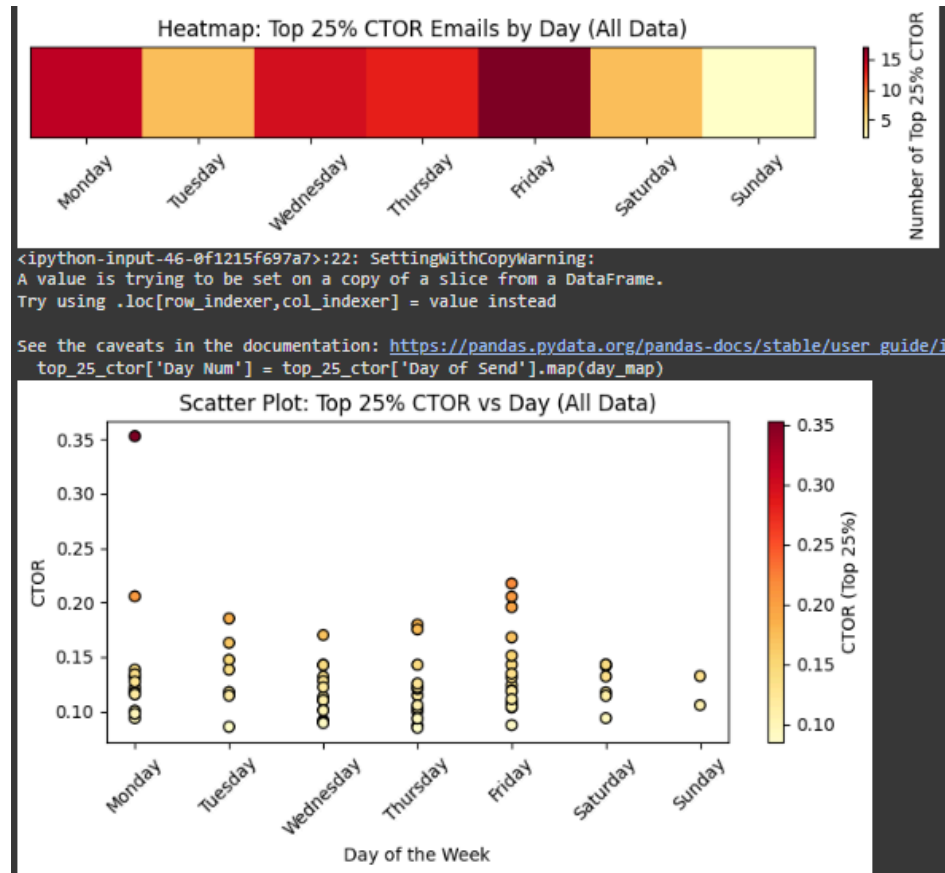| Metric	Mean	Median	25th Percentile	75th Percentile
Open Rate	0.4629	0.4483	0.4256	0.5043
CTOR	0.0661	0.0503	0.0354	0.0851
Conversion Rate	0.1164	0.0967	0.0655	0.1558
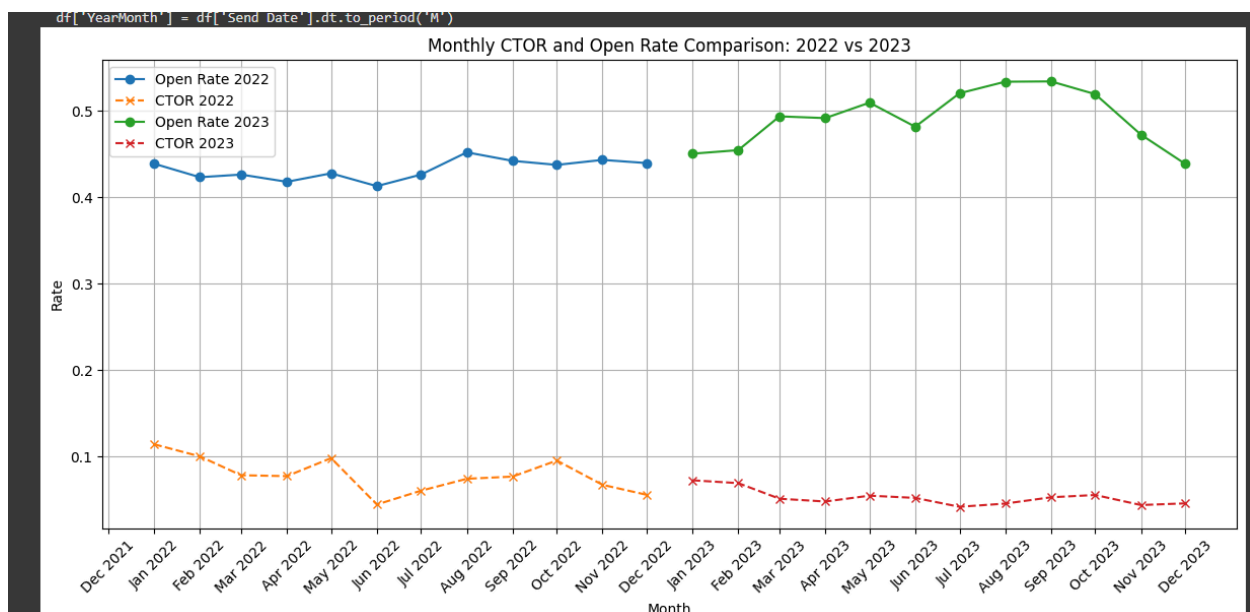
Key Interpretations:

- Open Rate (OR) → Reflects how compelling the subject line is.
- Click-To-Open Rate (CTOR) → Reflects how engaging the email body content is.
- Conversion Rate → Indicates how effective the landing page or offer is after the click.
- Converted Open Rate (OR) from percentage to decimal by dividing by 100 to ensure consistency in calculations.
- No time series relation as such.


```
<ipython-input-45-3ec4fe59d6a3>:31: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/i
  top_25_df['Day Num'] = top_25_df['Day of Send'].map(day_map)
```

Heatmap: Top 25% CTOR Emails by Day (All Data)

```
<ipython-input-46-0f1215f697a7>:22: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/i
    top_25_ctor['Day Num'] = top_25_ctor['Day of Send'].map(day_map)
```

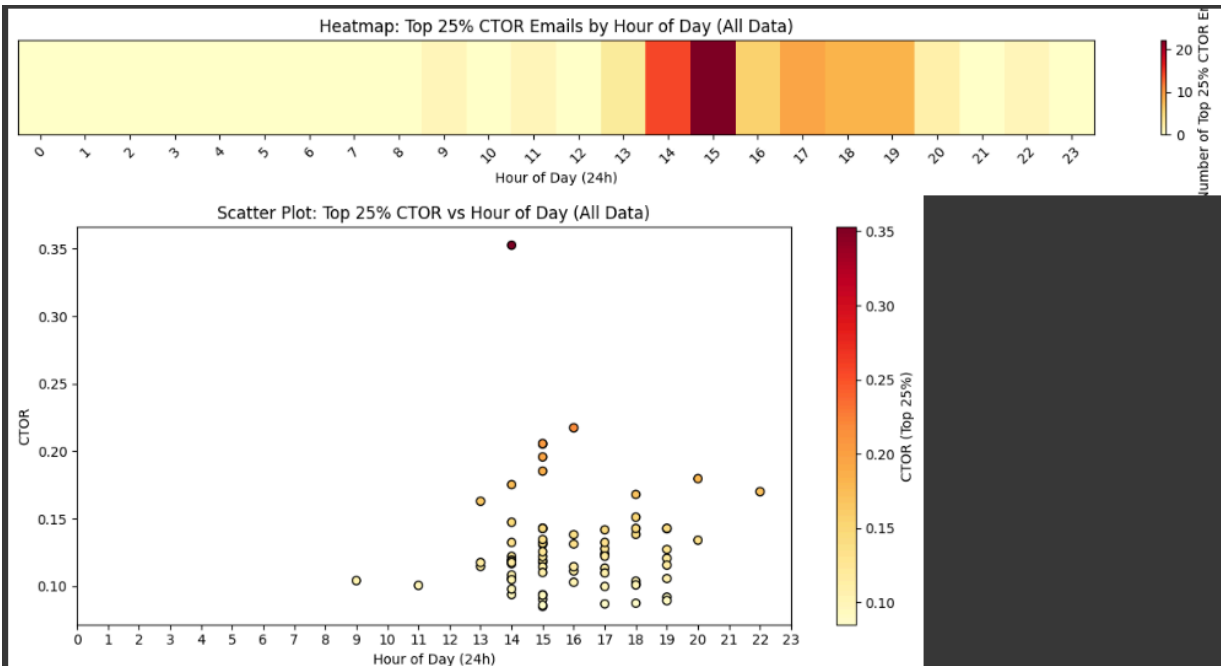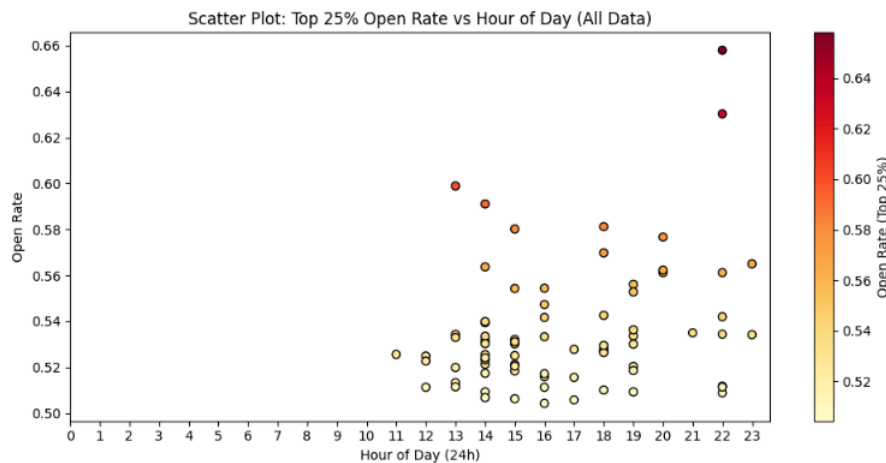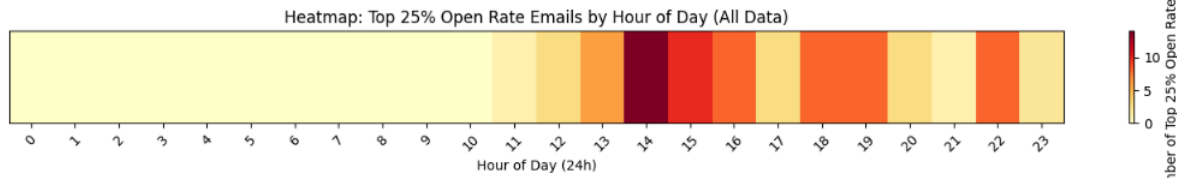Scatter Plot: Top 25% CTOR vs Day (All Data)

rom the Open Rate (Top 25%) Heatmap & Scatter Plot:

- **Monday**: Dominates both in count and intensity of high open rates.
- **Friday**: Also shows a strong presence of high open rates.
- **Wednesday**: Appears more frequently in the top 25% than Thursday, Saturday, or Sunday.

◆ **From the CTOR (Top 25%) Heatmap & Scatter Plot:**

- **Monday**: Again, shows highest presence in the top 25% — excellent email body engagement.
- **Friday**: Strong CTOR frequency.
- **Wednesday**: Appears moderately strong, slightly above Tuesday or Thursday in this top segment.

Heatmap: Top 25% Open Rate Emails by Hour of Day (All Data)

Scatter Plot: Top 25% Open Rate vs Hour of Day (All Data)

Heatmap: Top 25% CTOR Emails by Hour of Day (All Data)

Scatter Plot: Top 25% CTOR vs Hour of Day (All Data)

✅ Final Recommendation: Best Time to Send Emails

Rank	Hour (24h)	Justification
🥇	15:00	Peak performance in CTOR + strong presence in Open Rate (ideal blend)
🥈	14:00	Highest Open Rate concentration, also performs well in CTOR
🥉	16:00	Good engagement in both metrics, slightly behind 2–3 PM

Using chatgpt compare these two sheets

Top Performing Data – sheet 1.csv (📗 Top performing data)

- **Table 1**: Top 25% best-performing subject lines and email bodies (paired).
- **Table 2**: Top 5% best-performing subject lines (only subject lines).
- **Table 3**: Top 5% best-performing email bodies (only email bodies).

Worse Performing Data – Sheet1.csv (📗 worse performing data)

- **Table 1**: Bottom 25% worst-performing subject lines and email bodies (paired).
- **Table 2**: Bottom 5% worst-performing subject lines (only subject lines).
- **Table 3**: Bottom 5% worst-performing email bodies (only email bodies).

ANALYSIS: Key Differences Between High vs Low Performing Emails

1. Subject Line Characteristics

Feature	Top Performers	Bottom Performers
Tone	Positive, energetic, benefit-driven	Generic or vague
Urgency/Excitement	Uses urgency effectively (e.g., *"On NOW!"*, *"Early Access"*)	Overused or cliché urgency (*"Last Chance"*, *"Extended"*)
Discount Highlight	Clearly states % off and product context (e.g., "40% off puzzles")	Mentions discount without specificity (e.g., "Puzzle savings")
Personalization	Occasionally uses personal language (e.g., "Just for You")	Minimal personalization
Symbols/Emojis	Used sparingly and contextually (e.g., ♟, ❗)	Also used, but often feels forced or gimmicky

2. Email Body Patterns

Feature	Top Performers	Bottom Performers
Clarity & Focus	Clear offer upfront, concise messaging	Cluttered or overly generic intro content
Visual Layout	Better formatting, structured, prominent CTAs	Dense content or too much whitespace padding
CTA (Call-To-Action)	Highlighted and actionable (e.g., "Shop Now")	Less visually distinct or buried
Offer Reinforcement	Repeats key deal (e.g., "40% OFF Clearance")	Buries or doesn't repeat main value prop
Trust Elements	Includes support numbers, reassurance copy	May lack trust signals or helpful info

Subject Line Improvements

1. **Be Specific**: Mention the exact benefit (e.g., "40% Off Puzzles" beats "Puzzle Savings").
2. **Create Urgency — But Authentically**: Use terms like *"Early Access"*, *"Today Only"*, but avoid overused phrases like "Last Chance" unless it's truly final.
3. **Emojis = Spice, Not Sauce**: Use sparingly and only when aligned with message tone.

Email Body Optimization

1. **Front-load the Offer**: Lead with the deal and a short sentence on why it matters.
2. **Visual Hierarchy**: Make CTA buttons and key offers prominent.
3. **Trim the Fat**: Avoid long intros, repeated headers, or padded formatting that adds no value.

4. **Consistency**: Reinforce the subject line offer inside the body—ensure what was promised is the focal point.

Top 5% Performing Subject Line vs Bottom 5 percentile

Top 5% Performing Subject Line:

"9 Shop by Pieces 11.04"

📈 **Why it may have performed well:**

- **Action-Oriented**: Begins with a verb — *"Shop"* — prompting engagement.

- **Personal Relevance**: "Shop by Pieces" implies customization or control, which is attractive.

- **Time-Sensitive**: Includes a date (*11.04*), which may imply urgency or recency.

- **Specificity**: Suggests a clear theme or feature of the campaign.

Bottom 5% Performing Subject Line:

"December 2023 Image Survey"

📉 **Why it likely underperformed:**

- **Low Emotional Appeal**: "Survey" is generally low-engagement unless incentivized.
- **Lacks Urgency or Benefit**: Doesn't tell the user what they gain by opening.
- **Too Formal/Neutral**: Sounds more like internal communication or research rather than a promotional email.
- **No Call to Action**: Passive tone with no compelling hook or verb to drive interaction.

Top 5 percentile performing email body vs bottom 5

If we pair this with previous analysis (Top vs Bottom 25%), we know that

top-performing bodies:

- **Get to the point quickly**
- Use bold, action-driven CTAs (like "Shop Now")
- Highlight specific value or savings
- Reinforce the subject line message visually and textually
- Short, engaging lines

Whereas bottom performers:

- Bury the value or CTA
- Use too much whitespace padding

- Provide less visual contrast for CTAs
- Offer vague or unconvincing messaging

Campaign Date Selection for January 2025

- ○ Selected all Mondays, Wednesdays, and Fridays in January.
- ○ Supplemented with specific historical high-performing dates from previous Januarys, guided by top-performing trends in OR and CTOR.

Prompt forJanuary campaign - 📄 PROMPT FOR january Table

(📗 **Top performing data) (** 📗 **worse performing data)**

Date	Day	Time	Campaign Name	Theme / Purpose	Sample Subject Line	Sample Body Flow (outline)	Pred. OR	Pred. CTOR
Jan 01	Wed	15:00	New-Year Kick-off	Launch 2025 with a bang	Kick Off 2025: 35 % Off First-Day Puzzles 🎉	• 35 %-OFF hero banner → • 3 best-seller tiles → • CTA "Shop Now" → • Free-ship badge	0.55	0.15
Jan 03	Fri	15:00	Weekend Warm-Up	Early weekend offer	Your Puzzle Weekend Starts Now – 30 % Off ❄️	Hook strip → curated 6-pack grid → CTA (color block)	0.54	0.14
Jan 06	Mon	15:00	Top-Picks Reload	Showcase 2024 top sellers	Back by Demand: Best Puzzles of 2024 – 30 % Off	Collage hero → "You voted — we listened" copy → CTA → reviews snippet	0.55	0.15
Jan 08	Wed	15:00	Crowd-Choice Vote	Engagement / NPD input	Help Us Choose Our Next Puzzle ☀️ Vote & Save 20 %	One-sentence hook → image vote buttons → coupon reveal → CTA	0.53	0.13
Jan 10	Fri	15:00	Flash-40	40 % clearance blitz	24-Hour Deal: 40 % Off Clearance Puzzles – Ends Tonight!	Countdown bar → clearance grid → CTA repeated	0.55	0.16
Jan 13	Mon	15:00	New-Arrivals Drop	Fresh puzzle line	Just Landed: Brand-New Puzzles by Piece Count	Large hero → tabs by piece-count → CTA "See What's New"	0.54	0.14
Jan 14	Tue	15:00	Rewards Reminder	Loyalty nudge	Your Rewards Are Waiting – Redeem Today 🟥	Points summary → 20 % bonus → CTA "Use My Rewards"	0.52	0.12
Jan 15	Wed	15:00	VIP Early Access	Exclusive sale	Early Access 🎁 30 % Off Before Anyone Else	VIP code hero → urgency ticker → CTA	0.54	0.14
Jan 17	Fri	15:00	Weekend Kick-Start	Friday teaser	Unlock 25 % Off – Weekend Starts Now!	Offer bar → lifestyle hero → CTA → trust strip	0.53	0.14
Jan 19	Sun	15:00	Community Pulse	Quick survey + perk	Tell Us Your Dream Puzzle & Get 20 % Off	Single-question poll → coupon reveal → CTA	0.51	0.11
Jan 20	Mon	15:00	MLK-Day Special	U.S. holiday promo	Today Only: 20 % Off + Free Ship us	Patriotic banner → top SKUs → CTA mid+end	0.54	0.14
Jan 21	Tue	15:00	Clearance Boost	Mid-week push	You Heard It Here: 30 % Off Clearance Puzzles!	Bold hero → savings stack → CTA → urgency footer	0.52	0.13
Jan 22	Wed	15:00	Smart-Fun Focus	Educational line	Puzzles for Curious Minds 🧠 Save 25 %	Benefit headline → 4 edu puzzles → CTA → testimonial	0.53	0.13
Jan 23	Thu	15:00	Flash-24	24-hour surprise	Flash Sale: 24 Hrs • Up to 35 % Off!	Countdown → curated picks → CTA → social proof line	0.52	0.13
Jan 24	Fri	15:00	Weekend Countdown	48-hr mega sale	Weekend Deal: 40 % Off Starts Now!	40 % hero → tiered blocks → CTA in contrast	0.54	0.15
Jan 25	Sat	15:00	Top-10 Countdown	Content + commerce	Top 10 Puzzles You Can't Miss 🏆	Numbered list → image+rating → CTA under each	0.51	0.12
Jan 27	Mon	15:00	Fan-Favorites Return	Restock alert	Back in Stock: Your Favorite Puzzles – Shop Fast!	"Back in Stock" banner → carousel → CTA "Claim Mine"	0.54	0.14
Jan 29	Wed	15:00	Final-Jan Push	Clearance close	Final January Discounts – Up to 45 % Off	Savings headline → progress bar → CTA	0.53	0.14
Jan 31	Fri	15:00	Feb Sneak Peek	Tease next month	February Sneak Peek – Early Access Starts Today 🗓️	Preview banner → early-access code → CTA → Feb teaser	0.55	0.15

Link 📅 January_2025_Campaign_Calendar