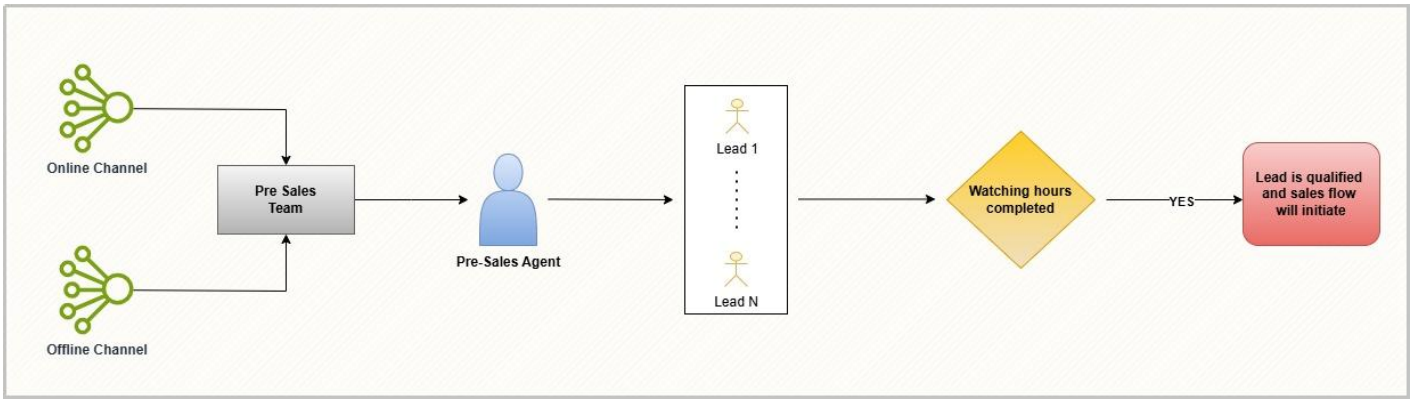
Lead Flow Analysis Report

Flow Diagram



Tools Used

- ✓ Python (Pandas, NumPy, Seaborn, Matplotlib) – for data cleaning, transformation, and analysis
- ✓ Microsoft Excel – for dashboard creation and interactive visualization

Analytical Process & Data Cleaning

Step-by-step breakdown of how the raw data was cleaned and transformed:

1. Removed non-ASCII characters from timestamp fields like “Lead creation date” to convert to datetime format.
2. Used pd.to_datetime with appropriate format strings.
3. Filled missing categorical fields where appropriate, or dropped rows when essential metrics (like webinar duration) were missing.
4. Created custom segments:
    - Engagement Segment based on webinar watch duration
    - Funnel Stage (Lead Created → Webinar Attended → Qualified for Sales)
5. Added flag columns to show if leads were contacted, qualified, or converted.
6. Converted columns like Pre-Sales and Sales team calls into integer values for aggregation.
7. Exported final cleaned dataset for Excel use.
8. Key Metrics & Visualizations

Visualization & Dashboard Metrics

Seven visualizations were created in Excel to support the analysis:

1. Leads by Channel
    - Columns Used: Channel, Lead ID
    - Rationale: Understand which marketing channels generate the most leads
2. Webinar Engagement Funnel
    - Columns Used: Webinar Attendance, Webinar Duration, Funnel Stage
    - Rationale: Show how leads progress from registration to qualification
3. Language-wise Conversion Rate
    - Columns Used: Language, Funnel Stage
    - Rationale: Analyze conversion performance across different language demographics
4. Sales Team Effectiveness
    - Columns Used: Qualified Leads, Sales Calls Attempted, Sales Success
    - Rationale: Evaluate follow-through rate and conversion after handoff to sales
5. Trend of Leads Over Time
    - Columns Used: Lead Creation Date
    - Rationale: Track peak periods of engagement; infer success of campaigns
6. Call Attempt Analysis
    - Columns Used: Total Calls by Pre-Sales and Sales
    - Rationale: Measure outreach effort and drop-off due to no contact

    Dynamic Slicers:

        - Filters were added to make dashboards interactive:
            - Language
            - Channel

Sales Trend vs Campaign Correlation

To understand whether marketing campaigns led to sales spikes:

- Analyzed lead creation and qualification by date.
- Overlaid sales assignment/closing dates on the same timeline.
- Inferred which campaigns likely influenced higher engagement or sales.

This helps optimize marketing spend towards successful campaigns.

Challenges and Resolution

1. Non-standard datetime format
    - Resolved with ASCII cleaning and custom datetime format parsing
2. Extensive missing webinar duration
    - Dropped those rows only in funnel-specific charts, retained for overall stats
3. Mixed-type call attempt fields (text & numbers)
    - Used pd.to_numeric() with error coercion
4. Funnel stage logic confusion
    - Reframed conditions to check attendance first, then duration threshold

Insights & Recommendations

1. Majority of leads drop after registering but not attending webinars. Increase webinar attendance by:
    - More reminder calls by pre-sales team
    - Better timing of sessions
    - Incentivized attendance
2. Hindi and regional language leads show higher conversion. Hence need to focus accordingly.
3. Leads qualified after longer webinars tend to convert better. Improve content and duration engagement. Our content should be interesting and engaging in the initial minutes.
4. The Sales team closes more deals when they receive qualified leads quickly — ideally within 1 to 2 days. So, speeding up the process of assigning leads to Sales can improve conversions.
5. Leads that watch webinars >35 mins convert 3x more. We need to extend duration and improve content delivery.
6. Mostly better engagement is done by channels like digital marketing, IM, offline workshops and outbounds.

Conclusion

This analysis can serve as a live dashboard foundation for sales/marketing heads and provide decision-makers with real-time visibility into their lead flow process.