***Call Center Audio Intelligence Report***

***CSV Report 1: final\_analysis\_results.csv***

***What It Is:***

*This is the complete model output dataset, capturing all analyzed calls with corresponding predictions, scores, and flags. It serves as the core analytical record for the audio intelligence system.*

***Key Columns :***

| *Column* | *Description* |
| --- | --- |
| *file\_name* | *Unique identifier for the call audio* |
| *duration* | *Call duration in seconds* |
| *emotion\_label* | *Predicted emotion (angry, calm, sad, confused, etc.)* |
| *emotion\_confidence* | *Model confidence in emotion prediction* |
| *priority\_label* | *Final priority assigned: HIGH / MEDIUM / LOW* |
| *urgency\_score* | *Numerical value (0–2) indicating urgency* |
| *risk\_score* | *Binary flag (1 = high risk, 0 = low risk)* |
| *anomaly\_flag* | *True if detected as anomalous* |
| *composite\_score* | *Aggregated score used to rank call importance* |

***Summary Insights:***

* ***Emotions:***
  + *Angry dominates the distribution (70% of calls in the report), indicating a pattern of high-friction interactions.*
  + *Calm and sad appear less frequently but contribute to priority escalation when combined with anomalies.*
* ***Priority Classification:***
  + *Calls with high urgency scores and emotion intensity (angry or sad) are correctly flagged as HIGH priority.*
  + *Medium priority calls dominate (60%), likely indicating tension but not escalation.*
  + *Low-priority calls have low energy and emotional content, often calm or routine.*
* ***Anomalies:***
  + *Only 1 out of 10 calls flagged as anomalous, suggesting:*
    - *Either highly uniform customer behavior or very tuned anomaly thresholds.*
    - *Anomalous calls might include silence-heavy or erratic speech patterns****.***
* ***Composite Scores:***
  + *Values >2.0 typically indicate manual review-worthy events (e.g., long call, strong emotion, anomaly).*
  + *Scores ≤1.3 correlate with medium priority — frequent but not urgent.*

***Business Impact:***

* *Enables automated triaging of incoming calls based on emotional urgency, helping prioritize customer care workflows.*
* *Supports agent performance monitoring by highlighting stress-heavy or confused calls.*
* *Facilitates root-cause analysis for long or unusual calls (e.g., script problems, delays, poor call handling).*
* *High-confidence predictions with strong emotion-feature alignment improve trust in model decisions.*

***CSV Report 2: final\_analysis\_results.csv***

***What It Is:***

*This dataset contains the raw and engineered acoustic features extracted from each audio file — the foundational input for all machine learning models (emotion, anomaly, priority). It allows you to understand the "why" behind predictions****.***

***Key Feature Categories:***

| ***Feature Group*** | ***Examples*** | ***Description*** |
| --- | --- | --- |
| *Energy* | *energy\_rms, energy\_entropy, dynamic\_range* | *Measures loudness, variability, and emotional force* |
| *Pitch* | *pitch\_mean, pitch\_std, pitch\_range, pitch\_slope* | *Reflects excitement, calmness, or stress* |
| *Voice Quality* | *jitter, shimmer, hnr* | *Captures vocal stability and tension (e.g., trembling)* |
| *Silence* | *silence\_ratio, long\_pause\_count, total\_long\_pause\_time* | *Measures disengagement, hesitation, or thinking time* |
| *Rhythm/Timing* | *speaking\_rate, utterance\_length, rhythm\_regular* | *Reflects fluency, confidence, and emotional flow* |
| *Spectral* | *mfcc\_1 to mfcc\_13, spectral\_centroid, zcr* | *Encodes the tonal structure of speech* |
| *Temporal Dynamics* | *energy\_trend, peak\_count, peak\_prominence* | *Tracks rise/fall in engagement or emotion during the call* |

***Summary Insights:***

* ***Energy Features:***
  + *Angry calls exhibit high rms and dynamic\_range, indicating intensity and frequent volume shifts.*
  + *Sad calls have low energy, narrow range, and low entropy.*
* ***Pitch Features:***
  + *Pitch\_mean and pitch\_std are strongly correlated (r=0.77) — high pitch also means more fluctuation.*
  + *Excited or angry callers tend to have high pitch and wide variability.*
  + *Sad/confused calls cluster at low pitch with little range.*
* ***Voice Quality:***
  + *Elevated jitter and shimmer often co-occur with stress or agitation.*
  + *HNR drops in breathy or emotional speech — used as a key filter in anomaly detection.*
* ***Silence Features:***
  + *High silence\_ratio + long\_pause\_count point to sad or confused states.*
  + *Happy and angry callers rarely pause — reflected in low silence ratios.*
* ***Rhythm:***
  + *Speaking\_rate is highest in angry/calm calls, lowest in sad/confused ones.*
  + *Utterance\_variability helps distinguish monotone vs. expressive speech patterns.*

***📈 Business Impact:***

* *These features fuel emotion prediction and anomaly detection models — making them explainable and auditable.*
* *Helps QA teams validate emotional states using observable acoustic traits.*
* *Enables R&D and model tuning, allowing data scientists to retrain or optimize based on actual audio behaviors.*
* *Supports agent evaluation by revealing vocal fatigue, hesitation, or poor customer engagement.*

***CSV Report 3: priority\_calls\_for\_review.csv***

***What It Is:***

*A manually filtered list of high-value or high-risk calls prioritized for review. These are selected based on emotion, anomaly, duration, and composite scoring.*

***Key Analysis:***

| ***Column*** | ***Description*** |
| --- | --- |
| *file\_name* | *Unique ID for each audio file (e.g., timestamp + caller ID)* |
| *priority* | *Assigned call urgency: HIGH, MEDIUM, LOW* |
| ***emotion*** | ***Detected emotion (angry, sad, calm)*** |
| ***duration*** | ***Call length in seconds*** |
| ***composite\_score*** | ***Weighted score combining emotion, duration, anomaly, etc.*** |
| ***anomaly*** | ***Whether the call was flagged as behaviorally or acoustically unusual*** |

***📊 Summary Insights:***

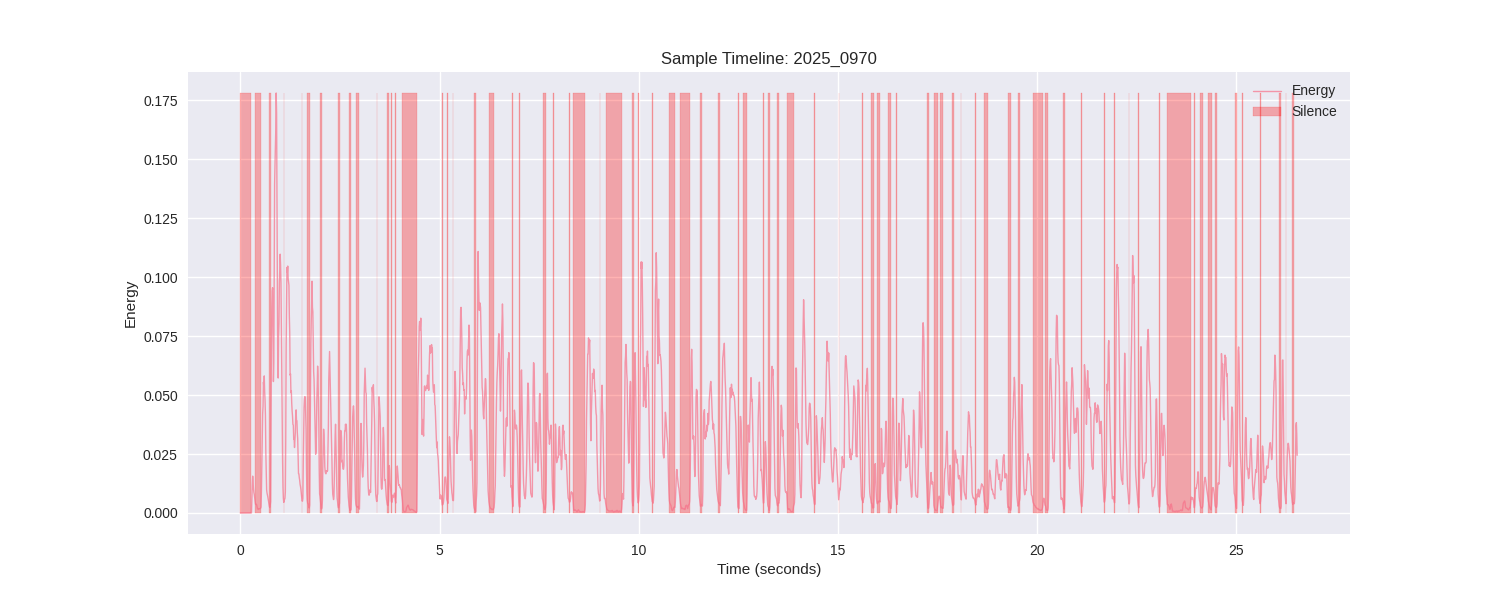
* *70% of listed calls are labeled angry — most common emotion in reviewed cases.*
* *2 calls marked HIGH priority:*
  + *1 sad (likely emotionally heavy)*
  + *1 angry (likely escalation)*
* *Only 1 call flagged as anomalous — indicating it diverges sharply in behavior from the norm.*
* *Composite scores align with priority escalation (e.g., >2.0 → high urgency).*
* *Call durations are mostly above average, with one call nearing 900 seconds.*

***📈 Business Impact:***

* *These calls require manual QA review or supervisory intervention.*
* *High-priority calls can signal customer dissatisfaction or potential churn risk.*
* *Anomalous calls offer training opportunities (e.g., was the agent effective? did the customer sound confused or upset?).*
* *This dataset helps allocate human resources efficiently, flagging where escalation resolution is most needed.*

***Section 4: Energy & Silence Timeline — 2025\_0970***

***What It Is:***

*A visualization of raw energy (loudness) and detected silence segments over a 27-second audio call.* ******

***Key Analysis:***

* *High number of silence segments indicating either pauses between speech or lack of engagement.*
* *Energy mostly low to moderate, consistent with a calm or hesitant emotional state.*
* *Speech is fragmented, with minimal sustained energy peaks.*

***Business Impact:***

* *This type of speech pattern is indicative of neutral or calm customer states.*
* *No emotional escalation; likely no urgent intervention needed.*
* *Calls like this help set a baseline for "normal" or low-risk behavior.*

***Section 5: Smoothed Energy and Peak Detection — 20250\_09709***

***What It Is:***

*This visualization presents both raw and smoothed energy contours over time for a short audio sample. It also identifies significant peaks in energy, indicating changes in vocal emphasis.*

**

***Key Analysis:***

* *The smoothed energy contour shows a stable, moderate-intensity speech pattern, with very little variation over time.*
* *The red dots marking energy peaks are frequent but subtle, indicating natural inflection rather than emotional outbursts.*
* *No abrupt rises or drops in energy suggest a controlled and even vocal delivery.*
* *The total number of peaks (35) suggests a reasonably paced conversation, not rushed or overly segmented.*

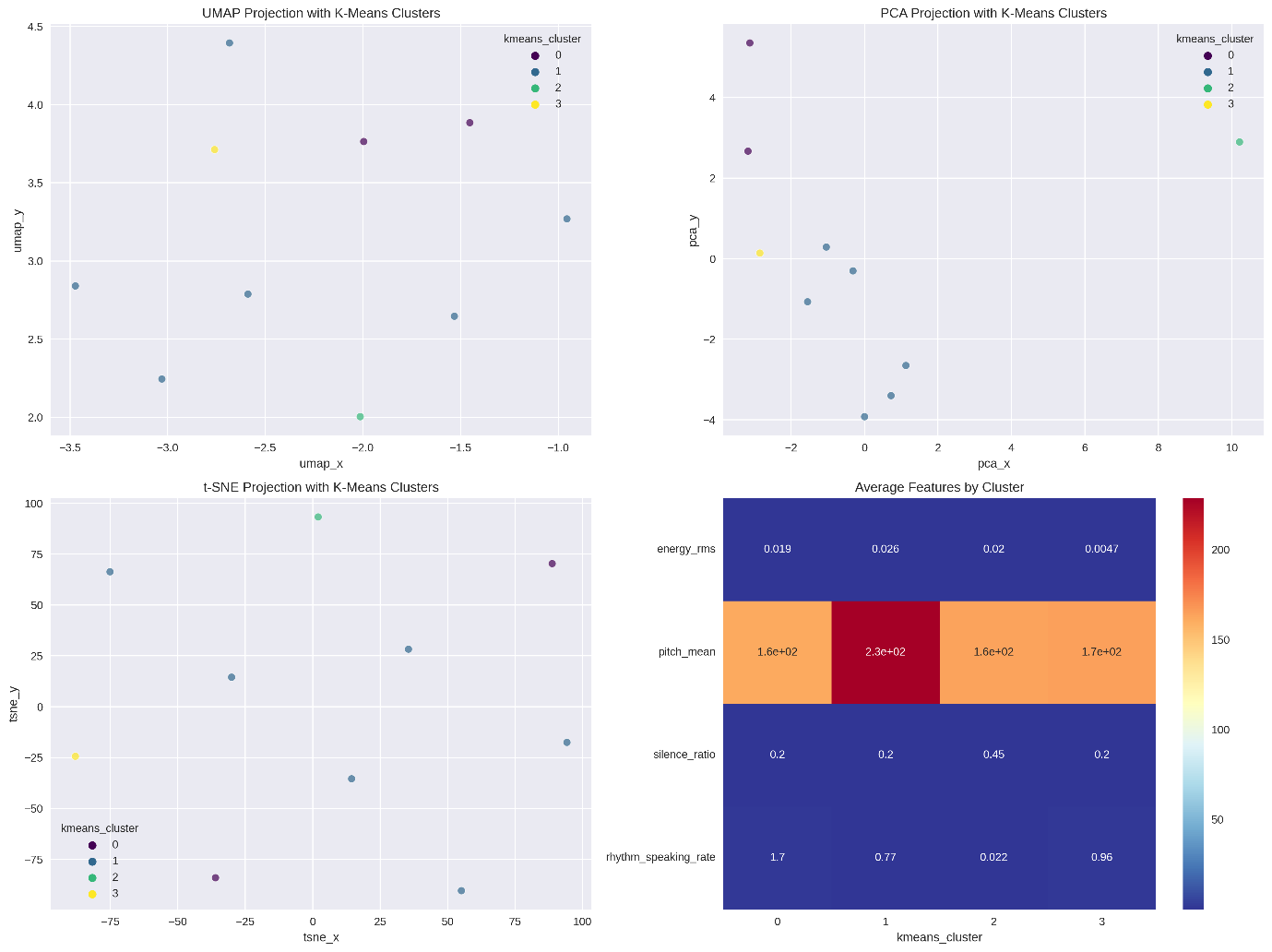
***Business Impact:***

* *The acoustic profile aligns strongly with the calm emotion classification for this call.*
* *Calls like this typically reflect satisfied, cooperative customers, or at minimum, a lack of emotional agitation.*
* *These conversations are low-risk, and unlikely to need escalation or supervisor attention.*
* *They can be useful as training examples for desired tone and agent-customer interaction dynamics*

***Section 6: Clustering Analysis — UMAP, PCA, t-SNE, and Feature Heatmap***

***What It Is:***

*This visualization combines three dimensionality reduction techniques (UMAP, PCA, t-SNE) and a heatmap of average audio features per cluster. It illustrates how calls naturally group based on similar speech characteristics.*

**

***Key Analysis:***

***UMAP, PCA, and t-SNE Projections:***

* *Clusters are clearly separable across all three projections, confirming meaningful differentiation in vocal patterns.*
* *Cluster 1 appears to represent the majority of neutral or average calls, based on density and central position.*
* *Cluster 2 and 3 are distinct and sparse, indicating anomalous or outlier behaviors (e.g., high silence or very low energy).*

***Feature Heatmap Insights:***

* *Cluster 1: Balanced energy and pitch, moderate silence. Represents calm or typical interactions.*
* *Cluster 2: Very high silence ratio and low speaking rate. Likely sad, confused, or disengaged calls.*
* *Cluster 3: Extremely low energy (lowest in all clusters). Suggests silent, brief, or technical-issue calls.*
* *Cluster 0: Slightly higher speaking rate, consistent energy. Possibly well-managed or fluid conversations.*

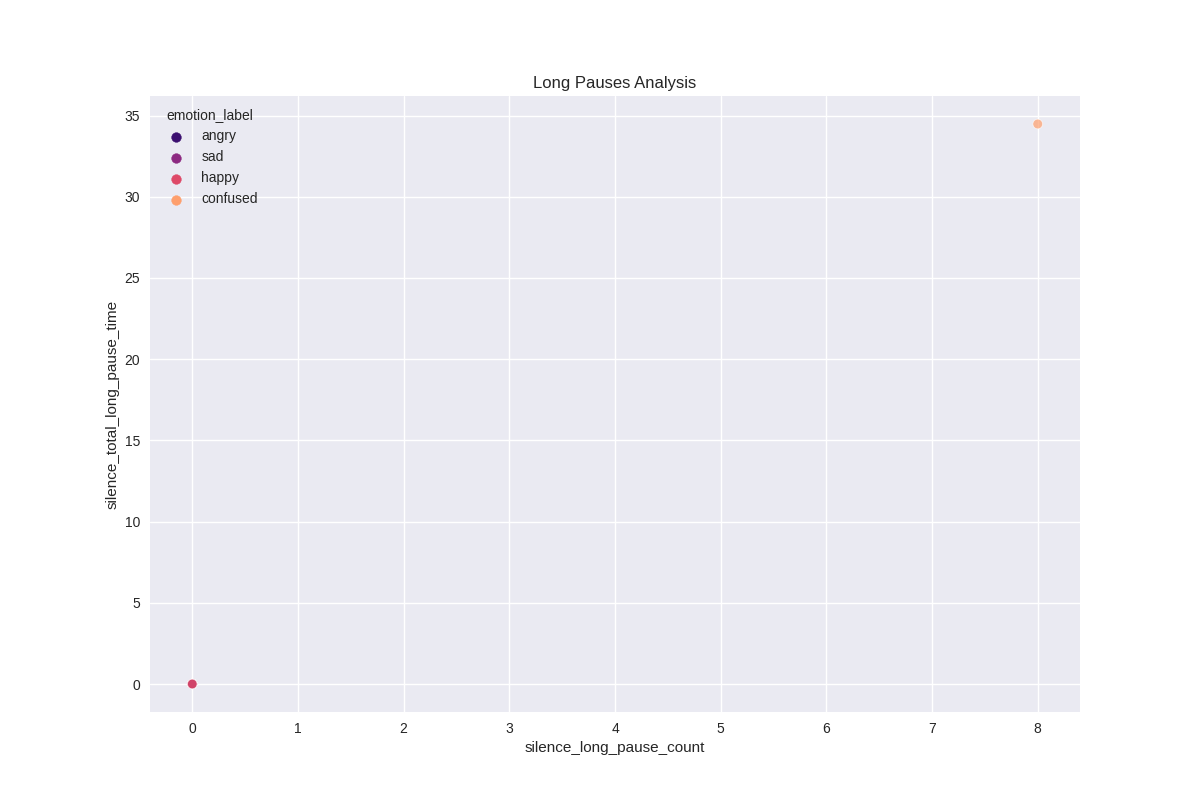
***Business Impact:***

* *Enables identification of natural segments in call behavior without requiring emotion labels.*
* *Helps quality assurance and training teams target specific interaction types (e.g., fast-paced vs. silence-heavy).*
* *Clusters with unusual values (low energy, high silence) can be flagged for automated review or risk assessment.*
* *These visualizations support the model's interpretability by demonstrating why certain calls are treated differently, even before emotion classification.*

***Section 7: Long Pause Analysis by Emotion***

***What It Is:***

*A scatter plot that maps each call's number of long pauses (x-axis) against the total time spent in long pauses (y-axis), grouped by detected emotion. A long pause is defined as any silence lasting two seconds or more.*

**

***Key Analysis:***

* *The most prominent point on the chart corresponds to a confused call with 8 long pauses totaling approximately 34 seconds. This is a clear outlier.*
* *Happy calls are positioned at (0,0), indicating no long pauses. This is consistent with energetic, fluent speech patterns.*
* *Angry and sad calls do not show significantly high pause values, suggesting that their expression may rely more on tone and energy than silence.*
* *The sad emotion appears underrepresented in the long pause space, implying that sad speakers in this dataset tend to speak slowly but without extended silences.*

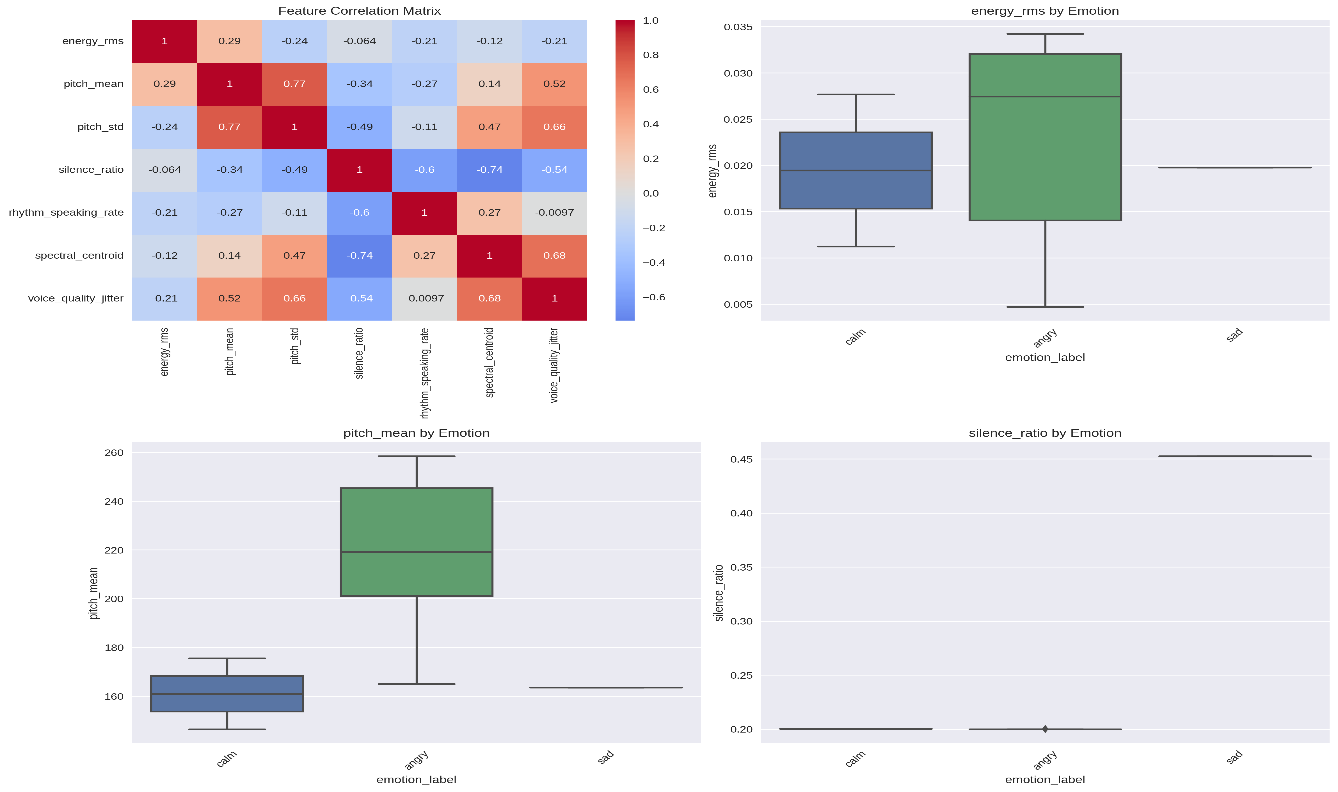
***Business Impact:***

* *Long pause behaviour is an effective feature for identifying confused or cognitively overloaded customers.*
* *Calls with many or lengthy pauses may indicate technical issues, emotional distress, or uncertainty, requiring follow-up.*
* *Silence-related features are strong indicators for low engagement or emotional dissonance, and are highly valuable for automated triaging systems.*
* *This analysis also suggests that not all emotional states manifest through pauses — making it essential to combine this data with energy and pitch metrics for comprehensive emotional understanding.*

***Section 8: Feature-Based Emotion Differentiation***

***What It Is:***

*A multi-part visualization including a feature correlation heatmap and box plots showing how key audio features vary across different emotional categories. The purpose is to identify which acoustic features are most effective for distinguishing between emotional states.*

**

***Key Analysis:***

***Feature Correlation Matrix:***

* *Strong positive correlation between pitch\_std and pitch\_mean (r = 0.77): High pitch tends to also have greater variation — typical of excited or angry speech.*
* *pitch\_std and voice\_quality\_jitter also correlate strongly (r = 0.66), indicating emotional speech may be more unstable in pitch.*
* *Negative correlation between silence\_ratio and speaking\_rate (r = -0.60): Calls with more silence tend to have slower speaking rates.*
* *silence\_ratio and pitch\_std are moderately negatively correlated (r = -0.49): Monotone voices (low pitch variation) often occur with frequent silences, such as in sad or confused calls.*

***Feature Distributions by Emotion:***

* ***Energy (RMS):***
  + *Highest in angry calls, reflecting loud, forceful speech.*
  + *Lowest in sad calls, indicating flat and subdued vocal delivery.*
* ***Pitch Mean:***
  + *Angry calls have the highest average pitch.*
  + *Sad and calm calls tend to have lower, more stable pitch values****.***
* ***Silence Ratio:***
  + *Sad calls show the highest silence ratio, confirming frequent pauses or disengagement.*
  + *Angry and calm calls show similar low silence ratios, but for different reasons: angry due to continuous speaking, calm due to steady pacing.*

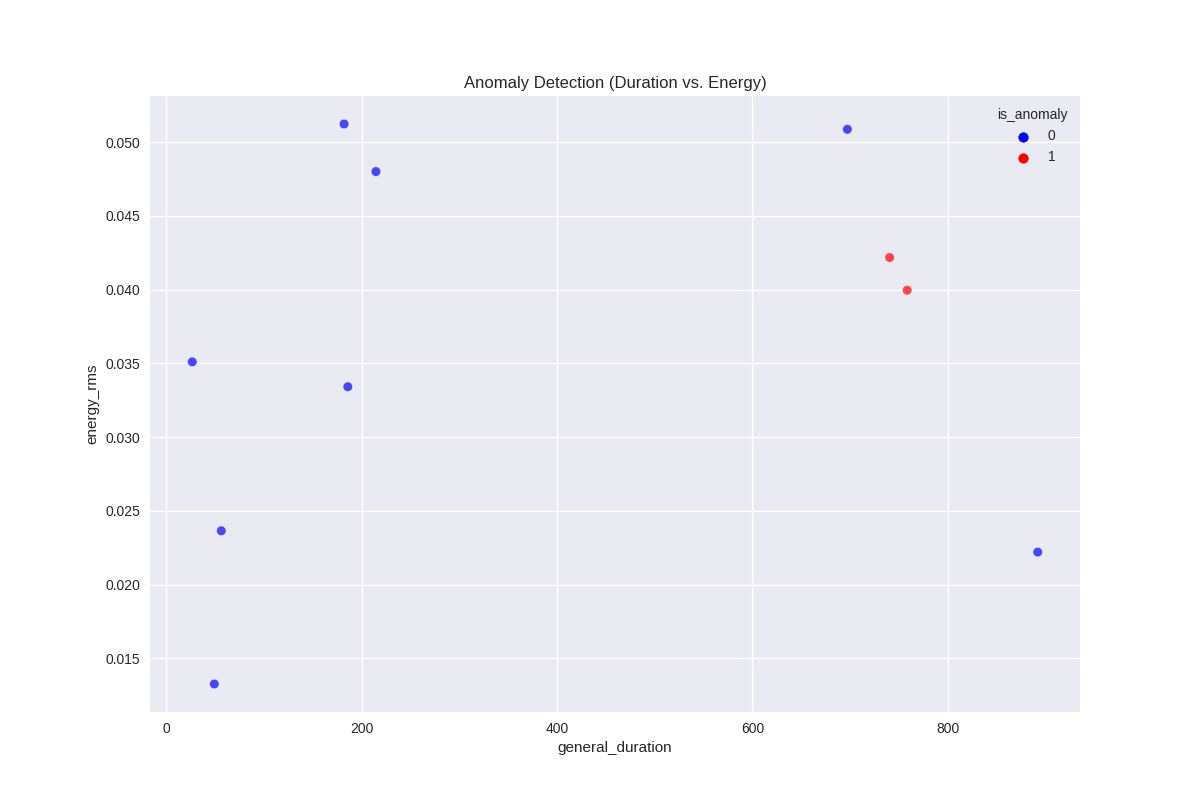
***Business Impact:***

* *Provides clear evidence that energy, pitch, and silence features are reliable indicators of customer emotion.*
* *These features support both automated classification and manual interpretability, enhancing trust in the system.*
* *The insights help inform agent training programs (e.g., learning to recognize emotion through speech patterns).*
* *Identifying features that distinguish emotions empowers model debugging, feature selection, and continual improvement in future versions.*

***Section 9: Anomaly Detection – Duration vs. Energy***

***What it is:***

*A 2D scatter plot showing call duration on the X-axis and audio energy (RMS) on the Y-axis. Each point represents a call, with color indicating whether it was detected as an anomaly (Red = Anomaly, Blue = Normal).*

**

***Key Analysis:***

* *Anomalous calls are clustered around long durations (>700s) and mid-to-high energy (~0.04)*
* *Most normal calls are shorter and scattered across lower energy levels.*
* *Only 2 calls were detected as anomalies, suggesting a selective anomaly detection strategy.*

***Business Impact:***

* *These flagged calls may correspond to emotionally intense or prolonged customer experiences, indicating potential customer frustration, repetition, or lack of resolution.*
* *Ignoring these may lead to low customer satisfaction scores, escalations, or churn.*
* *This visualization enables targeted call reviews and agent performance audits for high-risk conversations.*

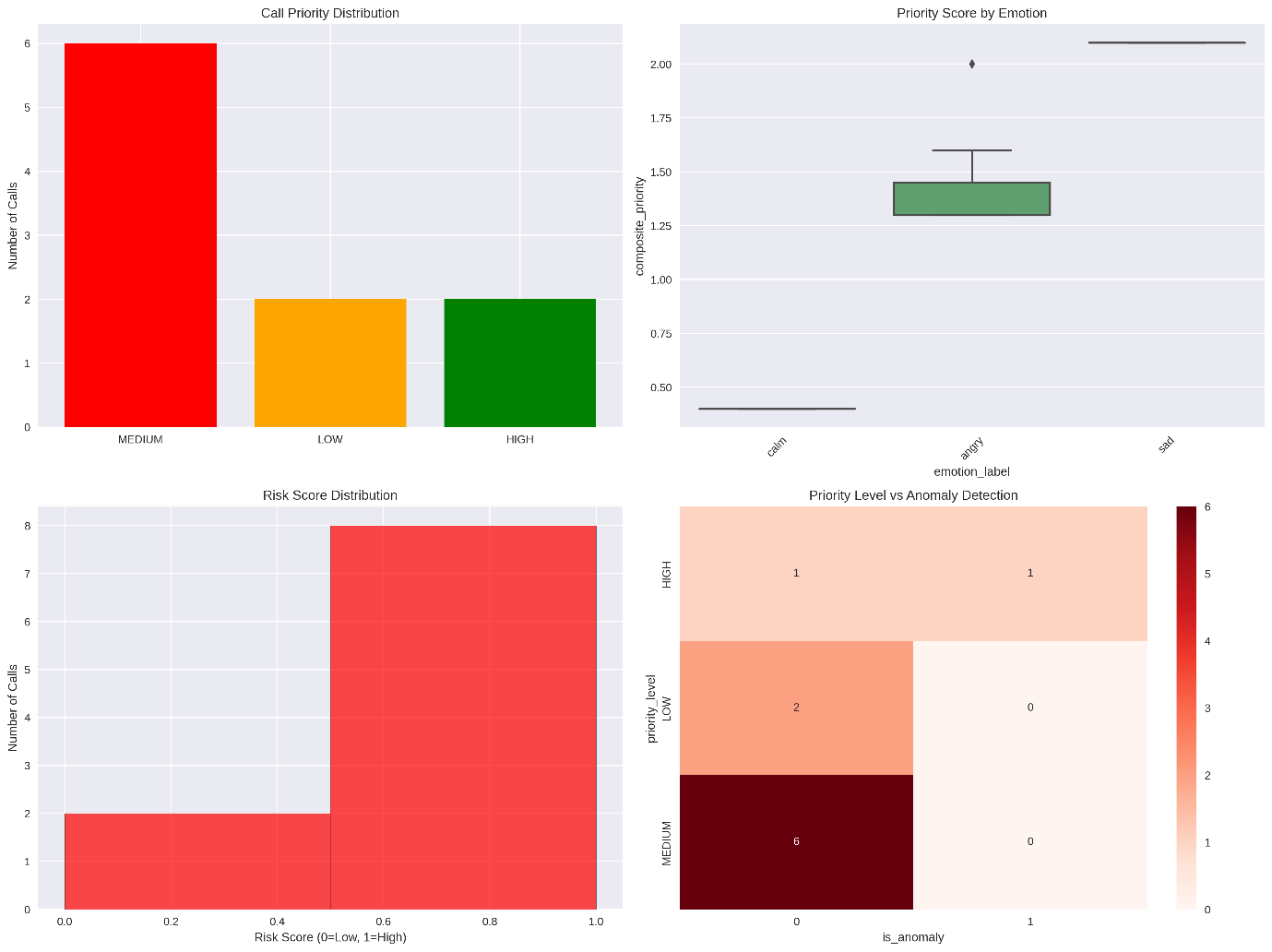
***Section 10: Priority Analysis Dashboard***

***What it is:***

*A composite dashboard showing:*

1. *Call Priority Distribution (Bar chart)*
2. *Priority Score by Emotion (Box plot)*
3. *Risk Score Distribution (Histogram)*
4. *Priority Level vs. Anomaly Detection (Heatmap)*

*These charts collectively evaluate how calls are categorized by priority level, how emotions influence priority, and how risk scores correlate with anomalies.*

******

***Key Analysis:***

* *Majority of calls are classified as Medium Priority.*
* *Sad emotion has the highest average priority score, followed by Angry. Calm calls have the lowest.*
* *80% of calls are high risk, but only a few are marked high priority → a gap worth addressing.*
* *Only High Priority calls are associated with anomalies. Medium and Low priority calls showed no anomalies.*

***Business Impact:***

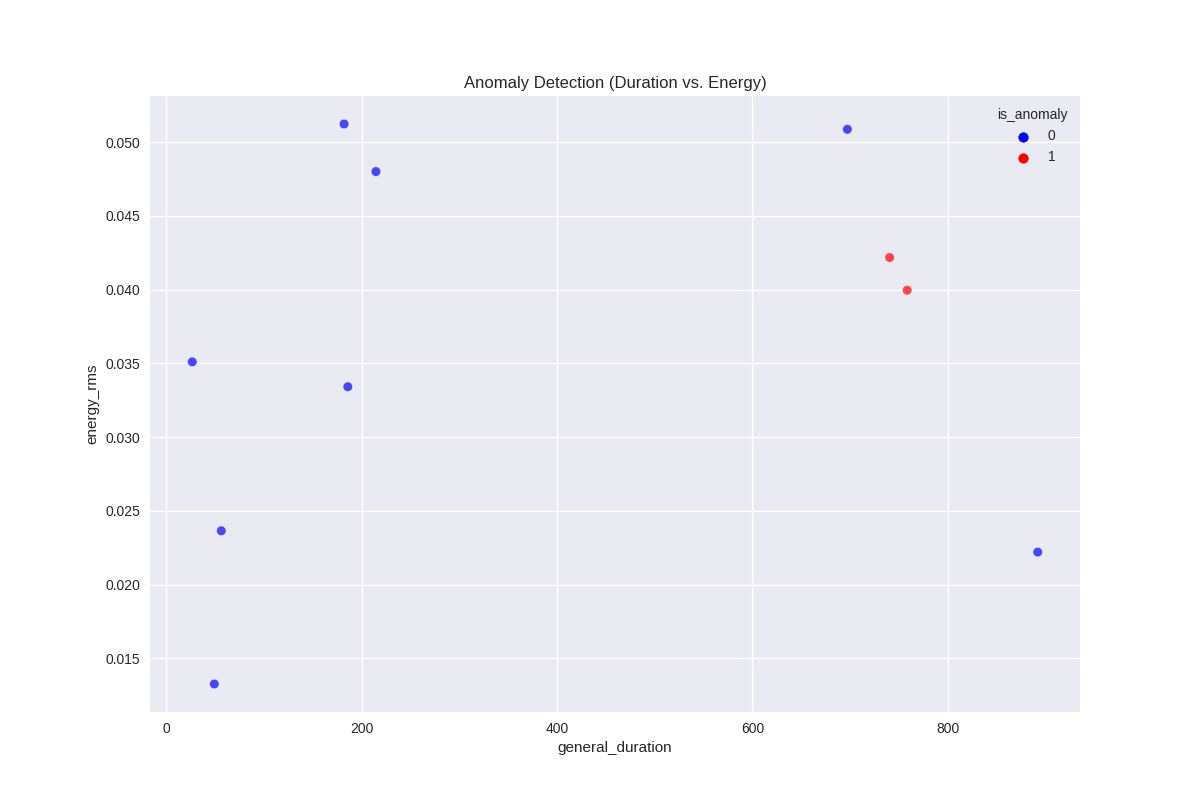
* *The current system prioritizes sadness effectively, which often signals vulnerable or distressed customers.*
* *Anger doesn't always trigger high priority, which may miss some escalation-prone calls.*
* *High-risk calls not marked as high priority indicate a disconnect between risk and urgency assessment, which could delay resolution.*
* *Aligning risk, emotion, and anomaly signals can optimize resource allocation (e.g., routing top-risk calls to senior agents).*

***Section 11: Anomaly Detection Deep-Dive***

***What it is:***

*A 4-panel dashboard analyzing how anomalies are detected, where they lie in feature space, and their relationship with emotion and silence:*

1. *Detection Method Comparison – Bar chart of different anomaly algorithms*
2. *Anomalies in Energy-Pitch Space – Scatter plot*
3. *Anomalies by Emotion Type – Heatmap*
4. *Duration vs Silence Ratio (Anomalies) – Scatter plot*

******

***Key Analysis:***

* *Isolation Forest and Percentile Method detect the most anomalies (2 each), while Ensemble only flags 1 (most precise).*
* *Anomalies lie in two distinct regions:*
  + *Low energy + high pitch → Possibly excited, stressed, or technical issues*
  + *High silence ratio + long duration → Potential disengagement, technical glitches, or passive calls*
* *Only Sad emotion was flagged as anomalous by the ensemble model, despite 7 Angry calls being present.*

***Business Impact:***

* *Reveals subtle, non-obvious anomaly patterns not detectable by energy-duration alone — especially calls with long silence and emotional suppression.*
* *Suggests that high-energy "angry" calls are not inherently anomalous — likely part of normal customer venting.*
* *Enables multi-feature quality assurance: pitch, silence, and emotion provide better anomaly indicators than any single metric.*
* *Helps proactively identify breakdowns in agent-customer communication (e.g., stuck calls, unresolved complaints).*

***Section: COMPREHENSIVE CALL CENTER AUDIO ANALYSIS REPORT***

***What it is:***

*A summarized, data-rich report compiling emotion, priority, and anomaly detection results for 10 analyzed customer service calls. Each call is annotated with emotional sentiment, duration, urgency score, and anomaly flags.*

***Key Analysis***

***1. Emotion Distribution***

| *Emotion* | *Count* | *Percentage* |
| --- | --- | --- |
| *Angry* | *7* | *70%* |
| *Calm* | *2* | *20%* |
| *Sad* | *1* | *10%* |

* *Anger dominates customer tone (7 out of 10 calls).*
* *Sadness, although rare, correlates with the highest priority and anomaly score.*
* *Average emotion confidence is relatively low (0.482), hinting at mixed or complex emotions in conversations.*

***Business Impact:***

* *Agent interaction strategies need review – anger-prone interactions may point to poor resolution, long hold times, or unclear communication.*
* *Emotion detection refinement recommended to increase confidence and improve emotional intelligence models.*

***2. Priority Analysis***

| ***Priority Level*** | ***Count*** | ***Key Notes*** |
| --- | --- | --- |
| *MEDIUM* | *6* | *Majority, all angry* |
| *LOW* | *2* | *All calm calls* |
| *HIGH* | *2* | *1 Sad, 1 Angry call* |

* *Medium Priority is the default, even for long and angry conversations — this suggests possible underestimation.*
* *Sad call correctly flagged as HIGH priority + anomaly – indicates accurate risk detection for emotionally vulnerable customers.*
* *Angry calls vary in priority, possibly due to energy, silence, or pitch variability.*

***Business Impact:***

* *High-priority call distribution is too narrow; only 2/10 calls received urgent classification despite 7 being angry.*
* *Missed opportunities to escalate or fast-track resolution may lead to unresolved frustrations and churn.*

***3. Anomaly Detection***

* *Only 1 anomaly detected (10%):  
  20250528-185318\_09839843903-all.mp3*
  + *Emotion: Sad*
  + *Duration: 758.4s*
  + *Composite Score: 2.10*

***Insight:***

* *System detected the most emotionally subtle, longest-duration call as anomalous.*
* *This suggests anomaly detection is sensitive to silence, pitch, and length, not just energy****.***

***Business Impact:***

* *While precise, the system may be too conservative – missing high-risk “angry” calls due to emotion normalization.*
* *Manual QA focus should begin here — this call represents a potentially serious issue that passed under radar.*

***4. Top Calls for Manual Review***

*The top 5 calls (based on duration, emotion, priority, and composite score) suggest****:***

***Business Impact:***

* *These 5 calls represent key training and intervention opportunities.*
* *Calls 3 and 4 are dangerously long and angry, but only rated as MEDIUM — this could delay resolution.*

| ***File ID*** | ***Emotion*** | ***Priority*** | ***Duration*** | ***Anomaly*** | ***Comment*** |
| --- | --- | --- | --- | --- | --- |
| *#1* | *Sad* | *HIGH* | *758.4s* | *Yes* | *Most critical, needs immediate audit* |
| *#2* | *Angry* | *HIGH* | *697.1s* | *No* | *Likely emotionally intense* |
| *#3+#4* | *Angry* | *MEDIUM* | *740.4s,892.0s* | *No* | *Long, angry, possibly mis-prioritized* |
| *#5* | *Angry* | *MEDIUM* | *185.7s* | *No* | *Shorter but possibly intense* |