Daily Ticket Spike Detection

Date Initiated: April 2025
Prepared For: Driver Excellence and whomsoever it may concern

---

Objective

This system identifies **unusual surges (spikes)** in complaint tickets raised by drivers, broken down by **issue category** and **driver zone**, and generates **daily alerts** for proactive intervention.
The aim is to distinguish **normal volume fluctuations** from **true anomalies** based on historical trends, ensuring timely alerts that prompt resolution and root-cause analysis.

Process Flow

1. Data Extraction

- Filters for selecting tickets:
    - Only those tickets that were raised by drivers were considered
    - Last 6 months of data (since October 2024)
    - Only active issue categories and issues were considered
    - Driver's zone was taken for counting tickets

---

2. Aggregation

- Tickets computed are calculated zone-wise and category-wise.
- Simultaneously, the top-contributing **issue within each category** is used based on the count of tickets and not max change in count of tickets.

---

3. Trend Modeling via Rolling Averages

For each Zone–Category pair:

- Historical daily ticket volumes are analyzed using **rolling windows approach**:
    - **7-day**, **30-day**, and **90-day**
- Metrics computed:
    - Average
    - Standard Deviation
    - Minimum and Maximum

Why Rolling Windows?
They capture short-term (weekly) and long-term (monthly/quarterly) behavior patterns, providing a dynamic, adaptive benchmark for what's "normal."

---

4. Anomaly Detection Logic

- **Safe Threshold**: Highest of three `(average + standard deviation)` bands
  - Rolling 7 days (average + std deviation)
  - Rolling 30 days (average + std deviation)
  - Rolling 90 days (average + std deviation)

- A **spike** is flagged if:
  - Ticket count exceeds the safe threshold
  - Count based threshold further reduce the number of alerts as:
    - Increase from safe zone should be at least of ≥ 3 tickets count
    - Ticket count of category should be ≥ 5 tickets to be considered

---

5. Severity Classification

Spikes are categorized as **High**, **Medium**, or **Low** based on:

- **Volume of activity** (High vs Low category historically)

- **Magnitude of increase** (absolute & % increase)

Volume Level	High Severity	Medium Severity	Low Severity
Low Volume (avg < 25)	≥20 tickets or ≥200%	≥10 tickets or ≥100%	<10 tickets or <100%
High Volume (avg ≥ 25)	≥50 tickets or ≥70%	≥30 tickets or ≥40%	<30 tickets or <40%

This adaptive thresholding ensures fair detection across both high-frequency and niche categories.

---

6. Reporting & Alerts

Google Sheets (Audit Trail)

- `raw`: Replaced daily
- `dailyTicketLogs`: Appended for traceability and checking past data

Slack Alerts

- Posted only for **High** and **Medium** spikes
- Sent to: `#central-alerts` (configurable)
- Message format:
  - Zone, Category
  - Ticket count vs Safe Threshold
  - % increase
  - Top Issue
  - Whether it's the **highest in 7/30/90 days**

What Makes This System Robust

Feature	Benefit
Adaptive windows (7/30/90)	Captures both short-term volatility and seasonal trends
Zone-level granularity	Enables targeted ops intervention
Top issue attribution	Speeds up RCA process
Severity tiers	Filters noise, ensures only action-worthy alerts
Multi-channel outputs	Google sheet + Slack alerts for real-time visibility

---

Edge Handling & Safeguards

- **Missing trend data**: Handled via historical backfill logic
- **False positives in low volume areas**: Controlled using thresholds on minimum counts and differences
- **Slack alert errors**: Fallback logic to avoid complete script failure
- **Safe zone definition**: Max of 3 bands ensures capturing of seasonal or program based increase in tickets efficiently to decrease chances of false alerts.

Slack message meaning

The Slack message includes key visual and textual elements that help teams quickly understand the severity and context of a ticket spike. Below are explanations of specific message components:

1. Peak Note (`peak_note`)
This line adds contextual information about how exceptional the current day's spike is compared to recent history:

- 📈 `Highest in L90!`
  → The current ticket count is the highest in the **last 90 days** — a strong indicator of a rare or serious spike.
- 📈 `Monthly high!`
  → The ticket count is the highest in the **last 30 days**, signaling a recent uptick.
- 📈 `Weekly high!`
  → The count is the highest in the **past week**, flagging short-term urgency.

2. Severity Icons (`severity_emoji`)
Each ticket spike is assigned a severity label and an accompanying **emoji** to make the alert **visually scannable**:

- 🔴 – **High Severity**
  Indicates a major spike requiring immediate investigation.
- 🟠 – **Medium Severity**
  Represents a moderate anomaly that should be reviewed soon.
- 🟡 – **Low Severity**
  A small but valid deviation, flagged for awareness.
- ⚪ – **Unknown**
  Used when severity couldn't be calculated due to missing or invalid data.

3. Ticket Delta Breakdown (`(+X, +Y%)`)
This part of the message quantifies the spike using two intuitive metrics:

- `+X` → **Absolute increase** in number of tickets compared to historical average
  (e.g., +14 means 14 more tickets than usual)
- `+Y%` → **Percentage increase** relative to average
  (e.g., +120% indicates the volume more than doubled)

Both metrics help balance perspective: large percentages in low-volume categories are not overhyped, and small percentages in high-volume zones are not ignored.