

BoutScout CSV Format Instructions

Input file requirements for local browser-based incubation inference

Authors

Jorge Lizarazo - Center for Animals on the Move, Western University, London, ON, Canada

Gustavo Londoño - Facultad Barberi de Ingeniería, Diseño y Ciencias Aplicadas, Universidad Icesi, Cali, Colombia

Languages included: English, Español, Português do Brasil, Français (Québec).

Core requirement	Meaning
One row per minute	Each complete day requires exactly 1,440 unique minute records.
Recommended timestamp	YYYY-MM-DD HH:MM:SS
Required columns	record_id, datetime, egg_temperature, ambient_temperature
Incomplete days	Skipped automatically. No interpolation, padding, or repeated values.

This document is designed to be downloaded before uploading CSV files to BoutScout. The app runs inference locally in the browser. The input CSV is not uploaded to an external server by the app.

English - CSV format instructions

Input file requirements for local browser-based incubation inference.

1. General rule

BoutScout expects one row per minute. Each complete day must contain exactly 1,440 unique minutes, from 00:00 to 23:59. The app processes the file locally in your browser, groups the data by calendar date, removes duplicated minutes, and runs inference only on complete days.

2. Incomplete days are excluded

Incomplete days are excluded automatically. BoutScout does not fill missing minutes, interpolate temperatures, or repeat previous values. This avoids artificial flat sections at the beginning or end of a day.

3. Required CSV structure

The current web version reads the CSV by column position. The names below are strongly recommended for reproducibility, but the app currently uses the order of the columns.

Position	Recommended name	Description
Column 1	record_id	Recommended. Any row number, logger ID, or record identifier. Currently ignored by the app.
Column 2	datetime	Required. Date and time for each minute. This is the most important column.
Column 3	egg_temperature	Required. Nest, egg, or sensor temperature used as the main signal for inference.
Column 4	ambient_temperature	Recommended. Ambient temperature. If missing or non-numeric, the app falls back to egg_temperature.

4. Date and time format

Use a clear timestamp in the recommended format: YYYY-MM-DD HH:MM:SS. Avoid ambiguous dates such as 06/15/24 or 15/06/24, and avoid AM/PM unless you have already tested it locally.

5. Complete-day rule

BoutScout keeps only days with exactly 1,440 unique minute records, where the first minute is 00:00 and the last minute is 23:59. Days with missing minutes, duplicated timestamps, incomplete edges, or unparseable values are skipped.

6. Temperature values

Temperature columns must contain numeric values only. Do not include degree symbols, units, text labels, or commas as decimal separators.

7. Recommended preprocessing

Sort the file by datetime, confirm a one-minute sampling interval, review duplicated timestamps, keep the required columns in the first four positions, and test first with a small file.

8. Outputs after inference and manual correction

BoutScout exports three local CSV files: minute-by-minute predictions, continuous event summaries, and daily summaries with nest attentiveness, bout durations, event counts, and temperature summaries.

Minimal valid example

```
record_id,datetime,egg_temperature,ambient_temperature
1,2024-06-15 00:00:00,31.82,21.75
2,2024-06-15 00:01:00,31.80,21.76
3,2024-06-15 00:02:00,31.79,21.77
...
1440,2024-06-15 23:59:00,32.10,22.40
```

Checklist

CSV file uses comma separators.

First row contains headers.

Column 2 contains datetime.

Column 3 contains nest, egg, or sensor temperature.

Column 4 contains ambient temperature when available.

Each complete day has 1,440 records from 00:00 to 23:59.

No units or text are mixed into temperature columns.

Dates are not ambiguous.

Español - Instrucciones de formato CSV

Requisitos del archivo de entrada para inferencia local de incubación en el navegador.

1. Regla general

BoutScout espera una fila por minuto. Cada día completo debe contener exactamente 1.440 minutos únicos, desde 00:00 hasta 23:59. La aplicación procesa el archivo localmente en tu navegador, agrupa los datos por fecha calendario, elimina minutos duplicados y ejecuta la inferencia solo en días completos.

2. Los días incompletos se excluyen

Los días incompletos se excluyen automáticamente. BoutScout no rellena minutos faltantes, no interpola temperaturas y no repite valores previos. Esto evita secciones planas artificiales al inicio o al final de un día.

3. Estructura requerida del CSV

La versión web actual lee el CSV por posición de columna. Los nombres de abajo son fuertemente recomendados para reproducibilidad, pero la aplicación usa actualmente el orden de las columnas.

Position	Recommended name	Description
Columna 1	record_id	Recomendada. Cualquier número de fila, ID del registrador o identificador del registro. Actualmente la app la ignora.
Columna 2	datetime	Requerida. Fecha y hora para cada minuto. Es la columna más importante.
Columna 3	egg_temperature	Requerida. Temperatura del nido, huevo o sensor usada como señal principal para la inferencia.
Columna 4	ambient_temperature	Recomendada. Temperatura ambiental. Si falta o no es numérica, la app usa egg_temperature como respaldo.

4. Formato de fecha y hora

Usa una marca de tiempo clara en el formato recomendado: YYYY-MM-DD HH:MM:SS. Evita fechas ambiguas como 06/15/24 o 15/06/24, y evita AM/PM salvo que ya hayas probado ese formato localmente.

5. Regla de día completo

BoutScout conserva solo los días con exactamente 1.440 registros únicos por minuto, donde el primer minuto es 00:00 y el último 23:59. Los días con minutos faltantes, marcas de tiempo duplicadas, bordes incompletos o valores que no puedan leerse se omiten.

6. Valores de temperatura

Las columnas de temperatura deben contener únicamente valores numéricos. No incluyas símbolos de grados, unidades, etiquetas de texto ni comas como separador decimal.

7. Preprocesamiento recomendado

Ordena el archivo por fecha y hora, confirma un intervalo de muestreo de un minuto, revisa marcas de tiempo duplicadas, deja las columnas requeridas en las primeras cuatro posiciones y prueba primero con un archivo pequeño.

8. Salidas después de la inferencia y corrección manual

BoutScout exporta tres archivos CSV locales: predicciones minuto a minuto, resumen de eventos continuos y resumen diario con atención al nido, duración de bouts, conteo de eventos y resúmenes de temperatura.

Ejemplo mínimo válido

```
record_id,datetime,egg_temperature,ambient_temperature
1,2024-06-15 00:00:00,31.82,21.75
2,2024-06-15 00:01:00,31.80,21.76
3,2024-06-15 00:02:00,31.79,21.77
...
1440,2024-06-15 23:59:00,32.10,22.40
```

Checklist

El CSV usa comas como separador.

La primera fila contiene encabezados.

La columna 2 contiene datetime.

La columna 3 contiene temperatura de nido, huevo o sensor.

La columna 4 contiene temperatura ambiental cuando esté disponible.

Cada día completo tiene 1.440 registros desde 00:00 hasta 23:59.

No hay unidades ni texto mezclados en las columnas de temperatura.

Las fechas no son ambiguas.

Português do Brasil - Instruções de formato CSV

Requisitos do arquivo de entrada para inferência local de incubação no navegador.

1. Regra geral

O BoutScout espera uma linha por minuto. Cada dia completo deve conter exatamente 1.440 minutos únicos, de 00:00 a 23:59. O aplicativo processa o arquivo localmente no navegador, agrupa os dados por data do calendário, remove minutos duplicados e executa a inferência apenas em dias completos.

2. Dias incompletos são excluídos

Dias incompletos são excluídos automaticamente. O BoutScout não preenche minutos ausentes, não interpola temperaturas e não repete valores anteriores. Isso evita trechos planos artificiais no início ou no final do dia.

3. Estrutura obrigatória do CSV

A versão web atual lê o CSV pela posição das colunas. Os nomes abaixo são fortemente recomendados para reprodutibilidade, mas o aplicativo atualmente usa a ordem das colunas.

Position	Recommended name	Description
Coluna 1	record_id	Recomendada. Qualquer número de linha, ID do registrador ou identificador do registro. Atualmente é ignorada pelo app.
Coluna 2	datetime	Obrigatória. Data e hora de cada minuto. É a coluna mais importante.
Coluna 3	egg_temperature	Obrigatória. Temperatura do ninho, ovo ou sensor usada como sinal principal para a inferência.
Coluna 4	ambient_temperature	Recomendada. Temperatura ambiental. Se estiver ausente ou não for numérica, o app usa egg_temperature como fallback.

4. Formato de data e hora

Use uma marca temporal clara no formato recomendado: YYYY-MM-DD HH:MM:SS. Evite datas ambíguas como 06/15/24 ou 15/06/24, e evite AM/PM a menos que esse formato já tenha sido testado localmente.

5. Regra de dia completo

O BoutScout mantém apenas dias com exatamente 1.440 registros únicos por minuto, em que o primeiro minuto é 00:00 e o último é 23:59. Dias com minutos ausentes, timestamps duplicados, bordas incompletas ou valores ilegíveis são ignorados.

6. Valores de temperatura

As colunas de temperatura devem conter apenas valores numéricos. Não inclua símbolos de grau, unidades, rótulos de texto ou vírgulas como separador decimal.

7. Pré-processamento recomendado

Ordene o arquivo por data e hora, confirme o intervalo de amostragem de um minuto, revise timestamps duplicados, mantenha as colunas necessárias nas quatro primeiras posições e teste primeiro com um arquivo pequeno.

8. Saídas após inferência e correção manual

O BoutScout exporta três arquivos CSV locais: previsões minuto a minuto, resumos de eventos contínuos e resumos diários com atenção ao ninho, durações de bouts, contagem de eventos e resumos de temperatura.

Exemplo mínimo válido

```
record_id,datetime,egg_temperature,ambient_temperature
1,2024-06-15 00:00:00,31.82,21.75
2,2024-06-15 00:01:00,31.80,21.76
3,2024-06-15 00:02:00,31.79,21.77
...
1440,2024-06-15 23:59:00,32.10,22.40
```

Checklist

- O CSV usa vírgulas como separador.
- A primeira linha contém cabeçalhos.
- A coluna 2 contém datetime.
- A coluna 3 contém temperatura do ninho, ovo ou sensor.
- A coluna 4 contém temperatura ambiental quando disponível.
- Cada dia completo tem 1.440 registros de 00:00 a 23:59.
- Não há unidades nem texto misturados nas colunas de temperatura.
- As datas não são ambíguas.

Français (Québec) - Instructions de format CSV

Exigences du fichier d'entrée pour l'inférence locale de l'incubation dans le navigateur.

1. Règle générale

BoutScout attend une ligne par minute. Chaque journée complète doit contenir exactement 1 440 minutes uniques, de 00:00 à 23:59. L'application traite le fichier localement dans votre navigateur, regroupe les données par date civile, retire les minutes dupliquées et lance l'inférence seulement sur les journées complètes.

2. Les journées incomplètes sont exclues

Les journées incomplètes sont exclues automatiquement. BoutScout ne remplit pas les minutes manquantes, n'interpole pas les températures et ne répète pas les valeurs précédentes. Cela évite les sections artificiellement plates au début ou à la fin d'une journée.

3. Structure CSV requise

La version web actuelle lit le CSV selon la position des colonnes. Les noms ci-dessous sont fortement recommandés pour la reproductibilité, mais l'application utilise actuellement l'ordre des colonnes.

Position	Recommended name	Description
Colonne 1	record_id	Recommandée. Numéro de ligne, ID de l'enregistreur ou identifiant du registre. Actuellement ignorée par l'app.
Colonne 2	datetime	Requise. Date et heure pour chaque minute. C'est la colonne la plus importante.
Colonne 3	egg_temperature	Requise. Température du nid, de l'oeuf ou du capteur utilisée comme signal principal pour l'inférence.
Colonne 4	ambient_temperature	Recommandée. Température ambiante. Si elle est absente ou non numérique, l'app utilise egg_temperature comme solution de repli.

4. Format de date et d'heure

Utilisez un horodatage clair dans le format recommandé : YYYY-MM-DD HH:MM:SS. Évitez les dates ambiguës comme 06/15/24 ou 15/06/24, et évitez AM/PM sauf si ce format a déjà été testé localement.

5. Règle de journée complète

BoutScout conserve seulement les journées avec exactement 1 440 enregistrements uniques par minute, où la première minute est 00:00 et la dernière est 23:59. Les journées avec minutes manquantes, horodatages dupliqués, bords incomplets ou valeurs illisibles sont ignorées.

6. Valeurs de température

Les colonnes de température doivent contenir uniquement des valeurs numériques. N'ajoutez pas de symboles de degré, d'unités, d'étiquettes textuelles ou de virgules comme séparateur décimal.

7. Prétraitement recommandé

Triez le fichier par date et heure, confirmez un intervalle d'échantillonnage d'une minute, vérifiez les horodatages dupliqués, gardez les colonnes requises dans les quatre premières positions et commencez avec un petit fichier de test.

8. Sorties après inférence et correction manuelle

BoutScout exporte trois fichiers CSV locaux : prédictions minute par minute, résumés des événements continus et résumés quotidiens avec attention au nid, durées des bouts, dénombrement des événements et résumés de température.

Exemple minimal valide

```
record_id,datetime,egg_temperature,ambient_temperature
1,2024-06-15 00:00:00,31.82,21.75
2,2024-06-15 00:01:00,31.80,21.76
3,2024-06-15 00:02:00,31.79,21.77
...
1440,2024-06-15 23:59:00,32.10,22.40
```

Checklist

Le CSV utilise des virgules comme séparateurs.

La première ligne contient des en-têtes.

La colonne 2 contient datetime.

La colonne 3 contient la température du nid, de l'oeuf ou du capteur.

La colonne 4 contient la température ambiante lorsque disponible.

Chaque journée complète contient 1 440 enregistrements de 00:00 à 23:59.

Aucune unité ni texte n'est mélangé dans les colonnes de température.

Les dates ne sont pas ambiguës.