

```

---
title: "PUC Weekly Bird Report"
author: "BirdWeather / PUC"
output:
  pdf_document:
    toc: false
    number_sections: false
fontsize: 11pt
geometry: margin=1in
---

``` r
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE)

library(readr)
library(dplyr)

```

```

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

```

```

library(lubridate)
library(ggplot2)
library(forcats)

tz_local    <- "Europe/Amsterdam"  # adjust if needed

no_data <- FALSE
reason  <- NULL

# --- Robust path lookup for master_detections.csv ---

# We try both from repo root and from scripts/ (one level below)
candidate_paths <- c(
  "data/master_detections.csv",
  "../data/master_detections.csv"
)

path_master <- NULL
for (p in candidate_paths) {
  if (file.exists(p) && file.info(p)$size > 0) {
    path_master <- p
    break
  }
}

```

```

}

if (is.null(path_master)) {
  no_data <- TRUE
  reason <- paste(
    "Master file missing or empty. Tried paths:",
    paste(candidate_paths, collapse = ", ")
  )
} else {
  df_raw <- read_csv(path_master, show_col_types = FALSE)

  # auto-detect species + time columns
  if (!all(c("timestamp", "species.commonName") %in% names(df_raw))) {
    no_data <- TRUE
    reason <- paste(
      "Expected columns 'timestamp' and 'species.commonName' not found. Columns present:",
      paste(names(df_raw), collapse = ", ")
    )
  } else {
    df <- df_raw %>%
      mutate(
        species = .data[["species.commonName"]],
        ts_utc = ymd_hms(.data[["timestamp"]], quiet = TRUE),
        ts_local = with_tz(ts_utc, tz_local),
        date = as_date(ts_local),
        hour = hour(ts_local)
      ) %>%
      filter(!is.na(species), !is.na(ts_local))

    if (nrow(df) == 0) {
      no_data <- TRUE
      reason <- "No valid detections after parsing timestamps and species."
    } else {
      # define 7-day window based on latest date in data
      last_date <- max(df$date, na.rm = TRUE)
      start_date <- last_date - days(6)
      end_date <- last_date

      df7 <- df %>% filter(date >= start_date & date <= end_date)

      if (nrow(df7) == 0) {
        no_data <- TRUE
        reason <- paste(
          "No detections in the last 7 days of data (",
          as.character(start_date), "-", as.character(end_date), ")."
        )
      } else {
        # global weekly metrics
        total_detections <- nrow(df7)
        unique_species <- n_distinct(df7$species)

        most_common <- df7 %>%
          count(species, sort = TRUE) %>%

```

```

slice(1)

# confidence if present
conf_col <- intersect(c("confidence", "score", "probability"), names(df7))[1]
avg_conf <- if (!is.na(conf_col)) {
  mean(df7[[conf_col]], na.rm = TRUE)
} else NA_real_

# top 10 species
top10 <- df7 %>%
  count(species, sort = TRUE) %>%
  slice_head(n = 10)

# BirdNET-style palette
birdnet_colors <- c(
  "#4e79a7", "#59a14f", "#f28e2c", "#e15759", "#76b7b2",
  "#edc949", "#af7aa1", "#ff9da7", "#9c755f", "#bab0ab"
)

# order species for plotting
top10 <- top10 %>%
  arrange(n) %>%
  mutate(species = factor(species, levels = species))

# time-of-day categories
df7_periods <- df7 %>%
  mutate(period = case_when(
    hour >= 0 & hour <= 4 ~ "Night (0-4)",
    hour >= 5 & hour <= 8 ~ "Dawn (5-8)",
    hour >= 9 & hour <= 11 ~ "Morning (9-11)",
    hour >= 12 & hour <= 16 ~ "Afternoon (12-16)",
    hour >= 17 & hour <= 19 ~ "Evening (17-19)",
    hour >= 20 & hour <= 23 ~ "Night (20-23)"
  ))

tod <- df7_periods %>%
  count(period) %>%
  mutate(period = factor(period, levels = c(
    "Night (0-4)", "Dawn (5-8)", "Morning (9-11)",
    "Afternoon (12-16)", "Evening (17-19)", "Night (20-23)"
  )))

# daily detection trends
daily_trend <- df7 %>%
  count(date) %>%
  arrange(date)

# new species in this 7-day window
first_all <- df %>%
  group_by(species) %>%
  summarise(first_date = min(date), .groups = "drop")

new_this_week <- first_all %>%

```

```

    filter(first_date >= start_date & first_date <= end_date) %>%
    left_join(
      df7 %>% count(species, name = "detections_this_week"),
      by = "species"
    ) %>%
    arrange(first_date)
  }
}

}
}
if (no_data) {
cat("No data available for this weekly report.\n\nReason:\n", reason)
} else {
summary_tbl <- tibble::tibble(
Metric = c(
  "Total detections",
  "Unique species",
  "Most common species",
  "Detections of most common",
  "Average confidence"
),
Value = c(
  total_detections,
  unique_species,
  most_common$species,
  most_common$n,
  if (!is.na(avg_conf)) sprintf("%.1f%%", avg_conf * 100) else "n/a"
)
)

cat(
  "Period: ",
  as.character(start_date), " - ", as.character(end_date), "\n\n",
  sep = ""
)

knitr::kable(summary_tbl, align = c("l","r"))
}

```

## Period: 2026-02-13 - 2026-02-19

Metric	Value
Total detections	3177
Unique species	63
Most common species	Eurasian Blue Tit
Detections of most common	854
Average confidence	82.9%

```

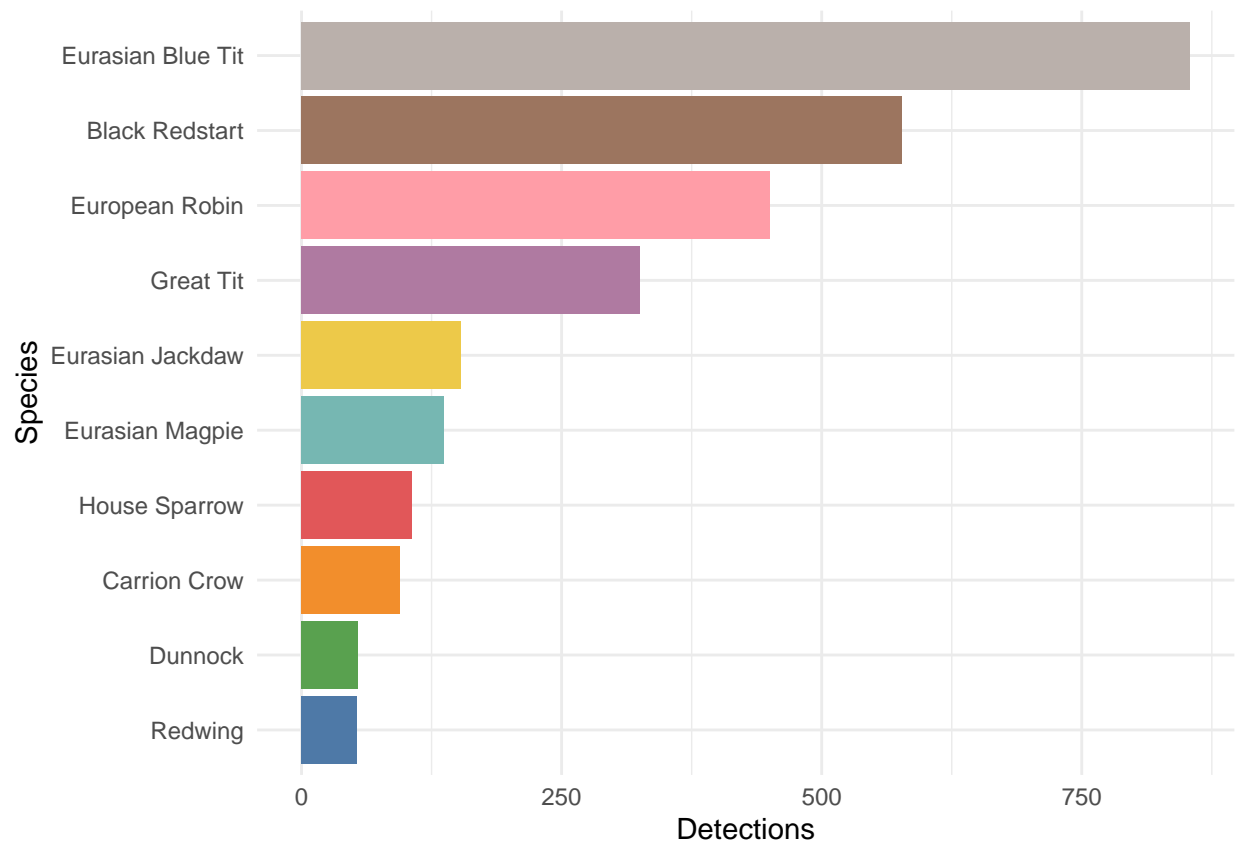
if (!no_data) {
ggplot(top10, aes(x = species, y = n, fill = species)) +

```

```

geom_col() +
scale_fill_manual(values = birdnet_colors) +
coord_flip() +
labs(
  x = "Species",
  y = "Detections"
) +
theme_minimal(base_size = 11) +
theme(legend.position = "none")
}

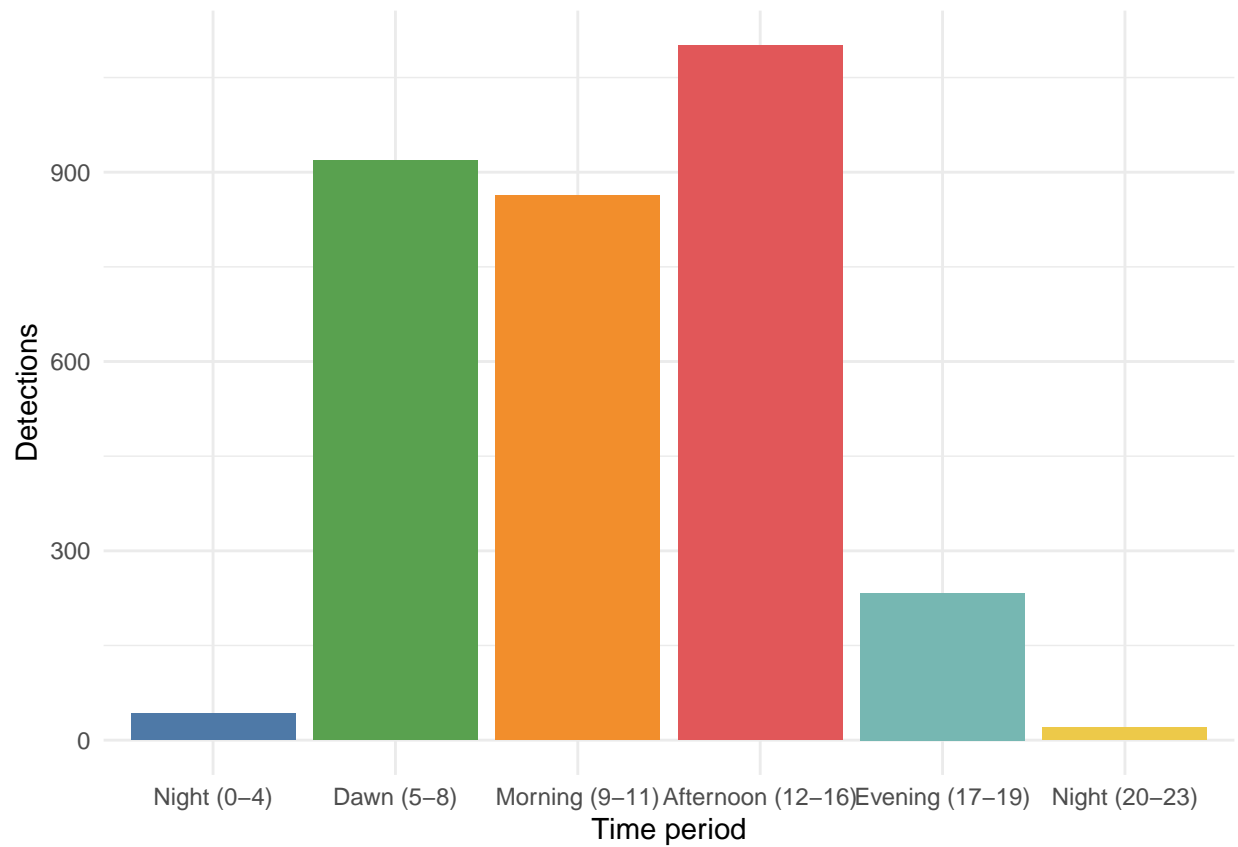
```



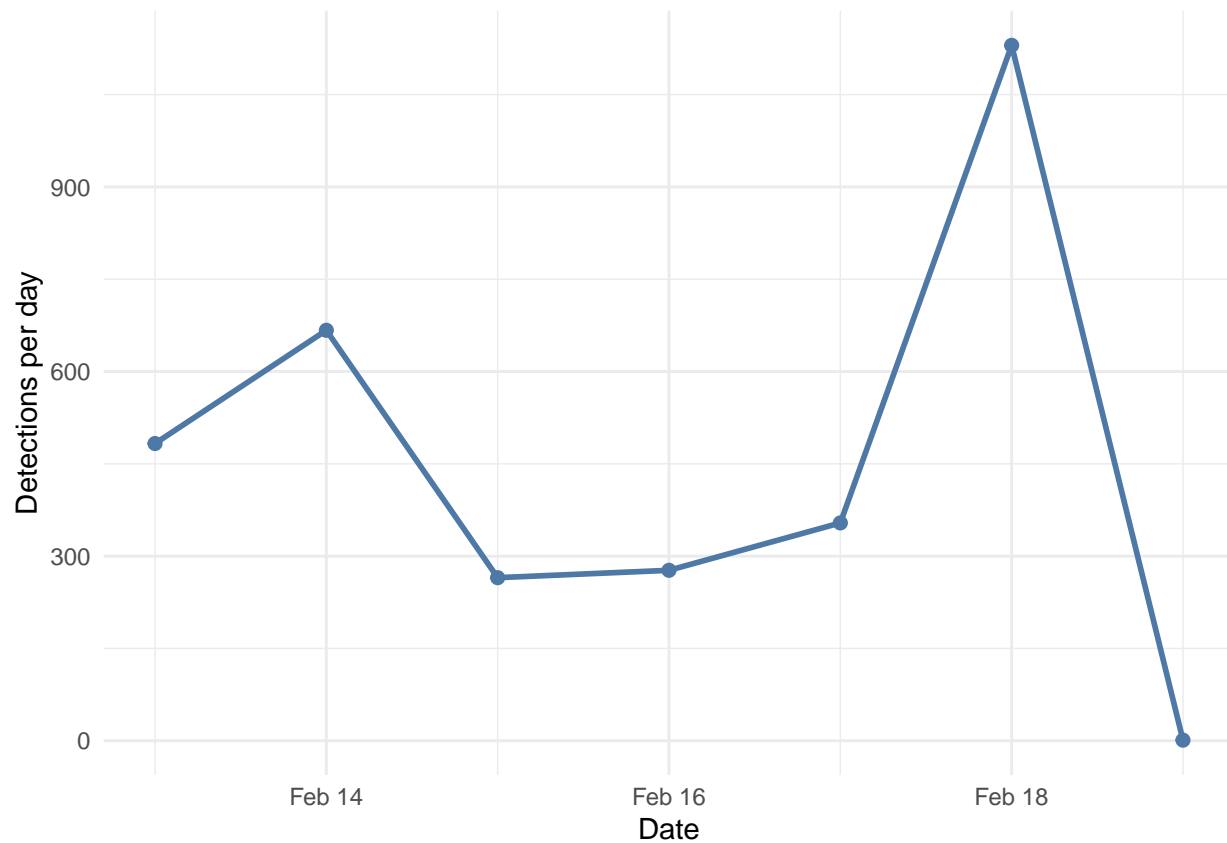
```

if (!no_data) {
  ggplot(tod, aes(x = period, y = n, fill = period)) +
  geom_col() +
  scale_fill_manual(values = birdnet_colors) +
  labs(
    x = "Time period",
    y = "Detections"
  ) +
  theme_minimal(base_size = 11) +
  theme(legend.position = "none")
}

```



```
if (!no_data) {  
  ggplot(daily_trend, aes(x = date, y = n)) +  
  geom_line(color = "#4e79a7", linewidth = 1) +  
  geom_point(color = "#4e79a7", size = 2) +  
  labs(  
    x = "Date",  
    y = "Detections per day"  
  ) +  
  theme_minimal(base_size = 11)  
}
```



```

if (no_data) {
  cat("No data available.")
} else {
  if (nrow(new_this_week) == 0) {
    cat("No new species detected in this 7-day period.")
  } else {
    knitr::kable(
      new_this_week %>%
      select(species, first_date, detections_this_week),
      col.names = c("Species", "First detection", "Detections this week"),
      align = c("l", "c", "r")
    )
  }
}

```

Species	First detection	Detections this week
Common Redshank	2026-02-14	1
Egyptian Goose	2026-02-18	1