```
# ============================================================
# Customer-hours from 15-minute outage records (2014–2024)
# ============================================================
suppressPackageStartupMessages({
  library(data.table)
})

# ---- Inputs / Outputs ----
infile        <- "eaglei_outages_2014_2024_miso_filtered.csv"
outfile_rows   <- "eaglei_outages_2014_2024_miso_filtered_with_customer_hours.csv"
outfile_county  <- "county_customer_hours_totals_2014_2024.csv"

# ---- Read ----
stopifnot(file.exists(infile))
DT <- fread(infile, na.strings = c("", "NA", "NaN"))

# ---- Clean stray column, ensure types ----
if ("X" %in% names(DT)) DT[, X := NULL]               # drop stray column if present
need <- c("fips_code","county","state","customers_out","run_start_time")
miss <- setdiff(need, names(DT))
if (length(miss)) stop("Missing required columns: ", paste(miss, collapse = ", "))

# customers_out numeric; keep NAs but ignore them in sums
DT[, customers_out := as.numeric(customers_out)]

# ---- Add customer-hours (15-min interval = 0.25 hours) ----
DT[, customer_hours := customers_out * 0.25]            # NAs propagate; fine

# (Optional) If you want to constrain to 2014–2024 strictly, uncomment this:
# DT[, ts := as.POSIXct(gsub("_", "-", run_start_time), format = "%Y-%m-%dT%H:%M:%SZ",
tz = "UTC")]
# DT <- DT[ts >= as.POSIXct("2014-01-01 00:00:00", "UTC") & ts <= as.POSIXct("2024-12-31
23:59:59", "UTC")]

# ---- Write the row-level file ----
fwrite(DT, outfile_rows)

# ---- County totals across the full period ----
county_totals <- DT[
  ,
  .(total_customer_hours_2014_2024 = sum(customer_hours, na.rm = TRUE)),
  by = .(fips_code, county, state)
]
```

```
# Nicely ordered (largest exposure first)
setorder(county_totals, -total_customer_hours_2014_2024, state, county)

# ---- Write the county totals file ----
fwrite(county_totals, outfile_county)

# ---- Console summary ----
cat("Wrote row-level file with customer_hours to:", outfile_rows, "\n")
cat("Wrote county totals to:", outfile_county, "\n")
cat("Counties in totals:", nrow(county_totals), "\n")
```