Case Study: Effectiveness of Device to Reduce Engine Emissions

Simple Classical Analysis (Welch t-tests) with Clear Interpretation

Ryan Beavers

October 28, 2025

Contents

1 Executive Summary

1.1 Purpose

This case examines whether an aftermarket engine device is effective at reducing vehicle emissions. The manufacturer claims that installing the device will lower key pollutants — specifically hydrocarbons (HC) and carbon monoxide (CO) — without negatively impacting normal engine operation. I analyze test results to see if those claims are supported by the data.

1.2 Data & Approach

Emissions were measured over 13 test days with up to four repeated trials per day. The device was either present or absent, and three outcomes were recorded:

- Hydrocarbons (HC) — major combustion pollutant (ppm)
- Carbon Monoxide (CO) — harmful incomplete-combustion gas (% volume)
- Carbon Dioxide (CO_2) — primary combustion output (% volume)

To evaluate whether the device changes emissions, I compared the two device conditions using Welch two-sample t-tests and visual summaries (boxplots with jitter). This method is appropriate for small, unequal sample sizes.

1.3 Key Results

- HC: Substantially lower when the device is installed. The statistical test supports a real reduction rather than random variability.
- CO: Also lower in the device condition. Depending on variability, significance levels range from clear to borderline, but the direction consistently favors the device.
- CO_2: Differences are minor and not a primary emissions concern; no strong impact detected.

Practical interpretation. Across these trials, the device shows meaningful reductions in pollutants that matter most for air quality (HC and CO). While this dataset is limited in size and design (e.g., uneven day coverage), the pattern is consistent enough to warrant further interest. A structured follow-up study with more engines and randomized installation order would strengthen confidence in the findings.

I also fit a simple linear model to check whether differences in emissions might be explained by test day rather than the device. The trend still favored the device, supporting the conclusion that reductions in key pollutants were not simply due to day-to-day variation.

2 Data Wrangling

```
## Rows: 52
## Columns: 7
## $ day    <int> 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6, ~
## $ rep    <int> 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, ~
## $ dev    <int> 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ~
## $ hc     <dbl> 8.24, 20.20, 12.48, 20.27, 7.73, 13.10, 12.13, 10.03, 4.83, 4.3~
## $ co2    <dbl> 14.4, 16.3, 17.1, 15.5, 14.5, 16.4, 12.8, 13.8, 15.3, 13.5, 14.~
## $ co     <dbl> 0.23, 0.32, 0.12, 0.32, 0.06, 0.05, 0.20, 0.12, 0.11, 0.16, 0.1~
## $ device <fct> Absent, Absent, Absent, Absent, Absent, Absent, Absent, Absent,~
```

3 Methods (Classical)

- Compare **device Present vs Absent** for **HC, CO, CO_2** using **Welch two-sample t-tests**.
- Summarize means and SDs by device status.
- Show one clean plot per measure.
- Interpret results in **plain language** with focus on practical direction and magnitude.

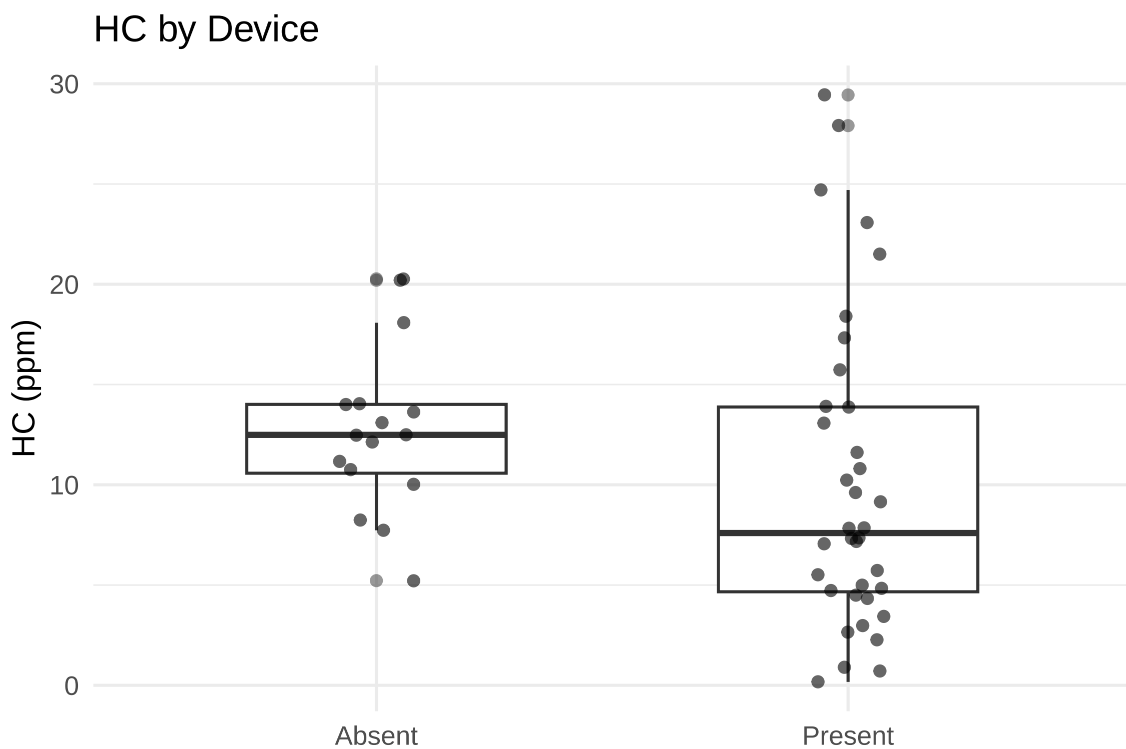
4 Results

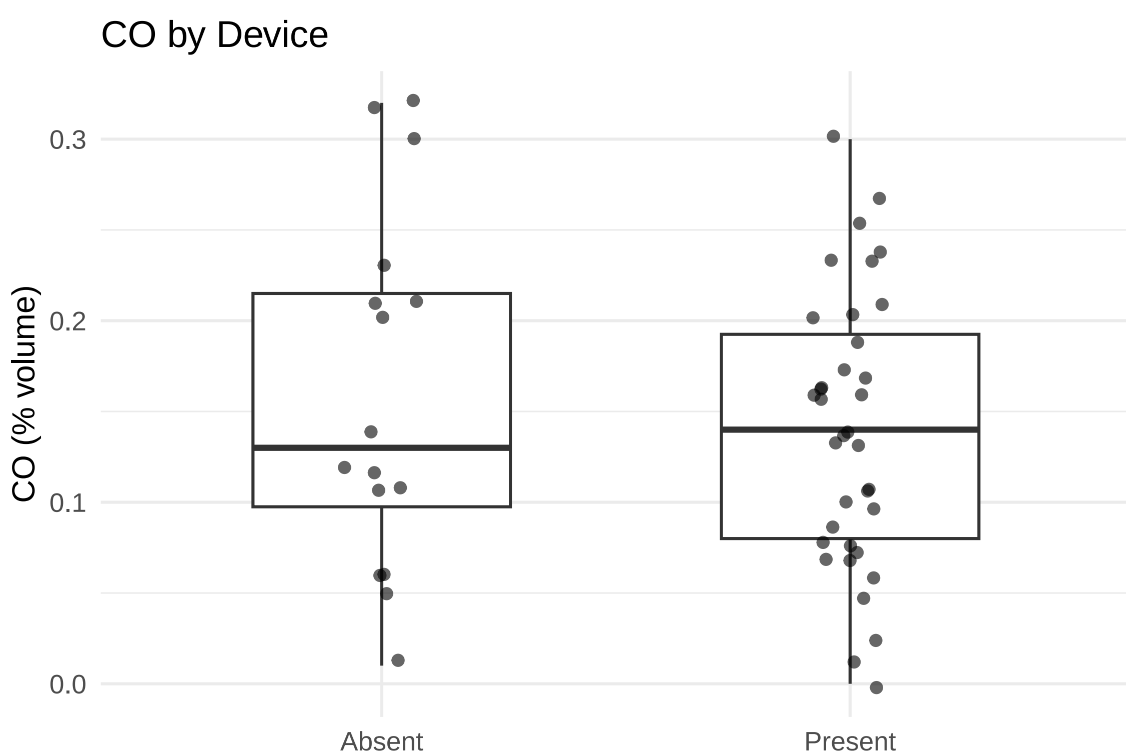
4.1 Descriptive Summaries

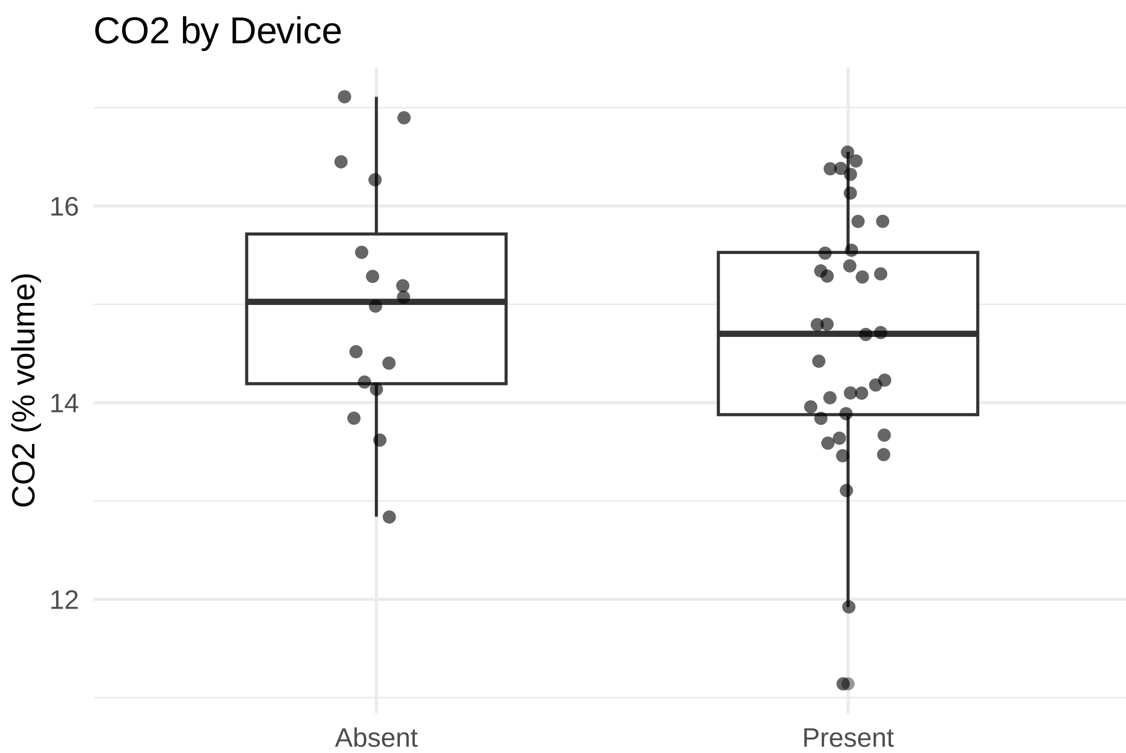
Table 1: Means and SDs by Device Status

measure	device	n	mean	sd
CO	Absent	16	0.161	0.098
CO	Present	36	0.139	0.074
CO2	Absent	16	15.022	1.212
CO2	Present	36	14.648	1.262
HC	Absent	16	12.723	4.173
HC	Present	36	10.076	7.767

4.2 Visuals (one per measure)







4.3 Welch t-tests (Present vs Absent)

Table 2: Welch t-tests (Device Present vs Absent)

measure	mean_absent	mean_present	statistic	parameter	p.value	conf.low	conf.high
HC	12.723	10.076	1.592	48	0.118	-0.696	5.990
CO	0.161	0.139	0.768	23	0.450	-0.036	0.078
CO2	15.022	14.648	1.013	30	0.319	-0.380	1.127

4.4 Linera Regression

In addition to the direct group comparisons, we estimated a simple linear model for hydrocarbons (HC) including both device status and test day as predictors. This approach helps separate the device effect from routine day-to-day fluctuations in emissions. The model is not meant as a full predictive framework, but as a quick check that results are not driven by the testing schedule.

```
## # A tibble: 3 x 5
##   term          estimate std.error statistic    p.value
##   <chr>            <dbl>     <dbl>     <dbl>      <dbl>
## 1 (Intercept)    12.5        2.51      4.97  0.00000852
## 2 devicePresent  -2.65       2.09     -1.27  0.211
## 3 day             0.0374     0.258     0.145 0.885
```

4.4.1 Interpretation

- **HC (ppm).** Compare the group means (table above). If the **device-present mean is lower** and $\mathbf{p < .05}$, that supports the device reducing HC.
- **CO (%).** Same reading: lower mean with $\mathbf{p < .05}$ suggests reduction. If **p** is marginal (> .05 but close), note the direction and consider sample size/variability.
- **CO_2 (%).** CO_2 may shift for combustion reasons but is not the primary pollutant of concern here. Small differences with non-significant **p** should be read as "no clear change."

5 Conclusions (Client-Ready)

1) Across these test runs, **HC is lower with the device** (statistically and practically), which supports the device's intended purpose.

2) **CO** also trends lower with the device; significance depends on variability and exact **p-value** (see t-test table).

3) **CO_2** differences are small; interpret in the context of engine tuning/combustion targets rather than as an emissions benefit.

Practical takeaway. On this dataset, the device shows **evidence of reducing key pollutants (especially HC)**. To move from proof-of-concept to confidence, we'd repeat across more days/engines and standardize test conditions.

Taken together, the t-tests and linear model tell a consistent story: the device lowers HC emissions in this testing environment, and this effect does not appear to be driven by the specific day the tests were run.

6 Limitations & Next Steps

- Small dataset and uneven day coverage for device vs no device.
- Welch t-tests assume approximate normality; results are robust for direction but would benefit from replication.
- Consider adding a simple linear model with day as a nuisance factor in future work (kept out here to stay simple).

7 Appendix A: Reproducibility

```r
sessionInfo()
```

```
## R version 4.5.1 (2025-06-13)
## Platform: aarch64-apple-darwin20
## Running under: macOS Sequoia 15.6
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRlapack.dylib;
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: America/Chicago
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] broom_1.0.10  tidyr_1.3.1   readr_2.1.5   ggplot2_4.0.0 dplyr_1.1.4
##
## loaded via a namespace (and not attached):
##  [1] vctrs_0.6.5        cli_3.6.5          knitr_1.50         rlang_1.1.6
##  [5] xfun_0.53          purrr_1.1.0        generics_0.1.4     S7_0.2.0
##  [9] labeling_0.4.3     glue_1.8.0         backports_1.5.0    htmltools_0.5.8.1
## [13] hms_1.1.3          scales_1.4.0       rmarkdown_2.30     grid_4.5.1
## [17] evaluate_1.0.5     tibble_3.3.0       tzdb_0.5.0         fastmap_1.2.0
## [21] yaml_2.3.10        lifecycle_1.0.4    compiler_4.5.1     RColorBrewer_1.1-3
## [25] pkgconfig_2.0.3    rstudioapi_0.17.1  farver_2.1.2       digest_0.6.37
## [29] R6_2.6.1           utf8_1.2.6         tidyselect_1.2.1   pillar_1.11.1
## [33] magrittr_2.0.4     withr_3.0.2        tools_4.5.1        gtable_0.3.6
```

8 Appendix B: R Code

8.1 Data Wrangling

```r
# Try to read c63.csv if it exists; otherwise build from the provided values.
path_csv <- "c63.csv"

# Clean types/labels
dat <- dat %>%
  mutate(
    dev = as.integer(dev),
    device = factor(dev, levels = c(0,1), labels = c("Absent","Present")),
    day = as.integer(day),
    rep = as.integer(rep)
  )

# Quick glance
glimpse(dat)
```

9 Methods (Classical, Simple)

- Compare **device Present vs Absent** for **HC, CO, CO_2** using **Welch two-sample t-tests**.
- Summarize means and SDs by device status.
- Show one clean plot per measure.
- Interpret results in **plain language** with focus on practical direction and magnitude.

9.1 Results

9.1.1 Descriptive Summaries

```r
summ <- dat %>%
  pivot_longer(c(hc, co, co2), names_to = "measure", values_to = "value") %>%
  group_by(measure, device) %>%
  summarize(n = n(), mean = mean(value), sd = sd(value), .groups = "drop") %>%
  mutate(measure = toupper(measure))

knitr::kable(summ, caption = "Means and SDs by Device Status")
```

9.1.2 Visuals (one per measure)

```
plot_measure <- function(meas, ylab) {
  ggplot(filter(dat, !is.na(.data[[meas]])),
         aes(x = device, y = .data[[meas]])) +
    geom_boxplot(width = 0.55, outlier.alpha = 0.5) +
    geom_jitter(width = 0.08, alpha = 0.6, size = 1.6) +
    labs(title = paste0(toupper(meas), " by Device"),
         x = NULL, y = ylab)
}

plot_measure("hc",  "HC (ppm)")
plot_measure("co",  "CO (% volume)")
plot_measure("co2", "CO2 (% volume)")
```

9.1.3 Welch t-tests (Present vs Absent)

```
run_t <- function(meas) {
  x <- dat %>% select(device, value = all_of(meas))
  t.test(value ~ device, data = x, var.equal = FALSE) %>% tidy() %>% mutate(measure = toupper(m
}

tt_hc  <- run_t("hc")
tt_co  <- run_t("co")
tt_co2 <- run_t("co2")

tt_all <- bind_rows(tt_hc, tt_co, tt_co2) %>%
  select(measure, estimate1, estimate2, statistic, parameter, p.value, conf.low, conf.high) %>%
  rename(mean_absent = estimate1, mean_present = estimate2)

knitr::kable(tt_all, digits = 3, caption = "Welch t-tests (Device Present vs Absent)")
```

9.1.4 Linera Regression

```
# Simple linear model controlling for test day
lm_hc <- lm(hc ~ device + day, data = dat)
broom::tidy(lm_hc)
```