

1 Experiment Set Difference: Transformer (25-50 words) v. Conv (25-50 words)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

cpu_hours	0.0275006 +/- 0.00929442 (stderr) (p=0.00513853, Welch's t-test)
gpu_hours	0.0625617 +/- 0.00191161 (stderr) (p=1.92615e-07, Welch's t-test)
estimated_carbon_impact_kg	0.00925278 +/- 0.00208604 (stderr) (p=0.00096178, Welch's t-test)
total_power	0.0243416 +/- 0.00659155 (stderr) (p=0.00231153, Welch's t-test)
kw_hr_gpu	0.00861176 +/- 0.000590999 (stderr) (p=1.76071e-07, Welch's t-test)
kw_hr_cpu	0.0067943 +/- 0.00426551 (stderr) (p=0.0487853, Welch's t-test)
exp_len_hours	0.0410923 +/- 0.00923026 (stderr) (p=0.00100031, Welch's t-test)
average_carbon_intensity	8.92749 +/- 8.68533 (stderr) (p=0.126273, Welch's t-test)

1.2 Cumulative Difference (Per Experiment Set)

cpu_hours	0.1375027777777775
gpu_hours	0.31280831280139854
estimated_carbon_impact_kg	0.04626388385575243
total_power	0.12170788432447954
kw_hr_gpu	0.04305878789640666
kw_hr_cpu	0.03397151863807409
exp_len_hours	0.2054614167080986
average_carbon_intensity	44.637450771513386