1 Experiment Set Difference: Normal Convolutions (32 filters) v. Normal Convolutions (64 filters)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

cpu_hours	-0.08195 +/- 0.000650501 (stderr) (p=2.63155e-15, Welch's t-test)
gpu_hours	-0.0712007 +/- 0.000650163 (stderr) (p=5.35095e-14, Welch's t-test)
estimated_carbon_impact_kg	-0.00394069 +/- 0.000266311 (stderr) (p=9.98248e-07, Welch's t-test)
total_power	-0.0226331 +/- 0.00244051 (stderr) (p=3.34334e-05, Welch's t-test)
kw_hr_gpu	-0.0119613 +/- 0.000494125 (stderr) (p=6.73188e-07, Welch's t-test)
kw_hr_cpu	-0.00236347 +/- 0.00160989 (stderr) (p=0.0548159, Welch's t-test)
exp_len_hours	-0.0757918 +/- 0.000678385 (stderr) (p=1.50027e-13, Welch's t-test)
average_realtime_carbon_intensity	-10.1543 +/- 4.44844 (stderr) (p=0.0123202, Welch's t-test)

1.2 Cumulative Difference (Per Experiment Set)

cpu_hours	-0.40974999999999995
gpu_hours	-0.35600329246557405
estimated_carbon_impact_kg	-0.01970346059815376
total_power	-0.1131653433632323
kw_hr_gpu	-0.05980626027284994
kw_hr_cpu	-0.011817374767170517
exp_len_hours	-0.3789589740170374
average_realtime_carbon_intensity	-50.77156196917599