1 Experiment Set Difference: Normal Convolutions (64 filters) v. Normal Convolutions (256 filters)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

cpu_hours	-0.760183 +/- 0.0921501 (stderr) (p=0.000788474, Welch's t-test)
gpu_hours	-0.689364 +/- 0.082923 (stderr) (p=0.000770558, Welch's t-test)
estimated_carbon_impact_kg	-0.0716765 +/- 0.00832957 (stderr) (p=0.000689732, Welch's t-test)
total_power	-0.213745 +/- 0.0371116 (stderr) (p=0.002202, Welch's t-test)
kw_hr_gpu	-0.115134 +/- 0.0138962 (stderr) (p=0.0007782, Welch's t-test)
kw_hr_cpu	-0.020148 +/- 0.0130289 (stderr) (p=0.0764214, Welch's t-test)
exp_len_hours	-0.699319 +/- 0.0846314 (stderr) (p=0.000784483, Welch's t-test)
average_realtime_carbon_intensity	-146.398 +/- 39.076 (stderr) (p=0.00733868, Welch's t-test)

1.2 Cumulative Difference (Per Experiment Set)

cpu_hours	-2.8852805555555556
gpu_hours	-2.624050278309054
estimated_carbon_impact_kg	-0.27999057343476763
total_power	-0.8146173383255365
kw_hr_gpu	-0.43904768682633977
kw_hr_cpu	-0.07653290705058205
exp_len_hours	-2.657002399696244
average_realtime_carbon_intensity	-418.94276948502807