1 Experiment Set Difference: Normal Convolutions (64 filters) v. Normal Convolutions (128 filters)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

cpu_hours	-0.769423 +/- 0.0106789 (stderr) (p=2.0155e-11, Welch's t-test)
gpu_hours	-0.677274 +/- 0.00499158 (stderr) (p=5.2642e-13, Welch's t-test)
estimated_carbon_impact_kg	-0.0987801 +/- 0.0051037 (stderr) (p=2.49044e-06, Welch's t-test)
total_power	-0.279981 +/- 0.00973518 (stderr) (p=1.87399e-05, Welch's t-test)
kw_hr_gpu	-0.166938 +/- 0.00547049 (stderr) (p=1.54887e-05, Welch's t-test)
kw_hr_cpu	-0.0102657 +/- 0.00116583 (stderr) (p=0.000590596, Welch's t-test)
exp_len_hours	-0.684917 +/- 0.00668176 (stderr) (p=1.08965e-09, Welch's t-test)
average_realtime_carbon_intensity	-36.663 +/- 13.1383 (stderr) (p=0.0119136, Welch's t-test)

1.2 Cumulative Difference (Per Experiment Set)

cpu_hours	-3.0776916666666665
gpu_hours	-2.7090950698658656
estimated_carbon_impact_kg	-0.39512039853688175
total_power	-1.1199254062635144
kw_hr_gpu	-0.6677506792904324
kw_hr_cpu	-0.04106286897761484
exp_len_hours	-2.7396674205197225
average_realtime_carbon_intensity	-146.65205422827216