This document provides the calculations for the results in the week 1 notes.

Background information: sd = 14.8 minutes/day

Precision approach:
Goal: Want se of a difference = 2 minutes/day
Need the formula for the se of a difference, using pooled sd: $goal = s\sqrt{2/n}$
Need to find n for which $2 = 14.8\sqrt{2/n}$
Algebra on the general equation: $n = 2\,(s/goal)^2$,
For this problem: $n = 2\,(14.8/2)^2 = 109.5$
Round up: $n = 110$

Confidence interval width approach:
Goal: Want 95% ci width for a difference = 3 minutes/day
Note: Not the problem in the notes - that was ci width for a mean
And I've changed the goal to 3 to clarify which value is which
The mean problem in the notes has exactly the same principles, changes are:
a different se formula: $se = s\sqrt{2/n}$
and a different df, since we will pool: $df = 2(n-1)$
where n = # obs per group
width = $2\, t_{1-\alpha/2,\,df}\; se$
95% confidence interval, so $\alpha = 0.05$, $1 - \alpha/2 = 0.975$
Need to find n such that $2\, t_{0.975,\,2n-2}\; s\sqrt{2/n} = 3$
As stated in lecture, n enters twice: t value (thru the df) and the se

Two ways to find n:
1) Try multiple values of n,
compute $2\, t_{0.975,\,2n-2}\; s\sqrt{2/n}$ for each n
see where drops below the goal, 3 here.
2) Iterative solution to a non-linear equation
Pick a starting n, compute df and find the T quantile
e.g. $n = 31$, $df = 2(31-1) = 60$, $T_{0.975,\,60} = 2.000$.
Now: need to solve: $2 \times 2.000 \times s\sqrt{2/n} = 3$,
That is $n = 2\,(2 \times 2.000 \times 14.8/3)^2 = 778.8$
$n = 779$ per group gives 2*(779-1) = 1556 df.
$t_{0.975,\,1556} = 1.961$
Now, need to solve: $2 \times 1.961 \times s\sqrt{2/n} = 3$,
That is $n = 2\,(2 \times 1.961 \times 14.8/3)^2 = 748.7$
$n = 749$ per group gives 2*(749-1) = 1496 df.
$t_{0.975,\,1496} = 1.961$
Same quantile, so that's the answer: 749 per group

Power approach:
Goal: 80% power to detect difference of 3
Could plug various values of n into the fundamental equation:

$$\delta = \left(t_{1-\alpha/2,\, df} + t_{power,\, df}\right) s\sqrt{2/n}$$

to see what is the smallest n that gives $\delta \le 3$

Or iterate
Worked example in week 2 notes, slightly different goal used here
Solving the fundamental equation for n gives:

$$n = 2\left(t_{1-\alpha/2,\, df} + t_{1-\beta,\, df}\right)^2 (s/\delta)^2$$

Problem specifies: $\alpha = 0.05$, $\beta = 0.80$, $s = 14.8$ (WWE data), $\delta = 3$
Start with $n = 31$ per group, for which df = 60
$t_{0.975,60} = 2.000$, $t_{0.8,60} = 0.847$, $s/\delta = 4.93$, $n = 2 \times 2.847^2 \times 4.93^2 = 394$ per group
New df = 2(394-1) = 786, so:
$t_{0.975,786} = 1.963$, $t_{0.8,786} = 0.842$, $n = 2 \times 2.805^2 \times 4.93^2 = 382$ per group
Similar to previous sample size (both are a bit less than 400).
If you continue one more time:
New df = 2(382-1) = 762, so:
$t_{0.975,762} = 1.963$, $t_{0.8,762} = 0.842$,
Quantiles are the same (to 3 decimal places), so n will be the same
If you carry more digits through the calculations, you get $n = 383.02$
Software (power.t.test(), proc power) using a non-central T distribution gives $n = 383.01$