

SAMPLE RStudio lab assignment: t-test

- 1) Run t-test analysis for “**weekswrk**” (outcome variable) and “**wrkslf**” (factor variable). Paste the table (10 points) and interpret the results in light of mean scores and statistical significance (10 points)

Code (-5):	t.test(weekswrk ~ wrkslf, data = gss) > parameters() > display(format="html")														
Table:	<table><tr><th>Outcome variable</th><th>wrkslf = Self employed</th><th>wrkslf = Working for someone else</th><th>Difference</th><th>95% CI</th><th>t(547.56)</th><th>p</th></tr><tr><td>Weeks respondents worked last year</td><td>33.58</td><td>30.69</td><td>2.88</td><td>(0.67, 5.10)</td><td>2.56</td><td>0.011*</td></tr></table>	Outcome variable	wrkslf = Self employed	wrkslf = Working for someone else	Difference	95% CI	t(547.56)	p	Weeks respondents worked last year	33.58	30.69	2.88	(0.67, 5.10)	2.56	0.011*
Outcome variable	wrkslf = Self employed	wrkslf = Working for someone else	Difference	95% CI	t(547.56)	p									
Weeks respondents worked last year	33.58	30.69	2.88	(0.67, 5.10)	2.56	0.011*									
Interpretation:	The average number of weeks respondents worked last year for those who are self employed is 33.58 weeks, the average number of weeks respondents worked last year for those who work for someone else is 30.69 weeks. Weeks respondents worked last year differs by respondents' employment status in a statistically significant way since the p-value is less than 0.05.														

- 2) Run t-test analysis for “**hrs2**” (outcome variable) and “**postlife**” (factor variable). Paste the table (10 points) and interpret the results in light of mean scores and statistical significance (10 points).

Code (-5):	<pre>t.test(hrs2 ~ postlife, data = gss) > parameters() > display(format="html")</pre>														
Table:	<table><tr><th>Outcome variable</th><th>born = Yes</th><th>born = No</th><th>Difference</th><th>95% CI</th><th>t(853.05)</th><th>p</th></tr><tr><td>Perceived personal health quality</td><td>2.17</td><td>2.14</td><td>0.02</td><td>(-0.04, 0.09)</td><td>0.76</td><td>0.446</td></tr></table>	Outcome variable	born = Yes	born = No	Difference	95% CI	t(853.05)	p	Perceived personal health quality	2.17	2.14	0.02	(-0.04, 0.09)	0.76	0.446
Outcome variable	born = Yes	born = No	Difference	95% CI	t(853.05)	p									
Perceived personal health quality	2.17	2.14	0.02	(-0.04, 0.09)	0.76	0.446									
Interpretation:	The average number of hours respondents usually work for those who believe in life after death is 42.93 hours; the average number of hours respondents usually work for those who do not believe in life after death is 30.50 hours. The number of hours respondents usually work does not differ by belief in life after death in a statistically significant way since the p-value is higher than 0.05.														

3) First, recode “educ” into two categories. The new variable name will be “**educbinary**”. (20 points)

- o 0 to 14 will be 1 [low education level]
- o 15 to 20 will be 2 [high education level]

Code (-5):	<pre>gss\$educbinary <- rec(gss\$educ, rec = "0:14=1 [Low education level]; 15:20=2 [High education level]", var.label = "Recoded respondents' education in years")</pre>
-------------------	--

4) Second, create a frequency table for “**educbinary**”. Paste (10 points for table) and interpret the table (10 points for interpretation).

Code (-5):	frq(gss\$educbinary, out = "v")																																				
Table:	<table><tr><th colspan="6">Recoded respondents' education in years (Variable label)</th></tr><tr><th>value</th><th>value label</th><th>frq</th><th>raw.prc</th><th>valid.prc</th><th>cum.prc</th></tr><tr><td>1</td><td>Low education level</td><td>2176</td><td>54.59</td><td>55.06</td><td>55.06</td></tr><tr><td>2</td><td>High education level</td><td>1776</td><td>44.56</td><td>44.94</td><td>100.00</td></tr><tr><td>NA</td><td>NA</td><td>34</td><td>0.85</td><td>NA</td><td>NA</td></tr><tr><td colspan="6">total N=3986 · valid N=3952 · $\bar{x}=1.45$ · $\sigma=0.50$</td></tr></table>	Recoded respondents' education in years (Variable label)						value	value label	frq	raw.prc	valid.prc	cum.prc	1	Low education level	2176	54.59	55.06	55.06	2	High education level	1776	44.56	44.94	100.00	NA	NA	34	0.85	NA	NA	total N=3986 · valid N=3952 · $\bar{x}=1.45$ · $\sigma=0.50$					
Recoded respondents' education in years (Variable label)																																					
value	value label	frq	raw.prc	valid.prc	cum.prc																																
1	Low education level	2176	54.59	55.06	55.06																																
2	High education level	1776	44.56	44.94	100.00																																
NA	NA	34	0.85	NA	NA																																
total N=3986 · valid N=3952 · $\bar{x}=1.45$ · $\sigma=0.50$																																					
Interpretation:	The recoded respondents' education in years variable shows that 55.05% of the respondents have a low education level and 44.94% of the respondents have a high education level.																																				

5) Finally, run t-test analysis for “**tvhours**” (outcome variable) and “**educbinary**” (factor variable). Paste the table (10 points) and interpret the results in light of mean scores and statistical significance (10 points).

Code (-5):	t.test(tvhours ~ educbinary, data = gss) > parameters() > display(format="html")																				
Table:	<table><tr><th>Outcome variable</th><th>educbinary = Low education level</th><th>educbinary = High education level</th><th>Difference</th><th>95% CI</th><th>t(2552.51)</th><th>p</th></tr><tr><td>Television screen time in hours</td><td>3.85</td><td>2.75</td><td>1.10</td><td>(0.85, 1.35)</td><td>8.57</td><td>0.000***</td></tr></table>							Outcome variable	educbinary = Low education level	educbinary = High education level	Difference	95% CI	t(2552.51)	p	Television screen time in hours	3.85	2.75	1.10	(0.85, 1.35)	8.57	0.000***
Outcome variable	educbinary = Low education level	educbinary = High education level	Difference	95% CI	t(2552.51)	p															
Television screen time in hours	3.85	2.75	1.10	(0.85, 1.35)	8.57	0.000***															

Interpretation:

The average television screen time in hours for those who have a low education level is 3.85 hours and the average television screen time in hours for those who have a high education level is 2.75 hours.

The average television screen time in hours differs by recoded respondents' education in years in a statistically significant way since the p-value is less than 0.05.