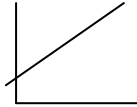# Homework #3d: Correlation & Regression Practice

From the website, get **Smoking & Four Lung Cancers** -- These are 1960s data relating Cigarettes smoked and deaths per 100k in 44 states.

1. Correlate Cigarettes Smoked & the four kinds of cancer. Report the number of unique sig. correlations in the matrix. 5	2. For the relationship between Cig. and B-Cancer, summarize the stat.	$r(42) = .704, p \le .05$ $r(42) = .697, p \le .05$
	3. Summarize the statistics for the three other relationships (between Cig and other cancers)............. ➔	$r(42) = .487, p \le .05$ $r(42) = -.068$, n.s.
4. How likely is it that the correlation between Lung-Cancer and K-Cancer is due to chance? What hypothesis testing conclusion do you reach? 6.3% chance, Retain Ho	5. How likely is it that the correlation between K-Cancer and B-Cancer is simply a fluke? What hypothesis testing conclusion do you reach? 1.7%, Reject Ho.	6. What percent of variance in Lung-Cancer is explained by Cigarettes? $r^2 = .4858$, so 49.59%
7. What percent of variance in B-Cancer is explained by K-Cancer? $r^2 = .1289$, so 12.89%	8. If appropriate, state the regression formula for predicting B-Cancer based on Cigarettes. $y' = bx + a = .122x + 1.086$	9. How much more accurate are you using the regression formula in the previous problem? $r^2 = .495$, so 49.5%
10. If appropriate, state the reg. formula for predicting Lung-Cancer based on Cigarettes. $y' = bx + a = .529x + 6.472$	11. What percent of variance in Lung-Cancer is explained by Cigarettes? What's the std err of the residual? $r^2 = .4858$, Sy' = 3.0661	12. Predict Lung-Cancer deaths based on 40 Cigarettes per capita. $y' = .529(40) + 6.472 = 27.632$
13. If appropriate, state the reg. formula predicting Leuk-Cancer based on Cigarettes. Not appropriate	14. Create a scatterplot with regression line predicting Lung-Cancer with Cigarettes. Sketch here ....➔	
Open the **employee selection** data file. Correlate (in this order) job perf, ass. center avg, cog abil, structured interview, & handwriting analysis.	15. How many unique sig. correlations? 2	$r(15) = .470$, n.s. $r(15) = .520, p \le .05$ $r(15) = .367$, n.s. $r(15) = -.183$, n.s.
	16. Summarize the four correlations with job performance here......... ➔	
17. How likely is it that the correlation between ass. center avg and job performance is due to chance? What hypothesis testing conclusion? .... ➔		5.7%, Retain Ho
18. How likely is it that the correlation between structured interview and job perf is due to chance? What hypothesis testing conclusion? .... ➔		14.7%
19. What percent of variance in job perform explained by cog abil? $r^2 = .2704$	20. Percent of variance in structured int score explained by cog abil? $r^2 = .3457$	21. Explain to manager the problem with using Ass Cntr avg to predict job perf. Greater than 5% chance that correlation is a fluke (i.e., not reliable)
22. If appropriate, state formula for predicting job perf based on cog ability. $y' = bx + a = .009x - 1.160$	23. Predict job perf with cog ability of 700. $y' = bx + a = .009(700) - 1.160 = 5.14$	24. For prior problem, how much overall error in predictions? How much var accounted for in job perf? Sy' = 1.303, $r^2 = .271$
25. If appropriate, state formula for predicting cog ability based on job perf. $y' = 29.832x + 444.781$	26. Predict cognitive ability with job perf scr of 7. $y' = 29.832(7) + 444.781 = 653.605$	27. If appropriate, state formula for predicting job perf based on assessment center average. Not appropriate