(1) Given a z test statistic of -1.40 what is the P-value for the test of statistical significance (two-tailed)? **(A)**

(2) Given a z test statistic of 1.80 is the result statistically significant (two-tailed)? **(C)**

(3) For Budget 7 for *Males*, what is the test statistic for the test of whether the difference in results between A&V and ECO220Y is statistically significant? **(E)**

(4) For Budget 4 for *Females*, what is the P-value for the test of whether the difference in results between A&V and ECO220Y is statistically significant? **(B)**

(5) The sample size has a huge influence on whether the results of a study are __. **(A)**

(6) For the CI estimator $\bar{X}_d \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$, what value should you plug in for n? **(B)**

(7) For the CI estimator $\bar{X}_d \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$, what value should you plug in for $t_{\alpha/2}$ if you wish to obtain a 95% CI? **(C)**

(8) What is the standard deviation of the *difference* (i.e. s_d) in the money passed comparing these two budgets? **(B)**

(9) If you ran a simple regression of *mp_budget6* on *mp_budget4* what would the R^2 be? **(A)**

(10) What is the P-value for the test of whether *overall* there are statistically significant differences among the three sessions? **(C)**

(11) How should you interpret t test statistic of 0.28 for *session_09am*? **(A)**

(12) On average how much money did students participating in the 10am session pass to the other player? **(C)**

(13) Suppose the experiment had been run over 50 one-hour sessions (instead of just three). Further, suppose that there are *no* systematic differences across the session times: i.e. any observed differences are just sampling error. If you ran a regression like the previous one but with 49 dummy variables, what is the chance that at least one coefficient is statistically significant at the 5% level? **(B)**

(14) When testing the statistical significance of the slope coefficient in simple regression, what is the P-value? **(C)**

(15) In seeking to answer a research question about how a person's income influences his/her propensity towards charitable giving, what is the difference between observational and experimental data? **(D)**

(16) The OECD collects data on students' performance on standardized tests across countries. "PISA, the OECD's Programme for International Student Assessment, evaluates the quality, equity and efficiency of school systems. Around 470,000 students participated in PISA 2009 from 65 participating countries." Considering the scatter diagram and OLS line shown, approximately what is the standard deviation of the residuals? **(C)**

(17) What is the R^2? **(C)**

(18) What is the standard deviation of piracy rates in 2011? **(D)**

(19) What is the standard deviation of piracy rates in 2007? (Hint: You need to use 0.0106568 in solving this.) **(E)**

(20) In Canada in 2007 the piracy rate was 0.33 and in 2011 it was 0.27. Given the regression, what is the value of the residual for Canada? **(C)**

(21) In 2007 the mean piracy rate is 0.60. What is the 90% prediction interval for 2011 piracy for a country with a piracy rate of 0.60 in 2007? **(A)**

(22) For multiple regression a rough guide for developing a prediction interval is to take the point prediction plus or minus two standard deviations of the residual. Suppose you tried this for simple regression. Which is <u>NOT</u> necessary for this quick calculation to be fairly close to the exact prediction interval? **(C)**

(23) Which is <u>NOT</u> an important underlying condition to consider when computing a prediction interval? **(B)**

(24) Again comparing A&V (2001) and ECO220Y (2014), consider another partial table of results for *Females* (making sure to read its title and headings). What is the test statistic for the test of whether the difference in results for Budget 4 is statistically significant? **(D)**

(25) How many separate regressions are reported in the table? **(B)**

(26) What do all the regressions have in common? **(A)**

(27) From Panel A what is the P-value for the test of statistical significance (i.e. $H_1: \beta \neq 0$) for the coefficient on *Unemployment rate* in the *Newspapers* regression? **(C)**

(28) In Panel A how should you interpret the coefficient on *Unemployment rate* for the *Banks* regression? When unemployment is ___ then the percent with either "a great deal" or "quite a lot" of confidence in *Banks* is ___ on average after controlling for a time trend. **(C)**

(29) In Panel B what is the *rejection region* for the test of statistical significance of the *Unemployment rate* coefficient if a 5% significance level is used? **(C)**

(30) In Panel B what is the *critical value* for the test of the overall statistical significance of the model if a 1% significance level is used? **(B)**

(31) Sheryl Sandberg, CEO of Facebook, wrote a book "Lean In: Women, Work, and the Will to Lead," which recently came out in paperback. In a radio interview she said "as a man gets more successful, he is better liked, and as a woman gets more successful, she is less liked." To test whether this is borne out in cross-sectional data on likeability, success, and sex, consider this model: $L_i = \beta_0 + \beta_1 S_i + \beta_2 S_i * Male_i + \beta_3 Male_i + \varepsilon_i$. The variable L_i is a quantitative variable measuring likeability, S_i is a quantitative variable measuring success, and $Male_i$ is a dummy variable for sex. What does it mean if the estimated coefficient b_2 is not statistically significant? **(D)**

(32) Which is the best way to check for a violation of the linearity assumption in simple regression analysis? **(D)**

(33) The abstract's description of the results suggests what about how the x and y variables may have been non-linearly transformed? **(B)**

(34) In the multiple regression analysis the ___ variables were standardized. **(A)**

(35) The unit of observation in the multiple regression models is ___. **(A)**

(36) The available data for "Can Facebook Predict Stock Market Activity?" are ___. **(B)**