	20	80	100	
HW-ID	HW3Q1	HW3Q2	HW3	Comments on HW3 from TAs
1003	20	74	94	
1017	18	67	85	Q2 Task2: that's not relatively normal, Task3: you should've subset in terms of observations (rows), Task4: if you decide to use logG as predictor, then have it on y axis when plotting, Task5: forgot to mention direction, Task6: regression line did not include the newly added variable
1085	20	80	100	
1112	20	74	94	Question 2: Demonstrated your thought process very clearly and discussed results of each plot/model. Also created new variable - eastern vs. western hemisphere which was an interesting analysis angle. However, you did not comment on the findings of each plot (e.g. the plots you created in task 3).
1276	20	74	94	
1280	20	68	88	Q2 Task2: might as well look at 'population' as a whole first, Task3: lack of plots for each subset, good job on comparing quantitive variables between two subsets, Task4: if the lm fit for your next task involves both groups, you should have plotted them together here, Task6: OK, didn't show the regression lines, although I can see them in task 4, Task7: might as well consider interaction. Task8: since you decided to have lifeExp as outcome of your regression, you may want to explore this variable at the begining.
1416	18	70	88	Q2 Task2: at least include one plot, Task6: I get what you mean, but no such thing as linear distribution or log distribution
1452	18	79	97	Q2 Task3: great, Task4: normally a simple linear regression is expected, but it's good to see something different, would've been nice to show how you landed on a cubic trend, Task5: forgot to mention direction, Task6: welcome back to linear regression, Task7: OK. Task8: really enjoyed reading your answers.
1478	18	72	90	Q2 Task4: necessity aside, box-cox transformation is not just exponentiating, see the help file for powerTransform for details
1514	20	68	88	Q2, Task2: shape of distribution for lifeExp is not unimodal. not include the numerical summary in the answer details. didn't try the Q-Q plot to check the distribution. Task3: ok, did more that 4 plot for this task. Task4: didn't mention the direction of relation and didn't try rescaling. Task6: ok but didn't rescale.
1590	18	70	88	Question 2: For task 3, the use of a bar chart (geom_bar) might have been a better way to visualize your data instead of using a scatter/box plots. Your faceted scatterplot comparing GDP per capitas vs. Life expectancy in African and European was an effective visualization, but you did not provide much interpretation of the findings other than reporting the regression and equation and R^2

	20	80	100	
HW-ID	HW3Q1	HW3Q2	HW3	Comments on HW3 from TAs
1685	14	68	82	Q1 NOT appropriate interpretation of the plot, Q2,Task1: no explanation for the code, Task2: shape of distribution for lifeExp is not normal. not include the numerical summary in the answer details. Task4: didn't mention the direction of relation and didn't try rescaling. Task5: didn't mention the direction of relation and statistical significance. no plot. Task7: didn't mention the direction of relation and statistical significance. no plot. Task8: did not address all questions.
1699	14	70	84	Q2, Use "anova(model1,model2)" to test the difference between two models before you decide to include interaction terms; Better to add plots to task 6
1704	20	80	100	
1713	14	70	84	Question 1: Even though age has the largest Spearman rho square value for a quantitative variable, the categorical variables sex and tobacco have larger values Spearman rho square values which would indicate that an interaction term of these variables may be something we would want to consider before adding a non-linear term to the quantitative variables. Question 2: Your visualization of life expectancy using boxplots stratified by year was interesting and demonstrated to the viewer that life expectancy appeared to increase with time. However, your labeling of the graphs and explanations associated with each graph was somewhat unclear. Also, your plot "Relationship between LifeExp and GDP per Capita" was a little busy. You could probably improve this graph by plotting fewer countries. It would also help if you discussed the implications of such a graph (e. g. increases in GDP per Capita appears to have a positive effect on Life expectancy in all countries, but the increase appears greater for some than others).
1716	12	68	80	Q2,Task2:Distribution plots missing; Task3:Good; Task4:Better to include confidence interval in plots;
1721	17	74	91	Q1 NOT include all variables in the plot, Task8: VERY GOOD
1785	20	66	86	Question 2: Your work was very clear and easy to follow and you created some interesting plots. However, you did not provide interpretation on the results observed in the plots.
1825	20	80	100	Excellent work, remeber to tranform variables if needed
1929	18	72	90	Q2,Task1-3:Good; Task6: "country" may not be a good choice for continent variables since there are too many categories
1991	18	68	86	Q2, Task3: no plot for quantitative variable. Task4: didn't try rescaling. Task5: didn't mention the statistical significance. no plot. Task7: not full model details.
2039	18	74	92	didn't specify sex as non-linear

	20	80	100	
HW-ID	HW3Q1	HW3Q2	HW3	Comments on HW3 from TAs
2064	20	74	94	
2137	20	74	94	
2171	18	72	90	Q2 please don't forget to label the tasks by numbers.
2230	18	75	93	Question 2: Really unique and great graphs - you obviously worked to incorporate techniques you learned outside of class. A few minor things - your labels for graphs are not always clear (e.g. x-axis should state log(GDP per Capita) when your taking the log) and you were not able to plot all the data of interest onto your geography plot (e.g. the united states appeared white even though you had united states data in your data set [this may have happened because the names did not match up between gapminder and the map_data dataframes)
2277	20	72	92	
2369	14	68	82	Q2,Task1-2:Good; Task3: Where is your subset analysis; Task4&5: Good try for cubic and polynomial model; Conclusion missing;
2384	20	74	94	
2389	18	68	86	Q2 Task1: good job spotting Kuwait's GDP, Task5: should've printed out the summary of lifeExp~GDP instead of lifeExp~pop as you interpreted the former, Task7: you fitted models with and without interaction, but didn't show why you chose to keep the one without interaction. Task8: please - next time - have each task in separate section.
2426	20	80	100	
2462	18	74	92	Q2 Task3: couldn't find your plot for lifeExp in 2007 without other variables, Task4: should've put your plot for GDP and lifeExp in this section, Task8: good work overall, although I didn't see why you did log10 transoformation instead of log.
2512	20	68	88	Not necessary to print the whole database in your file, missing analysis
2569	18	62	80	LATE, Q2, Task2: no explanation for the plots. Task3: no explanation for the plots. Task5: no model. Task7: no model details.
2688	20	70	90	Q2,Task1:Skim function is also optional; Task2:Distribution plot missing; Task3:Better to add conclusions to each question, plot of a single quantitative variable missing; Task4&Task5:Better to include residual plots and confidence intervals in plots; Task6: Nice plot

	20	80	100	
HW-ID	HW3Q1	HW3Q2	HW3	Comments on HW3 from TAs
2719	16	64	80	Q2,Task3: Subset analysis missing; Task6: Not clear what research question you want to ask; Task8: Again, what's the purpose to incorporate "year" into your model and what's the research question?
2741	18	67	85	Q2 dont forget to label tasks by numbers. Task4: OK but no explanation for the plot, Task5: didn't mention the statistical significance. no plot, Task7 :didn't mention the statistical significance, no plot.
2875	18	76	94	Question 2: Demonstrated your thought process very clearly and had thorough analysis of the results from each plot/model. Visualized quatitative + categorical variable data by both a violin plot and faceted plot. Also interpretted your residual plots of your final model
2915	18	73	91	Q2, Task4: use the difference variables plot instead of the real variable plot. Task5: no plot fot the model. Task7: not full model details
2916	20	68	88	Question 2: Quantitative plots: You plotted a histogram of GDP Per Capita for the entire data set and then plotted a GDP Per Capita in a violin plot to compare countries in the Americas and Africa. You could have used filter/subset function to narrow the dataset down to a more interesting cross-section of the data (e.g. histogram of GDP per capita of a certain year). Having multiple GDP Per Capita values for a single country, does not really teach us anything interesting about the distribution of GDP per Capita or how African countries compare to countries in the Americas
2989	18	70	88	Q2, Task5: didn't mention the statistical significance, no model plot. Task7: no model plot. Task8: very good.