HW-ID	HW6	HW6 Comments from the Teaching Assistants	Q1	Q2	Q3	Q4	Q5
1003	100	Excellent homework	20	20	20	20	20
1017	90	Q1: Including the county names make it more difficult to visualize. Q4: Good job on the model.	16	16	20	20	18
1085	94	Q3. The predictions need a better interpretation in terms of how close or how far were they from the original number. Q4. For the interpretation for example 1.79 means 80% higher odds.	20	18	18	18	20
1112	100	Q1: impressive graph, using circle size was very effective. Q2: very nice essay.	20	20	20	20	20
1276	95	Q1. Good graph, you can improve the labels, use complete words for titles (e.g.Percentage instead of %) Q4. Conclusions on your model need to be more specific. Q5. Missing rootogram	18	20	20	18	19
1280	96	Note: Please don't include the whole dataset in your .Rmd file. Q1. Good graph it looks too busy. Q2. Review your grammar. Good job	18	18	20	20	20
1416	96	Nice job	20	16	20	20	20
1452	100	Excellent homework	20	20	20	20	20
1478	94	Q4: very nice job. try not to include every step you dau before the final choosing of the variables.	16	18	20	20	20
1514	92	Q1: any finding on population size? Place your caption within the plot Q4: your outcome is not ordered Q5: the LogLik was larger for the Poisson model, not smaller	16	20	20	17	19
1590	92	Q1: if you used circle size with the color, the graph would be more strong and clear.	16	16	20	20	20
1685	84	Q1: it's not a good idea to label every county Q2: it's hard to tell if counties with water violation generally have higher mortality. I think you meant "loess", not "loose" Q3: Normal QQ plot is not for linearity, it's for normality. Q4: your outcome is ordered incorrectly ("low" in the middle) Q5: your dummy variable for obese_pct is incorrect	16	18	18	16	16
1699	75	Q1: lack of a plot describing three or more variables. Be consistent with your axis labels and your legend needs some fixing. Q3: your outcome "smoker_pct" isn't one of the outcomes defined in 2017 Ohio Summary report. "ment_days" and "phys_days" are highly correlated, I wouldn't include both. You had a residual plot, but didn't discuss if the model assumptions were met. Q4: your newly created outcome is not ordered. The model interpretation is vague. The model assumption is wrong. Q5: the count outcome created incorrectly. No prediction for the two counties.	16	18	15	14	12
1704	88	Q1: good effort, but your cluster map needs legend Q2: again, good effort and cool results, but too long (there is a word limit). Q3: your outcome "smoker_pct" is not one of the outcomes defined in 2017 Ohio Summary report. Typo. Q4: it's not clear to me how you divided the outcome (the way you created the variable seems pretty random to me, and it was not ordered). For the prediction, what are the observed levels for the two counties? Print them out. Q5: should have fitted another model using method for count outcome to compare with your Poisson model. Again, present the observed values for the two counties.	18	18	18	17	17
1713	88	Q2: nice essay, try to use number and percent next time to make the essay stronger. Q4: try to include assumption assessment and explain your results.	16	16	20	16	20
1716	95	Q1: nice graph. would it be better if you used the county names for some to show the highest of lowest. Q4: nice work.	19	18	20	20	18
1721	94	Q3: it is not so clear to me why you fitted a logistic regression. The results are not so off because your outcome was treated as probability. It'll bring difficulties in interpretation. Q5: would have been better if overdispersion was discussed	20	20	16	20	18
1785	95	Q1: very nice job, if you used another variable such population number for one of the visualization ways either different color of circle size would add more to your graph. Q4 and Q5: nice work.	17	18	20	20	20

HW-ID	HW6	HW6 Comments from the Teaching Assistants	Q1	Q2	Q3	Q4	Q5
1825	92	Q1: place your caption within the plot Q2: I wouldn't call an indication of weak correlation "not helpful" Q3: very good job overall, but no discussion on model assumptions Q4: no coefficient estimates interpretation	18	18	18	18	20
1929	90	Q3: your outcome "poor_kids" is not one of the outcomes defined in 2017 Ohio Summary report. There is no normality assumption for the predictors, you can google to see what the assumptions are for linear regression. It would be nice to assess the residual plot and then discuss whether the model assumptions are met. Q4: lack of model interpretation Q5: your count variable was created incorrectly. Your approach would have worked but you assigned the wrong 0 and 1. Pay attention to grammar mistakes.	20	20	17	18	15
1991	85	Q1: no caption Q2: Why is this particular visualization effective? Q3: your outcome "inactive_pct" is not one of the outcomes defined in 2017 Ohio Summary report. When discussing the model assumptions, I don't know what you mean by "not constant variance". In fact, linear regression assumes equal residuals across the regression line. And print out your prediction for the two counties. Q4: how well does your prediction agree to the observed value? Q5: again, how well does your prediction agree to the observed value?	16	16	17	18	18
2039	94	Q1. Nice graph, however the groups compared need to be revised. Q2. Grammar issues Q3. Make your conclusion specific to your linear model	17	17	20	20	20
2064	62	Q1. I like the idea of the graph but right now it looks very busy, it needs more cleaning (adjusting font size, etc...) Q2. Your explanation is too short and doesn't address the questions or your plot. Q3. Missing interpretation of model estimates. Q4. Missing interpretation of model estimates and predicted values.	17	10	18	17	0
2137	90	Q1: better to label how food access is measured as well Q4: address your prediction compared with the observed values Q5: should have fitted another model using method for count outcome to compare with your Poisson model. And why did you end up using Poisson for prediction? It's not clear.	18	18	20	18	16
2171	90	Q1: wrong legend label for unemployment Q2: again, what you colored is unemployment, not high school graduation Q3: typo. Should have stated what the assumptions are. Q5: you don't need to look at Normal QQ plot for model assumptions	18	16	18	20	18
2230	54	Q1: simple and different although pie graphs are very misleading. Another variable is needed to serve as a basis for comparison.	16	18	20	0	0
2277	95	Q1. Good Job creating the map, try to change the title and the labels to make them easier to understand and cleaner. We took the first plot for grading purposes Q2. Explanation for your first plot was very short Q3. Good job on the model.	18	17	20	20	20
2369	93	Q1: could have mentioned how you defined low and high smoking rate. Q3: very good job overall, note that sroh_fairpoor, you treated as predictor, is defined as one of the health outcomes, I'd be careful including it in a model predicting another health outcome Q4: No interpretation for the coefficients estimates? I don't get why you couldn't test with multinomial model. Why did you refit the model with your test set and then use predict()? Note that to test on the other two counties, just used the model you get from 86 counties and specify the newdata as the test set within predict().	20	20	19	14	20
2384	63	Q2: you should describe the findings from your plot. Capitalize the first letter of a sentence Q3: overall lack of efforts. No discussion on model assumptions. Q4: you were asked to fit an ordinal logistic regression which you didn't. And still a lack of efforts Q5: your binary variables were incorrectly created. You were supposed to fit different models using different methods, not different variables. No prediction.	20	15	14	6	8
2389	91	Q1. Nice graph, however the groups compared need to be revised Q5. You are missing one model.	16	20	20	20	15

HW-ID	HW6	HW6 Comments from the Teaching Assistants	Q1	Q2	Q3	Q4	Q5
2426	80	Q1. Nice graph, however the groups compared need to be revised/ The second plot is better the legends and the labels need to be improved. Only the first plot was taken into consideration. Q2. It's important to address the questions, this one was specific for this data not all visualizations. Q3. Your findings don't match your conclusion, it's important to start your sentences with capital letters, your prediction conclusion is very vague. Q4. Watch out for grammar errors.Conclusions need to be specific. Q5. Watch out for grammar errors.Conclusions need to be specific.	16	14	15	18	17
2462	88	Q1. Good graph try to pick either the color of the circle size or just color the circles and get 1 label, it look very busy. Q3. Missing conclusions about the goodness of fit of your model, estimates interpretation and predictions. Q4. The order of your categories is wrong, you got a wrong conclusion. Q5. Good job creating the models, but what is your final conclusion about which one you chose.	18	20	17	16	17
2512	92	Q1: no caption describing the key results Q2: good overall, but exceeded the word limit Q3: your outcome "sev_housing" is not one of the outcomes defined in 2017 Ohio Summary report. No need to transform predictors, the normality assumption is for the outcome. If you fit rcs within ols() instead of lm(), then you can predict for the two counties. Q4: typo	18	17	18	19	20
2569	93	Q1: exceeded caption word limit Q3: typo Q5: should have fitted another model for count outcome to compare with your Poisson. No need to check normal QQ plot	17	20	19	20	17
2688	80	Q1: your two plots can easily be integrated into one plot and the color and size didn't provide extra information. Q2: the advantages of visualization you mentioned are too general. Pay attention to grammar mistakes. Q3: your outcome "obese_pct" is not one of the outcomes defined in 2017 Ohio Summary report. No discussion on model assumptions Q4: your outcome is not ordered. Q5: you need to fit two models using the same predictors, not different predictors with the same method. No prediction for the two counties.	15	17	17	17	14
2719	81	Q1: your caption didn't highlight any key results Q2: I would disagree that your visualization would help the audience understand the relationship among three variables easily. Pay attention to grammar mistake. Q3: your outcome "uninsured" is not one of the outcomes defined in 2017 Ohio Summary report. Your model was built using the full data instead of holding out the two counties. No discussion on model assumptions. Q4: your outcome is not ordered. Again, the model was built using the full data. The df for your chi-squared test is 3, not 2. Q5: do not use the binary variables you created for the outcome.	16	18	16	15	16
2741	86	Q1: no caption. Interaction is not simply a visual intersection Q2: is it really obvious that exer_access has an effect on obesity? Q3: your outcome "obese_pct" is not one of the outcomes defined in 2017 Ohio Summary report. Normality is not the only assumption here Q4: your outcome is not ordered Q5: when looking at prediction, you should have stuck with the results from specifying type="response"	16	18	18	17	17
2875	98	Q1: very effective way to include more than two variable to come out with a very important conclusion. good job. Q2: very nice essay to describe the proceeding graph. Q4: nice work.	20	18	20	20	20
2915	89	Q1: exceeded caption word limit Q3: no discussion on model assumptions Q4: your outcome is not ordered Q5: you need to specify type="response" when predicting	16	20	18	17	18
2916	97	Q4: good job, try to be more specific in conclusions.	19	18	20	20	20
2989	95	Q1: good effort, but each graph only gives information on one variable Q3: should've discussed more about model assumptions	16	20	19	20	20