

Assignment 4 Rubric

Task	Criteria for Full Credit	Partial Credit	Points
1a. Histogram & QQplot (texture_mean)	Histogram and QQplot are correctly generated with appropriate axis labels and titles.	Plot(s) included but with missing labels/titles OR incorrect variable used.	0.25
1b. Interpretation of QQplot	Clear interpretation of normality based on QQplot (e.g., “deviates from line → not normal”).	Vague or partially correct interpretation.	0.25
1c. Logarithmic Transformation	Correct application of log transform, justification provided for whether to use transformed or untransformed data in linear model.	Log applied but no justification, or justification unclear.	0.5
2. Linear Model (texture_mean ~ area_mean)	Model run correctly, output inspected, significance properly interpreted (p-values, coefficients).	Model run but incomplete interpretation OR misinterpretation of results.	0.75
3. Visualization of Linear Model	Scatterplot with points and regression line included. Equation/formula displayed in title. Clear, professional formatting.	Plot included but missing regression line OR formula in title OR poorly formatted.	1.0
4. Filtering Genes in RNA-seq Data	Correctly filters out genes with ≤ 2 columns of counts. Reports accurate number of remaining features and samples.	Filtering attempted but incorrect numbers reported OR method unclear.	0.75
5. Run DESeq Function	Correct execution of DESeq() function to fit the object.	Code attempted but incomplete OR missing explanation.	0.25
6. DPN vs Control Differential Expression	Correct extraction of results for DPN group. Identifies top 5 DE genes. Reports number of up- and down-regulated genes.	Only partial results (e.g., top 5 genes but not up/down counts, or counts but not genes).	0.75
7. Gene Expression of Highest Log-Fold Change	Correctly identifies gene with max logFC and visualizes/compares across treatments.	Correct gene identified but incomplete comparison OR incorrect visualization.	0.5
8. Volcano Plot (DPN treatment)	Creates a volcano plot with proper axes (logFC vs $-\log_{10}(p)$), applies $p < 0.05$ cutoff, highlights DE genes clearly.	Plot generated but missing cutoff or incorrectly formatted.	0.75
Presentation & Code Quality (applied across all tasks)	Code is well-commented, plots labeled, answers concise and clear.	Code runs but poorly documented OR explanations unclear.	± 0.5 bonus/penalty