Name: _____ **Grade: ____/ 100**

Mid Semester Project Rubric (Written Portion)

Score: ____/52

	No Credit	Half Credit	Full Credit
Introduction (10 points)	Missing introduction.	Introduces breast cancer, but is missing key elements of an introduction such as citations, a section on genes/epidemiological factors, methods, and/or findings.	Accurately and thoroughly introduces the background of breast cancer, citing all articles used. Introduces a gene/epidemiological factor and connects it to breast cancer to set up the analyses of the paper. Briefly touches upon methods and findings.
Methods (10 points)	Missing methods.	Explains the methods, but either is missing key elements (R packages, plots, etc.) or goes far too in-depth for a methods section (lists out all arguments for each function, etc.). Layout is disorderly and difficult to follow.	Thoroughly explains the computational methods used in the analysis including packages and key arguments (ex: accession codes). The layout of the methods section is logical and the analysis could be replicated from these descriptions.
Results (10 points)	Missing results.	Results section is missing descriptions of key findings, does not refer to figures, or attempts to explain results instead of just stating them.	Explains what was found in the analysis, and does so objectively. Does not speculate on *why* those results were found, simply lists key findings. Refers to figures to back up each statement.
Figures (10 points)	Missing figures.	Figures are separate from the results section, are missing captions, or do not follow the guidelines in the directions.	Figures are embedded in the results section and have descriptive captions. Includes at least one Kaplan-Meier plot, at least one generic R plot, and at least one plot created using maftools. Plots accurately and intuitively describe the data, and the chosen figures are logical for the type of data that is being shown.

	No Credit	Half Credit	Full Credit
Discussion (10 points)	Missing discussion.	Does not compare results to previous analyses in the field or does not connect key results to physiology. Does not analyze implications or is missing a statement on future work.	Makes logical interpretations of results based on findings from this analysis and from background literature. Speaks clearly and thoroughly about background biology and connects findings to what is known and unknown about cancer physiology. Cites all papers used. Ends the discussion section by analyzing the implications of this research and brainstorming future directions to continue in.
References (2 points)	Missing citations.	Either in-text citations or a references section is missing, or citations are messy and unstandardized. Statements that should be cited are not.	Citations are included both in a references section and directly in the text itself. Citations are standardized and correctly formatted.

Comments:

Mid Semester Project Rubric (Code Portion)

Score: ____/32

	No Credit	Half Credit	Full Credit
Libraries (2 points)	Missing library statements.	Either installs libraries or loads libraries, but not both.	All necessary libraries are installed and loaded in at the top of the file.
Cleaning Data with Boolean Indexing (5 points)	Data is not cleaned of NAs, or this is done using a technique other than boolean indexing.	An attempt at cleaning the data is made, but there are issues such as flipped boolean masks, flipped rows/columns, etc.	Data is cleaned using methods taught in class (ex boolean indexing). NAs are removed when necessary and subsetting of data frames is correctly done. Variables with relatively low NA counts are chosen.
Plots (10 points)	Missing plots, or the code does not accurately create plots when run.	Plots axes are messy or unreadable, and the code includes substantial issues or inconsistencies.	Includes at least one Kaplan-Meier plot, one generic R plot, and one plot created using maftools. Plot axes are descriptive and properly formatted. Plots accurately and intuitively describe the data. Plots are saved to a folder called “outputs” inside of “midsemester_project_lastname”
Executable (10 points)	When run from a clean environment, the analysis does not run as it should. Errors are either substantial (ex: NAs are not removed, boolean mask is in the wrong column/row, etc.) or the file contains numerous smaller errors.	When run from a clean environment, there are only a few relatively small errors (typos, swapped lines of code, etc.)	The entire coding file can be run from a clean environment with no errors. This includes creating the “outputs” folder, installing/loading all packages, querying, downloading, and preparing all data, and saving all data/plots to “outputs”.
Readability (5 points)	Code is messy and overall unreadable. No comments are included.	Most variables are named ambiguously, some lines are too long, or comments are sparse and unhelpful.	All variables are named descriptively using either camelCase or snake_case. Explanatory comments are included

Comments:

Mid Semester Project Rubric (Review Questions)

Score: ____/16

	No Credit	Half Credit	Full Credit
General Concepts (4 points - 2 points per question) Q1: Q2: Total:	Missing answers.	Answers are incomplete or not fully thought out.	Responses are thoughtful and show a developed understanding of public data sets and multi-omic analyses.
Coding Skills (12 points - 2 points per question) Q1: Q2: Q3: Q4: Q5a: Q5b: Total:	Missing answers.	Answers are incomplete or not fully thought out.	Responses are thoughtful and show clear and accurate knowledge of coding skills.

Comments: