|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Criteria** | **High distinction (HD)**  **85 – 100%** | **Distinction (D)**  **75 – 84%** | **Credit (C)**  **65 – 74%** | **Pass (P)**  **50 – 64%** | **Fail (FL)**  **<50%** | **Feedback to student** |
| Undertake EDA to understand data and select appropriate covariate sets (30%)  **[ / 30 ]** | Extensive and detailed set of covariate justifications present, with thorough EDA and clear and extensive interpretation, going beyond replication. | Detailed set of covariate justifications present, with thorough EDA, with clear and correct interpretation. | Covariate selection present but lacking in detailed justification, EDA mostly presented, with some inconsistencies. | Limited justification of covariate justification present, some EDA shown, but lacking interpretation or implication. | No covariate justifications present, lacking or totally missing EDA. |  |
| Constructs univariable and multivariable models and appraises model fit (40%)  **[ / 40 ]** | Statistical method correctly applied, model selection is justified, and correct model fits presented and systematically interpreted. Final model diagnostics fully presented and correctly interpreted. All code has a clear purpose and runs without errors. | Statistical method correctly applied, model selection is justified, and correct model fits presented and interpreted. Final model diagnostics presented and correctly interpreted. All code has a clear purpose and runs without errors. | Model specifications and diagnostics are explored or are presented without demonstrating an understanding of why they are performed. Some code has unclear purpose or produces an error. | Limited model specifications and diagnostics are explored, without interpretation. Some code has unclear purpose or produces an error. | Inappropriate statistical method is performed, or correct method is performed but with major errors invalidating the results; large sections of code does not run or has unclear purpose; appropriate diagnostic tests are not performed or are misinterpreted. |  |
| Interpretation  (20%)  **[ / 20 ]** | Statistical output is clearly presented and interpreted correctly and thoroughly. Implications follow clearly from the analysis, are nuanced, and demonstrate appreciation for the strengths and limitations of the analysis. | Statistical output is clearly presented and interpreted correctly. Implications follow clearly from the analysis, and demonstrate appreciation for the strengths and limitations of the analysis. | Analysis output is presented adequately but lacks some important features. Interpretation lacks nuance or has some inaccuracies or misinterpretations. Implications, strengths and limitations are discussed at a superficial level. | Correct output is presented indiscriminately, without highlighting the important statistics. Interpretation is lacking, implications are absent or do not follow from the analysis. Strengths or limitations of the analysis are not described. | Correct output is not presented. Interpretation is incorrect or lacking entirely; implications are absent or do not follow from the analysis. Strengths or limitations of the analysis are inaccurate or absent. |  |
| Presentation  (10% )  **[ / 10 ]** | R code is well-written and clearly annotated and documented; right packages included for installation, tables and figures are presented to a publication-quality, with clear titles, labelling and legends; references are included as needed. | R code is well-written and documented; tables and figures are presented to a publication-quality, with clear titles, labelling and legends. | R code is written correctly but with some features missing.  Tables/Figures are interpretable but lack some details to reach publication standard, such as inadequate labelling/legends/scaling etc. References are present but have minor errors. | R code is poorly written and presented, with minimal comments. Tables/figures are difficult to interpret, with insufficient headings/ labels/ legends etc; references are missing. | R code is poorly written and presented, with no comments. Tables/figures are not presented; references are absent. |  |
| **Overall mark for Question 1 [ / 100 ]** | | | | | | |

**Question 1: Logistic regression**

**Question 2: Survival analysis**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Criteria** | **High distinction (HD)**  **85 – 100%** | **Distinction (D)**  **75 – 84%** | **Credit (C)**  **65 – 74%** | **Pass (P)**  **50 – 64%** | **Fail (FL)**  **<50%** | **Feedback to student** |
| Undertake EDA to understand data and select appropriate covariate sets (30%)  **[ / 30 ]** | Extensive and detailed set of covariate justifications present, with thorough EDA and clear and extensive interpretation, going beyond replication. | Detailed set of covariate justifications present, with thorough EDA, with clear and correct interpretation. | Covariate selection present but lacking in detailed justification, EDA mostly presented, with some inconsistencies. | Limited justification of covariate justification present, some EDA shown, but lacking interpretation or implication. | No covariate justifications present, lacking or totally missing EDA. |  |
| Constructs multivariable models and appraises model fit (40%)  **[ / 40 ]** | Statistical method correctly applied, model selection is justified, and correct model fits presented and systematically interpreted. Final model diagnostics fully presented and correctly interpreted. All code has a clear purpose and runs without errors. | Statistical method correctly applied, model selection is justified, and correct model fits presented and interpreted. Final model diagnostics presented and correctly interpreted. All code has a clear purpose and runs without errors. | Model specifications and diagnostics are explored or are presented without demonstrating an understanding of why they are performed. Some code has unclear purpose or produces an error. | Limited model specifications and diagnostics are explored, without interpretation. Some code has unclear purpose or produces an error. | Inappropriate statistical method is performed, or correct method is performed but with major errors invalidating the results; large sections of code does not run or has unclear purpose; appropriate diagnostic tests are not performed or are misinterpreted. |  |
| Interpretation  (20%)  **[ / 20 ]** | Statistical output is clearly presented and interpreted correctly and thoroughly. Implications follow clearly from the analysis, are nuanced, and demonstrate appreciation for the strengths and limitations of the analysis. | Statistical output is clearly presented and interpreted correctly. Implications follow clearly from the analysis, and demonstrate appreciation for the strengths and limitations of the analysis. | Analysis output is presented adequately but lacks some important features. Interpretation lacks nuance or has some inaccuracies or misinterpretations. Implications, strengths and limitations are discussed at a superficial level. | Correct output is presented indiscriminately, without highlighting the important statistics. Interpretation is lacking, implications are absent or do not follow from the analysis. Strengths or limitations of the analysis are not described. | Correct output is not presented. Interpretation is incorrect or lacking entirely; implications are absent or do not follow from the analysis. Strengths or limitations of the analysis are inaccurate or absent. |  |
| Presentation  (10% )  [ / 10 ] | R code is well-written and clearly annotated and documented; right packages included for installation, tables and figures are presented to a publication-quality, with clear titles, labelling and legends; references are included as needed. | R code is well-written and documented; tables and figures are presented to a publication-quality, with clear titles, labelling and legends. | R code is written correctly but with some features missing.  Tables/Figures are interpretable but lack some details to reach publication standard, such as inadequate labelling/legends/scaling etc. References are present but have minor errors. | R code is poorly written and presented, with minimal comments. Tables/figures are difficult to interpret, with insufficient headings/ labels/ legends etc; references are missing. | R code is poorly written and presented, with no comments. Tables/figures are not presented; references are absent. |  |
| **Overall mark for Question 2 [ / 100 ]** | | | | | | |