WINE DATA MODELLING
==========

1. Build a statistical model that estimates wine rating points.
2. Model whether the wine is suprior (points >= 90).
3. Identify factors associated with a good rating.
4. Can I identify whether some wine varieties in the data are better than others?
5. Use appropriate methods (hierarchical logistic regression).
6. Use either points or "superior_rating".
7. Feature "description" is free text use NLP techniques.
8. Group by the "variety" feature.
9. Investigate the skewed "price" variable.
10. Clearly explain models. Summary statistics are not models.
11. Decisions should be discussed and justified.
12. If you fit two models and one is better than the other, spend more time discussing the better model.
13. Central figure can be composite of several figures/diagrams (https://www.scss.tcd.ie/arthur.white/Teaching/STU33011/cluster_poster_dublin_nice_visualisation.pdf). It should include:
        # What model is specified?
        # Variable transformations and summaries.
        # Use of complementary skills, etc.
        # Use of model output.
14. Relavent case study: https://www.scss.tcd.ie/arthur.white/Teaching/CS7DS3/Case-Study-8-Hierarchical-Logistic-Regression.html
15. Use Hierarchical Logistic Regression models.
16. Case Studies:
        # https://www.scss.tcd.ie/arthur.white/Teaching/CS7DS3/Case-Study-6-Linear-Regression.html
        # https://www.scss.tcd.ie/arthur.white/Teaching/CS7DS3/Case-Study-7-Logistic-Regression.html
        # https://www.scss.tcd.ie/arthur.white/Teaching/CS7DS3/Case-Study-8-Hierarchical-Logistic-Regression.html