Exercise: AIS dataset

1. Visualize the distribution of red blood cell count (rcc) by sex (sex).
2. Is there a significant difference in red blood cell count between the two groups of sex?
3. Produce a correlation matrix of the relevant variables in this dataset.
4. Visualize the relationship of height (ht) and weight (wt).
5. Regress ht on wt. Find the equation of the fit regression line. Is there a significant relationship?
6. Split your regression model into training and testing subsets. What is the R-squared and MSE on your test model?