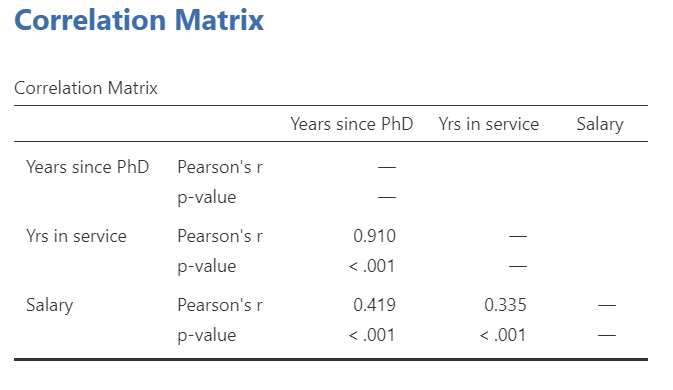
CHAPTER 6

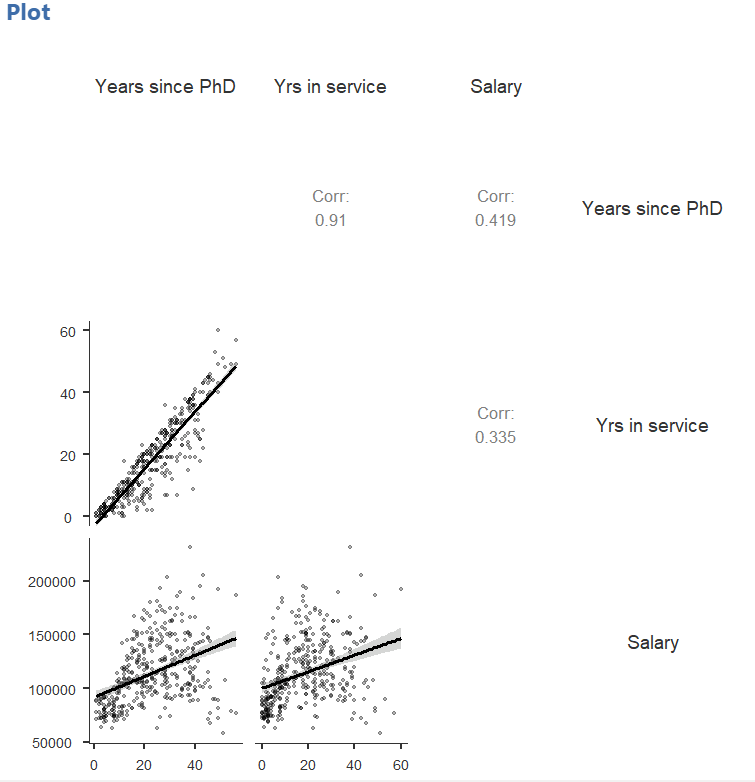
For this chapter, use the Salaries.csv file, opening the file in Jamovi.

**Correlations**

Find the correlations for all of the numeric variables with each other. Interpret the results. Which are strongly correlated and which are not? Are the correlations significant?

Use the Correlation Matrix, as shown in the book, using the three variables 'Years since PhD', 'Yrs in Service' and 'Salary'. The hypothesis should be 'Correlated'. Show also 'Report significance', 'Correlation matrix' and 'Statistics'. The results are from the Pearson option.





Note: Make sure that Years since PhD, Years in service, and Salary are input as continuous variables in order to get their correlation.

The results shows that years in service and years since PhD are strongly correlated. The correlation coefficient is 0.91 and it is significant since the *p* value is less than 0.05.

Salary and years since PhD are also positively correlated and this correlation is significant. However, it is not as strong as the first two since the coefficient is just 0.419. The significant correlation shows that as years since PhD increases, the professors' salaries increase as well.

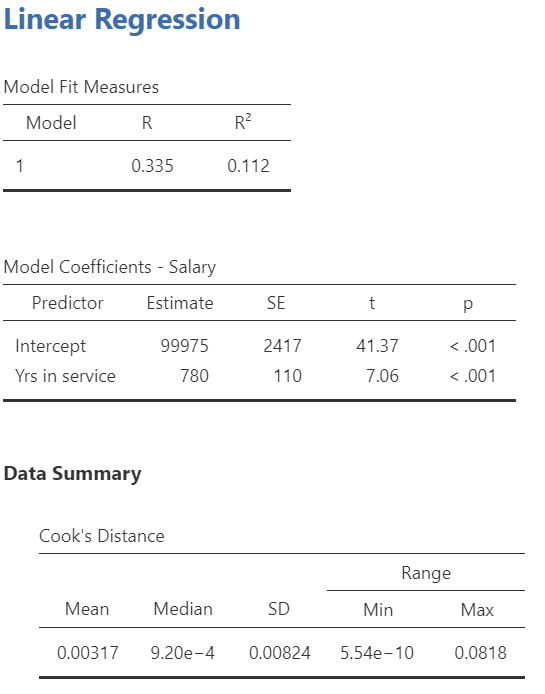
Lastly, years in service and salary have a significant positive correlation but is seen as the weakest since the correlation coefficient is 0.335. The salary increases with years in service.

The plot visualizes these positive relationships between the pairs of variables. The steepest line belongs to the plot of years since PhD vs years in service which has the highest correlation coefficient.

**Simple Linear Regression**

Fit a linear model that predicts a professor’s salary given his or her years in service.

Use the Linear Regression dialog box in Jamovi, as shown in the book. Salary should be the dependent variable; Yrs in service should be in the 'Covariates' box.



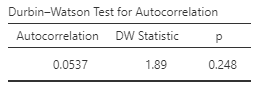
Model interpretation:

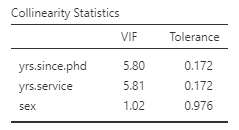
Based on the model fit measures, years in service can interpret 11.2% of the variability of a professor’s salary. Moreover, the estimates show that for every year increase in a professor’s tenure, his or her salary increases by $780.

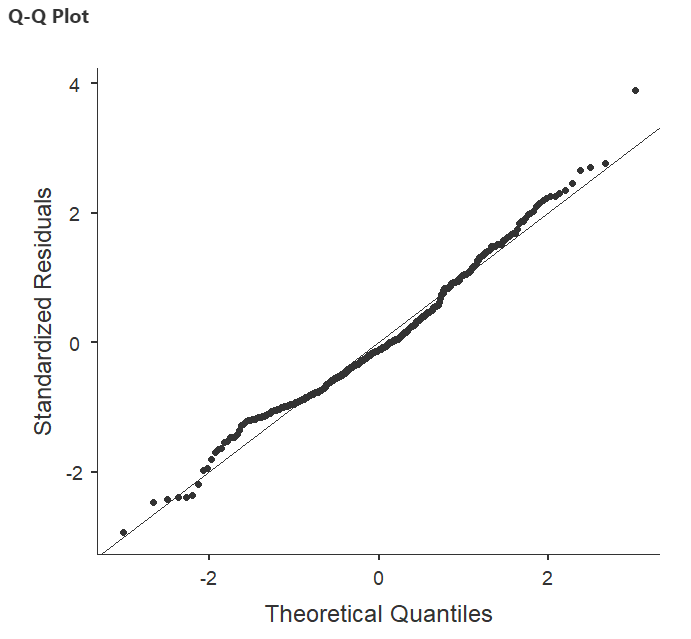
**Multiple Regression**

Fit a multiple linear regression model that predicts a professor’s salary using his or her years since graduating PhD, years in service, and sex. Analyze to see if the assumptions were satisfied and do the necessary corrections if not.

Salary remains the dependent variable, but this time the covariates are years since PhD, Years in service and Sex. From the Assumptions dialog, choose Autocorrelation, Collinearity and Q-Q plot.







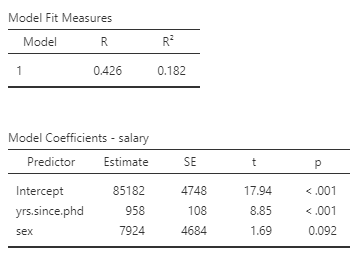
Assumption checking:

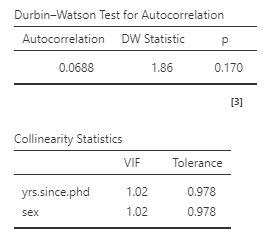
The Durbin Watson statistic should be between 1.5 and 2.5 to say that autocorrelation is not present. This was satisfied in the model since the DW Statistic is equal to 1.89. Moreover, the null hypothesis in this test statistic is that there is no autocorrelation between the residuals; since the *p* value is greater than 0.05, we do not reject the null hypothesis.

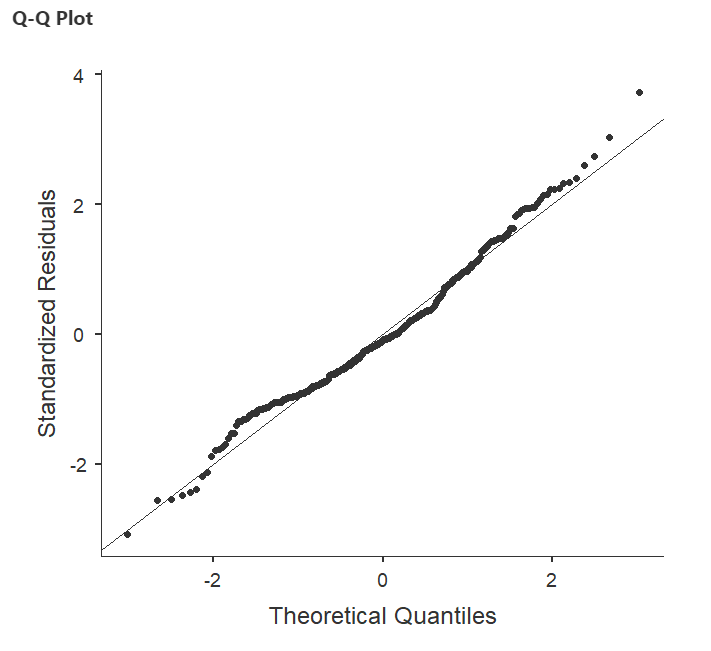
To test for multicollinearity, look at the VIF and tolerance statistics results. You want a low VIF and high tolerance. In this case, the opposite can be observed for the years since PhD and years in service variables. These show that they are more likely to be correlated with each other. Note that from the correlation results before this section, the two variables have the highest Pearson correlation coefficient, providing further confirmation that the assumption was not met.

For the test of normality, looking at the Q-Q plot, there is not much deviation from the normal line.

Since there is multicollinearity between years of service and years since PhD, remove one of these variables and retain the other within the data set. The basis for deciding on which to remove depends upon the researcher but for simplicity, retain Years since PhD since it has a higher correlation to salary based on the previous output.







Assumption checks:

Looking at the assumptions in the same manner as we did before, you can verify that all were met this time. There is high tolerance and low VIF for both years since PhD and sex. Thus, the collinearity assumption was met. We can proceed to interpret the model results.

The model fit shows that 18,2% of the variability of salary can be explained by years since PhD and salary. For the estimates, for every 1-year increase since being a PhD graduate, annual salary increases by $958 on the average. A male professor earns $7924 more than a female professor.