# The Coefficient of Determination

- The coefficient of determination $R^2$ is used in the context of statistical models whose main purpose is the prediction of future outcomes on the basis of other related information.
- It is the proportion of variability in a data set that is accounted for by the statistical model.
- It provides a measure of how well future outcomes are likely to be predicted by the model. $R^2$ is a statistic that will give some information about the goodness of

  t of a model.
- In regression, the $R^2$ coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An $R^2$ of 1.0 indicates that the regression line perfectly ts the data.
- In the case of simple linear regression, the coefficient of determination is equivalent to the squared value of the Pearson correlation coefficient. (Consider this to be co-incidental, rather than a definition)

## The Adjusted Coefficient of Determination

- Adjusted $R^2$ (often written as and pronounced "R bar squared") is a modi

  cation of $R^2$ that adjusts for the number of predictor terms in a model. Adjusted $R^2$ is used to compensate for the addition of variables to the model. As more independent variables are added to the regression model, unadjusted $R^2$ will generally increase but there will never be a decrease.
- This will occur even when the additional variables do little to help explain the dependent variable.
- To compensate for this, adjusted $R^2$ is corrected for the number of independent variables in the model, increases only if the new term improves the model more than would be expected by chance. If too many predictor variables are being used, this will be reflected in a reduced adjusted $R^2$.
- The adjusted $R^2$ can be negative (unlikely, but not impossible), and will always be less than or equal to $R^2$.
- The result is an adjusted $R^2$ than can go up or down depending on whether the addition of another variable adds or does not add to the explanatory power of the model. Adjusted $R^2$ will always be lower than unadjusted.
- Adjusted R square is generally considered to be a more accurate goodness-of-

  t measure than R square. It has become standard practice to report the adjusted $R^2$, especially when there are multiple models presented with varying numbers of independent variables.

## The Coefficient of Determination

- The coefficient of determination R2 is used in the context of statistical models whose main purpose is the prediction of future outcomes on the basis of other related information.
- It is the proportion of variability in a data set that is accounted for by the statistical model.
- It provides a measure of how well future outcomes are likely to be predicted by the model. R2 is a statistic that will give some information about the goodness of t of a model.
- In regression, the R2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R2 of 1.0 indicates that the regression line perfectly ts the data.
- In the case of simple linear regression, the coefficient of determination is equivalent to the squared value of the Pearson correlation coefficient. (Consider this to be co-incidental, rather than a definition) The Adjusted Coefficient of Determination
- Adjusted R2 (often written as and pronounced "R bar squared") is a modi cation of R2 that adjusts for the number of predictor terms in a model. Adjusted R2 is used to compensate for the addition of variables to the model. As more independent variables are added to the regression model, unadjusted R2 will generally increase but there will never be a decrease.
- This will occur even when the additional variables do little to help explain the dependent variable.
- To compensate for this, adjusted R2 is corrected for the number of independent variables in the model, increases only if the new term improves the model more than would be expected by chance. If too many predictor variables are being used, this will be re ected in a reduced adjusted R2.
- The adjusted R2 can be negative (unlikely, but not impossible), and will always be less than or equal to R2.
- The result is an adjusted R2 than can go up or down depending on whether the addition of another variable adds or does not add to the explanatory power of the model. Adjusted R2 will always be lower than unadjusted.
- Adjusted R square is generally considered to be a more accurate goodness-of- t measure than R square. It has become standard practice to report the adjusted R2, especially when there are multiple models presented with varying numbers of independent variables.