15 Sept 2018 – Linear Regression

NLP –

Corpus - word2vec

Images are pixels – each pixel has R, B and G levels – 3 matrices overlayed on each other

Get\_dummies – one-hot encoding

Ordinal categories (grades A,B,C etc) might not need get\_dummies

**Interpreting the intercept (β0):**

* It is the value of y when all independent variables are 0.
* Here, it is the estimated number of rentals when the temperature is 0 degrees Celsius.
* **Note:** It does not always make sense to interpret the intercept. (Why?)

**Interpreting the "temp" coefficient (β1):**

* **Interpretation:** An increase of 1 degree Celcius is *associated with* increasing the number of total rentals by β1.
* Here, a temperature increase of 1 degree Celsius is *associated with* a rental increase of 9.17 bikes.
* This is not a statement of causation.
* β1 would be **negative** if an increase in temperature was associated with a decrease in total rentals.
* β1 would be **zero** if temperature is not associated with total rentals.

sklearn.preprocessing.StandardScalar