**API - Challenge**

James Milne

USC – Data Science Bootcamp

Three observable trends based upon the data:

1. City latitude vs Max temperature scatter plot data appears to be non-linearly related. Max temperature might even be related to the square of latitude. Note that this relationship is for cities in both hemispheres.
2. If we split cities into northern and southern hemispheres, there is a strong negative linear relationship (R2 ~ 0.7) between maximum temperature and latitude when a regression line is passed through the northern hemisphere datapoints. Southern hemisphere cities have a weakly positive relationship (R2 = 0.3) between latitude and maximum temperature. In both cases, maximum temperature falls the further away we are from the equator. It isn’t clear what is driving this. Possible explanations are that more of the world’s landmass (and therefore more cities) is concentrated in the northern hemisphere so there are more datapoints to build a stronger relationship.
3. Both northern and southern hemispheres exhibit weakly positive relationships when latitude is regressed against humidity (R2 ~ 0.1 for northern hemisphere and R2 ~ 0.2 for southern hemisphere). The low R2 values indicate a poor explanatory power of humidity due to latitude. There may be other factors involved e.g. forest coverage, presence/absence of water, which may yield better explanatory power.