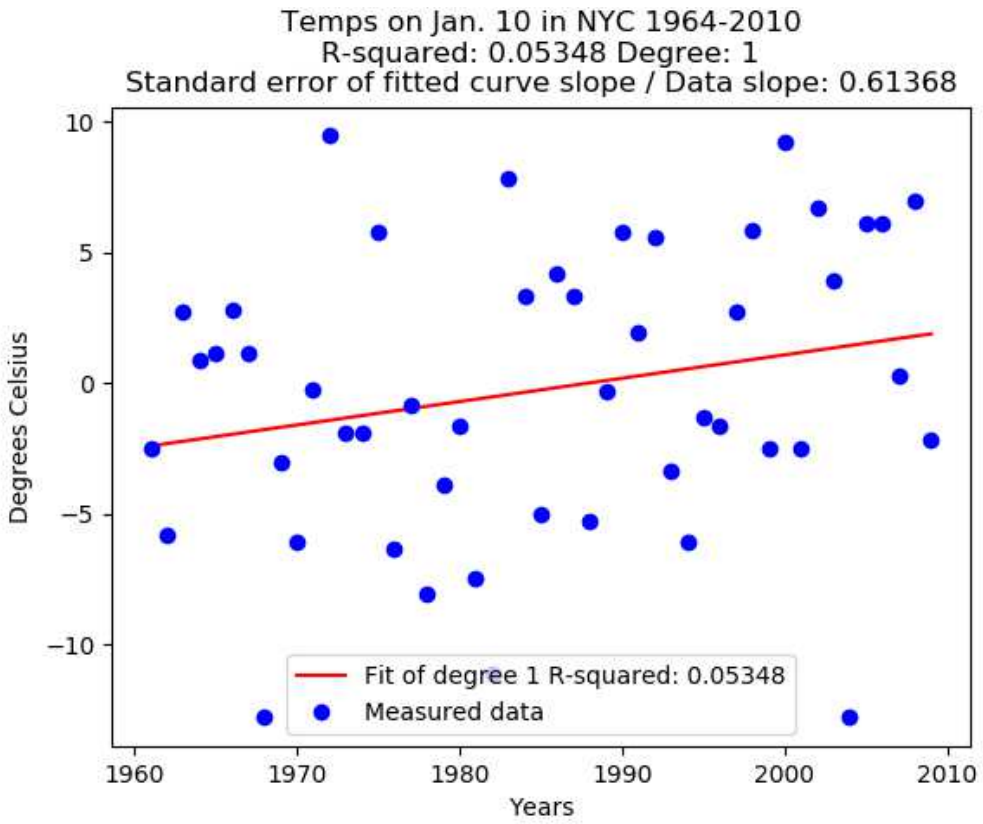
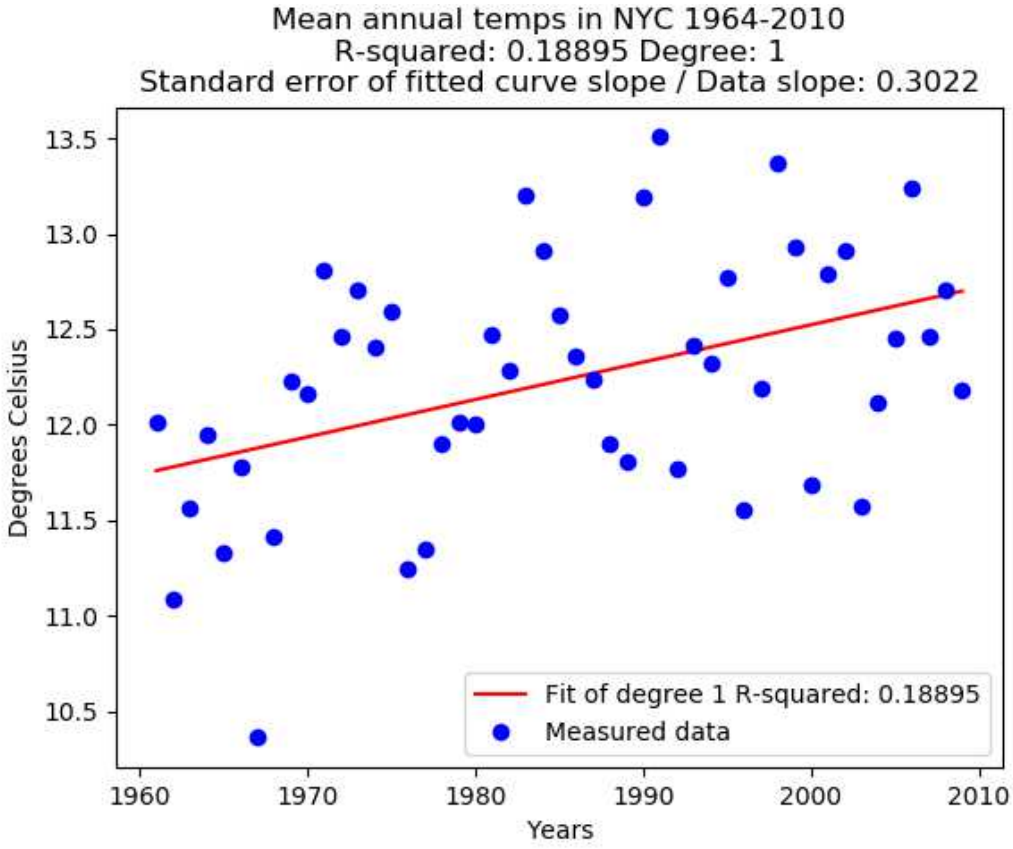
Ethan Fulbright

# 6.0002 – Problem set 5

## A 4.1




## A 4.2



## A4.1 / A4.2 writeup

**What difference does choosing a specific day to plot the data for versus calculating the yearly average have on our graphs (i.e., in terms of the $R^2$ values and the fit of the resulting curves)? Interpret the results.**

*Using the yearly average seems to smooth out some of the noise in the data, resulting in a higher $R^2$ and lower standard error to slope ratio. As a result, the resulting curves are a tighter fit (although, due to the high noise, the $R^2$ value is still quite low). This is because the yearly data consists of mean values, which results in statistical outliers influencing the overall data less.*

**Why do you think these graphs are so noisy? Which one is more noisy?**

*The daily data is much noisier than the annual mean data. I believe the noise is present for a variety of reasons, including experimental and equipment variability as well as the fact that weather is extremely variable and subject to wide change based on many different factors. The mean annual data smooths out some of this noise because taking the mean for an entire year removes some of the outlier data points. Another reason for the observed noise is that the Y-axis values are relatively small – although the graph appears noisy, it's important to note that the actual variation in the annual data is less than 3° Celsius.*

**How do these graphs support or contradict the claim that global warming is leading to an increase in temperature? The slope and the standard error-to-slope ratio could be helpful in thinking about this.**

*They definitely support the idea that global warming is leading to an increase in temperature – the fitted curves in both cases show a significant upward slope, indicating that temperatures are increasing. Although the $R^2$ values are low, the standard error to slope ratio (i.e., the mean distance between the data points and the fitted curve) is quite low in both cases, meaning that the curve is a relatively good fit for the provided data points (since the average distance those data points lie from the fitted curve is less than 1° Celsius in both graphs).*