Lecture 10

Statistics Continued: Error

Not This Kind of Error

```
6 adult_df([[adult_df['workclass'] == '?'],
7           [adult_df['occupation'] == '?'],
8           [adult_df['native_country'] == '?']])
```

```
---------------------------------------------------------------
TypeError                                 Traceback (mos
<ipython-input-24-e2ff9883893b> in <module>()
      4 adult_df([[adult_df['workclass'] == '?'],
      5           [adult_df['occupation'] == '?'],
----> 6           [adult_df['native_country'] == '?']])

TypeError: 'DataFrame' object is not callable
```

Although, here is a some info anyway

OverflowError	Means the output is too large Ex. 5.2**1000
TypeError	Means an operation was applied to an object that it does not work with passing a list into a function when it expects an integer
IndexError	Out of range Ex. ‘hello’[10]
NameError	Variable isn’t defined x + 2
FileNotFoundError	File is most likely not downloaded to the right path or not downloaded at all Ex. open(‘ajsclos.txt’)

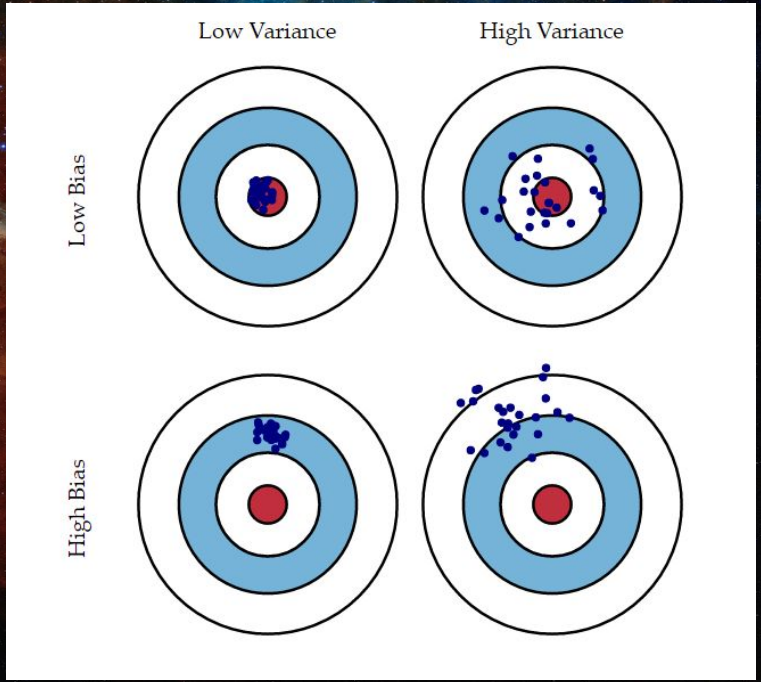
Last bit about coding errors

To learn more about errors here is a link to an error documentation:

https://docs.python.org/3/library/exceptions.html

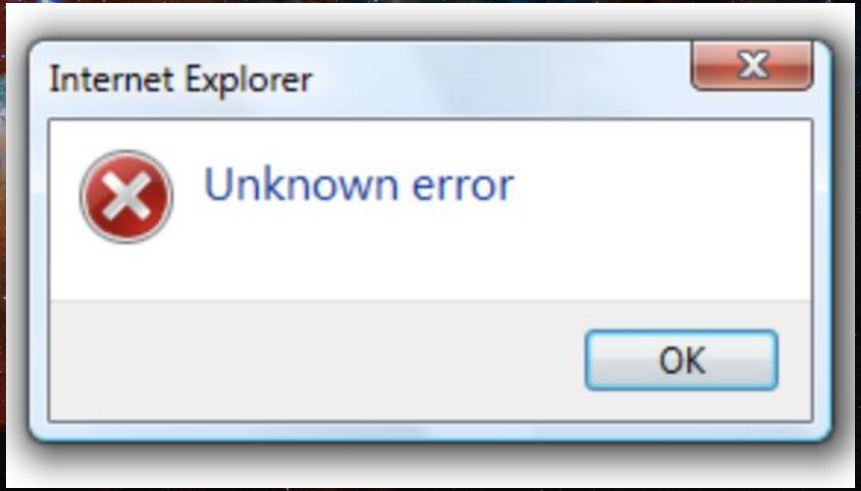
Error and Uncertainty

- Error is the quantitative measure of how much your measured values deviate from a standard or expected value.
- For example, if someone asked you to provide the mileage on your car, you might say that it is 45,000 miles, plus or minus 500 miles.
- All measurements contain some amount of uncertainty.



Types of Errors

- Observational Error
  - The observational uncertainty results from a judgment call made during an observation. You can estimate it by taking half the difference between the bounds of your judgment.
- Counting Error
  - When dealing with occurrence counts of a random process (such as the decay of a radioisotope) there is an inherent statistical counting uncertainty that grows as the square root of your count. That is, given a count of N, the counting uncertainty is .



Types of Uncertainties

- Reading Uncertainties
  - An uncertainty that is due to the finite resolution of our instruments. The reading uncertainty Δy read is plus-or-minus one-half the resolution of the measurement.
- Random Uncertainties
  - An uncertainty that is due to random fluctuations in our instruments or readings (very common with digital instruments).

Standard Deviation vs Standard Error

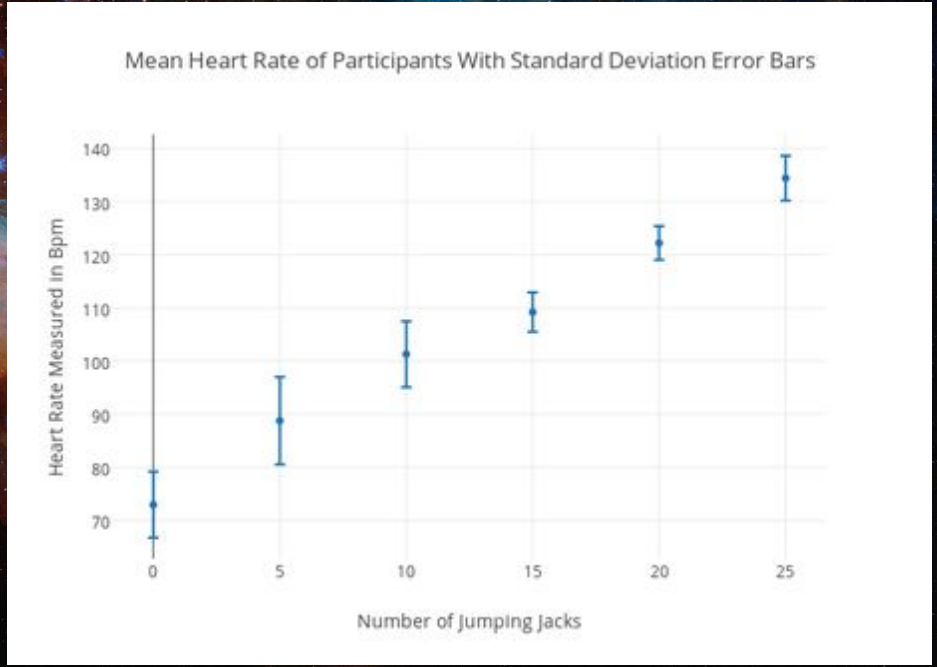
- The standard deviation represents the uncertainty of a single measurement and the standard error represents the uncertainty in the mean of multiple measurements.
- Standard Error is defined by Standard deviation and the size of your dataset

$$\mathrm{SD} = \sqrt{\frac{\sum |x - \bar{x}|^2}{n}}$$

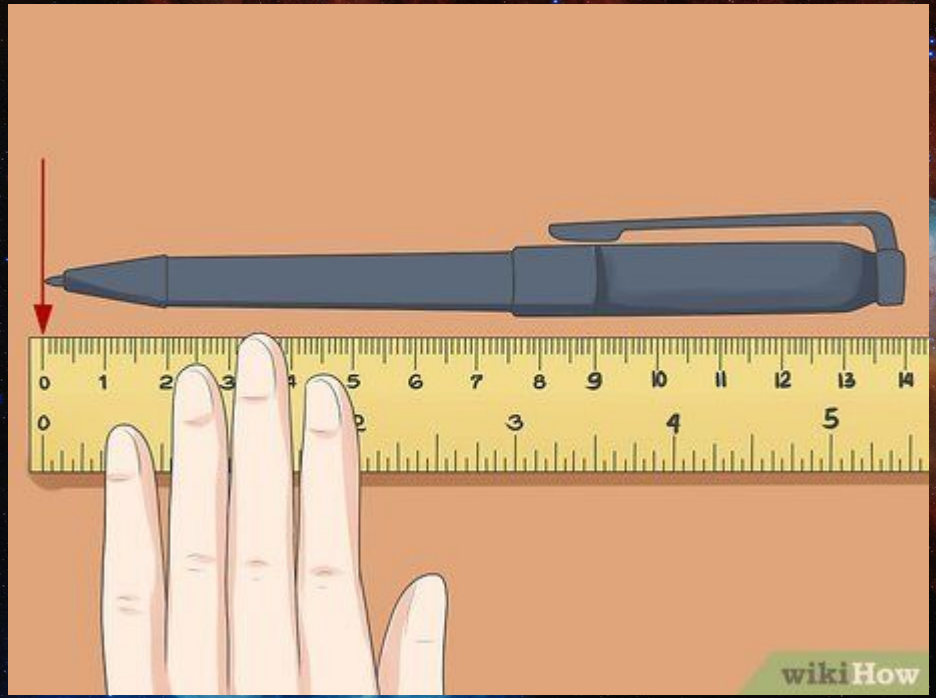
$$SE = \frac{\sigma}{\sqrt{n}}$$

Reporting Error in Python

- New function: plt.errorbar
  - Works in a very similar way to plt.plot
- Adds two new argument to apply to each data point
  - xerr & yerr
- These can vary per data point or just be one value applied to the whole dataset



How will you use Error

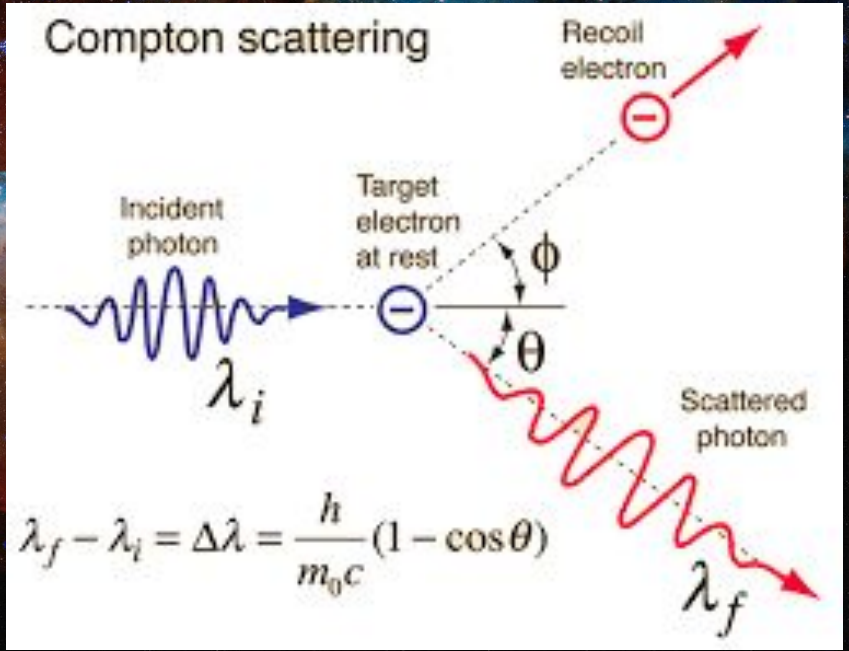


- As we have gone through there are many ways to calculate error and a lot of different types of error to keep track of
- Your job is to do your best keeping track of the different errors that could be introduced in your experiment to report as uncertainty
- When all else fails: report std or std err as your uncertainty


Coding Demo Time

Compton Scattering

- Compton Scattering is a quantum phenomenon of a photon scattering of an electron
- Once the photon scatters off of the electron, it loses energy and scatters off at a new angle
- The difference in the energy of the photon before and after it scatters can be directly measured by the scattering angle




Back to the Coding Demo