1. At the bottom of the page, there are 5 data sets (only a fraction of each is shown). Descriptions of each data set and goals of our analysis are also provided. For **each** of these data sets, write a brief description of an effective figure you could make to display these data, including the type of figure, any visual elements to include, and factors to consider in your design to enable the goals described.

   (a) Measurements of the kinetic energy and potential energy (both in J) at many different times for a single golf ball as it flies through the air. *Goal: understand how these types of energy change over time during flight.*

   (b) Measurements from tests of golf clubs, including speed of the golf club head during its swing and the resulting distance traveled by the ball after being struck. *Goal: understand the relationship between swing speed and distance.*

   (c) The same as the previous, but with two different brands of golf club. *Goal: understand whether the brand of golf club affects the relationship between swing speed and distance.*

   (d) Measurements of deflection angle (θ, in degrees) and wavelength shift ($\Delta\lambda$, in pm) for X-rays scattered off of a piece of graphite, as well as uncertainties on each quantity. *Goal: check whether our measurements agree with a theoretical model $\Delta\lambda(\theta) = A(1 - \cos\theta)$, where A is a constant.*

   (e) NASA-funded telescopes with costs in USD and primary type of light studied by each. *Goal: understand NASA's investment in different types of observations.*

(a)

	t	K	Ug
0	0.00	47.52	0.00
1	0.10	47.05	0.47
2	0.20	46.62	0.90
3	0.30	46.23	1.29
4	0.40	45.89	1.63
5	0.50	45.59	1.93
6	0.60	45.34	2.18
7	0.70	45.13	2.39
8	0.80	44.97	2.55
9	0.90	44.85	2.67
...	...	...	...

(b)

	v_head	dist
0	80	238
1	97	252
2	107	281
3	91	258
4	92	224
5	114	323
6	100	279
7	97	238
8	104	269
9	85	208
...	...	...

(c)

	brand	v_head	dist
0	Turtleist	104	279
1	Turtleist	98	234
2	Turtleist	81	209
3	Ballyway	113	289
4	Ballyway	100	278
5	Turtleist	91	197
6	Ballyway	102	250
7	Ballyway	104	249
8	Turtleist	100	281
9	Ballyway	87	212
...	...	...	...

(d)

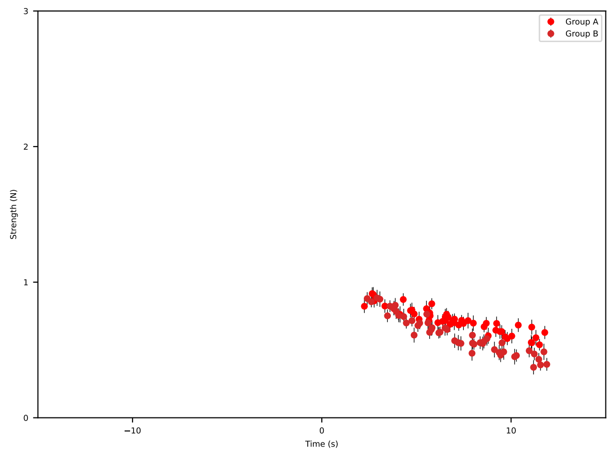
	theta	unc_theta	dlambda	unc_dlambda
0	41.9	1.1	0.98	0.31
1	5.9	1.8	0.15	0.35
2	62.5	1.1	1.31	0.24
3	44.3	1.8	1.02	0.40
4	16.9	1.2	0.27	0.41
5	40.0	1.7	0.53	0.39
6	15.8	2.0	0.50	0.42
7	44.5	1.2	0.62	0.38
8	9.5	1.6	0.16	0.45
9	139.1	1.7	4.05	0.49
...	...	...	...	...

(e)

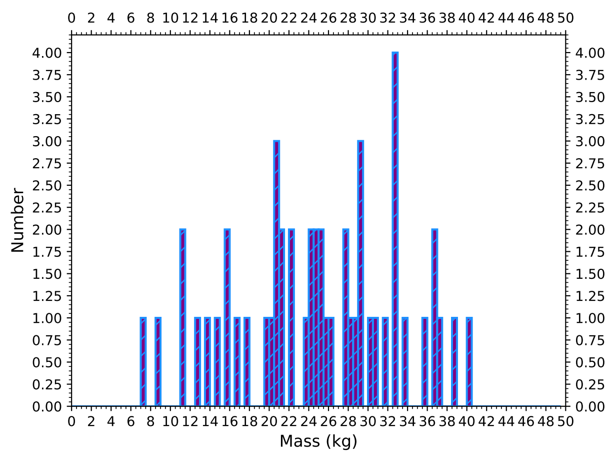
	project	type	cost
0	Hubble	Optical	16e9
1	JWST	Infrared	9.7e9
2	Chandra	X-ray	1.6e9
3	NuSTAR	X-ray	180e6
4	Fermi	Gamma-ray	2.5e9
5	IXPE	X-ray	188e6
6	Swift	X-ray	530e6
7	GALEX	UV	150e6
8	TESS	Optical	287e6
9	Spitzer	Infrared	1.36e9
...	...	...	...

2. For each of the following figures, identify at least three ways in which they could be improved.

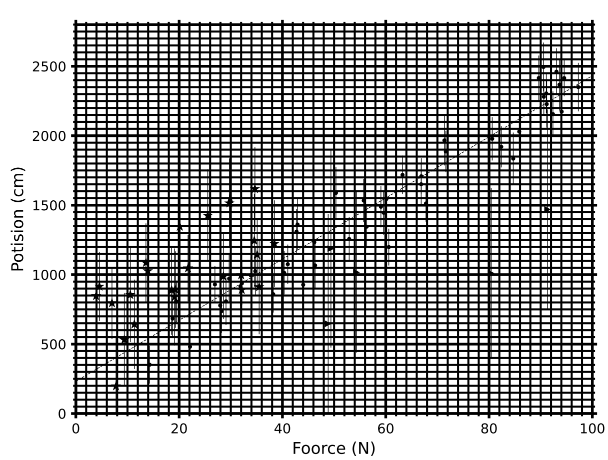
(a)



(b)



(c)



3. Imagine that you have a set of N measured values $\{x\} = x_1, x_2, ..., x_N$.

   (a) The mean value of $\{x\}$ is defined as $\bar{x} \equiv \frac{1}{N}\sum_{i=1}^{N} x_i$. If we define the set of values $\{y\}$ by applying the formula $y_i = ax_i + b$ for each x_i (where a and b are constants), show that the mean value of $\{y\}$ will satisfy $\bar{y} = a\bar{x} + b$.

   (b) The sample standard deviation of $\{x\}$ is defined as $\sigma_x \equiv \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$. If we define $\{y\}$ as in (b) above, show that the sample standard deviation of $\{y\}$ will satisfy $\sigma_y = a\sigma_x$.

4. The following questions are a review of ways we report and propagate uncertainties in physics, which you hopefully saw in PHY111 or other physics labs. You may wish to review the relevant sections of the PHY111 lab manual: Sec. 2.6 for significant figures, and Secs. 10.1-10.2 for error propagation. However, see the references below for the more detailed descriptions of these methods in our textbook.

   (a) **Significant figures:** For each of the values and uncertainties below, round the value and uncertainty appropriately based on methods described in Taylor's Ch. 2.2.

      i. $29.74892241 \pm 0.00139745\,\mathrm{s}$

      ii. $1085.2308563845 \pm 53.0002\,\mathrm{kg}$

      iii. $3.721 \times 10^6 \pm 4.2 \times 10^4\,\mathrm{N}$

      iv. $4.7437664\,\mu\mathrm{m} \pm 12.75\,\text{Å}$

   (b) **Addition & Subtraction:** For each of the calculations below, calculate the result (including the uncertainty with correct significant figures) in two cases: assuming the maximum possible uncertainty, and assuming the expected uncertainty if the individual measurements are independent. See Taylor 3.3 and 3.5.

      i. The total height above the ground of a chair ($h_c = 1.421 \pm 0.008\,\mathrm{m}$) on top of a table ($h_t = 0.953 \pm 0.009\,\mathrm{m}$).

      ii. The difference in finishing time between two swimmers who completed their laps in $t_A = 52.34\,\mathrm{s}$ and $t_B = 52.39\,\mathrm{s}$, respectively. The timing equipment has an uncertainty of $0.03\,\mathrm{s}$.

      iii. The total mass of a collection of pebbles with masses $5.2 \pm 0.4\,\mathrm{g}$, $10.3 \pm 0.5\,\mathrm{g}$, $7.1 \pm 0.4\,\mathrm{g}$, $12.6 \pm 0.5\,\mathrm{g}$, and $17.878 \pm 0.005\,\mathrm{g}$.

   (c) **Types of error:** Give an example of a scenario where it is appropriate to assume that the errors on several added values are independent, as well as an example where assuming the maximum possible error is more appropriate.

5. Complete the Introduction to Python 1 and 2 notebooks linked on Canvas.