Discrete Probability Distribution of Penguin Weights

I created a discrete probability distribution (Figure 1) and histogram (Figure 2) for the weights of penguins in the PalmerPenguins dataset. The mean weight in the data is approximately equal to 4202 grams, with P(4202) approximately equal to 3.5%. The probabilities and weights are not normally distributed.

A tibble: 94 × 2

weights	probabilities
<dbl>	<dbl>
2700	0.0029
2850	0.0058
2900	0.0117
2925	0.0029
2975	0.0029
3000	0.0058
3050	0.0117
3075	0.0029
3100	0.0029
3150	0.0117
3175	0.0058
3200	0.0146
3250	0.0146
3275	0.0029
3300	0.0175
3325	0.0146
3350	0.0146
3400	0.0234
3425	0.0058
3450	0.0234
3475	0.0088
3500	0.0205
3525	0.0058
3550	0.0263
3575	0.0029
3600	0.0205
3625	0.0029
3650	0.0175
3675	0.0058
3700	0.0322
⋮	⋮
4775	0.0029
4800	0.0088
4850	0.0117
4875	0.0088
4900	0.0058
4925	0.0058
4950	0.0058
4975	0.0029
5000	0.0175
5050	0.0088
5100	0.0088
5150	0.0058
5200	0.0117
5250	0.0088
5300	0.0117
5350	0.0088
5400	0.0146
5450	0.0029
5500	0.0146
5550	0.0175
5600	0.0058
5650	0.0088
5700	0.0146
5750	0.0029
5800	0.0058
5850	0.0088
5950	0.0058
6000	0.0058
6050	0.0029
6300	0.0029

Figure 1

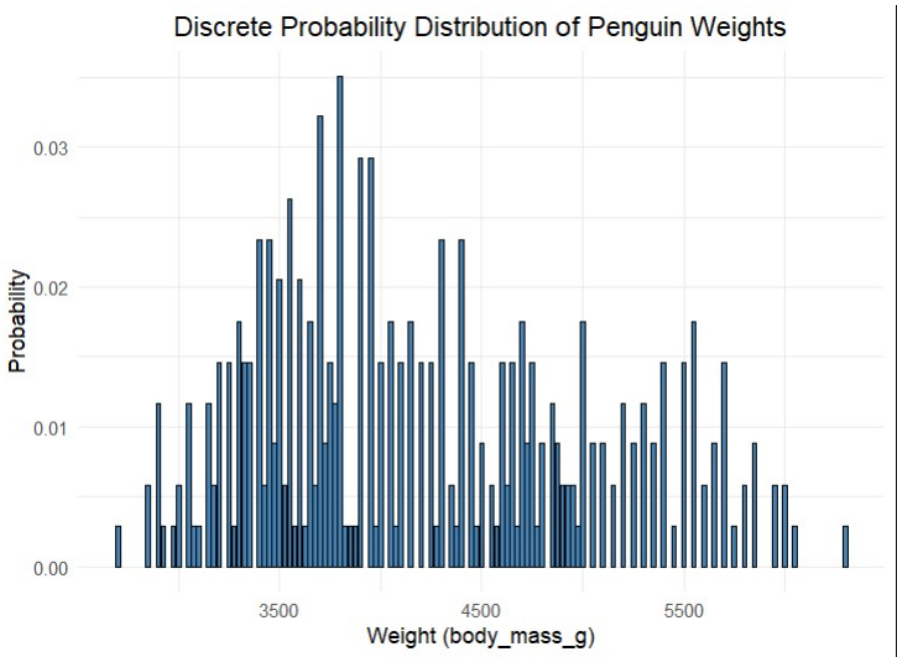

Figure 2

Since this initial visualization is a little challenging to interpret, I also created a density plot of weight (body_mass_g) across the dataset (Figure 3).

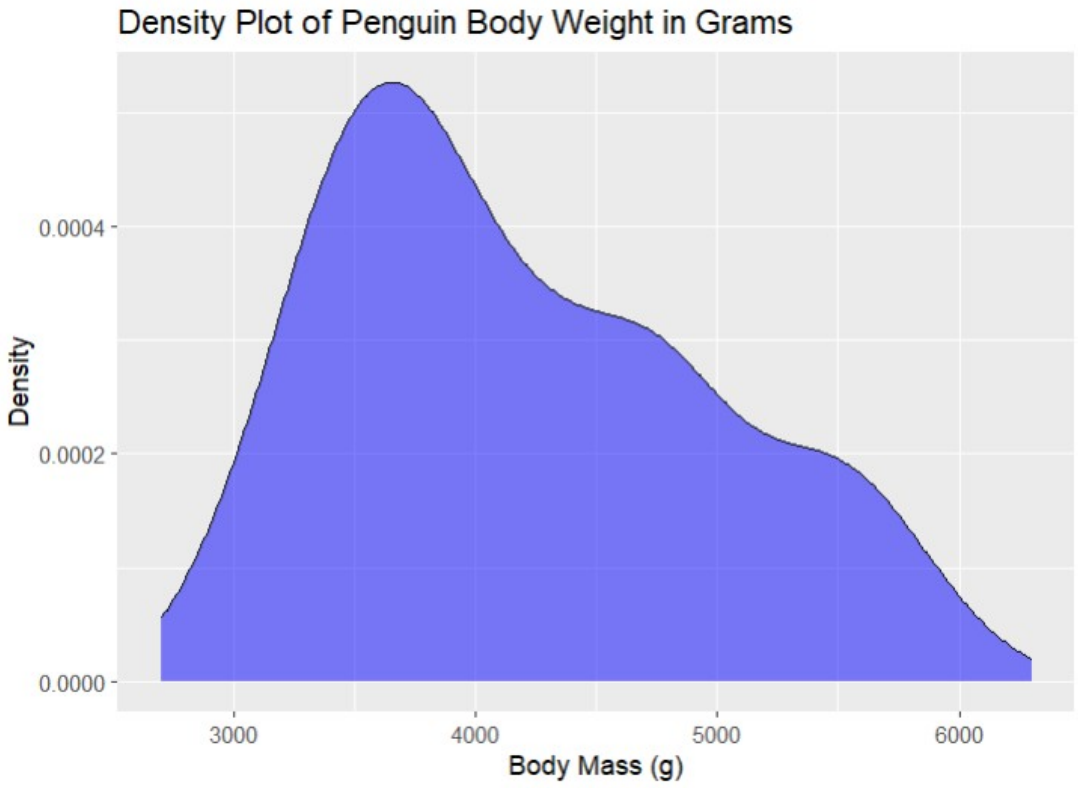

Figure 3

Binomial Probability Distribution and Histogram

I made a binomial probability distribution (and histogram) for the species Adelie in the PalmerPenguins dataset. See Figure 4. This shows the probabilities of the first 1-10 penguins selected being the species Adelie.

A tibble: 10 × 2

x	binomial_probabilities
<int>	<dbl>
1	0.0232
2	0.0827
3	0.1747
4	0.2420
5	0.2299
6	0.1517
7	0.0686
8	0.0204
9	0.0036
10	0.0003

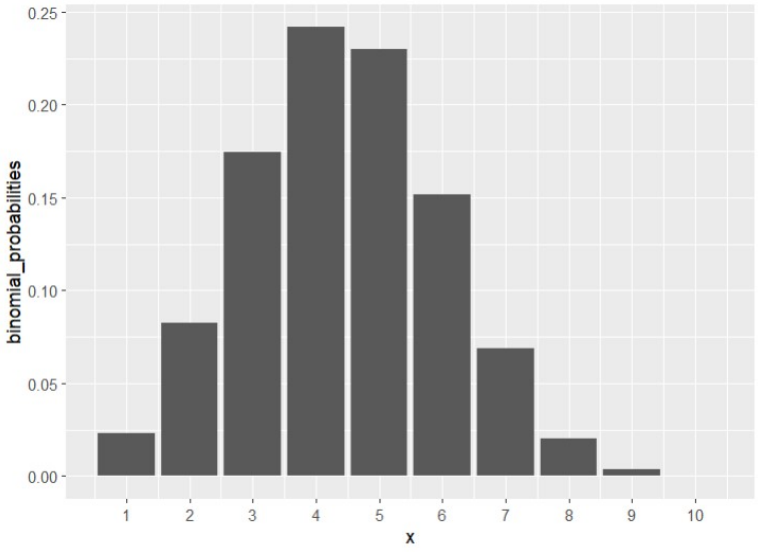

Figure 4

Table:

- The first column, labeled "x," represents the number of successes in the binomial distribution. In this context, "success" could be defined as a particular characteristic or outcome being observed in a penguin.
- The second column, labeled "binomial_probabilities," shows the probability of observing exactly x number of successes. For example, the probability of observing exactly 1 success (one penguin with the characteristic) is 0.0323, and the probability of observing 2 successes is 0.0827, and so on.

Visualization:

- The histogram visualizes the data from the table, with the x-axis representing the number of successes and the y-axis representing the probability of that number of successes.
- Each bar's height corresponds to the probability of that number of successes occurring. For instance, the tallest bar at x=4 corresponds to the highest probability (about 0.2420) among the probabilities listed in the table, meaning that observing exactly 4 successes is the most probable outcome within this dataset.
- This visualization is particularly useful in showing the distribution of probabilities across different numbers of successes. It provides a quick way to see which outcomes are most and least likely and how the probabilities are distributed (e.g., if they are skewed towards lower or higher numbers of successes).

Combined Interpretation:

- Together, the table and histogram provide a comprehensive view of the binomial probability distribution for the specified characteristic within the Adelie species. They illustrate not only the individual probabilities for each number of successes but also the overall shape of the distribution.
- This kind of analysis is valuable for understanding variability and the likelihood of different outcomes within a population. For researchers, it can inform predictions about the Adelie species and contribute to broader ecological and biological studies.