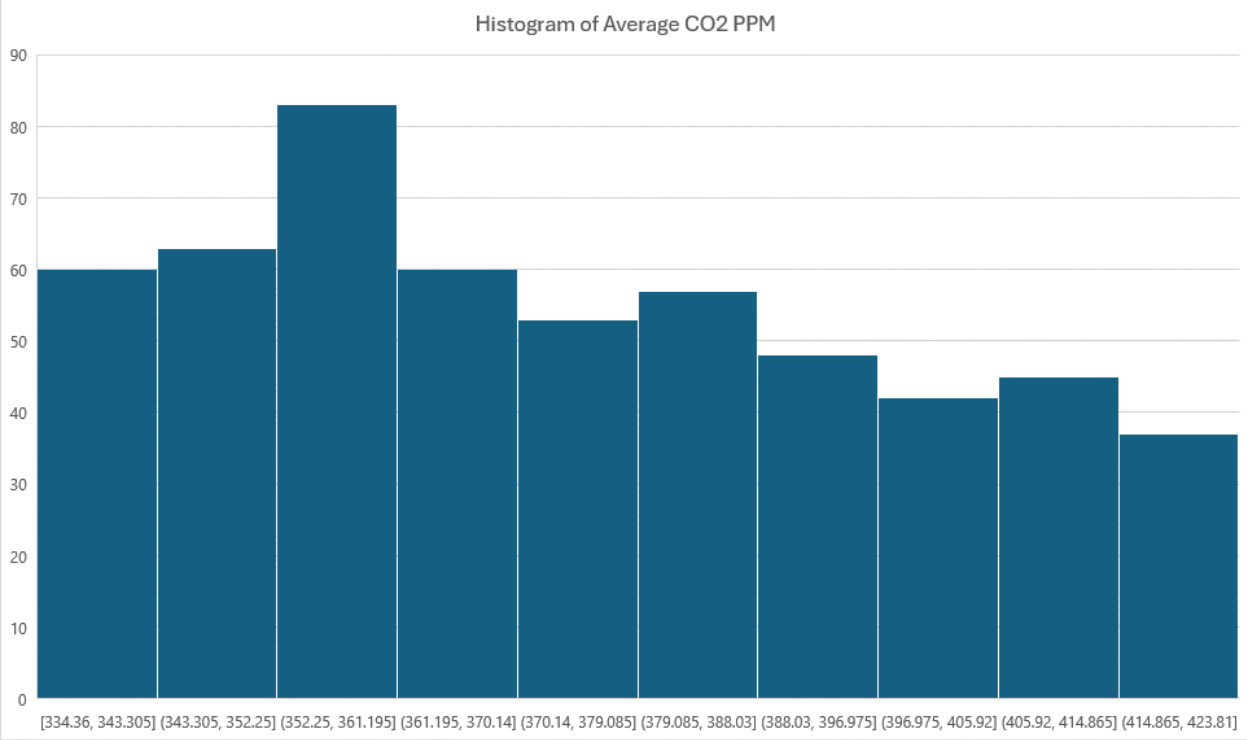
Intro:

The dataset I chose to use was a measurement over every month from 1979-2024. In this dataset, there are 4 columns: Date, Decimalized version of the date, Average, and Trend. I will be using this dataset to form problems for each section below.

Chapter 1

Section 1.2 – Histogram

Create a histogram of the "Average" CO2 ppm values from the dataset. What does this histogram tell you?



This histogram shows that there seems to be a pretty evenly spread window of these averages, with more of the averages reflecting lower numbers.

Section 1.3 – Numerical Methods

Calculate the mean, median, and standard deviation of the "Average" CO2 ppm values for the dataset.

mean	374.5002
median	371.295
Standard Dev	24.77563

Using my excel functions, I was able to find the mean, median, and standard deviation.

Chapter 2

Section 2.2 - Probability and Inference

If we were to randomly select a month from the dataset, what is the probability that the CO2 ppm value is above 340? Use the distribution of CO2 ppm values to make an inference.

To do this, we have to count the number of months where the average is above 340, then divide by the total number of entries in the dataset.

CO2 PPM > 340	517
Probability:	0.937876

Section 2.3 - Set Notation

Let A be the set of months where the "Average" CO2 ppm value exceeds 338, and let B be the set of months where the trend is greater than 0.1. What is the intersection of sets A and B? In other words, what is the probability that the CO2 ppm is above 338 and the trend is greater than 0.1?

First I made a helper column to indicate which numbers met these requirements. I took the sum of the set and got the probability by dividing by the total number of numbers in the set.

Probability:	0.306569

Section 2.4 - A Probabilistic Model for an Experiment: The Discrete Case

Assume we have a discrete probabilistic model based on the CO2 ppm levels, where each month is an experiment. What is the probability of observing a month with a "Trend" value greater than 0.2?

I used the following excel function to express this:

```
=COUNTIF(D2:D549, ">0.2") / COUNTA(D2:D549)
```

Probability:	0.007299

Section 2.5 - The Sample-Point Method

Using the sample-point method, list all possible outcomes for the "Trend" values (e.g., less than 0.1, between 0.1 and 0.2, and greater than 0.2). What is the probability of each outcome?

I used another helper column to display the results of whether each was in the specified range. I then made a frequency table and then calculated the probabilities.

		Probabilities
"Less than 0.1"	270	0.49270073
"0.1 to 0.2"	274	0.5
"Greater than 0.2	4	0.00729927

Section 2.7 - Conditional Probability and the Independence of Events

Given that a month has a "Trend" greater than 0.1, what is the conditional probability that the "Average" CO2 ppm value is greater than 338? Are the events "Trend > 0.1" and "Average CO2 ppm > 338" independent?

First I counted how many match the conditions for both sets, then got the probability of the two:

Set A (Both Conditions Met)	168
Set B (Trend > 0.1)	180
Probability:	0.933333

Section 2.9 - The Event-Composition Method

Use the event-composition method to find the probability that in a randomly selected month, either the "Average" CO2 ppm is greater than 338 or the "Trend" is greater than 0.1.

Using this function in excel, I was able to reach my answer:

```
=(COUNTIFS(C2:C549, ">338") + COUNTIFS(D2:D549, ">0.1") - COUNTIFS(C2:C549, ">338", D2:D549, ">0.1")) / COUNTA(C2:C549)
```

Probability:	0.99635

Section - 2.10 The Law of Total Probability and Bayes' Rule

Use Bayes' Rule to calculate the probability that a randomly selected month has a "Trend" greater than 0.1 given that the "Average" CO2 ppm is greater than 338.

Using Baye's Rule in the function shown below, I was able to get my answer:

```
=COUNTIFS(C2:C549, ">338", D2:D549, ">0.1") / COUNTIF(C2:C549, ">338")
```

Baye's Rule:	0.314607

Chapter 3

Section 3.2 - The Probability Distribution for a Discrete Random Variable

Treat the "Trend" values as a discrete random variable. Create the probability distribution for the "Trend" values and plot it. What does the distribution suggest about the variability in trends?

Section 3.3 - The Expected Value of a Random Variable or a Function of a Random Variable

Calculate the expected value of the "Trend" variable from the dataset. What does this tell you about the average change in CO2 ppm over the months?

Expected:	0.099088

It is shown on average from the trend that the CO2 PPM increases 0.099 on average each month.

Section 3.4 - The Binomial Probability Distribution

Assume that the probability of a "Trend" value being greater than 0.1 is p. If you randomly select 5 months, what is the probability that exactly 3 months have a "Trend" greater than 0.1?

First, I got the probability of the trend having a value greater than 0.1. After this, I used the binomial distribution function in excel to get the answer.

Probability of Trend > 0.1	0.328467
Binom. Dist:	0.159812

```
=BINOM.DIST(3, 5, H28, FALSE)
```

Where H28 is the probability.

Section 3.5 - The Geometric Probability Distribution

Assume the probability of a "Trend" greater than 0.1 is p. What is the probability that the first month in which the "Trend" exceeds 0.1 is the 4th month?

Using the probability from the previous question, I was able to use the function below to derive the answer:

Probability of Trend > 0.1	0.328467
Binom. Dist:	0.159812
Geometric Prob Dist.	0.09947

```
=((1-H28)^3) * H28
```

Where H28 is the probability.

Section 3.6 - The Negative Binomial Probability Distribution

Suppose you are counting the number of months needed before you observe 3 months with a "Trend" greater than 0.1. What is the probability distribution for this scenario?

Again, using the same probability from Section 2.4, I was able to derive this answer using the Negative Binomial Distribution function in excel:

Probability of Trend > 0.1	0.328467
Binom. Dist:	0.159812
Geometric Prob Dist.	0.09947
Negative Binomial Prob Dist.	0.101632

```
=NEGBINOM.DIST(5, 3, H28, FALSE)
```

Where H28 is, again, the probability.

Section 3.7 - The Hypergeometric Probability Distribution

From a sample of 12 months, 5 months have a "Trend" greater than 0.1. If you randomly select 3 months without replacement, what is the probability that exactly 2 of them have a "Trend" greater than 0.1?

Using the Hypergeometric distribution function in excel, I was able to solve this:

Hypergeometric Dist.	0.318182	=HYPGEOM.DIST(2, 3, 5, 12, FALSE)

Section 3.8 - The Poisson Probability Distribution

Assume the "Trend" values follow a Poisson distribution with a mean of λ = 0.1. What is the probability that in a randomly selected month, the "Trend" value is exactly 0.15?

Using the Poisson Probability function in excel, plugging in the values I was able to come up with this answer.

Poisson Dist.	0.904837	=POISSON.DIST(0.15, 0.1, FALSE)

Section 3.9 - Moments and Moment-Generating Functions

Calculate the first and second moments (mean and variance) of the "Trend" variable. Then, use the moment-generating function to derive the distribution.

Chapter 4

Section 4.2 - The Probability Distribution for a Continuous Random Variable

Assume the "Average" CO2 ppm values follow a continuous probability distribution. Estimate the probability that the "Average" CO2 ppm is between 337 and 340.

Using the excel function below, I was able to derive the answer:

Prob. Dist. For Continuous Random Var.	0.01682

```
=NORM.DIST(340, H2, H4, TRUE) - NORM.DIST(337, H2, H4, TRUE)
```

Where H2 is the Mean of "Averages" and H4 is the Standard Deviation of "Averages".

Section 4.3 - Expected Values for Continuous Random Variables

Calculate the expected value of the "Average" CO2 ppm using the probability density function (PDF) for the dataset. What does this value represent?

Section 4.4 - The Uniform Probability Distribution

Assume that the "Average" CO2 ppm values follow a uniform distribution between 334 and 341. What is the probability that a randomly selected month has a CO2 ppm value less than 336?

I was able to calculate this in excel using this function:

Uniform Prob. Dist:	0.285714	`=(336-334)/(341-334)`

Section 4.5 - The Normal Probability Distribution

Assume the "Average" CO2 ppm values are normally distributed with a mean of 338 and a standard deviation of 1. Calculate the probability that a randomly selected month has a CO2 ppm greater than 340.

I was able to calculate this in excel using this function:

Normal Probability Dist.	0.02275	`=1 - NORM.DIST(340, 338, 1, TRUE)`

Section 4.6 - The Gamma Probability Distribution

Assume that the "Trend" values follow a Gamma distribution. Given the data, estimate the shape and rate parameters of the distribution. Then, calculate the probability that the "Trend" is greater than 0.2.

Section 4.7 - The Beta Probability Distribution

Assume that the "Trend" values follow a Beta distribution. Using the dataset, estimate the parameters of the Beta distribution and calculate the probability that the "Trend" is between 0.1 and 0.2.

Chapter 5

Section 5.2 - Bivariate and Multivariate Probability Distributions

Consider the "Average" CO2 ppm and "Trend" as two random variables. Plot their joint distribution and estimate the correlation between these two variables. What does the correlation tell you about the relationship between CO2 ppm levels and trends?