Modeling Coney Beach's Tides Using a Cosine Function

About the beach

Coney Island Beach in Brooklyn, New York, is a famous destination known for its wide sandy shoreline and vibrant atmosphere . Stretching nearly three miles along the Atlantic Ocean, the beach is perfect for relaxation, recreation, and seaside fun (*ChatGPT*). Visitors flock here to enjoy sunbathing, swimming, and strolling along the boardwalk (*ChatGPT*). The tides at Coney Island Beach play an important role in shaping the daily experience. At low tide, the shoreline extends further out, revealing more space for beachgoers to set up umbrellas or explore the wet sand. High tide brings the waves closer, making it perfect for swimming or simply watching the rhythmic movement of the ocean. The waters are generally calm, providing a safe environment for families. The beach is bordered by the Riegelmann Boardwalk, where visitors can find food stands, souvenir shops, and beautiful ocean views. Beyond the sand, people enjoy fishing from the nearby pier or watching boats pass by in the distance. Accessible via subway, Coney Island Beach attracts a mix of locals and tourists. The flow of the tides, combined with its location, make Coney Island Beach a unique and cherished part of New York City's coastal charm (*ChatGPT*).

Why Tide Predictions are Crucial for Boat Entry

Boat navigation and docking rely heavily on understanding tide levels, especially in areas like Coney Island where shallow waters can limit access. Tide predictions help:

- **Safe Navigation**: Boats entering or leaving harbours must account for water depth to avoid grounding. Low tides can expose sandbars and other underwater hazards that pose significant risks.

- **Efficient Scheduling**: Vessels often plan their movements based on tide schedules to minimize delays and maximize safety. Knowing when high tide occurs allows boats with deeper drafts to enter the harbour without issues

Significance of Modeling Tides with the Cosine Function

Tides are periodic and follow a predictable rhythm, influenced by the gravitational pull of the moon and the sun, as well as the Earth's rotation. The **cosine function** is useful for modeling tides because it captures the wave-like, sinusoidal pattern of rising and falling water levels over time. By fitting a cosine curve to tidal data, it becomes possible to create a mathematical model that reflects the regular intervals of high and low tides.

This approach is easy and efficient because:

- **Precision and Predictability**: By accurately modeling tides, scientists and engineers can predict future tide levels with better precision.
- **Simplicity and Efficiency**: A cosine model reduces the complexity of tide prediction to a simple mathematical equation, making it efficient for calculations without extensive computational technology.

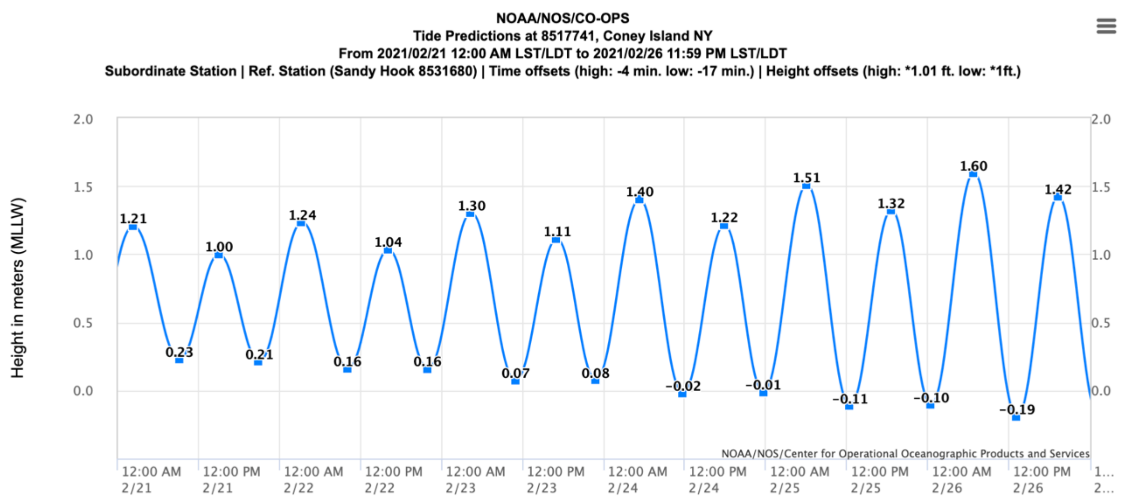


Image 1: Data taken from http://tidesandcurrents.noaa.gov/ reflecting the measurements from 21nd February 2021 – 26th February 2021 of the measurements of the high tides (m) and what time they rose and fell (AM and PM). To make the function, we will be taking the data from 22nd February 2021-24th February 2021

Day		Measurement of the tide (m)	Hours since first high tide (hours)
22nd February	High 1	1.24	0
	Low 1	0.16	6.75
	High 2	1.040	12.783
	Low 2	0.16	18.617
23rd February	High 1	1.30	25.05
	Low 1	0.07	31.6
	High 2	1.11	37.733
	Low 2	0.08	43.483
24th February	High 1	1.40	49.967
	Low 1	-0.02	56.383
	High 2	1.22	62.55
	Low 2	-0.01	68.317

Table 1: Measurement of the tides on 3 different days along with the hours since the first high tide on 22nd February 2021

Finding h=acos(bt)+c

Variables:

h=Height of the tide in metres

t=time of the tide from the first high tide on November 22nd 2021

Constants

a=amplitude

To find the amplitude, we will use $\frac{max-min}{2}$

Taking the average of the high tides (maximum) and low tides (minimum) to increase accuracy (in metres)

High 1	1.24	Low 1	0.16
High 2	1.04	Low 2	0.16
High 3	1.30	Low 3	0.07

Average of Highs (maximum):

$\frac{1.24+1.04+1.30}{3}$= 1.193

Average of Lows (minimum):

$\frac{0.16+0.16+0.07}{3}$= 0.13

Now, max=1.193 and min=0.13

$\frac{1.193-0.13}{2}$= 0.5315

a=0.5315

Finding b:

b= $\frac{360}{period}$

The period is the time in hours between every high tide and the next high tide or low tide and the next low tide. Taking the average of the periods to increase accuracy.

Taking the time between every high tide (hours):

H1	1.24
H2	1.04
H3	1.30
H4	1.11

12.783
12.267
12.683

$\frac{12.783+12.267+12.683}{3}$= 12.577

Period=12.577

b=$\frac{360}{period}$

b=$\frac{360}{12.577}$

b=28.62

Finding c

The value of c is the vertical shift

c= $\frac{max+min}{2}$

max=1.193 (refer to "Finding a")

min=0.13 (refer to "Finding a")

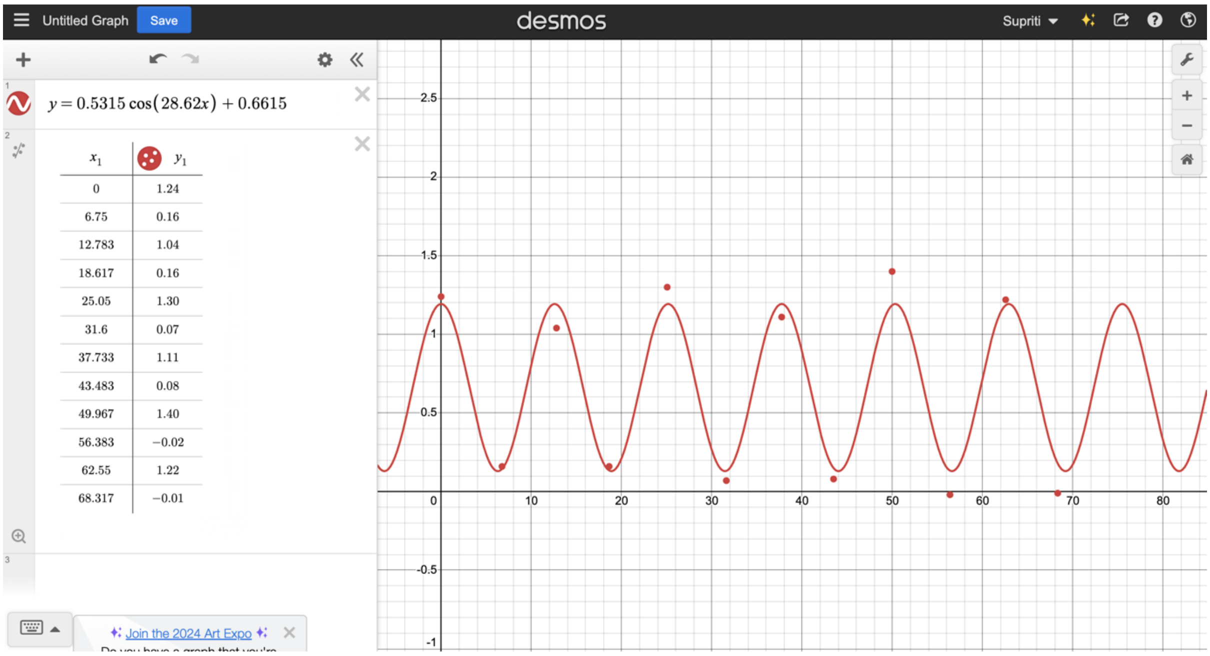
c=$\frac{1.193+0.13}{2}$

c=0.6615

Final equation:

h=0.5315 cos (28.62t) +0.6615

Let's input the equation into Desmos. Taking **t** as **x** and **h** as **y**, I made a data table with the data I was already provided with (Table 1)



Mean Absolute Error

Mean Absolute Error (MAE) measures how close a model's predictions are to the actual values. It calculates the average of the absolute differences between the predicted values and the actual values. MAE is easy to understand because it shows the average size of the errors in the same units as the data. Smaller MAE means the model's predictions are more accurate.

How to find MAE:

First, I will find the absolute error for all the different observations:

$$|Actual\ observation - graph\ projection|$$

Then,

I will find the mean of the absolute errors:

$$\frac{|Actual\ observation - graph\ projection|}{Number\ of\ observations}$$

High tides: For this, I will find the absolute error for each high tide then average it.

	Measurement of actual tides	**Graph projection**	**Absolute error**
H1	1.24	1.193	0.047
H2	1.04	1.193	0.153
H3	1.3	1.193	0.107
H4	1.11	1.193	0.083
H5	1.4	1.193	0.207

H6	1.22	1.193	0.027

Now, I fill find the average of the absolute errors:

$$\frac{0.047 + 0.153 + 0.107 + 0.083 + 0.207 + 0.027}{6}$$

=0.104

The mean absolute error for the high tides is 0.104

Low tides: For this, I will find the absolute error for each low tide then average it.

	Measurement of actual tides (m)	**Graph projection**	**Absolute error**
L1	0.16	0.13	0.03
L2	0.16	0.13	0.03
L3	0.07	0.13	0.06
L4	0.08	0.13	0.05
L5	-0.02	0.13	0.15
L6	-0.01	0.13	0.14

$$\frac{0.03 + 0.03 + 0.06 + 0.05 + 0.15 + 0.14}{6}$$

=0.077

The mean absolute error for the low tides is 0.077

Percentage Error

Percentage error measures the accuracy of predictions by expressing the error as a percentage of the actual value. It is calculated as:

$\left|\frac{\boldsymbol{estimated\ value - actual\ value}}{\boldsymbol{actual}}\right|$**x100%**

This formula gives the size of the error relative to the actual value, making it easy to compare errors across different scales or datasets. Smaller percentage errors indicate more accurate predictions.

Since I've already found the absolute error, I will divide the absolute error by the measurement of the actual tides

To calculate, I will do $\frac{\boldsymbol{Absolute\ error}}{\boldsymbol{Actual\ value}}\boldsymbol{x100}\%$

For high tides:

	Measurement of actual tides (m)	Graph projection	Absolute error	Percentage Error (%)
H1	1.24	1.193	0.047	3.7
H2	1.04	1.193	0.153	14.7
H3	1.3	1.193	0.107	8.2
H4	1.11	1.193	0.083	7.47
H5	1.4	1.193	0.207	14.7
H6	1.22	1.193	0.027	2.21

Now I will find the mean of the percentage errors:

$$\frac{3.7 + 14.7 + 8.2 + 7.47 + 14.7 + 2.21}{6}$$

=8.5%

The mean percentage error of the high tides is 8.5%

For low tides:

	Measurement of actual tides (m)	Graph projection	Absolute error	Percentage error (%)
L1	0.16	0.13	0.03	18.7
L2	0.16	0.13	0.03	18.7
L3	0.07	0.13	0.06	85.7
L4	0.08	0.13	0.05	62.5
L5	-0.02	0.13	0.15	750
L6	-0.01	0.13	0.14	140

Now, I will take the mean of the percentage errors:

$$\frac{18.7 + 18.7 + 85.7 + 62.5 + 750 + 140}{6}$$

=179.3%

The mean percentage error for the low tides is 179.3%

Which is better?

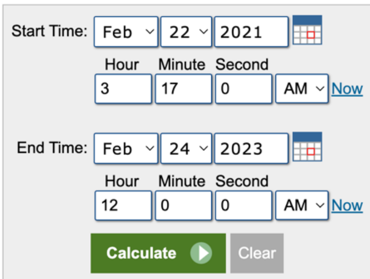
MAE calculates the average error in the same units as the data, making it easy to understand and straightforward, especially for datasets like this one with zero or near-zero values. In contrast, MPE shows errors as a percentage of actual values, which helps in comparing performance across datasets of different scales. One error they both have is that if there is an outlier, the average error tends to skew the result, leaning heavily on the outlier. However, one of the biggest problems with MPE is that it doesn't work well with zero or very small values because dividing by these can create undefined results or extremely large errors. This makes MPE less suitable in such cases, while MAE remains stable and easier to interpret. For data sets like mine, that have values such as (-0.02),(-0.01), calculating the MPE provided me with very high percentages such as 750% and 140%. Overall, MAE is better for datasets where scale matters and there are very small values approaching zero, while MPE is more useful for comparing relative errors across datasets**. For my dataset, I would prefer Mean Absolute Error (MAE)**

Predicting height and time of the high and low tides on:

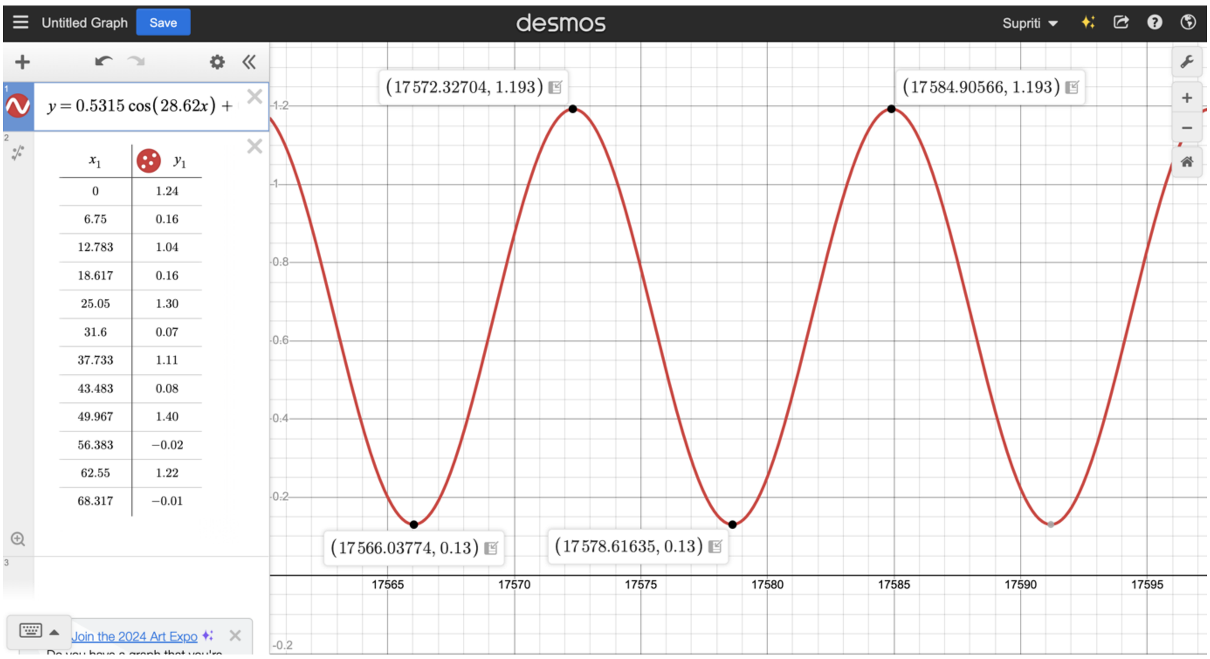
a) Wednesday 24th February 2023

With this date, I put it into a time calculator to see how many hours away it was from the first high tide in my data at 3:17 AM on February 22nd 2021:

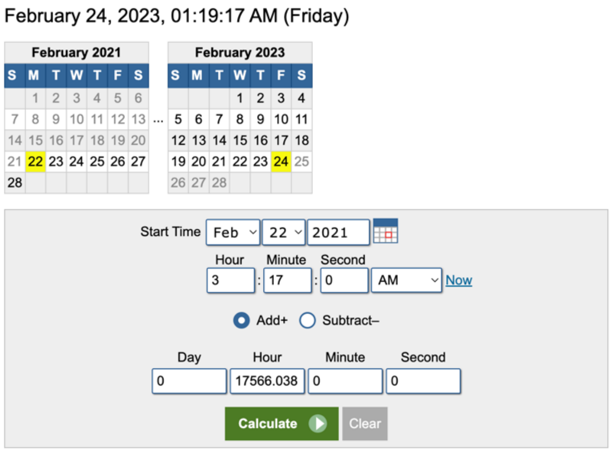
731 days, 20 hours, 43 minutes, and 0 second
731.8632 days
17,564.717 hours
1,053,883 minutes
63,232,980 seconds



After getting the number of hours, I went to 17,565 on my x axis on my data table since it represents the hours since the first high tide on February 22nd 2021. This point would be 12 AM of February 24th 2023 , marking the start of the day. After this, I looked at the highs and the lows of the graph after 17,565 on the x axis.



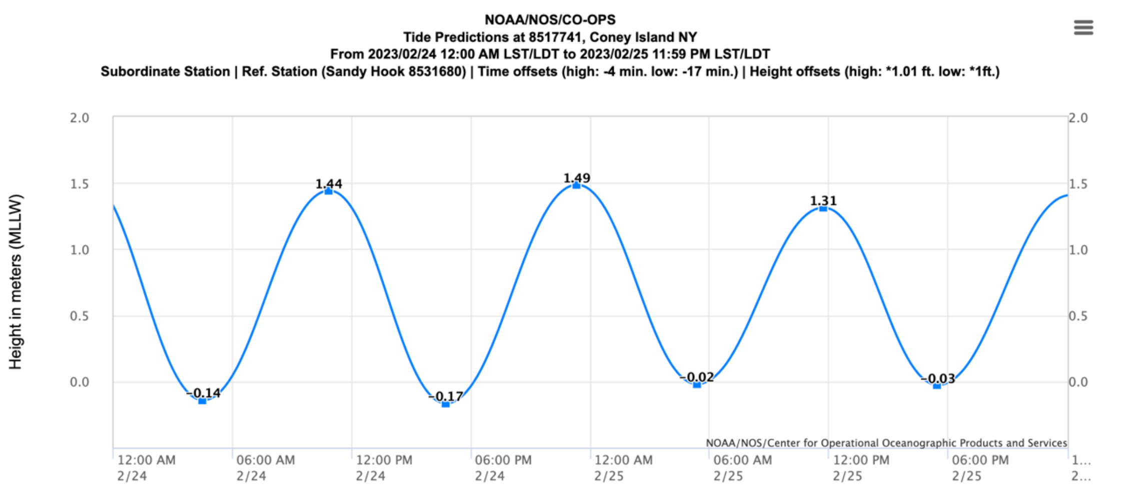
Now that I have the points, I can see that the Y coordinate resembles the predicted height and the x coordinate resembles the hours since 3:17 AM on February 22nd 2021. Now, to find the exact time of these high and low tides, I used the time calculator. For example, for the first low tide, I took the x coordinate, 17,566.038



With this, I could tell the first low tide on February 24th, 2023, was at 1:19 AM. Using this method, I calculated the time of the first and second high and low tides.

	Predicted measurement of the tides (m)	**Time (AM and PM)**
Low 1	0.13	1:19 AM
High 1	1.193	7:36 AM
Low 2	0.13	1:54 PM
High 2	1.193	8:11 PM

Now that I have the predicted times and heights, I used https://tidesandcurrents.noaa.gov to check the accuracy of my model

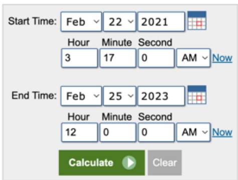


	Actual measurement of the tides (m)	**Time (AM and PM)**
Low 1	-0.14	4:30 AM
High 1	1.44	10:51 AM
Low 2	-0.17	4:43 PM
High 2	1.49	11:19 PM

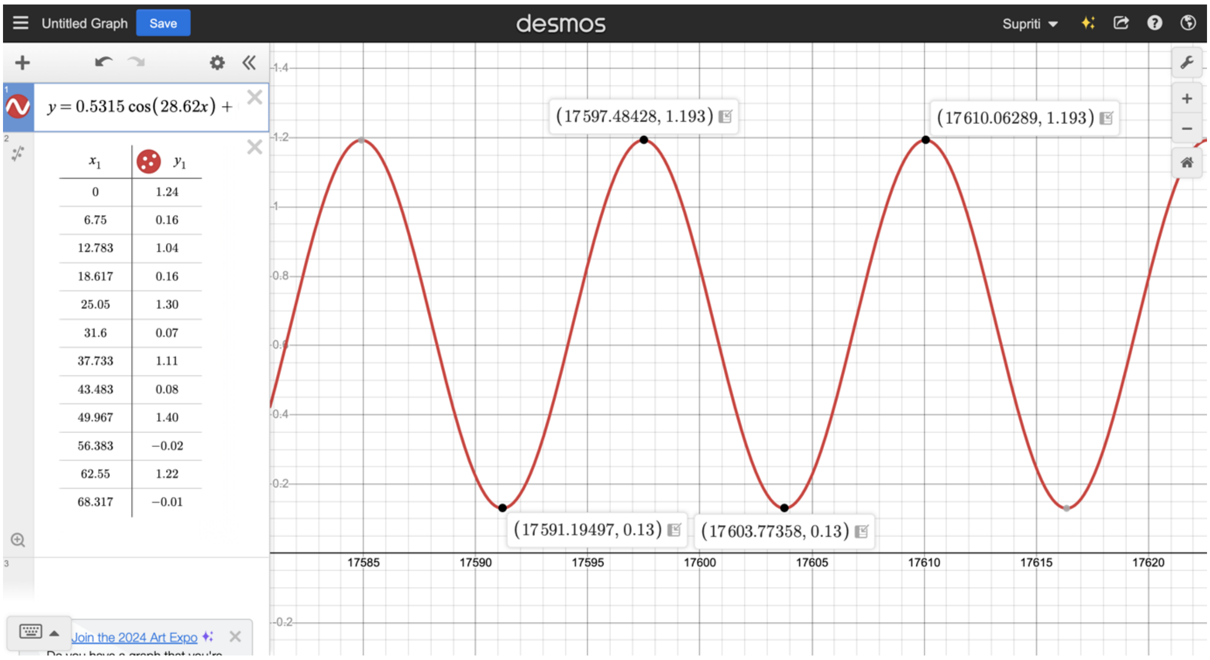
a) Thursday 25th February 2023

With this date, I put it into a time calculator to see how many hours away it was from the first high tide in my data at 3:17 AM on February 22nd 2021:

The time between Feb. 22, 2021, 3:17:00 AM and Feb. 25, 2023, 12:00:00 AM is:
732 days, 20 hours, 43 minutes, and 0 second
732.8632 days
17,588.717 hours
1,055,323 minutes
63,319,380 seconds



After this, I went to 17,588 on my x axis.

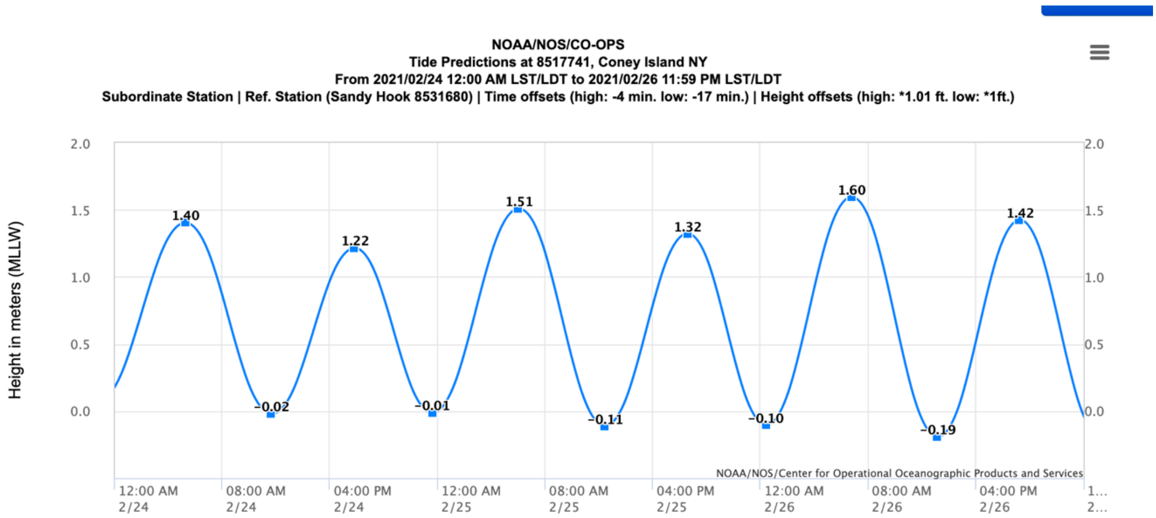


After this, I did the same thing I did previously for part a)

	Predicted measurement of the tides (m)	**Time (AM and PM)**
Low 1	0.13	2:28 AM
High 1	1.193	8:45 AM

Low 2	0.13	3:03 PM
High 2	1.193	9:20 PM

Now that I have the predicted times and heights, I used https://tidesandcurrents.noaa.gov to check the accuracy of my model



	Actual measurement of the tides (m)	**Time (AM and PM)**
High1	1.51	6:02 AM
Low 1	-0.11	12:25 PM
High 2	1.32	6:34 PM

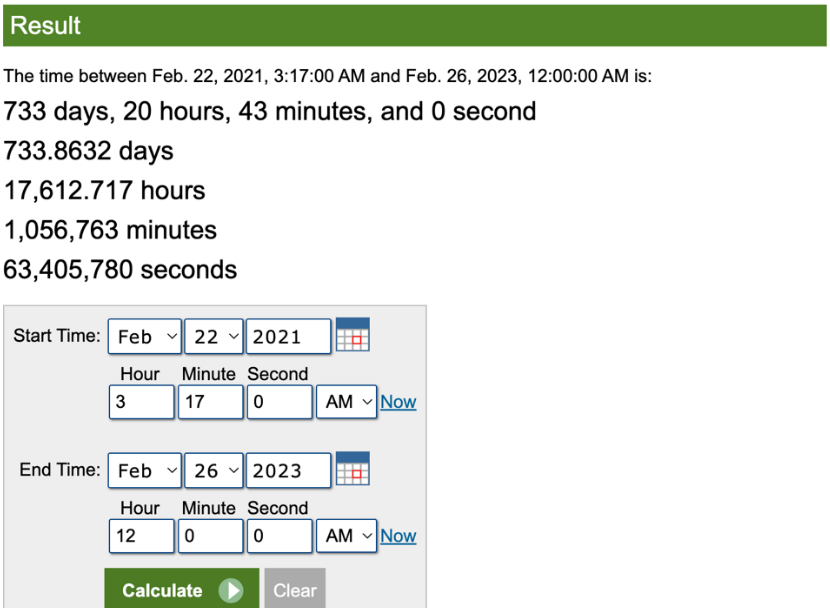
Note: For February 25th the second low tide after the second high tide fell on February 26th at 12:25 AM. That is why I didn't include it in the table.

Accuracy of the model:

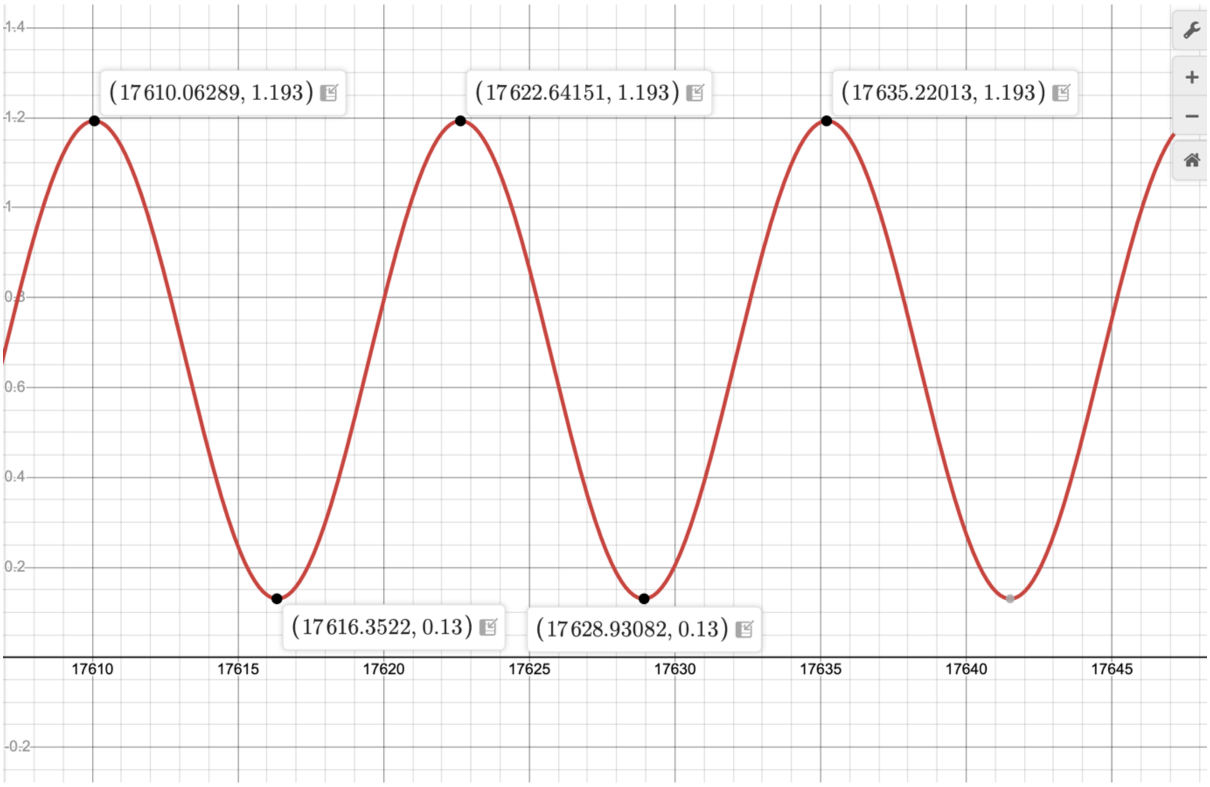
As seen above, the accuracy of my model was off by quite a few hours. The actual measurement of the tides also varied quite a lot. I think this was because of the decimal points, as well as the varying weather conditions. My graph does not take into account things such as weather, while in real life, this is something that would influence tide heights. Things such as gravitational forces and the season are also not taken into account.

When can the boat enter the harbor on February 26th, 2023?

With this date, I put it into a time calculator to see how many hours away it was from the first high tide in my data at 3:17 AM on February 22nd 2021:



After this, I went to 17,612 on the x axis of my data.



As shown above, since all the high tides would be approximately around 1.193, it would not be suitable for a normal boat to enter. Boats made for sailing in tides less than 2 meters deep are usually shallow-draft designs, for navigating coastal areas, estuaries, and shallow rivers (*ChatGPT*). They have flat or gently rounded bottoms to avoid grounding and ensure stability in these conditions (*ChatGPT*). Most of them also come with retractable keels or centreboards, making them more manoeuvrable and allowing for easy draft adjustments (*ChatGPT*). Examples of such boats include dinghies, small catamarans, and skiffs, which are lightweight but durable enough to handle changing tides. Such boats could enter when tides are highest, so at around 9:55 AM at the first high tide, or around 10:30 PM. I would recommend from 9:45 AM-10AM, since it would be brighter and there would be more visibility. The tides would also be pretty high for Coney Beach, from 1.13m-1.19m.

Main sources of error

a) Real life context

- **Irregularity of Tides**: Tides are influenced by things such as the positions of the moon, sun, and earth, local geography, and ocean currents. These things can create tidal patterns that are not perfectly sinusoidal.
- **Asymmetry in Tidal Cycles**: Tidal cycles often have unequal durations for rising (flood tide) and falling (ebb tide) phases, which can result to asymmetrical waveforms that a cosine model can't predict.
- **Mixed Tidal Patterns**: The beach I chose for this project experienced mixed semi-diurnal tides, where two high and two low tides occur daily but with varying amplitudes. A cosine model assumes a consistent pattern, which means since all the high tides have the same measurement and all the low tides have the same measurement, it cannot correctly predict the measurements of two high or two low tides on the same day because they vary.
- **Environment**: Coastal topography and bathymetry cause deviations in tidal heights and timings that a cosine function cannot account for.

b) Mathematical modeling:

- **Fixed Periodicity**: A cosine function assumes a constant period and amplitude, whereas tidal heights vary over time due to the changing positions of the moon and sun (for example, spring and neap tides)
- **Phase Shifts and Offsets**: Tidal patterns can sometimes include phase shifts and offsets caused by time delays in the response of the water body. Cosine function cannot stay up to date on these things so they can't account for them.
- **Nonlinear**: The interaction between tidal forces can create additional waves called overtides and compound. A single cosine function is too simple to capture these complex patterns, leading to less accurate results.

Conclusion:

While a cosine model can provide a rough approximation for tidal patterns, it oversimplifies the complex nature of real-world tides. To measure such things accurately we need advanced technology and methods such as harmonic analysis.