MATH 3800 F

Assignment 2

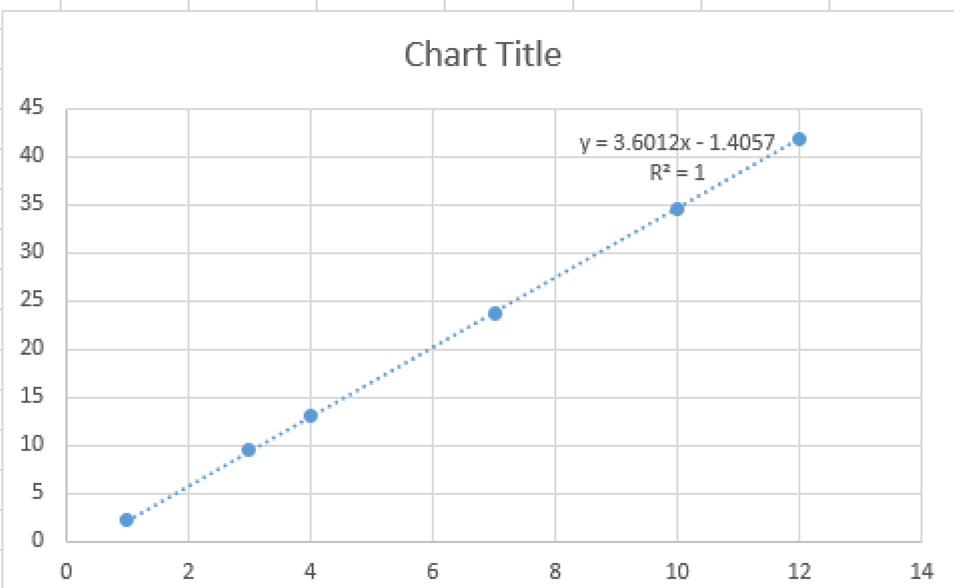
Krystian Wojcicki, 101001444

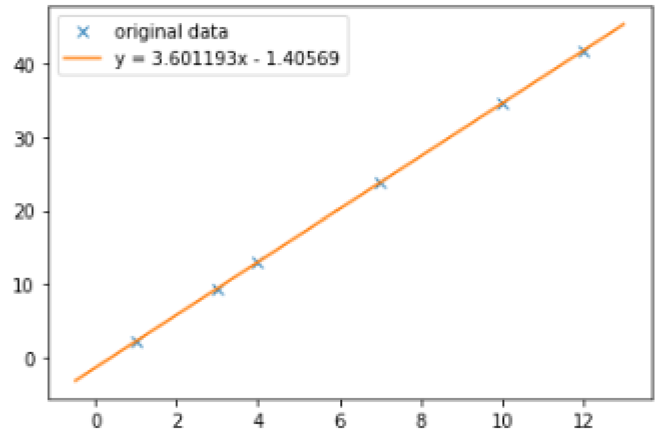
Winter 2020

1. **Use Least-Squares to fit a straight line, $y = ax + b$, to the data. Plot the data and the line. Caclulate R2 and comment on the quality of the fit.**

x	**1**	**3**	**4**	**7**	**10**	**12**
y	**2.16**	**9.42**	**13.05**	**23.77**	**34.58**	**41.83**

	x	y	x^2	y^2	xy	y_pred	residuals	residuals^2	(y-mean y)^2
	1	2.16	1	4.6656	2.16	2.195505	-0.0355	0.001260576	347.5117361
	3	9.42	9	88.7364	28.26	9.39789	0.02211	0.000488856	129.5423361
	4	13.05	16	170.3025	52.2	12.99908	0.050917	0.002592585	60.08833611
	7	23.77	49	565.0129	166.39	23.80266	-0.03266	0.001066712	8.811002778
	10	34.58	100	1195.776	345.8	34.60624	-0.02624	0.000688461	189.8424694
	12	41.83	144	1749.749	501.96	41.80862	0.021376	0.00045694	442.1908028
sums	37	124.81	319	3774.243	1096.77	124.81	6.22E-15	0.006554128	1177.986683
m	6								
	a	3.601193					SSE	0.006554128	
	b	-1.40569					SST	1177.986683	
							SSR	1177.980129	
							R^2	0.999994436	



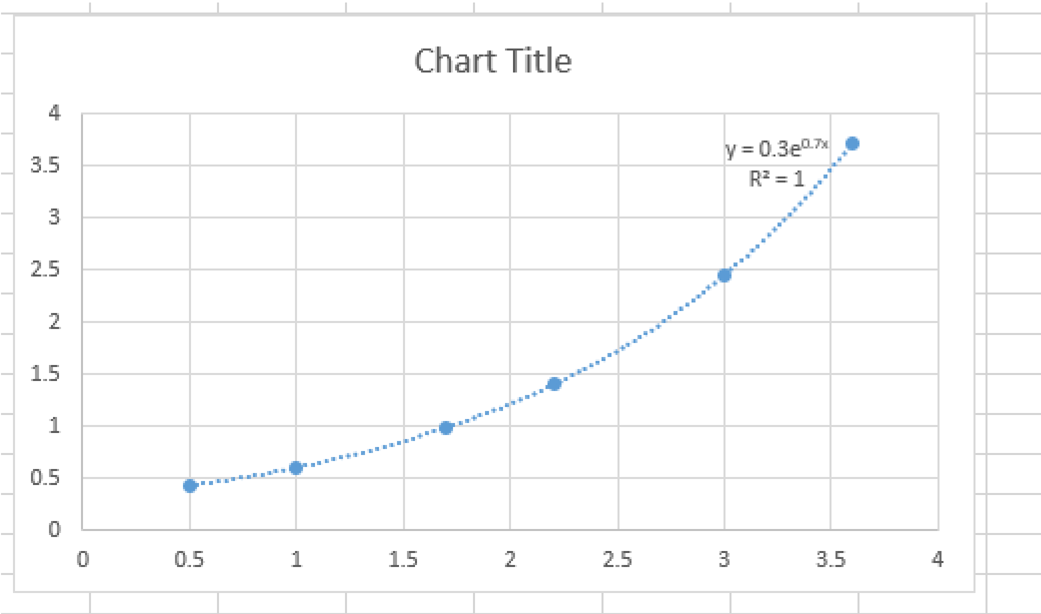


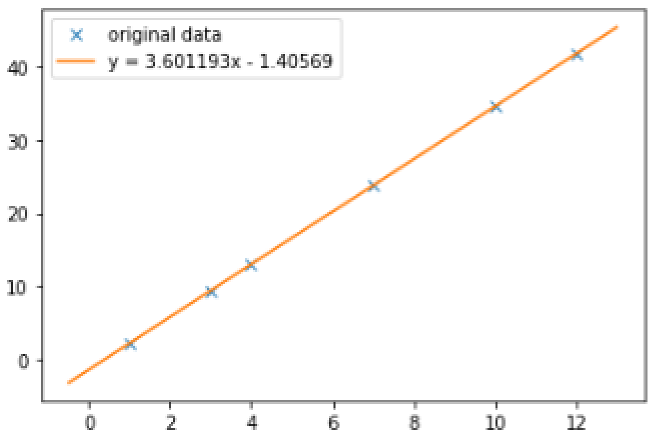
As we can see from the value of R^2 and the plot of the data, the fit is excellent.

2. **Use Least-Squares to fit a curve, $y = ae^{bx}$, to the appropriate transformation of the data. Plot the data and the curve. Caclulate R^2 and comment on the quality of the fit.**

x	**0.5**	**1**	**1.7**	**2.2**	**3**	**3.6**
y	**0.425**	**0.605**	**0.987**	**1.400**	**2.448**	**3.728**

	x	**y**	**lny**	**x^2**	**xlny**	**y_pred**	**residuals**	**residuals^2**	**(y-mean y)^2**
y=ae^(bx)	0.5	0.425	-0.85567	0.25	-0.42783	-0.85392	-0.00174	3.032E-06	1.106119862
lny=ln(ae^(bx))	1	0.605	-0.50253	1	-0.50253	-0.50393	0.001404	1.9722E-06	0.488018267
lny=lna + ln(e^bx)	1.7	0.987	-0.01309	2.89	-0.02224	-0.01394	0.000855	7.306E-07	0.043740151
lny=lna + bx	2.2	1.4	0.336472	4.84	0.740239	0.336054	0.000419	1.7518E-07	0.019716657
Y=b+ax	3	2.448	0.895271	9	2.685814	0.896044	-0.00077	5.9635E-07	0.488901817
	3.6	3.728	1.315872	12.96	4.737139	1.316036	-0.00016	2.6939E-08	1.253987147
sums	12	9.593	1.176337	30.94	7.210587	1.176337	-3.9E-16	6.5332E-06	3.400483901
m	6								
	a	0.699987					**SSE**	6.5332E-06	
	b	-1.20392					**SST**	3.4004839	
	e^b	0.300016					**SSR**	3.40047737	
							R^2	0.99999808	



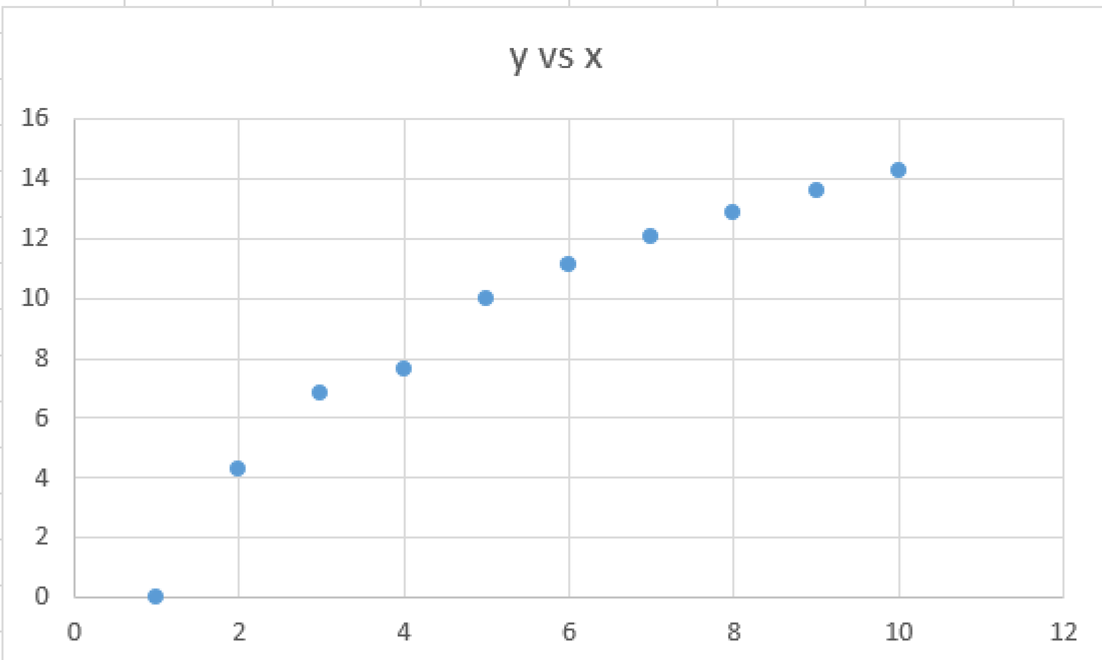


The R^2 on the transformed/linear data is excellent and once again the fit is excellent.

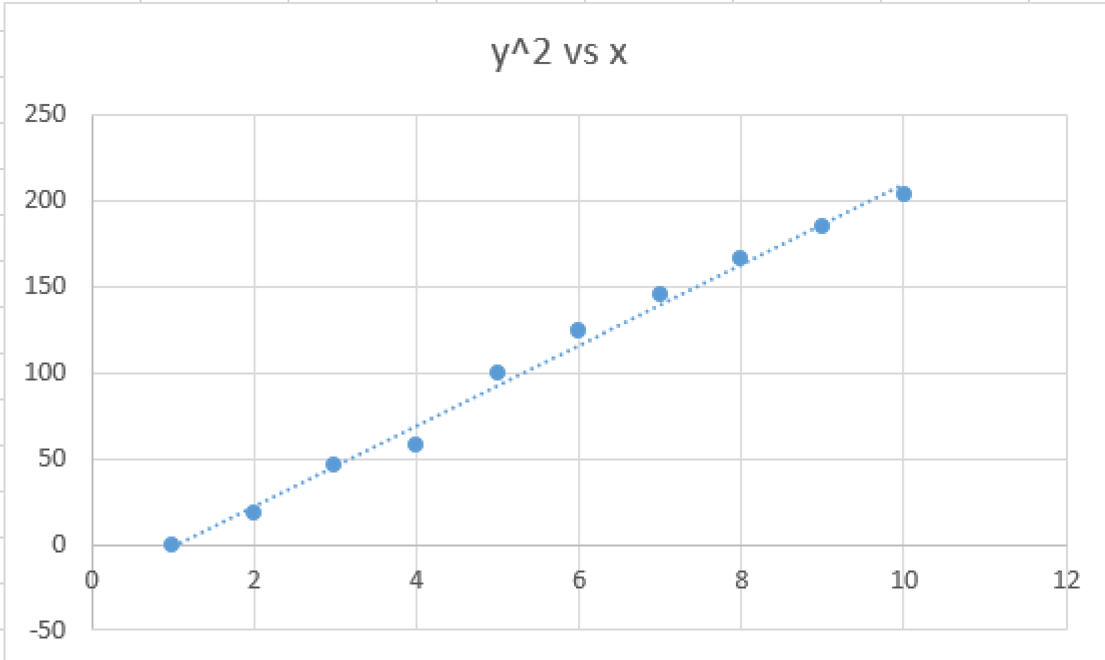
3. **Use the Ladder of Powers Transformations to determine if a simple/one-term model would be an appropriate fit. Which functional form seems best ? (You do not have to fit the curve.)**

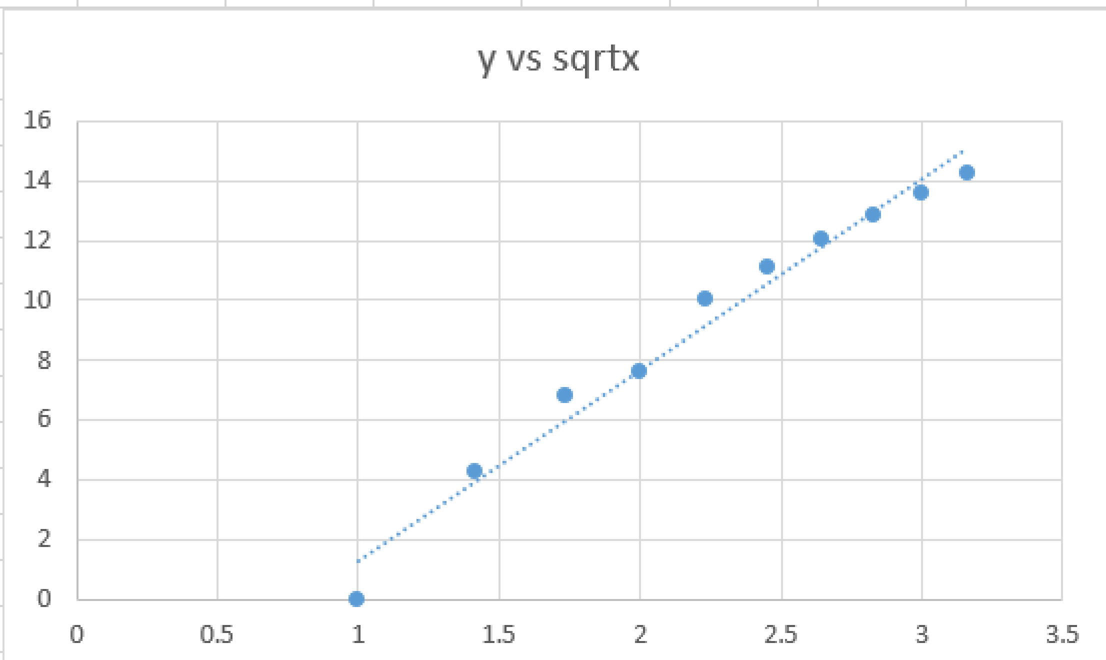
x	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
y	**0**	**4.27**	**6.82**	**8.60**	**10**	**11.13**	**12.04**	**12.88**	**13.61**	**14.25**

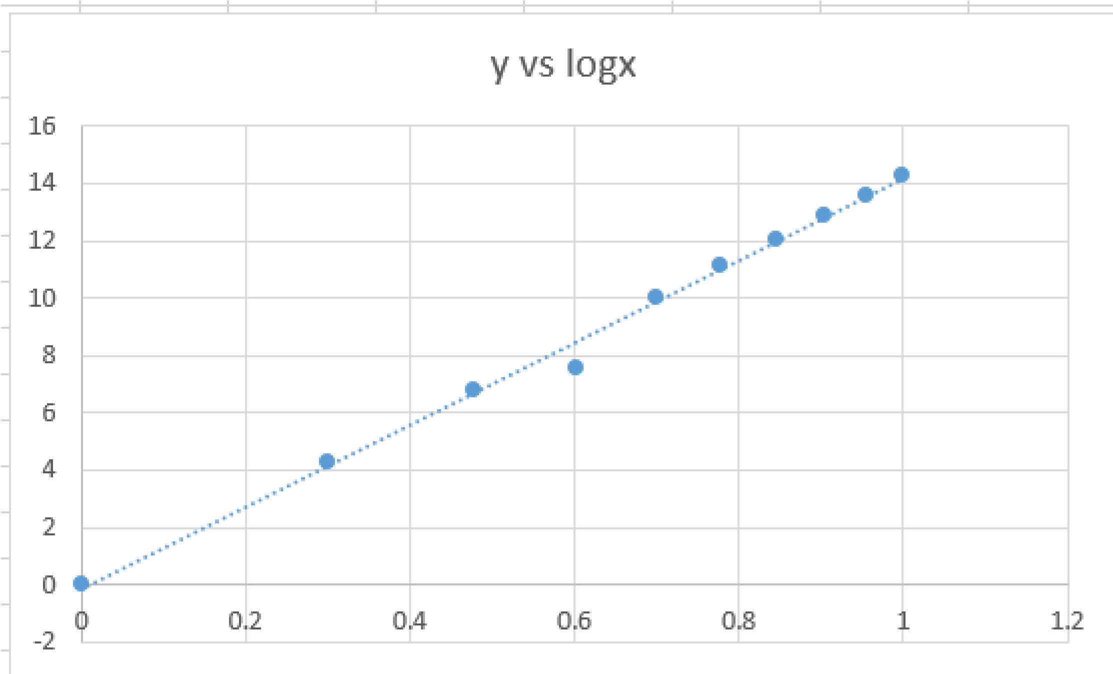
x	sqrtx	log x	minus 1/sqrtx	y	y^2
1	1	0	-1	0	0
2	1.41421	0.30103	-0.707106781	4.27	18.2329
3	1.73205	0.47712	-0.577350269	6.82	46.5124
4	2	0.60206	-0.5	7.6	57.76
5	2.23607	0.69897	-0.447213595	10	100
6	2.44949	0.77815	-0.40824829	11.13	123.877
7	2.64575	0.8451	-0.377964473	12.04	144.962
8	2.82843	0.90309	-0.353553391	12.88	165.894
9	3	0.95424	-0.333333333	13.61	185.232
10	3.16228	1	-0.316227766	14.25	203.063

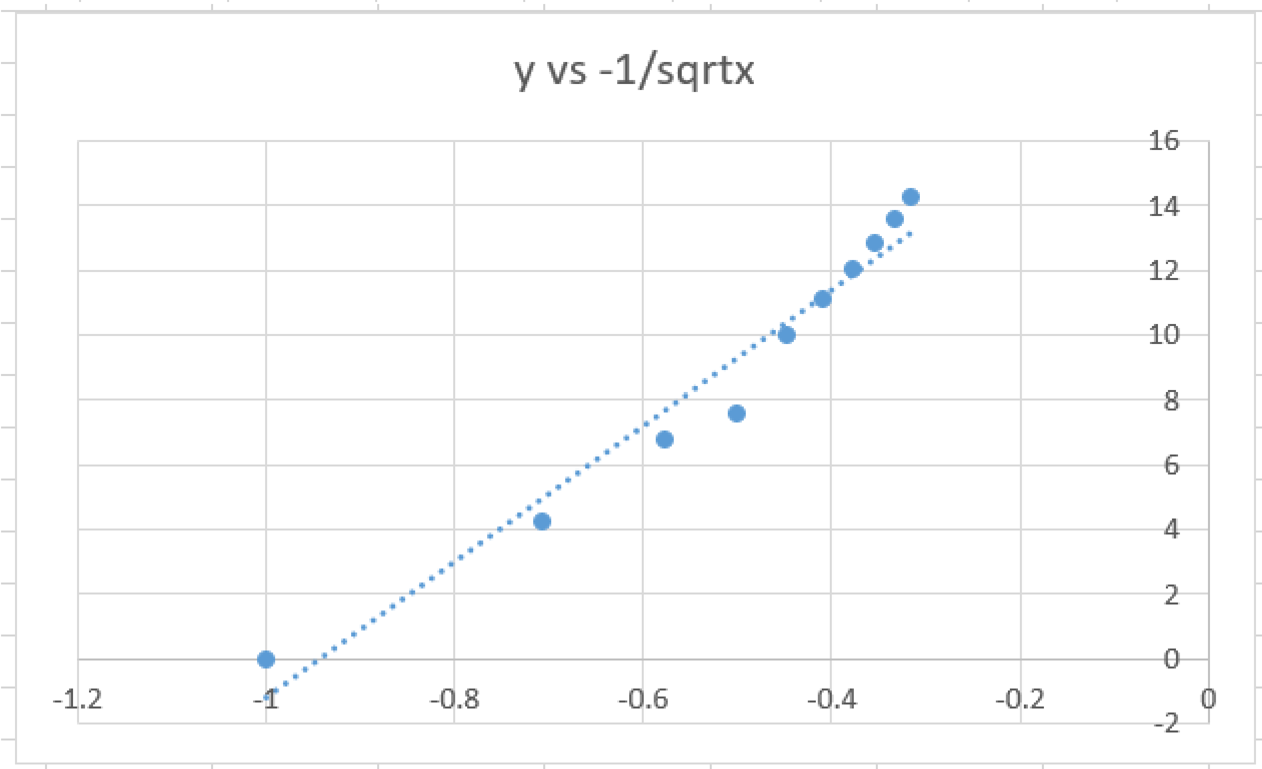


Data is increasing and concave down, so we will try $x \Rightarrow \sqrt{x}, log(x), \ldots$ on $y \Rightarrow y, y^2$


y^2 vs x
250
200
150
100
50
0
-50
0
2
4
6
8
10
12


y vs sqrtx
16
14
12
10
8
6
4
2
0
0
0.5
1
1.5
2
2.5
3
3.5





4. **Consider the data below. Construct a divided difference table and determine if a low-order polynomial would be appropriate as an empirical model.**

x	0	1	2	3	4	5	6	7
y	1	4.5	20	90	403	1808	8103	36316

x	y	first	second	third	fourth	fifth	sixth	seventh
0	1							
1	4.5	3.5						
2	20	15.5	6					
3	90	70	27.25	7.083333				
4	403	313	121.5	31.41667	6.083333			
5	1808	1405	546	141.5	27.52083	4.2875		
6	8103	6295	2445	633	122.875	19.07083	2.463889	
7	36316	28213	10959	2838	551.25	85.675	11.10069	1.233829

From the divided difference table we can see that there is no column that is relatively constant (or relatively close to 0). So we can conclude that a low order polynomial would not be appropriate for this data.

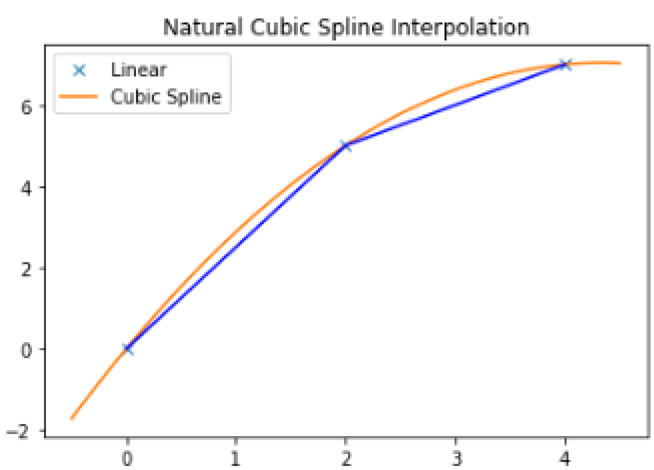
5. **Find the natural cubic spline for the points (0, 0),(2, 5) and (4, 7).**

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import interpolate

x = [0, 2, 4]
y = [0, 5, 7]

tck = interpolate.splrep(x, y, k=2)
xnew = np.linspace( -0.5, 4.5, 100 )
ynew = interpolate.splev( np.linspace(-0.5, 4.5, 100 ), tck)

plt.figure()
plt.plot(x, y, 'x', xnew, ynew, x, y, 'b')
plt.legend(['Linear', 'Cubic Spline'])
plt.title('Natural Cubic Spline Interpolation')
plt.show()
```



$$y_1 = S_1(x_1) : 0 = a_1 \tag{1}$$
$$y_2 = S_1(x_2) : 5 = a_1 + 2b_1 + 4c_1 + 8d_1 \tag{2}$$
$$y_2 = S_2(x_2) : 5 = a_2 + 2b_2 + 4c_2 + 8d_2 \tag{3}$$
$$y_3 = S_2(x_3) : 7 = a_2 + 4b_2 + 16c_2 + 64d_2 \tag{4}$$
$$S_1\prime(x) = b_1 + 2xc_1 + 3x^2 d_1$$
$$S_2\prime(x) = b_2 + 2xc_2 + 3x^2 d_2$$
$$S_1\prime\prime(x) = 2c_1 + 6xd_1$$
$$S_2\prime\prime(x) = 2c_2 + 6xd_2$$
$$S_1\prime(x_2) = S_2\prime(x_2) : b_1 + 4c_1 + 12d_1 = b_2 + 4c_2 + 12d_2 \tag{5}$$
$$S_1\prime\prime(x_2) = S_2\prime\prime(x_2) : 2c_1 + 12d_1 = 2c_2 + 12d_2 \tag{6}$$
$$S_1\prime\prime(x_1) = 0 : 2c_1 + 0 = 0 \tag{7}$$
$$S_2\prime\prime(x_3) = 0 : 2c_2 + 24d_2 = 0 \tag{8}$$

From (1) $\Rightarrow a_1 = 0$
From (7) $\Rightarrow c_1 = 0$
Substitute $c_1 = 0$ into (6) $\Rightarrow 12d_1 = 2c_2 + 12d_2 \Rightarrow d_1 = -d_2$ and $c_2 = -12d_2$
Substitute $c_1 = 0, d_1 = -d_2$ into (5) $\Rightarrow b_1 + 12d_1 = b_2 + 4c_2 + 12d_2 \Rightarrow b_1 - b_2 = -24d_2 \Rightarrow b_1 = b_2 - 24d_2, b_2 = b_1 + 24d_2$
Substitute $a_1 = 0, c_1 = 0, b_1 = b_2 - 24d_2, d_1 = -d_2$ into (2) and (3) $\Rightarrow 2b_2 - 48d_2 - 8d_2 = a_2 + 2b_2 - 40d_2 \Rightarrow a_2 = -16d_2$
Substittue $a_2 = -16d_2, c_2 = -12d_2$ into (4) $\Rightarrow 7 = -16d_2 + 4b_2 + 16(-12d_2) + 64d_2 \Rightarrow b_2 = \frac{7}{4} + 36d_2$
Substitute $b_2 = \frac{7}{4} + 36d_2$ into $b_2 = b_1 + 24d_2 \Rightarrow \frac{20}{7}(b_1 - 12d_2) = 5$
Substitute $\frac{20}{7}(b_1 - 12d_2) = 5, a_1 = 0, c_1 = 0, d_1 = -d_2$ into (2) $\Rightarrow \frac{20}{7}b_1 - \frac{240}{7}d_2 = 2b_1 - 8d_2 \Rightarrow b_1 = \frac{92}{3}d_2$
Substitute $b_1 = 92/3d_2$ into $b_1 - 12d_2 = \frac{7}{4} \Rightarrow d_2 = \frac{3}{32}$

Therefore $a_1 = 0, b_1 = (92/3) * (3/32) = \frac{23}{8}, c_1 = 0, d_1 = -\frac{3}{32}, a_2 = -16 * (3/32) = -1.5, b_2 = \frac{23}{8} + 24 * (3/32) = \frac{41}{8}, c_2 = -12 * \frac{3}{32} = -\frac{9}{8}, d_2 = \frac{3}{32}$
Giving us $\begin{cases} S_1(x) = \frac{23}{8}x - \frac{3}{32}x^3, 0 \leq x \leq 2 \\ S_2(x) = -1.5 + \frac{41}{8}x - \frac{9}{8}x^2 + \frac{3}{32}x^3, 2 \leq x \leq 4 \end{cases}$
Check to ensure that given points are in the cubic spine function created. And indeed $S_1(0) = 0, S_1(2) = 5, S_2(2) = 5, S_2(4) = 7$.