

PART I:

Women's 800m Freestyle

x	y
19	799
25	778.2
25	776
26	767.2
27	737.8
29	723.8
30	701.2
35	694.4
36	671.7
41	652.5
53	642.4
56	630.9
57	627.3
58	617.7
58	616.2
58	611.8
59	611.4
60	595.6
62	591.6
64	587.3
64	576.9
67	575.8
67	562.9
68	559
68	557.8
68	550.4
70	549.1
70	542.45
71	539.4
71	538.1
72	533.83

72	533.68
73	532.97
74	530.1
74	527.66
74	527.59
75	523.48
76	520.68
76	519.63
76	517.14
77	515.04
78	514.86
78	511.3
78	510.53
78	504.62
87	502.44
87	499.53
88	497.12
89	496.22
108	494.1

Men's 200m Freestyle

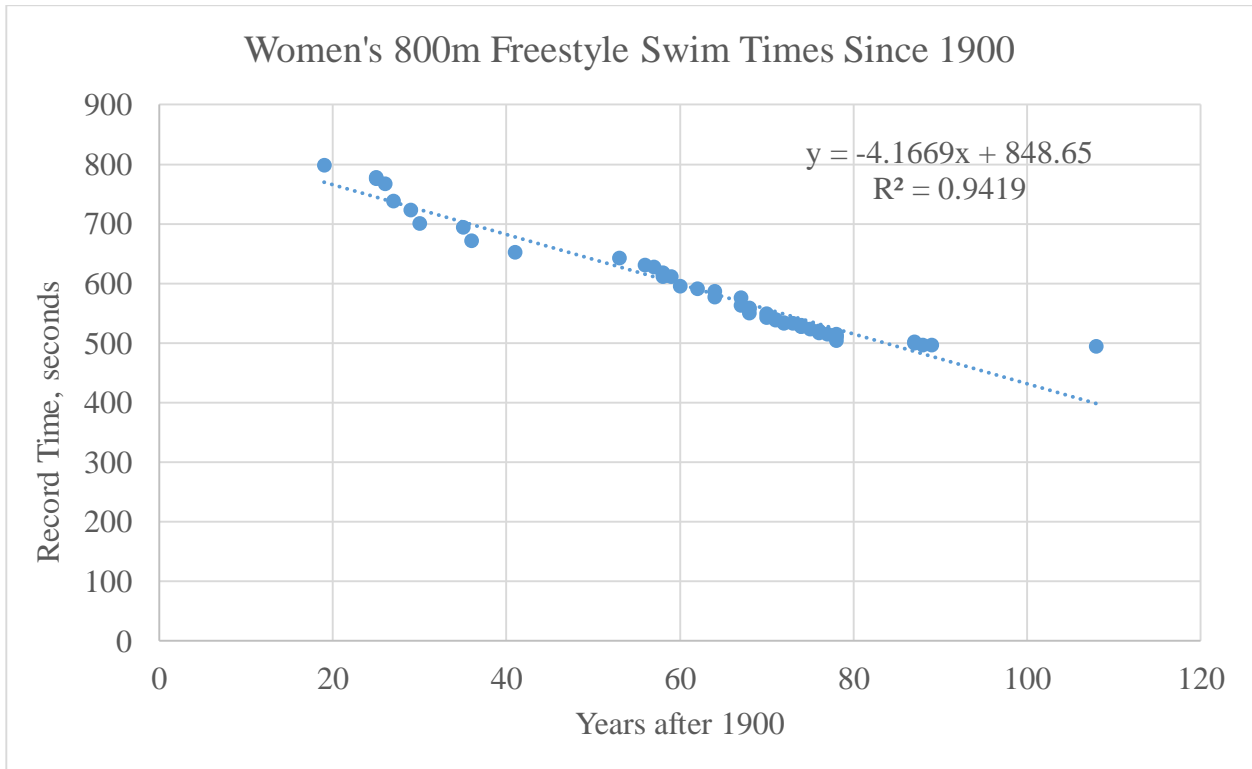
x	y
8	151.6
10	150
11	145.4
16	141.6
20	139.8
22	135.6
25	135.2
27	128
35	127.2
44	126.2
46	125.4
50	124.6
54	123.9
55	123.4
57	121.5
58	124.8
58	123.2

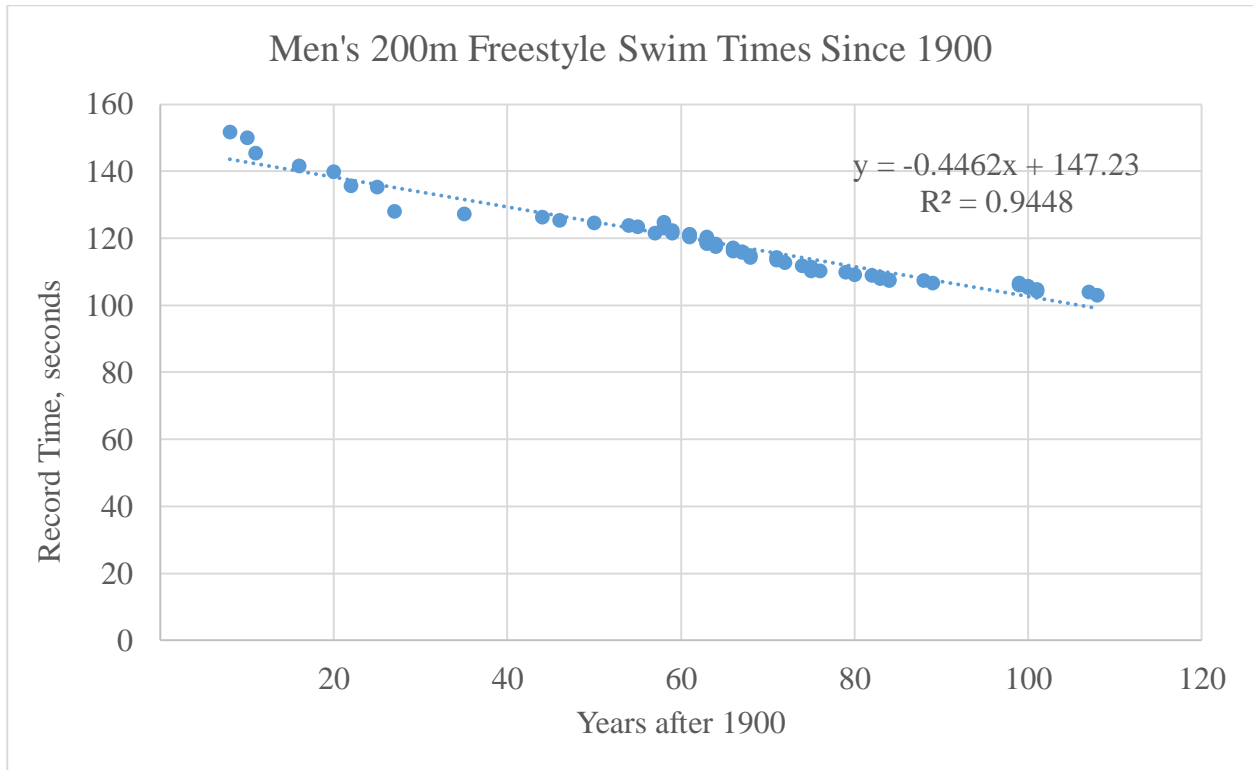
58	123
59	122.2
59	121.5
61	121.2
61	121.1
61	120.4
63	120.3
63	118.8
63	118.5
63	118.4
64	118.2
64	117.6
66	117.2
66	116.2
67	116
67	115.7
68	114.8
68	114.3
71	114.2
71	113.5
72	112.8
74	111.7
75	111.4
75	110.9
75	110.3
76	110.3
79	109.8
80	109.2
82	108.9
83	108.3
83	107.9
84	107.5
84	107.4
88	107.3
89	106.7
99	106.7
99	106.3
99	106
100	105.7
100	105.5
100	105.3

101	104.7
101	104.1
107	103.9
108	103

The conversion of minutes to seconds must occur, because it is easier and there is less room for error if the same units are being used.

PART II:





As seen from the scatterplots, the data appears to be asymptotic. The trendline for both scatterplots is linear. On the men's scatterplot, in approximately 1955 the data points appear to become more uniform. There is one outlier on the women's scatterplot in 2008 that shows that a linear trend is not the best fit, indicating a different trend might be more appropriate. Both plots have a negative correlation.

PART III:

Women's

Correlation coefficient
-0.971
Slope
-4.17
y-intercept
848.65

Men's

Correlation coefficient
-0.972
Slope
-0.45
y-intercept
147.23

In each event, 'r' represents the correlation coefficient, 'a' represents the slope, and 'b' represents the y-intercept. Using the slope intercept formula $y=mx+b$, or $t=ax+b$ in this case, and substituting the given year for x (independent variable), we get the record time, 't'(dependent variable). Applying the linear correlation coefficient formula gives us the correlation coefficient 'r'. This coefficient tells us how closely the data represents a straight line. The three variables, 'a', 'b', and 'r' are calculated the same way in both the men's and women's swim events.

PART IV:

Women's 800m Predicted World Records		
Year	Year Diff.	Record, sec
2012	112	381.96
2100	200	15.27
-776	-2676	11999.19

Men's 200m Predicted World Records		
Year	Year Diff.	Record, sec
2012	112	97.25
2100	200	57.98
-776	-2676	1341.36

There does not appear to be a causal relationship between the years and the record times outside of the 20th century. Unless the athletes at the first Olympic Games were terribly out of shape and only robots that swim as fast as boats are allowed to compete in the 2100 Olympics, there is no causal relationship between x and y . Athletes may have improved their times since the 1920's, but that is most likely attributed to the fact that it is their job to be in peak physical condition. Whereas, in the past, athletes may have held jobs during off-seasons and then started training for competitions close to the time of the Olympic Games.