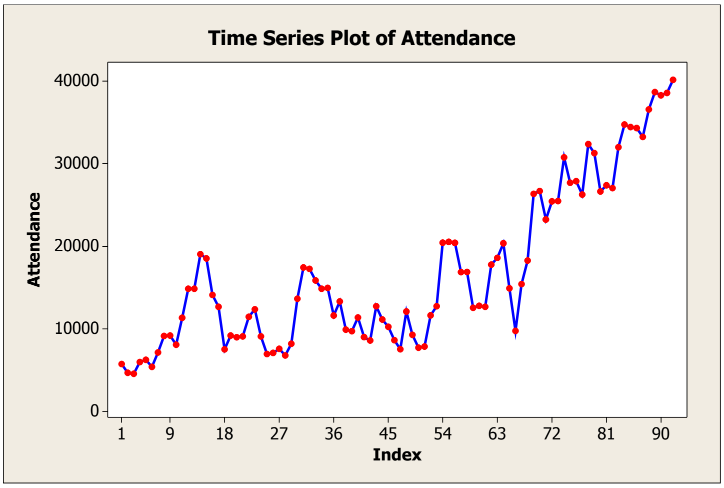
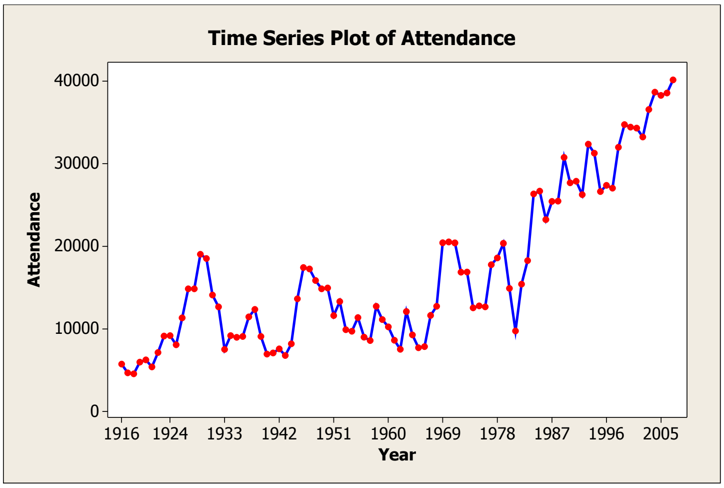
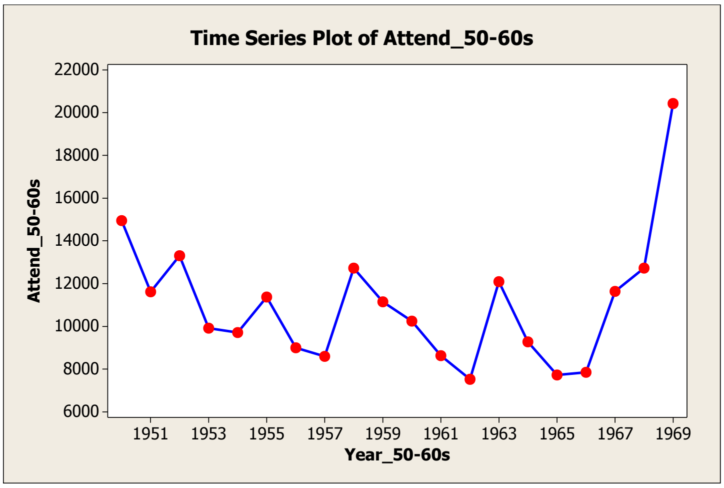
A.


Time Series Plot of Attendance
Attendance
Index
40000
30000
20000
10000
0
1
9
18
27
36
45
54
63
72
81
90


Time Series Plot of Attendance
Attendance
Year
40000
30000
20000
10000
0
1916
1924
1933
1942
1951
1960
1969
1978
1987
1996
2005


Time Series Plot of Attend_50-60s
Attend_50-60s
Year_50-60s
22000
20000
18000
16000
14000
12000
10000
8000
6000
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969

(a)

Year t	**Attendance** y_t	**MA(2)** $\widehat{y}_t$	**Prediction Error** $e_t = y_t - \widehat{y}_t$	**Squared Prediction Error**
1950	14,948	*	*	*
1951	11,616	*	*	*
1952	13,309	13,282	+27	729.00
1953	9,918	12,462.5	−2544.5	6,474,480.25
1954	9,717	11,613.5	−1896.5	3,596,712.25
1955	11,374	9817.5	+1556.5	2,422,692.25
1956	9,001	10,545.5	−1544.5	2,385,480.25
1957	8,598	10,187.5	−1589.5	2,526,510.25
1958	12,726	8799.5	+3,926.5	15,417,402.25
1959	11,146	10,662	+484	234,256.00
				33,058,262.50

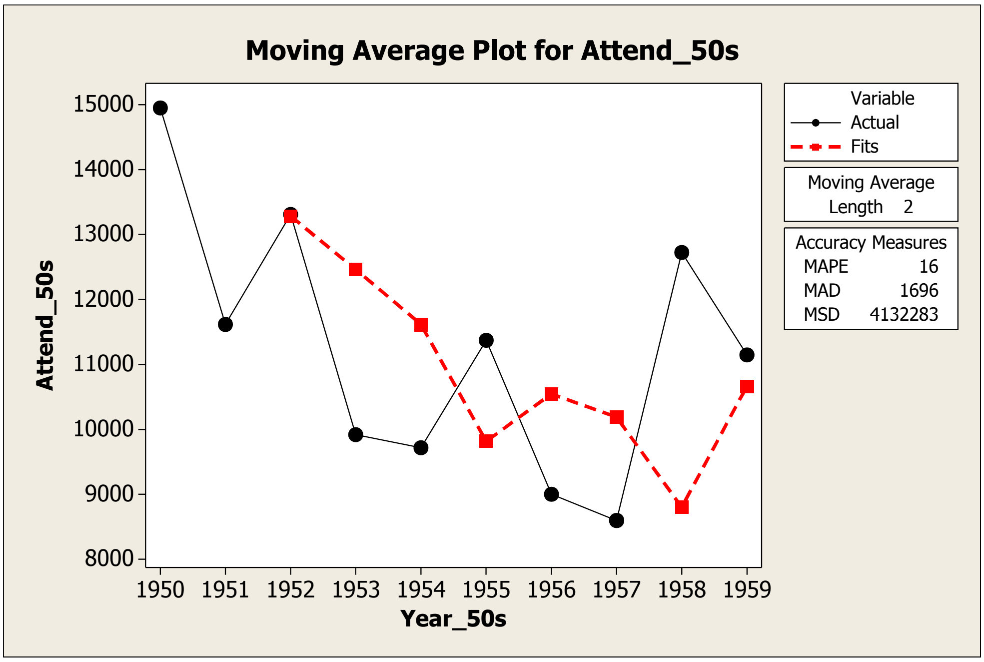
$$\text{MSE} = (33,058,262.50)/8 = \boxed{4{,}132{,}282.81}$$

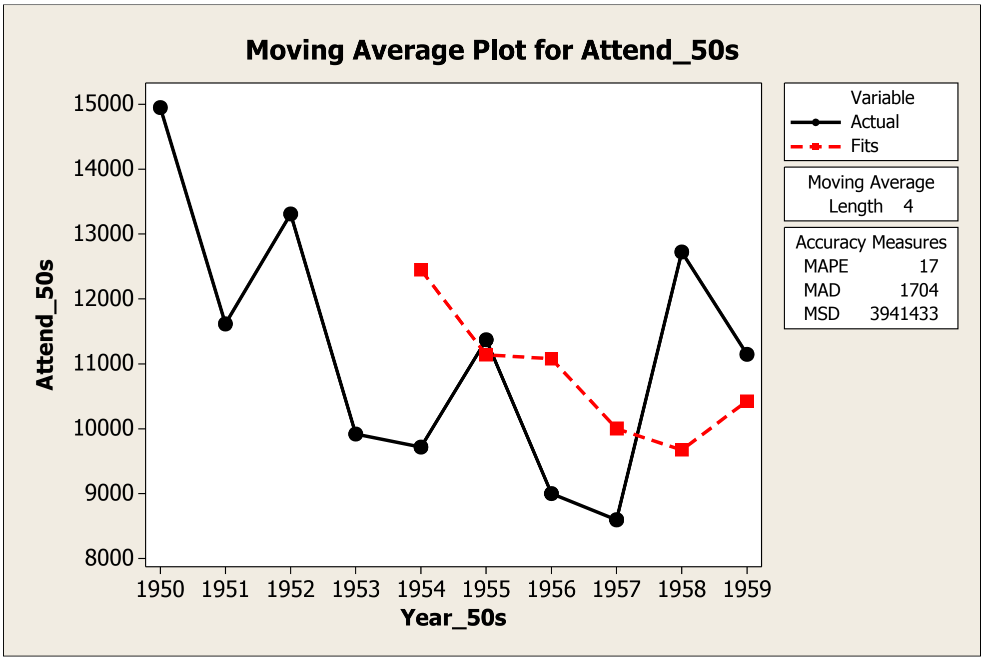
(b)

Year t	Attendance y_t	MA(4) $\widehat{y}_t$	Prediction Error $e_t = y_t - \widehat{y}_t$	Squared Prediction Error
1950	14,948	*	*	*
1951	11,616	*	*	*
1952	13,309	*	*	*
1953	9,918	*	*	*
1954	9,717	12,447.75	−2730.75	7,456,995.56
1955	11,374	11,140	+234	54,756.00
1956	9,001	11,079.5	−2078.5	4,320,162.25
1957	8,598	10,002.5	−1404.5	1,972,620.25
1958	12,726	9672.5	+3053.5	9,323,862.25
1959	11,146	10,424.75	+721.25	520,201.56
				23,648,597.87

$$\text{MSE} = (23,648,597.87)/6 = \boxed{3,941,432.98}$$

(c) Yes , except for minor roundoff error.





(d) MA(4)

(e)

$$\text{MSE} = \frac{\text{SSE}}{10} = \frac{28,554,838}{10} = \boxed{2{,}855{,}483.8}$$

(f) Yes since it has smaller MSE.

(g) MSE for MA(2) in 1960's = $\boxed{12{,}805{,}981}$

(h) MSE for MA(4) in 1960's = $\boxed{15{,}385{,}295}$

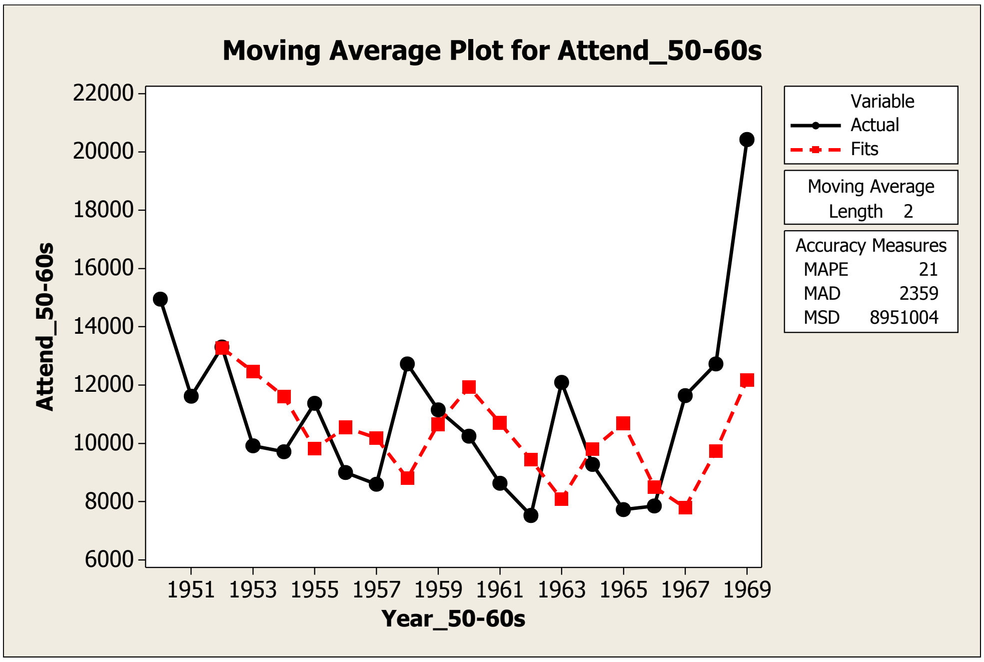
(i) MSE for Simple Regression Trend in 1960's = $\boxed{25{,}361{,}705}$

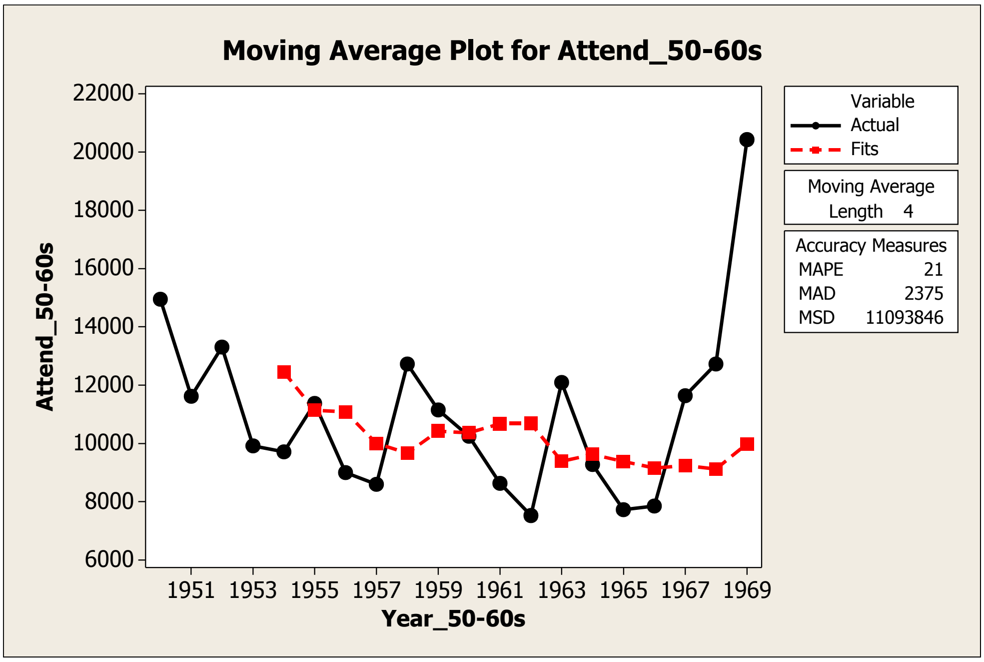
(j) MA(2) performs best in the 1960's.

No, Simple Regression Trend performed best historically (in the 1950's.)

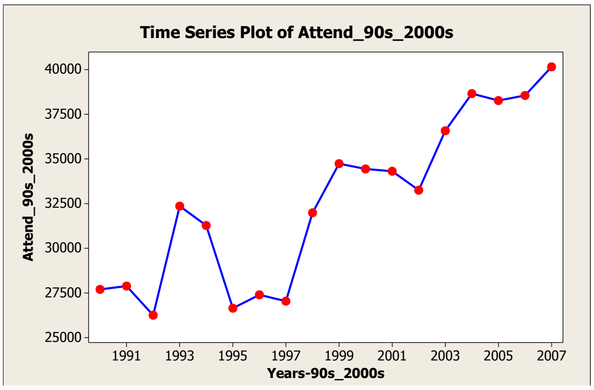
(continued)

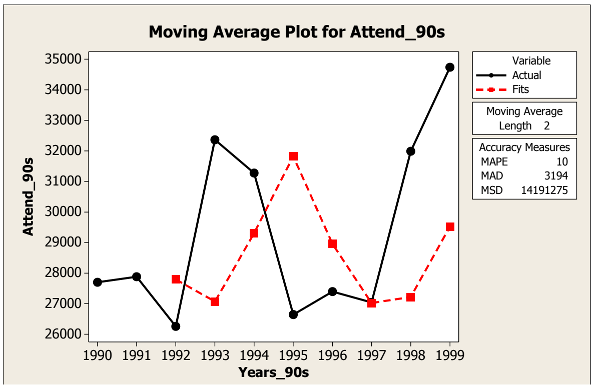
(k) MA(2) uses less past data than MA(4) and so is more flexible and responds more quickly to the dramatic fluctuations in Cubs attendance in the 1960's

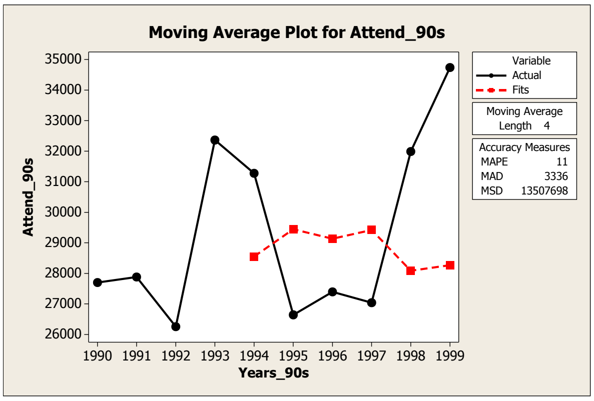




B.







(a) 14,191,275

(b) 13,507,698

(c) 61,985,540/10 = 6,198,554

(d) SRT

(e) 3,477,496

(f) 7,402,932

(g) 12,873,371

(h) MA(2)

(i) No

(j) SRT forecasts are consistently <u>too low</u>.

C.

(a)

Week t	**Gas Sales** y_t	**($w = 0.10$)** $\widehat{y}_t$	**Prediction Error** $e_t = y_t - \widehat{y}_t$	**Squared Prediction Error**
1	18	18	0	0
2	21	18	+3	9.00
3	19	18.3	+0.7	0.49
4	22	18.37	+3.63	13.18
5	17	18.73	−1.73	3.00
6	17	18.56	−1.56	2.43
7	20	18.40	+1.60	2.55
8	19	18.56	+0.44	0.19
9	11	18.61	−7.61	57.87
10	25	17.85	+7.15	51.18
				139.89

(b) MSE = 13.99

(continued)

(c)

Week t	**Gas Sales** y_t	($w = 0.90$) $\widehat{y}_t$	**Prediction Error** $e_t = y_t - \widehat{y}_t$	**Squared Prediction Error**
1	18	18	0	0
2	21	18	+3	9.00
3	19	20.7	−1.70	2.89
4	22	19.17	+2.83	8.01
5	17	21.72	−4.72	22.25
6	17	17.47	−0.47	0.22
7	20	17.05	+2.95	8.72
8	19	19.70	−0.70	0.50
9	11	19.07	−8.07	65.13
10	25	11.81	+13.19	174.05
				290.77

(d) MSE = 29.08

(e)

- MINITAB MSD for ($w = 0.10$) = 13.9884
- MINITAB MSD for ($w = 0.90$) = 29.0774

(f) MSE = 20.1484

(g) $w = 0.10$

D.

(a) ES($w = 0.90$) with MSE = 8,116,254 (or MSE = 8,115,921 if you use the option K = 1.)

(b) Here are the forecasts:

Model	$\widehat{y}_{2008}$
MA(5)	38,444
MA(10)	36,094.9
$w = 0.10$	32,640.6
$w = 0.50$	39,107.8
$w = 0.90$	39,991.7

So ES($w = 0.90$) also provides the closest forecast for the year 2008.

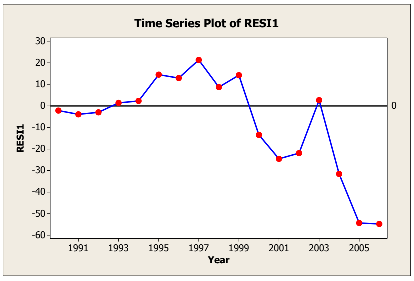
E.

(a)

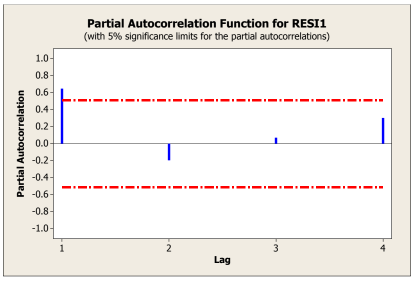
$$\widehat{y}_t = 20.0788 \times (1.306)^t$$

(b) Car ownership increases by approximately **30.6%** each year.

(c) The time plot of residuals does appear to show autocorrelation. The P-value from the Runs Test is 0.009. The conclusion is that the residuals are autocorrelated.



(d) Since residuals are autocorrelated, we need to predict them in order to improve the trend predictions. The Partial Autocorrelation plot indicates significance for a lag of a single time period in the past.



An AR(2) model from the ARIMA procedure shows a P-value of 0.877 for lag 2, not significant, so "drop" the second lag.

An AR(1) model shows a P-value of 0.000 for lag 1, significant. Therefore the correct value of p for the AR(p) model is $p = \boxed{1}$

The prediction equation for residuals is

$$\widehat{e}_t = (0.9514)e_{t-1}$$

(e)

$$\begin{aligned} F_t &= \widehat{y}_t + \widehat{e}_t \\ F_{2007} &= \widehat{y}_{2007} + \widehat{e}_{2007} \\ F_{2007} &= 2453.1 - 52.098 \\ &= 2401.002 \implies 2401.002 \times 10,000 = \boxed{\text{24,010,020 cars}} \end{aligned}$$

(f)

$$5464.4 - 99.653, 5464.4 + 9.931 = (5364.748, 5474.331)$$

$\Longrightarrow$ **Between 53,647,480 and 54,743,310 cars**

(g)

We are 95% certain that between 120,660,520 and 122,014,900 cars will be owned in China in the year 2013, if present trends continue.