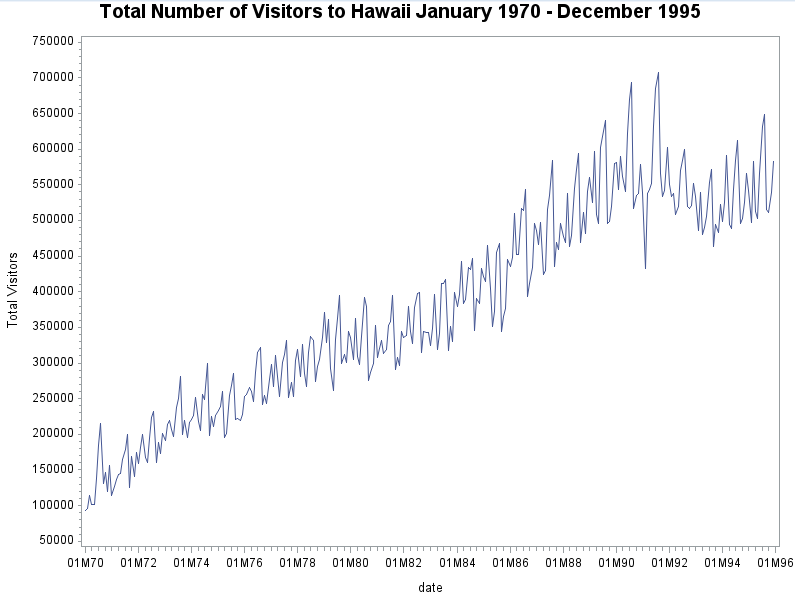
Vicki Bergquist

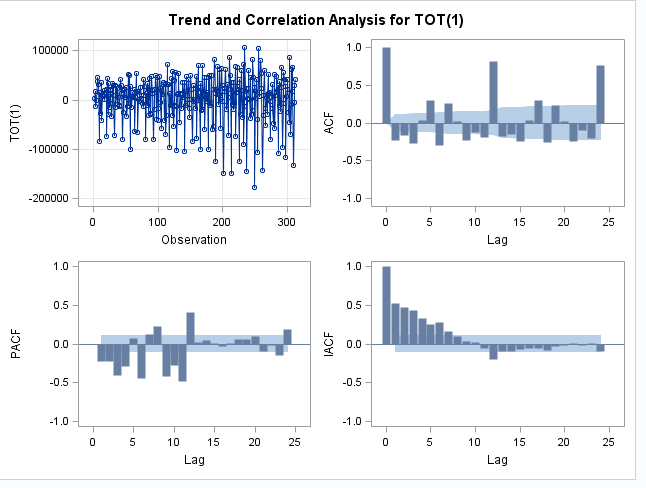
Take-Home Exam

“I have received no assistance on this exam except from Professor Fomby and/or Yixiang Zhang”

a)

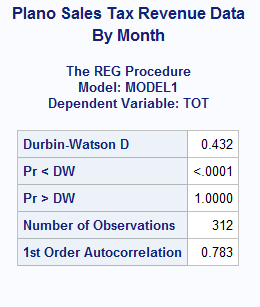
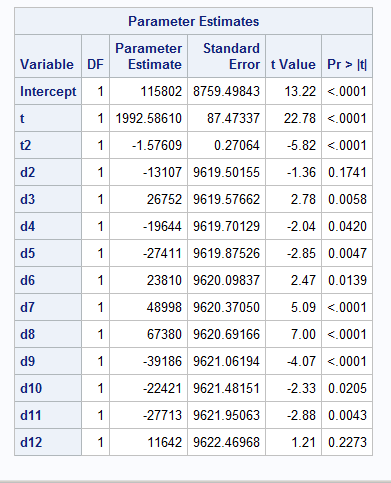
Plot of the Raw Data





Yes there seems to be seasonality in this data. After using differences to make the data stationary, the autocorrelation function shows there are sample ACF lags at 6, 12, 18 and 24. Based on the graph, it would be a statistically significant positive correlation on the 4th lag, follow by a negative correlation. This would be an indication of a smooth upward or downward movement, followed by a saw-tooth movement, which is repeated again with another statistically significant smooth upward or downward movement 8 periods later. The repeated smooth change in direction followed by a saw-tooth movement would indicates the data would respectively move upward or downward before changing directions again repeatedly over time, which is a strong indication of seasonality. This would mean within a 12 month period there would be the same repeated saw-tooth movement upward/downward, in other words, seasonality.

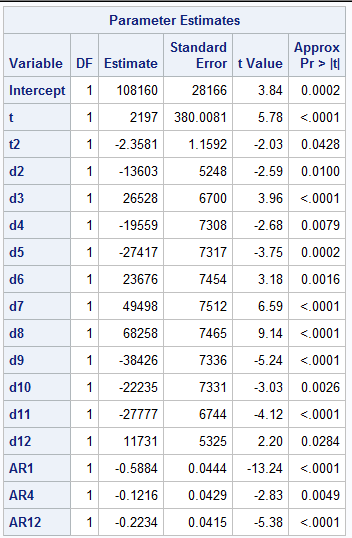
b)



Yes there seems to be autocorrelation based on the Durbin-Watson D. The Durbin-Watson D is an indication if there are autocorrelation in the data, otherwise, the errors aren’t just white noise but have a pattern. The test shows the autocorrelation is positive since the Durbin-Watson D is below the null hypothesis of D =2, with a statistically significant p-value of less than .05.

Since there is autocorrelation in this data, we can’t test the trend values because the data isn’t stationary. We would need to transform the data to make the data by pulling out the stochastic cycles from the errors. In this case, it would be using SAS Autoreg to transform the data by GLS and give us the AR(r) from the errors. Once the data is transformed, the estimates would be more reliable for testing statistical significance of the trend and seasonality.

c.



The above output is the estimated model from the Proc Autoreg. The model indicates that there is a statistically significant trend in the data. Both t and t2 have p-values that are less than the p-critical value of .05.

The maximum or minimum of this trend would be the following:

D = 108160 + 2197t -2.3581t^2

D’ = 2197-2\*2.3581t

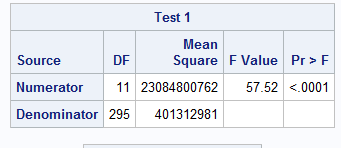
0 = 2197 – 4.7162t

-2197/-4.7162 = t

T = 465.8 th observation period of the trend

We know this is the maximum because D’’ = -4.7162, which is a negative number indicating D’ would give us the maximum.

d.



The null hypothesis of this test is that all the seasonal time period would be no different from January. This meaning all the monthly periods would be statistically insignificant. The alternative hypothesis would be there is at least one of the monthly dummies that is statistically significant from zero.

H0: d2=d3=d4… d12= 0

H1: At least one of the months is statistically significant from zero. There is seasonality in the data.

Based on this F Joint test above, I can reject the null hypothesis based on a p critical value of .05. There is seasonality in this data.

e.

yhat(t) = 108160+2197t - 2.3581t^2 - 13603\*d2 + 26528\*d3 -19559\*d4 -27417\*d5+23676\*d6

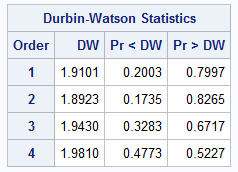
(28166) (380.0081) (1.1592) (5248) (6700) (7308) (7317) (7454)

+49498\*d7+68258\*d8 -38426\*d9-22235\*d10-27777\*d11+11731\*d12 + E(t)

(7512) (7465) (7336) (7331) (6744) (5325)

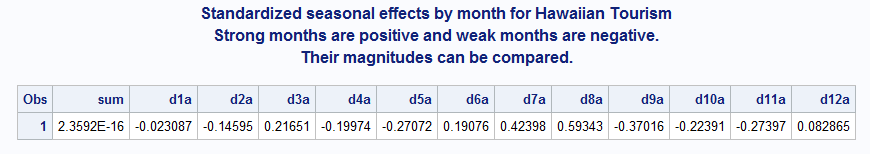
E(t) =.5884E(t-1) + .1216(t-4)+.2234(t-12)+a(t)

(.0444) (.0429) (.0415)



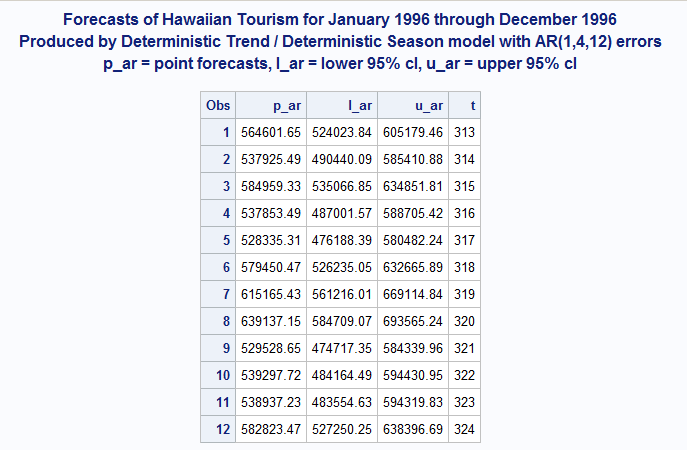
Based on the Durbin Watson Statistics, this model doesn’t have any statistically significant autocorrelations. The residuals of this model are white noise, otherwise, they are completely random in either direction of positive or negative. Another way of saying this is that there is no pattern in how a residual (error) would relate to another residual (error).

f.



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | D1a | D2a | D3a | D4a | D5a | D6a | D7a | D8a | D9a | D10a | D11a | D12a |
| (i) | Weak | Weak | Strong | Weak | Weak | Strong | Strong | Strong | Weak | Weak | Weak | Strong |
| (ii) |  |  |  |  |  |  | Strongest of all tourists months |  | Weakest of all tourist months |  |  |  |

g.



Total 1996 Forecasted Revenues = 6778015.3947

Total 1995 Actual Revenues = 6633840

Forecasted Change = 6778015.3947-6633840/6633840 = 2.17% Increase

The forecast of 1996 estimate a growth in tourism from the previous year of 1995. There was a forecasted increase of 2.17% in 1996 over the previous actuals of 1995.

APPENDIX

--- My Code Starts From Here ----

/\*question a \*/

**proc** **arima** data=travel;

identify var=TOT(**1**);

estimate q=(**0**)(**0**);

**run**;

/\*question b\*/

/\* added trend and seasonality coefficients to create DTDS model. After adding the coefficients, used proc reg to apply

an OLS method to get the DTDS model\*/

**DATA** travel;

SET travel;

t = \_n\_;

t2 = t\*t;

d1 = (month=**1**);

d2 = (month=**2**);

d3 = (month=**3**);

d4 = (month=**4**);

d5 = (month=**5**);

d6 = (month=**6**);

d7 = (month=**7**);

d8 = (month=**8**);

d9 = (month=**9**);

d10 = (month=**10**);

d11 = (month=**11**);

d12 = (month=**12**);

**run**;

**proc** **reg** data = travel;

model TOT = t t2 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12 / dwProb;

test d2, d3, d4, d5, d6, d7, d8, d9, d10, d11, d12;

**run**;

/\*question c & d & e\*/

/\*Test for curvature of the data. The above test shows that there is significant correlations in the model and

the trend can't be tested. Must applied a transformation of the data via ProcAutoreg to test the statistical

significance of the trend and white noise. In addition, test of seasonality by the joint F-test and get the output from

our forecast\*/

**proc** **autoreg** data = travel outest = coeff;

model TOT = t t2 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12 / nlag = **12** DW=**4** dwProb

method=ml backstep slstay=**0.05**;

test d2, d3, d4, d5, d6, d7, d8, d9, d10, d11, d12;

**run**;

/\*question f\*/

/\* based on the coefficents from the autoreg, generate standarized seasonal effects\*/

**data** coeff;

set coeff;

seasonsum = (**12**\*intercept + d2 + d3 + d4 + d5 + d6 + d7 + d8 + d9 + d10 + d11 + d12);

seasonave = seasonsum/**12**;

d1a = (intercept - seasonave)/seasonave;

d2a = (intercept + d2 - seasonave)/seasonave;

d3a = (intercept + d3 - seasonave)/seasonave;

d4a = (intercept + d4 - seasonave)/seasonave;

d5a = (intercept + d5 - seasonave)/seasonave;

d6a = (intercept + d6 - seasonave)/seasonave;

d7a = (intercept + d7 - seasonave)/seasonave;

d8a = (intercept + d8 - seasonave)/seasonave;

d9a = (intercept + d9 - seasonave)/seasonave;

d10a = (intercept + d10 - seasonave)/seasonave;

d11a = (intercept + d11 - seasonave)/seasonave;

d12a = (intercept + d12 - seasonave)/seasonave;

sum = d1a + d2a + d3a + d4a + d5a + d6a + d7a + d8a + d9a + d10a + d11a + d12a;

**run**;

title1 'Standardized seasonal effects by month. They sum to zero.';

title2 'Strong months are positive and weak months are negative.';

title3 'Their magnitudes can be compared.';

**proc** **print** data = coeff;

var sum d1a d2a d3a d4a d5a d6a d7a d8a d9a d10a d11a d12a;

**run**;

/\* question g\*/

/\* Pull estimated model into a base model. Add 12 additional months to the model and forecast out using AutoReg.\*/

**data** base;

set Travel;

keep TOT t t2 d1 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12;

**data** add;

input TOT t t2 d1 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12;

datalines;

. 313 97969 1 0 0 0 0 0 0 0 0 0 0 0

. 314 98596 0 1 0 0 0 0 0 0 0 0 0 0

. 315 99225 0 0 1 0 0 0 0 0 0 0 0 0

. 316 99856 0 0 0 1 0 0 0 0 0 0 0 0

. 317 100489 0 0 0 0 1 0 0 0 0 0 0 0

. 318 101124 0 0 0 0 0 1 0 0 0 0 0 0

. 319 101761 0 0 0 0 0 0 1 0 0 0 0 0

. 320 102400 0 0 0 0 0 0 0 1 0 0 0 0

. 321 103041 0 0 0 0 0 0 0 0 1 0 0 0

. 322 103684 0 0 0 0 0 0 0 0 0 1 0 0

. 323 104329 0 0 0 0 0 0 0 0 0 0 1 0

. 324 104976 0 0 0 0 0 0 0 0 0 0 0 1

;

**proc** **append** base = base data = add;

**run**;

**proc** **autoreg** data=base;

model TOT = t t2 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12/ nlag = **12** DW=**4** DWPROB method=ml backstep slstay = **0.05**;

output out=forecast alphacli = **0.05** p=p\_ar lcl=l\_ar ucl=u\_ar;

**run**;

**data** forecast;

set forecast;

if \_n\_ > **312**;

keep t p\_ar l\_ar u\_ar;

**run**;

title1 'Forecasts of Hawaiian Tourism for January 1996 through December 1996';

title2 'Produced by Deterministic Trend / Deterministic Season model with AR(1,4,12) errors';

title3 'p\_ar = point forecasts, l\_ar = lower 95% cl, u\_ar = upper 95% cl ';

**proc** **print** data=forecast;

**run**;

**Proc** **SQL**;

Create table Sum1996 as

Select Sum(p\_ar)as TotalForecasted

from forecast;

**Proc** **SQL**;

Create table Sum1995 as

select year, sum(TOT) as TotalActual

from Travel

where year = **95**

group by year;