Seasonality

- Recall that we said that it can be useful to describe the mean of a time series as the sum of components

$$\mu_t = T_t + S_t + C_t$$

  where S_t is the seasonal component.
- The seasonal component S_t is a repetitive cycle over the calendar year
- Seasonality S_t can be deterministic (predictable) or stochastic

Seasonality – Examples

- Gasoline consumption rises in summer due to increased auto travel
- International airline prices rise in summer due to increased tourism
- Natural gas consumption and prices rise in winter due to heating
- Electricity consumption increases in summer due to air conditioning
- Construction activity and jobs decrease in winter in the Midwest
- Consumer spending increases in November and December due to holiday shopping

Deterministic vs Stochastic Seasonality

- If the seasonal pattern repeats year after year, it is deterministic and predictable.
  - Christmas is always in December
- If the seasonal pattern roughly repeats itself, but evolves over the years, it is stochastic and only partially predictable
  - Holiday shopping as a percentage of income is not a fixed constant
- Seasonal patterns can change dramatically as the economy evolves
  - The spread of air conditioning shifted the seasonal pattern of residential electricity consumption from winter to summer

Seasonal Adjustment

- Most economic indicators reported by the government are **seasonally adjusted**.
- Roughly, the component S_t is estimated, and then what is reported is

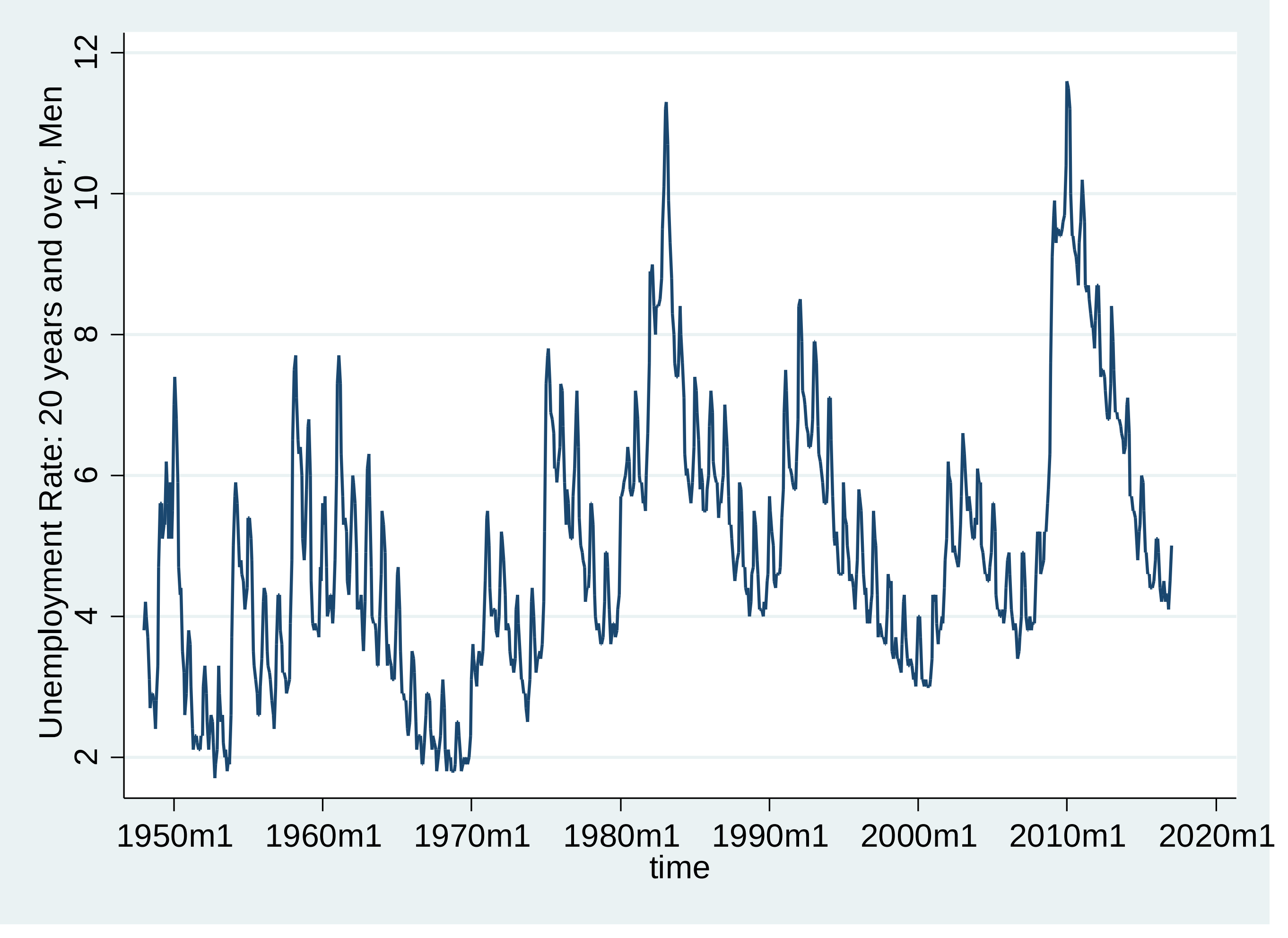
$$\begin{aligned} y_t^* &= y_t - S_t \\ &= T_t + C_t \end{aligned}$$

- The idea is that seasonality distracts from the main reporting purpose
  – Seasonally adjusted data allows users to focus on trend and business cycle movements
- Seasonal adjustment by central statistical agencies is sophisticated, allowing for evolving seasonal patterns.

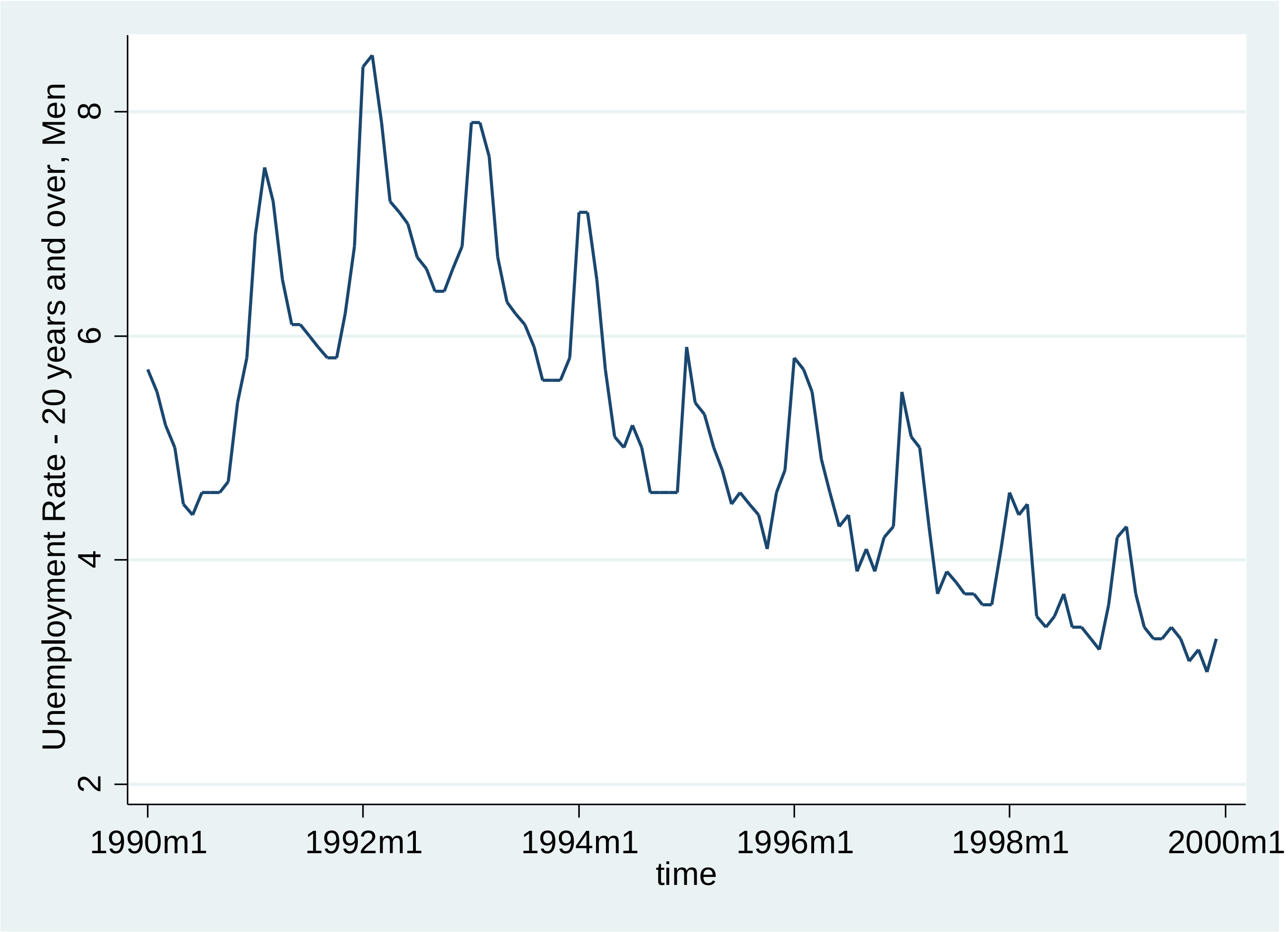
Examples of Seasonal Time Series

- Example 1:
- U.S. Unemployment Rate
  - Men, 20+ years
  - 1948-2009
  - Not seasonally adjusted

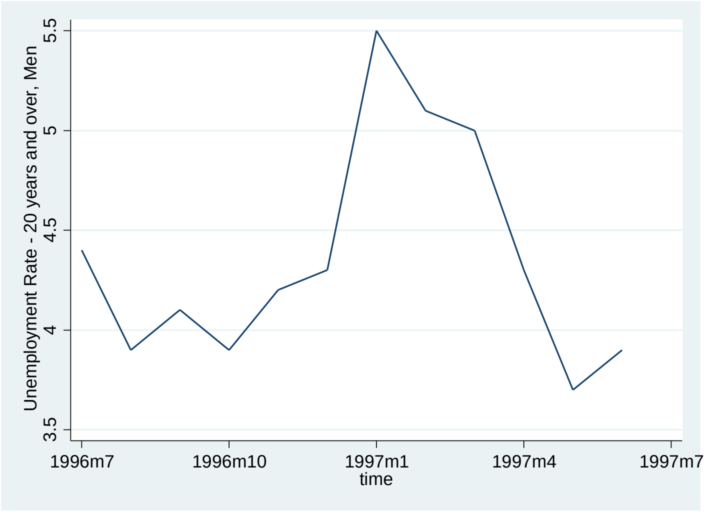
U.S. Unemployment Rate
Men, 20+ years, 1948-current

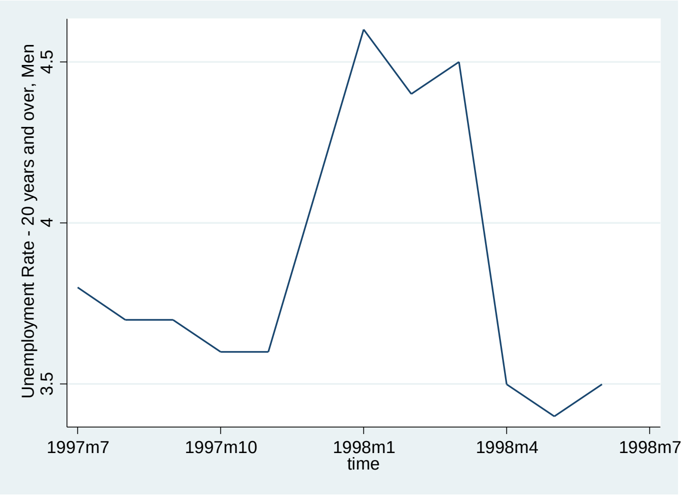


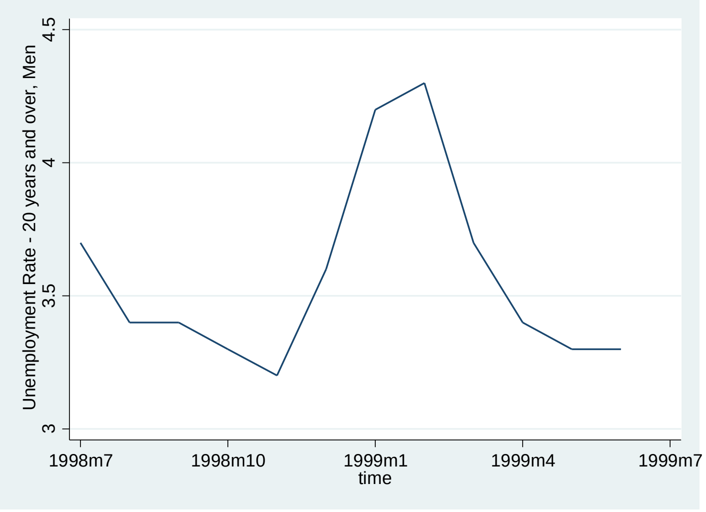
Unemployment Rate, 1990-1999

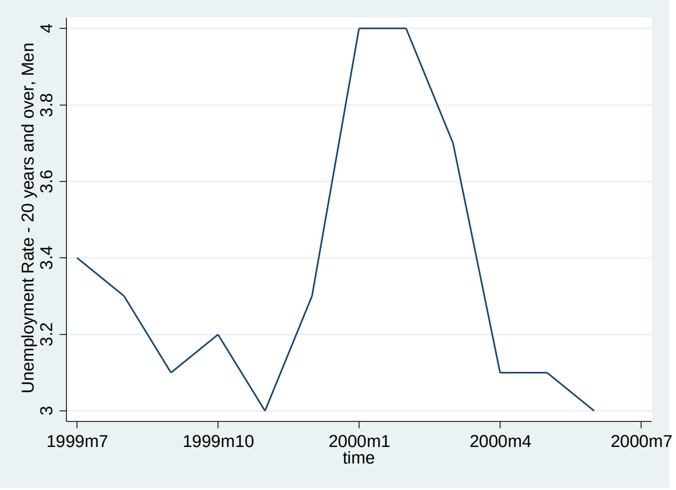


Unemployment Rate, by year





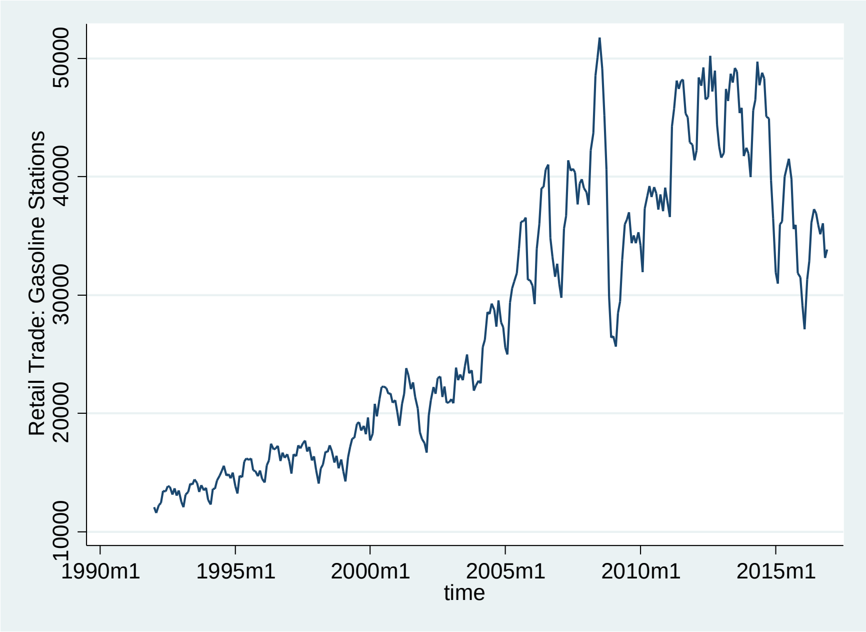


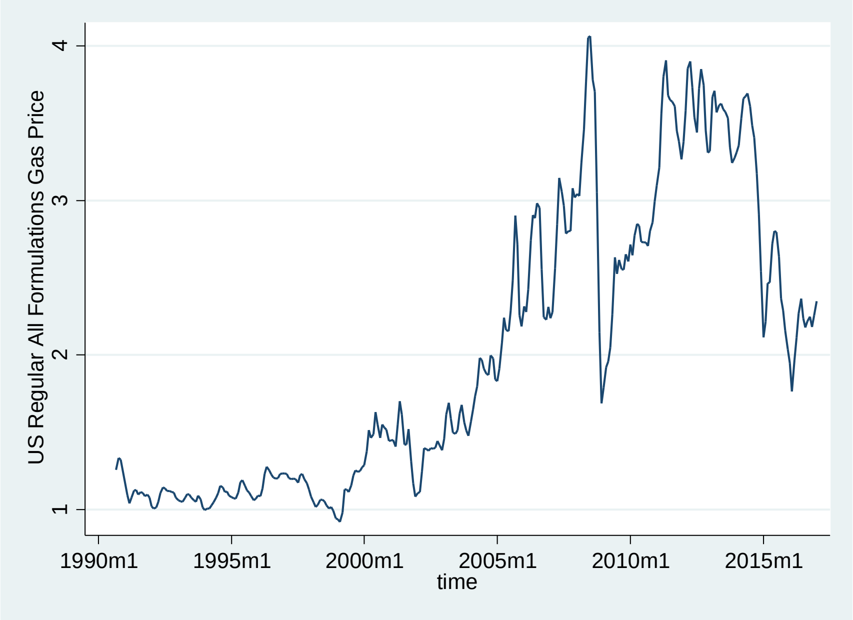


Examples of Seasonal Time Series

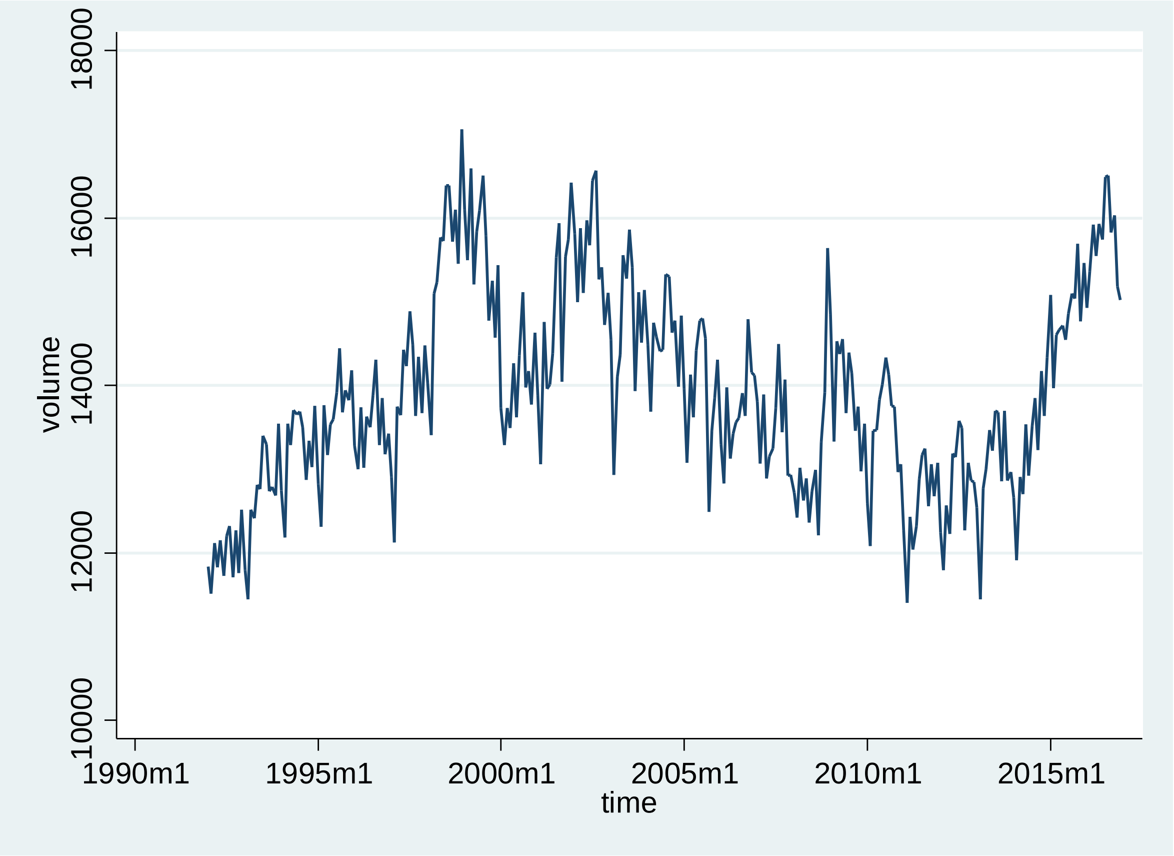
- Example 2:
- U.S. Gasoline Sales
  - RSGASSN, retail sales, $million
  - GASREGM, price ($) per gallon
  - volume= RSGASSN/ GASREGM (millions of gallons)
  - Monthly, 1992-current

U.S. Gasoline Retail Sales ($) & Price per Gallon

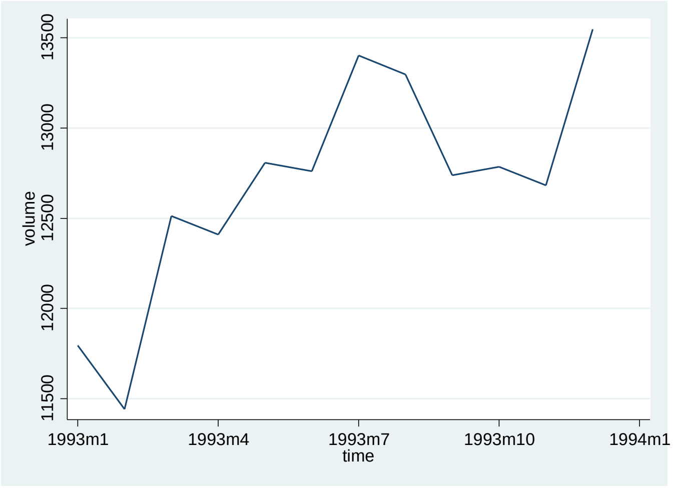


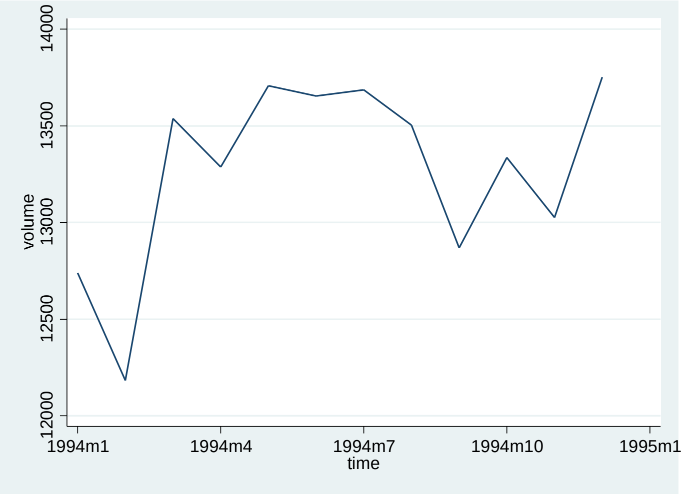


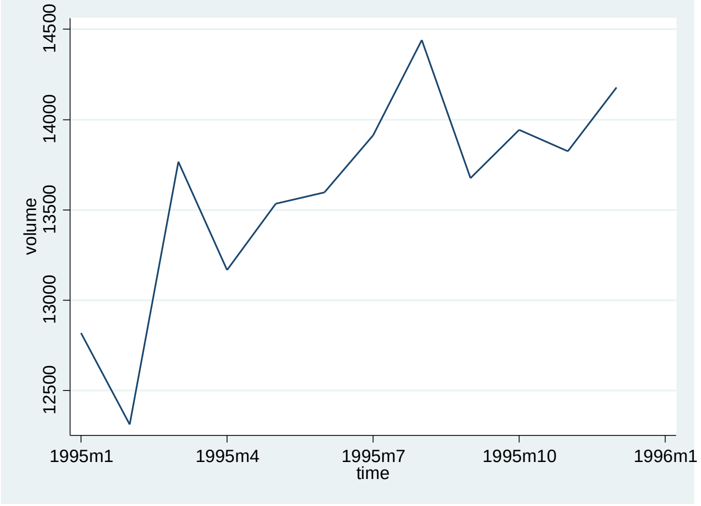
Gasoline Sales Volume (million of gallons)

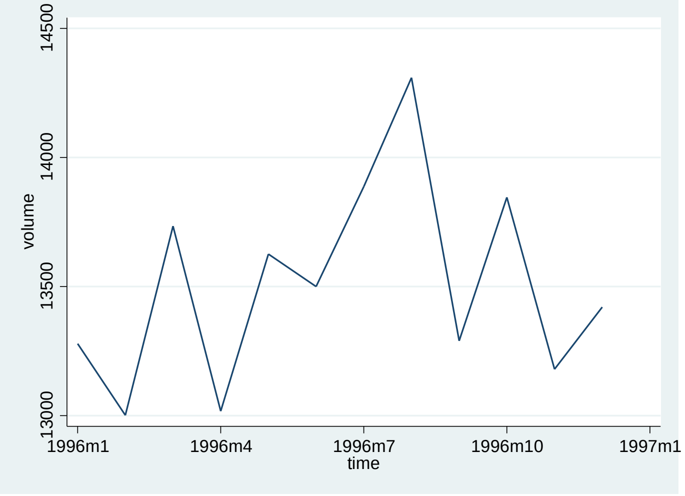


Gasoline Sales, by year

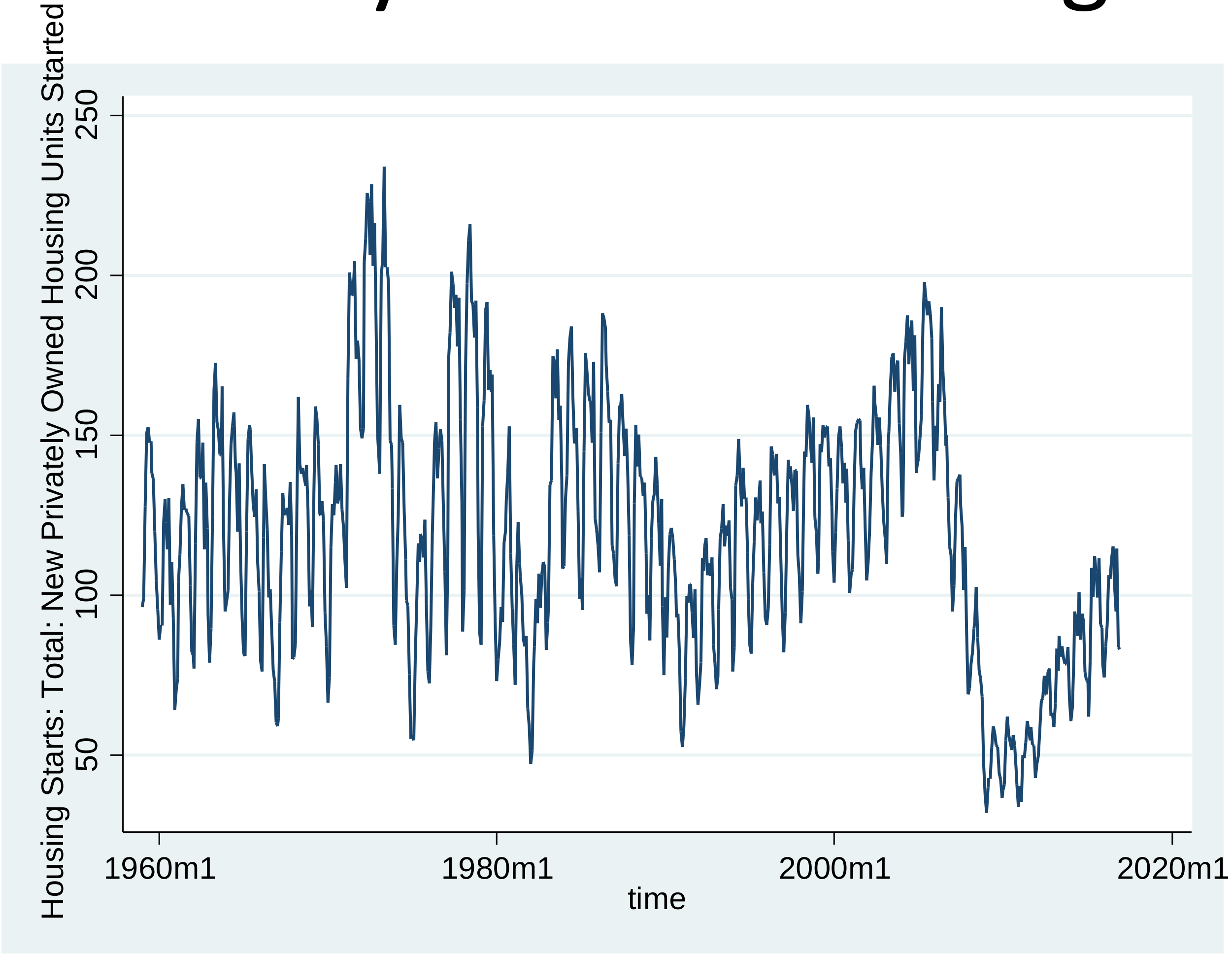




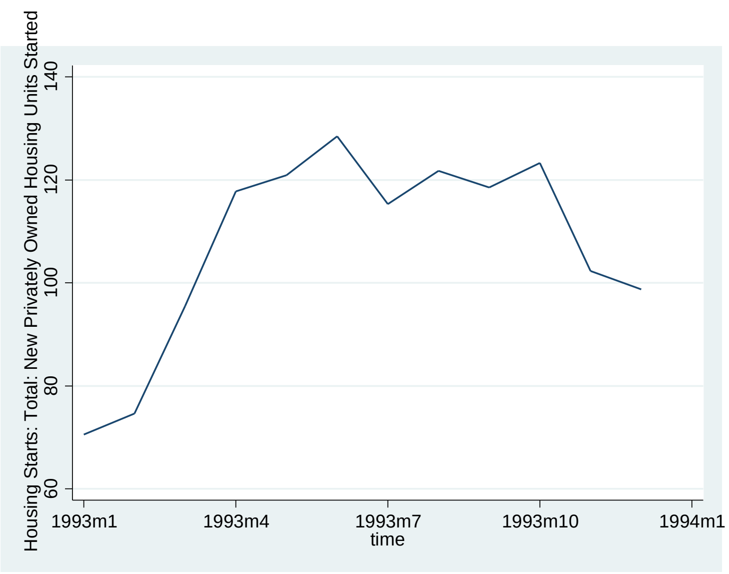


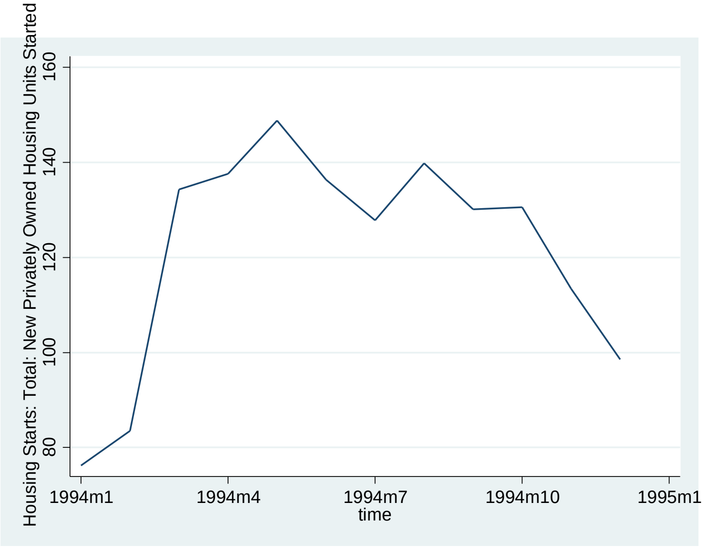


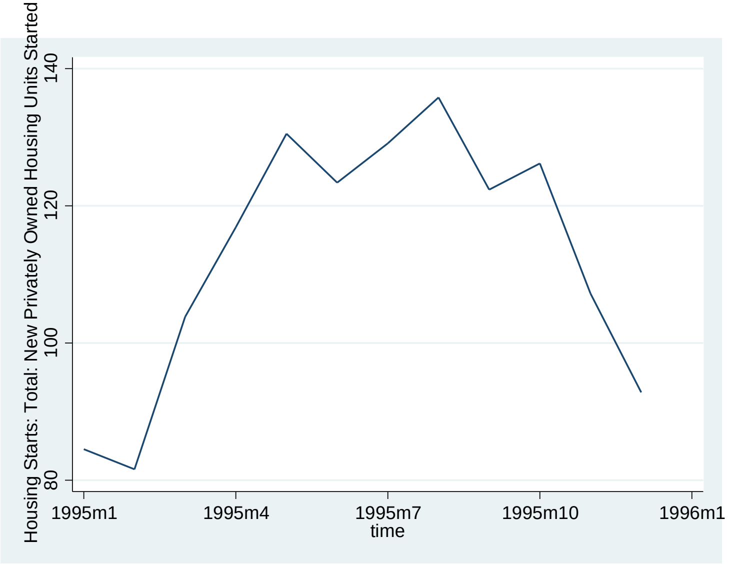
Example 3
U.S. Housing Starts
(New Privately Owned Housing Units)

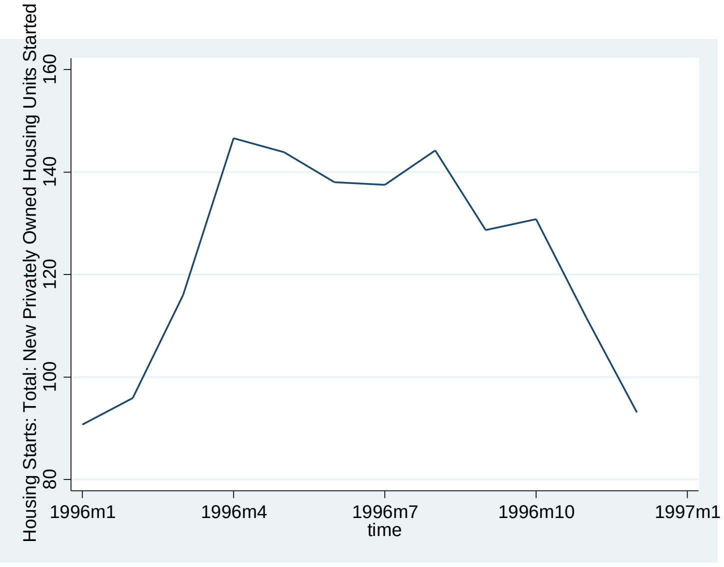


Housing Starts, by year









Deterministic Seasonality

- If seasonality is constant and deterministic then S_t is simply a different constant for each period
- For example, for monthly data

$$S_t = \begin{cases} \gamma_1 & if \quad t = January \\ \gamma_2 & if \quad t = February \\ \vdots & \qquad \vdots \\ \gamma_{12} & if \quad t = December \end{cases}$$

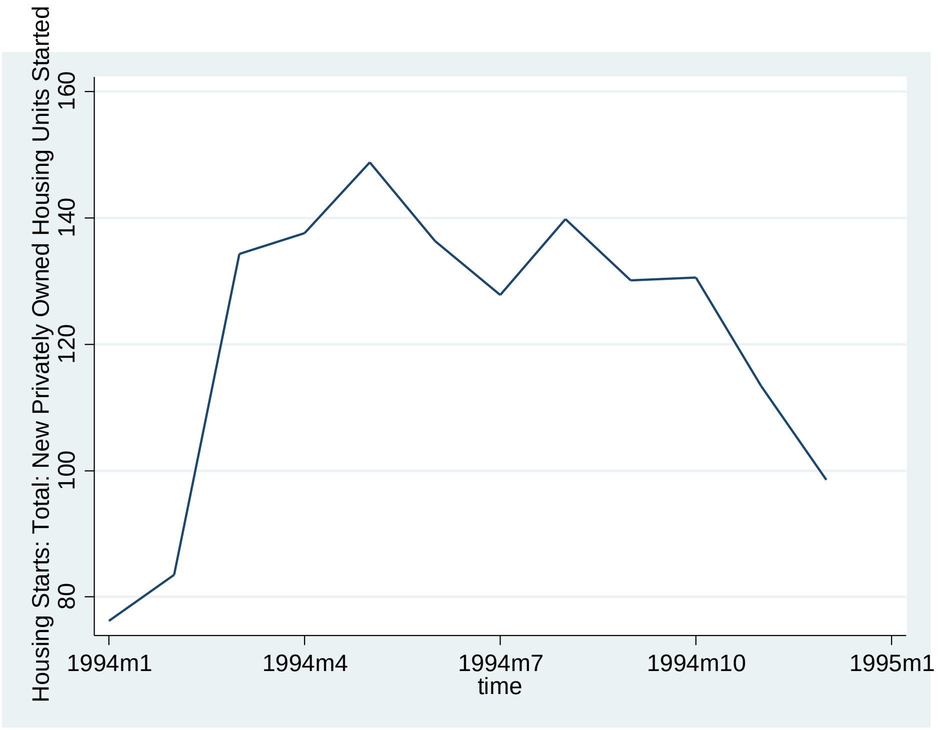
- Seasonality is a constant which varies by the calendar period (quarter, month, week, day, or time of day)

Fitted Values and Forecasts
Pure Deterministic Seasonality

- In the simple pure deterministic seasonality model, fitted values and forecasts are the simple seasonal pattern

Example – Housing Starts

January	76
February	83
March	134
April	138
May	149
June	136
July	128
August	140
September	130
October	131
November	113
December	98



Seasonal Dummy Model

- Deterministic seasonality S_t can be written as a function of seasonal dummy variables
- Let s be the seasonal frequency
  - s=4 for quarterly
  - s=12 for monthly
- Let D_{1t} , D_{2t} , D_{3t} ,…, D_{st} be seasonal dummies
  - D_{1t}= 1 if s is the first period, otherwise D_{1t}= 0
  - D_{2t}= 1 if s is the second period, otherwise D_{2t}= 0
- At any time period t, one of the seasonal dummies D_{1t} , D_{2t} , D_{3t} ,…, D_{st} will equal 1, all the others will equal 0.

Seasonal Dummy Model

- Deterministic seasonality

$$S_t = \begin{cases} \gamma_1 & if \quad t = January \\ \gamma_2 & if \quad t = February \\ \vdots & \qquad \vdots \\ \gamma_{12} & if \quad t = December \end{cases}$$

$$= \sum_{i=1}^{s} \gamma_i D_{it}$$

a linear function of the dummy variables

Estimation

- Least squares regression

$$y_{t+h} = \sum_{i=1}^{s} \gamma_i D_{it} + e_t$$

$$= \alpha + \sum_{i=1}^{s-1} \beta_i D_{it} + e_t$$

- You can either
  - Regress y on all the seasonal dummies, omitting the intercept, or
  - Regress y on an intercept and the seasonal dummies, omitting one dummy (one season, e.g. December)
- You cannot regress on both the intercept plus all seasonal dummies, for they would be collinear and redundant.

Interpreting Coefficients

- In the model

$$S_t = \alpha + \sum_{i=1}^{s-1} \beta_i D_{it}$$

  the intercept $\alpha = \gamma_s$ is the seasonality in the omitted season.
- The coefficients $\beta_i = \gamma_i - \gamma_s$ are the difference in the seasonal component from the s'th period.

STATA Programming

- If the time index is *time* and is formatted as a time index, you can determine the period using the commands

```
generate m=month(dofm(time))
generate q=quarter(dofq(time))
```

for monthly and quarterly data, respectively

(See dates and times in STATA Data manual)

Creating Dummies

- If *m* is the month (1 for January, 2 for February, etc.), then
  - **generate m1=(m==1)**
  - This creates a dummy variable "m1" for January
  - Then
  - **regress y m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11**
    or
  - **regress y m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12, noconstant**
- Easier
  - Type "b12.m" in the regressor list
  - **regress y b12.m**
  - This includes dummies for months 1 through 11, omits 12
  - Same as mechanically listing the eleven dummies, but easier.
  - It is important that "m" be the numerical month (1 for January, 2 for February, etc.)

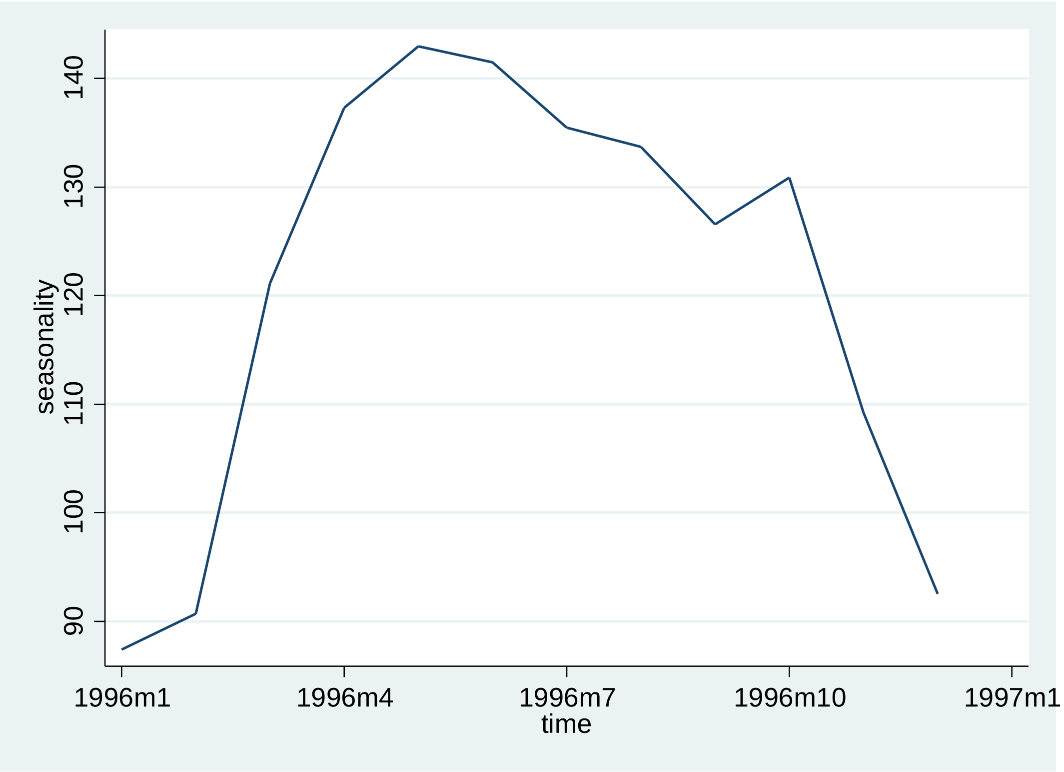
Example 3: Housing Starts

- gen m = month(dofm(time))
- regress starts b12.m

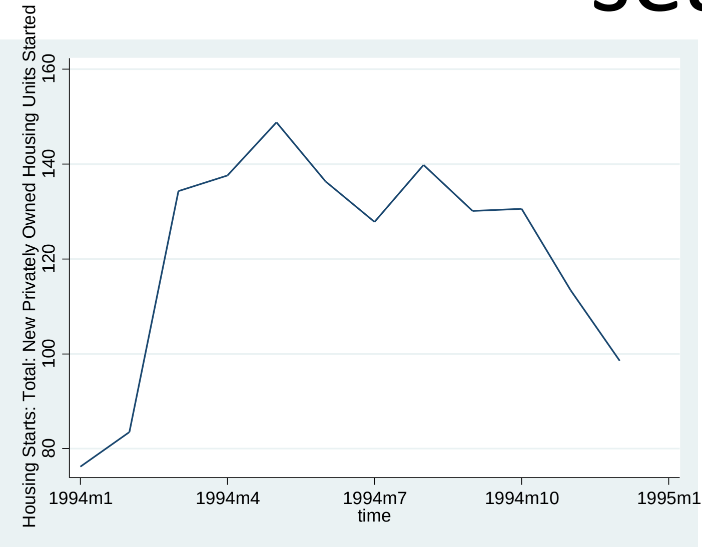
starts	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-5.125	6.515521	-0.79	0.432	-17.91865	7.668648
2	-1.841071	6.515521	-0.28	0.778	-14.63472	10.95258
3	28.59464	6.515521	4.39	0.000	15.80099	41.38829
4	44.76429	6.515521	6.87	0.000	31.97064	57.55793
5	50.42857	6.515521	7.74	0.000	37.63492	63.22222
6	48.99821	6.515521	7.52	0.000	36.20457	61.79186
7	42.95179	6.515521	6.59	0.000	30.15814	55.74543
8	41.19286	6.515521	6.32	0.000	28.39921	53.98651
9	34.03929	6.515521	5.22	0.000	21.24564	46.83293
10	38.35893	6.515521	5.89	0.000	25.56528	51.15258
11	16.71964	6.515521	2.57	0.011	3.925994	29.51329
_cons	92.52143	4.607169	20.08	0.000	83.47495	101.5679

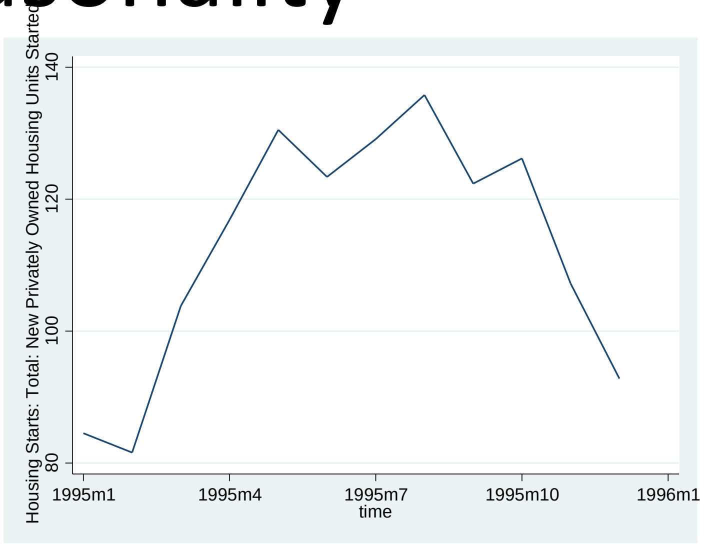
Estimated Seasonality – Housing Starts

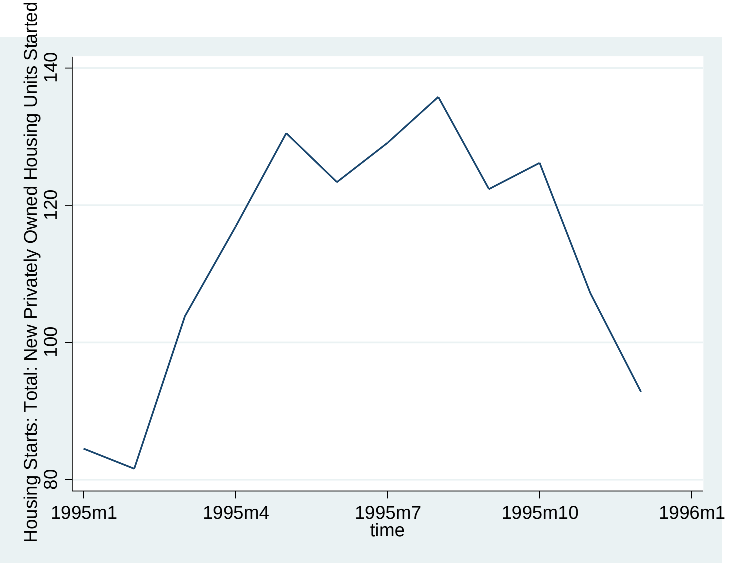
January	87
February	91
March	121
April	137
May	143
June	142
July	135
August	134
September	127
October	131
November	109
December	93

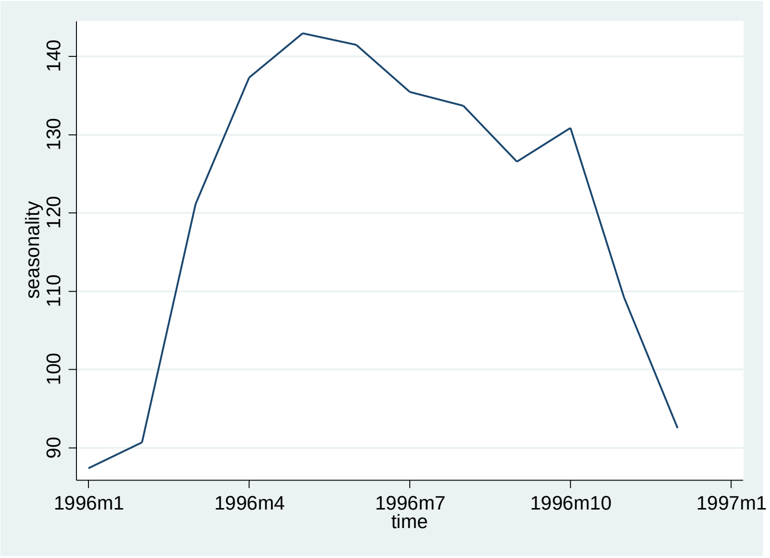


Housing Starts, by year and estimated seasonality

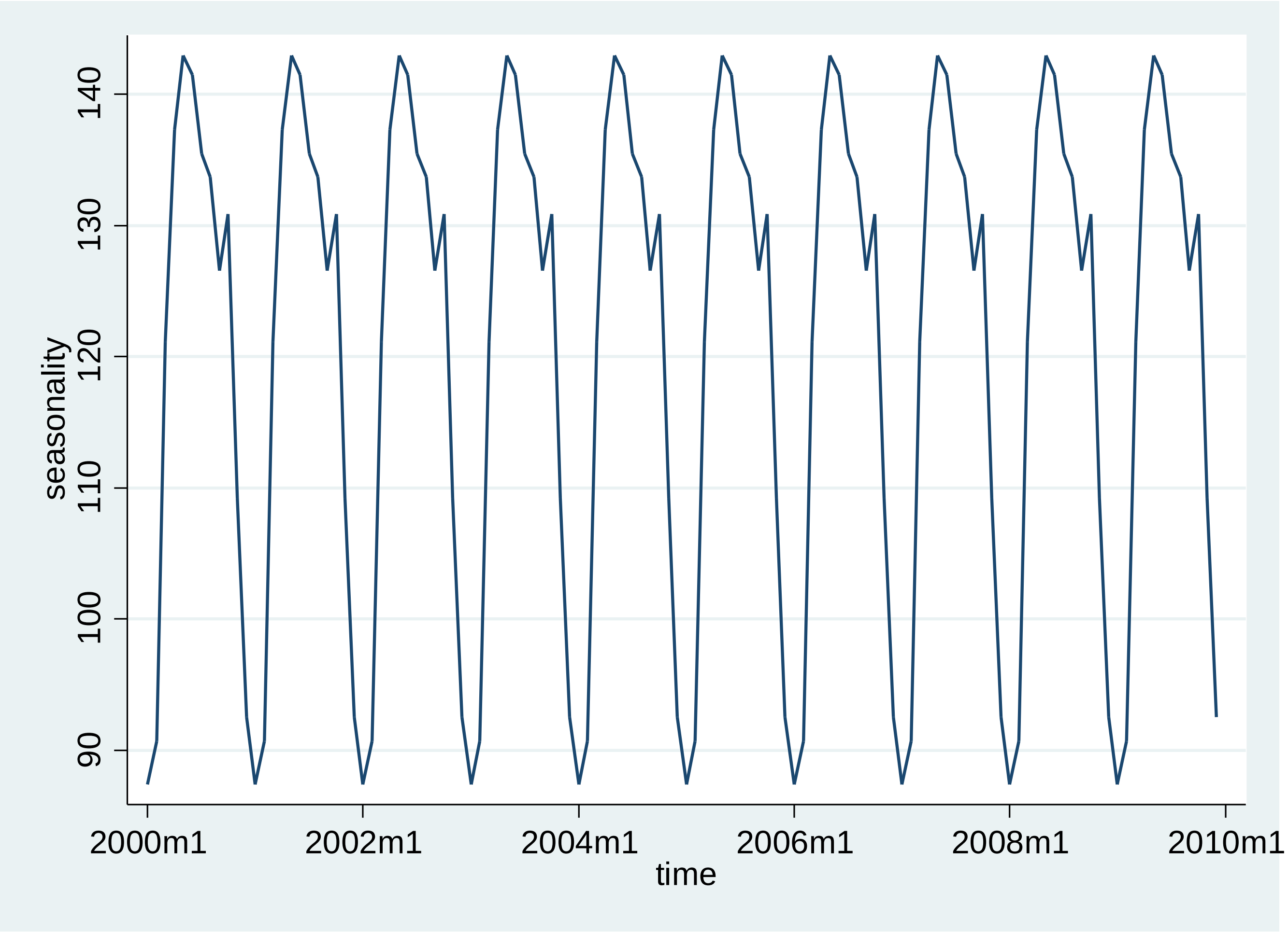






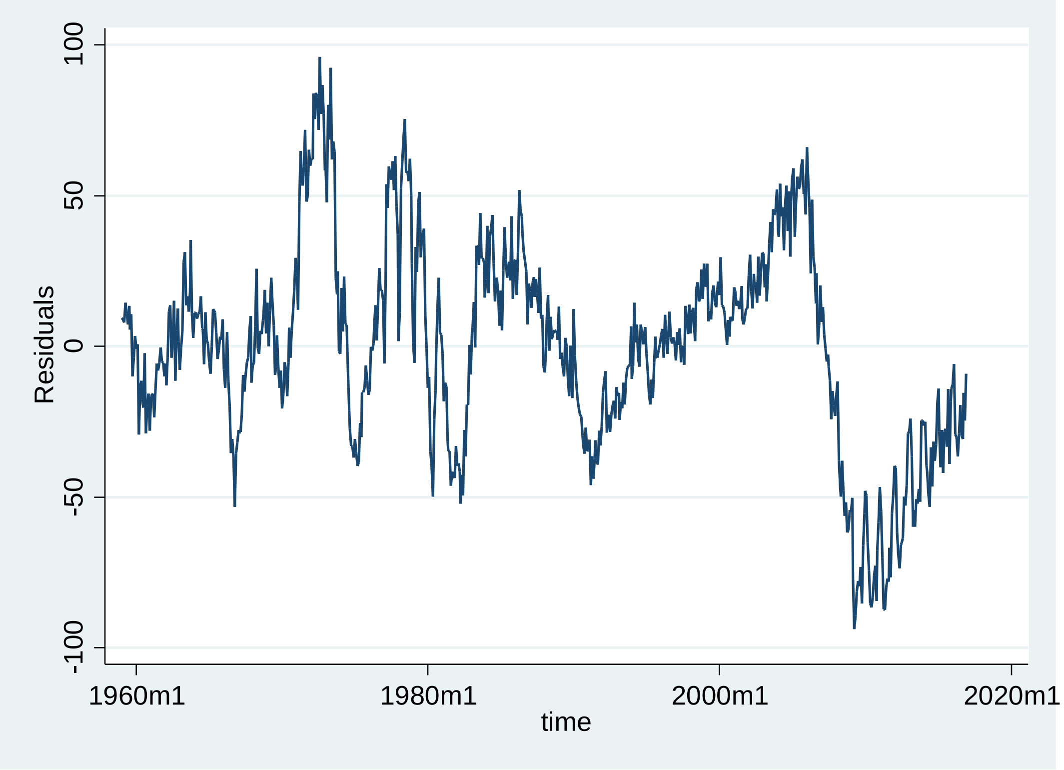


Predicted Values



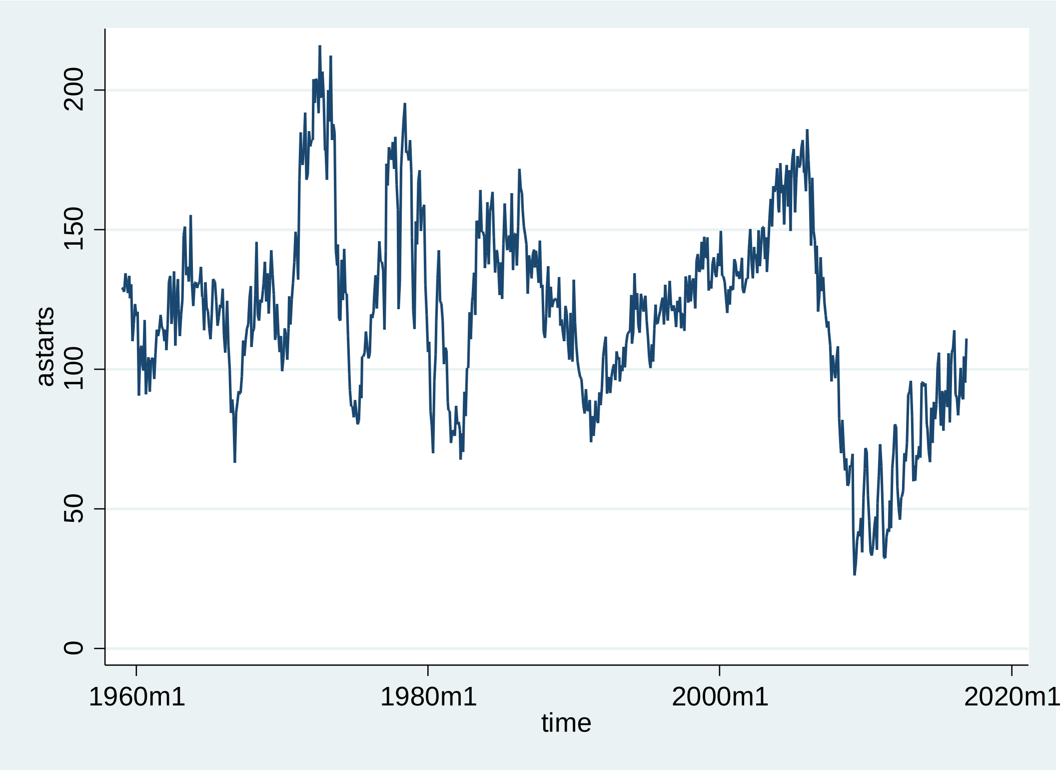
Residual Plots

- Plot the residuals after the seasonal dummy regression, to see if the seasonality has largely been modeled
- If seasonality has changed, the residuals will still display seasonality
  - predict e, residuals
  - tsline e



DeSeasonalized

- Original series less seasonal
- In seasonal dummy model, equals residual plus the sample mean
- egen mu = mean(starts)
- gen astarts = mu + e
- tsline astarts

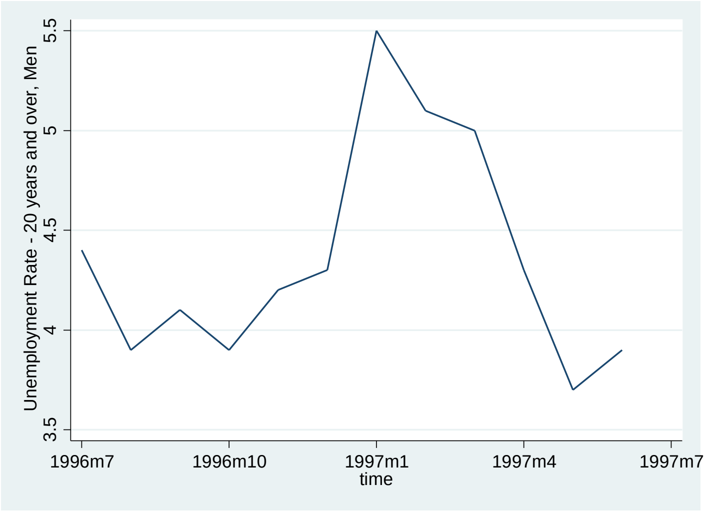


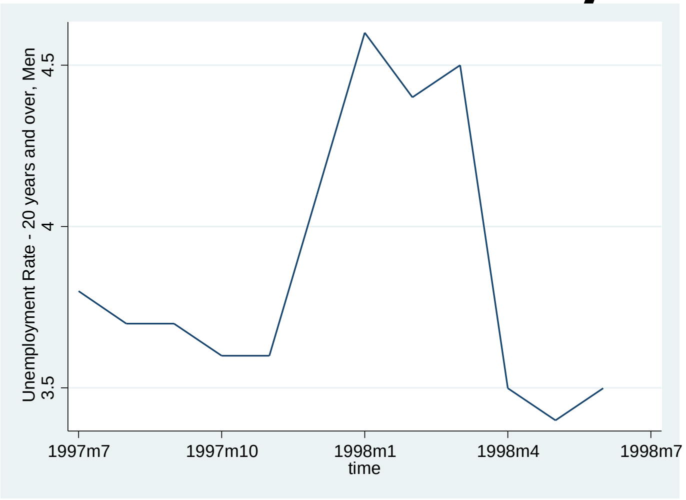
Example 1
Unemployment Rate

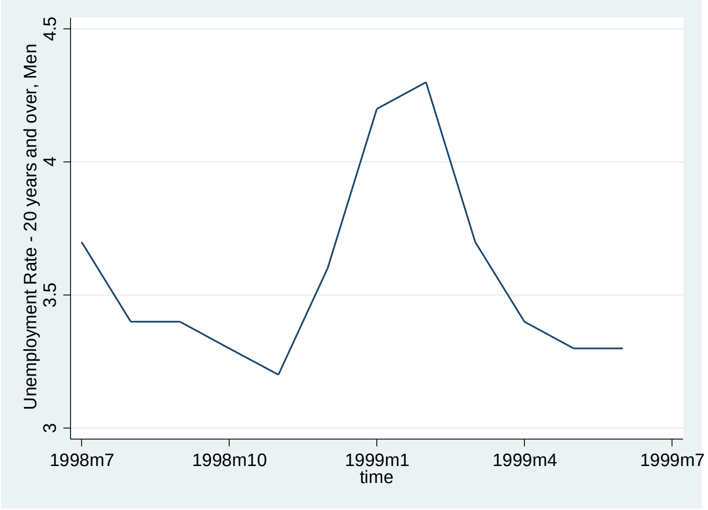
- gen m = month(dofm(time))
- regress men b12.m

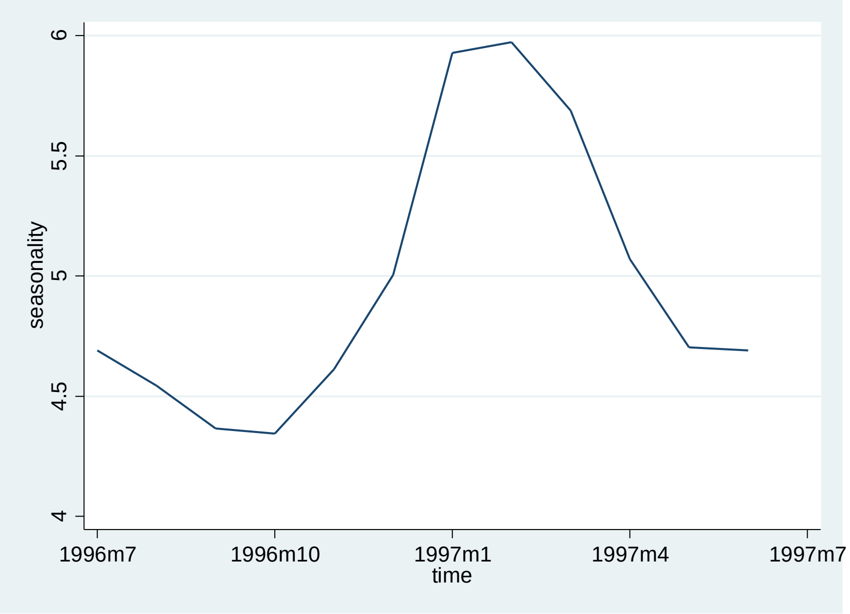
men	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	.9238806	.3142008	2.94	0.003	.3071158	1.540645
2	.9686567	.3142008	3.08	0.002	.351892	1.585421
3	.6835821	.3142008	2.18	0.030	.0668173	1.300347
4	.0656716	.3142008	0.21	0.834	-.5510931	.6824364
5	-.3014925	.3142008	-0.96	0.338	-.9182573	.3152722
6	-.3134328	.3142008	-1.00	0.319	-.9301976	.3033319
7	-.3149254	.3142008	-1.00	0.317	-.9316901	.3018394
8	-.461194	.3142008	-1.47	0.143	-1.077959	.1555707
9	-.638806	.3142008	-2.03	0.042	-1.255571	-.0220412
10	-.6597015	.3142008	-2.10	0.036	-1.276466	-.0429367
11	-.3925373	.3142008	-1.25	0.212	-1.009302	.2242274
_cons	5.004478	.2221735	22.53	0.000	4.568359	5.440596

Unemployment Rate, by year, and estimated seasonality

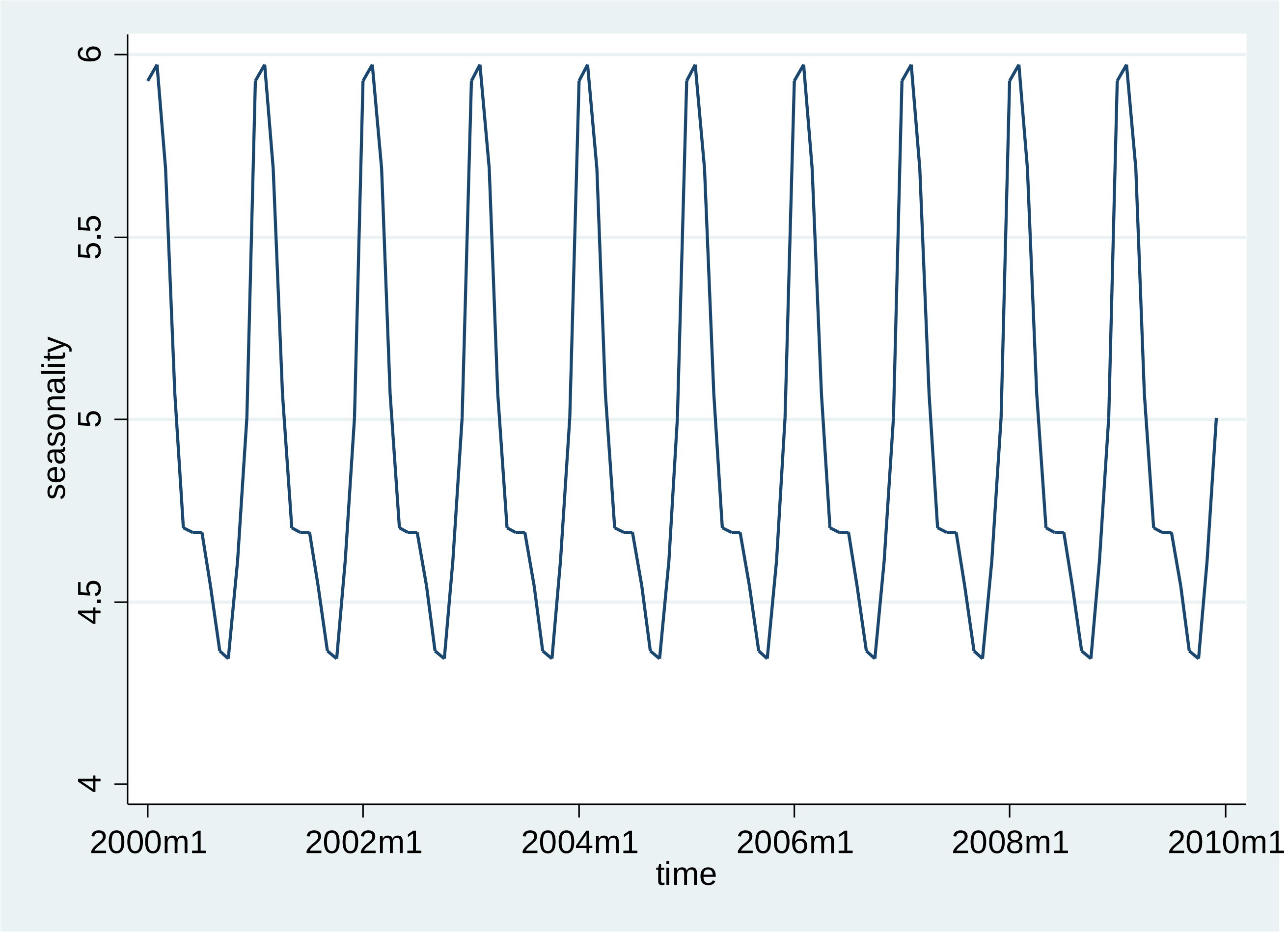




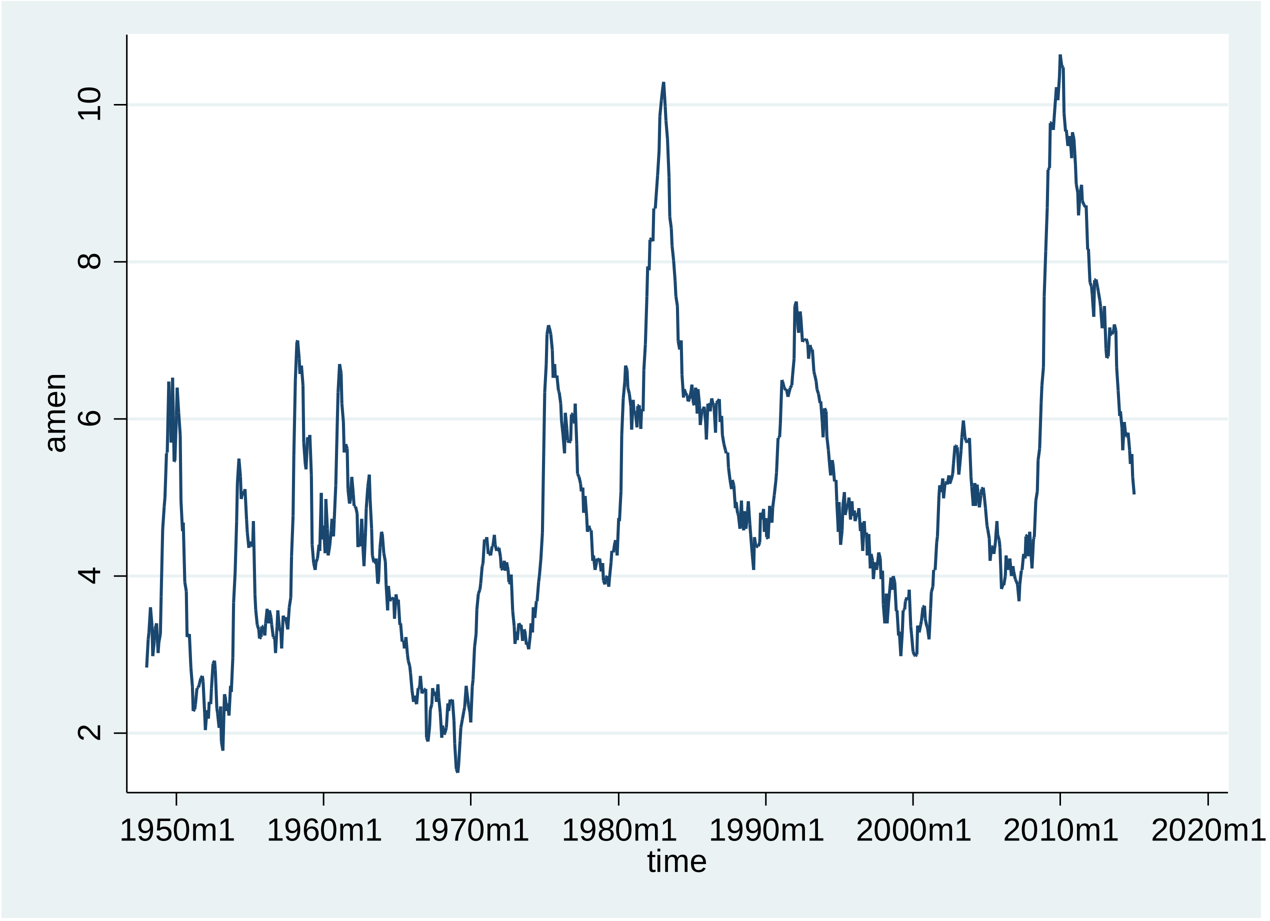




Predicted Values



DeSeasonalized

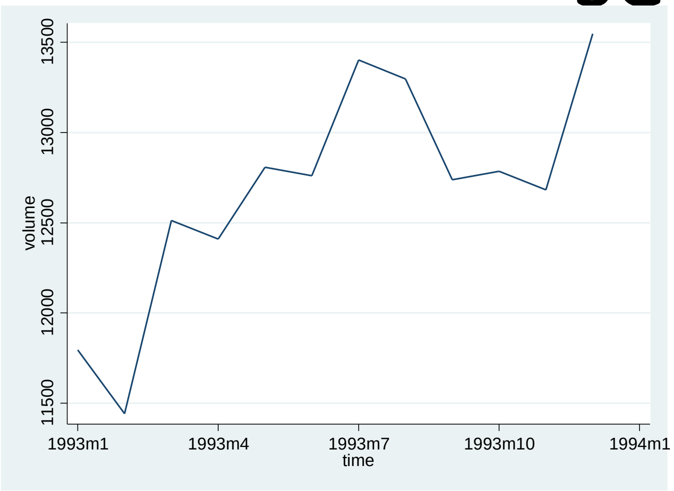


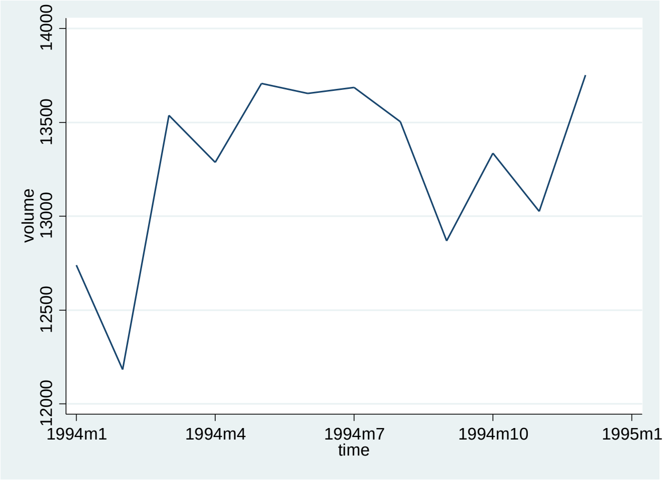
Example 2: Gasoline Sales

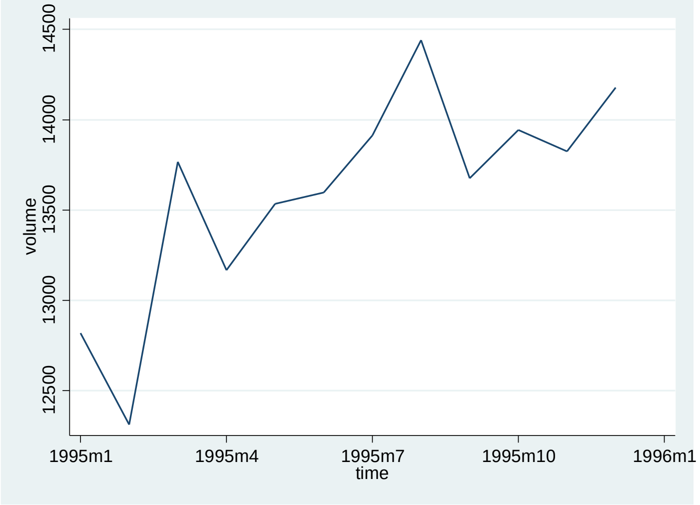
- gen m = month(dofm(time))
- regress volume b12.m

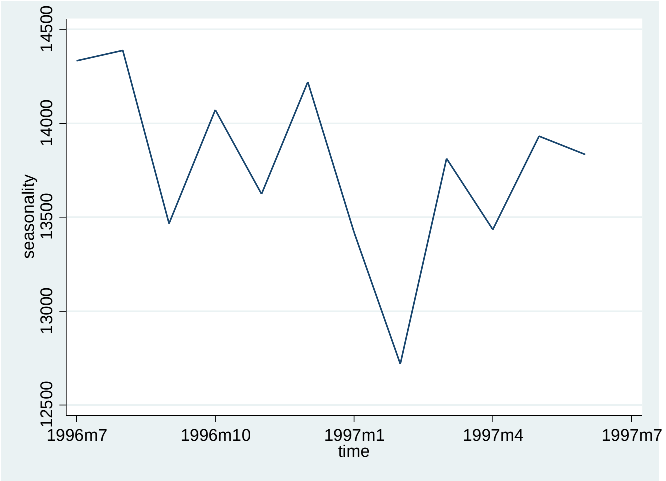
volume	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-800.4976	320.3884	-2.50	0.013	-1431.339	-169.6559
2	-1500.67	320.3884	-4.68	0.000	-2131.512	-869.8283
3	-407.3326	320.3884	-1.27	0.205	-1038.174	223.5091
4	-784.7925	320.3884	-2.45	0.015	-1415.634	-153.9508
5	-288.0735	320.3884	-0.90	0.369	-918.9152	342.7682
6	-386.9278	320.3884	-1.21	0.228	-1017.77	243.9139
7	113.6742	320.3884	0.35	0.723	-517.1675	744.5159
8	167.2559	320.3884	0.52	0.602	-463.5858	798.0976
9	-752.7424	320.3884	-2.35	0.020	-1383.584	-121.9006
10	-149.3386	320.3884	-0.47	0.642	-780.1804	481.5031
11	-597.5424	320.3884	-1.87	0.063	-1228.384	33.29936
_cons	14220.11	226.5488	62.77	0.000	13774.04	14666.18

Gasoline Sales, by year, and estimated seasonality

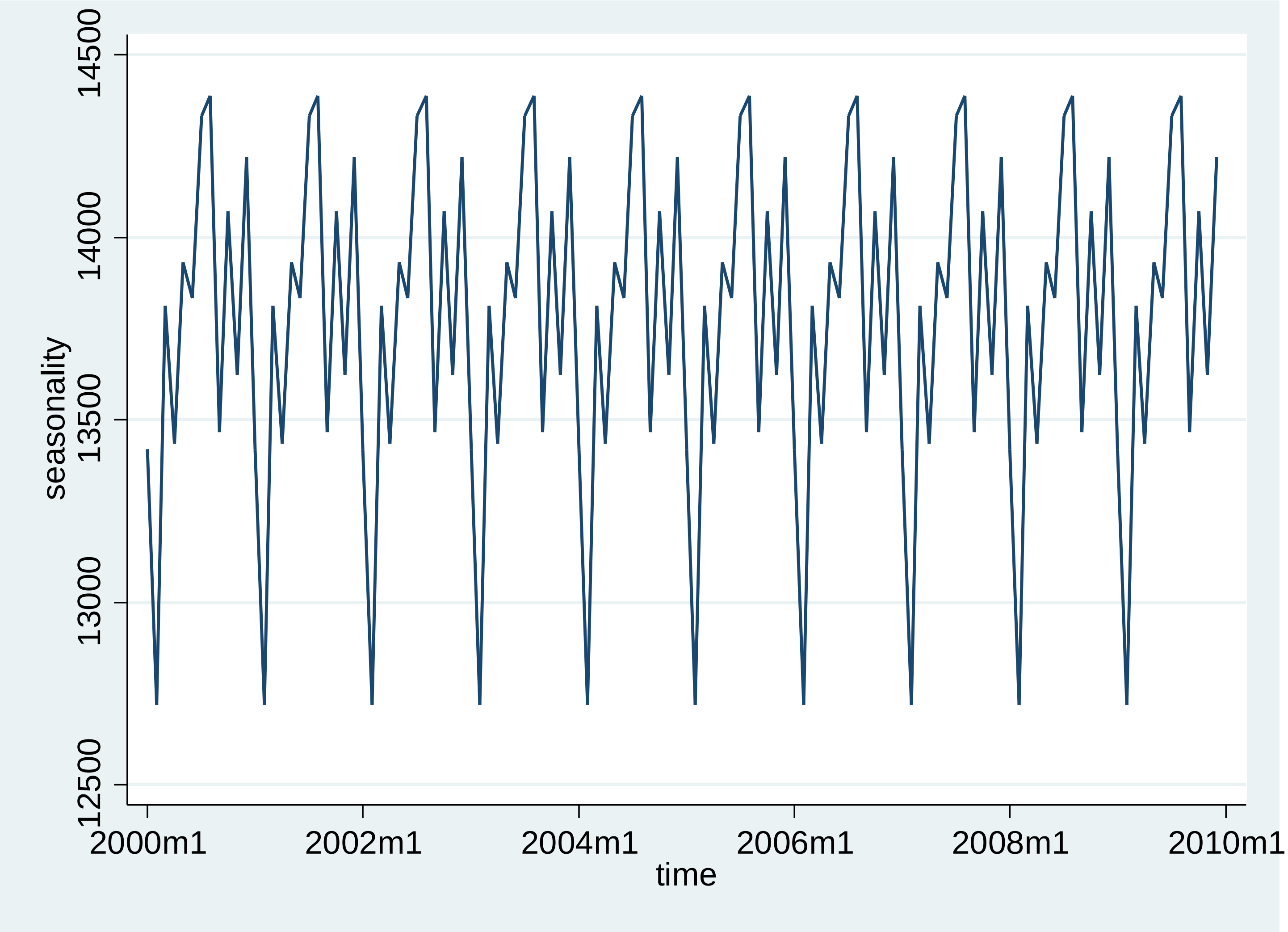




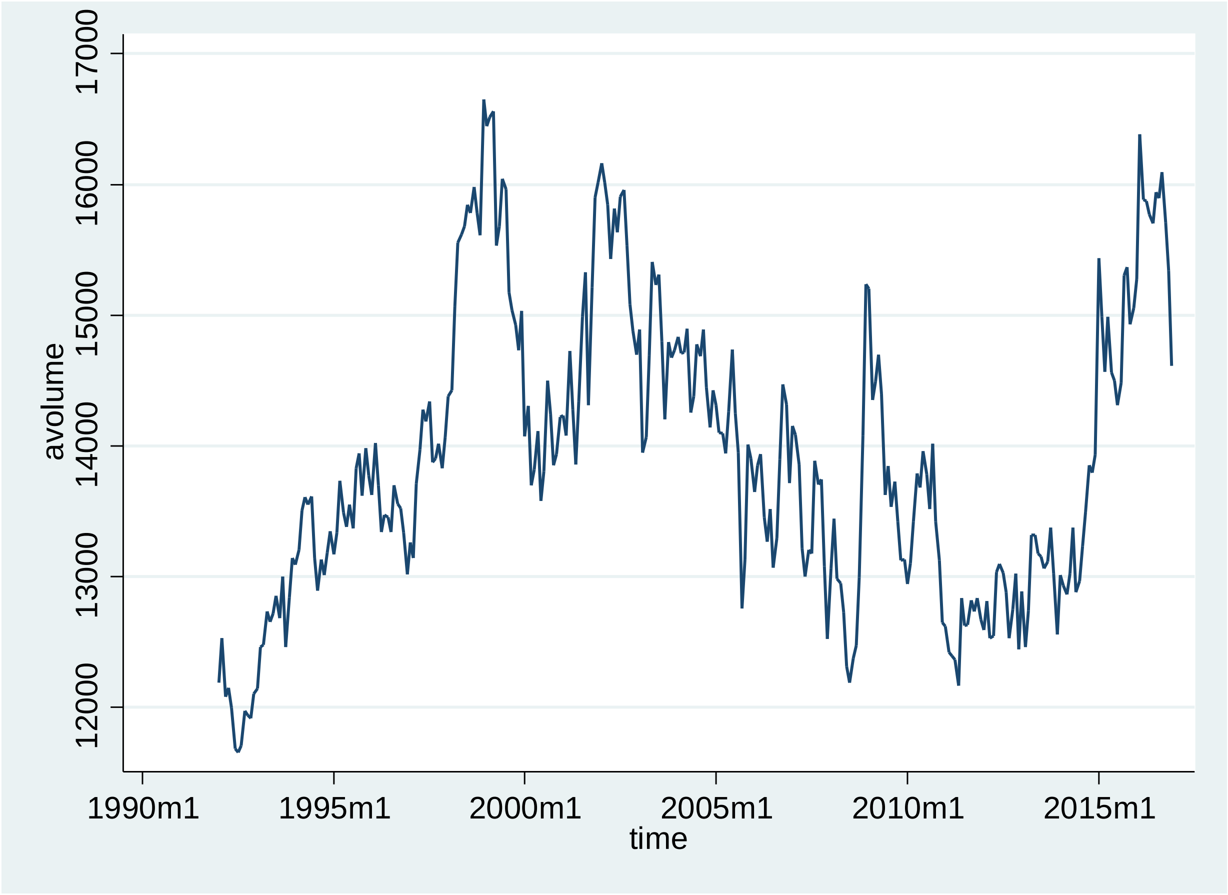




Predicted Values



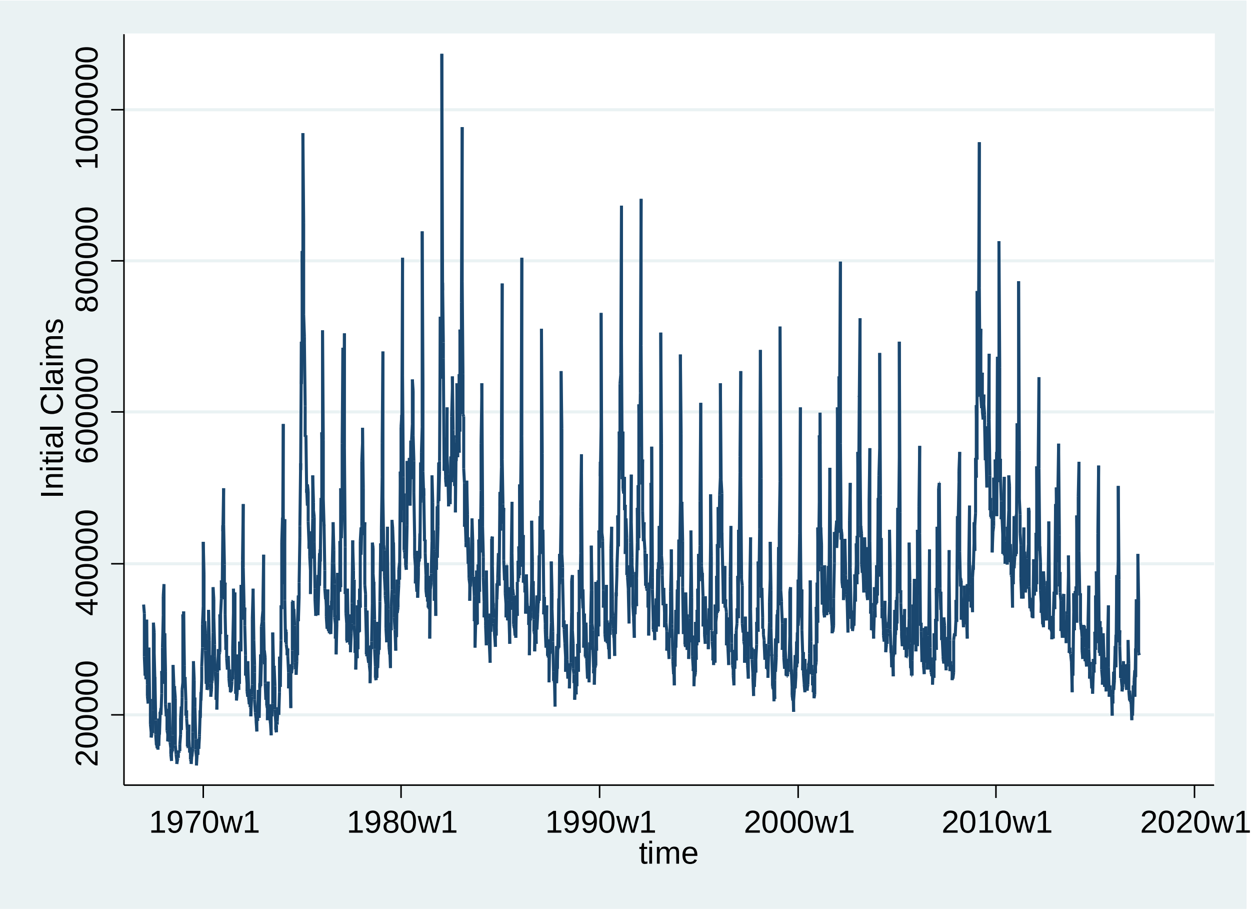
DeSeasonalized



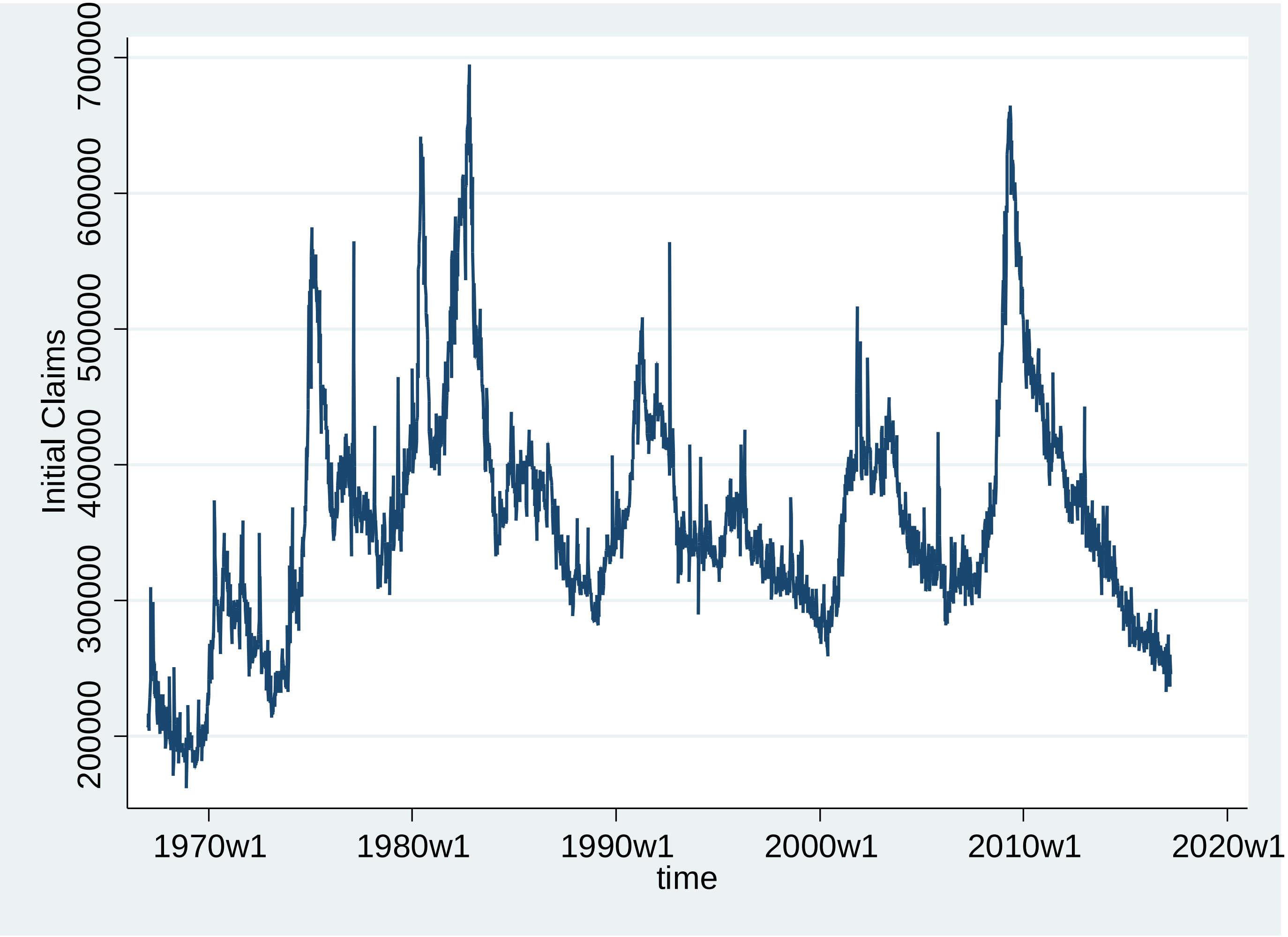
Example 4

- Unemployment Insurance Claims
- Department of Labor
- Weekly
- Important indicator for unemployment

Unemployment Claims
Not Seasonally Adjusted



Unemployment Claims
Official Seasonally Adjusted Series



Estimation

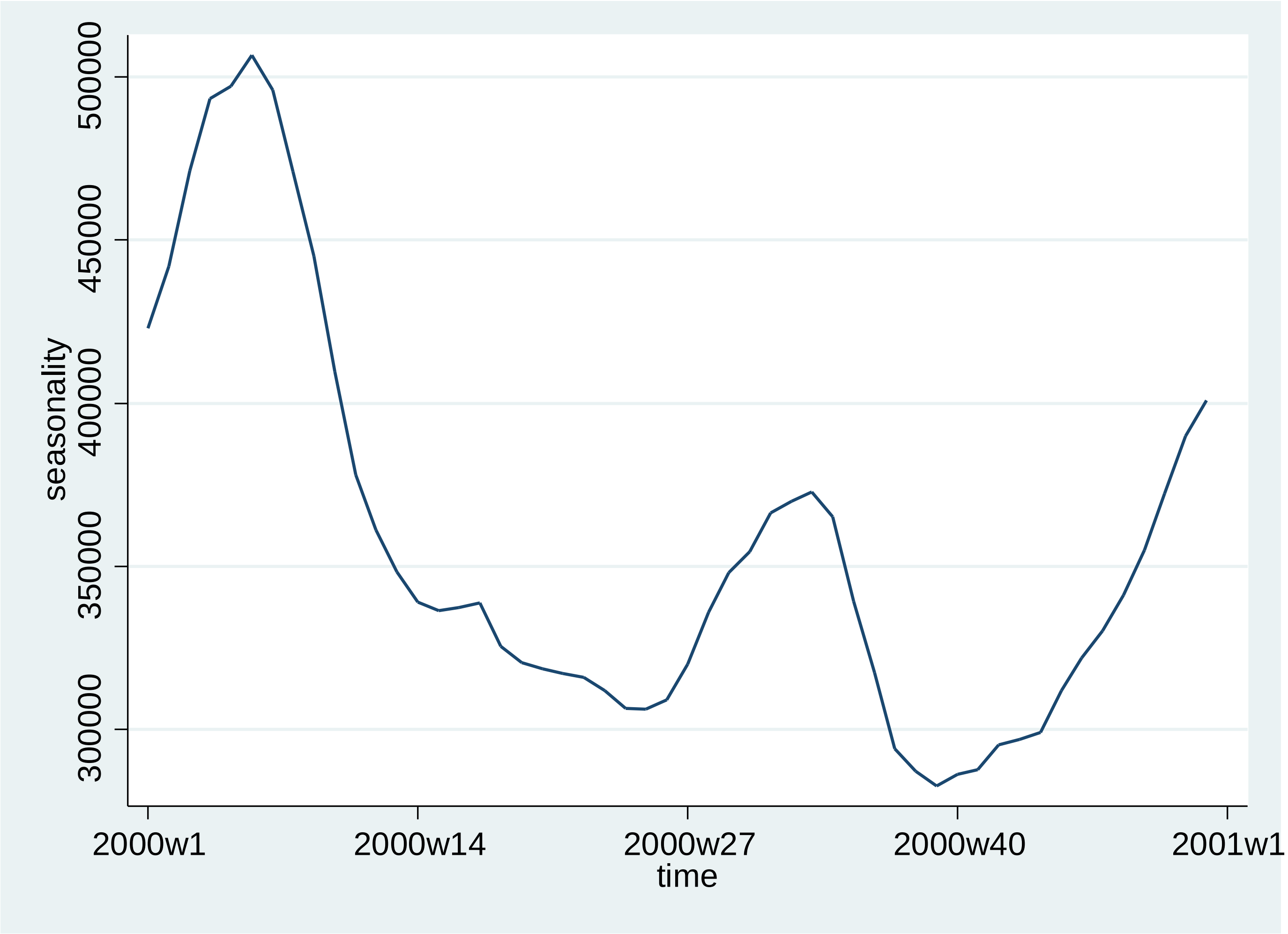
```
. regress iclaims_nsa b52.w
```

Source	SS	df	MS
Model	1.2804e+13	51	2.5105e+11
Residual	1.8865e+13	2186	8.6297e+09
Total	3.1668e+13	2237	1.4157e+10

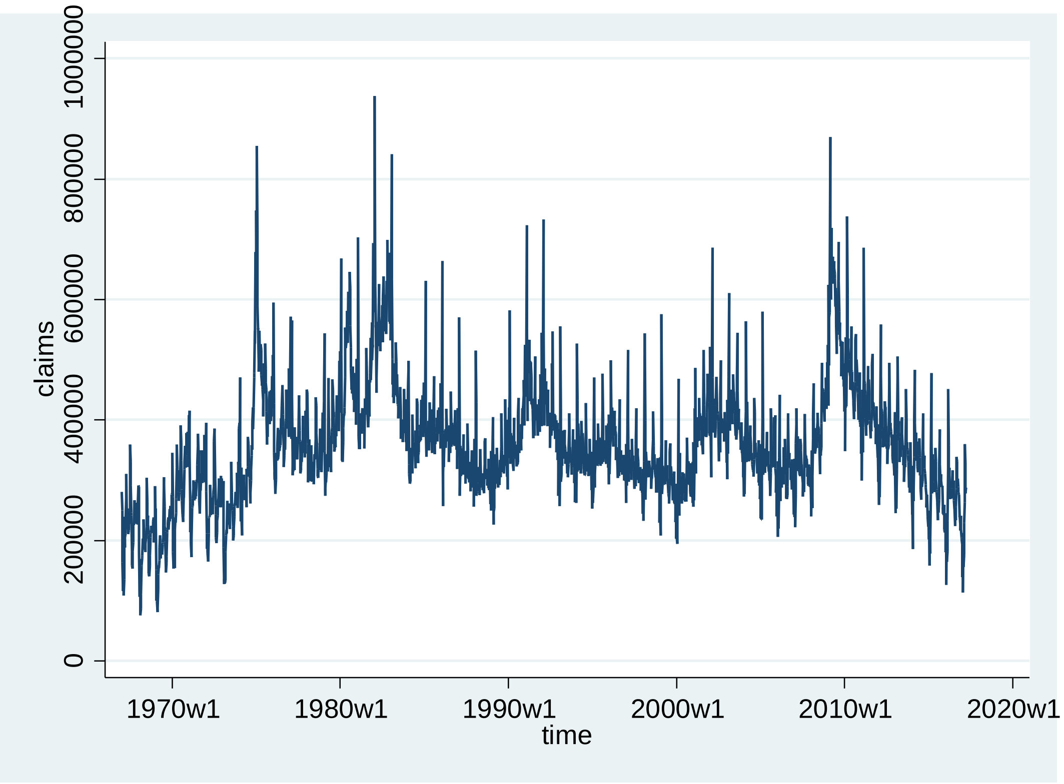
Number of obs	=	2238
F(51, 2186)	=	29.09
Prob > F	=	0.0000
R-squared	=	0.4043
Adj R-squared	=	0.3904
Root MSE	=	92896

iclaims_nsa	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
w						
1	35954.91	19920.38	1.80	0.071	-3109.952	75019.78
2	168185.3	19920.38	8.44	0.000	129120.4	207250.2
3	11166.58	20034.54	0.56	0.577	-28122.15	50455.32
4	-71213.05	20034.54	-3.55	0.000	-110501.8	-31924.31
5	-77421.58	20034.54	-3.86	0.000	-116710.3	-38132.85
6	-70397.88	20034.54	-3.51	0.000	-109686.6	-31109.15
7	-125853.5	20034.54	-6.28	0.000	-165142.3	-86564.8
8	-148580.9	20034.54	-7.42	0.000	-187869.6	-109292.2
9	-142809.6	20034.54	-7.13	0.000	-182098.4	-103520.9
10	-142668.5	20034.54	-7.12	0.000	-181957.2	-103379.8
11	-167656.8	20034.54	-8.37	0.000	-206945.6	-128368.1
12	-178125.4	20034.54	-8.89	0.000	-217414.1	-138836.7
13	-187898.8	20034.54	-9.38	0.000	-227187.6	-148610.1
14	-157631.2	20034.54	-7.87	0.000	-196919.9	-118342.5
15	-144329.8	20034.54	-7.20	0.000	-183618.5	-105041.1
16	-171520.5	20034.54	-8.56	0.000	-210809.2	-132231.8

Estimated Seasonal Process



DeSeasonalized



- Seasonality reduced, but not eliminated by dummy variables
- Indicates that seasonality has changed over time

Other types of seasonality

- Daily data
  - Day of the week
  - Handle by including dummy variables for each day
- High-frequency data
  - Include hourly or time-of-day indicators
- Holiday effects
  - Flower sales big on Valentines Day, Mothers Day, Easter, yet these days can move around
  - Trading-day/business-day variation
    - Number of trading days/business days varies across months
    - Can divide by number of trading days, or include as a regressor

Assignments

- Read Diebold through Chapter 6.4
- Problem Set # 4
  - Due Tuesday (2/14)
- Read Chapter 3 from *The Signal and the Noise*
  - Reading Reflection
  - Due Thursday (2/9)