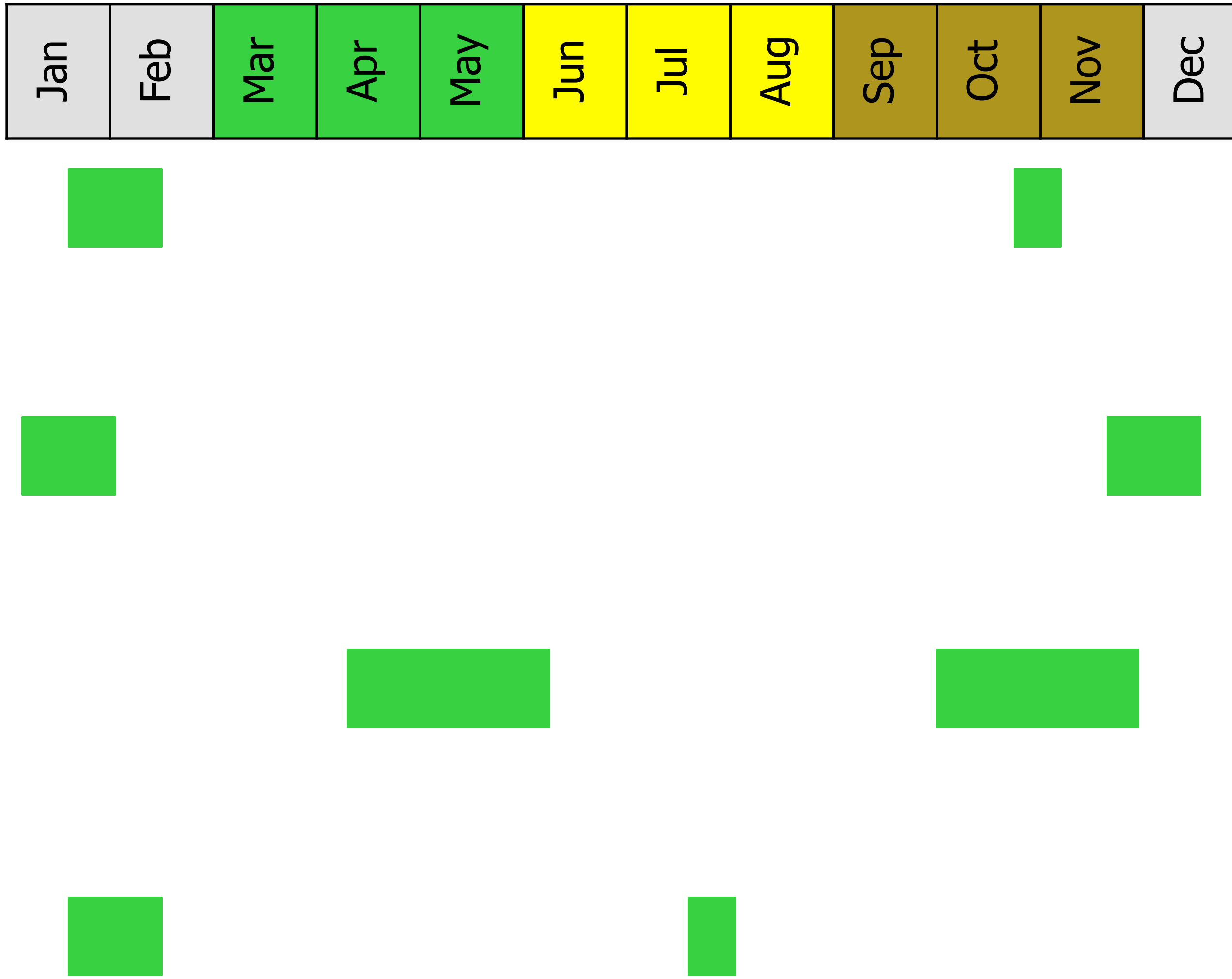
Exponential Smoothing: Seasonality

MIT Center for Transportation & Logistics

Annual Seasonality for Various Products



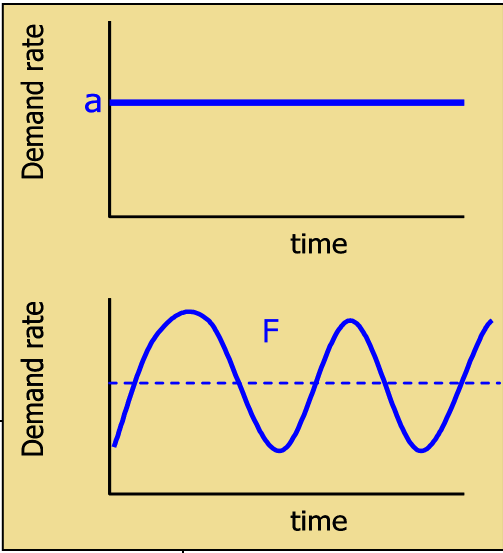
Agenda

- Double Exponential Smoothing Model
  - Level & Seasonality
- Holt-Winter Model
  - Level, Trend & Seasonality
- Initialization of Parameters
- Practical Concerns

Double Exponential Smoothing

Time Series Analysis

- Exponential Smoothing for Level & Seasonality
  - Double exponential smoothing
  - Introduces multiplicative seasonal term



Underlying Model: $x_t = aF_t + e_t$

where: $e_t \sim \text{iid}\ (\mu=0\ ,\ \sigma^2=V[e])$

Forecasting Model: $\hat{x}_{t,t+\tau} = \hat{a}_t \boxed{\hat{F}_{t+\tau-P}}$

The most current estimate of the appropriate seasonal index

Updating Procedure:

$$\hat{a}_t = \alpha \boxed{\left(\frac{x_t}{\hat{F}_{t-P}}\right)} + (1-\alpha)(\hat{a}_{t-1})$$

De-seasonalizing the most recent actual observation.

$$\hat{F}_t = \gamma \boxed{\left(\frac{x_t}{\hat{a}_t}\right)} + (1-\gamma)\hat{F}_{t-P}$$

This term "de-levels" the most recent actual observation.

F_t = Multiplicative seasonal index appropriate for period t

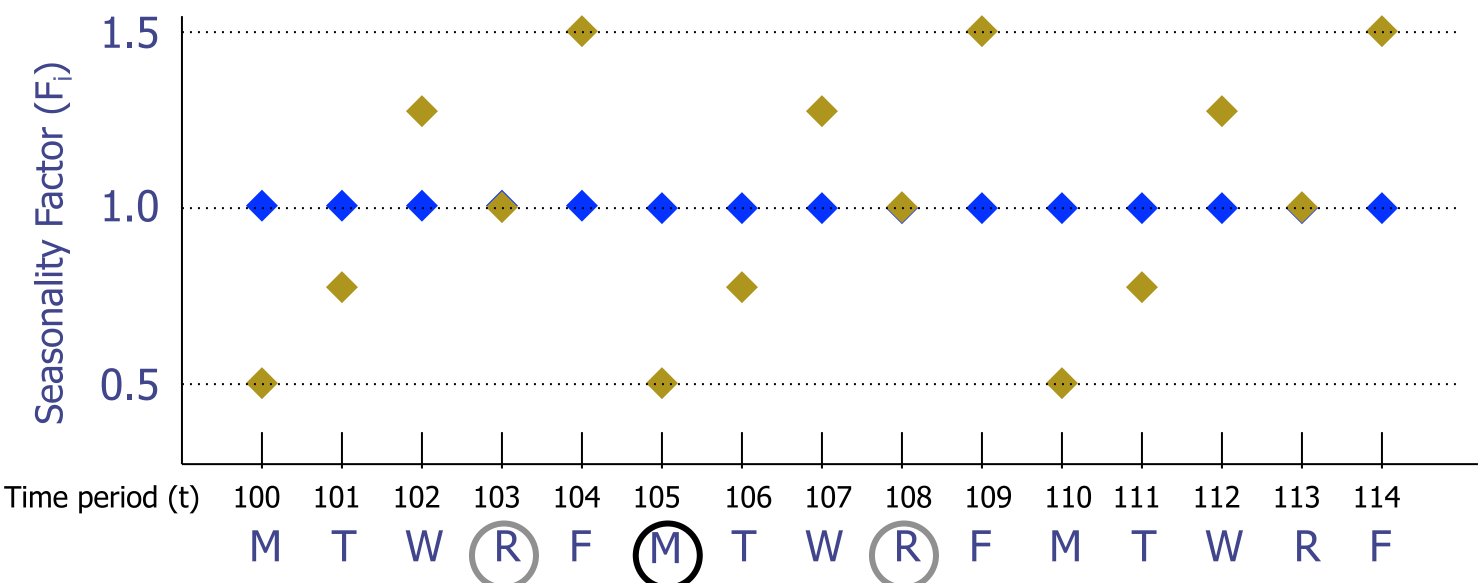
P = Number of time periods within the seasonality

γ = Seasonality smoothing factor

Note that: $\sum_{i=1}^{P} \hat{F}_i = P$

Seasonality

- For multiplicative seasonality, think of the F_i as "percent of average demand" for a period i
- The sum of the F_i for all periods within a season must equal P



	F_i
Monday	0.50
Tuesday	0.75
Wednesday	1.25
Thursday	1.00
Friday	1.50

τ

P

$$\hat{x}_{t,t+\tau} = \hat{a}_t \hat{F}_{t+\tau-P}$$

Suppose we are in Monday (t=105) forecasting for Thursday (t=108)? What is $F_{t+tau-P}$?

- This means t=105, P=5 and τ=3 and thus t+τ-P=105+3-5 = 103
- So, $x^{\wedge}_{105,108} = a^{\wedge}_{105} F^{\wedge}_{103}$
- We modify the most recent estimate of level by the most recent relevant seasonality index

Example: Forecasting Bagels

- We observe that demand is level with seasonality by day of week. It is now Friday (t=104) and we have estimated the level (a^$_{104}$) to be 121 bagels. The current daily seasonality factors are shown to the right and smoothing factors alpha=0.1 and gamma=0.05.

t	DOW	F_i
100	Mon	0.50
101	Tues	0.75
102	Wed	1.25
103	Thur	1.00
104	Fri	1.50

- What is your forecast for Monday (x^$_{104,105}$)?

Recall that: $\hat{x}_{t,t+\tau} = \hat{a}_t \hat{F}_{t+\tau-P}$ so x^$_{104,105}$ = (121)(0.50) = 60.5

- Suppose our actual bagel sales on Monday (t=105) was 76. What is our forecasted demand for Tuesday (t=106)?

	A	B	C	D	E	F
1	Alpha =		0.10			
2	Gamma =		0.05			
3						
4	t	DOW	x(t)	a^(t)	F^(t)	x^(t,t+1)
5	100	M			0.50	
6	101	T			0.75	
7	102	W			1.25	
8	103	R			1.00	
9	104	F		121.0	1.50	60.5
10	105	M	76	124.1	0.51	93.1

Current Seasonality Factors

$$\hat{a}_t = \alpha\left(\frac{x_t}{\hat{F}_{t-P}}\right) + (1-\alpha)(\hat{a}_{t-1})$$

$$\hat{F}_t = \gamma\left(\frac{x_t}{\hat{a}_t}\right) + (1-\gamma)\hat{F}_{t-P}$$

$$\hat{x}_{t,t+\tau} = \hat{a}_t \hat{F}_{t+\tau-P}$$

=D10*E6

=$C$1*(C10/E5)+(1-$C$1)*D9

=$C$2*(C10/D10)+(1-$C$2)*E5

Normalizing Seasonality Indices

Example: Forecasting Bagels

- The seasonality factors no longer sum to P!!
- So what? What happens if I ignore this?
- Need to update or normalize them.
- What is the new seasonality factor F^$_{106}$?

$$\hat{F}_t^{NEW} = \hat{F}_t^{OLD}\left(\frac{P}{\sum_{i=t-P}^{t}\hat{F}_t^{OLD}}\right)$$

t	DOW	F_i
100	Mon	0.500
101	Tues	0.750
102	Wed	1.250
103	Thur	1.000
104	Fri	1.500
105	Mon	0.506
106	Tues	

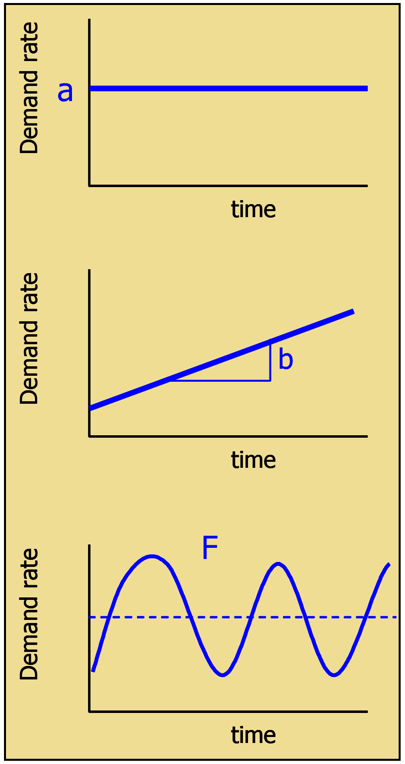
	Estimates	Normalized
Tues	0.750	0.749
Wed	1.250	1.249
Thur	1.000	0.999
Fri	1.500	1.498
Mon	0.506	0.505
Sum	5.006	5.000

Seasonality factors must be kept current or they will drift dramatically. This requires a lot more bookkeeping which is tricky to maintain in a spreadsheet.

Holt-Winter Model

Time Series: Level, Trend, & Seasonal Data

- Expanding exponential Holt's model to include trend
- Commonly called the Holt-Winter Method
- Multiplicative seasonality and additive trend



Underlying Model:

$$x_t = (a+tb)\,F_t + e_t$$

where: $e_t \sim \text{iid}\ (\mu=0\ ,\ \sigma^2=V[e])$

Forecasting Model: $\hat{x}_{t,t+\tau} = \left(\hat{a}_t + \tau\hat{b}_t\right)\hat{F}_{t+\tau-P}$

Where :

$$\hat{a}_t = \alpha\left(\frac{x_t}{\hat{F}_{t-P}}\right) + (1-\alpha)\left(\hat{a}_{t-1} + \hat{b}_{t-1}\right)$$

$$\hat{b}_t = \beta\left(\hat{a}_t - \hat{a}_{t-1}\right) + (1-\beta)\hat{b}_{t-1}$$

$$\hat{F}_t = \gamma\left(\frac{x_t}{\hat{a}_t}\right) + (1-\gamma)\hat{F}_{t-P}$$

De-seasonalizing the most recent observation

Old estimate of level and trend

No change from Holt Model!

No change from level & seasonality model!

Current observation is "de-leveled"

Seasonal term from most recent similar time period.

Example II: Forecasting Bagels

- We now observe that demand is level with additive trend and multiplicative seasonality by day of week. It is now Friday (t=104) and we have estimated the level ($a^{\wedge}_{104}$) to be 121 bagels with a trend ($b^{\wedge}_{104}$)=0.3 bagles per day. The current daily seasonality factors are shown to the right and smoothing factors are alpha=0.1, beta=0.08, and gamma=0.05.
- What is your forecast for Monday ($x^{\wedge}_{104,105}$)?

Recall that: $\hat{x}_{t,t+\tau} = \left(\hat{a}_t + \tau \hat{b}_t\right) \hat{F}_{t+\tau-P}$ so $x^{\wedge}_{104,105}$ = (121+0.3)(0.50) = 60.7

- Suppose our actual bagel sales on Monday (t=105) was 76. What is our forecasted demand for Tuesday (t=106)?

t	DOW	F_i
100	Mon	0.50
101	Tues	0.75
102	Wed	1.25
103	Thur	1.00
104	Fri	1.50

	A	B	C	D	E	F	G
1	Alpha =		0.10				
2	Beta =		0.08				
3	Gamma =		0.05				
4							
5	t	DOW	x(t)	a^(t)	b^(t)	F^(t)	x^(t,t+1)
6	100	M				0.500	
7	101	T				0.750	
8	102	W				1.250	
9	103	R				1.000	
10	104	F		121.0	0.3	1.500	60.7
11	105	M	76	124.4	0.5	0.506	93.7

=$C$1*(C11/F6)+(1-$C$1)*(D10+E10)

=$C$2*(D11-D10)+(1-$C$2)*E10

=$C$3*(C11/D11)+(1-$C$3)*F6

=(D11+E11)*F7

$$\hat{a}_t = \alpha\left(\frac{x_t}{\hat{F}_{t-P}}\right) + (1-\alpha)\left(\hat{a}_{t-1} + \hat{b}_{t-1}\right)$$

$$\hat{b}_t = \beta\left(\hat{a}_t - \hat{a}_{t-1}\right) + (1-\beta)\hat{b}_{t-1}$$

$$\hat{F}_t = \gamma\left(\frac{x_t}{\hat{a}_t}\right) + (1-\gamma)\hat{F}_{t-P}$$

$$\hat{x}_{t,t+\tau} = \left(\hat{a}_t + \tau\hat{b}_t\right)\hat{F}_{t+\tau-P}$$

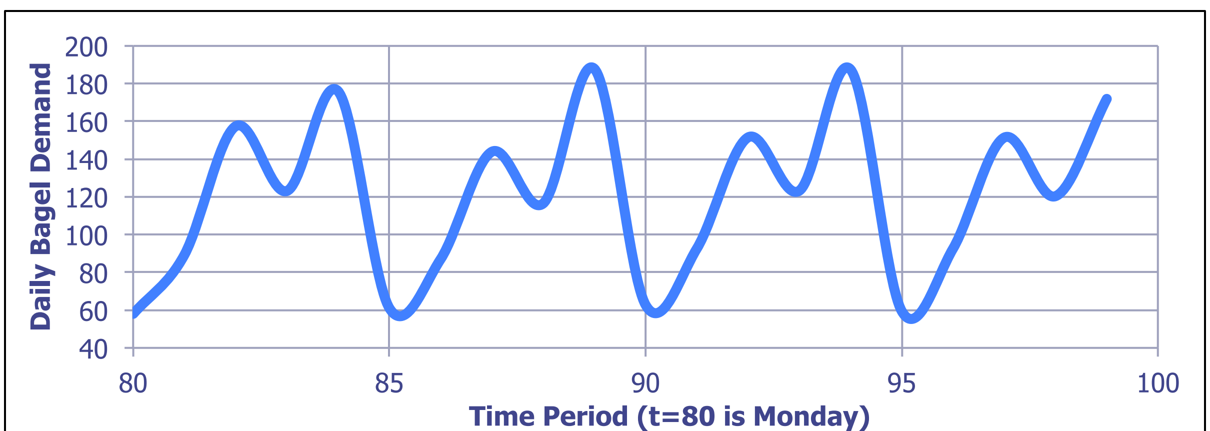
Initialization of Models

Initialization of Parameters

- Points to Note
  - No single best method – many good ones
  - Need to partition the data (initialization, training, and testing)
- Simple Exponential Smoothing Model
  - Estimate initial level parameter ($\hat{a}_0$)
  - Use average demand for first several periods
- Holt Model
  - Estimate initial level ($\hat{a}_0$) and trend ($\hat{b}_0$) parameters (a & b)
  - Find best fit linear equation to data in initial data set
  - Use ordinary least squares regression of demand for several periods
    - Dependent variable = demand in each time period = x_t
    - Independent variable = slope = β_1
    - Regression equation: $x_t = \beta_0 + \beta_1 t$
  - Note: this gives equal weight to each observation in the initialization data sample

Initialization of Parameters

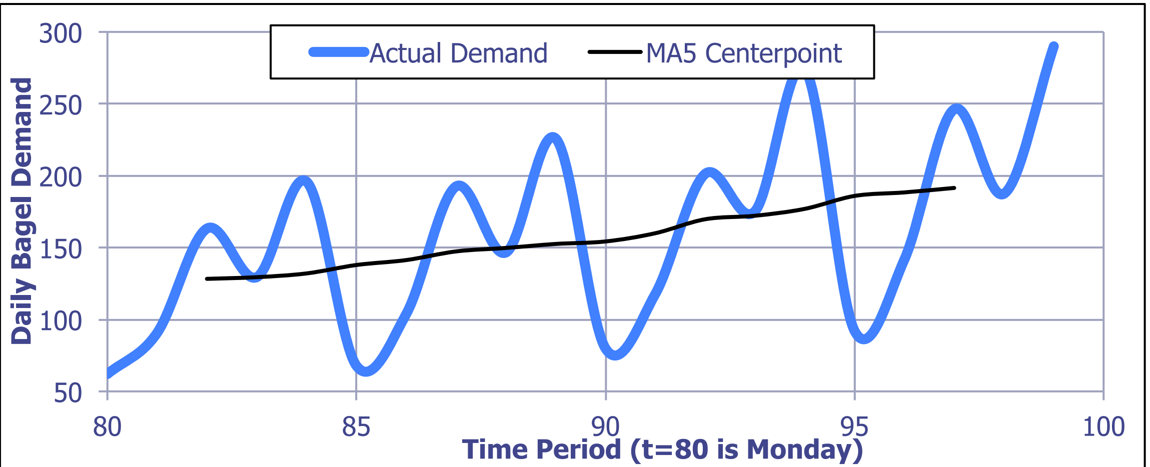
- Seasonality Models
  - These are much more complicated
  - Several different methods used in practice
  - You need lots of data – >2 seasons worth but prefer ≥4 seasons
- Double Exponential Smoothing Model
  - Estimate initial level parameter ($\hat{a}_0$) & seasonality indices ($\hat{F}_i$)
  - Find average demand for each common season period
  - Find average demand for all periods
  - Set initial seasonality indices to ratio of each season to all periods



Day	Average Daily Demand	Initial Seasonality Index
Mon	60	0.50
Tues	90	0.75
Wed	151	1.25
Thur	121	1.00
Fri	181	1.50
All Days	121	

Initialization – Holt-Winter

1. Estimate initial level for each season
   - Find P-Moving Avg centered in each season
   - Take ratio of actual to initial F_i estimate
2. Find initial season indices
   - Average F_i for common periods

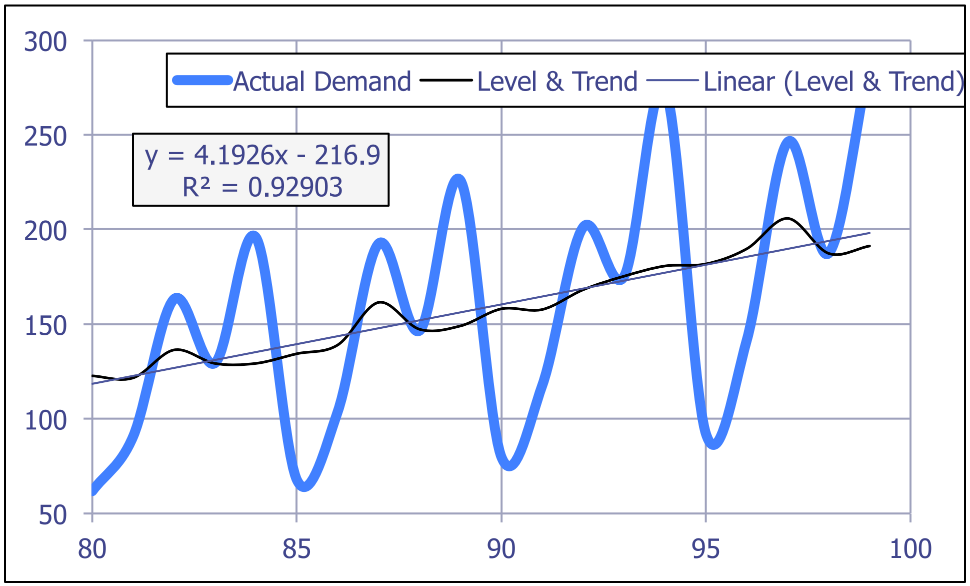


DOW	Average of F_i Estimates	Normalized F_is
M	0.501	0.506
T	0.741	0.748
W	1.195	1.195
R	0.998	0.998
F	1.516	1.516
Sum	4.952	5.000

Week	t	DOW	Actual Demand	MA5 Center point	Initial F_i Estimate
1	80	M	62		
	81	T	91		
	82	W	163	128.3	1.273
	83	R	129	129.4	1.000
	84	F	196	132.0	1.483
2	85	M	68	137.9	0.492
	86	T	104	141.3	0.735
	87	W	193	147.3	1.307
	88	R	147	149.8	0.979
	89	F	226	152.5	1.481
3	90	M	80	154.3	0.518
	91	T	118	160.0	0.736
	92	W	201	169.7	1.186
	93	R	175	172.2	1.018
	94	F	274	176.9	1.551
4	95	M	92	185.9	0.495
	96	T	142	188.4	0.752
	97	W	246	191.5	1.286
	98	R	187		
	99	F	290		

Initialization – Holt-Winter

3. Estimate initial level and trend values
   - De-season each observation by dividing by its estimated Seasonality Index
   - Use these values to estimate $\hat{a}_0$ and $\hat{b}_0$ using OLS regression



Wk	t	DOW	Actual Demand	Normalized F_is	Deseasoned Demand
1	80	M	62	0.506	122.53
	81	T	91	0.748	121.66
	82	W	163	1.195	136.40
	83	R	129	0.998	129.26
	84	F	196	1.516	129.29
2	85	M	68	0.506	134.39
	86	T	104	0.748	139.04
	87	W	193	1.195	161.51
	88	R	147	0.998	147.29
	89	F	226	1.516	149.08
3	90	M	80	0.506	158.10
	91	T	118	0.748	157.75
	92	W	201	1.195	168.20
	93	R	175	0.998	175.35
	94	F	274	1.516	180.74
4	95	M	92	0.506	181.82
	96	T	142	0.748	189.84
	97	W	246	1.195	205.86
	98	R	187	0.998	187.37
	99	F	290	1.516	191.29

So, for t=100 (start of our modeling)

$\hat{a}_0 = (4.1926)(100) - 216.9 = 202.36 \approx \underline{202.4}$

$\hat{b}_0 = 4.1926 \approx \underline{4.2}$

$\hat{F}_i = \{\underline{0.506}, \underline{0.748}, \underline{1.195}, \underline{0.998}, \underline{1.516}\}$ for {M,T,W,R,F}

Final Comments on Exponential Smoothing

Comments on Time Series Models

- Three phases of work on three different data sets
  - Initialize: Estimate level, trend, and/or seasonality factors
  - Train: Determine the smoothing parameters to use
  - Test: Evaluate the quality (accuracy & bias) of forecasts
- Selecting data for use in model formulation
  - Needs to be appropriately long but still relevant
  - Needs to be cleaned of all non-repeating events
- Picking appropriate smoothing factors
  - Level (α)
    - Stationary: ranges from 0.01 to 0.30 (0.1 reasonable)
    - Trend/Season: ranges from 0.02 to 0.51 (0.19 reasonable)
  - Trend (β)
    - Ranges from 0.005 to 0.176 (0.053 reasonable)
  - Seasonality (γ)
    - Ranges from 0.05 to 0.50 (0.10 reasonable)

Comments on Time Series Models

- Most of the work is bookkeeping
  - Initialization procedures are somewhat arbitrary
  - Adding seasonality greatly complicates calculations

- Measuring Bias in Forecasts
  - Track the cumulative sum of Forecast Errors
  - Normalize by mean square of errors
  - Should fluctuate around 0

$$C_t = C_{t-1} + e_t$$

$$C_t^N = C_t \Big/ \sqrt{MSE_t}$$

Comments on Time Series Models

- Variety of More Sophisticated Models
  - Seasonality: None, Additive, or Multiplicative
  - Trend: None, Additive, or Multiplicative
  - Trend Damping: None or Present
  - Box-Jenkins or Autoregressive Integrated Moving Average (ARIMA)

		Seasonality Component		
		None	Additive	Multiplicative
Trend Component	None	Simple		Double
	Additive	Holt		Holt-Winter
	Additive Damped	Holt Damped		
	Multiplicative			
	Multiplicative Damped			

Questions, Comments, Suggestions?
Use the Discussion!

"Ginger Belle"
Photo courtesy Yankee Golden Retriever Rescue (www.ygrr.org)

MIT Center for Transportation & Logistics

caplice@mit.edu

Image Credits

- slide 2 –
  - "Valentines Day Chocolates from 2005" by John Hritz from Ann Arbor, MI, USA - Flickr. Licensed under Creative Commons Attribution 2.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Valentines_Day_Chocolates_from_2005.jpg#mediaviewer/File:Valentines_Day_Chocolates_from_2005.jpg
  - "Johns Inc Salad Bar Buffet" by Tudokin - Own work. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Johns_Inc_Salad_Bar_Buffet.jpg#mediaviewer/File:Johns_Inc_Salad_Bar_Buffet.jpg
  - By Das Ant (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0) or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons. http://commons.wikimedia.org/wiki/File%3AGartengeraete.jpg
  - "Bandaż" by Original uploader was Reytan at pl.wikipedia - Transferred from pl.wikipedia; transferred to Commons by User:Pjahr using CommonsHelper.. Licensed under Public domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Banda%C5%BC.jpg#mediaviewer/File:Banda%C5%BC.jpg