Collection

Users can Collect posts into a printable, sortable format. Collections are a good way to organize posts for quick reading. A Collection must be created to tag posts. More Help

Thread: "Untidy" data example
Post: "Untidy" data example
Author: James Topor
Posted Date: February 28, 2016 4:44 PM
Status: Published
Overall Rating:

1. Data - a wide table. Each county has a fixed Land Area and Natural Amenity status (4 = NO; 3 = YES). Measurements were taken every 10 years on the proportion of county residents that had completed college and the number of total number of jobs in the county.

Source: http://www.theanalysisfactor.com/wide-and-long-data/

	County	LandArea	NatAmenity	College1970	College1980	College1990	College2000	Jobs1970	Jobs1980	Jobs1990	Jobs2000
1	Autauga	599	4	.064	.121	.145	.180	6853	11278	11471	16289
2	Baldwin	1578	4	.065	.121	.168	.231	19749	27861	40809	70247
3	Barbour	891	4	.073	.092	.118	.109	9448	9755	12163	15197
4	Bibb	625	3	.042	.049	.047	.071	3965	4276	5564	6098
5	Blount	639	4	.027	.053	.070	.096	7587	9490	11811	16503

2. Type of analysis one might do: determine whether number of college graduates in the county (and possibly either of the other the two fixed variables) is a predictor of the number of jobs in the county.

Tags: None

(Post is Read)

Thread: Untidy Auction Data
Post: Untidy Auction Data
Author: Christopher Martin
Posted Date: February 29, 2016 12:41 PM
Edited Date: March 3, 2016 8:08 PM
Status: Published
Overall Rating:

EDITED 3/3: Added column names and changed analysis request.

Untidy Data: Auction listings

	Lot#	Description		
	2001	New England Federal mahogany slant front desk, ca. 1810, with allover line inlay, 45" h., 40" w.	$300 - $500	$283
	2002	Federal brass and wire fire fender, early 19th c., 13 3/4" h., 30" w.	$80 - $120	$234
	2003	Pennsylvania or New Jersey late Chippendale applewood card table, ca. 1790, 29 1/2" h., 36" w.	$200 - $300	$357

Columns: Picture | Lot # | Description | Estimate | Price Sold

Analysis: Find the percent difference from the estimates (high and low) and the sold prices. (Sold Price - High Estimate)/High Estimate = %

Tags: None

(Post is Unread)

Thread: Oil Consumption data
Post: Oil Consumption data
Author: Kishore Prasad
Posted Date: March 1, 2016 9:37 AM
Status: Published
Overall Rating:

Data: Below is the oil consumption details of various brands in the first half of 2015 along with the opening balances.

* Numbers indicate "Purchased : Consumed" in Gallons

Month	Category	Caltex	Gulf	Mobil
Open	Engine Oil	140 : 000	199 : 000	141 : 000
	GearBox Oil	198 : 000	132 : 000	121 : 000
Jan	Engine Oil	170 : 103	194 : 132	109 : 127
	GearBox Oil	132 : 106	125 : 105	191 : 100
Feb	Engine Oil	112 : 133	138 : 113	171 : 101
	GearBox Oil	193 : 148	199 : 119	134 : 127

Mar	Engine Oil	184 : 100	141 : 141	114 : 108
	GearBox Oil	138 : 121	172 : 133	193 : 115
Apr	Engine Oil	149 : 150	117 : 118	117 : 118
	GearBox Oil	185 : 125	191 : 133	119 : 121
May	Engine Oil	170 : 139	104 : 119	200 : 117
	GearBox Oil	168 : 117	138 : 102	121 : 146
Jun	Engine Oil	159 : 129	170 : 138	169 : 105
	GearBox Oil	107 : 129	195 : 141	141 : 112

Analysis:

a) Give the closing balance of component + Brand

b) Give most consumed brand across the 2 categories of oil.

Note: I have also attached the excel file

Attachment: Project 2.xlsx (9.747 KB)

Tags: None

(Post is Unread)

Thread: Oil Consumption data
Post: RE: Oil Consumption data
Author: Chirag Vithalani
Posted Date: March 4, 2016 1:17 AM
Status: Published
Overall Rating:

This data & analysis more useful in current situation when oil prices are low and oil companies are running in losses.

BTW, from where you got this data ?

Tags: None

(Post is Unread)

Thread: Untidy MLB World Series Data
Post: Untidy MLB World Series Data
Author: Logan Thomson
Posted Date: March 1, 2016 4:15 PM
Status: Published
Overall Rating:

Based on the emphasis in statistics in Major League Baseball, I was a little surprised to find such an untidy table of World Series results directly from the MLB:

World Series History: Recaps and Results

Year	Results	MVP	Year	Results	MVP
2015	Royals 4, Mets 1	Salvador Perez	1999	NY Yankees 4, Atlanta 0	Mariano Rivera
2014	Giants 4, Royals 3	Madison Bumgarner	1998	NY Yankees 4, San Diego 0	Scott Brosius
2013	Red Sox 4, Cardinals 2	David Ortiz	1997	Florida 4, Cleveland 3	Livan Hernandez
2012	Giants 4, Tigers 0	Pablo Sandoval	1996	NY Yankees 4, Atlanta 2	John Wetteland
2011	Cardinals 4, Rangers 3	David Freese	1995	Atlanta 4, Cleveland 2	Tom Glavine
2010	Giants 4, Rangers 1	Edgar Renteria	1994	Not Held	N/A
2009	Yankees 4, Phillies 2	Hideki Matsui	1993	Toronto 4, Philadelphia 2	Paul Molitor
2008	Philadelphia 4, Tampa Bay 1	Cole Hamels	1992	Toronto 4, Atlanta 2	Pat Borders
2007	Boston 4, Colorado 0	Mike Lowell	1991	Minnesota 4, Atlanta 3	Jack Morris
2006	St. Louis 4, Detroit 1	David Eckstein	1990	Cincinnati 4, Oakland 0	Jose Rijo
2005	Chi. White Sox 4, Houston 0	Jermaine Dye			
2004	Boston 4, St. Louis 0	Manny Ramirez			
2003	Florida 4, NY Yankees 2	Josh Beckett			
2002	Anaheim 4, San Francisco 3	Troy Glaus			
2001	Arizona 4, NY Yankees 3	Schilling/Johnson			
2000	NY Yankees 4, NY Mets 1	Derek Jeter			
Year	Results	MVP	Year	Results	MVP
1989	Oakland 4, San Francisco 0	Dave Stewart	1979	Pittsburgh 4, Baltimore 3	Willie Stargell

For those not as familiar with the World Series, it is the championship series for Major League Baseball in the US (and one city in Canada). The numbers after the team names (or city in some cases) is the number of games won in the series, not the score. Also, MLB has divided the data up by decade, which is why there are separate headers for year, results, mvp - but apparently forgot that a new decade started in 2010. The winning team is obviously listed on the left, but it is odd that winning and losing teams are not separated.

Though not any kind of deep statistical analysis, one could tidy the data and then find out the teams with the most World Series appearances, those with the least, who has won the most Series games, won the least, what team has the highest percentage of series wins per World Series appearances, ratio of 4-5-6 and 7- game series (a few were more), etc.

source: http://mlb.mlb.com/mlb/history/postseason/mlb_ws.jsp?feature=recaps_index

Tags: None

(Post is Unread)

Thread:
Shapefile .dbf data National roadway files
Post:
Shapefile .dbf data National roadway files
Author: Robert Sellers

Posted Date: March 1, 2016 10:31 PM
Edited Date: March 3, 2016 9:55 PM
Status: Published
Overall Rating:

Original data from ftp://ftp2.census.gov/geo/tiger/TIGER2015/PRIMARYROADS/

&

http://www.scribekey.com/AttributeDomains/FEATNAMES.MTFCC.html

Link to modified file
: https://raw.githubusercontent.com/RobertSellers/R/master/data/tiger_file.tx

This shapefile is of road systems in the United States with miles calculated using ArcGIS "calculate geometry" and exported as a .txt comma-delimited file. Shapefiles contain various metadata files, vector information files, and an attribute file (.dbf). I have converted this to .txt for convenience. For your inconvenience I pasted the MTFCC lookup values at the top the txt file.

tl_2015_us_primaryroads_Proj1

Shape *	LINEARID	FULLNAME	RTTYP	MTFCC	Shape_Length	miles_length
Polyline	11022127866	I-91	M	S1100	2478.532707	1.540086
Polyline	11026514219	I 79	M	S1100	497.502507	0.309133
Polyline	11026514557	I 79	M	S1100	1494.49451	0.928634
Polyline	11018422617	Ih 45-Hov Fwy	M	S1100	28053.895796	17.431848
Polyline	11041500085	Indian Hill Rd	M	S1100	12.516267	0.007777
Polyline	11043283807	Indiana East-West Toll Rd	M	S1100	34086.238777	21.180164
Polyline	11022240099	Indiana East-West Toll Rd	M	S1100	14919.847166	9.270745
Polyline	11022240099	Indiana East-West Toll Rd	M	S1100	10278.487988	6.386744
Polyline	11042581255	Indiana East-West Toll Rd	M	S1100	1152.742995	0.71628
Polyline	11042582229	Indiana East-West Toll Rd	M	S1100	5853.732314	3.637333
Polyline	11042588024	Indiana East-West Toll Rd	M	S1100	23343.088258	14.504694
Polyline	11044484740	Indiana East-West Toll Rd	M	S1100	34087.240387	21.180787
Polyline	11044711419	Indiana East-West Toll Rd	M	S1100	40360.605431	25.078867
Polyline	11044711419	Indiana East-West Toll Rd	M	S1100	15034.344818	9.34189
Polyline	11044711429	Indiana East-West Toll Rd	M	S1100	10781.914458	6.699558
Polyline	11044712240	Indiana East-West Toll Rd	M	S1100	40360.660353	25.078901
Polyline	11044712240	Indiana East-West Toll Rd	M	S1100	39512.459424	24.551855
Polyline	11044714743	Indiana East-West Toll Rd	M	S1100	37936.124241	23.572368
Polyline	11044714743	Indiana East-West Toll Rd	M	S1100	37934.134836	23.571131
Polyline	11042581262	Indiana Eats West Toll Rd	M	S1100	751.071878	0.466693
Polyline	11045231793	Indiana Toll Rd	M	S1100	39668.211756	24.648635
Polyline	11045231793	Indiana Toll Rd	M	S1100	39670.613467	24.650127
Polyline	11037720093	Indiana Toll Rd	M	S1100	4229.687771	2.628201
Polyline	11037720093	Indiana Toll Rd	M	S1100	2802.600813	1.741452
Polyline	11022192107	Indiana Tollroad	M	S1100	18478.16541	11.481777
Polyline	11022192107	Indiana Tollroad	M	S1100	12149.948722	7.549613
Polyline	11022192107	Indiana Tollroad	M	S1100	3449.751261	2.143572
Polyline	11090175362	Inland Empire Hwy	M	S1100	65.168186	0.040494
Polyline	11068615115	Inland Fwy	M	S1100	791.858312	0.492037
Polyline	11015610603	Inner Belt	M	S1100	5928.296743	3.683665
Polyline	11033585809	Island Rd	M	S1100	1141.392783	0.709227
Polyline	11042592938	J F K Expy	M	S1100	11835.133119	7.353996
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	10847.909263	6.740565
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	26.48221	0.016455
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	10852.090751	6.743163
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	989.929148	0.615112
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	807.831421	0.501962
Polyline	11048243362	Jack Nicklaus Fwy	M	S1100	2828.426175	1.757499
Polyline	11048243363	Jack Nicklaus Fwy	M	S1100	3108.995705	1.931837
Polyline	11048243363	Jack Nicklaus Fwy	M	S1100	6775.813306	4.210287
Polyline	11048243363	Jack Nicklaus Fwy	M	S1100	14410.63839	8.954338
Polyline	11048243363	Jack Nicklaus Fwy	M	S1100	684.023063	0.425031

This data could be grouped (group_by), and statistics calculated upon the final field which represents miles, which would give us the total length of roadway per grouped vector. The data could also be aggregated and sorted according to its MTFCC attribute.

Tags: None

(Post is Unread)

Thread: Distribution of wealth in the US **Posted Date:** March 2, 2016 11:43 PM
Post: Distribution of wealth in the US **Status:** Published
Author:  **Ken Markus** **Overall Rating:**

Data: Distribution of Wealth in the US

Financial (Non-Home) Wealth			
	Top 1 percent	Next 19 percent	Bottom 80 percent
1983	42.9%	48.4%	8.7%
1989	46.9%	46.5%	6.6%
1992	45.6%	46.7%	7.7%
1995	47.2%	45.9%	7.0%
1998	47.3%	43.6%	9.1%
2001	39.7%	51.5%	8.7%
2004	42.2%	50.3%	7.5%
2007	42.7%	50.3%	7.0%
2010	42.1%	53.5%	4.7%

Source: http://www2.ucsc.edu/whorulesamerica/power/wealth.html

Analysis: Review how the distribution of wealth has changed for the three groups listed over time.

Tags: None

(Post is Read)

Thread: Distribution of wealth in the US **Posted Date:** March 4, 2016 1:23 AM
Post: RE: Distribution of wealth in the US **Status:** Published
Author:  **Chirag Vithalani** **Overall Rating:**

This is interesting , there is hardly any difference between first row and 2nd last row. i.e. not much changed in 24 years.

We see slight reduction in bottom 80%, but that is hardly any change.

Tags: None

(Post is Unread)

Thread: Rejections for firefarms
Post: Rejections for firefarms
Author: Gurpreet Singh
Posted Date: March 3, 2016 3:58 PM
Edited Date: March 3, 2016 4:01 PM
Status: Published
Overall Rating:

The dataset consists of number of applications received for firearms transfers, the rejections and rate of felonies from 1999 - 2010.

United States, 1999-2010

		Rejections				
		Total		Felons\a\		
	Applications received	Number	Percent	Number	Percent	Rate pe 1,000 applica
1999	8,621,000	204,000	2.4%	148,000	73%	
2000	7,699,000	153,000	2	88,000	58	
2001	7,958,000	151,000	1.9	87,000	58	
2002	7,806,000	136,000	1.7	65,000	48	
2003	7,831,000	126,000	1.6	53,000	42	
2004	8,084,000	126,000	1.6	53,000	42	
2005	8,278,000	132,000	1.6	57,000	43	
2006	8,612,000	135,000	1.6	52,000	39	
2007	8,658,000	136,000	1.6	49,000	36	
2008\b\	9,900,000	147,000	1.5	77,000	52	
2009	########	150,000	1.4	67,000	45	
2010	########	153,000	1.5	62,000	40	

Analysis: Rate of felonies per 1000 applications depends on number of applications rejected each year. Is there any correlation between two variables.

(or Can a linear model be created to predict one varible using other)

Attachment: background_check.csv (3.13 KB)

Tags: None

(Post is Unread)

Thread: Peer to Peer Loan Information
Post: Peer to Peer Loan Information
Author: Asher Meyers
Posted Date: March 3, 2016 5:37 PM
Status: Published
Overall Rating:

id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	term	int_rate	installment	grade	su
1077501	1296599	5000	5000	4975	36 months	10.65%	162.87	B	
1077430	1314167	2500	2500	2500	60 months	15.27%	59.83	C	
1077175	1313524	2400	2400	2400	36 months	15.96%	84.33	C	
1076863	1277178	10000	10000	10000	36 months	13.49%	339.31	C	
1075358	1311748	3000	3000	3000	60 months	12.69%	67.79	B	
1075269	1311441	5000	5000	5000	36 months	7.90%	156.46	A	

Lending Club publishes every single loan issued or rejected, and the corresponding information, with around 50 fields. https://www.lendingclub.com/info/download-data.action

Possible analysis: predicting likelihood of default based on borrower characteristics; identifying loans that have a low likelihood of default relative to their assigned grade and interest rate.

Tags: None

(Post is Unread)

Thread: National Income Acccounting **Posted Date:** March 3, 2016 9:34 PM
Post: National Income Acccounting **Status:** Published
Author: Shaun Smith **Overall Rating:**

From the BEA.gov website. The GDP, by year and by quarter, broken down by major categories of production. I've attached it here in Excel. The link is below in the citation.

Citation:

- U.S. Bureau of Economic Analysis, “Table 1.1.5. Gross Domestic Product http://bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&904=1929&903=5&906=a&905=2015&910= March 3, 2016).

Attachment: BEA_Data.xlsx (70.636 KB)

Tags: None

(Post is Unread)

Thread: Generator Capacity Prices
Post: Generator Capacity Prices
Author: Daniel Smilowitz
Posted Date: March 3, 2016 9:36 PM
Status: Published
Overall Rating:

The Data

Generators in the NYS electricity market are paid based on the outcome of competitive Unforced Capacity (UCAP) auctions.

There are 3 auctions:

- Monthly - takes place monthly before final values are known
- Spot - takes place monthly once values are finalized
- Strip - takes place twice a year (May and November) for the following six months

The prices for each auction are set in four different locations - NYC, Long Island (LI), Lower Hudson Valley (LHV), and Rest of State (ROS).

The past ~12.5 years worth of auction results are contained in a table, as depicted below:

Auction	Monthly				Spot				Strip			
Location	NYC	LHV	LI	ROS	NYC	LHV	LI	ROS	NYC	LHV	LI	ROS
November 2003	$ 6.67	$ 1.15	$ 0.50	$ 1.15	$ 6.98	$ 1.94	$ 8.14	$ 1.94	$ 6.55	$ 1.17	$ 4.00	$ 1.17
December 2003	$ 6.67	$ 1.58	$ 5.00	$ 1.58	$ 6.98	$ 1.78	$ 8.22	$ 1.78	$ 6.55	$ 1.17	$ 4.00	$ 1.17
January 2004	$ 6.67	$ 1.65	$ 8.10	$ 1.65	$ 6.98	$ 1.75	$ 7.99	$ 1.75	$ 6.55	$ 1.17	$ 4.00	$ 1.17
February 2004	$ 6.95	$ 1.67	$ 7.50	$ 1.67	$ 6.98	$ 1.73	$ 7.08	$ 1.73	$ 6.55	$ 1.17	$ 4.00	$ 1.17
March 2004	$ 6.25	$ 1.65	$ 7.00	$ 1.65	$ 6.98	$ 1.00	$ 7.72	$ 1.00	$ 6.55	$ 1.17	$ 4.00	$ 1.17
April 2004	$ 6.25	$ 0.99	$ 6.85	$ 0.99	$ 6.98	$ 0.80	$ 7.04	$ 0.80	$ 6.55	$ 1.17	$ 4.00	$ 1.17
May 2004	$ 11.16	$ 1.65	$ 8.00	$ 1.65	$ 11.42	$ 1.31	$ 9.83	$ 1.31	$ 11.15	$ 1.68	$ 8.00	$ 1.68
June 2004	$ 11.42	$ 1.39	$ 9.50	$ 1.39	$ 11.42	$ 1.27	$ 9.79	$ 1.27	$ 11.15	$ 1.68	$ 8.00	$ 1.68
July 2004	$ 11.42	$ 1.16	$ 8.75	$ 1.16	$ 11.42	$ 1.04	$ 8.42	$ 1.04	$ 11.15	$ 1.68	$ 8.00	$ 1.68
August 2004	$ 11.42	$ 1.00	$ 7.50	$ 1.00	$ 11.42	$ 1.17	$ 8.16	$ 1.17	$ 11.15	$ 1.68	$ 8.00	$ 1.68
September 2004	$ 11.42	$ 1.04	$ 7.50	$ 1.04	$ 11.42	$ 1.07	$ 8.15	$ 1.07	$ 11.15	$ 1.68	$ 8.00	$ 1.68
October 2004	$ 11.42	$ 1.08	$ 7.50	$ 1.08	$ 11.42	$ 1.12	$ 8.15	$ 1.12	$ 11.15	$ 1.68	$ 8.00	$ 1.68
November 2004	$ 6.96	$ 0.70	$ 4.00	$ 0.70	$ 7.12	$ 0.70	$ 6.34	$ 0.70	$ 6.68	$ 0.60	$ 4.00	$ 0.60
December 2004	$ 7.12	$ 0.70	$ 4.50	$ 0.70	$ 7.12	$ 0.61	$ 6.21	$ 0.61	$ 6.68	$ 0.60	$ 4.00	$ 0.60
January 2005	$ 7.12	$ 0.55	$ 4.50	$ 0.55	$ 7.12	$ 0.27	$ 6.16	$ 0.27	$ 6.68	$ 0.60	$ 4.00	$ 0.60
February 2005	$ 7.12	$ 0.30	$ 4.50	$ 0.30	$ 7.12	$ 0.25	$ 6.14	$ 0.25	$ 6.68	$ 0.60	$ 4.00	$ 0.60
March 2005	$ 7.12	$ 0.25	$ 4.25	$ 0.25	$ 7.12	$ 0.41	$ 6.16	$ 0.41	$ 6.68	$ 0.60	$ 4.00	$ 0.60
April 2005	$ 7.12	$ 0.35	$ 4.25	$ 0.35	$ 7.12	$ 0.27	$ 6.23	$ 0.27	$ 6.68	$ 0.60	$ 4.00	$ 0.60
May 2005	$ 11.86	$ 0.75	$ 8.00	$ 0.75	$ 12.00	$ 2.00	$ 12.15	$ 2.00	$ 11.68	$ 0.75	$ 8.00	$ 0.75
June 2005	$ 11.86	$ 1.55	$ 8.50	$ 1.55	$ 11.96	$ 1.96	$ 11.96	$ 1.96	$ 11.68	$ 0.75	$ 8.00	$ 0.75
July 2005	$ 11.86	$ 1.50	$ 9.00	$ 1.50	$ 11.95	$ 1.00	$ 10.48	$ 1.00	$ 11.68	$ 0.75	$ 8.00	$ 0.75
August 2005	$ 11.86	$ 0.85	$ 8.50	$ 0.85	$ 11.86	$ 1.00	$ 10.06	$ 1.00	$ 11.68	$ 0.75	$ 8.00	$ 0.75
September 2005	$ 11.86	$ 0.91	$ 8.50	$ 0.91	$ 11.70	$ 1.45	$ 9.90	$ 1.45	$ 11.68	$ 0.75	$ 8.00	$ 0.75
October 2005	$ 11.68	$ 1.35	$ 8.75	$ 1.35	$ 11.86	$ 1.45	$ 9.49	$ 1.45	$ 11.68	$ 0.75	$ 8.00	$ 0.75

All values are the price paid to generators in $/kW-month. I've attached the data in a messy .csv here.

Data source: New York Independent System Operator

The Analysis

Which month of the year sees the highest prices in each location?

What is the average difference between NYC and ROS prices?

Which calendar year saw the highest average price across regions (ignoring weighting)?

Is the monthly auction or the spot auction more volatile (i.e. which has the most variability in pricing)?

Attachment: UCAP.csv (16.262 KB)

Tags: None

(Post is Unread)

Thread:	Dirty Data - Done Dirt Cheap	**Posted Date:**	March 3, 2016 10:42 PM
Post:	Dirty Data - Done Dirt Cheap	**Status:**	Published
Author:	 **John Hodde**	**Overall Rating:**	

Here are some U.S. Amtrak train schedules:

Route	Cities Served	Region
Acela Express - see Northeast Corridor timetables.	Boston - New York - Philadelphia - Washington, DC	Northeast
Adirondack Effective January 11, 2016	Montreal - Albany - New York	Northeast
Amtrak Cascades Effective February 20, 2016	Vancouver, BC - Seattle - Tacoma - Portland - Salem - Eugene	Northwest
Auto Train Effective January 11, 2016	Lorton, VA (Washington, DC) - Sanford, FL (Orlando)	Northeast South
California Zephyr Effective January 11, 2016	Chicago - Denver - Salt Lake City - Emeryville / San Francisco	West Midwest California
Capitol Corridor Effective October 26, 2015	San Francisco - San Jose / Oakland - Sacramento / Auburn / Reno	California
Capitol Limited Effective January 11, 2016	Washington, DC - Pittsburgh - Cleveland - Chicago	Midwest Northeast
Cardinal & Hoosier State Effective January 11, 2016	Chicago - Indianapolis - New York	Midwest Northeast
Carolinian/Piedmont	New York - Raleigh - Charlotte	Northeast

Effective January 11, 2016		South
City of New Orleans Effective January 11, 2016	Chicago - Memphis - Jackson - New Orleans	Midwest South
Coast Starlight Effective January 11, 2016	Seattle - Portland - Los Angeles	West Northwest California
Crescent Effective January 11, 2016	New York - Washington - Atlanta - New Orleans	Northeast South
Downeaster Effective October 19, 2015	Brunswick - Portland - Boston	Northeast
Empire Builder Effective January 11, 2016	Chicago - Milwaukee - St. Paul / Minneapolis - Portland / Seattle	West Midwest Northwest
Empire Service Effective January 11, 2016	New York - Albany - Buffalo - Toronto	Northeast
Ethan Allen Express Effective January 11, 2016	Rutland - Albany - New York	Northeast
Heartland Flyer Effective January 8, 2016 Effective March 14, 2016	Oklahoma City - Fort Worth	West
Hiawatha Effective January 3, 2016	Milwaukee - Chicago	Midwest
Illinois / Missouri Services Effective January 11, 2016	Chicago - Quincy / Omaha, Carbondale, St. Louis / Kansas City	Midwest
Keystone Service Effective December 14, 2015	Harrisburg - Philadelphia - New York	Northeast
Lake Shore Limited Effective January 11, 2016	Chicago - Cleveland - Buffalo - Boston / New York	Midwest Northeast
Maple Leaf Effective January 11, 2016	New York - Buffalo - Toronto	Northeast
Michigan Services Effective January 11, 2016	Chicago - Detroit - Pontiac / Chicago - Grand Rapids	Midwest
Northeast Corridor 1 New York - Washington Effective January 11, 2016	New York - Philadelphia - Baltimore - Washington	Northeast South
Northeast Corridor 2 Boston / Springfield - Washington Effective January 11, 2016	Boston / Springfield - New Haven - New York - Philadelphia - Washington	Northeast South
Northeast Corridor 3 Boston - New York - Washington - Virginia Effective January 11, 2016	Boston - New York - Washington, DC - Charlottesville - Lynchburg - Richmond - Newport News - Virginia Beach - Norfolk	Northeast South
Northeast Regional - see Northeast Corridor	Boston - Springfield / Providence - New York - Washington, DC -	Northeast

timetables.	Richmond - Lynchburg - Newport News / Virginia Beach - Norfolk	South
Pacific Surfliner Effective October 5, 2015	San Luis Obispo - Santa Barbara - Los Angeles - San Diego	California
Pennsylvanian Effective January 11, 2016	New York - Philadelphia - Pittsburgh	Northeast
San Joaquin Effective January 11, 2016	San Francisco - San Joaquin Valley - Fresno - Bakersfield	California
Silver Service / Palmetto Effective January 11, 2016	New York - Washington, DC - Charleston - Savannah - Jacksonville - Orlando - Tampa / Miami	Northeast South
Southwest Chief Effective January 11, 2016	Chicago - Kansas City - Albuquerque - Los Angeles	West Midwest California
Sunset Limited Effective January 8, 2016 Effective March 14, 2016	New Orleans - San Antonio - Los Angeles	West South California
Texas Eagle Effective January 8, 2016 Effective March 14, 2016	Chicago - Dallas - San Antonio - El Paso - Tucson - Los Angeles	West Midwest South California
Vermonter Effective January 11, 2016	St. Albans - Springfield - New York - Washington	Northeast

https://www.amtrak.com/train-schedules-timetables

1. Clean up columns, remove extraneous noise

2. Determine most popular regions and cities

Attachment: Amtrak.xlsx (13.616 KB)

Tags: None

(Post is Unread)

Thread: Food additives.
Post: Food additives.
Author: Chirag Vithalani

Posted Date: March 4, 2016 1:09 AM
Edited Date: March 4, 2016 1:11 AM
Status: Published
Overall Rating:

1. Data

Food additive like "Carrageenan" often found in organic food can cause cancer (although it is approved by FDA).

I found out that for all food additives there is number system being followed

(EU-European Union follows it, US doesn't).

and here is table
from https://en.wikipedia.org/wiki/International_Numbering_System_for_Food_Additives

INS #	Approvals			Names	Type
100	A	E	U	curcumins	colour
100(i)	A	E	U	curcumin	colour (yellow and orange)
100(ii)	A	E	U	turmeric	colour (yellow and orange)
407	A	E	U	carrageenan	thickener, vegetable gum, stabilizer, gelling agent, emulsifier
428	A	E	U	gelatin, gelatine (not classified as an additive)	carrier, emulsifier, gelling agent, stabiliser, thickener
441			U	Superglycerinated hydrogenated rapeseed oil, Hydrogenated rapeseed oil superglycerinated, Superglycerinated fully hydrogenated rapeseed oil[5]	Emulsifier

and we click on " carrageenan", we see below text somewhere.

"Some animal studies indicate tumor promotion or initiation by carrageenan."

2. Analysis

1. Which food additives are harmful and approved.

2. Which food additive are approved by US (FDA) and not by EU (and vice versa). + If not approved then why ?

3. Which food additives are found in organic food.

PS: I have missed time line (of putting this reply) by an hour.

Tags: None

(Post is Unread)

← OK