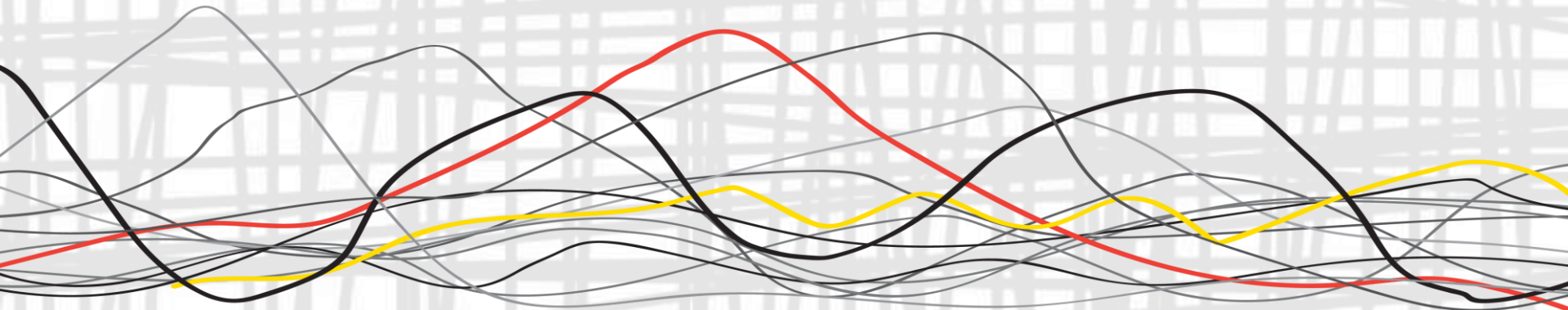# statgraphics 18®

2018's Top Ten Ways to Visualize Your Data

Presented by
Dr. Neil W. Polhemus

Webinar Outline

- We will examine 10 of the most useful graphs for data visualization that were popular in 2018.

- We will examine them in the context of selected datasets, showing their ability to display the information in the data.

- Subjective list (my choices) in no particular order.

- Concentrated on descriptive graphs.

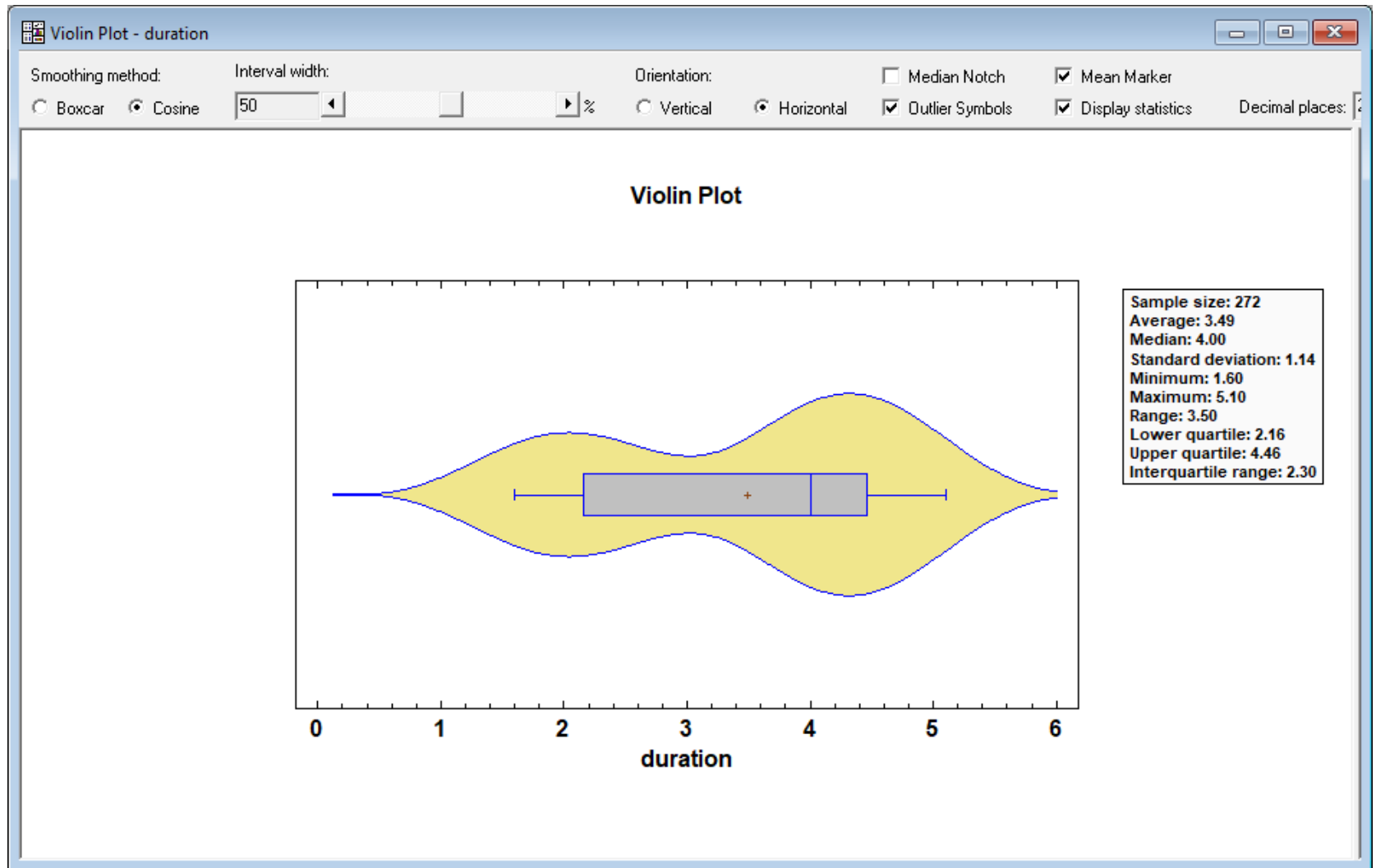# First Dataset: Eruptions of the Old Faithful Geyser in Yellowstone National Park

C:\Data\Webinars\Top Ten\old faithful.sgd

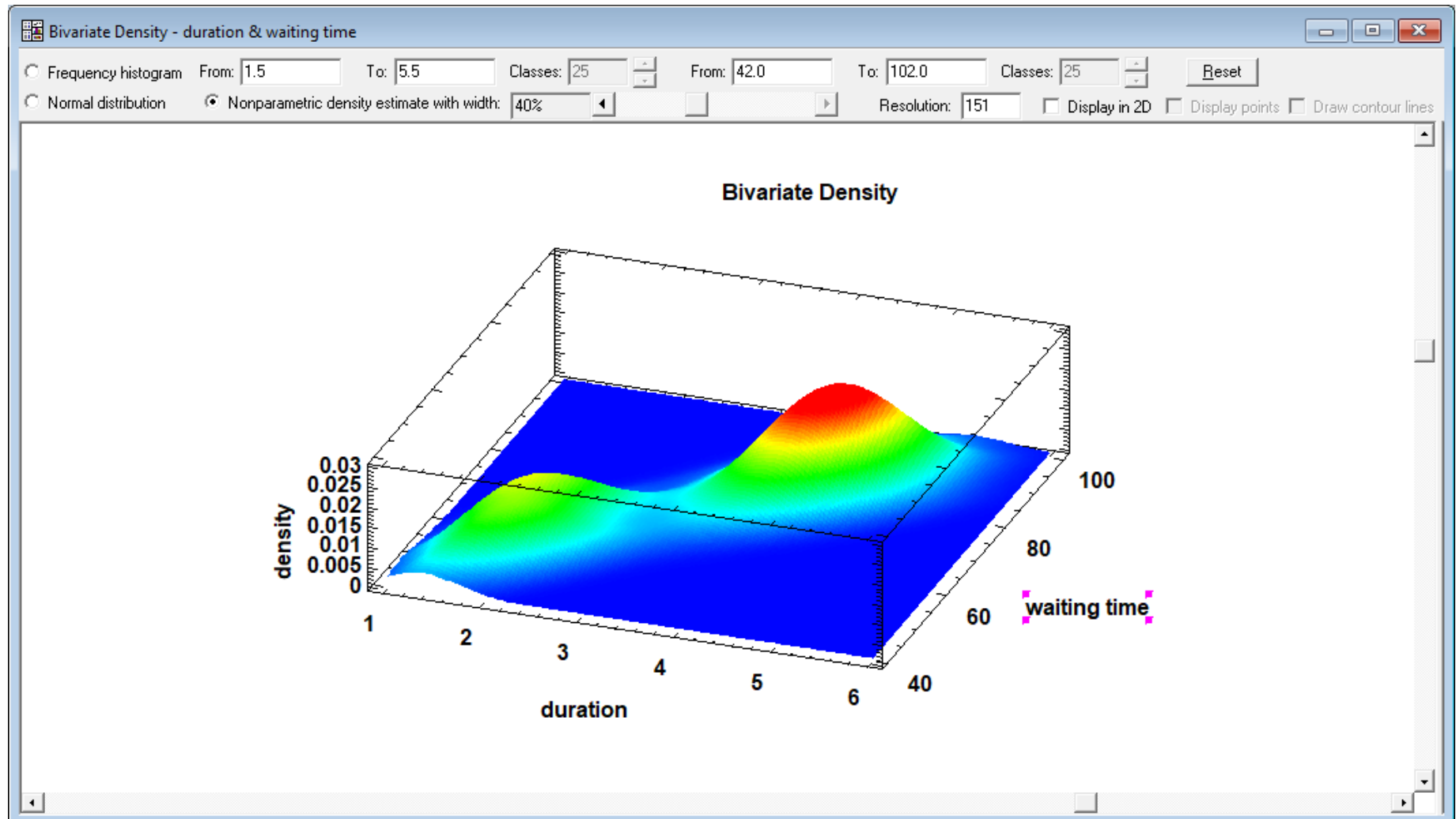|	duration	waiting time	subsequent duration
	minutes	minutes until next eruption	
	Numeric	Numeric	Numeric
1	3.600	79	1.800
2	1.800	54	3.333
3	3.333	74	2.283
4	2.283	62	4.533
5	4.533	85	2.883
6	2.883	55	4.700
7	4.700	88	3.600
8	3.600	85	1.950
9	1.950	51	4.350
10	4.350	85	1.833
11	1.833	54	3.917

old faithful

statgraphics 18

Questions

- What is the distribution of eruption duration?

- What is the distribution of waiting time between eruptions?

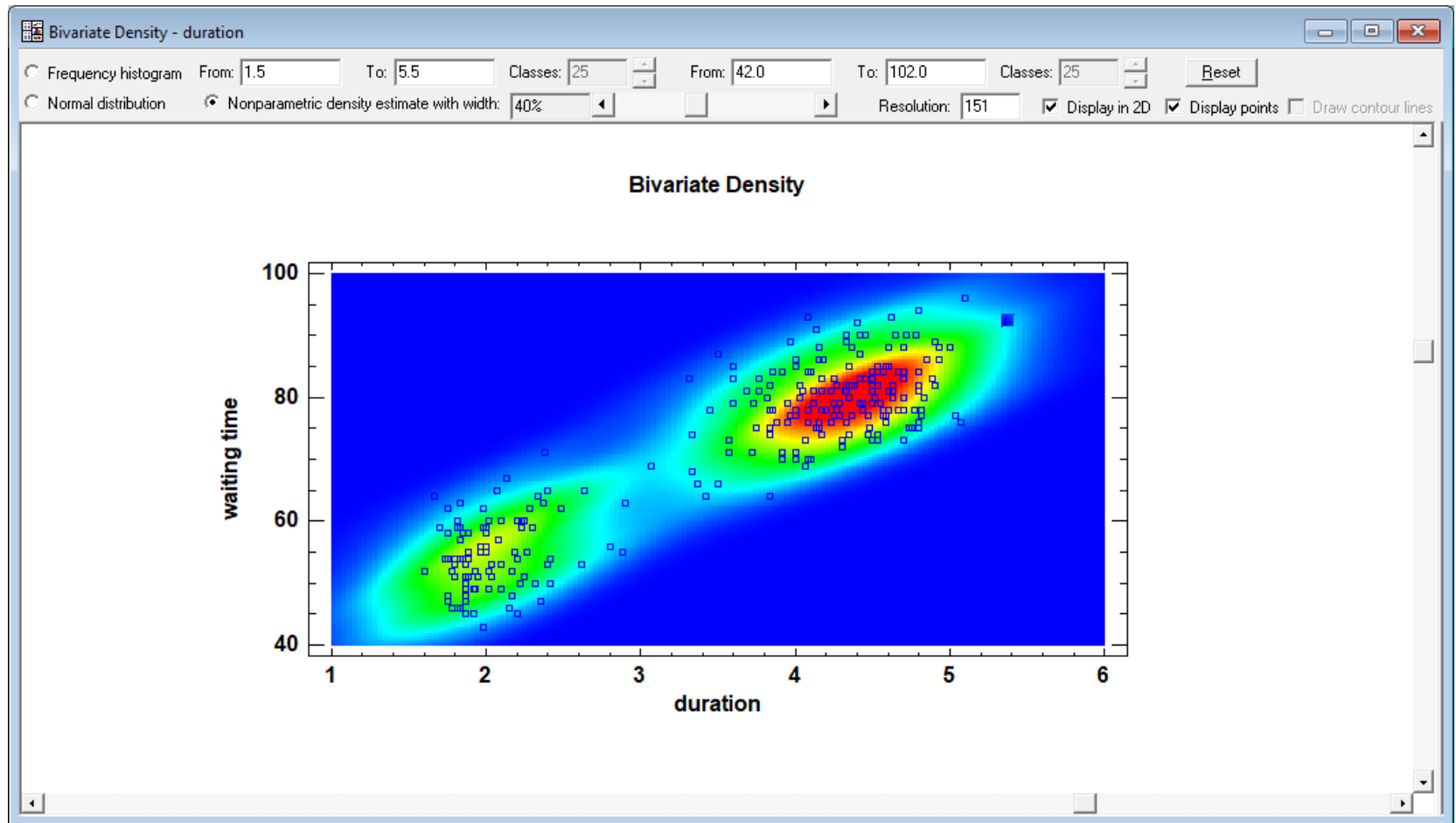- Is there a relationship between duration and waiting time?

Graph #1: Violin Plot

Graph #2: Bivariate Density Plot

Graph #2A: Bivariate Density Plot

statgraphics 18®

Second Dataset: U.S. Population

POPULATION CLOCK
December 15, 2018

USA
329,181,754

World
7,539,693,140

Source: census.gov

C:\Data\Webinars\Top Ten\us population.sgd

	Year	Age	Male	Female
	Numeric	Character	Numeric	Numeric
1	1950	0-4	8236164	7990587
2	1950	5-9	6714555	6485130
3	1950	10-14	5660399	5458869
4	1950	15-19	5311342	5305256
5	1950	20-24	5606293	5875535
6	1950	25-29	5972078	6270182
7	1950	30-34	5624723	5892284
8	1950	35-39	5517544	5728842
9	1950	40-44	5070269	5133704
10	1950	45-49	4526366	4544099
11	1950	50-54	4128648	4143540
12	1950	55-59	3630046	3605074
13	1950	60-64	3037838	3021637
14	1950	65-69	2424561	2578375
15	1950	70-74	1628829	1783120
16	1950	75-79	1506756	1770995
17	1950	80-84	500345	624225
18	1950	85+	236828	340073

us population

statgraphics 18®

Questions

- How is the U.S. population distributed by age?

- How do men compare to women?

- How has the distribution changed in recent years?

Graph #3: Population Pyramid
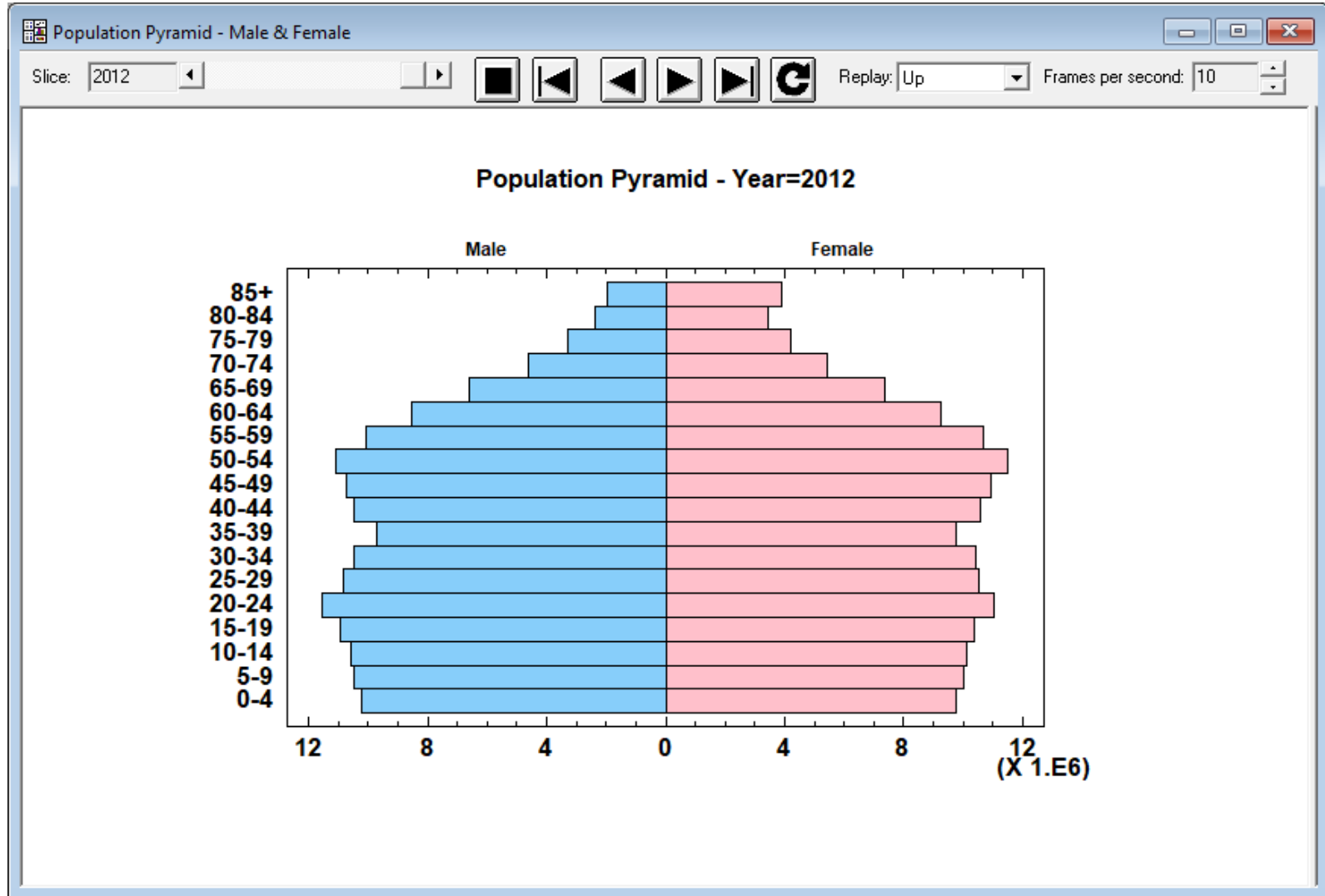
Third Dataset: Wind Speed and Direction

	Date	Time	Wind speed	Wind direction
			mph	degrees
	Date	Time (HH:MM)	Numeric	Numeric
1	1/1/2018	0:53	8	300
2	1/1/2018	1:53	10	300
3	1/1/2018	2:53	11	310
4	1/1/2018	3:53	11	310
5	1/1/2018	4:53	9	300
6	1/1/2018	5:53	13	290
7	1/1/2018	6:53	10	280
8	1/1/2018	7:53	9	280
9	1/1/2018	8:53	10	290
10	1/1/2018	9:53	11	290
11	1/1/2018	10:53	10	270
12	1/1/2018	11:53	10	270
13	1/1/2018	12:53	11	280
14	1/1/2018	13:53	11	280
15	1/1/2018	14:53	13	260
16	1/1/2018	15:53	11	270
17	1/1/2018	16:53	9	270

C:\Data\Webinars\Top Ten\midway airport.sgd

midway airp

Source: Midwestern Regional Climate Center
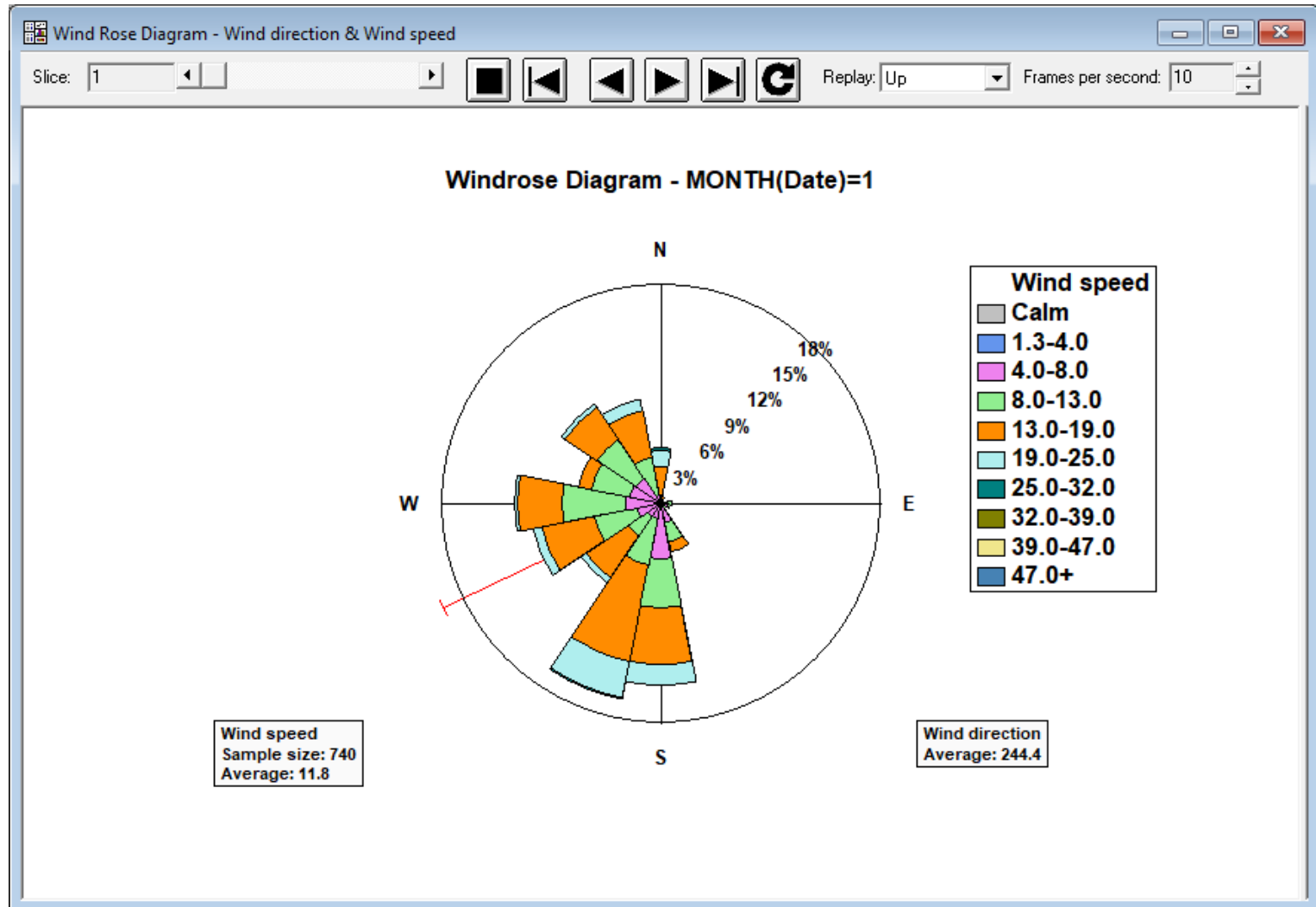
statgraphics 18®

Questions

- What is the dominant wind direction?

- What is the average and maximum wind speed?

- How do direction and wind speed change throughout the day and year?

Graph #4: Wind Rose
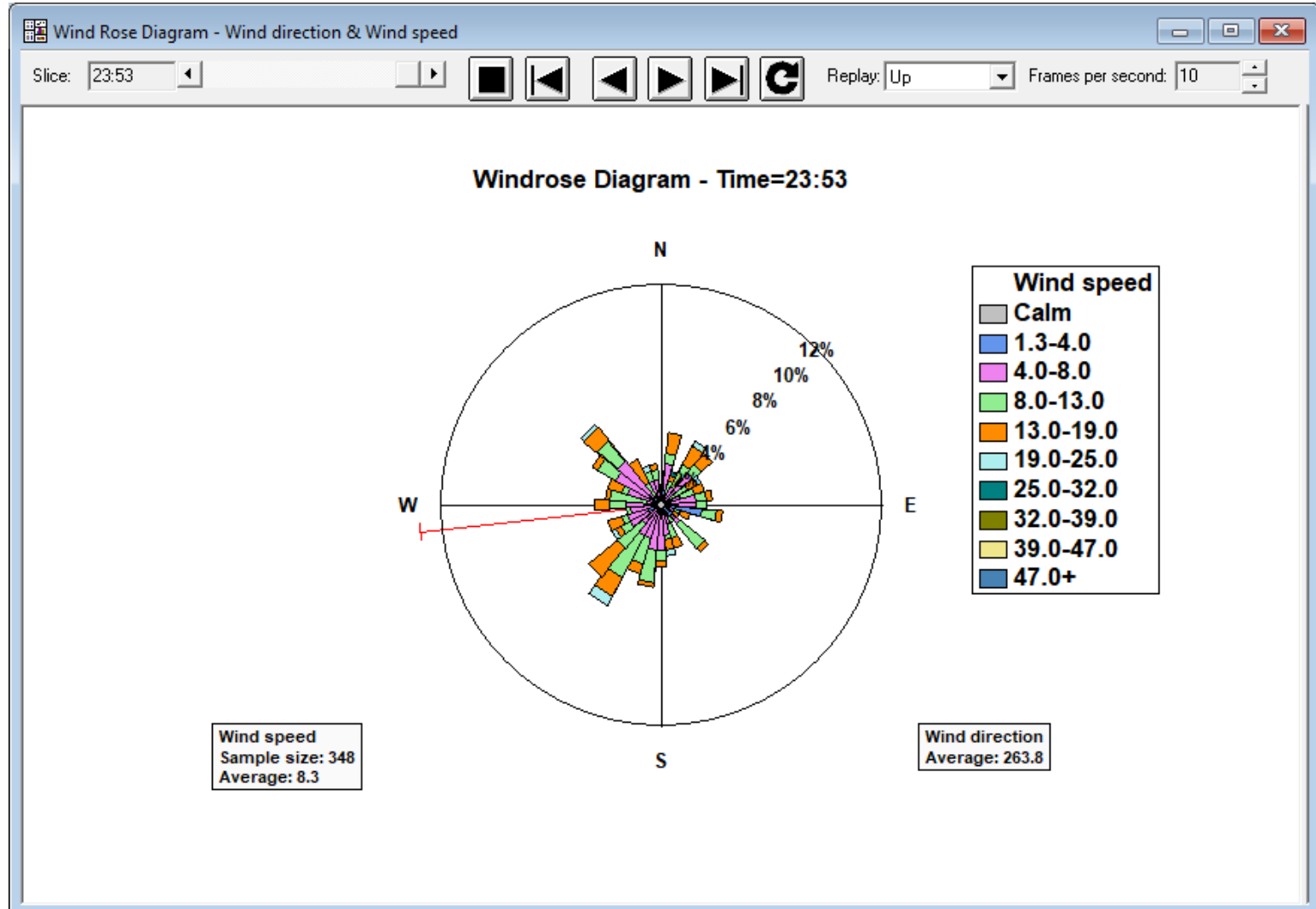
Graph #4A: Wind Rose

Fourth Dataset: World Demographics

	Country Code	Country	Year	Population	Pop. Density	Rural Population	Female Percentage	Age Dependency Ratio	Life Expectancy (Total)
				total	people per sq. km of land area	% of total population	% of total population	% of working-age population	years at birth
	Character	Character	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
1	ABW	Aruba	1961	55436	307.98	49.22	50.97	85.56	65.99
2	ABW	Aruba	1962	56227	312.37	49.24	50.97	84.44	66.37
3	ABW	Aruba	1963	56698	314.99	49.26	50.99	83.02	66.71
4	ABW	Aruba	1964	57031	316.84	49.28	51.01	81.42	67.04
5	ABW	Aruba	1965	57362	318.68	49.3	51.03	79.76	67.37
6	ABW	Aruba	1966	57714	320.63	49.32	51.04	78.03	67.7
7	ABW	Aruba	1967	58052	322.51	49.34	51.04	76.26	68.03
8	ABW	Aruba	1968	58388	324.38	49.36	51.05	74.4	68.38
9	ABW	Aruba	1969	58725	326.25	49.38	51.07	72.47	68.73
10	ABW	Aruba	1970	59066	328.14	49.4	51.1	70.44	69.09

C:\Data\Webinars\Top Ten\worldbank.sgd

worldbank B C

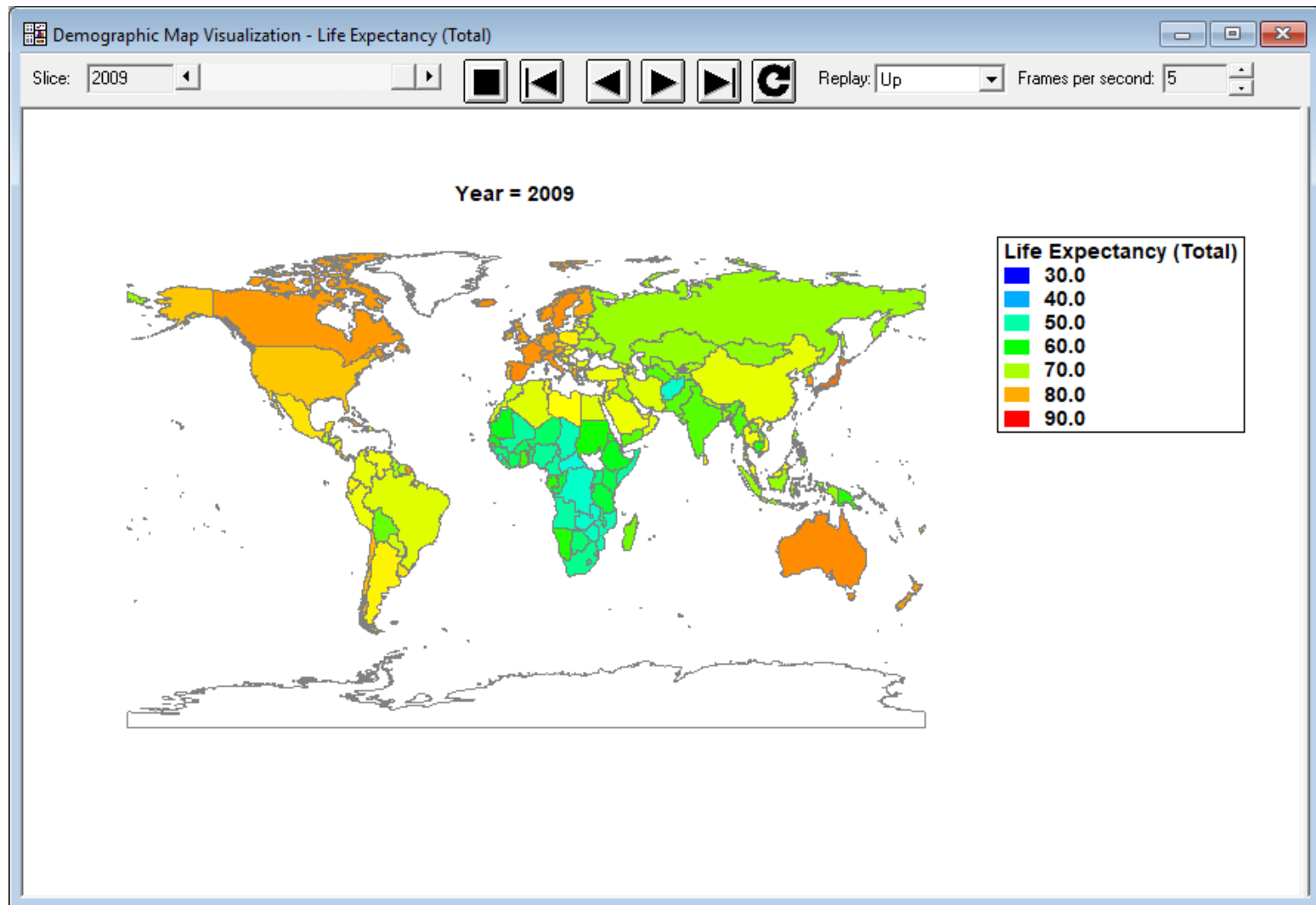Source: worldbank.com

statgraphics 18®

Questions

- How does life expectancy vary throughout the world?

- Is it related to other demographic factors?

- How has life expectancy changed in recent years?

Graph #5: Dynamic Bubble Chart
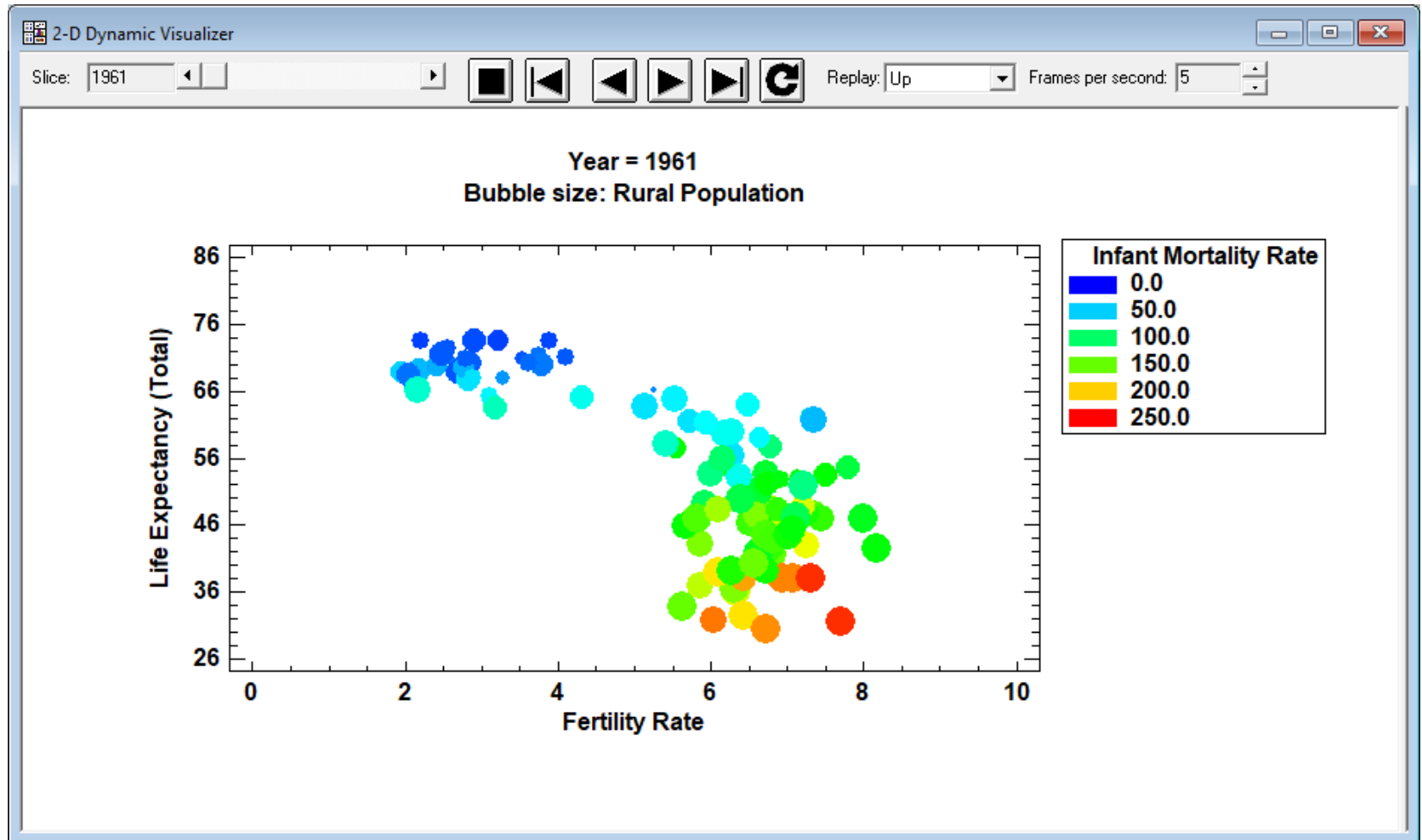
Graph #6: Demographic Gradient Map

Fifth Dataset: Sea Water Temperatures

	SEAS	YR	TOTAL	ONI	Month
	Character	Numeric	Numeric	Numeric	Month
815	NDJ	2017	25.65	-0.97	11/2017
816	DJF	2018	25.73	-0.86	12/2017
817	JFM	2018	26.02	-0.75	1/2018
818	FMA	2018	26.60	-0.60	2/2018
819	MAM	2018	27.19	-0.41	3/2018
820	AMJ	2018	27.61	-0.13	4/2018
821	MJJ	2018	27.65	0.06	5/2018
822	JJA	2018	27.38	0.11	6/2018
823	JAS	2018	27.19	0.20	7/2018
824	ASO	2018	27.26	0.44	8/2018
825	SON	2018	27.48	0.72	9/2018

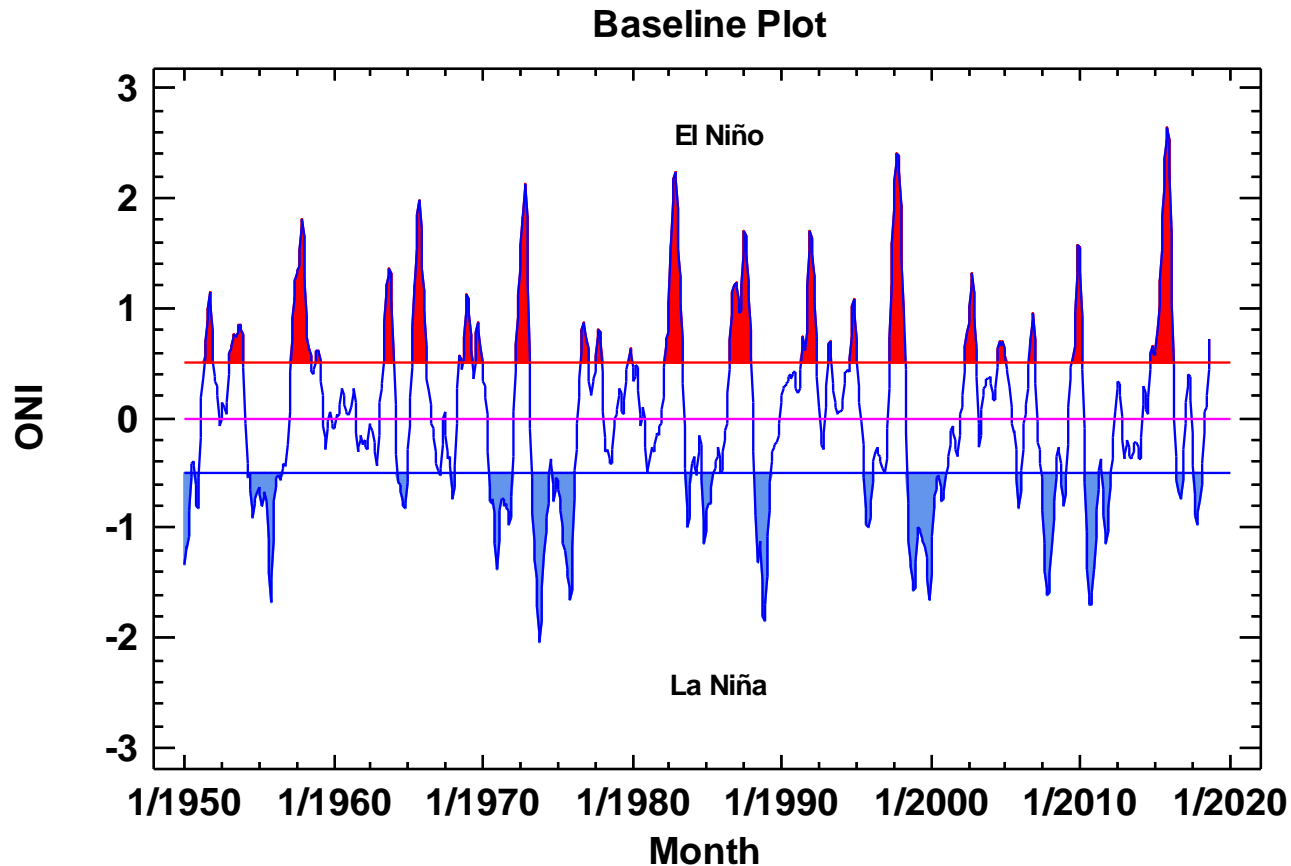C:\Data\Webinars\Top Ten\oni2.sgd

Source: NOAA Climate Prediction Center

statgraphics 18®

Questions

- What does the cyclical pattern of La Niñas and El Niños look like?

- How long do such events usually last?

- What type of period are we in currently?

Graph #7: Time Series Baseline Plot

Baseline Plot

Sixth Dataset: Bitcoin Prices
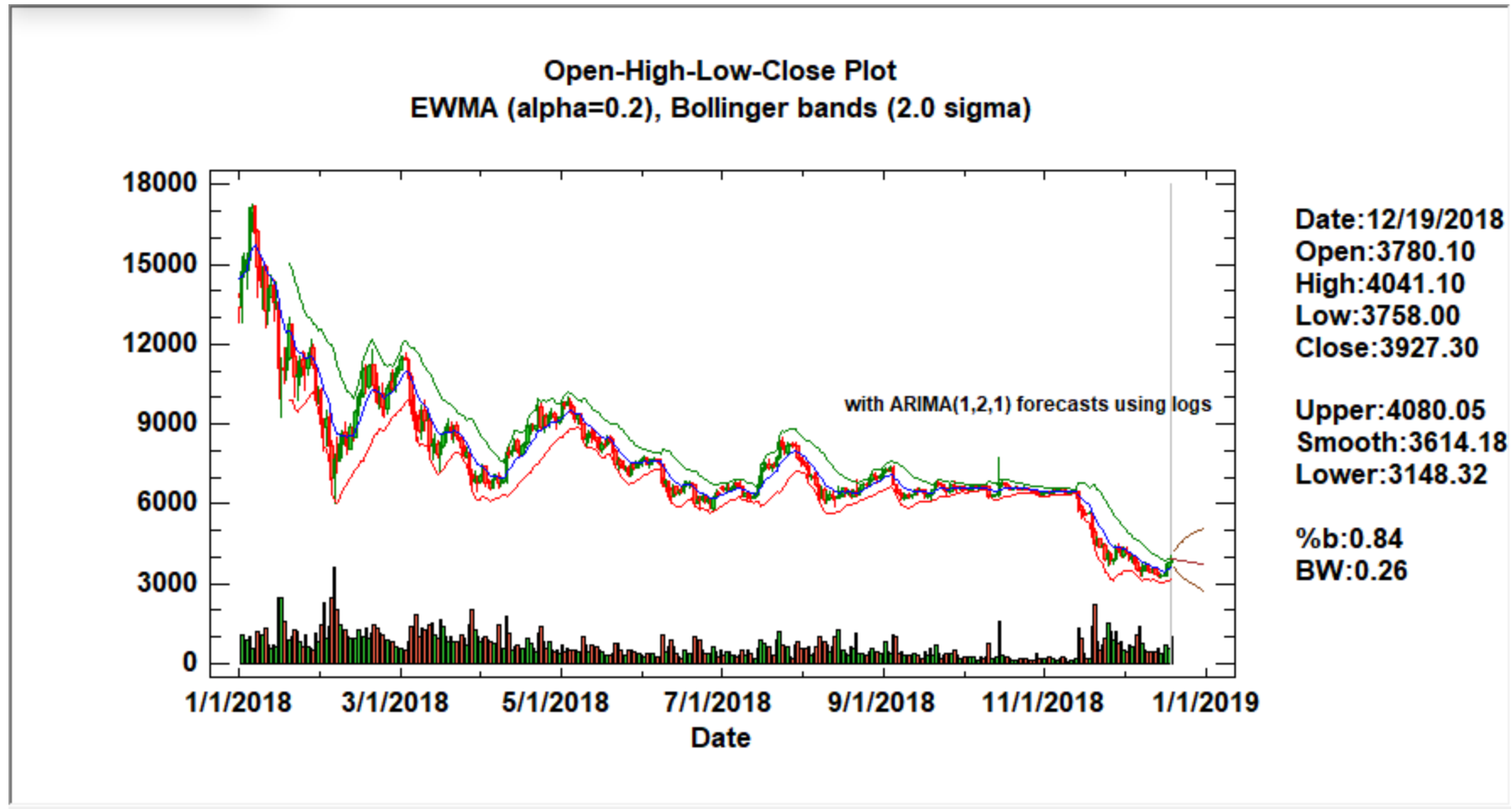

C:\Data\Webinars\Top Ten\bitcoin price new.sgd	Date	Price	Open	High	Low
	Date	Numeric	Numeric	Numeric	Numeric
338	12/4/2018	3981.9	3903.8	4131.1	3793.1
339	12/5/2018	3765.3	3982.2	4000.1	3751.1
340	12/6/2018	3531.3	3770.5	3908	3506
341	12/7/2018	3460	3528.6	3624.8	3278.4
342	12/8/2018	3506	3460	3784.8	3316.2
343	12/9/2018	3639.6	3506	3766.6	3469.7
344	12/10/2018	3524.9	3642.1	3697.9	3475.8
345	12/11/2018	3435.3	3523.5	3549.6	3384
346	12/12/2018	3535.6	3435.3	3611.2	3415.1
347	12/13/2018	3354.4	3535.6	3539.1	3322.2
348	12/14/2018	3282.8	3354.4	3390.8	3226.4
349	12/15/2018	3244.9	3281.6	3319.8	3219.2

bitcoin price new   B   C

Questions

- How has the price of Bitcoin changed during 2018?

- Have there been periods with more volatility than others?

- What is likely to happen to the price over the next few weeks?

Graph #8: OHLC Plot with Forecasts

Open-High-Low-Close Plot
EWMA (alpha=0.2), Bollinger bands (2.0 sigma)

with ARIMA(1,2,1) forecasts using logs

Date:12/19/2018
Open:3780.10
High:4041.10
Low:3758.00
Close:3927.30

Upper:4080.05
Smooth:3614.18
Lower:3148.32

%b:0.84
BW:0.26

Seventh Dataset: Fish Counts in Gulf of Maine

C:\Data\Webinars\Top Ten\Gulf of Maine.sgd

|	year	Bay of Fundy	Browns Bank	Central Gulf of Maine	Continental Slope	Eas Coasta
	Numeric	Numeric	Numeric	Numeric	Numeric	Nu
1	1963	0	4	92	14	23
2	1964	5	7	123	23	34
3	1965	7	8	107	23	33
4	1966	4	7	74	20	15
5	1967	2	2	31	4	8
6	1968	2	7	78	17	19
7	1969	3	7	108	24	33
8	1970	3	6	86	16	21
9	1971	1	7	80	22	27
10	1972	5	9	86	26	22

Gulf of Maine   B

Questions

- How have fish counts changed over the years?

- Which years and locations have the highest counts?

- Is there a correlation between the counts at different locations?

Graph #9: Heat Map

Heat Map
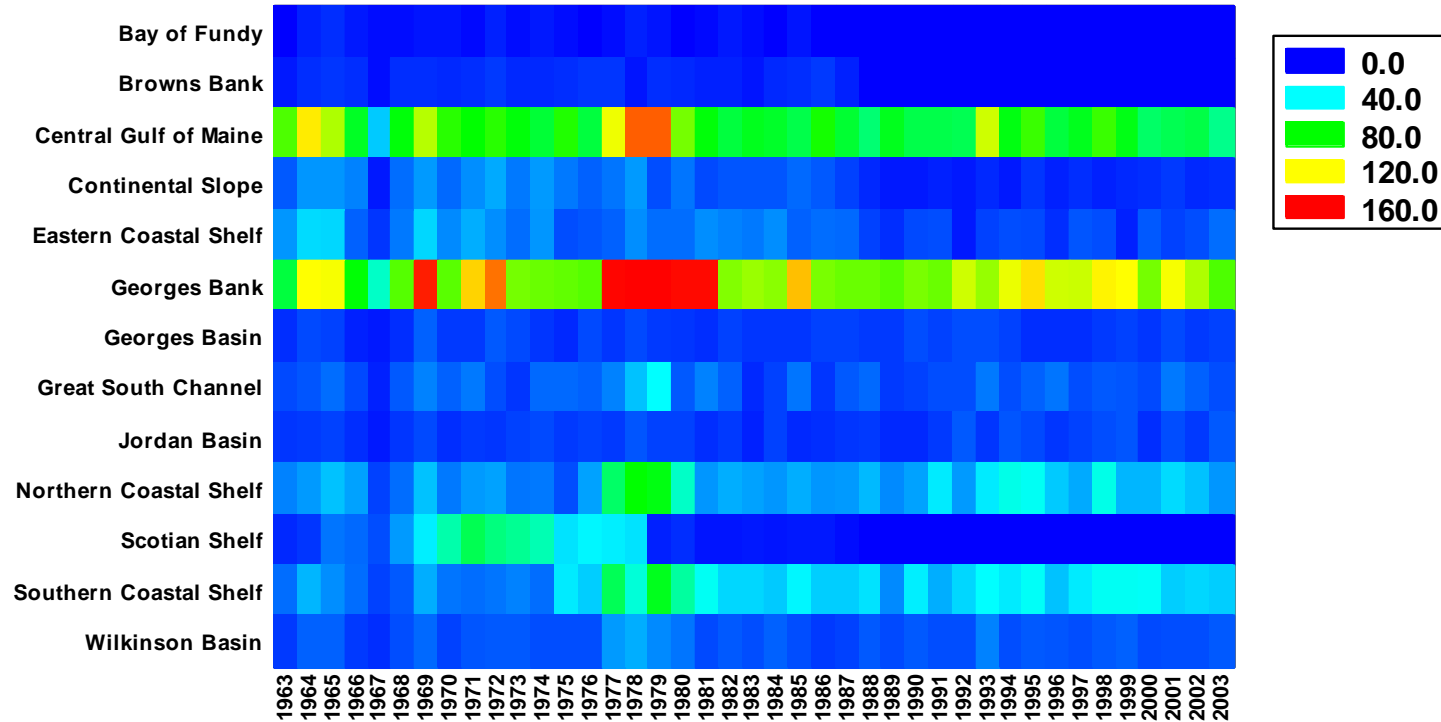
Eighth Dataset: U.S. Commercial Airline Flights

	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime	ArrDelay	DepDelay	Origin	Dest	D
	plane tail number	in minutes	in minutes	in minutes	arrival delay, in minutes	departure delay, in minutes	origin (IATA airport code)	destination (IATA airport code)	i
	Character	Integer	Integer	Integer	Integer	Integer	Character	Character	
6858192	N313US	150	164	130	78	92	MCO	DTW	957
6858193	N592NW	172	165	139	22	15	MCO	DTW	957
6858194	N594NW	168	165	136	-3	-6	MCO	DTW	957
6858195	N581NW	181	165	132	40	24	MCO	DTW	957
6858196	N594NW	157	165	128	-8	0	MCO	DTW	957
6858197	N590NW	154	165	125	-5	5	MCO	DTW	957
6858198	N590NW	177	165	135	67	55	MCO	DTW	957

C:\Data\Webinars\Top Ten\2008 Flight Data.sgb

2008 Flight Data / B / C

Questions

- What is the distribution of departure delays?

- Are arrival delays related to departure delays?

- How clean is the data?

Graph #10: Hexagon Plot

Plot of Fitted Model
ArrDelay = -1.94045 + 1.01865*DepDelay
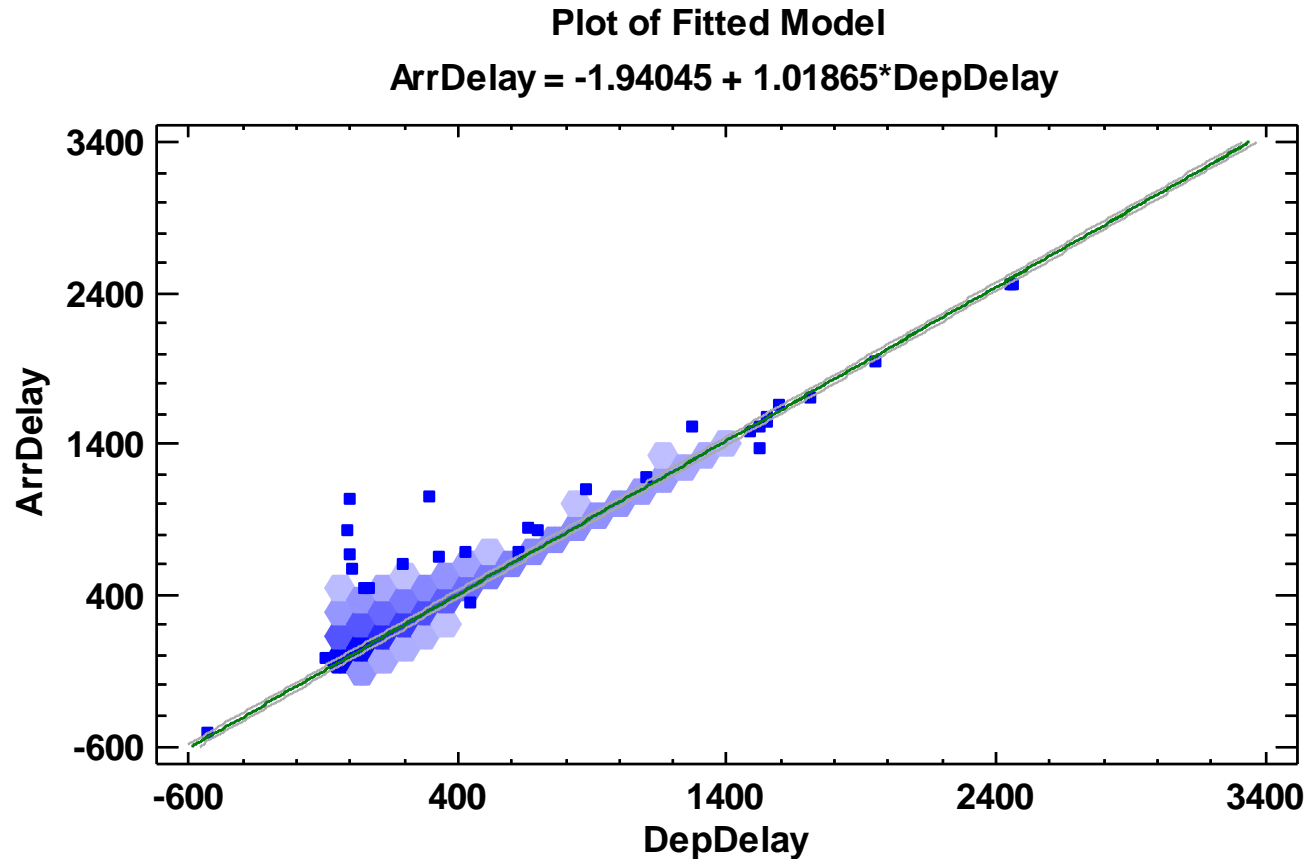
The Top Ten

1. Violin Plot
2. Bivariate Density Plot
3. Population Pyramid
4. Wind Rose
5. Dynamic Bubble Chart
6. Demographic Gradient Map
7. Time Series Baseline Plot
8. Open-High-Low-Close Plot with Forecasts
9. Heat Map
10. Hexagon Plot

References

- StatFolios and data files are at:
  www.statgraphics.com/webinars