

**STATISTICA ECONOMICA PER L'ANALISI
MICROECONOMICA**

Prof. Alessandro ZINI

**ANALISI STATISTICA DELLA SERIE STORICA
DEI NUMERI INDICE DEI VOLUMI DI
FATTURATO MANIFATTURIERO IN ITALIA, DAL
2002 AL 2016**

Tesi di: Andrea Rossi Matr.N. 750339

INDICE

Presentazione dei dati	p.3
Analisi dei dati	p.6
Previsioni	p.12
Appendice	p.17

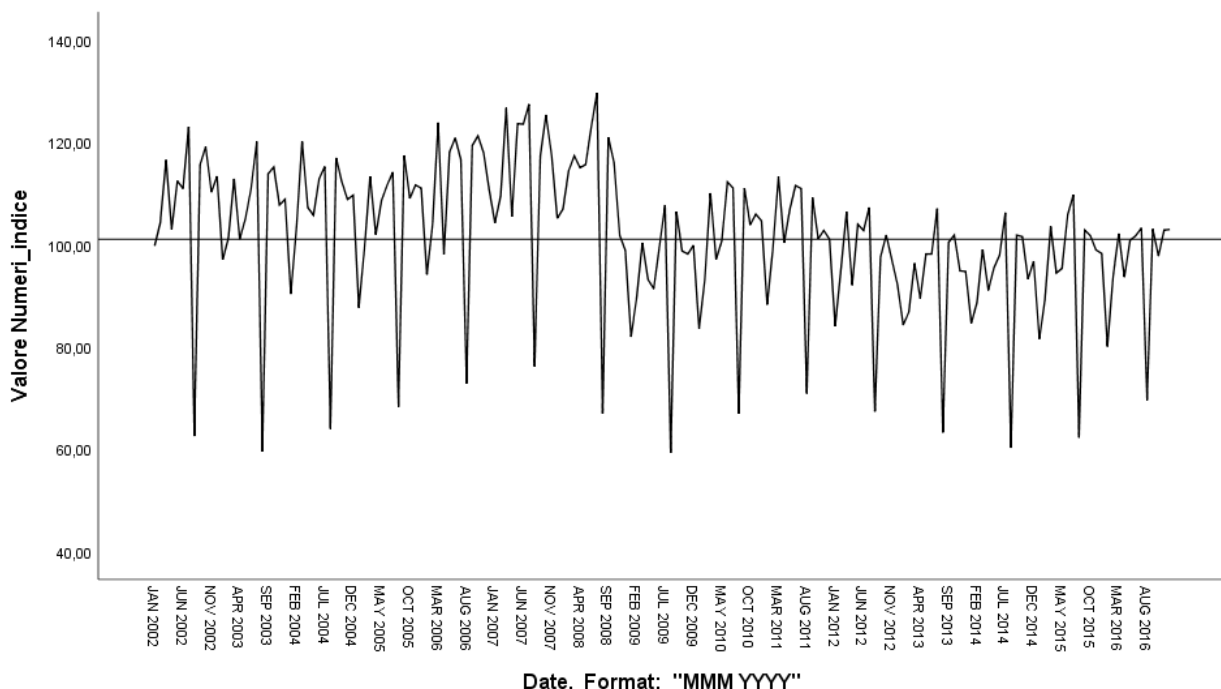
1. PRESENTAZIONE DEI DATI

I dati utilizzati¹ per questo studio sono i numeri indice relativi al volume del fatturato manifatturiero in Italia dal Gennaio 2002 al Dicembre 2016 prendendo come base (100) i dati relativi all'anno 2010.

L'obiettivo di questa analisi è stimare i valori relativi all'anno 2016 e di verificarne la bontà di approssimazione con i valori effettivamente osservati per quell'anno.

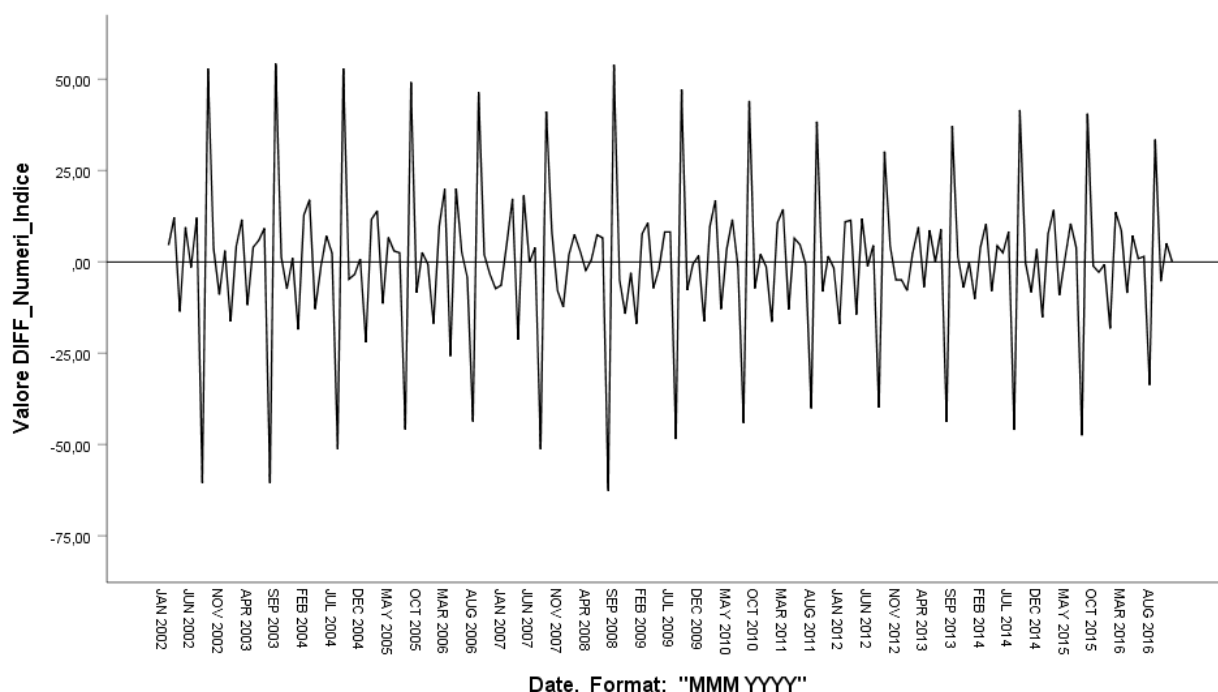
Come prima cosa dobbiamo verificare che i dati scelti siano stazionari sia in media che in varianza, possiamo fare questo osservando la spezzata che unisce i punti di coordinate $(x;t)^2$

Di seguito il grafico:



Immediatamente si nota che l'andamento della spezzata non è stazionario rispetto alla media³.

In queste condizioni i dati non possono essere analizzati. Quindi, per risolvere questo problema, si trasformano i valori osservati attraverso il calcolo delle differenze prime⁴, per renderli stazionari in medi e se ne rappresenta la relativa spezzata:

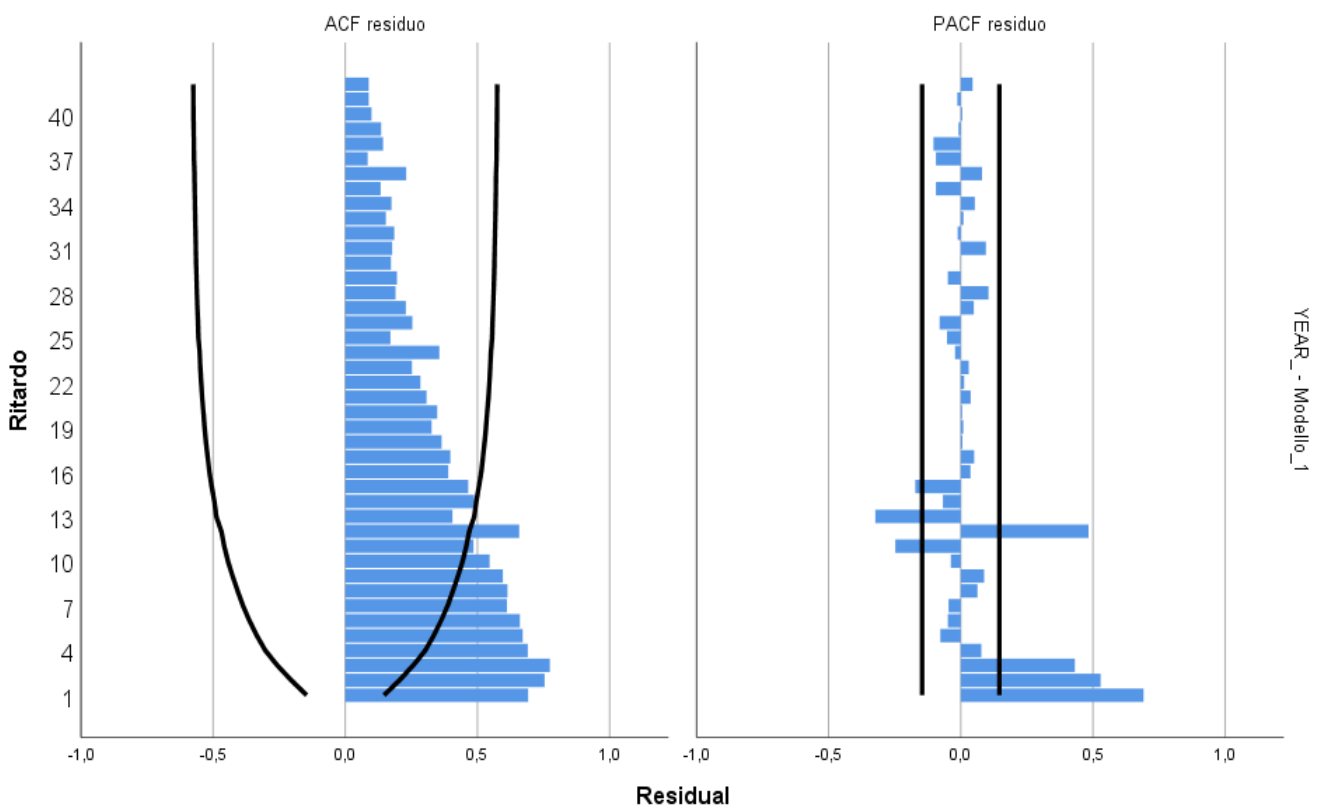


L'andamento della spezzata è notevolmente migliorato e si presta meglio al nostro studio, infatti si nota adesso la presenza di una stazionarietà sia in media che in varianza⁵ dei valori osservati. E' ora possibile effettuare le successive analisi sui dati.

2. ANALISI DEI DATI

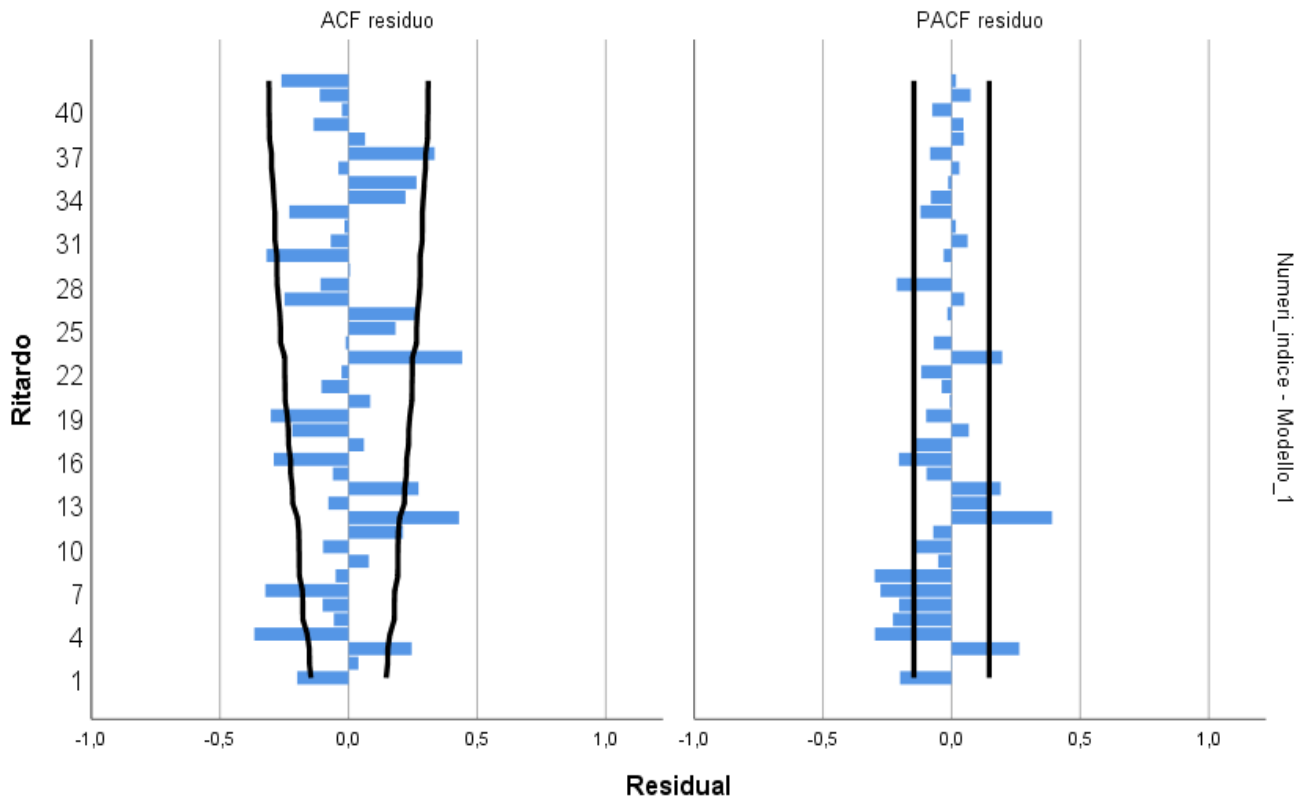
Terminata la presentazione e le analisi preliminari si procede con l'analisi delle funzioni di autocorrelazione con 42 ritardi⁶, applicando i modelli ARIMA (p, d, q) (P, D, Q)⁷.

Come prima cosa si ricavano le funzioni di autocorrelazione globale e parziale dei dati originari non ancora differenziati⁸:



La non stazionarietà in media della serie originale è appurata dalla funzione di autocorrelazione globale, la quale mostra un decadimento molto lento verso lo zero.

Procediamo nella nostra analisi differenziando⁹ la serie e rappresentandone le autocorrelazioni dei residui:



Osservando il grafico dell'autocorrelazione parziale si nota un ritardo positivo e maggiore delle bande di confidenza, precisamente al ritardo 12. Inoltre dal grafico dell'autocorrelazione globale si osservano picchi significativi al di fuori delle bande di confidenza ogni 12 mesi.

Utilizzando il programma statistico SPSS, in particolare l'opzione di creazione dei modelli Expert Modeler, questo individua come modello ARIMA migliore il seguente: ARIMA (2, 0, 0) (1, 1, 1)

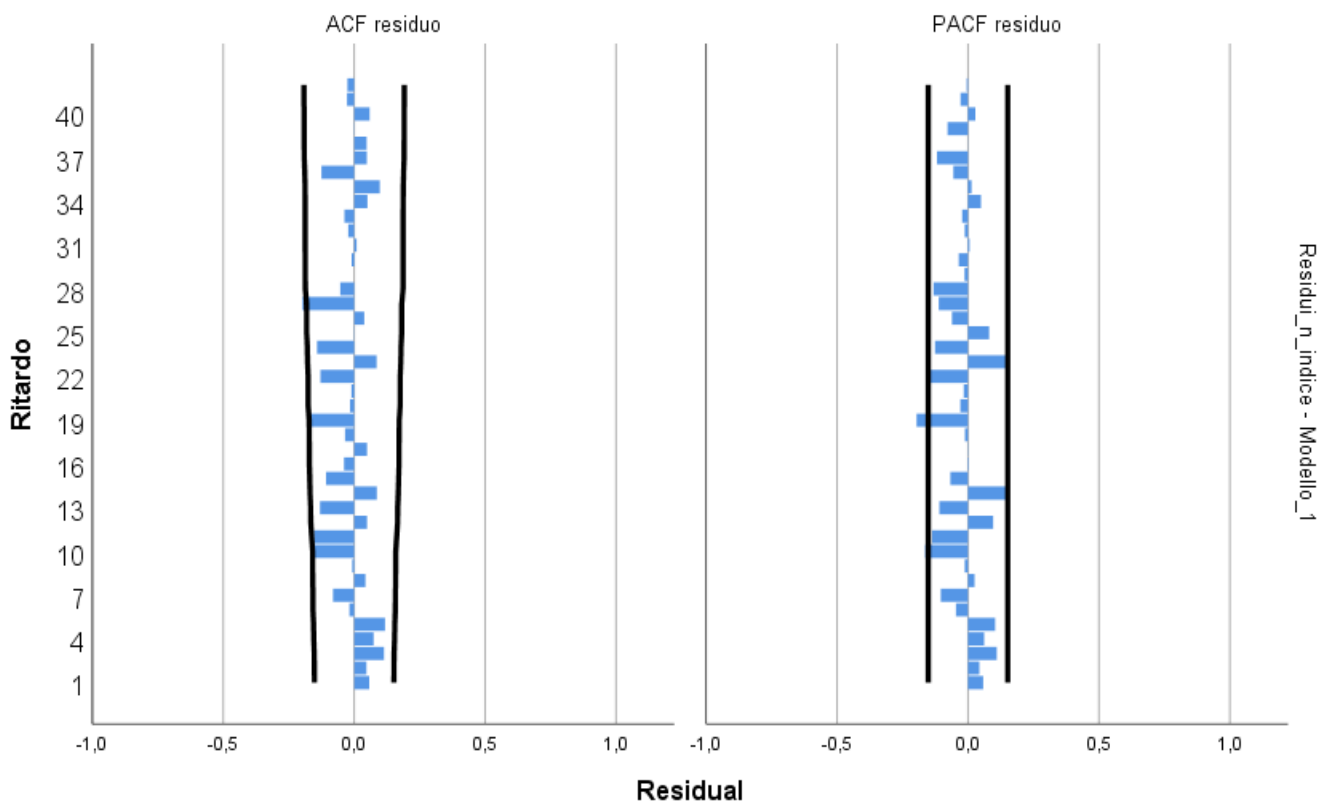
Questo modello presenta i seguenti valori¹⁰:

Statistiche adattamento modello

Test di Ljung Q

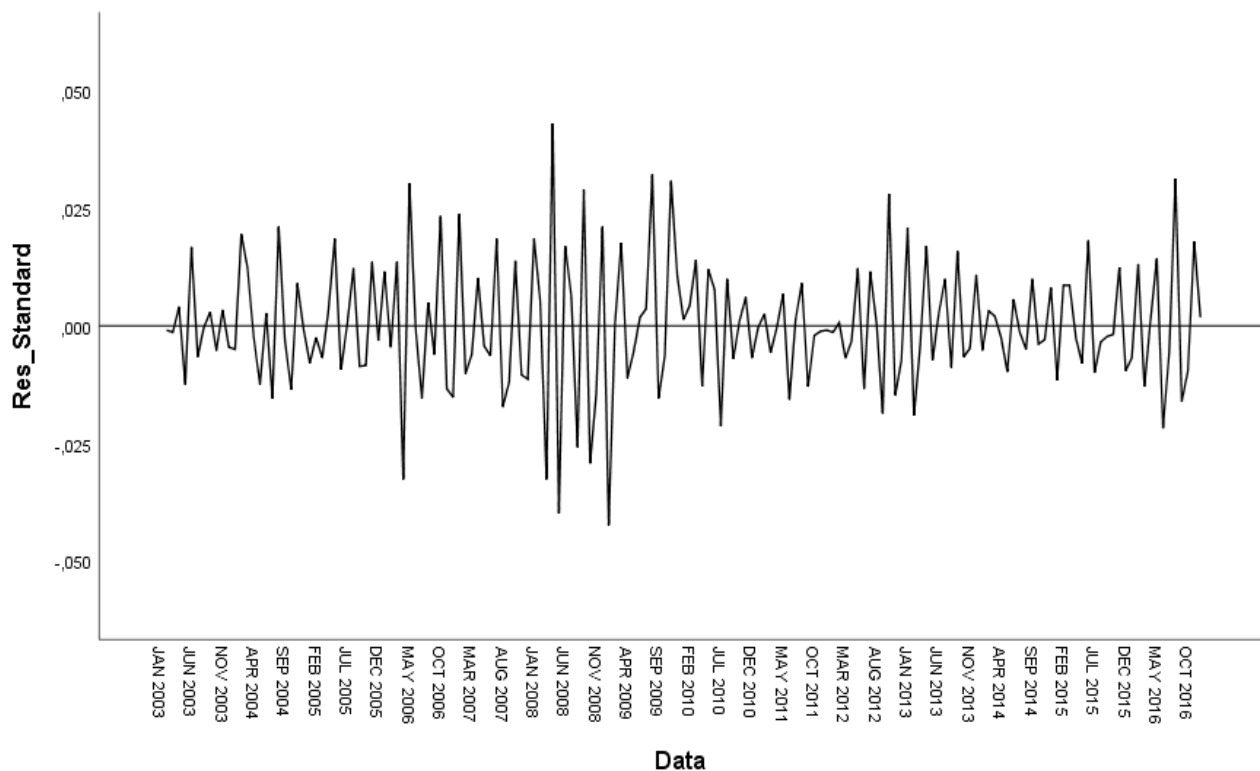
RMSE	MAE	Q	p-value
4,157	3,230	25,474	0,030

E' necessario verificarne la bontà di adattamento, considerando i grafici delle autocorrelazioni dei residui.



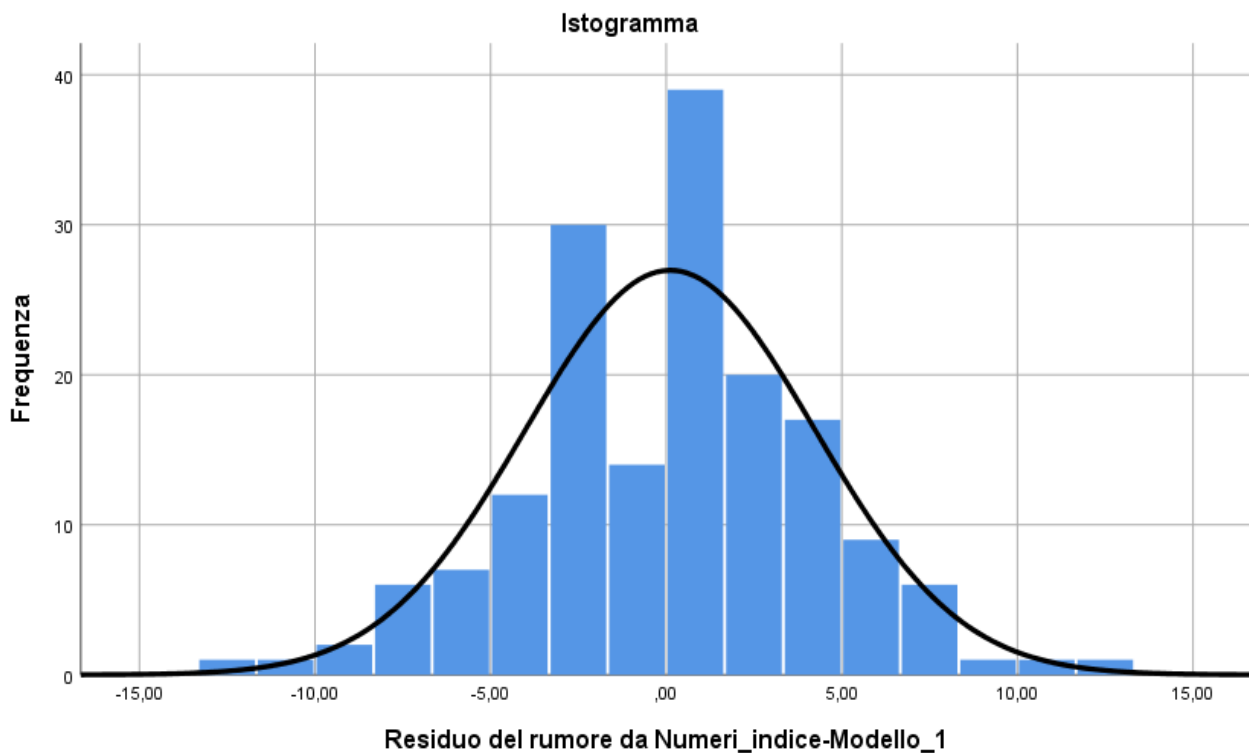
I coefficienti di correlazione sono contenuti all'interno delle due bande di confidenza, questo ci indica che i residui non sono tra loro correlati.

Per effettuare un'ulteriore verifica si calcolano i residui standardizzati¹¹ e se ne rappresenta una spezzata di punti $(x;t)$ ¹²:



I residui non sono anomali, hanno media non significativamente diversa da zero e varianza costante; costituiscono quindi la realizzazione di un processo "white noise"¹³, come doveva risultare sulla base delle ipotesi assunte.

E' inoltre necessario verificare se questi si distribuiscono come una normale. A questo fine si ricorre alla rappresentazione di un istogramma delle frequenze relative dei residui.

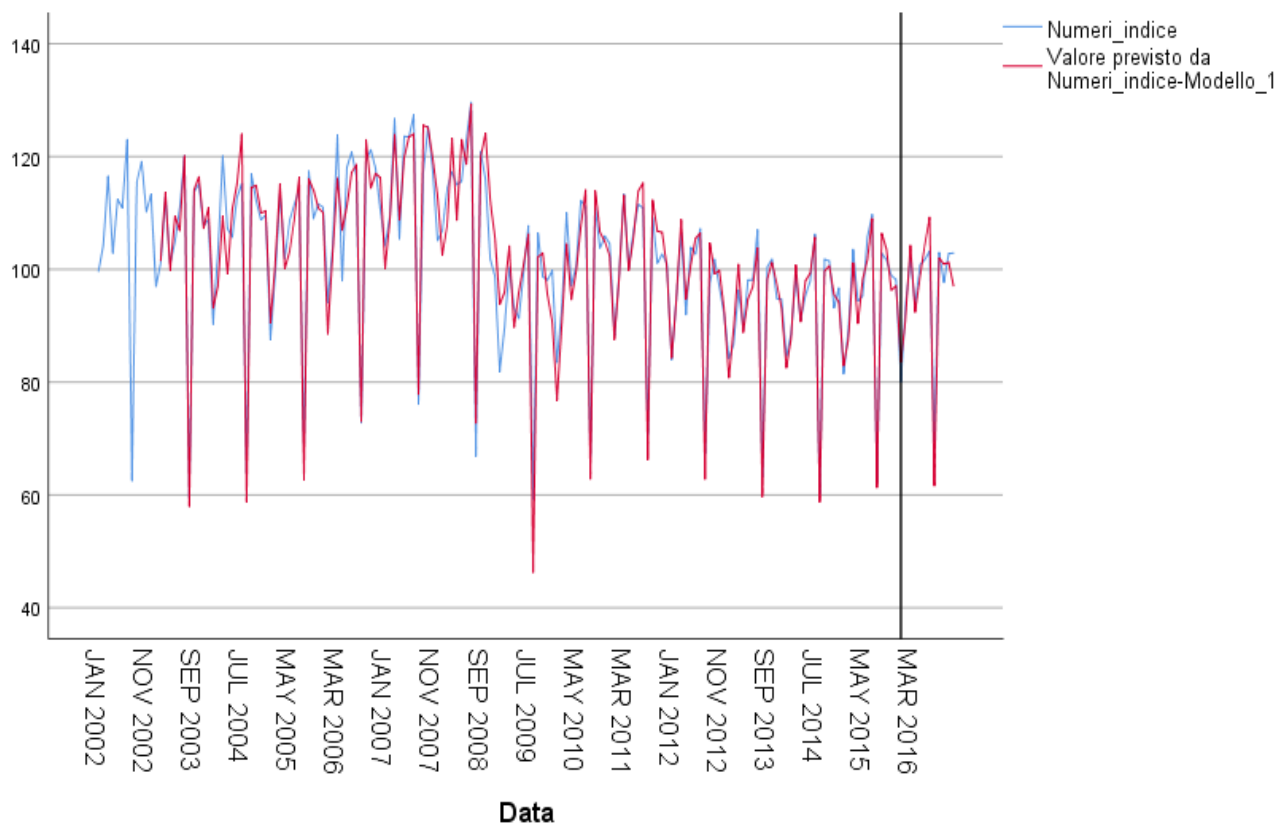


Il grafico presenta una distribuzione normale dei residui con la classica forma a campana; possiamo quindi dire che il modello si presta bene alla spiegazione del fenomeno oggetto di studio.

3. PREVISIONI

Questa analisi si conclude con il confronto tra i valori effettivamente osservati e quelli previsti.

Confrontando le due spezzate dei punti ci si accorge che il modello scelto spieghi bene il fenomeno studiato:

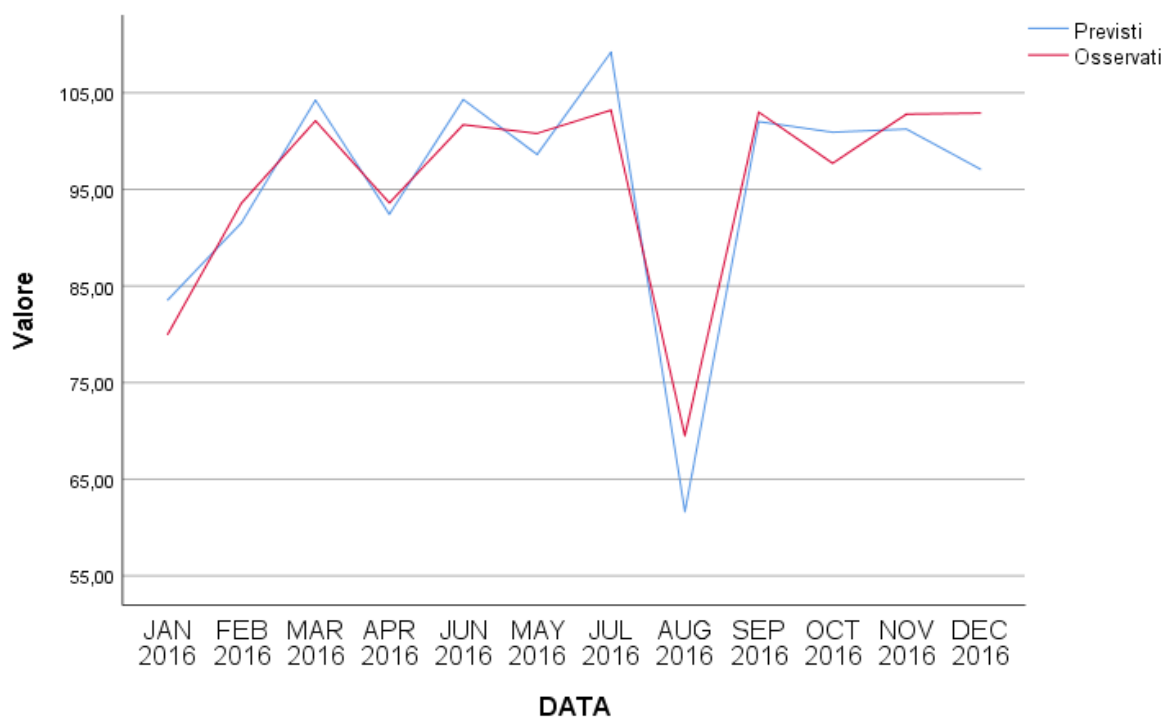


Tuttavia in alcuni momenti, come si può notare dal grafico, i valori stimati e osservati non coincidano; a volte vengono sovrastimati altre sottostimati.

Di seguito vengono riportati i dati stimati e i rispettivi limiti superiori e inferiori, ottenuti applicando un intervallo di confidenza del 95%.

DATA	Gen 2016	Feb 2016	Mar 2016	Apr 2016	Mag 2016	Giu 2016	Lug 2016	Ago 2016	Set 2016	Ott 2016	Nov 2016	Dic 2016
PREVISTI	83,57	91,56	104,24	92,43	98,62	104,3	109,2	61,64	102,01	100,91	101,24	97,1
LIM INF	75,52	83,52	96,2	84,39	90,57	96,26	101,16	53,6	93,97	92,87	93,2	89,06
LIM SUP	91,61	99,6	112,28	100,47	106,66	112,35	117,24	69,68	110,05	108,96	109,28	105,14

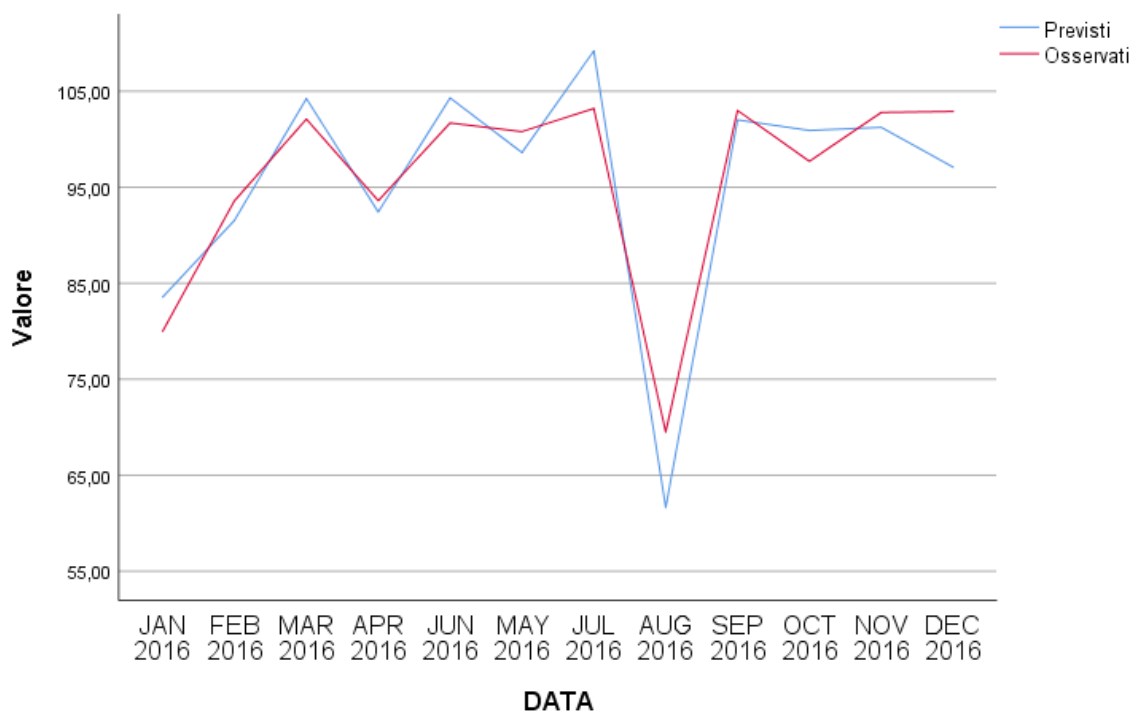
E se ne rappresentano le spezzate:



Successivamente si rappresentano in tabella il confronto fra valori osservati e previsti per l'anno 2016¹⁴.

2009	A_t	\hat{A}_t	$A_t - \hat{A}_t$	$(A_t - \hat{A}_t)/\hat{A}_t$
Gennaio	80,00	83,57	-3,57	-0,042718679
Febbraio	93,60	91,56	2,04	0,022280472
Marzo	102,10	104,24	-2,14	-0,020529547
Aprile	93,60	92,43	1,17	0,012658228
Maggio	100,80	98,62	2,18	0,02210505
Giugno	101,70	104,30	-2,60	-0,024928092
Luglio	103,20	109,20	-6,00	-0,054945055
Agosto	69,50	61,64	7,86	0,127514601
Settembre	103,00	102,01	0,99	0,009704931
Ottobre	97,70	100,91	-3,21	-0,031810524
Novembre	102,80	101,24	1,56	0,015408929
Dicembre	102,90	97,10	5,80	0,059732235

Questi sono più facilmente osservati con una rappresentazione grafica:



Il dislivello tra i dati osservati e previsti può essere dovuto a variazioni di mercato verificatesi nell'anno 2016, i modelli ARIMA per loro semplicità e natura non possono catturarli.

4. APPENDICE

1. I dati utilizzati sono i numeri indice relativi al volume del fatturato manifatturiero in Italia dal Gennaio 2002 al Dicembre 2016 prendendo come base (100) i dati relativi all'anno 2010.

I dati sono stati reperiti dal sito dell'ISTAT:

http://dati.istat.it/Index.aspx?DataSetCode=DCSC_ORDFATT

t	X_t	X_t-X_{t-1}	t	X_t	X_t-X_{t-1}
GEN 2002	99,70		LUG 2004	115,20	2,40
FEB 2002	104,30	4,60	AGO 2004	63,90	-51,30
MAR 2002	116,50	12,20	SET 2004	116,90	53,00
APR 2002	102,90	-13,60	OTT 2004	112,20	-4,70
MAG 2002	112,40	9,50	NOV 2004	108,80	-3,40
GIU 2002	110,90	-1,50	DIC 2004	109,60	0,80
LUG 2002	123,00	12,10	GEN 2005	87,60	-22,00
AGO 2002	62,50	-60,50	FEB 2005	99,30	11,70
SET 2002	115,50	53,00	MAR 2005	113,30	14,00
OTT 2002	119,10	3,60	APR 2005	101,90	-11,40
NOV 2002	110,20	-8,90	MAG 2005	108,60	6,70
DIC 2002	113,30	3,10	GIU 2005	111,60	3,00
GEN 2003	97,00	-16,30	LUG 2005	114,10	2,50
FEB 2003	101,20	4,20	AGO 2005	68,20	-45,90
MAR 2003	112,80	11,60	SET 2005	117,40	49,20
APR 2003	101,00	-11,80	OTT 2005	109,00	-8,40
MAG 2003	105,00	4,00	NOV 2005	111,60	2,60
GIU 2003	110,90	5,90	DIC 2005	111,00	-0,60
LUG 2003	120,10	9,20	GEN 2006	94,10	-16,90
AGO 2003	59,50	-60,60	FEB 2006	103,80	9,70
SET 2003	113,80	54,30	MAR 2006	123,80	20,00
OTT 2003	115,10	1,30	APR 2006	98,00	-25,80
NOV 2003	107,70	-7,40	MAG 2006	118,10	20,10
DIC 2003	108,80	1,10	GIU 2006	120,80	2,70
GEN 2004	90,30	-18,50	LUG 2006	116,50	-4,30
FEB 2004	103,10	12,80	AGO 2006	72,80	-43,70
MAR 2004	120,10	17,00	SET 2006	119,30	46,50
APR 2004	107,20	-12,90	OTT 2006	121,20	1,90
MAG 2004	105,70	-1,50	NOV 2006	117,90	-3,30
GIU 2004	112,80	7,10	DIC 2006	110,60	-7,30

t	X_t	X_t-X_{t-1}	t	X_t	X_t-X_{t-1}
GEN 2007	104,20	-6,40	MAG 2010	100,60	3,50
FEB 2007	109,40	5,20	GIU 2010	112,20	11,60
MAR 2007	126,70	17,30	LUG 2010	111,00	-1,20
APR 2007	105,40	-21,30	AGO 2010	66,90	-44,10
MAG 2007	123,60	18,20	SET 2010	111,00	44,10
GIU 2007	123,50	-0,10	OTT 2010	103,80	-7,20
LUG 2007	127,40	3,90	NOV 2010	105,90	2,10
AGO 2007	76,10	-51,30	DIC 2010	104,60	-1,30
SET 2007	117,30	41,20	GEN 2011	88,20	-16,40
OTT 2007	125,30	8,00	FEB 2011	98,90	10,70
NOV 2007	117,40	-7,90	MAR 2011	113,30	14,40
DIC 2007	105,10	-12,30	APR 2011	100,30	-13,00
GEN 2008	106,90	1,80	MAG 2011	106,80	6,50
FEB 2008	114,40	7,50	GIU 2011	111,50	4,70
MAR 2008	117,30	2,90	LUG 2011	110,90	-0,60
APR 2008	115,00	-2,30	AGO 2011	70,80	-40,10
MAG 2008	115,60	0,60	SET 2011	109,20	38,40
GIU 2008	123,00	7,40	OTT 2011	101,10	-8,10
LUG 2008	129,60	6,60	NOV 2011	102,70	1,60
AGO 2008	66,90	-62,70	DIC 2011	101,00	-1,70
SET 2008	120,90	54,00	GEN 2012	84,00	-17,00
OTT 2008	116,00	-4,90	FEB 2012	95,00	11,00
NOV 2008	101,80	-14,20	MAR 2012	106,40	11,40
DIC 2008	98,80	-3,00	APR 2012	92,00	-14,40
GEN 2009	81,90	-16,90	MAG 2012	103,90	11,90
FEB 2009	89,60	7,70	GIU 2012	102,70	-1,20
MAR 2009	100,30	10,70	LUG 2012	107,20	4,50
APR 2009	93,10	-7,20	AGO 2012	67,30	-39,90
MAG 2009	91,30	-1,80	SET 2012	97,50	30,20
GIU 2009	99,50	8,20	OTT 2012	101,80	4,30
LUG 2009	107,70	8,20	NOV 2012	96,90	-4,90
AGO 2009	59,20	-48,50	DIC 2012	92,00	-4,90
SET 2009	106,40	47,20	GEN 2013	84,20	-7,80
OTT 2009	98,70	-7,70	FEB 2013	86,80	2,60
NOV 2009	98,10	-0,60	MAR 2013	96,30	9,50
DIC 2009	99,80	1,70	APR 2013	89,40	-6,90
GEN 2010	83,50	-16,30	MAG 2013	98,10	8,70
FEB 2010	93,10	9,60	GIU 2013	98,10	0,00
MAR 2010	110,00	16,90	LUG 2013	107,00	8,90
APR 2010	97,10	-12,90	AGO 2013	63,20	-43,80

t	X_t	X_t-X_{t-1}	t	X_t	X_t-X_{t-1}
SET 2013	100,40	37,20	MAG 2015	95,30	0,90
OTT 2013	101,80	1,40	GIU 2015	105,80	10,50
NOV 2013	94,80	-7,00	LUG 2015	109,70	3,90
DIC 2013	94,70	-0,10	AGO 2015	62,20	-47,50
GEN 2014	84,60	-10,10	SET 2015	102,80	40,60
FEB 2014	88,60	4,00	OTT 2015	101,70	-1,10
MAR 2014	99,00	10,40	NOV 2015	98,90	-2,80
APR 2014	91,00	-8,00	DIC 2015	98,20	-0,70
MAG 2014	95,40	4,40	GEN 2016	80,00	-18,20
GIU 2014	97,90	2,50	FEB 2016	93,60	13,60
LUG 2014	106,20	8,30	MAR 2016	102,10	8,50
AGO 2014	60,20	-46,00	APR 2016	93,60	-8,50
SET 2014	101,80	41,60	MAG 2016	100,80	7,20
OTT 2014	101,50	-0,30	GIU 2016	101,70	0,90
NOV 2014	93,20	-8,30	LUG 2016	103,20	1,50
DIC 2014	96,70	3,50	AGO 2016	69,50	-33,70
GEN 2015	81,50	-15,20	SET 2016	103,00	33,50
FEB 2015	89,30	7,80	OTT 2016	97,70	-5,30
MAR 2015	103,50	14,20	NOV 2016	102,80	5,10
APR 2015	94,40	-9,10	DIC 2016	102,90	0,10

2. Il grafico rappresenta la spezzata che unisce i punti di coordinate $(x;t)$, dove x è il valore mensile osservato al tempo t .

3. Se fosse stato stazionario in media, i valori osservati si sarebbero aggirati intorno al valore medio nel tempo

4. Calcolare le differenze prime significa sostituire ai valori originari x_t le differenze date da $x_t - x_{t-1}$, ottenendo la serie storica di ordine

1. In questo modo si rende la serie stazionaria in media

5. Stazionarietà in media significa che la serie ha varianza costante.

6. Il numero max di ritardi è stabilito secondo la regola:

$$\text{MaxR} = N/4$$

Dove N è il numero di osservazioni.

Essendo $N = 168$, il numero max di ritardi è $168/4$, cioè 42

7. Le lettere minuscole contenute nella parentesi rappresentano il numero di parametri auto regressivi (p), l'ordine di differenziazione (d) e il numero di parametri media mobile (q).

Le lettere maiuscole rappresentano gli equivalenti parametri stagionali.

8. Alla serie originale applico un modello ARIMA (0, 0, 0) con lo scopo di analizzare i grafici delle autocorrelazioni dei residui.

9. Alla serie viene applicato un modello ARIMA (0, 1, 0), con il quale i dati vengono differenziati una volta.

10. I valori riportati in tabella hanno il seguente significato:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - \hat{y})^2}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N \frac{|y - \hat{y}|}{|y|}$$

Q = valore del test di Ljung-Box, applicato ai residui

p-value = relativo al test di Ljung-Box

11. I residui sono standardizzati in questo modo:

$$a_{\text{tsi}} = (a_{\text{ti}} - \mu) / \sigma^2$$

Dove :

a_{ti} = residui

μ = media dei residui

σ^2 = varianza dei residui

12. Il grafico rappresenta la spezzata dei punti di coordinate (x,t), dove x è il residuo standardizzato osservato al tempo t.

13. I processi stocastici “white noise” sono caratterizzati da media $\mu = 0$ (circa) e varianza σ^2 costante nel tempo. Tali caratteristiche fanno sì che i residui tra i valori osservati e quelli stimati siano non correlati, casuali, non anomali e normalmente distribuiti.

14. Di seguito sono riportati in tabella i valori osservati, quelli previsti e i relativi residui:

t	A_t	\hat{A}_t	$A_t - \hat{A}_t$
JAN 2002	99,70		99,70
FEB 2002	104,30		104,30
MAR 2002	116,50		116,50
APR 2002	102,90		102,90
MAY 2002	112,40		112,40
JUN 2002	110,90		110,90
JUL 2002	123,00		123,00
AUG 2002	62,50		62,50
SEP 2002	115,50		115,50
OCT 2002	119,10		119,10
NOV 2002	110,20		110,20
DEC 2002	113,30		113,30
JAN 2003	97,00		97,00
FEB 2003	101,20	101,6	-0,40
MAR 2003	112,80	113,61	-0,81
APR 2003	101,00	99,87	1,13
MAY 2003	105,00	109,41	-4,41
JUN 2003	110,90	106,86	4,04
JUL 2003	120,10	120,09	0,01
AUG 2003	59,50	57,98	1,52
SEP 2003	113,80	114,13	-0,33
OCT 2003	115,10	116,4	-1,30
NOV 2003	107,70	107,31	0,39
DEC 2003	108,80	110,9	-2,10
JAN 2004	90,30	93,17	-2,87
FEB 2004	103,10	97,14	5,96
MAR 2004	120,10	109,41	10,69
APR 2004	107,20	99,23	7,97
MAY 2004	105,70	110,62	-4,92
JUN 2004	112,80	115,17	-2,37
JUL 2004	115,20	124	-8,80
AUG 2004	63,90	58,79	5,11
SEP 2004	116,90	114,52	2,38
OCT 2004	112,20	114,89	-2,69
NOV 2004	108,80	109,98	-1,18
DEC 2004	109,60	110,38	-0,78
JAN 2005	87,60	90,56	-2,96

FEB 2005	99,30	101,17	-1,87
MAR 2005	113,30	115,08	-1,78
APR 2005	101,90	100,13	1,77
MAY 2005	108,60	103,16	5,44
JUN 2005	111,60	108,99	2,61
JUL 2005	114,10	116,32	-2,22
AUG 2005	68,20	62,73	5,47
SEP 2005	117,40	116,01	1,39
OCT 2005	109,00	114,07	-5,07
NOV 2005	111,60	110,87	0,73
DEC 2005	111,00	110,11	0,89
JAN 2006	94,10	88,57	5,53
FEB 2006	103,80	102,61	1,19
MAR 2006	123,80	116,16	7,64
APR 2006	98,00	106,95	-8,95
MAY 2006	118,10	111,27	6,83
JUN 2006	120,80	117,23	3,57
JUL 2006	116,50	118,58	-2,08
AUG 2006	72,80	73,04	-0,24
SEP 2006	119,30	122,94	-3,64
OCT 2006	121,20	114,43	6,77
NOV 2006	117,90	117,01	0,89
DEC 2006	110,60	116,24	-5,64
JAN 2007	104,20	100,19	4,01
FEB 2007	109,40	108,78	0,62
MAR 2007	126,70	123,91	2,79
APR 2007	105,40	108,78	-3,38
MAY 2007	123,60	119,83	3,77
JUN 2007	123,50	123,51	-0,01
JUL 2007	127,40	124,01	3,39
AUG 2007	76,10	77,87	-1,77
SEP 2007	117,30	125,57	-8,27
OCT 2007	125,30	125,13	0,17
NOV 2007	117,40	119,57	-2,17
DEC 2007	105,10	112,94	-7,84
JAN 2008	106,90	102,58	4,32
FEB 2008	114,40	107,92	6,48
MAR 2008	117,30	123,22	-5,92
APR 2008	115,00	108,74	6,26
MAY 2008	115,60	122,99	-7,39
JUN 2008	123,00	118,64	4,36
JUL 2008	129,60	129,25	0,35
AUG 2008	66,90	72,77	-5,87
SEP 2008	120,90	119,87	1,03

OCT 2008	116,00	124,07	-8,07
NOV 2008	101,80	112,3	-10,50
DEC 2008	98,80	105,66	-6,86
JAN 2009	81,90	93,82	-11,92
FEB 2009	89,60	95,93	-6,33
MAR 2009	100,30	104,12	-3,82
APR 2009	93,10	89,74	3,36
MAY 2009	91,30	95,95	-4,65
JUN 2009	99,50	100,88	-1,38
JUL 2009	107,70	106,23	1,47
AUG 2009	59,20	46,27	12,93
SEP 2009	106,40	102,2	4,20
OCT 2009	98,70	102,9	-4,20
NOV 2009	98,10	95,49	2,61
DEC 2009	99,80	90,67	9,13
JAN 2010	83,50	76,8	6,70
FEB 2010	93,10	90,21	2,89
MAR 2010	110,00	104,47	5,53
APR 2010	97,10	94,67	2,43
MAY 2010	100,60	99,55	1,05
JUN 2010	112,20	107,56	4,64
JUL 2010	111,00	114,09	-3,09
AUG 2010	66,90	62,86	4,04
SEP 2010	111,00	113,9	-2,90
OCT 2010	103,80	106,6	-2,80
NOV 2010	105,90	104,93	0,97
DEC 2010	104,60	102,43	2,17
JAN 2011	88,20	87,59	0,61
FEB 2011	98,90	98,71	0,19
MAR 2011	113,30	113,15	0,15
APR 2011	100,30	99,85	0,45
MAY 2011	106,80	105,7	1,10
JUN 2011	111,50	113,84	-2,34
JUL 2011	110,90	115,31	-4,41
AUG 2011	70,80	66,28	4,52
SEP 2011	109,20	112,29	-3,09
OCT 2011	101,10	106,76	-5,66
NOV 2011	102,70	106,59	-3,89
DEC 2011	101,00	100,28	0,72
JAN 2012	84,00	84,34	-0,34
FEB 2012	95,00	95,04	-0,04
MAR 2012	106,40	108,8	-2,40
APR 2012	92,00	94,73	-2,73
MAY 2012	103,90	100,8	3,10

JUN 2012	102,70	105,5	-2,80
JUL 2012	107,20	106,41	0,79
AUG 2012	67,30	62,81	4,49
SEP 2012	97,50	104,64	-7,14
OCT 2012	101,80	99,24	2,56
NOV 2012	96,90	99,8	-2,90
DEC 2012	92,00	92,77	-0,77
JAN 2013	84,20	80,8	3,40
FEB 2013	86,80	89,6	-2,80
MAR 2013	96,30	100,87	-4,57
APR 2013	89,40	88,83	0,57
MAY 2013	98,10	94,81	3,29
JUN 2013	98,10	96,95	1,15
JUL 2013	107,00	103,8	3,20
AUG 2013	63,20	59,72	3,48
SEP 2013	100,40	98,2	2,20
OCT 2013	101,80	101,21	0,59
NOV 2013	94,80	97,49	-2,69
DEC 2013	94,70	93,11	1,59
JAN 2014	84,60	82,56	2,04
FEB 2014	88,60	87,81	0,79
MAR 2014	99,00	100,75	-1,75
APR 2014	91,00	90,73	0,27
MAY 2014	95,40	97,95	-2,55
JUN 2014	97,90	99,27	-1,37
JUL 2014	106,20	105,7	0,50
AUG 2014	60,20	58,76	1,44
SEP 2014	101,80	99,71	2,09
OCT 2014	101,50	100,6	0,90
NOV 2014	93,20	95,5	-2,30
DEC 2014	96,70	94,25	2,45
JAN 2015	81,50	82,91	-1,41
FEB 2015	89,30	87,65	1,65
MAR 2015	103,50	101,1	2,40
APR 2015	94,40	90,43	3,97
MAY 2015	95,30	98,2	-2,90
JUN 2015	105,80	101,76	4,04
JUL 2015	109,70	108,98	0,72
AUG 2015	62,20	61,31	0,89
SEP 2015	102,80	106,35	-3,55
OCT 2015	101,70	103,53	-1,83
NOV 2015	98,90	96,33	2,57
DEC 2015	98,20	97,07	1,13
JAN 2016	80,00	83,57	-3,57

FEB 2016	93,60	91,56	2,04
MAR 2016	102,10	104,24	-2,14
APR 2016	93,60	92,43	1,17
MAY 2016	100,80	98,62	2,18
JUN 2016	101,70	104,3	-2,60
JUL 2016	103,20	109,2	-6,00
AUG 2016	69,50	61,64	7,86
SEP 2016	103,00	102,01	0,99
OCT 2016	97,70	100,91	-3,21
NOV 2016	102,80	101,24	1,56
DEC 2016	102,90	97,1	5,80