

Forecasting Crypto Data

Roger Garcia

Mississippi State University

rsg169@msstate.edu

April 17, 2019

- 1 About me
- 2 Motivation and Background
- 3 Importing Data
- 4 Data Exploration and Data Cleaning
- 5 Visualize
- 6 Modeling
- 7 Conclusion and Future Work

About Me



About Me



Ph.D Student (*Graduate Teacher Assistant*)

- B.S/M.S Science and Technology (*Comp. Math*) part of New Jersey Center for Science, Technology, and Mathematics at Kean University.



About Me



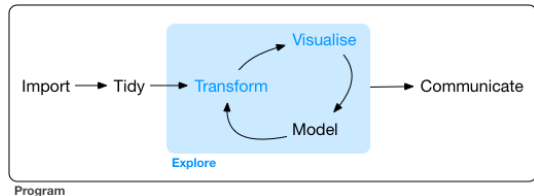
Ph.D Student (*Graduate Teacher Assistant*)

- B.S/M.S Science and Technology (*Comp. Math*) part of New Jersey Center for Science, Technology, and Mathematics at Kean University.

Career interests

- Artificial Intelligence
- Reinforcement Learning
- Cryptoeconomics
- Blockchain for Social Good
- Learning how to learn

Personal Website: rogersgarcia.github.io

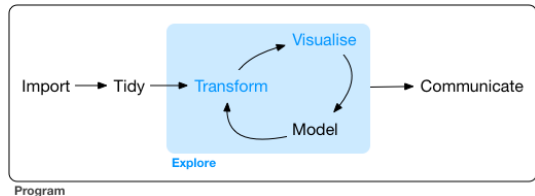


Project

- Find Dataset
- The Data Science Workflow

Motivation

- Learn about time series data
- R packages and functions for time series data



Project

- Find Dataset
- The Data Science Workflow

Motivation

- Learn about time series data
- R packages and functions for time series data
- Trading crypto:
100 to 0, real quick

Project Proposal and Motivation



Background

Time Series Dataset data that changes over time

Time Series Dataset data that changes over time

Autoregressive (AR) Model

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \dots + \beta_p y_{t-p}$$

Time Series Dataset data that changes over time

Autoregressive (AR) Model

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \dots + \beta_p y_{t-p}$$

Moving Average (MA) Model

$$y_t = \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots + \alpha_q \epsilon_{t-q}$$

Time Series Dataset data that changes over time

Autoregressive (AR) Model

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \dots + \beta_p y_{t-p}$$

Moving Average (MA) Model

$$y_t = \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots + \alpha_q \epsilon_{t-q}$$

Autoregressive Moving Average (ARMA) Model

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} \dots + \beta_p y_{t-p} \\ + \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots + \alpha_q \epsilon_{t-q}$$

Importing Data

Package 'crypto'

January 13, 2019

Type Package

Title Cryptocurrency Market Data

Description Retrieves crypto currency current and historical information as well as information on the exchanges they are listed on. For current and historical it will retrieve the daily open, high, low and close values for all crypto-currencies. This retrieves the historical market data by web-scraping tables provided by 'Cryptocurrency Market Capitalizations' <<https://coinmarketcap.com>>.

Version 1.1.1

Date 2019-01-13

Maintainer Jesse Vent <cryptopackage@icloud.com>

URL <https://github.com/JesseVent/crypto>,
<https://CRAN.R-project.org/package=crypto>

BugReports <https://github.com/JesseVent/crypto/issues>

Depends R (>= 3.4.0), rvest, xml2

License MIT + file LICENSE

Encoding UTF-8

LazyData yes

Imports dplyr, tidble, jsonlite, lubridate, xtr, curl, utils, cli,
crayon, progress, stats, tidy, ltr, rmarkdown

Suggests R6

SystemRequirements libxml2-devel, libcurl-devel, openssl-devel,
libsecret-devel, libodium-devel

RoxygenNote 6.1.1

NeedsCompilation no

Author Jesse Vent [aut, cre]

Repository CRAN

Date/Publication 2019-01-13 16:20:08 UTC

Package 'crypto'

January 13, 2019

Type Package

Title Cryptocurrency Market Data

Description Retrieves crypto currency current and historical information as well as information on the exchanges they are listed on. For current and historical it will retrieve the daily open, high, low and close values for all crypto-currencies. This retrieves the historical market data by web-scraping tables provided by Cryptocurrency Market Capitalizations <<https://coinmarketcap.com>>.

Version 1.1.1

Date 2019-01-13

Maintainer Jesse Vent <crypto@jessevent.com>

URL <https://github.com/JesseVent/crypto>,
<https://CRAN.R-project.org/package=crypto>

BugReports <https://github.com/JesseVent/crypto/issues>

Depends R (>= 3.4.0), rvest, xml2

License MIT + file LICENSE

Encoding UTF-8

LazyData no

Imports dplyr, tidble, jsonlite, lubridate, xtr, curl, utils, cli,
crayon, progress, stats, tidy, ltr, rmarkdown

Suggests R6

SystemRequirements libxml2-devel, libcurl-devel, openssl-devel,
libsecret-devel, libsofium-devel

RoxygenNote 6.1.1

NeedsCompilation no

Author Jesse Vent [aut, cre]

Repository CRAN

Date/Publication 2019-01-13 16:20:08 UTC

1

`install.packages("crypto")`

- Current and historical data
- CoinMarketCap API

Other packages

- `library("crypto")`
- `library("tidyverse")`
- `library("lubridate")`
- `library("tseries")`
- `library("forecast")`

Importing Dataset

```
> # get historic crypto currency market data (crypto_history() see crypto.pdf)
> # top ten coins. last request: Mon Apr 15 5pm
> topten_coins <- as_tibble(crypto_history(limit = 10))
♥ If this helps you become rich please consider donating
```

```
ERC-20: 0x375923BF82F0b728d23A5704261a6e16341fd860
XRP: rK59semLsuJZEWftxBFHwUeN6uhznjz2bK
```

```
> Scraping historical crypto data
```

```
- [6 / 10] [=====]-----] 60% in 00:00:38 ETA: 26s
?crypto
```

	\$data						
	id		name symbol	slug	circulating_supply	total_supply	max_supply
1	1		Bitcoin BTC	bitcoin	1.764740e+07	17647400	21000000
2	1027		Ethereum ETH	ethereum	1.056738e+08	105673782	NA
3	52		XRP XRP	ripple	4.197075e+10	99991649568	100000000000
4	1831		Bitcoin Cash BCH	bitcoin-cash	1.773036e+07	17730362	21000000
5	1765		EOS EOS	eos	9.062451e+08	1006245120	NA
6	2		Litecoin LTC	litecoin	6.135623e+07	61356234	84000000
7	1839		Binance Coin BNB	binance-coin	1.411755e+08	189175490	NA
8	825		Tether USDT	tether	2.458100e+09	2750057493	NA
9	512		Stellar XLM	stellar	1.930294e+10	104882366062	NA
10	2010		Cardano ADA	cardano	2.592707e+10	31112483745	45000000000
11	1958		TRON TRX	tron	6.668207e+10	99281283754	NA
12	328		Monero XMR	monero	1.692132e+07	16921323	NA
13	3602		Bitcoin SV BSV	bitcoin-sv	1.772832e+07	17728323	21000000
14	131		Dash DASH	dash	8.752614e+06	8752614	18900000
15	1720		IOTA MIOTA	iota	2.779530e+09	2779530283	2779530283
16	2011		Tezos XTZ	tezos	6.638304e+08	790789822	NA
17	1376		NEO NEO	neo	6.500000e+07	100000000	100000000
18	1321		Ethereum Classic ETC	ethereum-classic	1.097473e+08	109747278	210000000
19	2566		Ontology ONT	ontology	4.948544e+08	1000000000	NA
20	1518		Maker MKR	maker	1.000000e+06	1000000	NA
21	873		NEM XEM	nem	9.000000e+09	8999999999	NA
22	1437		Zcash ZEC	zcash	6.324619e+06	6324619	NA
23	3635		Crypto.com Chain CRO	crypto-com-chain	4.997717e+09	100000000000	NA
24	3077		VeChain VET	vechain	5.545473e+10	86712634466	NA
25	1697		Basic Attention Token BAT	basic-attention-token	1.249417e+09	1500000000	NA
26	74		Dogecoin DOGE	dogecoin	1.190999e+11	119099934457	NA
27	2083		Bitcoin Gold BTG	bitcoin-gold	1.751392e+07	17513924	21000000
28	1274		Waves WAVES	waves	1.000000e+08	100000000	NA
29	3408		USD Coin USDC	usd-coin	2.593593e+08	259927862	NA
30	1808		OmiseGo OMG	omiseGO	1.402454e+08	140245398	NA
31	1684		Qtum QTUM	qtum	8.939982e+07	101399816	NA
32	1168		Decred DCR	decred	9.631155e+06	9631155	21000000

Data Exploration and Data Cleaning

\$data	id	name	symbol	slug	circulating_supply	total_supply	max_supply
1	1	Bitcoin	BTC	bitcoin	1.764740e+07	17647400	21000000
2	1027	Ethereum	ETH	ethereum	1.056738e+08	105673782	NA
3	52	XRP	XRP	ripple	4.197075e+10	99991649568	100000000000
4	1831	Bitcoin Cash	BCH	bitcoin-cash	1.773036e+07	17730362	21000000
5	1765	EOS	EOS	eos	9.062451e+08	1006245120	NA
6	2	Litecoin	LTC	litecoin	6.135623e+07	61356234	84000000
7	1839	Binance Coin	BNB	binance-coin	1.411755e+08	189175490	NA
8	825	Tether	USDT	tether	2.458100e+09	2750057493	NA
9	512	Stellar	XLM	stellar	1.930294e+10	104882366062	NA
10	2010	Cardano	ADA	cardano	2.592707e+10	31112483745	45000000000
11	1958	TRON	TRX	tron	6.668207e+10	99281283754	NA
12	328	Monero	XMR	monero	1.692132e+07	16921323	NA
13	3602	Bitcoin SV	BSV	bitcoin-sv	1.772832e+07	17728323	21000000
14	131	Dash	DASH	dash	8.752614e+06	8752614	18900000
15	1720	IOTA	MIOTA	iota	2.779530e+09	2779530283	2779530283
16	2011	Tezos	XTZ	tezos	6.638304e+08	790789822	NA
17	1376	NEO	NEO	neo	6.500000e+07	100000000	100000000
18	1321	Ethereum Classic	ETC	ethereum-classic	1.097473e+08	109747278	210000000
19	2566	Ontology	ONT	ontology	4.948544e+08	1000000000	NA
20	1518	Maker	MKR	maker	1.000000e+06	1000000	NA
21	873	NEM	XEM	nem	9.000000e+09	8999999999	NA
22	1437	Zcash	ZEC	zcash	6.324619e+06	6324619	NA
23	3635	Crypto.com Chain	CRO	crypto-com-chain	4.997717e+09	100000000000	NA
24	3077	VeChain	VET	vechain	5.545473e+10	86712634466	NA
25	1697	Basic Attention Token	BAT	basic-attention-token	1.249417e+09	1500000000	NA
26	74	DogeCoin	DOGE	dogecoin	1.190999e+11	119099934457	NA
27	2083	Bitcoin Gold	BTG	bitcoin-gold	1.751392e+07	17513924	21000000
28	1274	Waves	WAVES	waves	1.000000e+08	100000000	NA
29	3408	USD Coin	USDC	usd-coin	2.593593e+08	259927862	NA
30	1808	OmiseGO	OMG	omiseGO	1.402454e+08	140245398	NA
31	1684	Qtum	QTUM	qtum	8.939982e+07	101399816	NA
32	1168	Decred	DCR	decred	9.631155e+06	9631155	21000000

Data

- data.frame
- Coin symbol and name
- Market date
- Current Rank
- OHLC
- Volume 24 hours
- USD Marketcap
- Close rate
- spread

Data Cleaning

Sdata	id	name	symbol	slug	circulating_supply	total_supply	max_supply
1	1	Bitcoin	BTC	bitcoin	1.764740e+07	17647400	21000000
2	1027	Ethereum	ETH	ethereum	1.056730e+08	105673782	NA
3	52	XRP	XRP	ripple	4.197075e+10	99991649568	100000000000
4	1831	Bitcoin Cash	BCH	bitcoin-cash	1.773036e+07	17730362	21000000
5	1765	EOS	EOS	eos	9.062451e+08	1006245120	NA
6	2	Litecoin	LTC	litecoin	6.135623e+07	61356234	84000000
7	1839	Binance Coin	BNB	binance-coin	1.411755e+08	189175490	NA
8	825	Tether	USDT	tether	2.458100e+09	2750057493	NA
9	512	Stellar	XLM	stellar	1.930294e+10	104882366062	NA
10	2010	Cardano	ADA	cardano	2.592707e+10	31112483745	45000000000
11	1958	TRON	TRX	tron	6.668207e+10	99281283754	NA
12	328	Monero	XMR	monero	1.692132e+07	16921323	NA
13	3602	Bitcoin SV	BSV	bitcoin-sv	1.772832e+07	17728323	21000000
14	131	Dash	DASH	dash	8.752614e+06	8752614	18900000
15	1720	IOTA	MIOTA	iota	2.779530e+09	2779530283	2779530283
16	2011	Tezos	XTZ	tezos	6.638304e+08	790789822	NA
17	1376	NEO	NEO	neo	6.500000e+07	100000000	100000000
18	1321	Ethereum Classic	ETC	ethereum-classic	1.097473e+08	109747278	210000000
19	2566	Ontology	ONT	ontology	4.948544e+08	1000000000	NA
20	1518	Maker	MKR	naker	1.000000e+06	1000000	NA
21	873	NEM	XEM	nen	9.000000e+09	8999999999	NA
22	1437	Zcash	ZEC	zcash	6.324019e+06	6324019	NA
23	3635	Crypto.com Chain	CRO	crypto-com-chain	4.997717e+09	100000000000	NA
24	3077	VeChain	VET	vechain	5.545473e+10	86712634466	NA
25	1697	Basic Attention Token	BAT	basic-attention-token	1.249417e+09	1500000000	NA
26	74	Dogecoin	DOGE	dogecoin	1.190999e+11	119099934457	NA
27	2083	Bitcoin Gold	BTG	bitcoin-gold	1.751392e+07	17513924	21000000
28	1274	Waves	WAVES	waves	1.000000e+08	100000000	NA
29	3408	USD Coin	USDC	usd-coin	2.593593e+08	259927862	NA
30	1808	OmiseGO	OMG	omiseGO	1.402454e+08	140245398	NA
31	1684	Qtum	QTUM	qtrum	8.939982e+07	101399816	NA
32	1168	Decred	DCR	decred	9.631155e+06	9631155	21000000

Data Cleaning

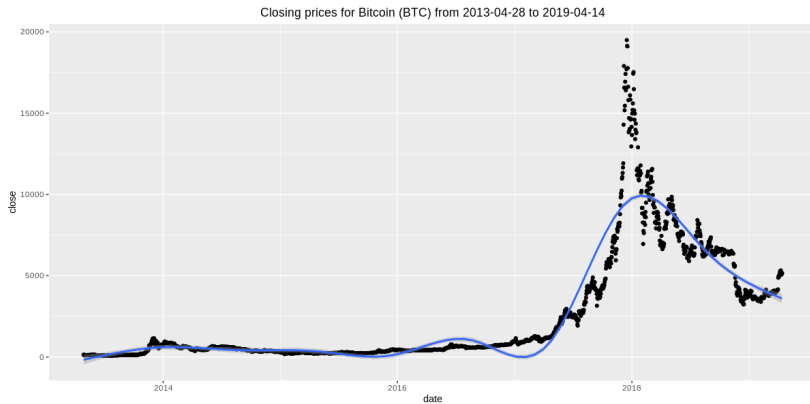
Sdata						
id	name	symbol	slug	circulating_supply	total_supply	max_supply
1	Bitcoin	BTC	bitcoin	1.764740e+07	17647400	21000000
2	Ethereum	ETH	ethereum	1.056730e+08	105673782	NA
3	XRP	XRP	ripple	4.197075e+10	99991649568	100000000000
4	Bitcoin Cash	BCH	bitcoin-cash	1.773036e+07	17730362	21000000
5	EOS	EOS	eos	9.062451e+08	1006245120	NA
6	Litecoin	LTC	litecoin	6.135623e+07	61356234	84000000
7	Binance Coin	BNB	binance-coin	1.411755e+08	189175490	NA
8	Tether	USDT	tether	2.458100e+09	2750057493	NA
9	Stellar	XLM	stellar	1.930294e+10	104882366062	NA
10	Cardano	ADA	cardano	2.592707e+10	31112483745	45000000000
11	TRON	TRX	tron	6.668207e+10	99281283754	NA
12	Monero	XMR	monero	1.692132e+07	16921323	NA
13	Bitcoin SV	BSV	bitcoin-sv	1.772832e+07	17728323	21000000
14	Dash	DASH	dash	8.752614e+06	8752614	18900000
15	IOTA	MIOTA	iota	2.779530e+09	2779530283	2779530283
16	Tezos	XTZ	tezos	6.638304e+08	790789822	NA
17	NEO	NEO	neo	6.500000e+07	100000000	100000000
18	Ethereum Classic	ETC	ethereum-classic	1.097473e+08	109747278	210000000
19	Ontology	ONT	ontology	4.948544e+08	1000000000	NA
20	Maker	MKR	maker	1.000000e+06	1000000	NA
21	NEM	XEM	nem	9.000000e+09	8999999999	NA
22	Zcash	ZEC	zcash	6.324019e+06	6324019	NA
23	Crypto.com Chain	CRO	crypto-com-chain	4.997717e+09	100000000000	NA
24	VeChain	VET	vechain	5.545473e+10	86712634466	NA
25	Basic Attention Token	BAT	basic-attention-token	1.249417e+09	1500000000	NA
26	Dogecoin	DOGE	dogecoin	1.190999e+11	119099934457	NA
27	Bitcoin Gold	BTG	bitcoin-gold	1.751392e+07	17513924	21000000
28	Waves	WAVES	waves	1.000000e+08	100000000	NA
29	USD Coin	USDC	usd-coin	2.593593e+08	259927862	NA
30	OmiseGO	OMG	omiseGO	1.402454e+08	140245398	NA
31	Qtum	QTUM	qtum	8.939982e+07	101399816	NA
32	Decred	DCR	decred	9.631155e+06	9631155	21000000

Closing Price

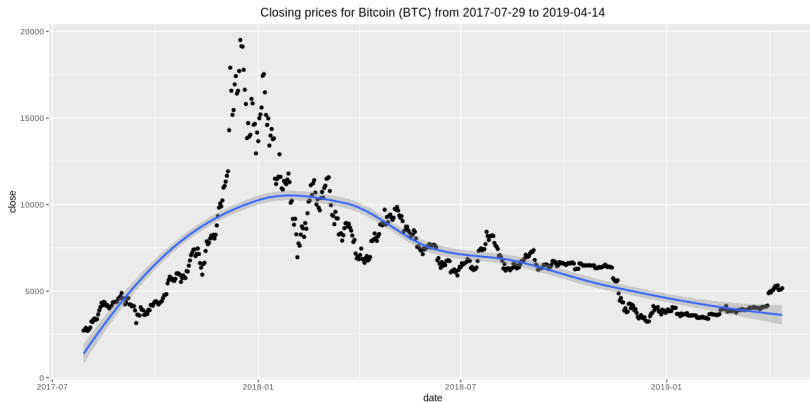
- end of trading day
- performance over time frame
- make decision

Visualize

Visualizing Data

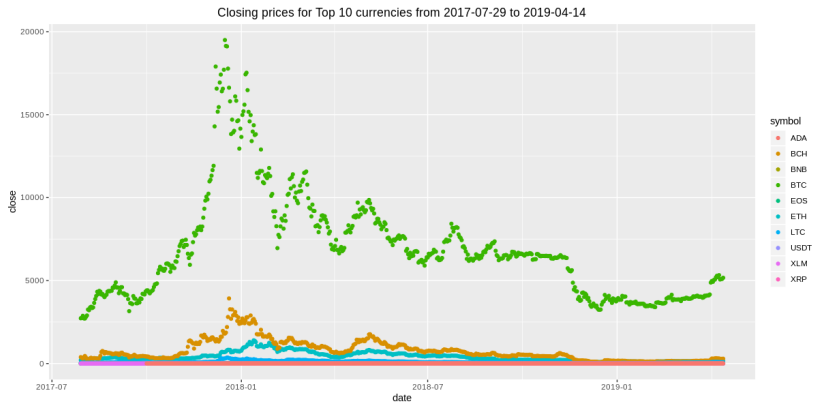


Visualizing Data

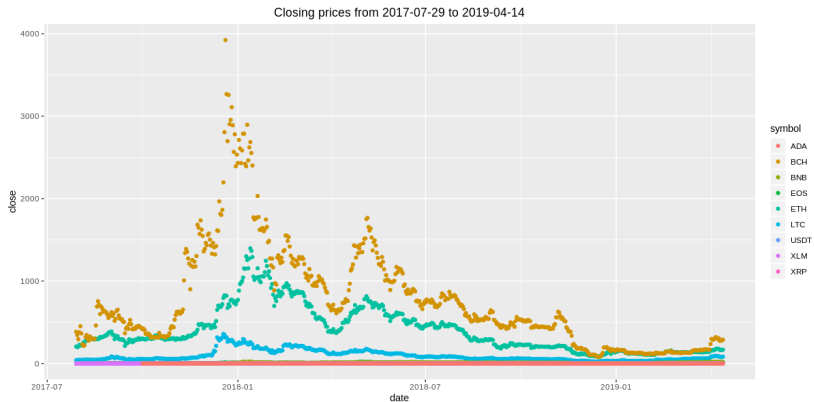


Other Cryptocurrencies (Altcoins)?

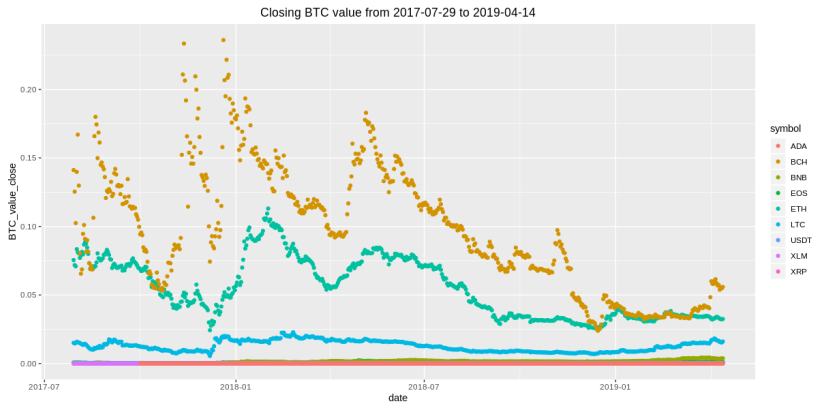
Visualizing Data



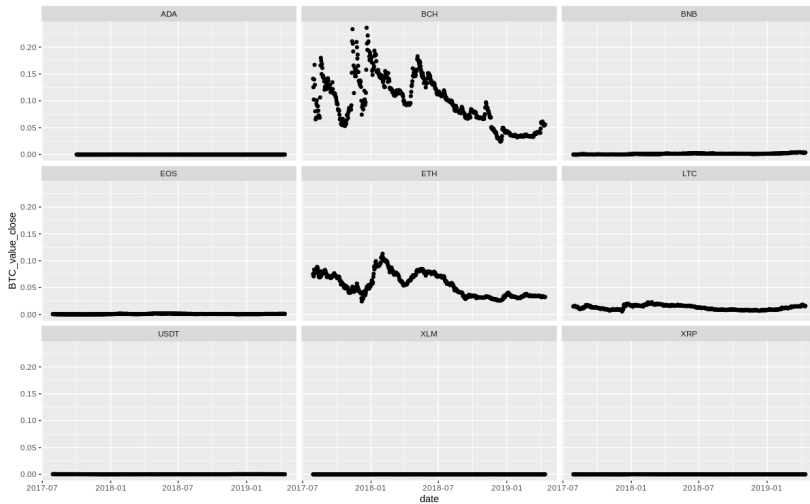
Visualizing Data



Visualizing Data



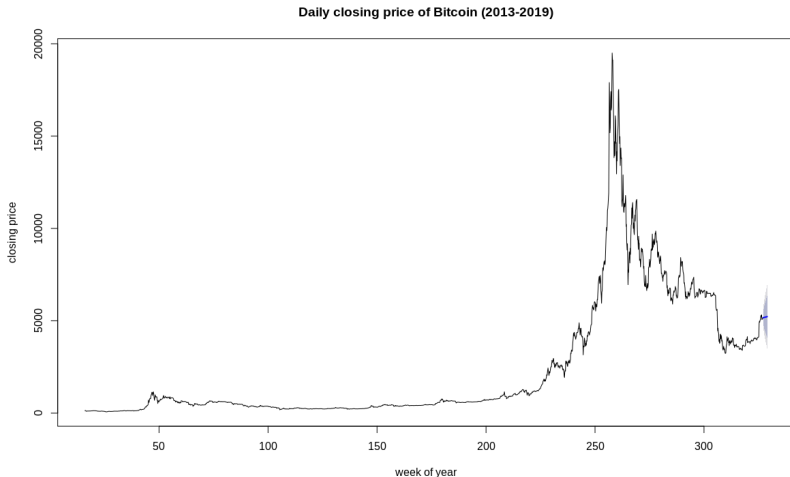
Visualizing Data



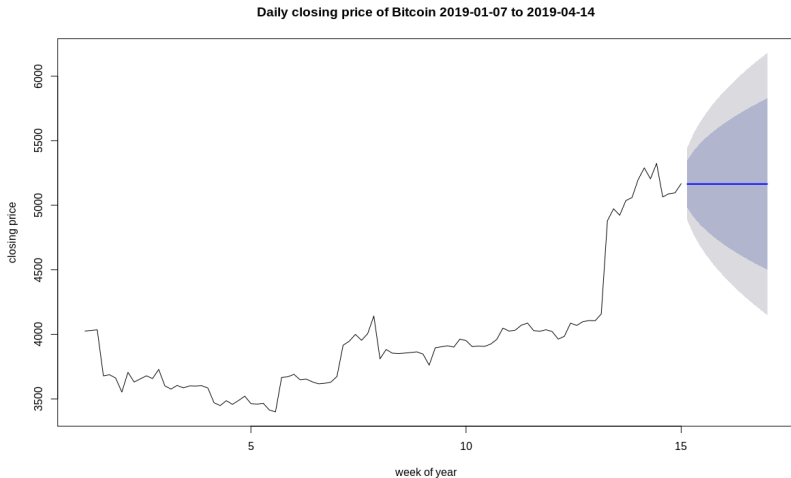
Model

Exponential Smoothing

Forecasting method, but model explicitly uses exponentially decreasing weight for past observations



Exponential Smoothing




```
Series: btc_close_ts_freq1  
ARIMA(1,1,1)
```

```
Coefficients:
```

	ar1	ma1
	-0.6626	0.7394
s.e.	0.0805	0.0716

```
sigma^2 estimated as 52540:  
log likelihood=-14919.3  
AIC=29844.6    AICc=29844.61  
BIC=29861.65
```

Modeling

- `auto.arima()`
- Stationary Data Assumption
- *is differencing required?*
- Forecast from the ARIMA model
- Visualize forecast (prediction)

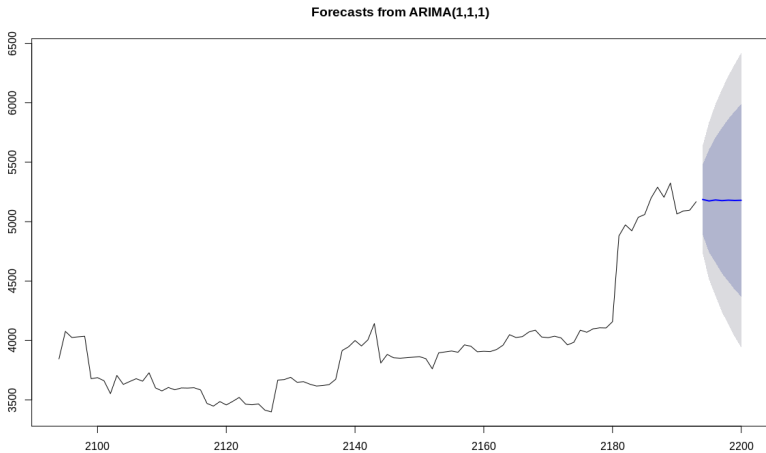
Autoregressive Moving Average (ARMA) Model

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} \dots + \beta_p y_{t-p} \\ + \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots + \alpha_q \epsilon_{t-q}$$

```
ar1  
-0.6625775  
ma1  
0.7394288
```

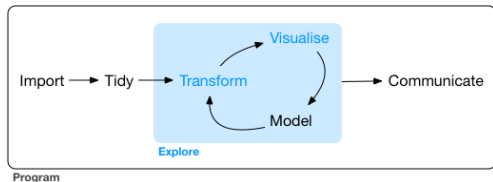
```
$pred  
Time Series :  
Start = 2194  
End = 2200  
Frequency = 1  
[1] 5186.526 5174.066 5182.322 5176.852  
      5180.476 5178.075 5179.666  
$se  
Time Series :  
Start = 2194  
End = 2200  
Frequency = 1  
[1] 229.2152 336.8457 410.8100 477.2437  
      533.1936 585.1968 632.0880
```

ARMA Model



Conclusion and Future Work

Conclusion



What I learned

- *Data Science Workflow*
- *no straightforward process*
- *underlying structure and functions*

Presentation Shortcoming → (Report)

- highercharter
- stationary vs non stationary model
- ACF and PACF for model residuals

Future work

- self-reflection on trade decisions
- continue working with financial time series data in R

Thoughts or suggestions
moving forward ?

Thank you for your time