PREDICTING UFC PAY PER VIEW BUYS

JENNIFER

Quick Background

- Ultimate Fighting Championship founded in 1993
- Pay Per View Buys make up 30% of UFC’s revenue
- Today PPV costs are between $49.99 to $59.99

Data collection

UFC 214: Cormier vs. Jones 2

Rematch

DATE: July 29, 2017 **LOCATION:** Anaheim, California, USA **ATTENDANCE:** 16,610

Click on a row below to see in-depth event stats. **Fight, Perf, Sub, and KO of the Night Bonuses:** *FIGHT *PERF *SUB *KO

Rivalry

W/L	FIGHTER	STR	TD	SUB	PASS	WEIGHT CLASS	METHOD	ROUND	TIME
NC NC	Daniel Cormier Jon Jones	58 95	0 0	0 0	0 0	Light Heavyweight *PERF	Overturned Kick	3	3:01
WIN	Tyron Woodley Demian Maia	57 28	0 0	0 0	0 0	Welterweight	U-DEC	5	5:00
WIN	Cris Cyborg Tonya Evinger	74 32	0 1	0 0	0 0	Women's Featherweight	KO/TKO Knees	3	1:56
WIN	Robbie Lawler Donald Cerrone	77 82	0 1				U-DEC	3	5:00
WIN	Volkan Oezdemir Jimi Manuwa	12 2	0 0	0 0	0 0	Light Heavyweight *PERF	KO/TKO Punches	1	0:42
WIN	Ricardo Lamas Jason Knight	42 8	1 0	0 0	2 0	Featherweight	KO/TKO Punches	1	4:34
WIN	Aljamain Sterling Renan Barao	71 19	1 2	0 0	1 3	Catch Weight	U-DEC	3	5:00
WIN	Brian Ortega Renato Moicano	65 109	0 2	1 0	0 0	Featherweight *FIGHT	SUB Guillotine Choke	3	2:59

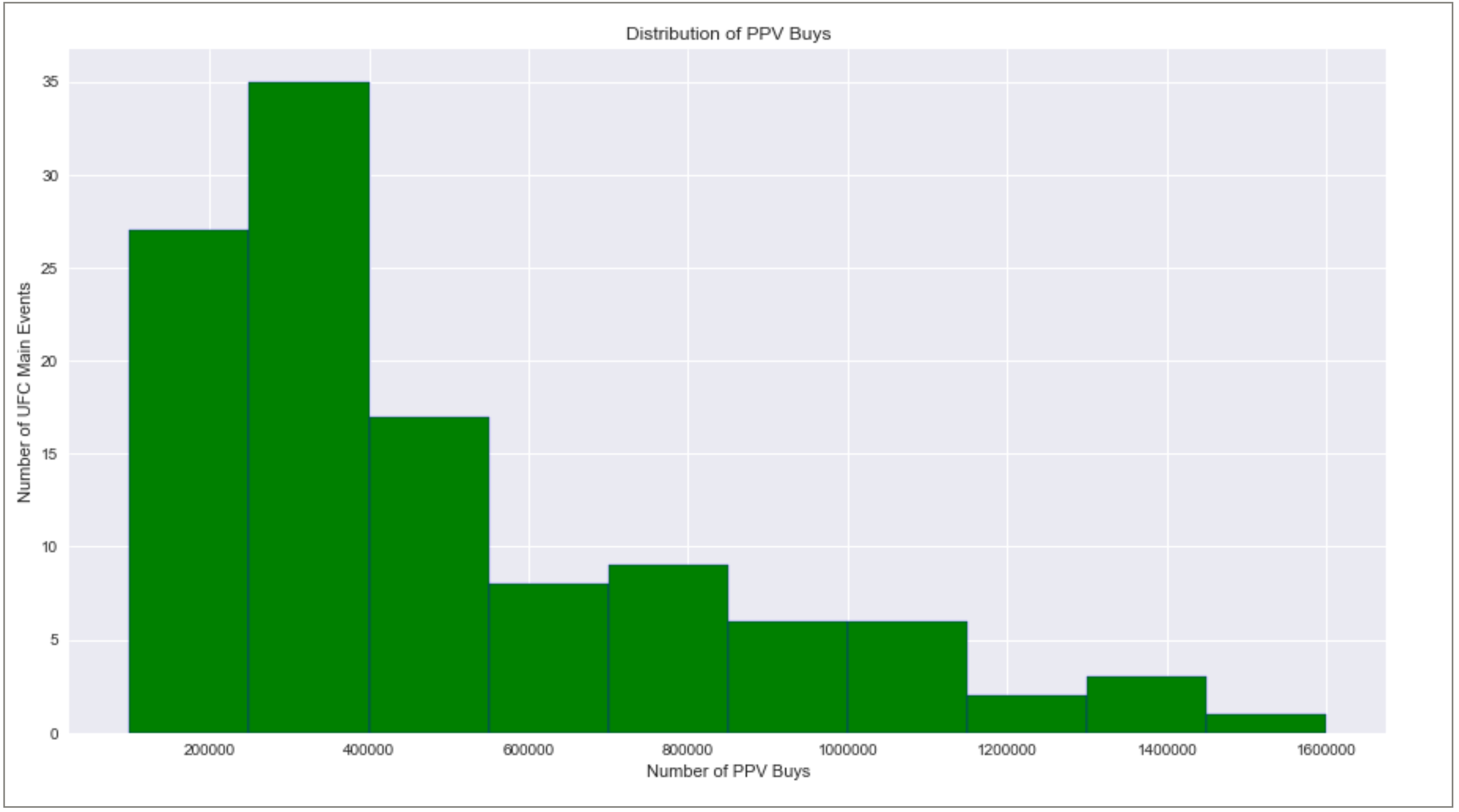
Belt

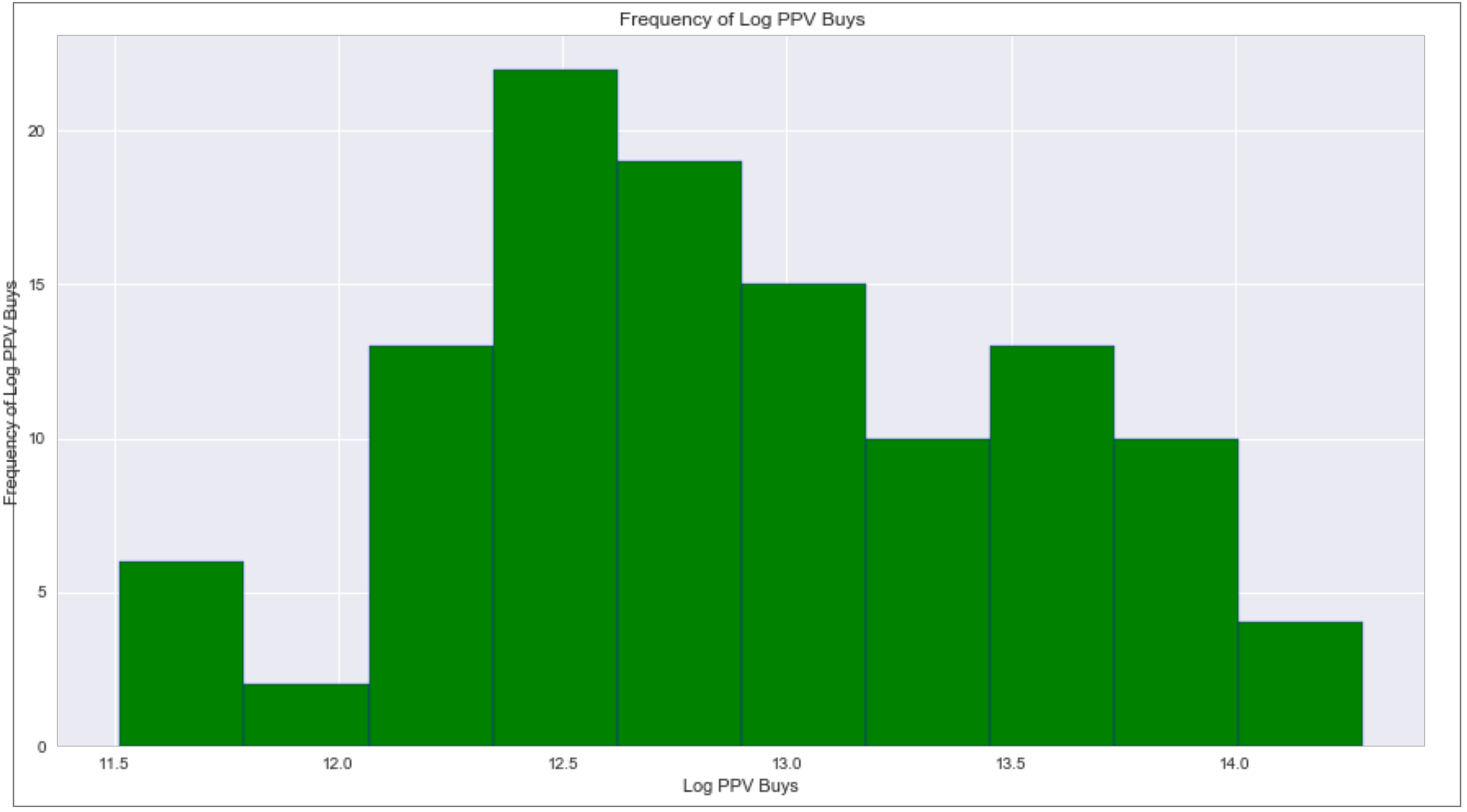
Women's fight

Source: FightMetric

Explored Features

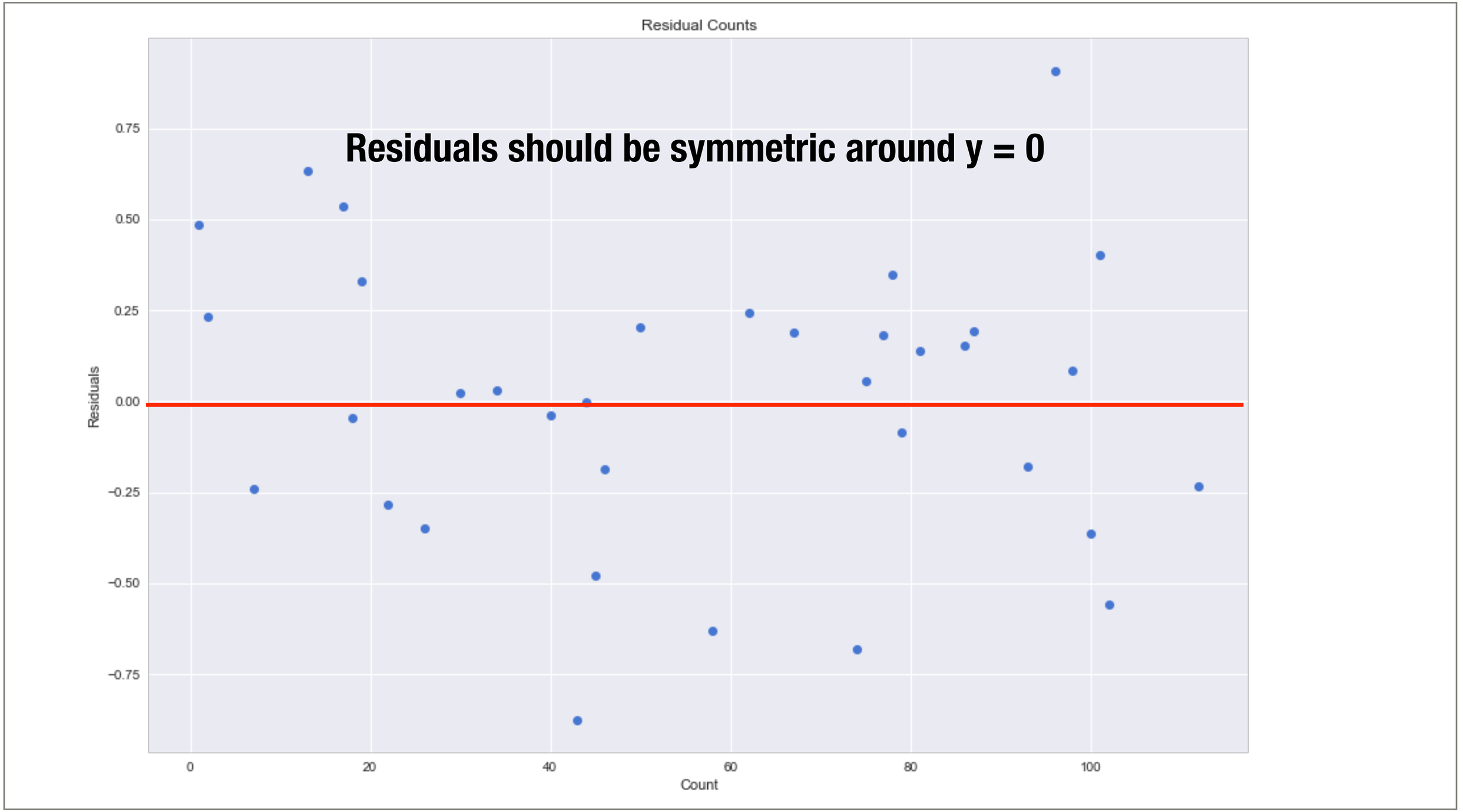
- Ticket Revenue (Live Gate)
- Attendance
- TV Ratings (viewers)
- Personal Rivalry
- Total # of Strikes, Submissions, Passes
- Rematch
- Number of Fights Per Event (Bouts)
- Women Fights
- Title Shot
- Technical Fighters
- Number of Rounds
- Fight Time
- Fan Favorites - McGregor, Silva, Rousey, Liddell, Lesnar, GSP
- Finishes : KO, Sub., U-Dec
- Lightweight vs Heavyweight

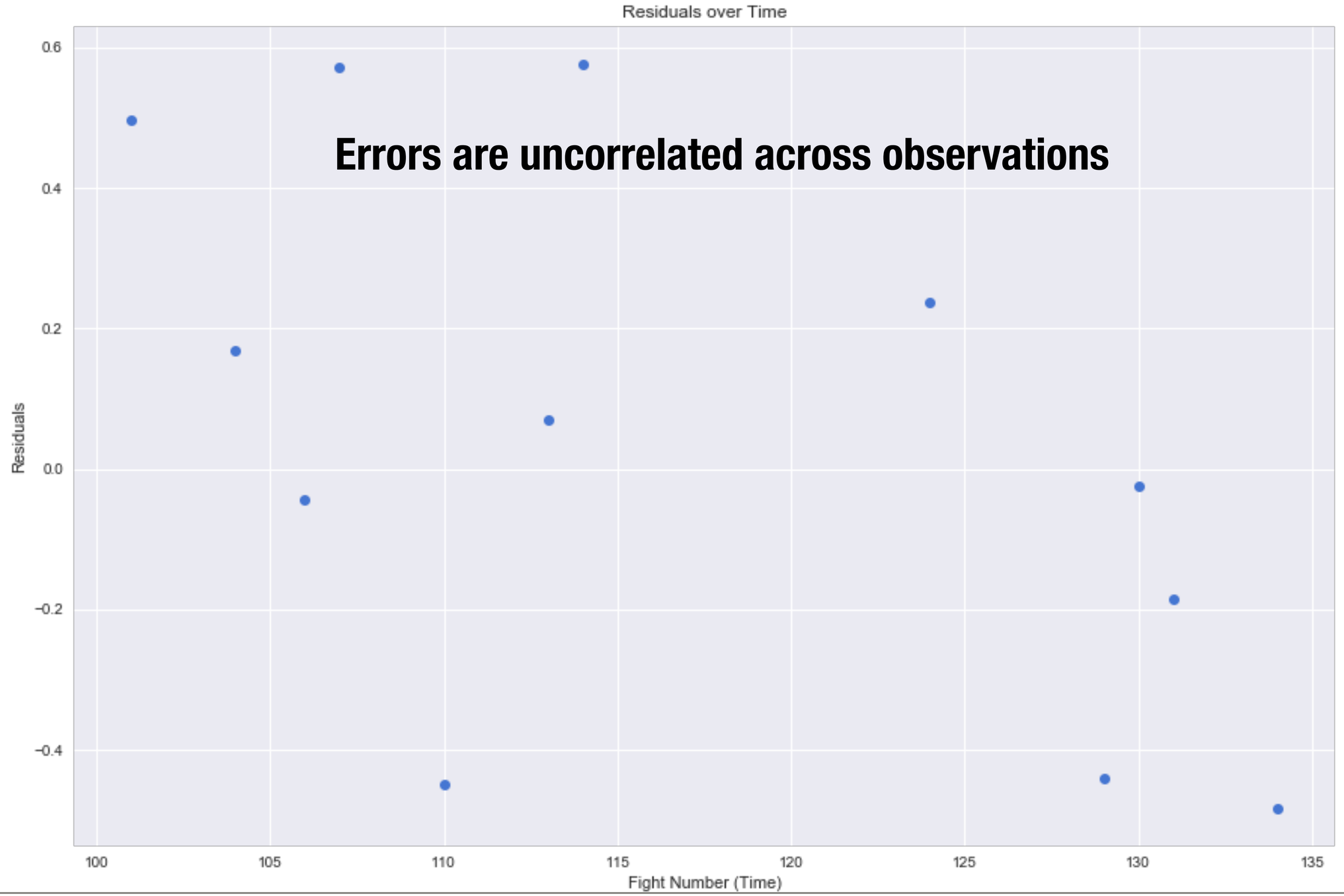

Distribution of PPV Buys
Number of UFC Main Events
0
5
10
15
20
25
30
35
200000
400000
600000
800000
1000000
1200000
1400000
1600000
Number of PPV Buys


Frequency of Log PPV Buys
Frequency of Log PPV Buys
0
5
10
15
20
11.5
12.0
12.5
13.0
13.5
14.0
Log PPV Buys

Linear Assumptions


Actual vs Predicted PPV Buys
Regression is linear in parameters
Predicted PPV Buys
14.5
14.0
13.5
13.0
12.5
12.0
12.5
13.0
13.5
14.0
Actual PPV Buys


Residual Counts
Residuals should be symmetric around y = 0
Residuals
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
0
20
40
60
80
100
Count


Residuals over Time
Errors are uncorrelated across observations
Residuals
0.6
0.4
0.2
0.0
−0.2
−0.4
100
105
110
115
120
125
130
135
Fight Number (Time)

Regression Modeling

Correlations

log_ppv	1.000000
ppv	0.945756
fan_fav	0.684376
tvrate	0.636035
ticketrev	0.596661
heavyfights	0.410474
rival	0.263060
belts	0.206176
attend	0.203378
tech_star	0.121231
men	0.120958
KO	0.115108
Other	0.114884
bouts	0.071050
submission	0.046193
takedowns	0.036711
DEC	-0.028123
rounds	-0.037491
tottime	-0.048620
rematch	-0.050113
passes	-0.056578
women	-0.082990
SUB	-0.103044
strikes	-0.115967
fightnum	-0.191897
lightfights	-0.352749

Dep. Variable:	log_ppv	**R-squared:**	0.745
Model:	OLS	**Adj. R-squared:**	0.720
Method:	Least Squares	**F-statistic:**	30.08
Date:	Wed, 31 Jan 2018	**Prob (F-statistic):**	3.13e-26
Time:	20:45:00	**Log-Likelihood:**	-31.470
No. Observations:	114	**AIC:**	84.94
Df Residuals:	103	**BIC:**	115.0
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	11.1607	0.426	26.178	0.000	10.315	12.006
attend	-6.201e-07	5.54e-06	-0.112	0.911	-1.16e-05	1.04e-05
ticketrev	5.79e-08	1.99e-08	2.909	0.004	1.84e-08	9.74e-08
tvrate	6.324e-07	1.14e-07	5.554	0.000	4.07e-07	8.58e-07
lightfights	0.0181	0.039	0.458	0.648	-0.060	0.096
heavyfights	0.0980	0.041	2.416	0.017	0.018	0.178
women	0.0131	0.044	0.300	0.765	-0.073	0.100
belts	0.1053	0.054	1.958	0.053	-0.001	0.212
tech_star	-0.0483	0.022	-2.238	0.027	-0.091	-0.006
rival	0.1959	0.068	2.878	0.005	0.061	0.331
fan_fav	0.3577	0.066	5.444	0.000	0.227	0.488

Omnibus:	1.515	**Durbin-Watson:**	2.150
Prob(Omnibus):	0.469	**Jarque-Bera (JB):**	1.462
Skew:	-0.175	**Prob(JB):**	0.481
Kurtosis:	2.569	**Cond. No.**	5.71e+07

Train Test Split Test

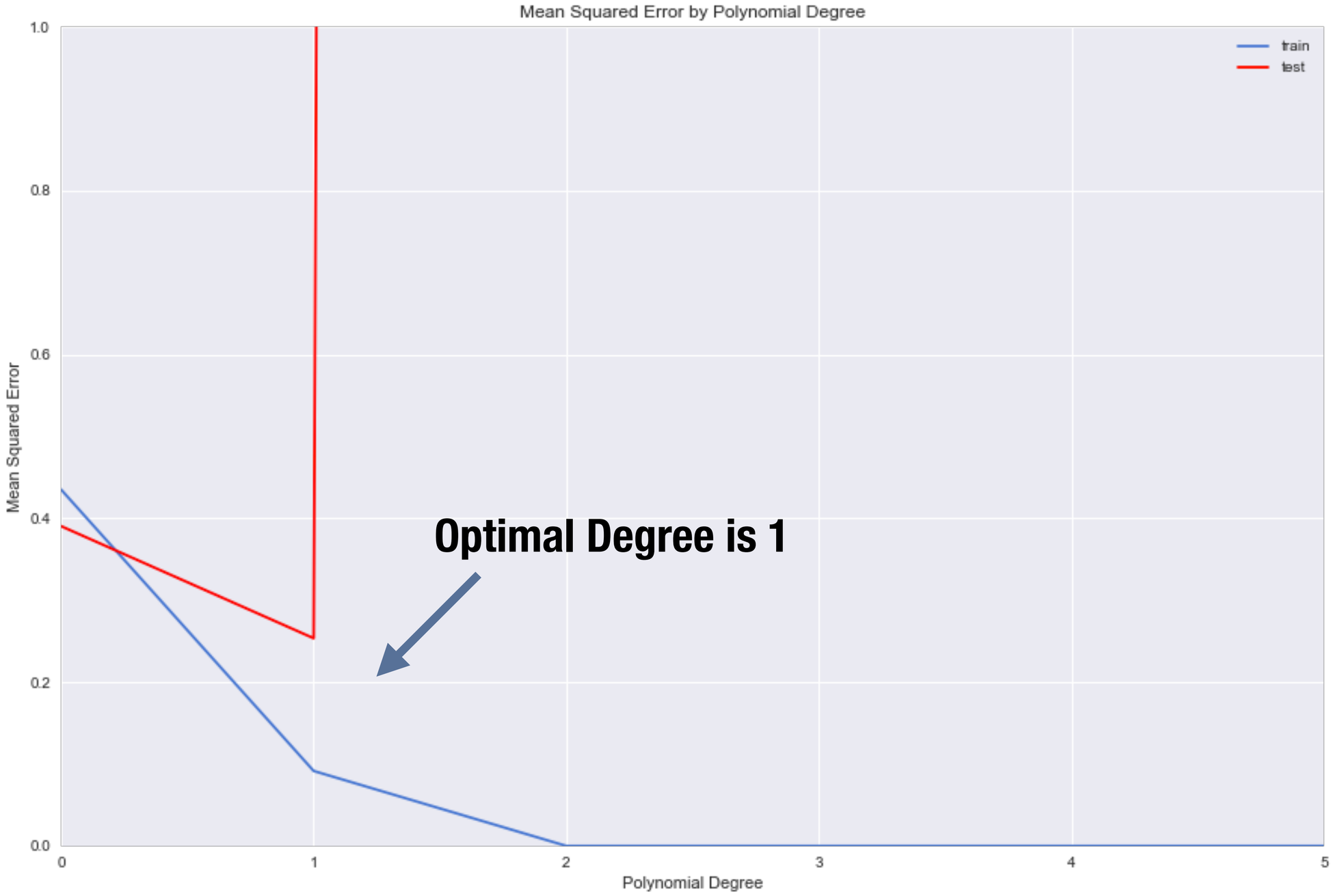
- Test model against an out of sample test set
- 30% data will designated as test set
- R^2 = 0.6608839390946249 5
- MSE for Train Set = 0.09073735001478117 1
- MAE for Train Set = 0.24997090345011294
- MSE for Test Set = 0.11580326376199075
- MAE for Test Set = 0.26870784486413135

K-Fold Regression

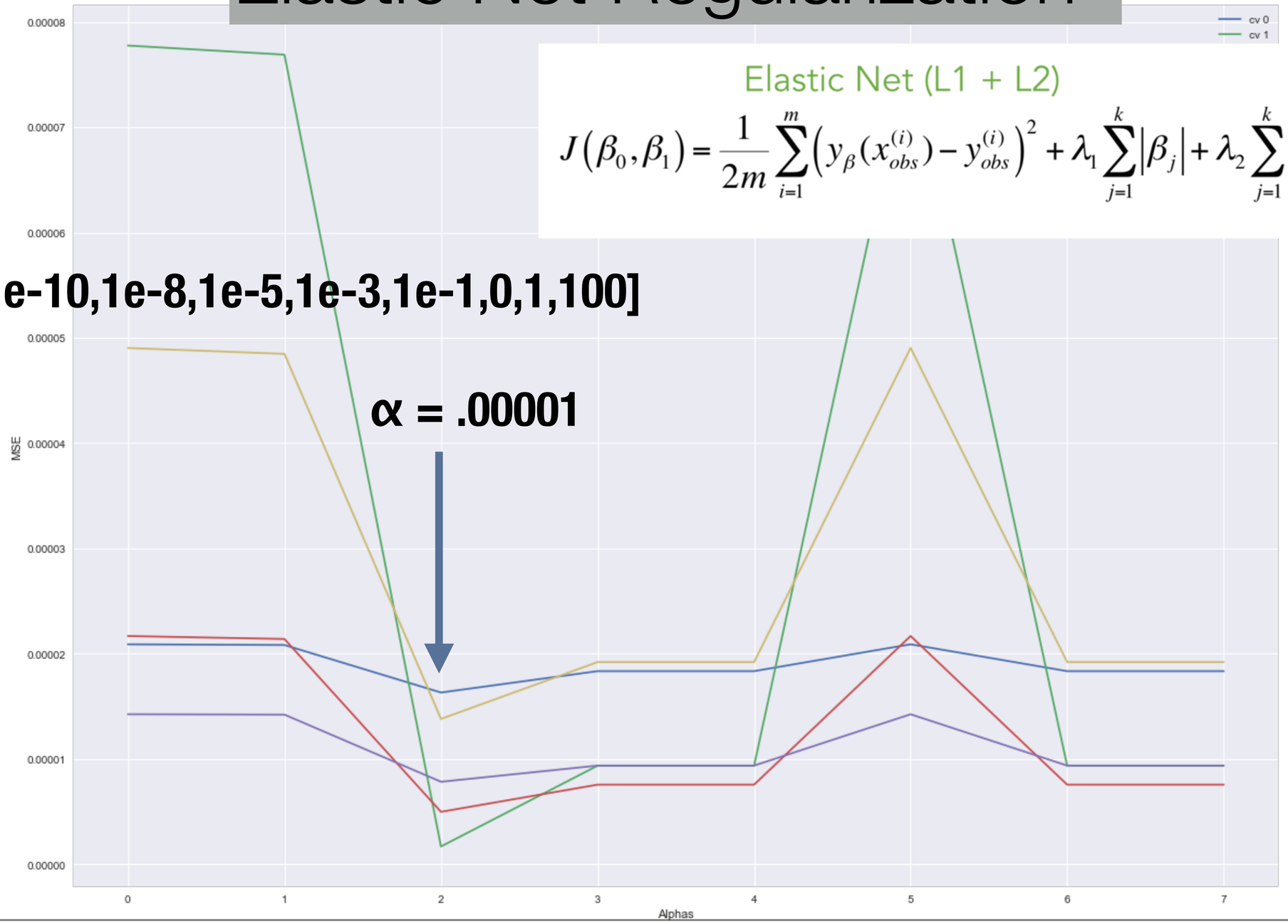
- Split data into 10 folds, each fold will have a chance to be tested against
- Identify most statistical significant features with p-value < 2%
- Mean R^2 = 0.6464631107004711 3

Significant Features	Coefficients
ticketrev	4.42E-08
tvrate	5.95E-07
lightfights	-8.44E-03
heavyfights	7.02E-02
belts	1.03E-01
rival	1.78E-01
fan_fav	3.71E-01
Intercept	1.14E+01

Can I improve my model?


Mean Squared Error by Polynomial Degree
train
test
Mean Squared Error
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
Polynomial Degree
Optimal Degree is 1

Elastic Net Regularization



Elastic Net (L1 + L2)

$$J(\beta_0,\beta_1)=\frac{1}{2m}\sum_{i=1}^{m}\left(y_\beta(x_{obs}^{(i)})-y_{obs}^{(i)}\right)^2+\lambda_1\sum_{j=1}^{k}|\beta_j|+\lambda_2\sum_{j=1}^{k}\beta_j^2$$

α=[1e-10,1e-8,1e-5,1e-3,1e-1,0,1,100]

α = .00001

Elastic Net Cross Validation

- Confirmed α = .00001 minimizes MSE
- Score the model against test sets
  - R^2 = 0.6388818636528790 2
- CV against training sets
  - MSE = 7.5015622479941857e-06
- Lasso Ratio = .5

bouts	0
attend	0
ticketrev	0.0165
tvrate	0.0523
strikes	0
takedowns	0.003
submission	0
passes	-0.01
lightfights	-0.006
heavyfights	0.0385
women	0.002
men	0
belts	0.0074
KO	0
DEC	0
SUB	0.004
Other	0
rematch	0.0009
rival	0.0098
rounds	0
tottime	0
tech_star	-0.009
fan_favorite	0.0243

Lasso Regularization

- Score the model against test sets
  - R^2 = 0.6780955607372309 8
- CV against training sets
  - MSE = -6.78565721e-06
- Eliminates Lightweight, passes, sub

bouts	0
attend	0
ticketrev	0.018
tvrate	0.039
strikes	0
takedowns	0
submission	0.0031
passes	0
lightfights	0
heavyfights	0.0142
women	-0.002
men	0
belts	0.0039
KO	-0.005
DEC	0
SUB	0
Other	0
rematch	-0.0042
rival	0.0085
rounds	0
tottime	0
tech_star	-0.0083
fan_favorite	0.0256

Thank you.

Saturday 07.29.2017 at 10:00 PM ET

U.S. Broadcast: Pay Per View | **Prelims:** FX

Promotion: Ultimate Fighting Championship

Ownership: WME-IMG

Venue: Honda Center

Location: Anaheim, California, United States

Enclosure: Octagon

TV Announcers: Jon Anik, Joe Rogan, Dominick Cruz

Ring Announcer: Bruce Buffer

Post-Fight Interviews: Joe Rogan

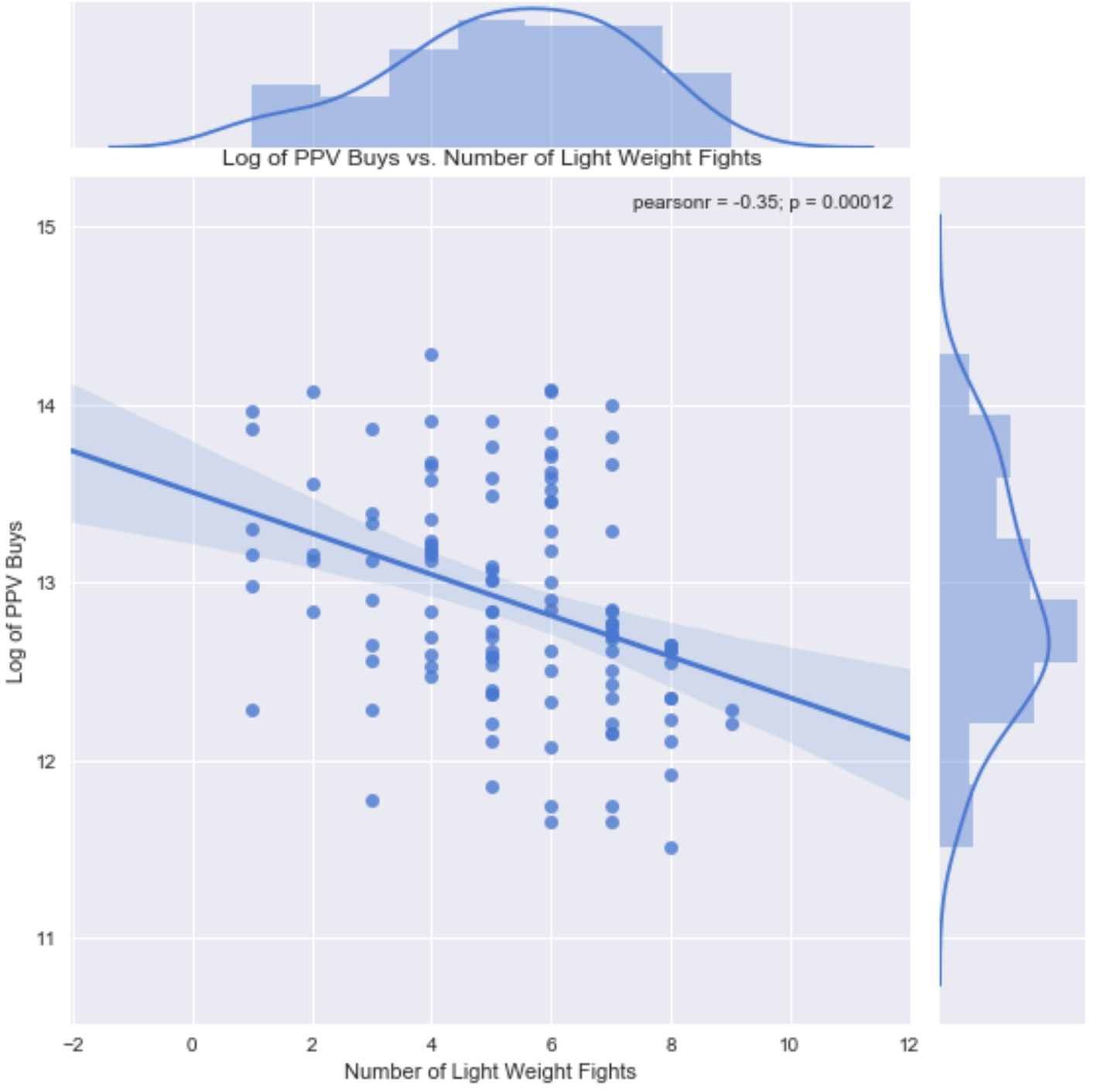
Ticket Revenue (live gate): $2,448,870

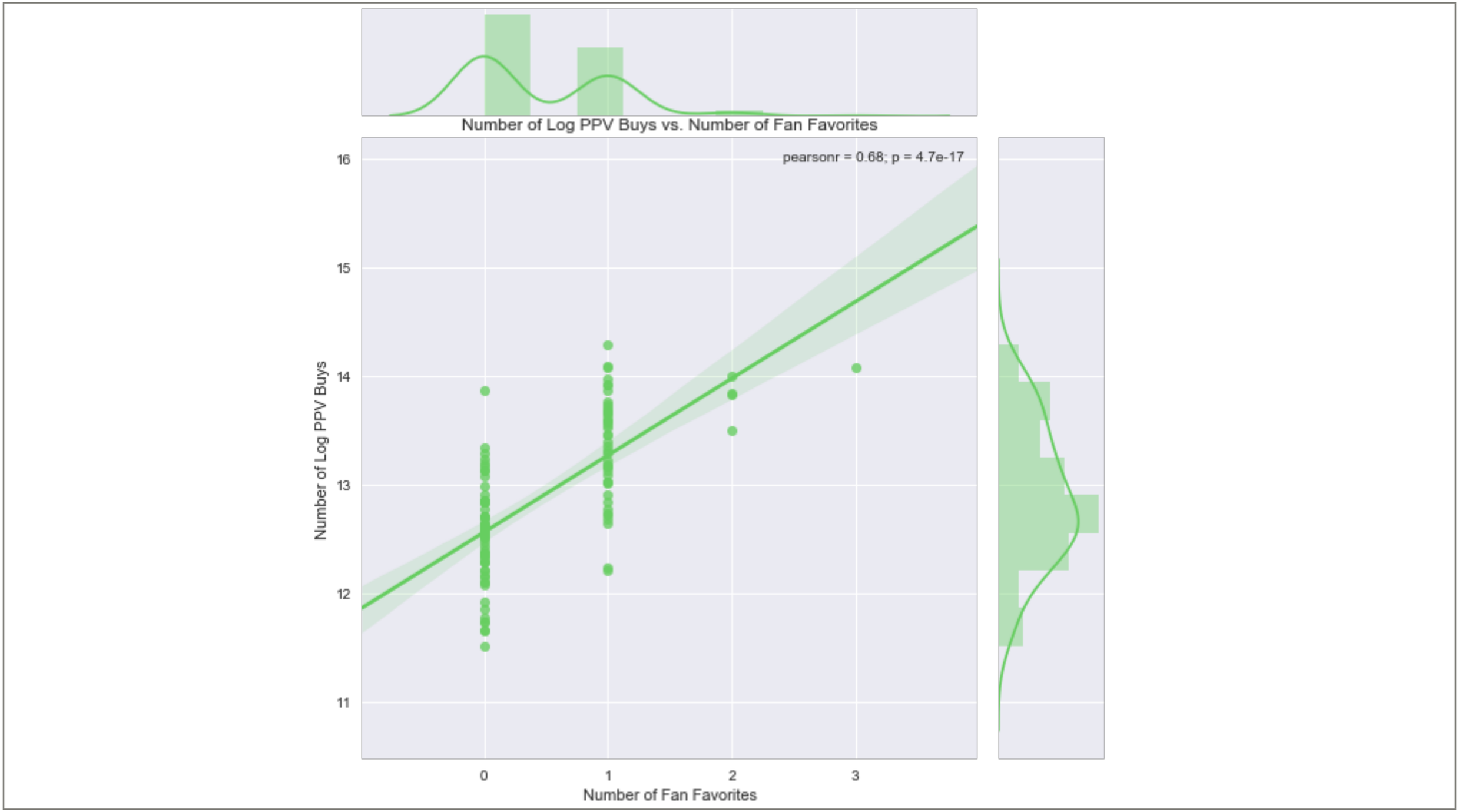
Attendance: 16,610

PPV Buys / Buyrate: 860,000 | **TV Ratings:** 886,000 avg. viewers (FXX prelims, 1.07M peak)

Number of MMA Bouts: 12

Event Pages: Sherdog | Wikipedia | The UG | Promoter


Log of PPV Buys vs. Number of Light Weight Fights
pearsonr = -0.35; p = 0.00012
Log of PPV Buys
Number of Light Weight Fights
15
14
13
12
11
−2
0
2
4
6
8
10
12


Number of Log PPV Buys vs. Number of Fan Favorites
pearsonr = 0.68; p = 4.7e-17
Number of Log PPV Buys
16
15
14
13
12
11
0
1
2
3
Number of Fan Favorites