ANALYSIS OF SOLAR POWER COST OVER TIME

05MAY2018

MARK SAUSVILLE
SAUS@IEEE.ORG

THE PROBLEM

Produce a model that will allow prediction of the Fair Market Value cost of installing a solar electrical system so the customer can decide to install now or defer to enjoy more favorable pricing.

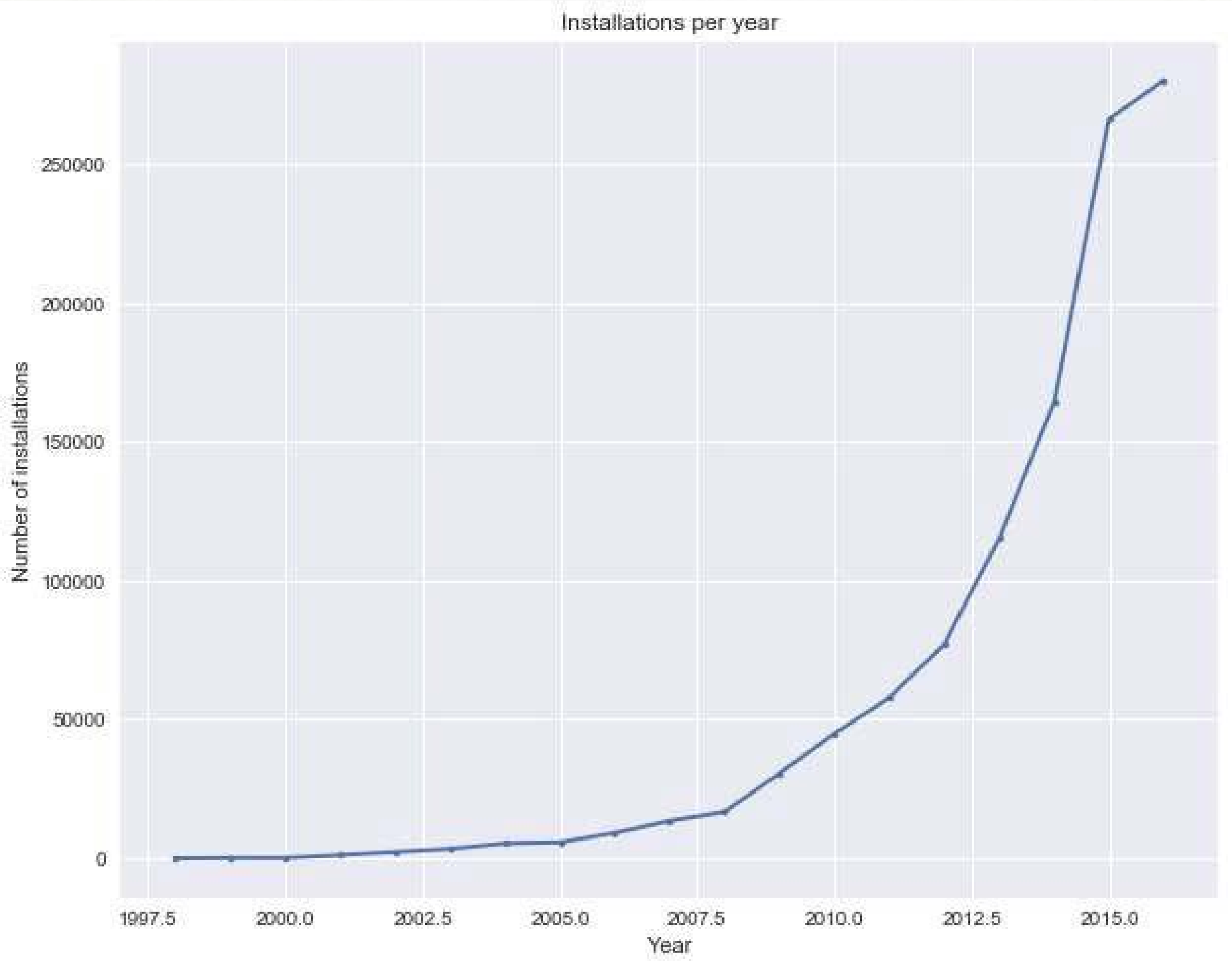
THE CUSTOMER

- Jane Q. Public, homeowner
- Direct purchaser of solar power
- 3 – 8 kwatts
- Roof-mounted

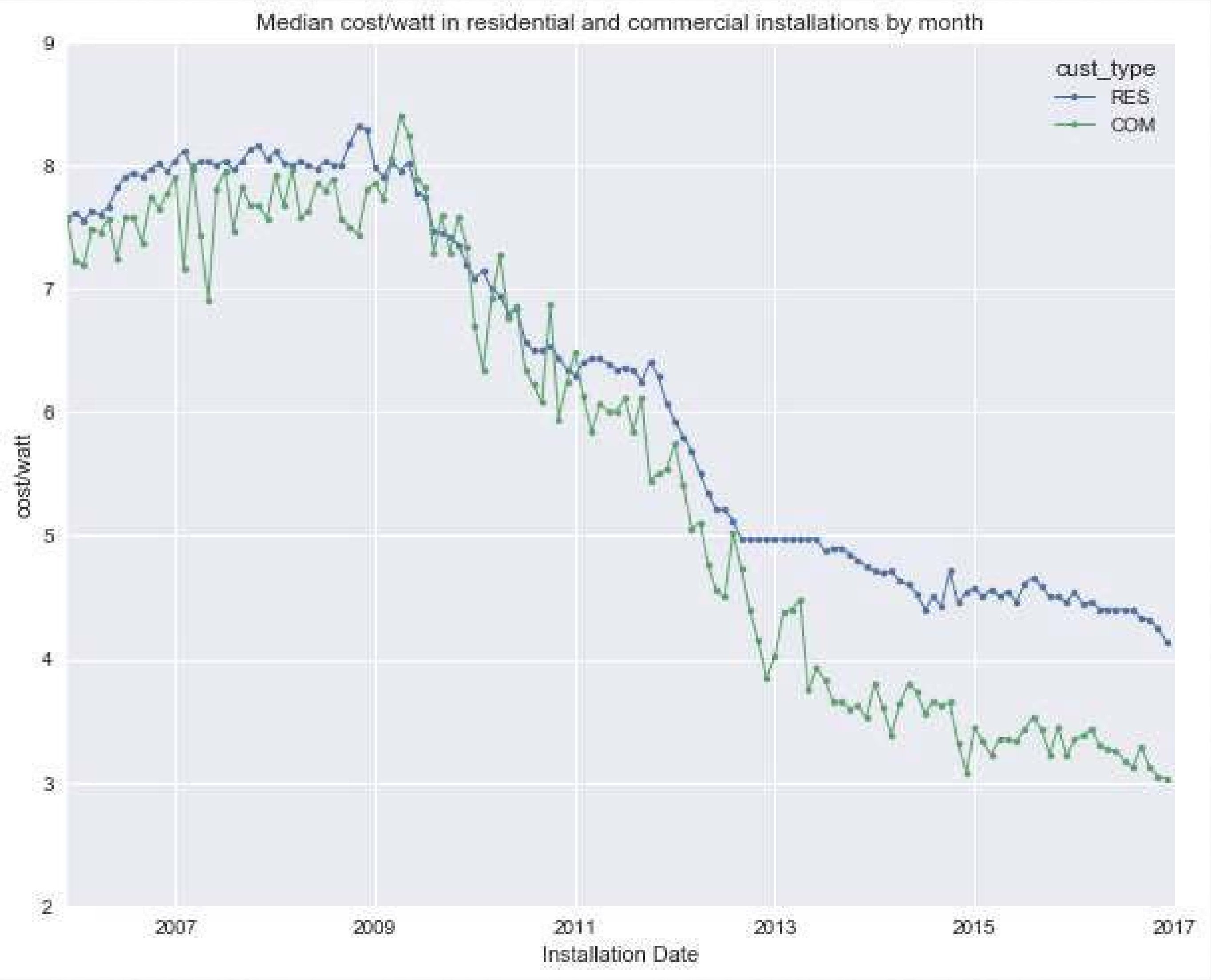
THE DATA

- https://openpv.nrel.gov/search
- ~1,000,000 rows, ~80 features
- Two datasets
  - NREL – National Renewable Energy Laboratory
    - had many duplicates
  - LBNL – Lawrence Berkeley National Laboratory
    - had fields to id info source ➔ No duplicates

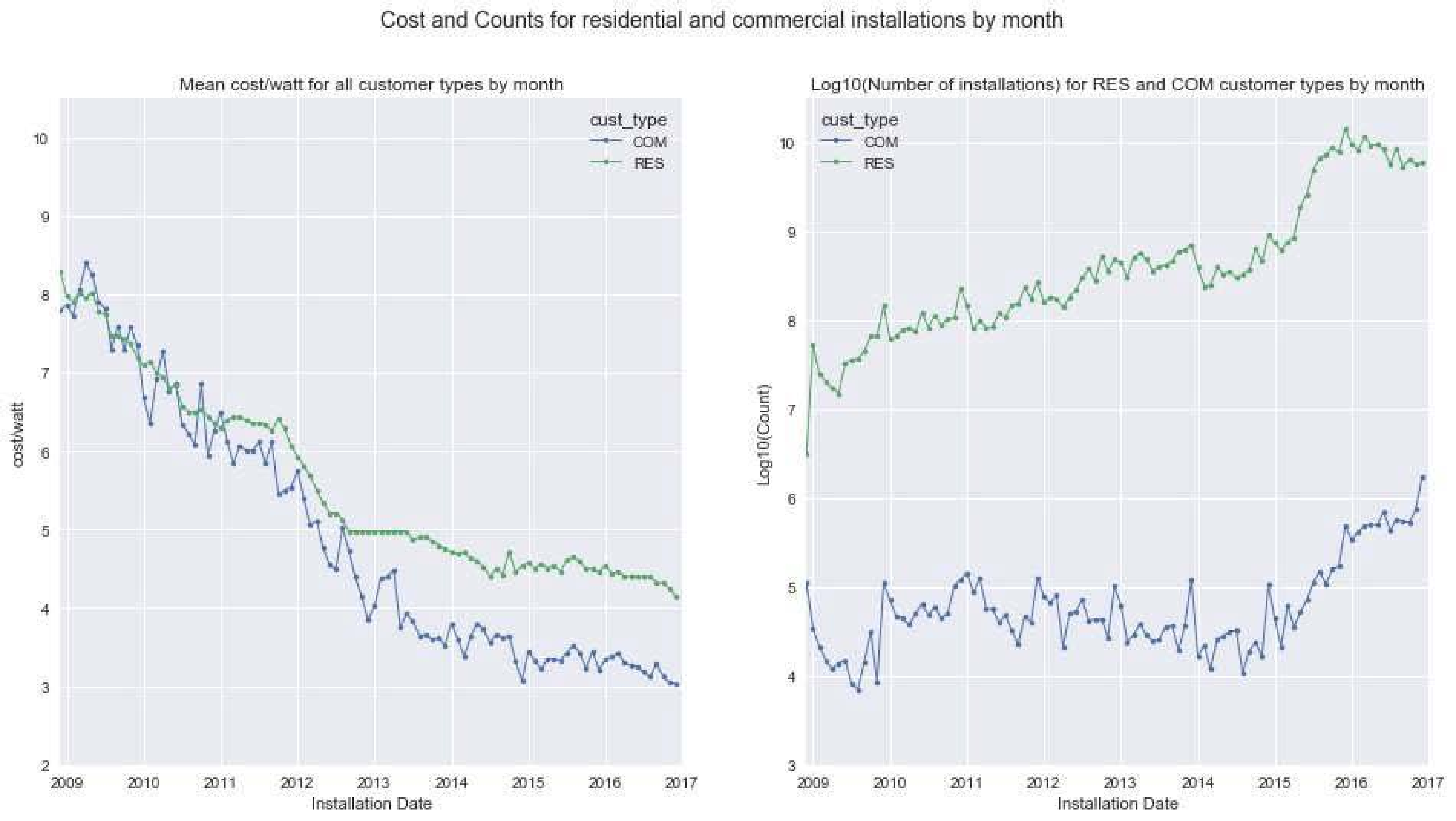
GROWTH OF HOME SOLAR POWER



COST/WATT OVER TIME



RESIDENTIAL VS COMMERCIAL

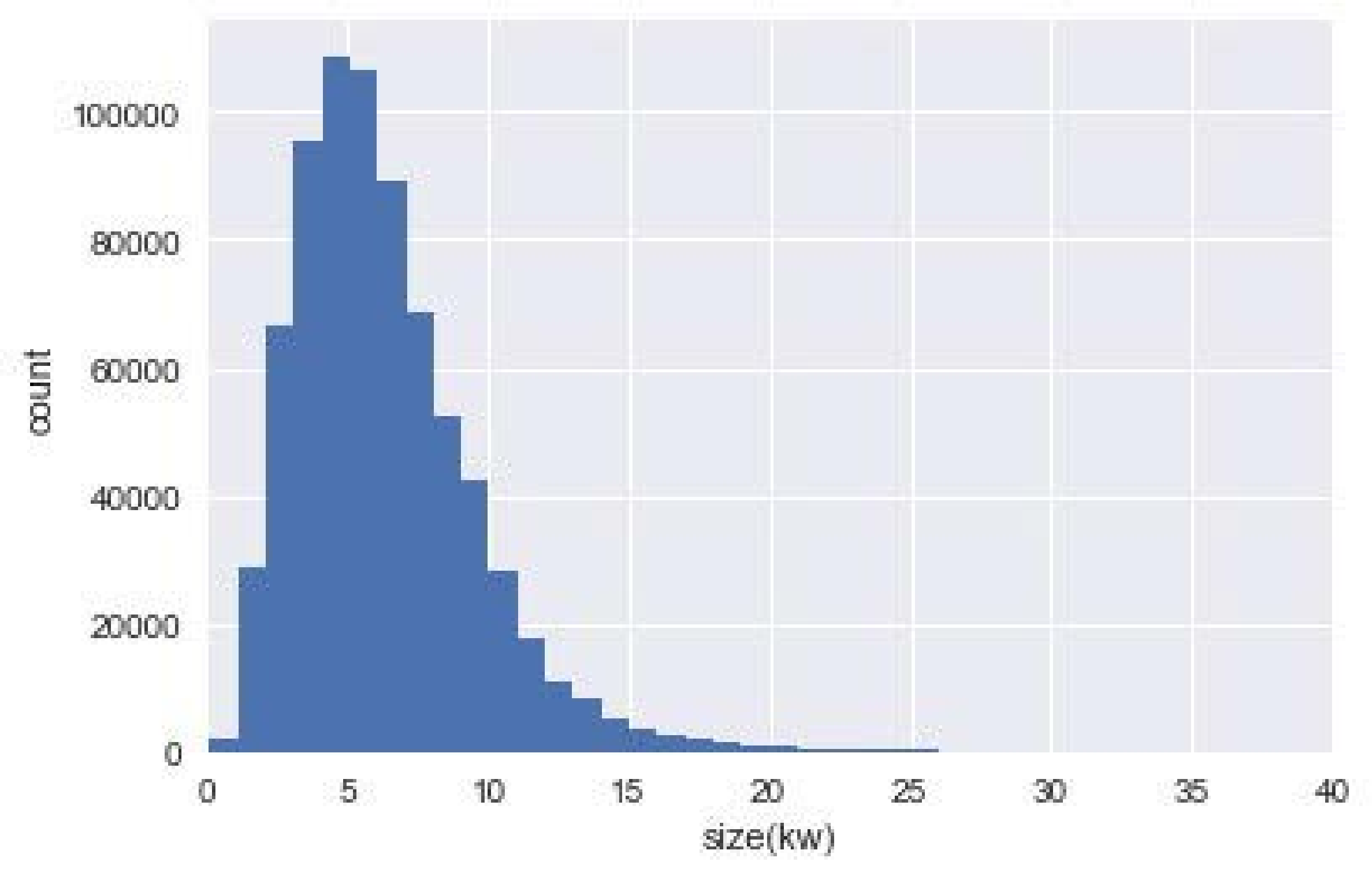


RESIDENTIAL VS COMMERCIAL SIZES

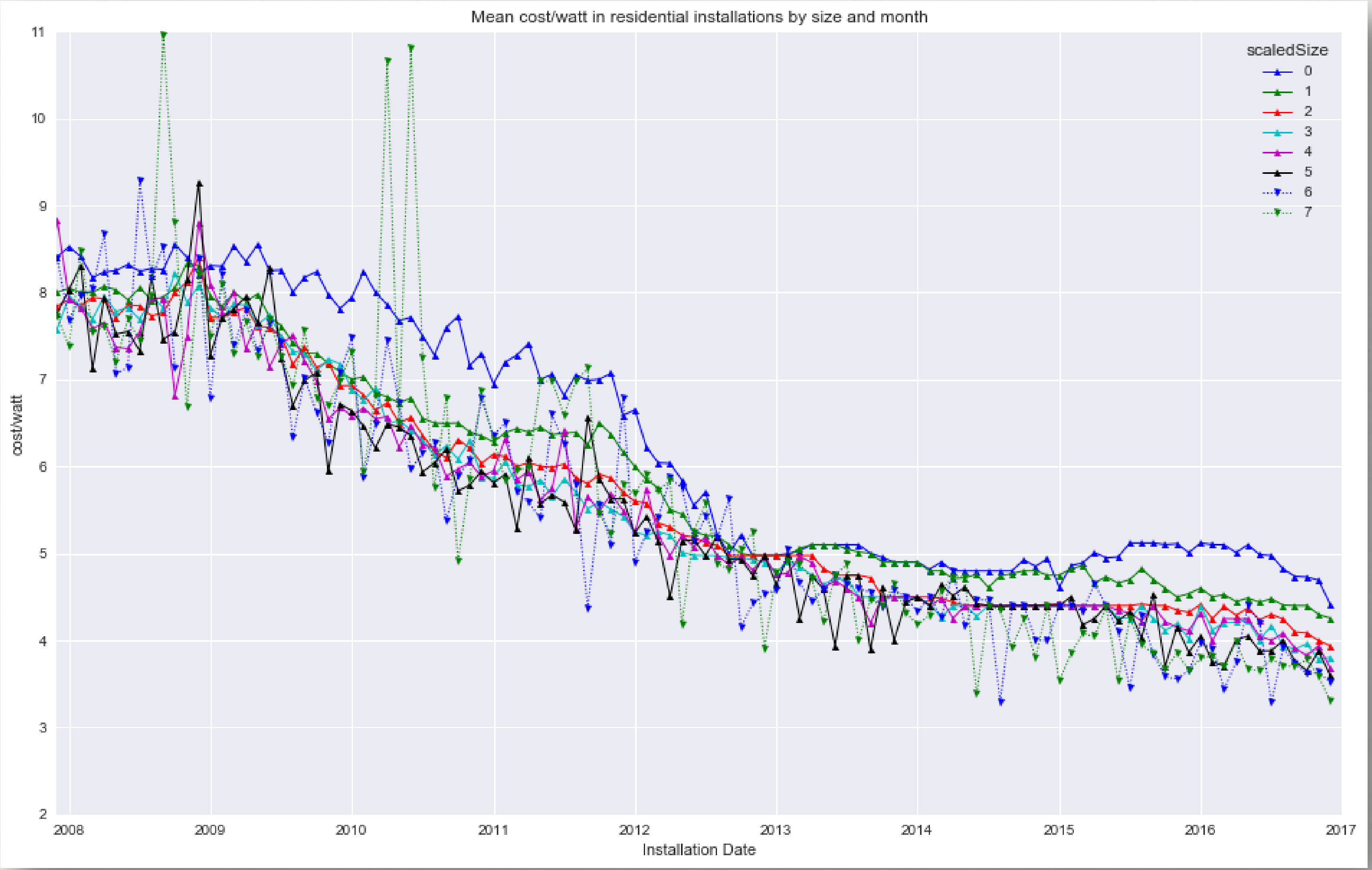
Table 3 Summary statistics for size (kw) in commercial and residential installations

	count	mean	std	min	50%	max
Com size(kw)	15199	131.3	373.8	0.33	25.2	5999
res size(kw)	745688	6.45	7.1	0.11	5.75	1989

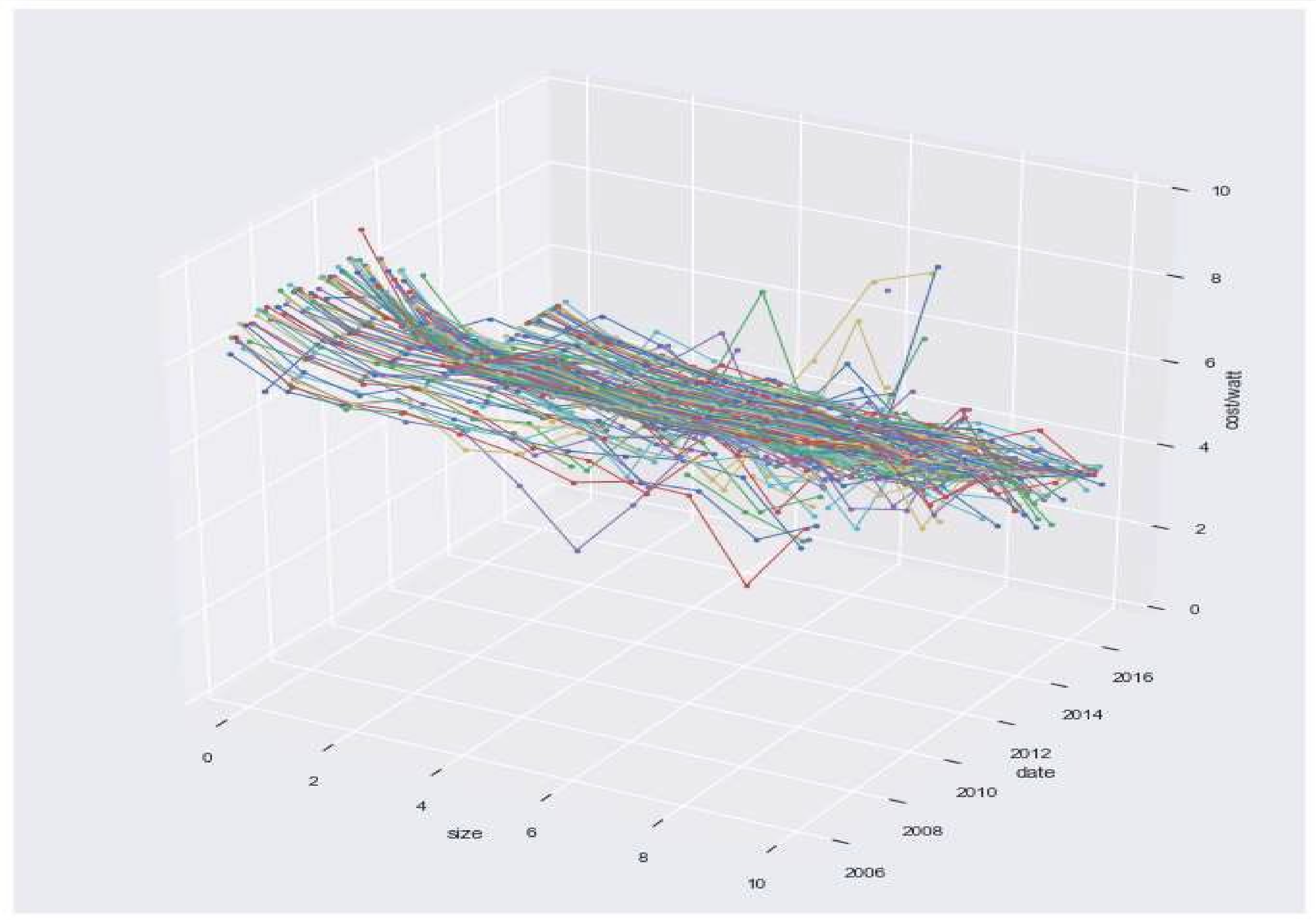
RESIDENTIAL SIZE DISTRIBUTION



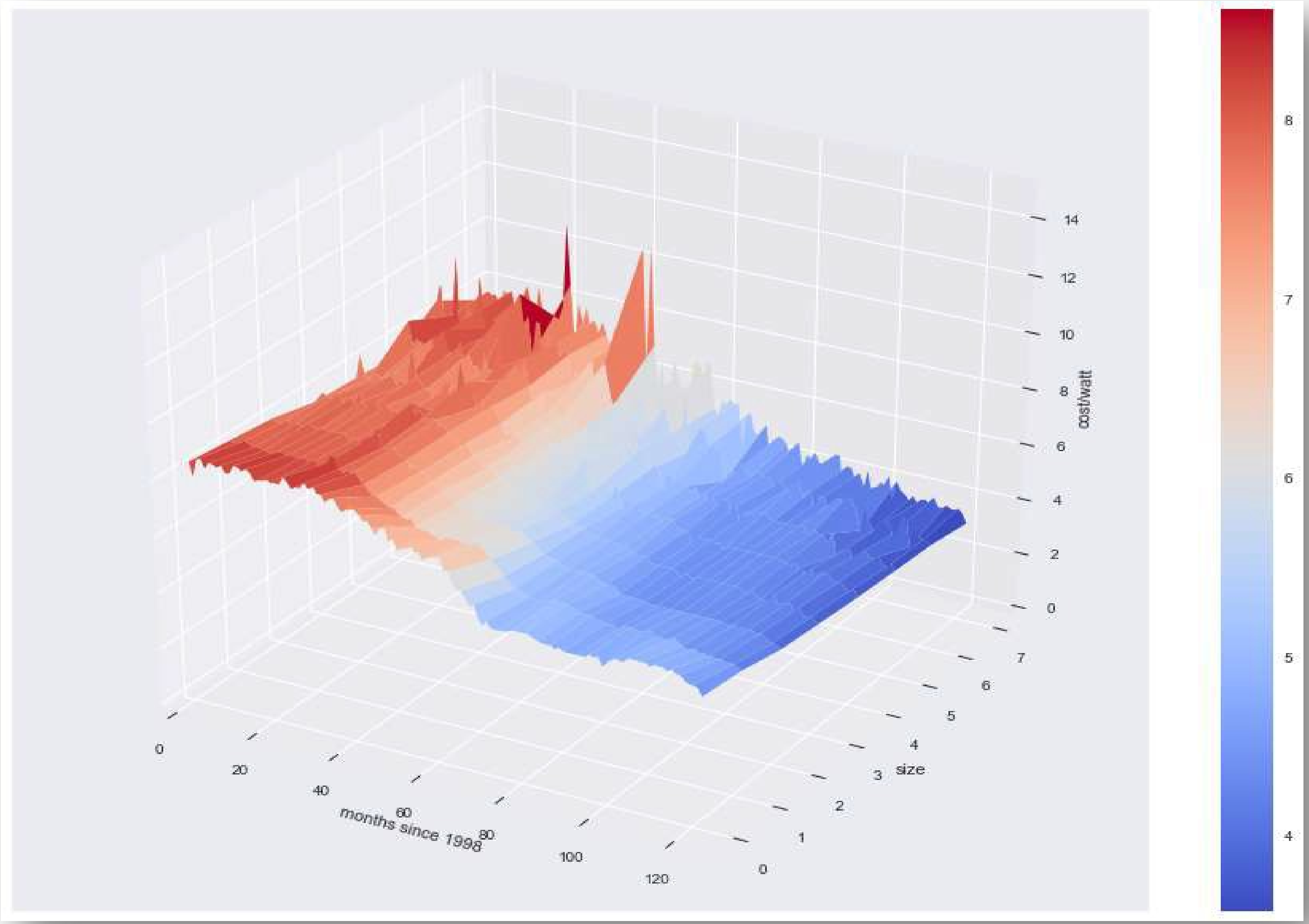
COST VS SIZE 1



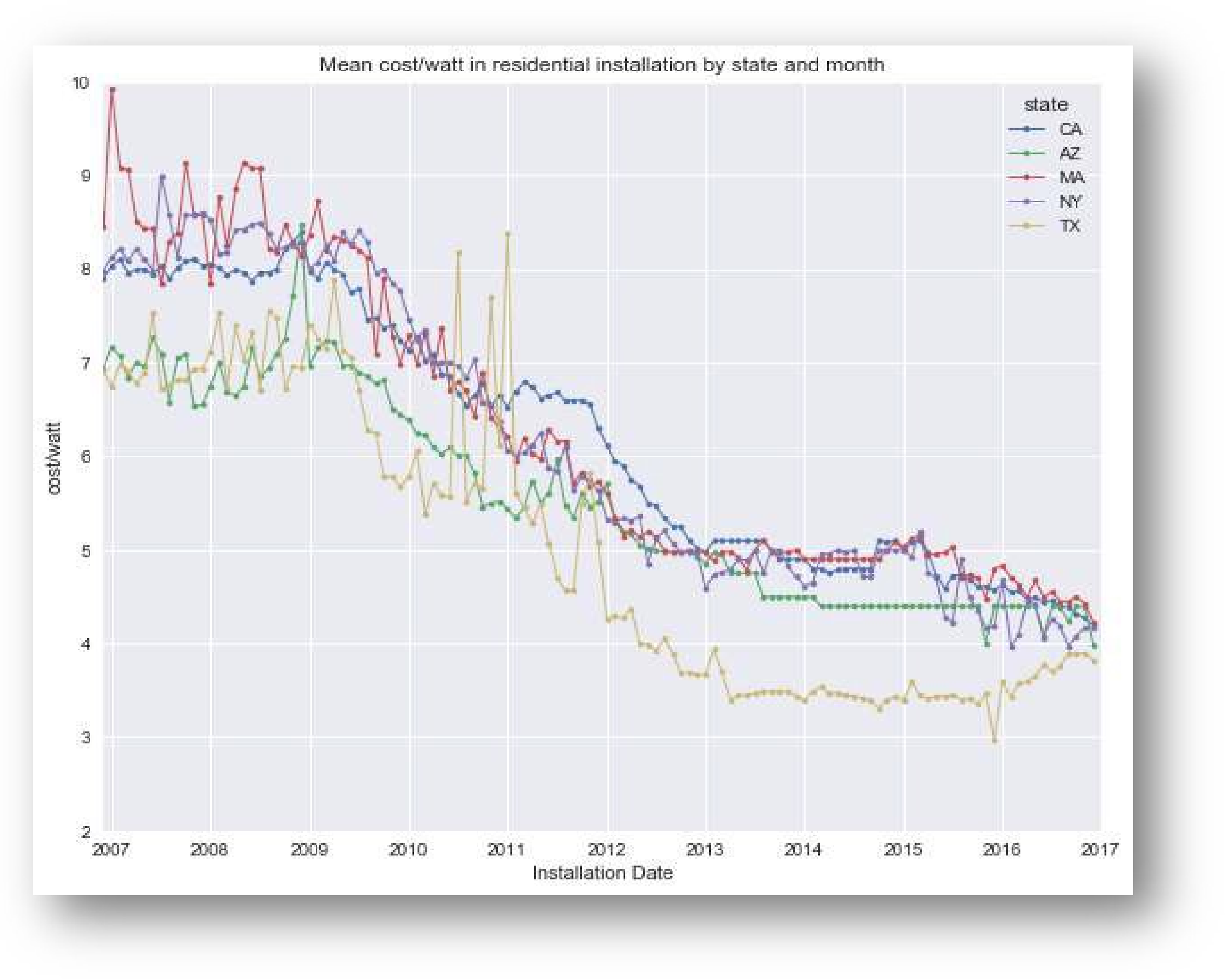
COST VS SIZE 2



COST VS SIZE 3



DOES COST/WATT VARY BY REGION?



DRIVERS OF VARIATION IN COST/WATT

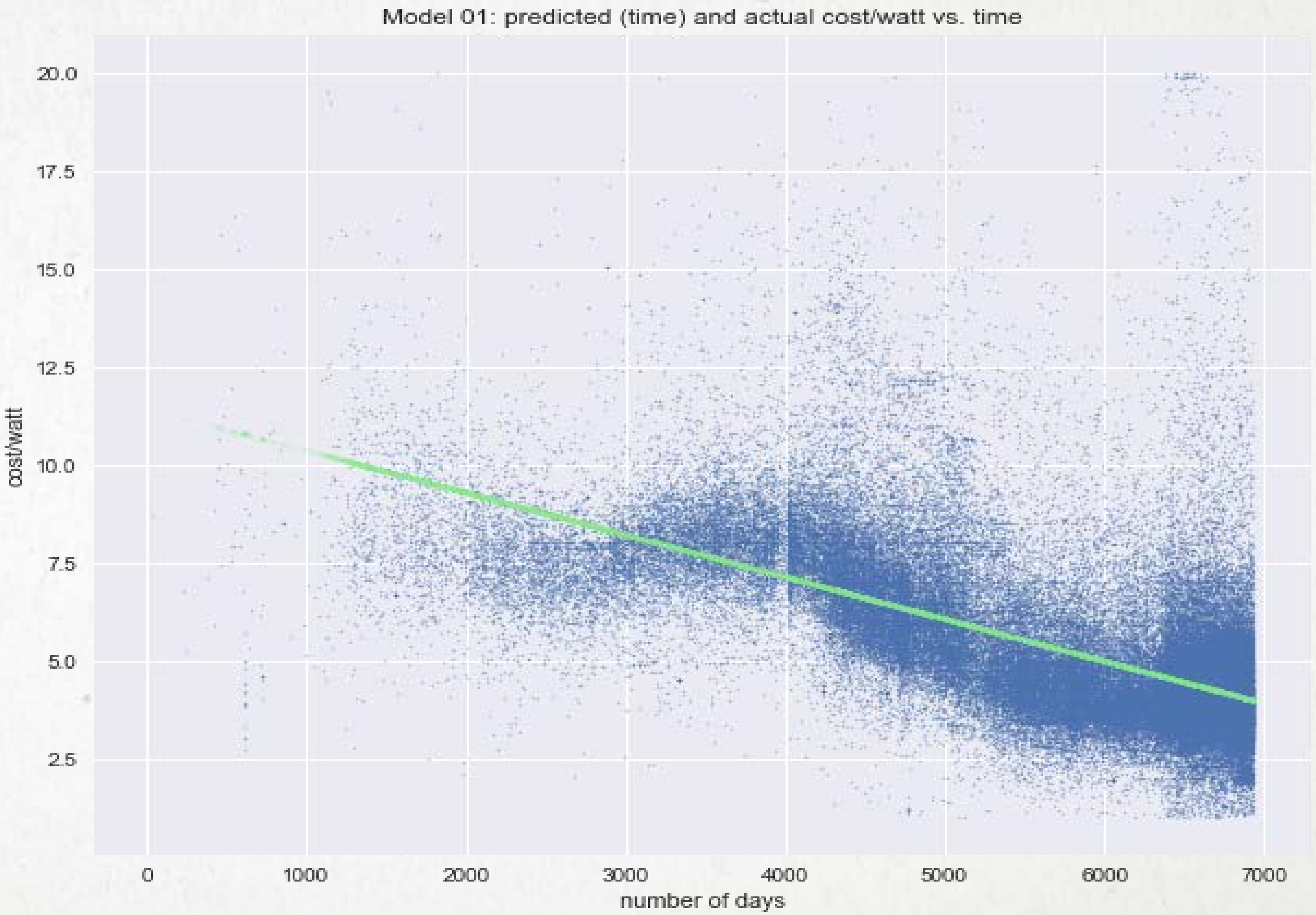
The exploration show that the following features have a relationship with cost.

- Time (strong)
- Size (moderate)
- Location (moderate)

We will use these features to model the cost.

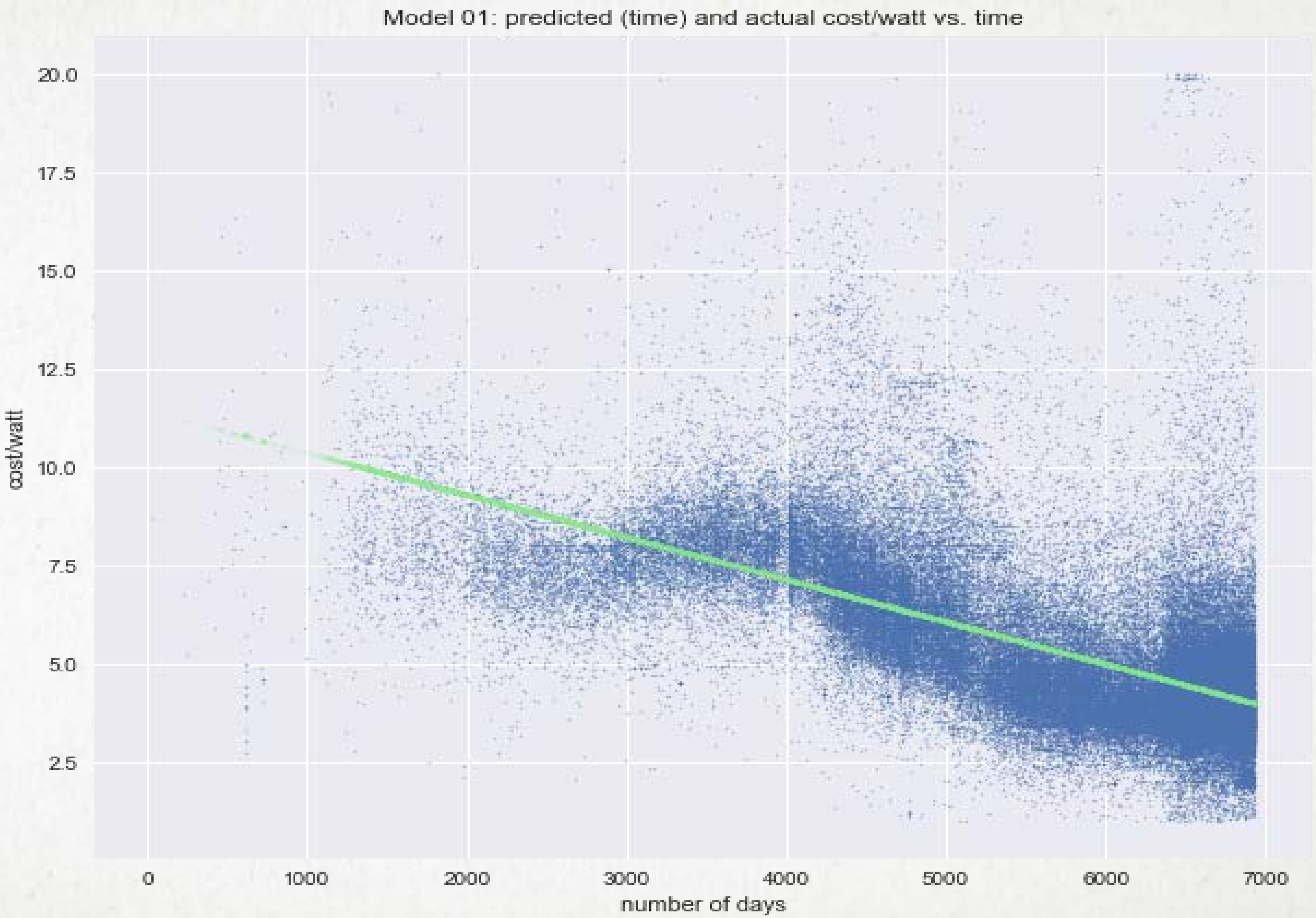
MODEL WITH STATISTICAL/MACHINE LEARNING

- Noisy Data
- Baseline Model
- Add features
- Transform features



MODEL 1: LINEAR (TIME)

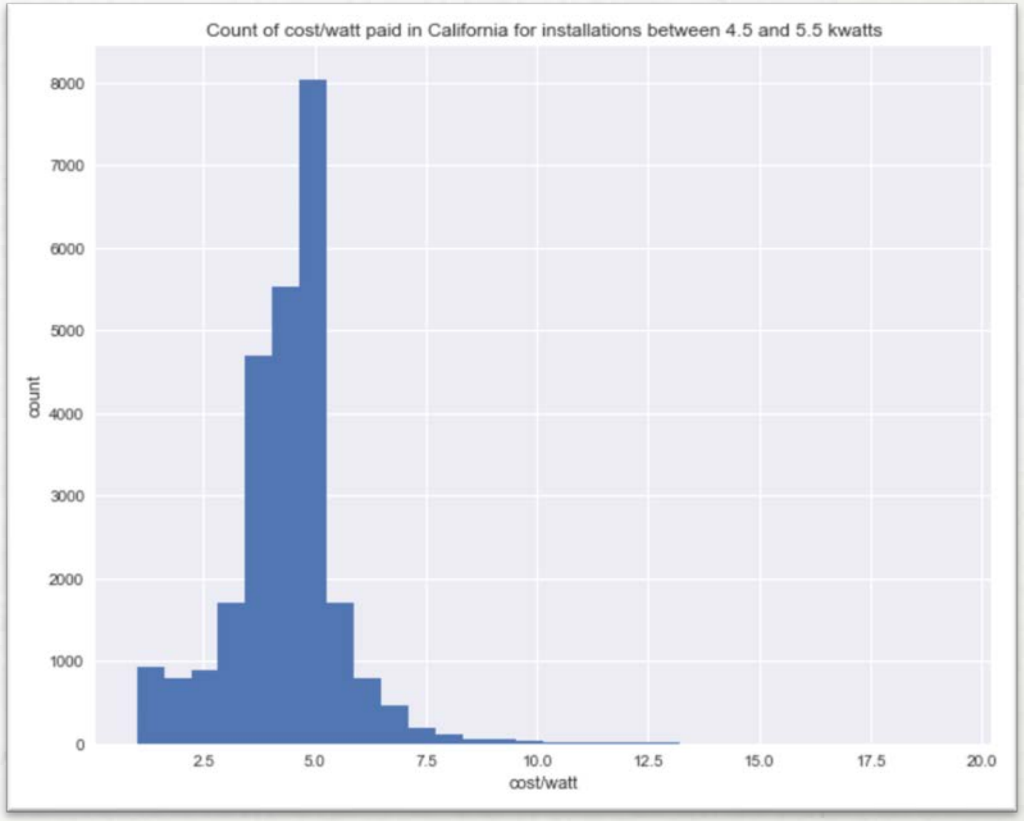
- R2: 0.404



INTERLUDE – NOISY DATA

- California
- 2016
- 4.5-5.5 kw
- 26k installations

	count	mean	std	50%	max
cost_per_watt	26107	4.41	1.33	4.50	19.30

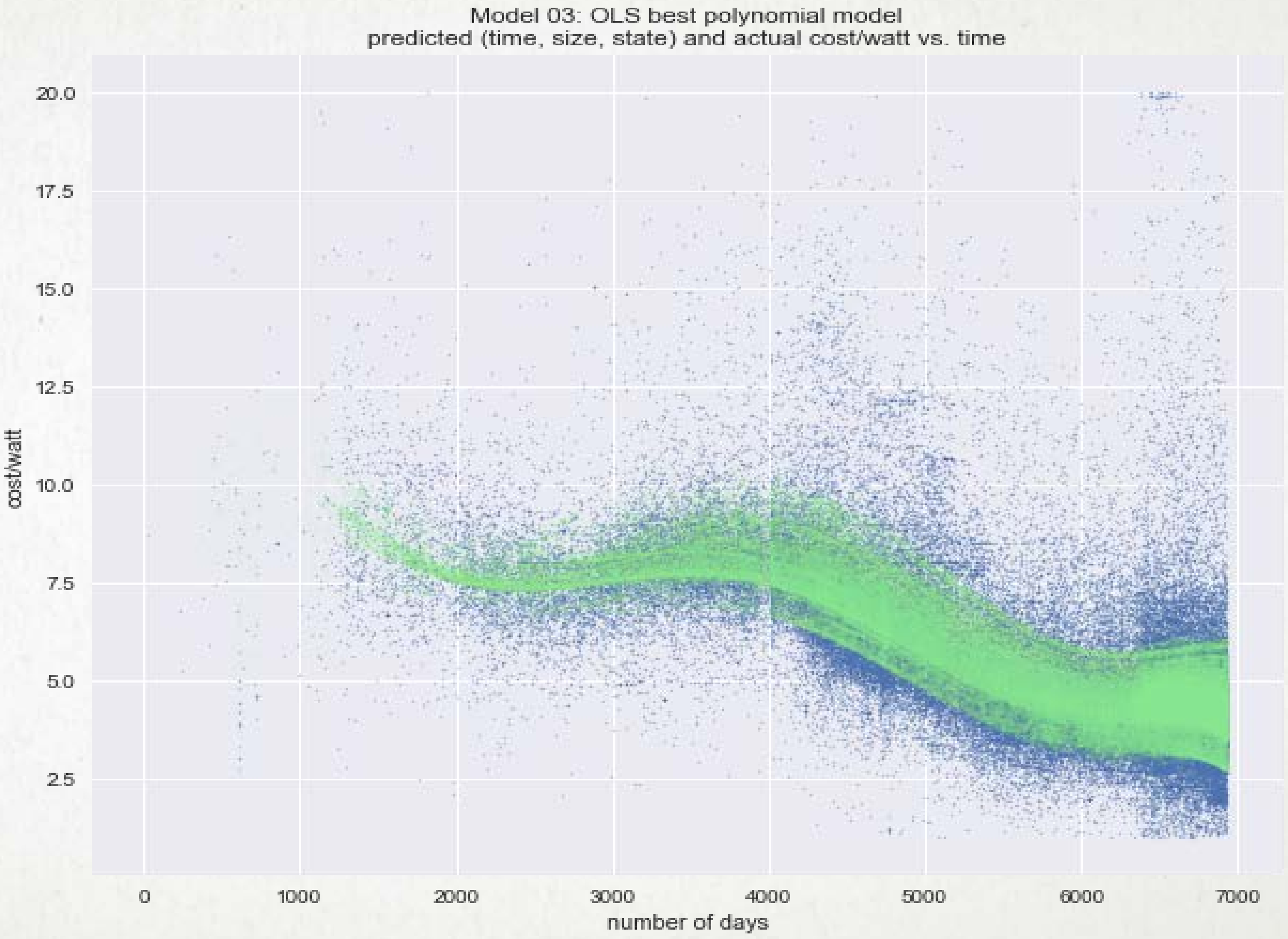


MODEL 2: LINEAR (TIME, SIZE, STATE)

- $R2$: 0.4613

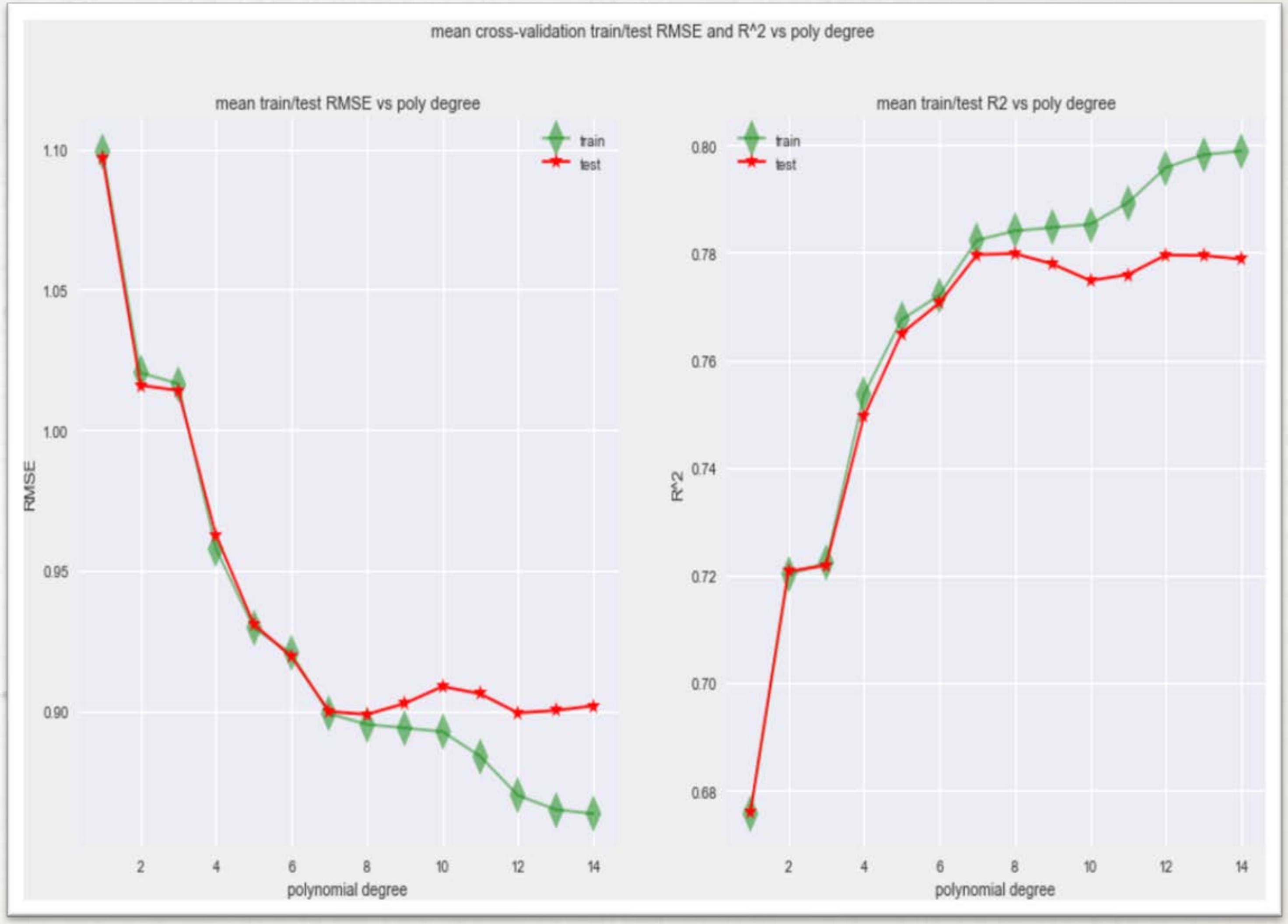
MODEL 3: POLYNOMIAL (TIME, SIZE, STATE)

- R2: 0.520
- Degree: 14



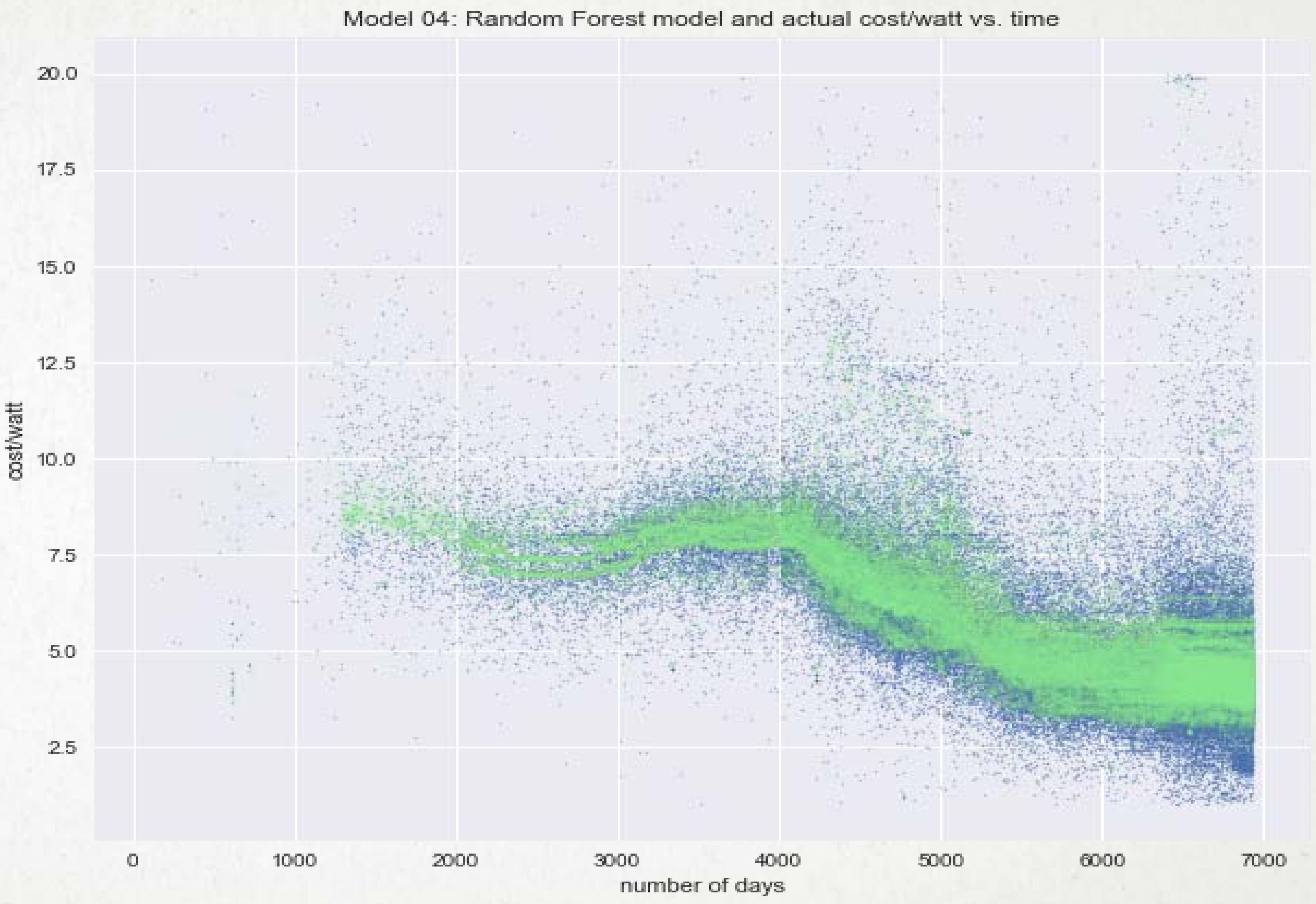
INTERLUDE: HYPER-PARAMETER TUNING

- Approach:
  - Split into train and test
  - Grid-seach with cross-validation on training set.
  - Apply best parameters to the test set.



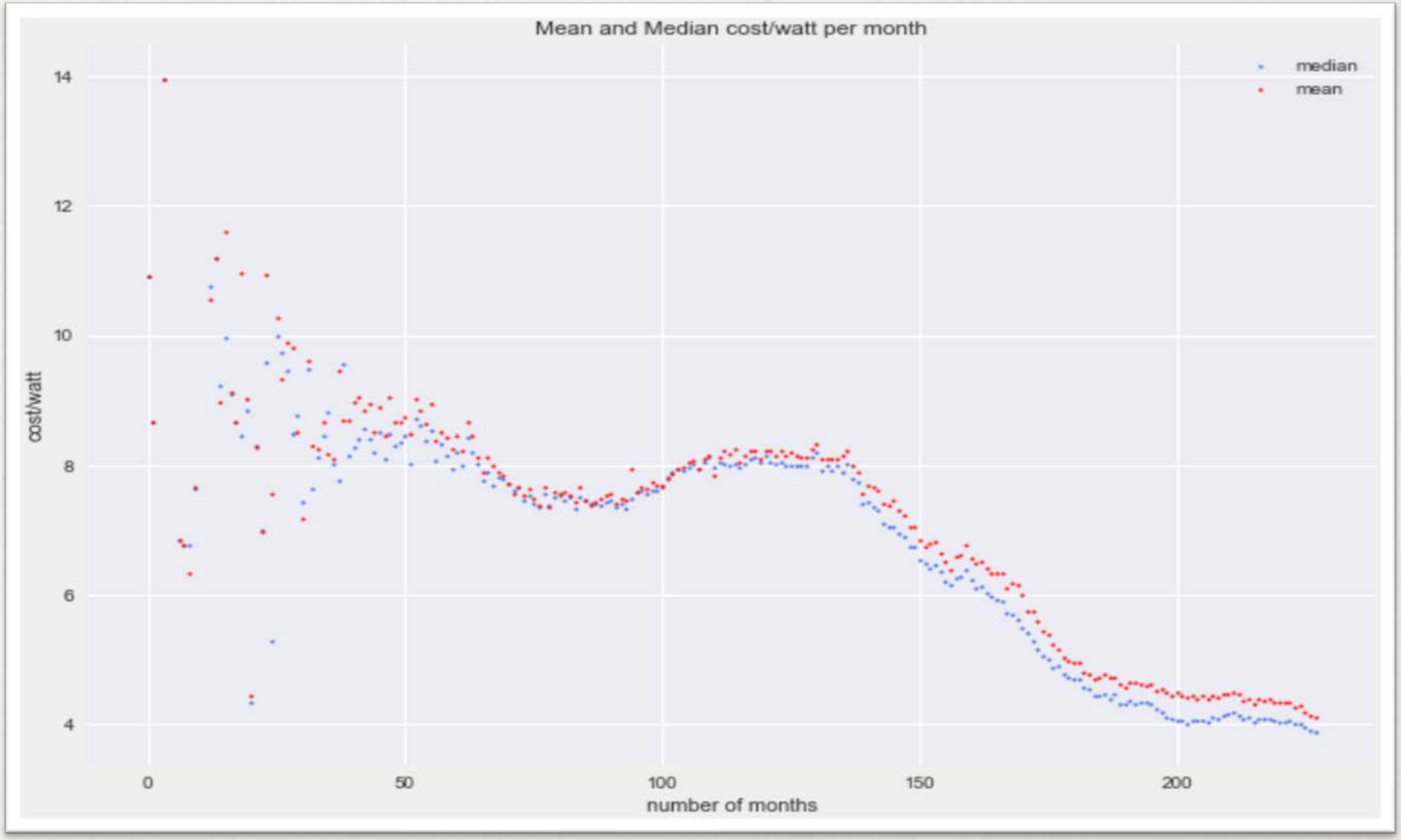
MODEL 4: RANDOM FOREST (TIME, SIZE, STATE)

- R2: 0.5993
- Predicts noise!



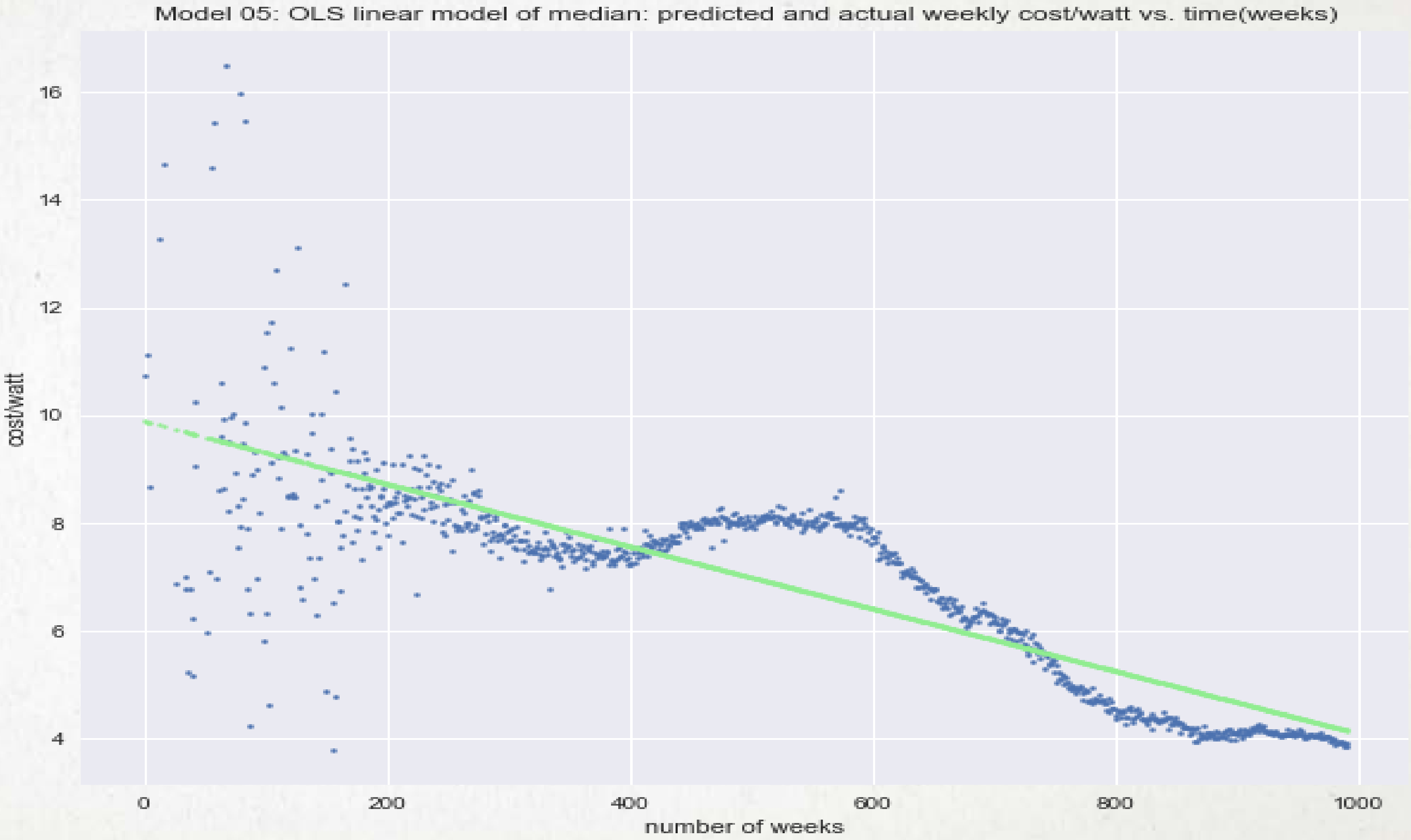
INTERLUDE: MEAN AND MEDIAN

- Mean rejects noise
- Median rejects noise
- Mean sensitive to outliers
- Median is better proxy for Fair Market Value (FMV)



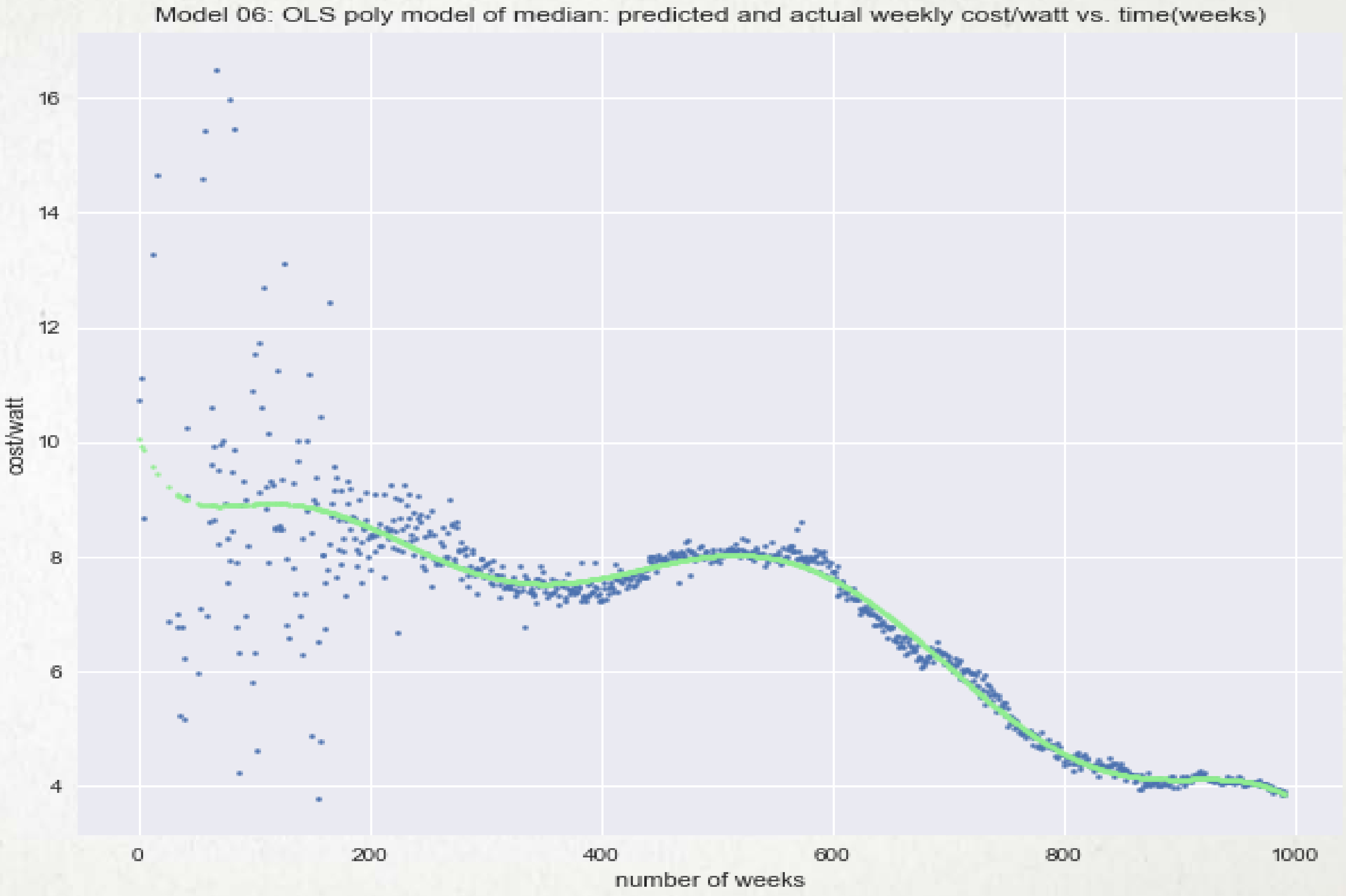
MODEL 5: LINEAR (MEDIAN(TIME))

- R^2: 0.715



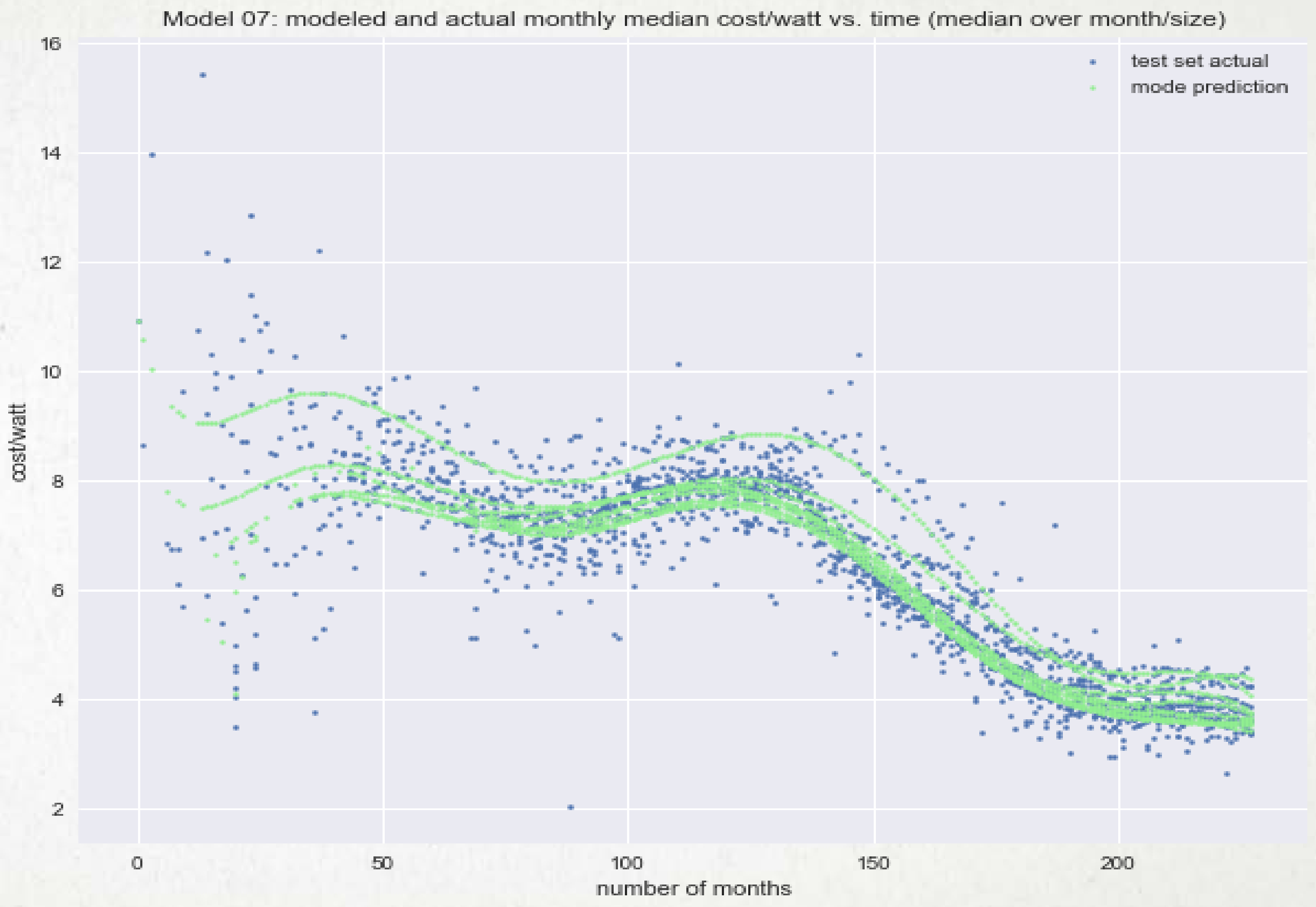
MODEL 6: POLYNOMIAL (TIME (MEDIAN))

- R2: 0.834



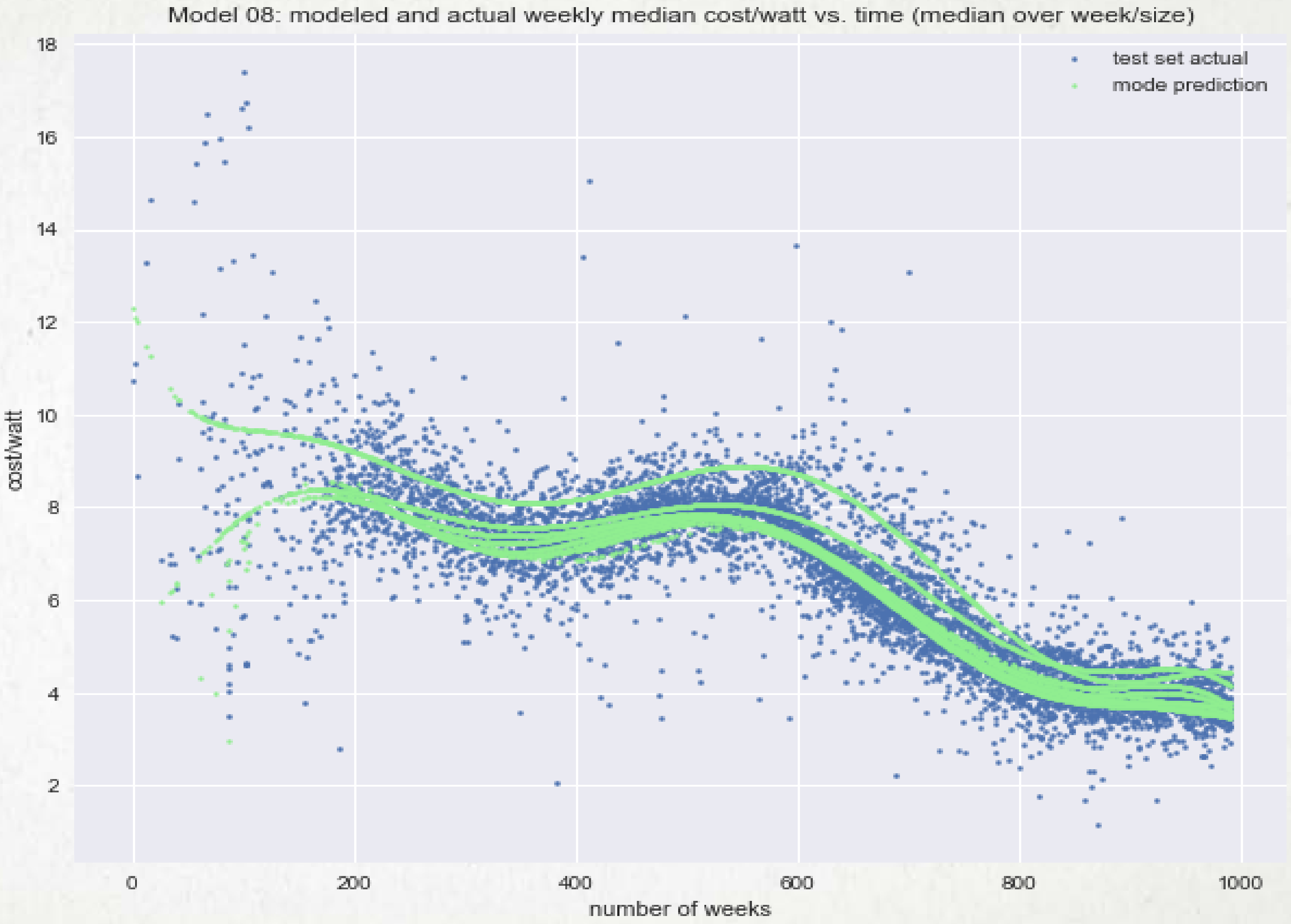
MODEL 7: POLYNOMIAL (TIME, SIZE (MEDIAN))

- Time (months)
- Degree



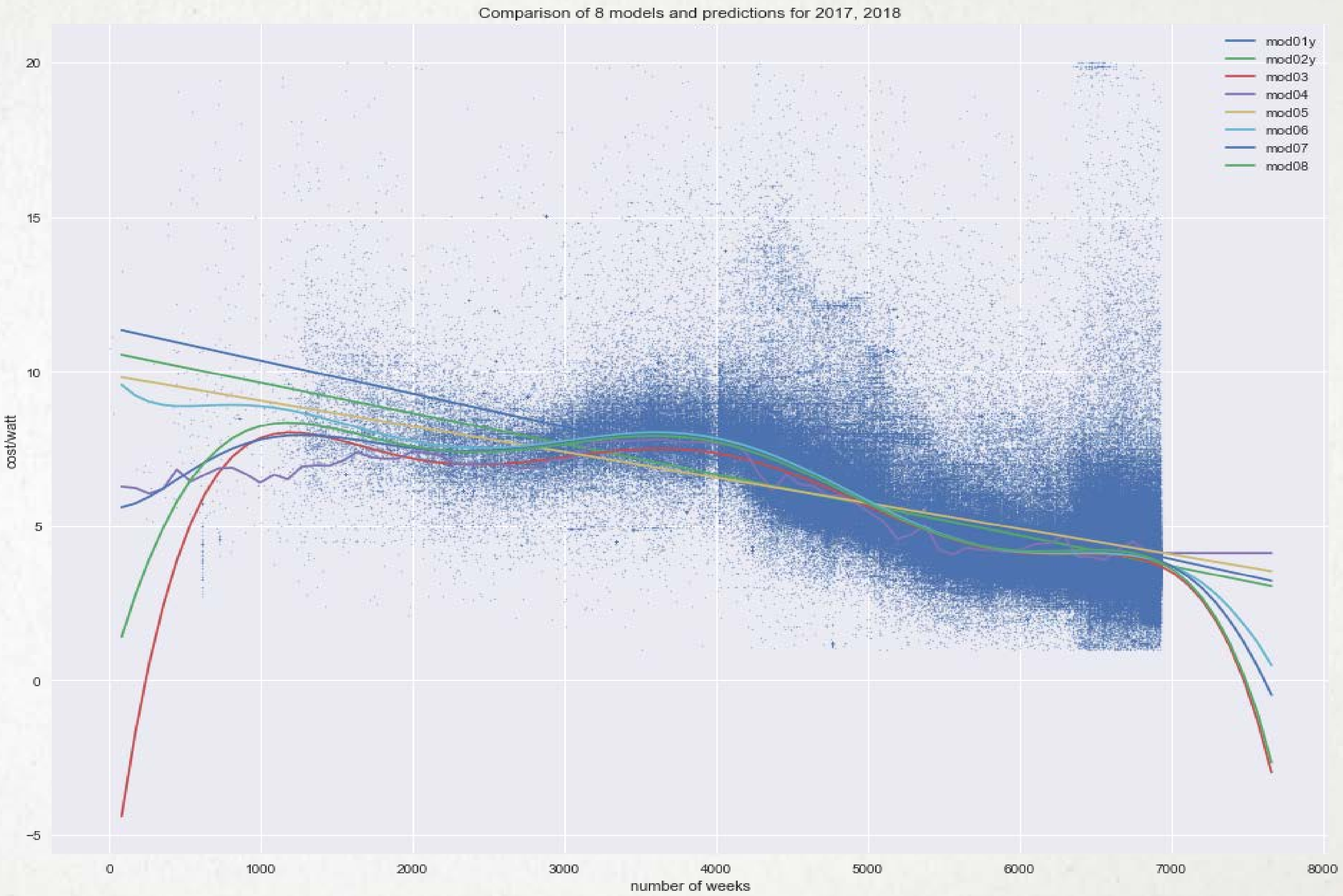
MODEL 8: POLYNOMIAL (TIME (MEDIAN))

- Time (weeks)
- Degree
- R^2



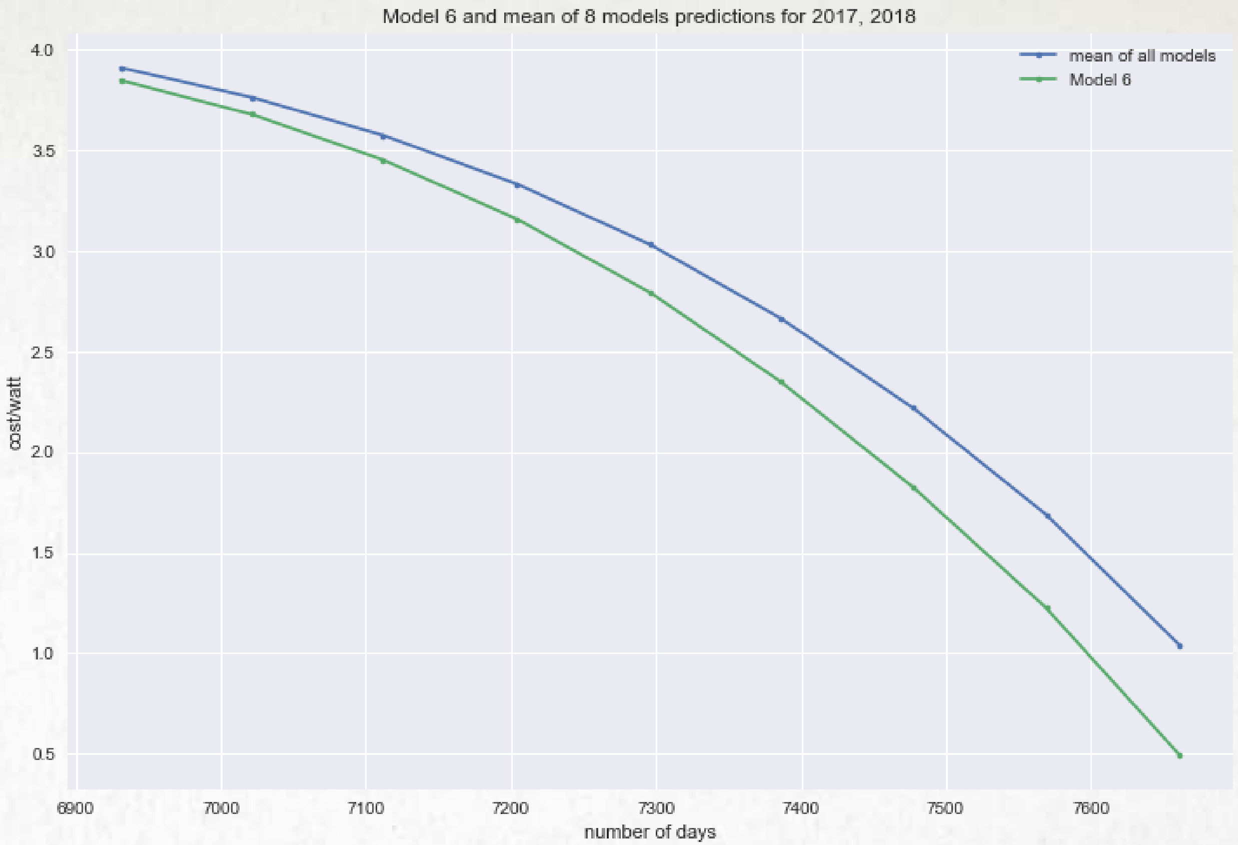
PREDICTIONS(I)

- Extrapolate all models – reject outlying predictions.
  - Linear models
  - Polynomial models
  - Random Forest



PREDICTIONS(II)

- 2017/2018
- 8 quarters past EOD
- Upper bound – mean()
- Lower bound – model 6



	20161231	20170331	20170630	20170930	20171231	20180331	20180630	20180930	20181231
Mean prediction	3.91	3.76	3.57	3.33	3.03	2.66	2.22	1.69	1.04
Model 6 Prediction	3.85	3.68	3.45	3.16	2.79	2.35	1.83	1.22	0.49

RECOMMENDATIONS

"WILL IT BE MORE COST-EFFECTIVE TO INSTALL NOW, OR WILL I SAVE MONEY BY WAITING A YEAR OR TWO?"

- **Get several competitive quotations for the solar installation**

  The data show that prices for similar solar installation vary dramatically (in some case by a factor of 10). A selection of vendors is likely to provide the best price. Getting a competitive price is probably more important than optimal timing of the purchase.

- **Given the shape of the predicted cost curve, it is reasonable to expect that median cost will drop by about $1.00/watt over 2017.**

  It is also reasonable to expect at least this rate of decline in 2018 and there are grounds to believe that the rate of price decline will accelerate in 2018. It would be prudent to revisit this analysis incorporating more data as it becomes available.

Given these predictions the customer can weigh the current cost of installation and the expected price decrease and make an informed decision.

FUTURE WORK

- Validate with 2017 data, when it becomes available.
- Experiment with other feature transformations (logistic function).
- Apply time-series methods.
- Measure other regressors.