

CSE 519 -- Data Science
Quant Shop Video Review Form

Name: **Bipul Islam**

Email: **ibipul@cs.stonybrook.edu**

ID Number: **111578726**

Video episode title: **Winning Baby Pool** Episode number :**3**

Of the eight episodes, how many have you previously watched?: **0**

Give the month and year you started your graduate studies at Stony Brook: **August, 2017**

Did you learn anything by watching it? **Yes**

Describe stuff you did learn below:

- *Coming up with a model evaluation parameter was interesting. Using error distribution to select whether baseline model actually needs improvement against.*
- *Better visualization of data answers a lot of questions in itself.*

What did you learn from looking at their project report?

- *They did not adjust for overdue/early delivery of the child. As the mother who is a delivery nurse herself says baby can gain upto 0.5-1 ounce per day (from video). Would have been nice to see some questions about it.*
- *Would have liked to see some tweaking with delivery date based on the hospital's own statistics as they seem to have a facility to care for earlyborn babies – often times this has financial angle associated with it as medical care is quite expensive in the US. Maybe they could have gotten a better delivery date approximation.*
- *Pre-natal visits seem to deliver healthier babies, if that data was available could have been put to use. This is often associated with financial standing of the parents. Using a broad brush location data (discussed in the video) might have washed this out.*
- *Not sure whether the raw data could have been scraped to get an idea about average birth weights in families. Might have been interesting to see whether that figures into anything.*

Rate program quality from 1 (terrible) to 10 (amazing): 7

Was it worth watching? **Yes**

What could we do to make the videos better, and more educational?

- *Include more video footage about research on the data. How interesting questions about the probable causal factors figure towards the answer they are looking for.*
- *Include footage that describes the datasets.*

Name: **Bipul Islam**

Email: **ibipul@cs.stonybrook.edu**

ID Number: **111578726**

Video episode title: **The Art of Auction** Episode number :**4**

Of the eight episodes, how many have you previously watched?: **0**

Give the month and year you started your graduate studies at Stony Brook: **August, 2017**

Did you learn anything by watching it? **Yes**

Describe stuff you did learn below:

- Painting Auction is a rigged game. It's not at all an appreciation of genius.
- Prices of painting are more dependent on what auction houses are willing to bet on it, how middle men are going to play it and how people are going to willingly fall for it.

What did you learn from looking at their project report?

- Not really convinced that Art-specialists do unbiased pre-sale estimates. The paper, ***Do Art Specialists Form Unbiased Pre Sale Estimates? An application for Picasso paintings'*** by ***Corina Czujack and Maria Fraga O. Martins***, focuses primarily on Picasso. There is not much wiggle room of manipulation because he is an acclaimed master. They trend may not continue for painters who are newer, or are old and have a specific niche. The report does talk about various hypothesis tests done, but on what data set?— would have been good to know.
- Provenance is mentioned but the feature set doesn't seem to cover it properly. In effect provenance is the factor that may cause wild fluctuations to the pre-sale estimates. Maybe number of times it has changed hands, number of acclaimed collector who owned it, price-increments in each change of hands would have been interesting to explore.
- Gravitas--a measure of excellence is mentioned but not clearly explained, how is it computed. How that adds to the overall appeal how much it captures provenance.
- Kudos on adding Auction house as a feature variable.

Rate program quality from 1 (terrible) to 10 (amazing): 8-9

Was it worth watching? **Yes**

What could we do to make the videos better, and more educational?

- *Make the footage about interacting with the domain expert snappier.*

Name: **Bipul Islam**

Email: **ibipul@cs.stonybrook.edu**

ID Number: **111578726**

Video episode title: **Playing the market** Episode number : **8**

Of the eight episodes, how many have you previously watched?: **0**

Give the month and year you started your graduate studies at Stony Brook: **August, 2017**

Did you learn anything by watching it? **Yes**

Describe stuff you did learn below:

- It is better to follow the data on money bets like 'futures' in this case when predicting something.
- Changes in prices for commodities like oil and gold that depend heavily on the market, it's often a good idea to start with some simple auto-regressive models.

What did you learn from looking at their project report?

- Relationship between oil and gold prices not explored properly, some internet reading threw conflicting information about how these prices interact over observational time-frames. One being used as a predictor of other is hence a bit disconcerting.

Rate program quality from 1 (terrible) to 10 (amazing): 6

Was it worth watching? **Yes**

What could we do to make the videos better, and more educational?

- *No comments*

Name: **Bipul Islam**

Email: **ibipul@cs.stonybrook.edu**

ID Number: **111578726**

Video episode title: **Finding Miss Universe** Episode number : **1**

Of the eight episodes, how many have you previously watched?: **0**

Give the month and year you started your graduate studies at Stony Brook: **August, 2017**

Did you learn anything by watching it? **Yes**

Describe stuff you did learn below:

- Clever use of human perception data to augment the scarce feature-set and creating a score out of it.

What did you learn from looking at their project report?

- Figure 3 where the correlation matrix of all features is shown, it seems most factors have very minimal correlations with each other whether positive or negative ($|\text{corr}| < 0.3$). Only factor with a significant correlation is country_score to winner, and our_score (human rated) to winner-- .4 and .39 respectively. Top choices by the model were USA and Venezuela both of which have highest number of Miss Universe titles, till date.
- Liked the concept of weighted ranking scheme to normalize the contestants over their performances in various other competitions.

Rate program quality from 1 (terrible) to 10 (amazing): 5

Was it worth watching? **Yes**

What could we do to make the videos better, and more educational?

- *No comments.*