

Jed Dougherty  
STAT 4010

- 1.) I am one of two writer/editors with Adam Obeng. I met with the “Next Big Thing” team on Sunday and am pushing hard for interviews from class members with domain expertise. I am editing drafts for both articles as well. Finally I am interviewing Chris Wiggins.
- 2.) P you really hit the nail on the head. The NYT, every so often, feels the need to spout off articles from this real estate developer perspective that are just ridiculous. (If you are wondering why, check out the amount of real estate ads in the Saturday paper.)

This is an issue that ties together the definition of community, how data science can be used to track the framework and interaction of communities, and who is left empowered or disenfranchised by data science.

There are opposing views on ROC. From the perspective of the real estate developer or city government, ROC is largely a factor of property value. This isn't an entirely bad thing. Higher property values often go hand-in-hand with infrastructure development, increased public space, and new business creation.

On the other side, there is a localized view of ROC that is very different. In this view, ROC would be largely a function of length of stay and network diameter. A long average length of stay and an short network diameter imply a deep-rooted bond within a community - even if that community is not necessarily massively revenue generating for the city or developers.

Often - as you mentioned in Las Vegas - cities destroy communities with a high latter level of ROC in order to encourage the creation of communities with a high level of the former. It is funny that the article uses Brooklyn as a model. Of course the Brooklyn the author of the article is talking about is really the few costal neighborhoods that have boomed in recent years at the expense of the prior residents - places like Park Slope, Downtown Brooklyn, Fort Greene, and Williamsburg.

Ever since the Supreme Court Case Kelo V. New London in 2005, cities have been able to enact eminent domain as long as the community enjoyed economic growth as a result. This means that cities can now enact eminent domain without taking into consideration the second ROC view - and indeed NYC has repeatedly enacted it in conjunction with Forest Ratner in Downtown Brooklyn. Basically it works by NYC kicking out the former (read poorer) residents and allowing development companies to build high rises in their place.

Finally, data science ties into all this by empowering those who are inputting data into the system. That is - it empowers the economically well off who use credit cards and smart phones, who have bicycle rental stations in their neighborhoods, who use the internet every day, etc. We see the data of the well off, while the poor disappear even more than before. Thus when we analyze the data, we ignore those who are already being disenfranchised by it.

3.) Kaggle Username: JedIV  
Kaggle Teamname: Team Edward

4.) Kaggle Visualization showing estimated output As we can see there is a lot of work still to be done on tightening up the data:

