Who is your client / user / audience ?
Broadly speaking, people who are interested in text analysis, however more specifically people who are economist readers/subscribers.

What is the central question your project addresses?
The Economist is well known for its editorial anonymity. The reasoning behind this unique practice is to channel a collective voice. Is it possible to identify the authors of most/certain economist articles solely using the actual texts?

What is theme of your project?
Can one identify writers of articles based on their texts.

What would make a good, informative title (or subtitle) for your project?
Economist authors REVEALED!!!!

Write a prioritized list of more specific sub-questions, if applicable, or an elaboration of the central question.
I'm looking at the blog posts in each sub-section i.e (American politics/Middle East) which do have the initials of the authors. Then based on how well I can predict those authors, I'm then finding the corresponding sections in the magazine and predicting those authors.
If the models aren't good in general, I might look at specific authors or even look for specific phrases that appear with only one author to predict certain articles.

a. What data will you use to address the question, from which source(s)?
Various blogs on the economist.com, I started with the american politics section because it had the most posts in the last two years.

4.. b. How will you acquire the data?
Web Scraping (selenium)

4.. c. How much data will there be?
I'm going back two years for each blog, ~500-100 blog posts and somewhere around the same number of articles in each section

4.. d. How will you organize and store the data?
MongoDB, I'm storing the headlines, texts, and dates for the articles, the same plus the author initials for the blog posts.

How will you approach the analysis of your data?
(This can be a complete plan or a description of planned first steps and conditional later steps)
Building models for each section, then use those models on the economist articles
Going to be using Naive Bayes Model and maybe Logistic Regression, it all depends on what works best empirically.

What will be the form of your final deliverable? How will you present your findings?
Anticipated formats are blog post or dashboard.
Not sure of a good dashboard, but after discussing with you, I might have a better idea, if not a blog post where I go through the process of what I've done, optimally both.

What challenges do you anticipate?
(eg: Where might surprise challenges be lurking? Are their any weak points in your plan? Are you counting on something coming out statistically significant? Are you using something fuzzy like sentiment analysis to create sub-populations? Can you think of a failsafe? Something interesting to do or say if your "gamble" comes up bust?)

Initial challenge was whether I can predict authors at all. I started with the US Politics section and got an accuracy score of .8 with five authors, much higher than anticipated. My big concern is external validity. I'm assuming that the same authors who are writing blog posts are writing articles, as well as the texts of blog posts and articles being similar.