I wanted to present this topic because it is something that comes up frequently at my workplace and I’m starting to think about how I am going to handle this issue in the future. Our shop is relatively small and the analytic demands of our organization is staggering. To keep up with the demands, we have some cases where we send out data to other firms to process such as an identify management solution offered by one of the credit agencies (not the one in the news thankfully). But in the case of the identify management I came across this great python library called dedupe <https://dedupe.io/developers/library/en/latest/How-it-works.html> that might have worked for us or at least got us closer to our goal with some modifications and would likely have been an amazing learning experience.

At the same time, there simply isn’t enough hours in the day for each of use on the team to build out custom products for each of our stakeholder’s needs. Besides the dedupe option could have been a spectacular failure in which no amount of work could have done better that the credit agency, which also has a lot more data on customers. So, in these types of scenarios I’ve been finding the path to a decision challenging to work through, particularly since I started this program to build data products and each time we pass on something it feels like a lost opportunity. As a side note if anyone wants to use dedupe, this use case came and went before I got a chance to test the library so I’m not entirely sure how accurate this solution would be for your use.

One thing that stood out for me in reviewing articles on this topic was in a piece by Blue Triangle Tech - <https://www.bluetriangletech.com/blog/data-science-build-vs-buy-dilemma/>. They mentioned a mental trap - “When we talk to companies, frequently their data scientist will chime into the conversation and announce, “We build all our own tools” or “I have tools, big data structures, and can develop what you do”.” I liked this wording because I often want to also say the same thing but it’s clearly not an attitude to take into every scenario. I also thought this quote from Dataiku’s blog - <https://blog.dataiku.com/data-science-at-scale-make-or-buy-in-house-or-outsource> touched on the same issue in a great way - “But companies need to be careful, as hiring underused specialists in-house can get as expensive as outsourcing”. Having data scientists build everything is simply a bad strategy, there are somethings that are just not worth having a senior data scientist work on.

I’ll add a few more articles for and against the debate, but I did find that it was a lot easier to stumble on articles against building analytics since a lot of them are written by firms offering analytic solutions.

1. The first use case in this article explains a communication breakdown between an analytic supplier and the customer which was a Canadian Bank, the customer wasn’t aware of the analytic solution nuances and used the results incorrectly. While this could happen for anything I would think that if the solution was built in-house this kind of nuance wouldn’t have been missed.

<http://www.kdnuggets.com/2012/06/business-analytics-failures.html>

1. I felt like this article had a great balance of for and against arguments for software but it just as easily been about data products.

<https://www.forbes.com/sites/chuckcohn/2014/09/15/build-vs-buy-how-to-know-when-you-should-build-custom-software-over-canned-solutions/#72a8102fc371>

1. Against building analytics but from a firm offering analytics

<https://amplitude.com/blog/2016/04/21/thinking-building-analytics-dont/>

I really am interested in everyone’s thoughts on this topic, especially in any personal experiences that you would be willing to share. In case you don’t have any experiences on the subject I throw out a few questions but whatever comes to mind please feel free to post. What are the analytics that you wouldn’t mind passing along to another firm to develop? What are analytics you wish you could do but know you will have to pass along because either a firm can do it better/faster/cheaper or it’s just too basic for your skill level? What kind of things do you think you’ll need to take into consideration at your current position when it comes to a build or buy scenario?