Who did the most with the least in Sochi?

Никто не верит олимпийскому девизу “главное участие”. Все знают, что главное - медали, по возможности - золотые. Эффективность команд на Олимпиаде - обсуждаемый вопрос (типа “Вдогонку посту на USAToday и etc etc”  Неофициальный “командный зачет”, сайт о популяции и GDP, куча сайтов о вознаграждении олимпийцам. Кто же прав?

Существует масса литературы по DEA. Мы используем ее чтобы посмотреть кто же на самом деле был лучше всех на Олимпиаде и насколько эффективность коррелирует с вознаграждением.

**Что такое DEA и почему ее все используют?**

Обычно, эффективность измеряют как простое отношение одного выхода ко входу. Если использовать различные входы и выходы их нужно (а) взвесить, (б) компоненты числителя должен быть одной размерности, (в) компоненты знаменателя должны быть одной размерности. Этого можно добиться введя цены и сделав предположение о важности (пропорциях) inputs and outputs. Все это вызывает множество вопросов - правильны ли цены? правильны ли веса? Но можно использовать DEA, и тогда большинство вопросов отпадут сами собой.

Существует множество сайтов, посвященных DEA, есть [определение](http://en.wikipedia.org/wiki/Data_envelopment_analysis) и в Wikipedia. Вполне удачно, по моему, высказался [Vaninsky (2013)](http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1130&context=numeracy): “DEA evaluates the performance of functionally similar objects by weighting  outcomes against inputs.It uses Linear Programming (LP) to assign an efficiency score scaled between 0 and 1 to each object in the group”. Преимущества DEA в том, что это

* непараметрический метод, не требующий предположений о том, каким именно образом inputs трансформируются в outputs.
* может быть сколько угодно inputs и outputs, причем они могут отражать конфликтующие key performance indicators
* inputs и outputs могут измеряться в любых единицах: мы можем “сравнивать” kilograms/kilometers/dollars/population, и т.п. - все что можно измерить, в любых комбинациях.
* неэффективным единицам показывают, сколько осталось до полной эффективности, за счет каких именно inputs/outputs они неэффективны и кто ближайшие бенчмарки

Это и определяет чрезвычайную распространенность метода. В Wikipedia есть простой [пример](http://en.wikipedia.org/wiki/Data_envelopment_analysis), иллюстрирующий DEA на примере трех DMU с двумя inputs и одним output. Прекрасная реализация DEA в R с детальным разбором есть у [Pessanha, Marinho, Laurencel, do Amaral](http://professorjf.webs.com/DEA%202013.pdf) (2013). Но это для любителей verbatim. Мы же воспользуемся R project, пакетом Benchmarking и несколькими строками R кода чтобы ответить на вопросы, поставленные в первом абзаце.

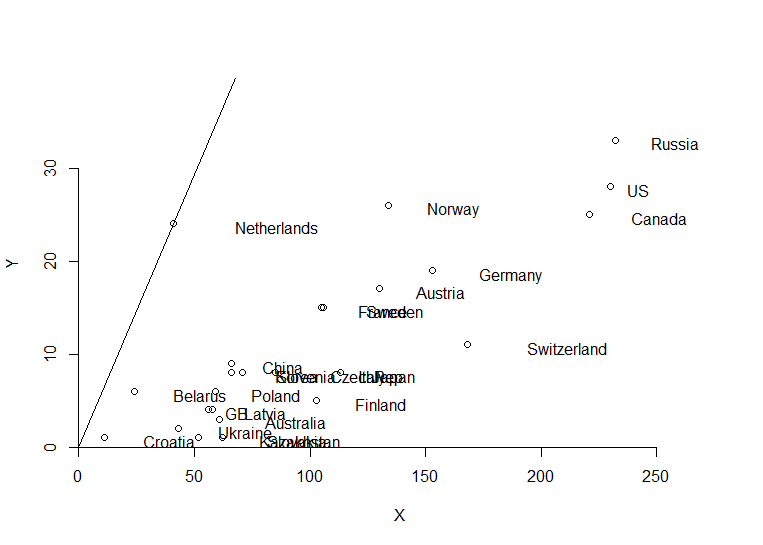
**Простой пример: does size really matters, finally, and how exactly matters**

USA Today [says](http://ftw.usatoday.com/2014/02/olympics-medal-count-russia-sochi-united-states-most/) “the Dutch did the most with the least”, предлагая измерять эффективность команды, используя только одно отношение - medals to athlets count. Putting key output in enumerator and key input in denominator is an old trick. For example, in ancient times investors deciding on how high market pays for a stock, used to put market price as enumerator, and nominal, or book, as denominator (Graham, 1949), thus yielding famous book-to-market ratio. Tolstoy () suggested to measure the value of an individual as ratio: "A man is like a fraction whose numerator is what he is and whose denominator is what he thinks of himself. The larger the denominator, the smaller the fraction".

Farrel () noted, that this approach could be extended to more general case, which was subsequently developed as “data envelopment analysis” in seminal paper by Charnes (), which in turn eventually became a hunting license for taking any set of anything, allowing principal unification, and calculating it’s relative efficiency by dividing multiple weighted outputs to multiple weighted inputs.

There are numerous variations of DEA, and later on we will use several, but first we’ll try to implemetn and visualise the most basic one, which - surprise! - coincides with approach described in USAToday article.

Let’s plot all teams on 2d space, where OX axis is responsible for athletes count, while OY - for medals count.

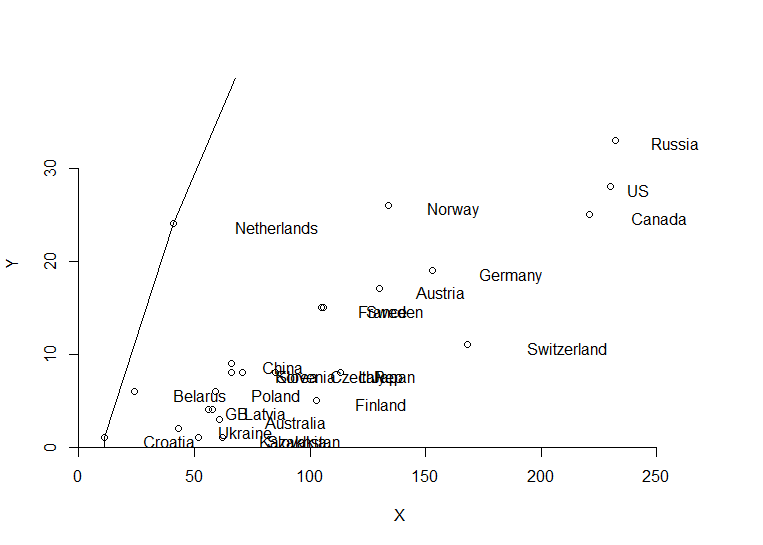


This case perfectly fits USAToday’s logic. In this setting the team with the highest ratio of medals to athletes will be the most efficient (Dutch in this case). All other teams are lying to the right and lower than straight line connecting zero and Dutch. In DEA they say that Netherlands are dominating all other teams: no one is better than they are, i.e. nobody is situated to the left of the line. The line here represents “best technology”, currently available to any olympic team: the team, lying on the line (hence possessing the same “best technology”) would have the same ratio as Dutch team, and hence, will be at least as efficient as they are.  This also should be quite familiar: in fact, any linear function Y(X) of the form y = 0 + k\*x is in fact a law (which could be represented visually). This law describes how X is reflected into Y (or how Y is produced by X).

But there is several significant flaws in such approach. First is that it implicitly assumes, that if an efficient team would, say, double its athletes count, it would automatically double medals output. If, on the contrary, it would scale down it’s size, the outputs would be scaled down in linear way.

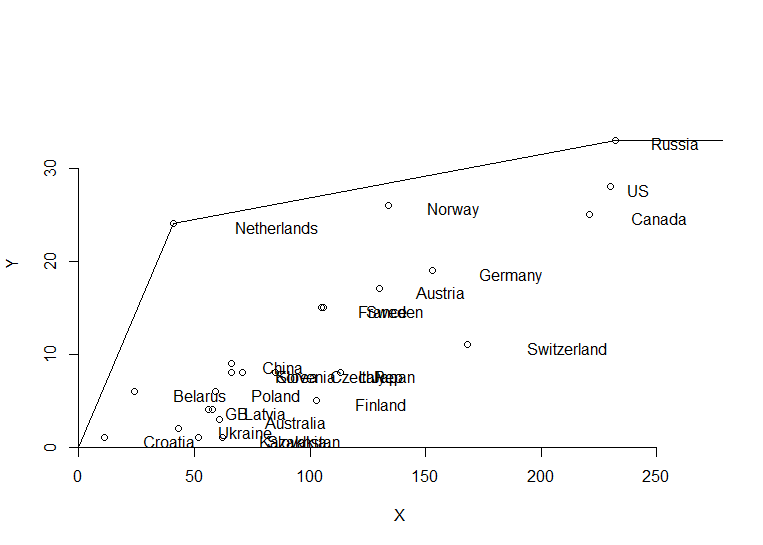
It stands to reason that in reality that’s not always the case. For example, we may speculate that the smaller the team, the harder it is for it to win at least one medal. First, small team has less opportunities to get medals in team sports (like hockey), or team events. Second, single athlete in a small team is exposed to more pressure as tactical support from other teammates performing in the same sport is reduced (which is big deal in disciplines like short-track).

DEA could elegantly solve this issue by introducing various forms of return-to scale effects. Skipping the math (if you prefer not to skip I would recommend perfect introduction in topic in Bogetoft, Peter, …) let’s write the following lines in R:

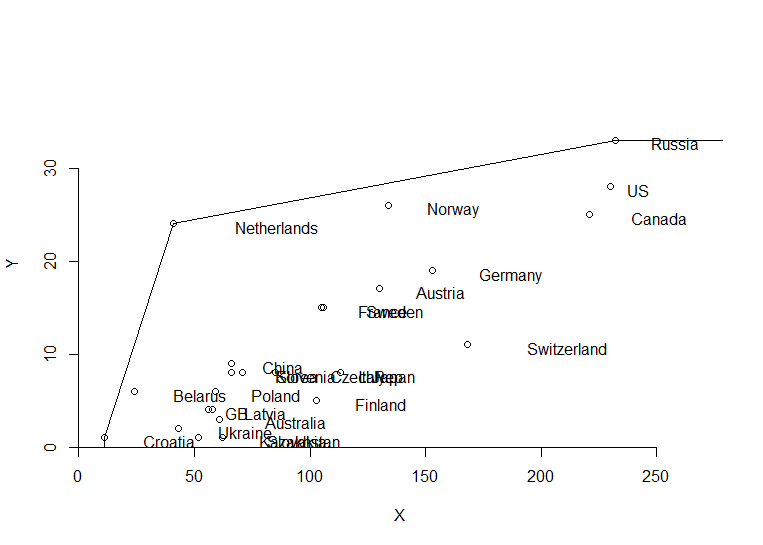


As you see, under increasing return to scale (i.e. when bigger teams have more than linear payout to its size) small teams would have more chances to be efficient, and the number of efficient teams is increasing. In our simplest case, Croatia shoots ahead other small teams, thus becoming new champion in addition to Netherlands. Croatia took only one silver medal, but it was the smallest of all teams with at least one medal in Sochi.

Somebody would object that small teams are in fact in better position than large ones. First, in small team athletes are less likely to be involved in team sports (like, again, hockey), consuming dozens of athletes to get just one medal. Second, participation in small team puts more personal responsibility on the result. Athlete even could display personal heroism, participating in several disciplines, knowing that “no one except him” etc., etc. Last, but not the least, in small team athletes are more likely to be treated on individual basis by supporting team and trainers. DEA could take this into account by introducing decreasing returns to scale (DRS):



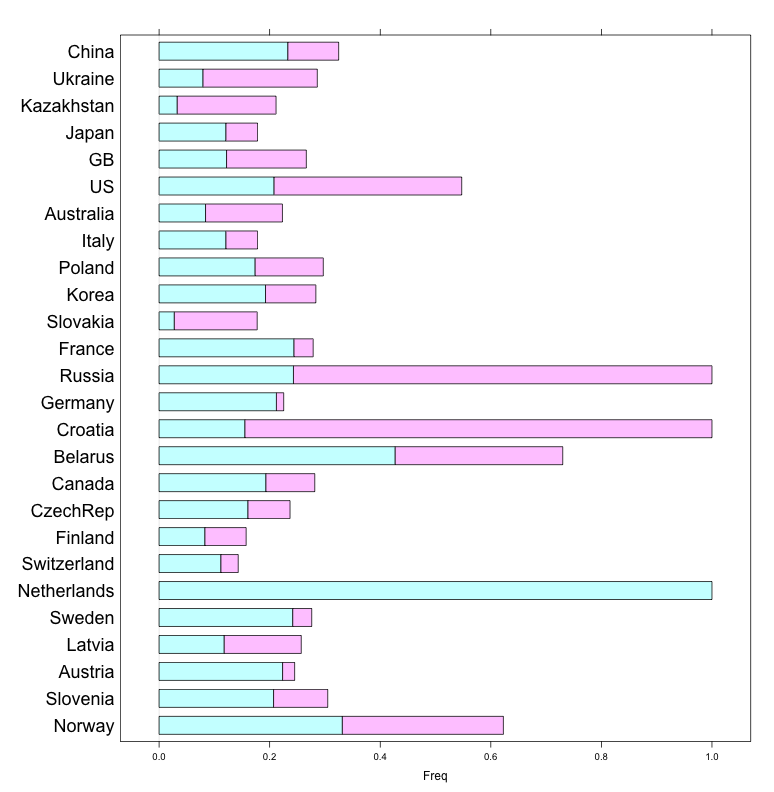
Under DRS Croatia falls back where it was under CRS assumption, but instead we have a new king - Russia, which held in Sochi an absolute record not only in medals count, but also in athletes count.  Elephant team makes Russians lag far behind Dutch under CRS assumptions, but under DRS it makes all other “elephants” (USA, Canada) with handicap. We also see improved efficiency of Norwegia, Austria, Germany, China, etc. They had quite big teams too and perhaps faced DRS effect due to reasons discussed above. Still all of them are not on the efficiency frontier, although Norwegia improved so much that it is almost efficient (which is not surprising given the second place Norwegia had in total medal count).

DEA имеет возможность сочетать IRS and DRS assumptions under VRS assumptions. Мы опять опустим the math magic fueling it, просто покажем картинку

Этот вариант позволяет маленьким командам иметь свои характерные преимущества, при этом оставляя большим командам их преимущества. В этом демократичном варианте эффективными становятся и Нидерланды, и Россия, и Хорватия.

**Improvements due to scale effect**

After we allowed to small abd big teams to be efficient due to scale, we want to know just exactly how much everyone improved its results due to scale effects. I.e. what part of VRS model improvement relative to CRS model is provided by relaxing CRS assumptions and allowing teams to enjoy IRS/DRS benefits.



**A little math formalism: on free disposability and convexity**

Сравнивая страны между собой по одному KPI, как это делает USAToday, мы фактически сравниваем “технологии” перевода чемпионов в медали с лучшей (Dutch). В DEA мы сравниваем команды не между собой, а с наилучшей технологией (сплошная линия на рисунках выше). Мы выбираем эффект масштаба и говорим, что теперь команде нужно или уменьшить количество атлетов или (например, копируя практики лучших команд) сделать так, чтобы они зарабатывали больше медалей. Теперь мы знаем кто на самом деле did the best with the least.

Выше я говорил, что DEA делает минимум предположений. Но какие? Мы уже определились с эффектом масштаба, поэтому остается еще два: free disposability and convexity.

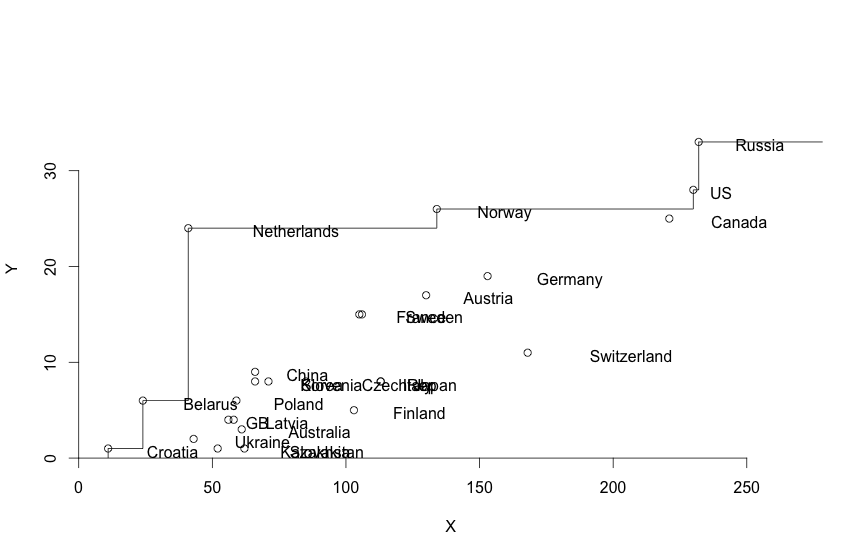
Let’s define input-oriented efficiency of olympic team consisting of X athletes earning Y medals, as the smallest factor e by which we can multiply the input X (number of athletes) so that E athletes still can bring home Y medals.

E(Athletes, Medals) = min{ e | e \* X can produce Y }

Определяя эффективность таким образом, мы неявно предполаем, что для любой страны возможно увеличить количество атлетов и при этом уменьшить количество получаемых медалей. Sounds crazy, но если подумать, есть случаи когда это невозможно (скажем, мы не можем увеличить количество туристов в Sochi и одновременно уменьшить количество GHG emissions, производимых для их комфортного проживания там). If we speak about unwanted outputs, they are not always freely disposable, but luckily this is not our case.

Another assumption is convexity. We call a technology (solid line) convex (and this is indeed the case) if any mixture of two input/output combinations is also a feasible combination. It’s like saying that if I’m Austria, and I don’t need to change the size of my team, but I want to be effective, I need to take some Netherlands times W, and then some Russia times (1 - W), and then sum it and enjoy my effectiveness. We even can calculate the W. Sounds even more crazy? Well, in fact, sometimes technologies are indeed convex. We can think of it as of portfolio of “best practices”, whatever it might be. Then W would measure relative weight of best practices of best teams that should be in portfolio of any team of given size, that wants to be effective.

But can we combine this practices in linear “weighted” manner? Maybe yes, maybe not - at least we have to be conscious about our assumptions. And what it would be, if we can not combine it? I.e. if any team is unique combination of athletes, having unique structure, unique psychological climate, unique mix of superpowers, interacting in a unique unreproducible way? Then we have to stick to free disposability hull assumptions, and relax convexity assumption. What would change then?



Relaxing convexity would yield even more new champions: we have Croatia, Russia and Netherlands, plus Belarus, Norway, and US. The more relaxed we are in our assumptions, the more efficiency we allow to teams.

**How much athlets could stayed at home**

The equation (1) could be formulated it in output-oriented form, by saying that for a team it is possible to leave at home (1 - e) \* X of the athletes and still earn Y medals.

So now we can calculate how much athletes could stay at home, if all teams were efficient:

sav <- (1- dea(olympics$Athletes, olympics$TotMedals, RTS=’vrs’)$eff)\*olympics$Athletes

rownames(sav) <- rownames(olympics)

sum(sav)/sum(olympics$Athletes)

Striking 62% or 1607 of 2563.

**Более сложный случай: множественные inputs/outputs**

Подход с одним входом и одним выходом хорош своей простотой, и тем, что его легко визуализировать, при этом поиграв с эффектами масштаба. Но он является только оценкой partial efficiency. По сути, мы используем всего лишь один из множества возможных KPI. Насколько правильно использовать размер команды в качестве входа? Может, мы должны смотреть на экономическую мощь нации? Или оценивать, какой выбор у тренеров, выбирающих спортсменов в команды? DEA тем и хорош, что позволяет использовать сразу несколько inputs и outputs.

Я воспользовался данными (сайт по олимпийским играм) для того, чтобы расширить свой датасет. Теперь я могу использовать как inputs не просто количество атлетов в команде, но и размеры GDP и population of the corresponding country. Это позволит мне, с одной стороны, отразить экономическую сторону дела (wealth of nation and level of living impacts sports financing and level of involvement of population in sport activities), а с другой - генетическую (nations population size is connected with genetic diversity and probability of champion being born in that nation).

Одним из выходов по прежнему будет общее число медалей. Но я дополню output set with one more KPI, reflecting team “super-success”, by counting quantity of gold medals, earned by team, as additional separate output. Таким образом, “золото” будет фактически считаться дважды: в общем счете, и в своем собственном.  Множественные inputs/outputs уже не дают нам роскошь визуализации эффективной границы, но зато позволят увидеть новых чемпионов.

Under VRS we’ll have three more champions - Slovenia, Latvia, Kazakhstan - as well as good old Dutch, Russians, Croats, US, Belarus, Norway. If we would further relax convexity assumption, we’ll have a bunch of new guys: Austria, Sweden, Slovakia, Poland and China. Having big population and GDP not always pays off: the most inefficient is Germany, following Japan, France and Italy. They haven’t “earned” enough medals to compensate for relative prosperity.

**What about peers?**

Peers are best teams, to which our team should нацелиться, чтобы стать эффективной. Benchmarking пакет предлагает специальную команду peers() чтобы показать peers. К сожалению, он показывает только номера команд, поэтому чтобы получить названия стран, мы используем следующее элегантное выражение:

\

**Slacks: sources of inefficiency**

**Is success well-rewarded?**