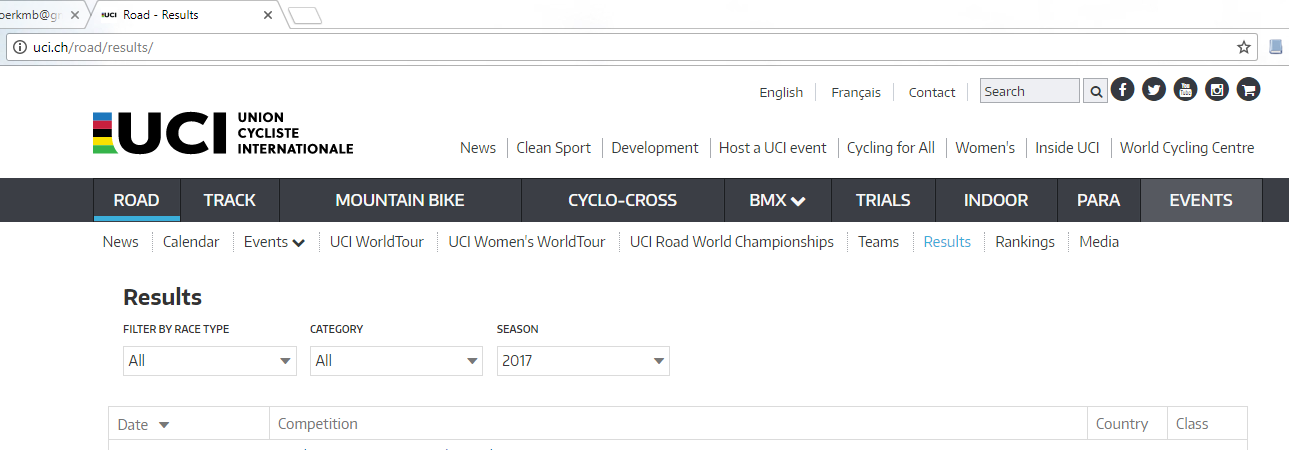
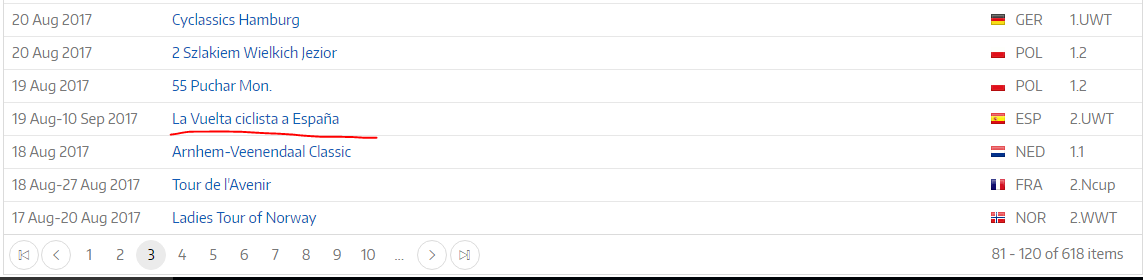
**Grand Tour Cycling Results Data**

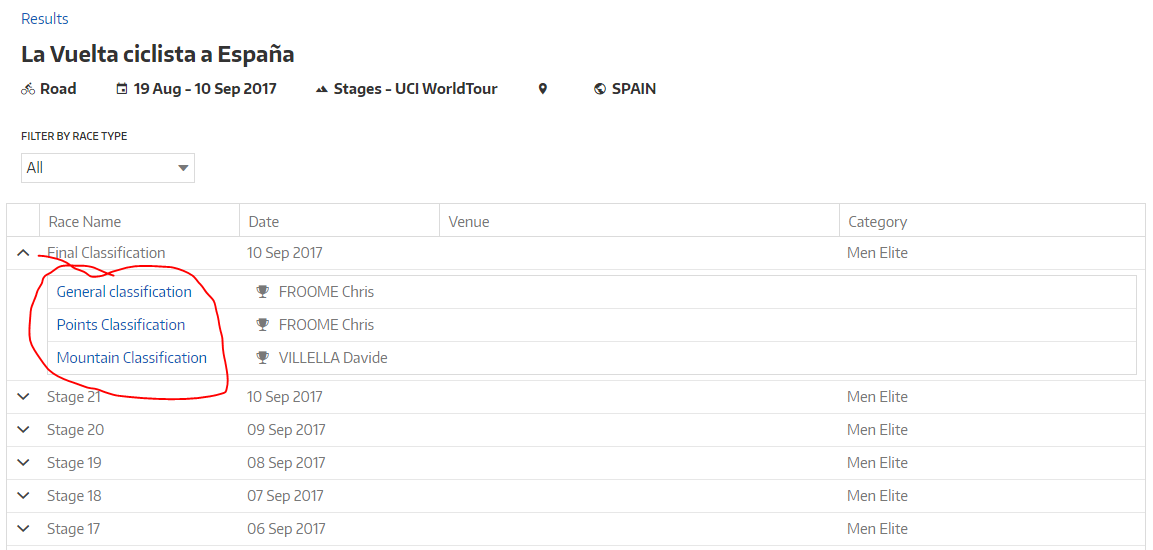
This is for a project on Moral Hazard in Cycling, where some authors found a rule change induced more risk taking, and where I’ve found another rule change that would induce a similar behavioral response. The idea behind getting this data is to see whether these races became more competitive and whether the size of time and point gaps between different riders (top 3, top 10, top 20 etc.), predict crashes.

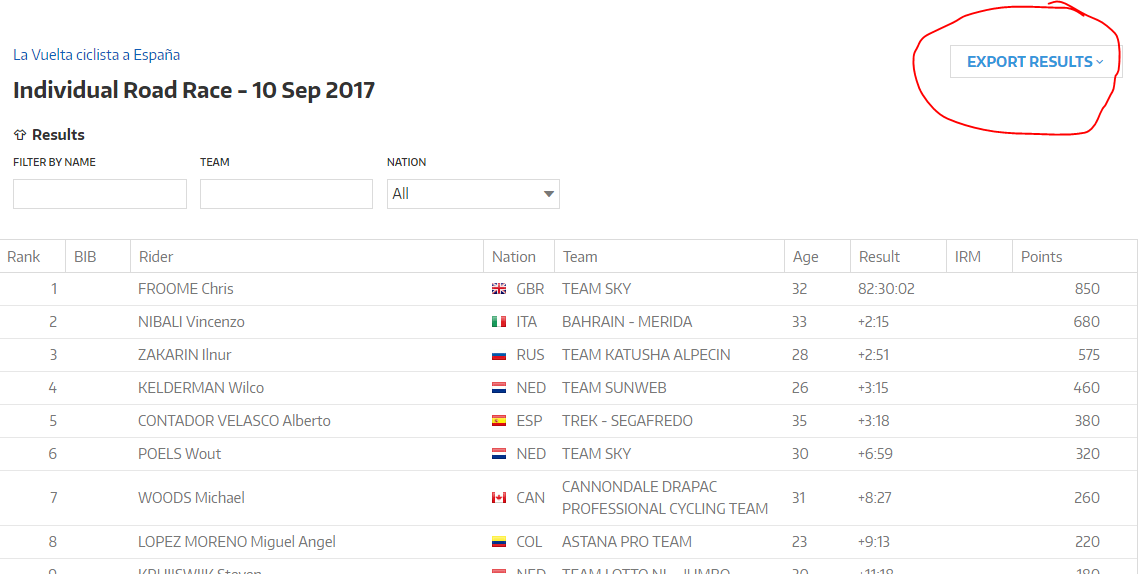
The idea is to gather the result files(mostly interested in 2 kinds of result files, but it may be easier to just get all and then I later throw away what’s useless), from every stage (20-22 per race) for each race (Giro D’Italia, Tour de France, Vuelta Espana), for all the available years (2009-2017). The website which hosts these is <http://uci.ch/road/results/> . As you already understand, if it’s not too complicated to write code to do this, it may be worth it, relative to manually downloading the, at least, 2x21x3x9= 1134 excel files. I should say, I assume the easiest way to do this is to download the excel files and append them later, if it happens to be easier to scrape the results I obviously don’t have a preference..

The website is pretty consistent, but there are sometimes more than 3 result files, and sometimes the race names change ever so slightly (unlike the snippets below, in 2009 the Vuelta D’Espana is simply named that..), and there are different number of total races each year which makes the races we care about appear on different pages.









Markus: So for every stage, I'd like the general classification and points classification.

The only thing, is that I know some of the older races lists the competitions kinda weirdly, so the easiest thing for you may be to write code to grab them all and then I can throw away the useless ones

k. That approach is easy for me, but is it less work for you?

the alternative is I look for only valid ones and you notice certain years are missing, then choose them by hand

That actually is a great call!

The general classification should be fine. They sometimes use the phrase points competition and sprint competition inconsistently. I suggest you set it to download Points competition and I can go from there