732A94 Advanced R Programming
Computer lab 5: API R projects

Krzysztof Bartoszek
(designed by Leif Jonsson and Måns Magnusson)

18 September 2017

Seminar date: **3 October 08:15** (P18)
Lab deadline: **6 October 23:59**

Below are some API that can be used as a part of the assignment to implement an R package to connect to an API. In connection to each API there is a short description and the difficulty of implementation is indicated using one, two or three stars with one star (simpler) to three stars (difficult). Some projects are about implementing XML parsers for more complex XML files of interest.

In general there exist many different web API:s and you can choose any API you may find interesting. Most web API:s in Sweden can be found in **API-katalogen** `(http://apikatalogen.se/api)`. If you choose any of these other than the ones listed below you need to confirm your choice with the teacher, the only restriction is that the API should not already have an R package.

*** Valmyndigheten XML** After each election the Swedish election agency (Valmyndigheten) release a detailed file with all election resuls in an XML format. Implement a package to get the election data (that is published as excel files) out of these files direct.

More information can be found **here** `http://www.val.se/val/val2014/statistik/index.html`.

*** Google geocode API** is used to do simple convertions of geodata between coordinates and adresses/zip codes. This is a relative straightforward API. More information can be found **here** `https://developers.google.com/maps/documentation/geocoding/intro`.

*** Kolada API** is an API to the Kolada database that contains key indicators regarding the Swedish municipalities. More information can be found **here** `http://www.kolada.se/appspecific/rka/download/api/Kolada-API-test.html`.

**** Riksdagen** or the Swedish parliment have an API to get documents, debates and information about members of parliment. Information about the API can be found **here** `http://data.riksdagen.se/`.

The data is mainly textual data so the R package should convert the data to the datastructures used in the `tm` R package.

**** Protokollen.net** is an API to get all the meting protocols from Swedish municipalities using elastic search queries (more information **here** `https://en.wikipedia.org/wiki/Elasticsearch`. As with the Riksdagen API the documents need to be stored in the `tm` package format. More information can be found **here** `http://www.protokollen.net/#/` and api information can be found **here** `https://github.com/rotsee/protokollen/blob/master/README-database-api.md`.

**** Språkbanken XML** is a web resource with corpuses in XML format. As with the other textual resources the purpose with this corpus is to read any XML file into R as a `tm` type corpus object with all textual metadata included. The corpuses can be found **here** `http://spraakbanken.gu.se/swe/resurser/corpus`.

***** Thenmaps** Historical map in geoJSON format. Need to handle spatial data, see spatial.ly for more information on geographical data. More information can be found **here** `http://www.thenmap.net/`.