Task description

The RACF currently produce a fuel factsheet (https://www.racfoundation.org/wp-content/uploads/RACF_fuel_fact_sheet.pdf) on a weekly basis. This is designed by our design firm (Javelin, Nick is cc'ed), and each week, they make edits based on a spreadsheet of new numbers (see attached) which we send them.

This spreadsheet of new numbers is produced by an R script, which we aren't happy with. The unit testing is buggy and overly-complex. The layout of the emailed spreadsheet is too complicated (Nick is very kind and says it isn't, but ignore that).

We would like this first (outbound) script re-made, keeping some elements where appropriate (downloading data, emailing, cropping images, etc.). It would:

1. Download the latest raw fuel data (all on Google Drive) and conduct unit testing to ensure it is correct. Stop if unit testing fails.
2. Do the necessary calculations to provide editing instructions (list of required edits included attached) for our design firm.
3. Create a list of edits (copy-able for the designers) and email that list to the designers (code provided). This could be email text and attached tables of data, etc.
4. Save this list of edits to an existing folder structure.

In addition, we would like a second script (inbound) produced. This would "read" an edited PDF and then:

1. Determine whether the edits had been made correctly. This would be a largely automated process.
2. Email any un-done or incorrectly done edits back to the designers, comparing the PDF to the list of edits originally sent.
3. When all edits appear done, email the PDF to a member of RACF staff with a message of approval.
4. When all edits appear done, email the designers with a message of approval.
5. Crop a series of images from the PDF (code provided) and email them to an RACF member of staff.

I do not expect PDF "reading" to work immediately and without trouble. We would want you to provide a list of unreadable elements of the PDF, and then RACF would work with designers to produce alternatively designed elements which could be read by machine. So there would be some iteration of the design until you have a readable PDF to work with.