Analysis Journal—RAC Parking Report Automation

mz

Saturday 9.3.2019

- Initialised public github repo.
- Setup folder structure, currenlty looks like this (output of `tree --charset=ascii -d`):

```
|-- code
|-- data
|   |-- 01-raw
|   |-- 02-interim
|   `-- 03-processed
|-- docs
|   |-- admin
|   |-- journal
|   `-- original-reports
|-- figures
`-- outputs
    |-- instructions
    |-- reports
    `-- technical-appendix
```

- available reports used for development, downloaded from RAC press website, pdfs stored in `/docs/original-reports`, see appendix for full table

Next steps: find data

Thursday 14.3.2019

- Look through back reports available, saving all and saving links here, in appendix 1
- Start query file for Ivo.
- OK, got all reports, despite missing years, it looks like most of the data is included in other reports as they have a window of usually 4 years previous that they include data for.
- Now figure out data sources for England, and detail them
- Outline typical England report

England data requirements and availability?

- 2017/18 uses "MHCLG tables on parking income and expenditure"

This looks like the data here.

- For each year there is normally budget, provisional outturn and outturn data, the latter needs to be for individual local authorities, which does not seem to exist for 2007/08, but is available for all other years.
- The files are grouped under the title *Local authority revenue expenditure and financing England: 2010 to 2011 individual local authority data - outturn* for each year

- The relevant Excel file is usually called `Revenue outturn RO2 (highways and transport services) + year.xls`.
- There are three columns we are interested in: expenditures, incomes and capital charges, these are usually on separate sheets. There is one of these for both on street and off street parking. So that's 6 columns in total.
- There are also summaries, but I don't need them, maybe only for manually double-checking.
- Appendix 2 lists all the files, their links, the number or rows (local authorities) in the table and the sheets on which the data can be found.
- TODO Add the starting row and the actual columns where the data in question can be found and I can write a funciton to extract the data automatically.
- Download all `.xls` files to `/data/raw`
- "Earlier, councils also submitted budgets for the 2018-19financial year but on a less detailed basis." where are these budgets? Do I need to include them as well? Ah, yes.
- At one point the England data stops including capital charges..2013/14 is last one. This means current reports are less data rich, and if I stick to the brief, it means that the code won't be able to replicate the old reports.
- At one point the Englad data starts including penalty charges for individual Local authorities, although the reports don't do anything special with that, just use the aggregates.
- England totals do not include the national parks and the Nottingham levy. Not sure what to do with that!?

Friday 15.3.2019

- OK, so England outturn data is tabled in Appendix 2.
- Now England Budget data
  - not in *Local authority revenue expenditure and financing tables*
  - under *Budget estimates of local authority revenue expenditure and financing for the financial year* the table *Revenue account budget (RA)* has the parking surplus down as 792, but the report says 782. There have been a few other differnces, not sure if these are typos/errors?
- download data into `\data\raw` and fill out appendix two table with relevant info.
- careful because at one point there are capital charges and then not. So make sure you've got the correct column. Because these are sums, I would even set up a check?
- OK, England budget files are all saved and relevant info extracted manually.
- Now back to outlining the England report as it stands.
- E17/18 says national park surplus is 1.7 million, sum of cells DF405:411 is 1.774.
- Comparing Table 1 for E17/18 to excel file 17/18:
  - on street parking is all the same
  - off street parking: there is a discrepancy. Income is the same, but Leibling's expenditure value is higher than the Excel one (making the surplus lower): 362 vs 356?!

Saturday 16.3.2019

- Comparing Table 1 for E17/18 to excel file 16/17
  - on street parking is all the same
  - off street parking: there is a discrepancy. Leibling's income is 693 vs 689 in the table, and Leibling's expenditures are are 349 vs 343 in the table.

- Comparing Table 1 for E17/18 to excel file 15/16
  - on street parking is all the same
  - off street parking: there is a discrepancy. Leibling's income is 670 vs 674 in the table, and Leibling's expenditures are are 340 vs 337 in the table.
- Trying to figure out where these discrepancies come from: It must be the nottingham workplace levy? Which I don't have data for.
- Second thing, presumably the difference is also due to not including the national parks. Although I have the data for that, I cannot confirm, because I don't know the levy numbers.
- TODO write (find) a funtion to convert excel column names to column numbers.
- Oh, congestion charge data needs to be read as well, that's in a different couple of cells than what i've been looking at until now, so add that to the appendix.

Monday 18.3.2019

- Clean up queries for Ivo for England.
- Look at Scotland data. 2015/16 is the last year with a report but the data on [this link](https://www2.gov.scot/Topics/Statistics/Browse/Local-Government-Finance/PubScottishLGFStats] seems to go to 2016/17
- I'm looking at `Annex A by LA` excel files, which have LA level data.
- OK, so data checking: compare surplusses in individual sheets of `sco-15-16.xlss` with table 9, parking surpluses in report 15/16.
  - Edinburg is same, so is Glasgow, Dundee, Renfrewshire, Fife, South Lank, East Ayr, Highland, Argyle, Aberdeenshire, Angus, South Ary, Moray,Perth, Stirling,
  - Aberdeen city is 4.89 vs 0?
  - If I add Aberdeen city form Leibling to the Excel totals get the Leibling totals for Surplus.OK
- But if I add the Aberdeen city incomes to the Excel ones, it's still not enough. And same for expenditures... So there is another LA that has a discrepancy.
  - Clackmanshire discrepancy: .014 -> .02, but that's minor, could be poor rounding. + South Ayershire discrepancy: .0 -> .80
  - And of course Aberdeen city: .0 -> 9.20
  - This fixes the total income discrepancies.
- OK, now for the expenditure discrepancies:
  - Aberdeen city: 0 -> 4.32
  - South Ayershire expenditure discrepancy: -.327 -> .470
  - ANd these now also add up.
- Start Scotish data query file for Ivo
- Start outline of Scotland report based on *newly discovered* 16/17 report
- Penalty notice charge data is apparenlty out of a scanned pdf in 2015/16.... And 16/18 data in a regular pdf. But perhaps the tabulizer package could help me out? Unfortunatley there are some sort of java issues in installing the package...
- solved issue using this link, installed rJava, but first had to also run `sudo R CMD javareconf`. This is presumably only an ubuntu issue, so should work fine on Windows. * Tried tabulizer on the 17/18 pdf and it works!

PNC data

Decriminalised Parking Enforcement - Local Authorities - Income and Expenditure

- pdf's exist for 2016/17 and 2017/18. Additionally a scanned pdf exists for 2013/14/15/16. All have:
  - a table of which LAs are doing DPE, and which are considering it etc, this is in the report

  - a table of PNC incomes. However this last table also has other incomes and expenditures in it, so you'd think it would match with the gov.sco data in the Excel tables, but it doesn't. SOme are quite close, e.g. Glasgow, but others not.

- Added table of PCN data to the appendix, created table to summarise the sources.

*Todo: outline scotland data.

Tuesday 19.3.2019

- Found the 16/17 scotland files, they're not under press, but under publications...
- Add them to app1
- Actually clean up the data source tables and move them to code, these are manual data entry files.
- Saving the tables into rds? Probably, it is serialised, and as opposed to csv it retains the data types. Let's me cleanly load individual objects and assign then names, and there's no danger of overwriting an object of the same name. But is not human readable.
- This also involves using here::here() to load the data in the child .rmd documents for the appendices.
- **mapping England**: OK, so the data is for 353 "local authorities", or are they actually "councils"? Shapefiles seem to have 326 shapes. Wikipedia says there are 27 county and 201 district councils, *which cover the same physical area.* plus 55 UAs, 36 met burroughs and 32 London ones, plus the city, plus scilly. But they overlap..
- SO e.g. you have Oxford council and Oxfodshire council. The first only does off street parking, while the county one does both. Do I add them up? If I want to map them I have to. But in the tables are they separate?
- OK, found lookup table for counties and districts, that will be useful!
- There are counties in the lookup table that are not county councils though, these eight: Greater Manchester, Inner London, Merseyside, Northamptonshire, Outer London, Tyne and Wear, West Midlands, West Yorkshire.
- Download UK map shapefile: *Local Administrative Units Level 1 (January 2018) Super Generalised Clipped Boundaries in United Kingdom* this link
- install `sf` package.
- Get basic maps to work, hooray.
- Probably want ultra generalised, this is too fine. Yeah.
- OK, now back to outlining Scotland data.
- OK, so there is crossover in the scottish report, and requires picking up data from the other two reports...
- Check conditional formatting of tables that work in pdf. `condformat` might work. Do they want it?
- Actually cell_spec in kableExtra probably does the trick see here.
- OK, now queries for Ivo.
- I also now have a much better oveview of how much manual tweaking the reports will require (not too much), so I have the idea of highlighting all the spots that need a manual double check (e.g. in red) so that nothing like that gets accidentally published.

- Add the four Nottingham workplace levy files to data. Need to try if extracting the tables works there as well.

Wednesday 20.3.2019

- OUTLINE the whole project scheme.
- Try extracting the WPL data. No, it doens't work. There are two other tables on the page and this table is too small I guess, or sth, so it will have to be manually input.
- Actually, I need to have a look at Wales quickly before I do this. For 2017/18
- OK, summary looks cool: income, expenditure and total transport as well, all match
- Participation in the *Wales Penalty Processing Partnership* and councils where "on-street parking controlled by Gwent Police who are planning to transfer it to the councils" probably requires an external source, this is in a map legend.
- Population data is available on same platform on this link
- There is additional penalty charge notice data, but that seems to be two years behind i.e. the 17/18 report has data for 15/16
- source for PCNs is here
- Unfortunately `tabulizer` does not extract the tables properly.
- meanwhile have a look at importing json objects..
- OK, so figured out how to extract the json file, and even the table, but that's only the first 1000 records. In order to get what I want I need to query it.
- hmm, ok, massive issues with fromJSON giving different results each time. If i don't figure it out, then it will have to me manual downloading of csv files. Not the worst thing in the world, it's still probably better than the ENgland and Scotland data..

Thursday, 21.3.2019

- created json-reprex.Rmd file, tested by Anneka and Martin as well, all get same non-deterministic results. Posted on rstudio community, tomorrow will crosspost on stackoverflow if there isn't any solution.
- OK, start Wales outline including replies from Ivo.
- worry about tables being too large. this stackoverflow answer seems useful
- OK, Wales outline complete.
- Now back to plotting out the whole thing.
- Download Wales data. csvs directly from the wales stats website.
- cross-post the question to stackoverlfow

Downloading Wales Data

- Export to .csv, without headers, without metadata!
- Gross Expenditure X Parking of vehicles -> wal-exp-17-18.csv link
- Total Income X Parking of vehicles -> wal-inc-17-18.csv link
- Net current cost X Total transport planning, highways, rows and transport -> wal-trans-17-18.csv link

Consolidating all data

- Start new file `02-import-data-original.R`
- Design table that will accommodate all 3 countries.
- figure out how to rename all files in a folder
- `for file in *; do mv "$file" "$(basename "$file").orig"; done;` this didn't work
- `rename 's/\.orig$//' *.orig` this fixed it.
- haha `for file in *; do mv "$file" "$orig.(basename "$file")"; done;` also fucked it up
- christ, anyway, now manually fixed..

MASTER TABLE

- Start by listing all variables required
- Start by importing Wales data
- Decide that the year variable will simply be the first year e.g. for 13/14 it is 2013. Anything can be done with that if needed.
- Wales data done!
- Scotland Excel files. OK, so looks like cell-specification for readxl cannot handle noncontiguous cells.
- Also cannot handle single cells - they are treated as the column header.. just leads to ugly code.
- Now to loop everything it would be nice to have all the data in the meta summary tables, including the year (not fiscal year) and the file name. So I'll fix that
- A, ffs, the name of the local authority also changes cell, and in 16/17 doesn't even have its own cell
- OK, it's not pretty, but it's working: Scotland income and expenditure original data is done!
- Next up: Scotland PCN data

Friday 22.3.2019

OK, back to the Scotland PNC data scraped from the pdf

- Becasue the 13-16 data is aleady transcribed in the old Leibling reports, i can try using tabulizer on them instead of dealing with the scanned pdf file. Actually, no they're not htere, and it isn't even important.
- Next the PCN tables with tabulizer
- Nah, actually these all have to be written as funcitons, so they can be reused..
- So extra file with functions now. Which will be reused in the templates
- OK, so extracting PCN numbers. The reusable funciton will only pick out the last year. But I also need to pick out the previous years. So I need a function that picks one or more columns based on the year passed as an argument?
- And of course the numbers have commas in them, so they need to be regexed to become numeric. Actually remove all non-numeric characters.
- OK, PCN tables done, phew.
- Next, PCN income from pdf
- OK, now all the tables from the pdfs are done
- Merge them by year, for 14 and 15 there is only one, and there are three for the years 16 and 17
- Then merge all the pdf tables together
- The add the excel data.
- Ah, there are some spaces in the authority names.. in the DPE tables. Fixed
- Additionally five LAs have non-standard name spellings. Need to figure out which ones are standard..
- Additionally there is an empty row in the ie table. Damn, this is getting kludgy..
- OK, no3 there is a funciton for changing spellings of scottish LAs, I'll add it to all the funcitons. Fixed
- Now go back to scotland excel import and rewrite as funcitons. I mean move them to the funcitons.R file
- OK, all of scotland works now, there are only some column type issues . fixed
- Oh, and welsh income is negative .

** ENGLAND IMPORT ORIGINAL DATA**

- First add actual file names to the metadata table.
- Also, start with 2009, not 2008, because there were 388 LAs then
- So the trick now is to write code to import the original data that will also be useful for updating new data for RAC..
- Turns out the columns with the la names and classes also change from one year to another, have to add them to the metadata
- WHY doesn't this work? Because one cell has "..." in it instead of numbers..

- FFS, ok, this now works for expend.on and expend. off. Now extend to incomes as well. Done.

Satyrday 23.3.2019

- prealocating a data.frame does not need to list columns, this are small applications anyway, so it's not necessary to worry about efficiency.
- Now add the totals for ENgland, transport and PCN totals, so only one row for the whole country, for each year. `auth.type` classed as "X"
- Now again, because these are single cells, they get exported as the column headers..Same as the scotland data.
- Totals import OK.
- Forgot to add country var to England. OK.
- Add the penalty ranges to the I.E. outturn code as well. This possibly requires checking if the data is there..because it's only there for the last three years. OK. **this turns out was not true!? not sure how i missed this!!**
- Remove authorities we don't need. That's all the ones with an O type, except ones with Naitonal park in the name?OK
- Congestion charges.
- OK, added congestion charges from the main otturn import, so the whole column is imported for everyone, it doesn't matter, since we only use Greater London.
- But now i see that some other authorities are kept, ones with WM and YH, and one that's even n/a, this is all 2014 data... Ah, I had the auth.name and auth.type wrong in the metadata, fixed.
- Double check no of authorities: 9 * 353 = 3177, that's perfect! In addition also 9* Greater London and The national parks, 9 of them each year, except in 2008/09, so that's another 80 records, total of 3266 records.
- Budget data next
- actually the auth.name columns change as well, so need to add them to the metadata. * Also have to add file.names to the metadata. OK
- Budget data all OK.
- Manual input of WPL data OK

** Merge England **

- Full_join not working - too many rows result.. is it a name thing?
- Yeah, in 2014 the outturn data e.g. says "Worcestershire CC" instead of "Worcestershire" etc.. SO probably safest thing to do is to remove all capitalised two letter words at the end of `auth.name` with some regex magic. or three letter words.
- Still 30 doubles... variant spellings.
- SO new FunEnglandFixNames for manually changing names...
- so 3266 authorities in the outturn data, and another 366 from the budget data, since it has an extra year, and finally 4 wpl years plus 9 england totals = 3642 england rows.
- plus scotland and wales is 4046.
- DONE

Sunday 24.3.2019

- Change Scotland lookup tables to the same style as the England one. OK
- move both to separete rds files. OK
- rename data sources file to 00-manual-data-input.R. OIK
- move the .rds files for metadata into /data/01-raw. OK
- also rename all of these to start with orig. OK.
- **ToDo : Wales doens't have a metadata file, should update that**
- **ToDo: for consistency, all the original data files names should end with the single year, not the double ones as they do now.. that means also changing all the metadata...**
- Clean up code to read all meta .rds files at once? OK
- Looks like this mass import using assign and lapply is changing the character columns into factors. fixed (stringsasfactors option was off in 00-manual-data-input.R)
- move WPL manual entry to 00-manual-data-input.R . OK
- Also add year to the manual WPL entry instead of wpl.year, because it gets added using bind.row anyway. Oh, actually it should get added using full_join, so they are not separate rows. OK
- And now full join the wpl data by auth.name and year. but wpl logical actually doesn't need to be there anymore, you can tell who has wpl simply if they have a wpl.income value. OK
- So now there are 4 fewer rows than before.
- so 3266 authorities in the outturn data, and another 366 from the budget data, since it has an extra year, plus 9 england totals = 3638 england rows.
- Scotland has 6 years of i.e. data * 32 = 192, pdf data is for 96 of the same rows, so 192 in total.
- Wales has 308 rows, that's 14 years by 22 LAs.
- Now that should be 3638 + 192 + 308 = 4138.
- OK!!

** start wales report template ** * how do i get the year into the title? * get colours to work (include in header) * now, i import the whole table at the start, from rds? * where are the surpluses calculated? maybe i should do all the calculations in the reports.Let's try that. * creating tables.. + empty cells? + column headers that are numbers (2018-19) + hlines + alignment + multirows + alingment if you have mutlirows...

Monday 25.3.2019

- Write up technical background - metadata for original data file.
- Continue the Wales report
- ** Add % symbols to table**
- metadata for master - turns out it would be great to have greater london authority separated out from the other london ones.. So i need to fix that in the import. OK, it's now GLA. OK
- Ah, collapsing rows in kable extra is not working as expected.. like buggy. OK
- But updating the package worked! OK
- OK, the row table took 1:26 h haha. But it's done now. And I learned a tonne..
- Now column names. OK
- Oh, dpe.status is empty, and the relevant info is in key instead..fix it!OK
- Oh, and also wpl.expend should be expend.wpl, fix it. OK
- So make sure all the income data is in thousands.. i.e. the Scotland PCN data isn't, so fix it. OK
- AND the England total PCN incomes don't need to be in their own column, htey should be in the income.pcn one. OK

- Finished rows and columns in tehcnical documentation for master data. Commit.
- **The data for 2013-16 is in a scanned pdf..**
- **List different definitions?**
- **TODO maybe figure out thos NA warnings in the import..**
- cross-referencing tables.. bookdown seems to be the easiest way.OK
- **ToDO: sources in tecnical documentation**
- OK, back to Wales, add % symbols
- Conditional text, seems to work fine. A bit wooden though..
- Now have to pull out england's surplus as proportion of transport costs.
- OK, I get England surplus/transport to be 21% not Leibling's 12 %. Additionally if I go directly to the 17/18 file I can confirm that my totals are correct, or rather that the Excel files have an error: the off-street column does not add up to the total they have down (cell DF455 is wrong), probably something to do with Copeland? But the difference is 250 only (thoushands), so it doens't affect the 21 % I get.
- Next the plot.
- Ahh, the ugly fonts.. was it install.packages("extrafont") that did the trick before?
- I think if you use extrafonts you have to upon installation also run `import_fonts()`, but only once, this needs to be in the instructions!
- This here says you need to specify dev = cairo_pdf to make sure the fonts are embedded.
- Hmm, can't figure out how to use extrafont in knitr, only directly to pdf.
- OK, shotext seems to be working. But you have to add family to each text bit..OK
- the labels are still too large. OK
- legned. OK
- OK, also fixed axis ticks so they are all there. OK
- Now remove ToC from bookdown OK
- gridlines?OK
- fontsize is too tiny. OK
- But now table font is too large. OK

map wales

- OK, another detour. First I'll have to manually add the info.
- Let's do some more text first.
- Now table 2, OK
- Thousands separator. OK
- Actually, might as well do all three tables while i'm at it :)
- There are these odd gaps in the table. turns out there is an `\addlinespace` added randomly because of booktabs beign true. but i have striped anyway, so it looks ugly. using `linesep = ""` fixed that, thanks stackoverflow! OK.
- Forgot the total line. OK
- Make sure the percent change for the total is also calculated. OK
- add hline on last row. OK
- arraystretch to loosen it up a bit. OK
- Expenditure tables is OK, although last column needs to be changed

- Also getting exactly two decimal points in R is tricky! You need to use `format(round(x, 2), nsmall = 2)` apparenlty !
- have to find a better solution than the font_add_google function, since this downloads it every time..OK, font now in data/01-raw and automatically loaded, all good.
- Expenditure table done. Added change on previous year.
- Now the surplus table. OK
- Conditionally format cells where the change is in deficit, not surplus.OK
- Add footnote to make that clear. OK
- Now some text..
- Oh, just in case anyone wants to redo old reports I have to limit full data to stop at current.year. OK
- Now I need to convert numbers to words. `xfun::numbers_to_words()` does the trick.
- Now I need to capitalise these words as well
- Just did a summersault to exclude Torfaen or any other LA with an income under 30,000 from a condiotnal text tree.
- More conditional trees...
- Damn, the auth.names have spaces at the end.. I need to strip them. Update function and rerun data import. OK
- maybe a page break would be nice for each new section? OK
- had to move the surplus calculations to the top because they are used in the intro. OK

Tuesday 26.3.2019

- Finish the surplus text in the Wales report. OK
- Because Leibling splits up the surpluses into surpluses and deficits, I can add two extra rows at the bottom. Hm, but that's not that simple since it is different ones every year... OK.
- OMG, that was a heroic effortl, and it worked! To get the sub group sums conditional on each year's surpluses or deficits, and additionally keep the transport data only for the last year in there.. crazy. OK.
- Now the text requires I extract the proportion that the top three councils contribute to the positive surplus.. OK
- Foudn an error in Leibling, the 56% is from the total surlpus, not the 15m positive surplus.
- Wales report complete. OK
- Committ
- check the 2015/16 report/
- weird error in the expend table 15-16: Wrexham down as 583 instead of 304 in 2013/14
- oh, this format round stuff is so ugly, need to make a fucntion.OK
- surplus as proportion of total transport expenditure is different in 15/16 compared to Leibling. But that's because he's using gross expenditure before parking surplus, although he doesn't say that in the text. error?
- OK, go through queries and see what is missing.
- Saving the reports, can i pick the name and folder?
- OK, folder is fixed in yaml.
- What about name? Looks like you can't change that dynamically if you want to use rstudio - knit.
- I'm getting an ! LaTeX Error: Unknown float option 'H'. errror.. didn't have it before. looks like this answe is the solution, i need to load the latex packages in the yaml

- Send off Wales reports.
- parametrize the report. but doesn't work - Error in eval(expr, envir, enclos) : object 'params' not found . Sorted
- Hmm, let's try some updates and read up on param to make sure i;ve got this right, otherwise an repex will be in order.
- Aah, RTFM worked. The params can be passed from the rmarkwdown::render function, but they also have to first be specified in the YAML. OK
- OK, what next?
- Process, figure out how the whole thing will work!
- The trickiest bit is getting to add rows to master, but if you fuck up being able to overwrite what you've just added.
- If what you're adding always has the same year-auth.name combinations? And then just issue a warning when you are overwriting them? repex.

1. If the anti_join with the master has the same number of rows as master, then you are adding X new rows, and not overwriting anything.
2. If the anti_join is smaller, you are overwritting. 2.1. If the difference is smaller than the update, then you are overwritting X with X+Y rows. 2.2. If the difference is larger, than the update, then I can't know. Instead I can just make sure that the update is always the correct nrows. e.g. 22 for Wales, 32 for Scotland? What about england? if there are different number of national parks, that will mess things up?

- clean up original wales file names, and replace in 01-orig-data-import.R OK.
- Wales update template ready and setup!!! Works great.
- Commit
- Oh, and delete the annoying empty folder that keeps cropping up. OK
- Todo Now clean up the Wales references as well! OK
- OK, now Scotland
- TODO: shoudl I move all the code out of the Rmds? It would look nicer for sure. Done
- OK, DPE table done! And that was not easy!
- let's try a map. The scotland map has 41 polygons, I need to colapse them into 32..
- More lookup table updating..
- map works, but for some reason the font won't work - if i leave the par command there it stops working. *todo*?
- Ugh, East dunbartonshire has lost it's year.. that's the import of the DPE table..Fixed
- UGH, turns out i also dont have Scotland total transport data. OK
- i had originally entered both income and expenditure cells for the total transport into the metadata, but that seems unnecesary, so i'll change that now to just the net revenue expenditure one. fixed in 00-orig-data-entry. OK
- OK, added funciton for importing scotland transport totals, added the code to 01-orig-data-import, and now reran the whole thing. OK
- Now need to update the technical doc, to take itno acount new lines, six of them? Aah, nice, code just works! OK

- OK< back to Scotland summary.
- After 2013 Scotland started reporting their incomes as negative. LOL. OK, where do I fix this? In the import function I think. best. Yeah, added an abs there. OK
- Summary table looks great - but (error) it massively disagrees with the Leibling report. No idea how that's possible, although the fact that there is no data for Aberdeen city in the excel files could be the cause!?
- Still haven't added the percentages to the bottom row. need to fix that in the Wales one as well. OK
- Wales as well. OK
- Instructions: ammending tables with a footnote!
- scotland summary table and text are now OK
- plot. OK
- issues: floats of tables and figures, little control..use "H". OK
- text and code seem to work fine from Wales, a bit of adjusting.. Like the inf and Nans, I still need to deal with them a bit more. OK
- longtable is maybe better here, trying it out now. OK.
- todo: there are still a lot of old format/round combos need replacing, OK
- OK, scotland Expenditures done.
- todo: deal with the nans and infs.OK
- just not sure why clangmanshire is in the text, check that!. OK
- commit

Wednesday 27.3.2019

- OK, the top/bottom code needs cleaning up.
- Should i consider moving all the code out of the .Rmd files? OK
- top/bottom. first use deframe() instead of pull in order to get a named vector. nice tip!
- top/bottom: exclude from list if they have under 30,000 income/expenditure.
- *clean up code: stop using column indices!*
- in dplyr nrow(.) can be replaced wiht n() .tip
- excluding rows for under 30.000 is now only done for largest decreases, but should also be for largest increases... That needs to then be fixed in the text.
- change that in wales as well..
- OK, done for sco.expend. OK
- OK, repeated for sco.income. OK
- Hm, sidenote: chunk names where there are tables (used for referencing) cannot have dots or underscores in them? tip!
- Next: copy to Wales. OK
- Next: fix that one sentence top three incomes. OK
- group a df by rownumber: group_by(row_number() <= 3). tip!
- OK < a bit of wrangling, but i got the top 3 incomes out. OK
- add to Wales. OK.
- need to create function for 2017-18 etc..OK
- Note: i'm getting latex warnign; label multiply defined, but it seems that's because of longtable meaning the table caption is on two pages. So you can ignore it. tip!
- OK, income and expenditure codes and texts OK in both Scotland and Wales.

- ToDo funfisc, replace all. OK
- fundec, replace all. OK
- !!as.name.. automated using code snippets. tip! OK
- Next: surpluses?
- I've moved the sco surplus code to the top, because in wales it is used in the text there, don't forget.
- surplus totals for deficits and sufficits are different from Wales, since in Scotland we don't have individual level transport totals. So that needs to be rewritten
- but that isn't actually true, the information is in the IE excel files... Need to rewrite the import.. OK
- fixed in funcitons.R and 01-import. create new original/master data. OK
- Now first rerun the technical.rmd. OK
- OK, back to surplus. table OK
- First fix the binning function to let you have 0 as well as poz and neg/.OK
- And do that in all 6 cases. OK
- Now get table to work. OK
- width of first column.. OK.
- now surplus text.
- top three surplus councils need to be named..OK
- do that in Walse as well. OK
- Now add the top 3 bottom 2 thing that we did everywhere else, but seems to be missing here. OK
- HMM, when i exclude the extremes because they have too little income/expenditure/surplus, then I am only looking at the top/bottom 2/3. But what if after excluding them there are still very low values left? OK
- OK, turn it around. first find the top three and bottom two. Then anti_join, but just below/above, to get the excluded ones out... let's try this. OK
- OK, this now works. But how do i make the conditional tree simple if there is potentially X number of excluded councils? Where do I do `, and`. OK wrote a function. OK
- This all works now, need to replicate in all 6..
- Replicated for Wales surpluses. OK
- still need to replicate for incomes and expenditures. in scotland OK.
- still need to replicate for incomes and expenditures. in wales.OK
- SHit, if surplus goes from poz to neg or vv, there should be no change. OK.
- Also, the transport numbers for scotland are definitely wrong. I am using net expenditure, but I should presumably be using gross expenditure? OK, I'll fix that now... BUT still no correspondence with the Leibling report..
- TODO: table of expenditure as proportion of income over the years? That removes the need for the final column in the expend.table
- SO if a change or proportion is nan or inf it should become NA, right? Yeah, but now the problem seem to be the italics... If it's NA and italic, it won't make it disappear. Because the whole column is now character.. So I need to fix that as well. This is done in Scotland. OK
- OK, surplus tables with change in sign. Scotland and Wales. OK
- OK, let[s do a expenditure as proportion of income table for Scotland first. OK
- OMG, this conditional colouring is super hard!!! Kinda works now, but still have NAs, NANs and Infs....They need different colour codes? OK
  - NA is empty and whiteOK
  - Inf means no income. So also white? OK
  - then gray out the text for the ones with too small numbers? Nah, a footnote will do.
- Now repeat for Wales. OK
- Now the PCNs. first need to extract the old pdf data from the Leibling report.NO, they're not there. The data would have to be manually input... Maybe later.OK
- LOL, i had a [1-9] gsub line, which stripped out all the zeros from the Scotland PCN data, eeek! OK
- OK, just the last two years for PCN incomes then. OK
- Manually correct the 2016/17 Argyll and Bute no on PCNs from 3018 to 13018? Make sure it's in the tech doc.
- OK, need to fix the top/bottom code for income and expenditure in Scotland. OK for income.OK for

expenditure.
- todo: write a function for numbers to words. OK
- change in all scotland and wales reports. ok.
- Now DPE east dumbartonshire's lost year. OK, fixed.
- also cleaned up some of the warnings. all of them actually. OK
- OK, table 3 - comparing the 3 nations summaries. Can it be generalised, so it can be in every report. Most recent year available.
- oops, there might be an error in the england calc in wales forgot to group? NO, it;s OK.
- error: the england and london figures in table 3 of scotland report 16/17 don't quite add up, specifically the non London income is about 2mill smaller and the expenditure about 7 million smaller.
- OK, table 3 is now done, for both Scotland and copied to Wales as well. OK
- Next: table 4 missing in Scotland and Wales. Then send it off again?

Thursday 28.3. 2019

- OK, now for the last of the tables in Scotland.
- Summaries for all three tables make the fact that on and off are not added up already seem a bit odd. And requires more ugly code. Maybe I should do that in the import function. Then I should also do the surpluses? Probably. It would be nicer anyway for giving out the open data.
- OK, import adds up england totals and calculates surplus now.OK
- OK, import calculates surplus for Scotland now. OK
- could remove the superflous claculations of it later - although it makes no difference. OK
- Now wales: OK, added that, and also changed that in the wales update.R file.OK
- Reran wales update, make sure the copy.file has overwrite swithced on!
- technical: add this: only on+off and surplus have been calculated. Despite it being bad practice probably.
- Now clean up all the unnecessary calculatoins of surpluses.. ok.
- OK, now fix that in Wales as well.OK.
- Commit.
- OK, the summary changes over past 4 years work now, but odn't quite match up with Leibling's ones...
- Also, he has a GB total as well inthere... Add that. OK
- actually, add GB total to table 3 as well.., OK
- This is such a stupid table. It should be annual change, not over 4 years, who can interpret that!? I'll do both. OK, done.
- Now need to add formatting ..OK
- All tables for scotland done. Commit
- write up queries.
- Add this annual table to Wales as well.
- Now the table is printed off the page!?
- also r chunks can't have overlapping names, haha. tip!
- OK, clean up a bit, Wales and Scotland ready to go. Kinda.

- send off queries 5, commit.
- OK, next job: penalty as proportion of DPE income. That's the table I'm missing, because the data needs to be manually input..
- Income PCN and income TFS for 2013-2015... OK,
- and double check? OK.
- Also PCn number for 2013.OK
- Now all these 3 sets of data are input manually in 00-orig-data-entry. OK
- They get automatically picked up by 01-orig-data-import.
- Now they have to be merged with the existing data in 01-orig-data-import.
- double check everything. OK
- OK, scotland report: PCN numbers table now goes back to 2013.
- Next: the PCN income as prop of DPE income table (table 7 in the original report).
- ffs, turns out there is LA level PCN data for england for before 2015 afterall, somehow. i just missed it... better sort that out now... OK
- *todo*: in the end re-download all the original data files!!
- OK, the PCN columns for England are now input, although I'm a bit worried about the 08/09 one, not that that matters much tbh.. But also it isn't added up in the total income column there? weird. not important. OK
- so, add sheet and column for PCN data for england 09/10 to 14/15, which i had somehow missed before. OK
- rerun 00 and 01. OK. PCN data for England is all there. OK
- check if the sums are OK. perfect match!
- finally got rid of the last of the warnings in the import - it was because of a separate call that was dropping empty. OK
- OK, back to the PCN table, it's similar to the surplus/income table, so let's reuse that..
- Not sure if i'm using TFS income or regular? try both i guess. OK, it's the regular.OK
- Oh, OK, income.pch is not in 1000s. OK, fixed in 00 data entry. OK
- Now table is ok, but have to add england and london.. and wales?
- *check that using data.full in the annual, not data*
- just a reminder, "X" in auth. type means England totals for transport and PCN. OK
- OK, the sorting of the table should be alphabetical. OK
- and instead of scotland it should say dpe. OK
- table done. OK

Monday 1.4.2019

- respond to Ivo and Phill's answers. OK
- Incorportate the following:

- change visible in text is described as such in text. (difference smaller than 0.05 is under FunDec threshold). OK.
- add 30.000 explanation sentence> OK
- Table 2 units.OK
- There are still some format's that haven't been FunDec-ed.OK
- changed all the 0.02 to 0.05 when saying two things are the same, since 0.05 at FunDec disappeares. OK.
- delete for in tfs.ok
- unit in table 2. OK
- "Income has .. sentence". OK, so there are 12 possible income/expenditure/surplus combinations... I should write a function I guess. Later. Nah, simplified it.
- million/m. OK
- all other typos. OK
- The surplus/deficit top/bottom 3 are wrong because some are deficits.. query.OK
- also in the surlus table Clackmanshire has no change, but it should, -100%, consistent with the others. AH, comparing signs. but zero has a zero sign. so instead of signs being different, i have to have signs being 2 apart. Fixed.
- oops, looks like the FunMulti text was not done in Wales? OK
- add tfs table.OK
- *add total to tfs table.OK
- Error: Leibling's comparison table for Scotland: last row sums up the totals for all the LAs, including the ones not reported on by TS. So in the last three columns the total of the differences is inflated by including as differences also ones where there was no DPE reporting. Instead I've calculated the totals as just that: the totals of the columns as they are in the table. OK
- clean up millions. in Wales as well.
- table 7 13-14 data is in thousands..? FixedOK
- add references.
- Mapping reference and copyright notice. OK
- OK, references:
  - some are already input into the metadata tables. What will happen with new ones?
  - Additionally there is cross-country comparisons that use a bunch of other references as well. OK
  - Should I list them all? Yes
  - OK, so I need to create a master bib file/table. Of all england, wales and scotland sources. OK, manual input in 00-orig-data-entry. OK. Save as master.bib.rds. OK
- OK, so now when you compile a report, you must first create the new updated bib.master.rds, and from there create a bespoke bib.file.
- Additionally you will need to update the bib.master file when you update all the rest of the data - todo. OK
- Now the update creates a new bib. file for this report only: the relevant country references are saved into a .bib file in the same folder, with the same name as the report.
- OK, the report.Rmd has to have the correct .bib file in the Yaml. That actually means the name cannot be dynamic. so instead it is just the country name, and it will be overwritten every time, that's OK anyway, because when the .Rmd is created, the correct .bib file will also be created.
- but the no-cite references also have to be created in the .Rmd file, and this luckily can happen in the second part of the YAML, meaning that the names can already be picked up from another object based on the current.date.
- Man, took a while, but finally got it to work! OK
- OK, Wales references are now all cool, only problem is this (!?) has somehow broken the table cross-referencing!?
- Let's try Scotland and see how that fares. Yeah, as soon as bibliography is added to the yaml, the other references seem to stop working...
- No idea how this is happening!? ALso no idea how it suddenly started working again?! Anyway, refs are in Scotland, needs cleaning up
- rename rds bibs, to have bib in the name. OK

- ONS for maps name needs {}.OK

Tuesday 2.4.2019

- make sure the references work for Wales as well
- cleaning up: turns out i need to be careful distinguishing between the fiscal year and year published.. in the bibs it's not the publihsed year that's called year, because that's what the ref manager picks up it seems. And so the key actually has the fiscal year in the name because that's the one I can pass to it from current.year. OK
- referencing style
  - date accessed. update APA, OK.using this style editor
  - But change text of date accessed in refmanager. meh, this seems a bit involved.. i think i would need to switch from bibtex to biblatex, and then update the style again... Not important anyway. it happens in the WriteBib function.
  - capitalisation. solve by using {}. OK.
  - italics instead of quotations, OK APA.
  - tip: when using WriteBib form a dataframe, column names shouldn't have "." in them.
- add conditional formatting to change columns. OK
- crossing zero requires a bit more work... But done for income, now expenditure. This one is tricky. Actually it's the bloody South Ayershire that's wrong that's fucking everything up. I should manually fix it to what Leibling has. OK
- font size of table captions changes. why? oh, because they are longtable... hmm, that means i best make them all longtables. OK
- ALso striped rows don't work if you have colapsed. There's an unspported hack, but it doesn't work with Table 2 in Scotland I'm afraid. OK
- add note on comparison table.OK
- maps? OK, done
- OK, mapping means figuring out palattes, make them nice and make them match the ones in the tables. It also means centering on zero. So probably two palattes diverging. OK this link was useful.
- but the example there is silly, if the data is skewed the zero will be way off. Maybe it isn't silly, just need to fix the scale/legend?
- kinda works, but i need to figure out the scale legend as well. OK
- and then split it into two palettes.
- aah, viridis is great, does the interpolation for you, colorbrewer has only 11 max and then you have to use colorramp to interpolate.
- yeah, palettes work!!OK
- Now need to be able to swap D and C if i want.
- phew, some messines, but it seems to work.
- Not as pretty as before, but at least consistent..
- font for scale legend can't be fixed?
- adding % to scale.. means need to get nice axis `at` . OK
- narrower scale, means the

- Now add same palettes to the conditional cell colouring.
- *maybe inset shetland isles at some point..* Probably need ggplot for that though..

Wednesday 3.4.2019

- change plot line colours and map. OK
- add aplha to the conditional tables at least? Hmm, so the findColours function drops the alpha values. This sucks.
- OK, but i can use findCols instead. I get that to work, the palettes now have 8-character colours, the last two being the alpha. And it works in the maps. But in the conditional formatting it doesn't come through. No, it won't, seems LaTeX doesn't handle it. Could try mixing with white? But too much effort. Nope, sorted it out. `FunLigthen()` But i don't like it, so revert to normal..
- instructions: install Latex, tinytex maybe?
- fix palette function for non-negative scales
- change plot line colours.OK
- maybe maps look nicer without black lines? OK.
- references to maps in text. OK
- Do wales maps? OK
- add text reference to maps. OK
- change line pplot colours.OK
- add boundary data licence.OK
- add income sentence to wales.ok
- efficiency error somewhere.OK
- instructions: you do have to read the text through. It will say things like 14% and 14% respectively, which you might want to smooth over..
- todo: millions wales.ok
- instructions: sometimes the float cause text to ovwerflow... not sure why, but the alternative you don't want: where floats are actually floating. This won't work well in a document like this with so many tables and figures, you want them to stay in the correct order with the text. So if this happens manually insert `\pagebreak` at any location that is overflowing to force a new page.
- fix words to numbers function so that it only works if number is > 10. killin it.OK
- add conditional formatting of tables.OK
- move code out of rmd. OK for wales
- OK, testing external code:
- knitr options moved out work
- packages installed work
- objects are in .rmd environment. so all good.
- todo: when compiling remove log file. OK
- todo save csv files of individual tables! *Wales OK
- these csv's mean splitting tables into csv-able and formatted for kable.. but all good. OK
- *prepare scotland update file*
- *move code out of rmd for scotland* and add csv writing

Wales (based on 17/18)

1. Introduction

- Text, only fiscal year
- Text with named councils that have free parking, manual input.

2. Summary

- Table 1 based on Wales statistics data, all checks out
- Text based on table 1
- Crossover value from England report.
- Figure 1 from table 1
- Map, is a map, but potentially unnecessary.

3. Income (there is no separate section?)

- text based on table 2,
- Some conditionals.
- Sentence on proportion of the population.. Do I have to get the data for that?
- Paragraph on income per population removed, since data also removed.
- Table 2, data from Wales statistics data, ignore the population measures, cf Ivo: it is a meaningless measure

3.1 PNCs

This data is massively time lagged, Ivo suggested removing it?

4. Expenditures

- Text based on table 4, some conditionals too much for automation
- Table 4 based on Wales stats data.

5. Surpluses

- Table based on Tables 3 and4
- Text based on table
- Some text has explanations for the reducitons/increases that are too tricky to write conditionally. But could potentially be done.

6. reporting

- Non automatable, can leave skeleton text/table if required.

Scotland (based on 16/17)

- Mainly pure text chunk.
- Available data: fiscal year, link to income and expenditure data, link to PNC data, link to SPRECC report with previous data.
- Paragraph on data availability on Penalty Charne Notices, incl. source, which is a scanned pdf...

1. Introduction

Text

- Text based on table 1. Conditional tree could get complicated as the table becomes simpler, but can make it simple I think.

Table 1

- exact same table as in pcn.pdf, can be extracted directly from pdf and cleaned up.

Text

- Pure text

2. Summary

Text

- based on table 2. Conditoinal tree must be able to describe temporal trend..

Table 2

- Based on first summary sheet in Scotland income and expenditure data files. But it is inconsistent, partially because of Aberdeen, but something else is also going on..
- Based also on total transport expenditure from same sheet, but numbers don't match. ##### Figure 1
- Based on Table 2, OK

Table 3

- Based on summary data in table 2, BUT also data for England and Wales, which historically is already available by the time Scotland's report is due.

Table 4

- Again, crossover with other two countries, additionally this one is 12/13 to 16/17, not sure what to do next window wise.

Text

- Based on calculation based on data used in table 4.

3. Income

Text

- Based on data in table 5

Table 5

- Based on data in income and expendidure files, except for Aberdeen City.. And possibly some other local authorities?
- Also includes data from Table 1, i.e. the pdf scraped table.

Text

- Text based on table 6.

Table 6

- NUmbers of PNCs are from the table scraped in the pdf.
- Additionally the calculated average cost of a PNC is from the income table also scraped from the PNC.
- 3 year window in 16/17

Text

- Based on numbers in table 7

Table 7

- Proportions calculated from PCN data in pdf, and income data from xls.
- Except for Aberdeen..
- Including crossover with England data

4. Expenditure

Text

- Text based on table 8, potentially tricky conditional tree...

Table 8

- Expenditure data from excel files only.
- Proportion calculated from income also from xls files.
- Additionally DPE in operaction data from first scraped table from pdf.

5. Surpluses

Text

- Text based on nice conditional tree from table 9
- Also national ranking!!

Table 9

- Surplus data calculated from incomes and expenditures in xls
- changes also.
- again DPE data drom table 1 in the pdf

Table 10

- Contribution of surplus towards total transport expenditure.
- OK, but which one is the total transport expenditure!?

6 Comparison between Local Government Finance figures and Transport Scotland decriminalised parking enforcement figures

- OOh, interesting! But should it really be in this report, haha
- So combine data from xls and pdf and look at discrepancies, which should be mainly off street parking, but maybe are not just that.

Appendix

- Map based on Table 1

England

1. Introduction

- Mainly pure text chunk.
- Available data: fisc.year
- Missing data: national parks, Notthingham workplace parking levy? (17/18)

2. Summary

Table 1 - Parking income and expenditure England

- Unit: whole country
- Available data:
  - Off-street: expenditures (sum from LA tables)
  - Off-street: incomes (sum from LA tables)
  - On-street: expenditures (sum from LA tables)
  - On-street: incomes (sum from LA tables)
  - On-street: penalties - single cell on summary sheet
  - Net expenditure on transport - single cell on summary sheet
  - Budget estimates for next year
- That's all, everything else is calculated from this

Text

- Conditional tree based on data from Table 1.

Figure 1

- Simple line chart based on data from Table 1.

Table 2

- Data from Table 1 for current year, split by London and non-London.

Text

- Conditional tree based on data from Table 2.

3. Income

Text

- Conditional tree based on data from Tables 1 and 2
- Includes footnote on penalty tarrifs.

4. Expenditure

Text

- Conditional tree based on data from Tables 1 and 2

5. Surpluses

Text

- Conditional tree based on data from Tables 1 and 2
- Additionally there is the total for on national parks, that data is also available
- Missing data: the total for Nottingham workplace levy.

Table 3a surpluses for London councils

- Individual LA data for 33 London councils

Table 3b surpluses for top 20 non London councils

- Individual LA data for 33 London councils

Text

- Conditional tree based on data from Tables 3a and 3b

6. Comparison with budgets for2017-18and 2018-19

- Conditional text based on comparison of indovidual LA surpluses with individual LA budgets

7. Congestion charge

Text

- Conditional text based on table below

Table

- Simple table based on data from outturn table.

Appendix 1

- Table of all LA surpluses for 5 years, alphabetical

Appendix 2

- Table of all LA surpluses for 5 years, by surplus size.

Appendix 1

links to all reports online

country	year	htmls	pdfs	tab	window
England	2010/11	link		link	
	2011/12	link	link	link	2(4)
	2012/13	link	link	link	3(5)
	2013/14	link	link	<-	4
	2014/15	link*	link	link	4
	2015/16	link	link	link	4
	2016/17	link	link	link	4
	2017/18	link	link	link	4
Scotland	2010/11				
	2011/12	link	link		1
	2012/13				
	2013/14	link	link		3
	2014/15	link	link		4
	2015/16	link	link		4
	2016/17	link	link		
	2017/18				
Wales	2010/11				
	2011/12				
	2012/13	link	link		5
	2013/14				
	2014/15				
	2015/16	link	link		4
	2016/17	link	link		4
	2017/18	link	link		4

∗ link missing from RAC press releases page

Columns headings:

- *country*
- *year*–period covered in report
- *html*–link to press release on RAC website
- *pdf*–link to pdf report if exists
- *tab*–link to additional tables if exist (only England)
- *window*–number of previous years' data included in report, number in brackets is for aggregates, not at LA level

Appendix 2

data overview - England

England–Outturn

fisc.year	year	link	rows	las	first	e.sh	e.on	e.off	la.name	la.type	i.sh	i.on	i.off	e.cc	i.cc	pen.sh
2008/09	2008	xls	405	388	0	3	CU	DB	B	D	3	CX	DE	BZ	CC	3
2009/10	2009	xls	370	353	12	3	BH	BL	B	D	4	BH	BL	AV	AV	4
2010/11	2010	xls	371	353	12	3	BH	BL	B	D	4	BH	BL	AV	AV	4
2011/12	2011	xls	371	353	12	3	BH	BL	B	D	4	BH	BL	AV	AV	4
2012/13	2012	xls	371	353	12	3	BH	BL	B	D	4	BH	BL	AV	AV	4
2013/14	2013	xls	371	353	12	3	BH	BL	B	D	4	BH	BL	AV	AV	4
2014/15	2014	xls	371	353	12	3	BE	BI	C	E	4	BE	BI	AS	AS	4
2015/16	2015	xlsx	444	353	8	3	CU	DB	C	E	3	CX	DE	BS	BV	3
2016/17	2016	xlsx	446	353	8	3	CU	DB	C	E	3	CX	DE	BS	BV	3
2017/18	2017	xlsx	445	353	8	3	CU	DB	C	E	3	CX	DE	BS	BV	3

- *year*–period covered in report
- *link*–link to data file
- *rows*–number of rows (not all are local authorities though!)
- *las*–number of local authorities (i.e. not ones with class "O"
- *first*–first row of data on sheets 3,4 and 5
- *e.sh*–which sheet are expenditures on
- *e.on*–which column on street expenditures
- *e.off*–which column are off street expenditures
- *e.cc*–which column are congestion charge expenditures on (row = Greater London Authority)
- *i.sh*–which sheet are incomes on
- *i.on*–which column on street incomes
- *i.off*–which column are off street incomes
- *i.cc*–which column are congestion charge incomes on (row = Greater London Authority)
- *pen.sh*–which sheet are penalty charge incomes for LAs on
- *pen.on*–which column are penalty charge incomes for LAs
- *pen.1*–which cell is the totaL Penalty charge notice income for on street parking data on in sheet 2
- *tot.1*–which cell is the Net current expenditures for Total transport and Highways services on in sheet 2

England–Budget

fisc.year	year	file.name	link	rows	first	las	budg.tot	budg.la	la.name	la.ty
2008/09	2008			0	0	0	-	-	B	D
2009/10	2009			0	0	0	-	-	B	D
2010/11	2010	data/01-raw/orig.eng-10-11-budget.xls	xls	443	10	353	E27	U	B	D
2011/12	2011	data/01-raw/orig.eng-11-12-budget.xls	xls	444	10	353	E28	V	B	D
2012/13	2012	data/01-raw/orig.eng-12-13-budget.xls	xls	444	10	353	E28	V	B	D
2013/14	2013	data/01-raw/orig.eng-13-14-budget.xls	xls	444	10	353	E28	V	B	D
2014/15	2014	data/01-raw/orig.eng-14-15-budget.xls	xls	444	9	353	F34	U	B	D
2015/16	2015	data/01-raw/orig.eng-15-16-budget.xls	xls	444	9	353	F34	U	B	D
2016/17	2016	data/01-raw/orig.eng-16-17-budget.xlsx	xlsx	443	8	353	E35	U	B	D
2017/18	2017	data/01-raw/orig.eng-17-18-budget.xlsx	xlsx	446	8	353	E35	V	C	E
2018/19	2018	data/01-raw/orig.eng-18-19-budget.xlsx	xlsx	443	8	353	E35	V	C	E

- *year*–period covered in report
- *link*–link to data file
- *rows*–number of local authorities
- *budg.tot*–wich cell on sheet 2 is the total budgeted surplus in
- *budg.la*–which column on sheet 3 are the LA parking budgeted surpluses on

England–Nottingham WPL

year	link	page
2014/15	link	69
2015/16	link	79
2016/17	link	84
2017/18	link	86

- *year*–period covered in report
- *link*–link to data file
- *page*–which page is the WPL income table on

Appendix 2

data overview - Scotland

fisc.year	year	report	file.name	link	start.sh	end.sh	exp.cell	inc.cell	t.exp.cell	autl
2011/12	2011	yes		–	NA	NA	-	-	-	-
2012/13	2012	-	data/01-raw/orig.sco-12-13.xlsx	xlsx	2	33	F57	G57	F50	A2
2013/14	2013	yes	data/01-raw/orig.sco-13-14.xlsx	xlsx	2	33	F40	G40	F33	A2
2014/15	2014	yes	data/01-raw/orig.sco-14-15.xlsx	xlsx	2	33	E40	F40	E33	A2
2015/16	2015	yes	data/01-raw/orig.sco-15-16.xlsx	xlsx	2	33	B41	C41	B34	A2
2016/17	2016	-	data/01-raw/orig.sco-16-17.xlsx	xlsx	2	33	B41	C41	B34	A1
2017/18	2017	-	data/01-raw/orig.sco-17-18.xlsx	xlsx	3	34	L13	L23	U13	E2

- *year*–period covered in report
- *report*–does the report exist for this year?
- *file.name*–name on link to file on the website
- *link*–link to data file
- *start.sh*–number of first sheet with LA data
- *end.sh*–number of last sheet with LA data
- *exp.cell*–cell with expenditure data
- *inc.cell*–cell with income data
- *t.exp.cell*–total expenditure for all transport and roads cell
- *t.inc.cell*–total income for all transport and roads cell

DPE and PNC data

fisc.year	year	file.name	file.type	link	dpe.tab	pcn.tab	e.i.tab
2011/12	2011	–	–	–	NA	NA	NA
2012/13	2012	–	–	–	NA	NA	NA
2013/14/15/16	2013	data/01-raw/orig.sco-13-14-15-16-pcn.pdf	scan	pdf	NA	NA	NA
2016/17	2016	data/01-raw/orig.sco-16-17-pcn.pdf	pdf	pdf	4	5	6
2017/18	2017	data/01-raw/orig.sco-17-18-pcn.pdf	pdf	pdf	4	5	6

- *year*–period covered in report
- *report*–does the report exist for this year?
- *file.name*–name on link to file on the website
- *link*–link to data file
- *dpe.tab*–which table (including the table of contents) contains the table of LAs that have/haven't DPE
- *pcn.tab*–which table (including the table of contents) contains data on no of PCNs issued
- *i.e.tab*–which table (including the table of contents) contains data on incomes and expenditures of DPE LAs

stackoverflow question:

alpha channel in 8-digit HEX code stripped in kableExtra cell_spec()?

I'm (conditionally) formatting cells in a LaTeX `kable` table, and attempting to pass 8-digit HEX colours in order to get some transparency.

It seems however, that `cell_spec()` strips away the last two characters of the colour, leaving it opaque.

NB: In fact even if I use the built in `viridis` functionality via the helper function `spec_color()` that has an alpha argument, `cell_spec()` doesn't seem to pass it on.

See repex below: you can see the colours are 6-digit instead of 8.

```
library(kableExtra)
library(dplyr)
x <- data.frame(a = 1:4)

x %>%
  mutate(a = cell_spec(a, "latex",  background = c("#AADC3280"))) %>%
  kable()

x %>%
  mutate(a = cell_spec(a, "latex",  background = spec_color(1, alpha = 0.5))) %>%
  kable()
```