



HM Revenue  
& Customs

# Interactive data visualisation in HMRC

Presentation to the GSS Regional Roadshow  
11<sup>th</sup> July 2018 Defra York

Paul Pinnington and Andy Schofield,  
Centre for Data Exploitation (CoDE)

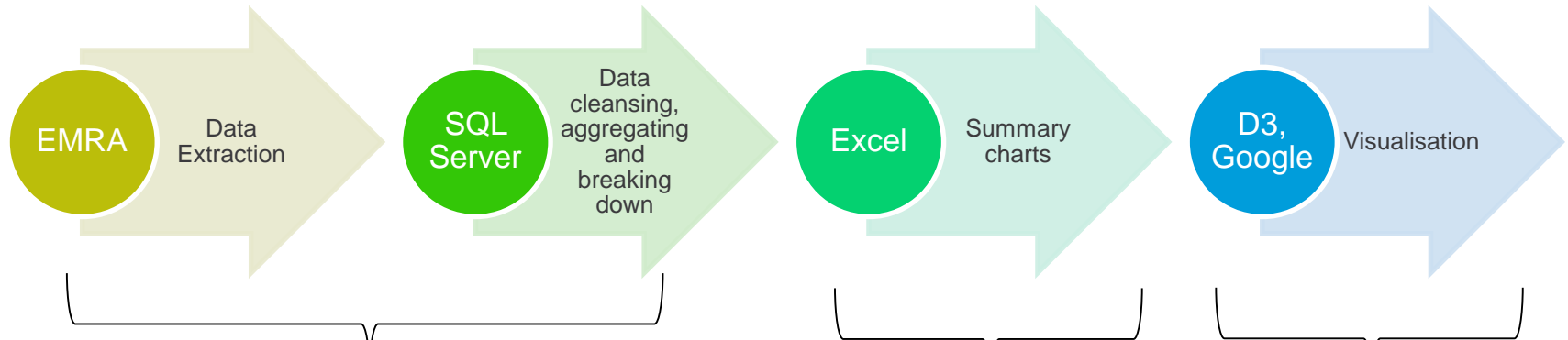
# Using visualisations

- Today's showcase session covers:
  - Some D3 visualisation tools used for EMRA analysis internally, running in JavaScript.
  - Weaving R with HTML using R Studio, R Markdown, KnitR, and Leaflet applications to produce reproducible interactive HTML for external statistical publications.
- To allow easier sharing with internal and external customers we have developed web services (Google tools / GitLab / GitHub) to host some of our applications.



# EMRA Visualisations (1)

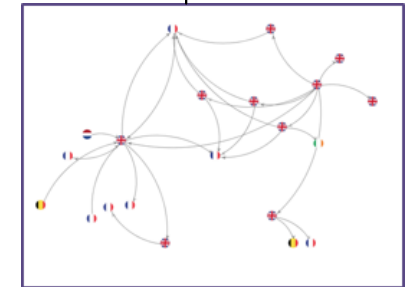
## Process:



## Products:



Documentation  
on best practice



HM Revenue  
& Customs

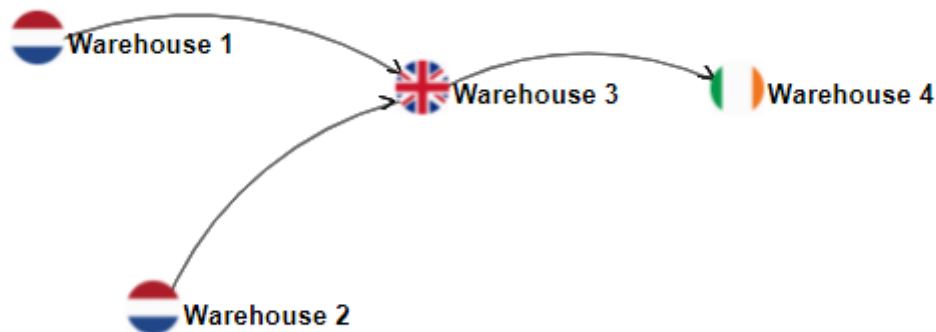
**EMRA** = “Excise Movement Risk Analysis”

– A tool developed by HMRC supporting bulk extracts of data from EU-wide **EMCS** system which tracks the movement of Excise goods (e.g. alcohol, tobacco, oils).

# EMRA Visualisations (2)

- Cleaned data and JavaScript/D3 code hosted in Google Drive.
  - Allows easy sharing of visualisations and for us to restrict permissions to named individuals.

Source	sourceCountry	Target	targetCountry	Producer	Volume
Warehouse 1	NL	Warehouse 3	GB	BEER A	15360
Warehouse 1	NL	Warehouse 3	GB	BEER B	15760
Warehouse 2	NL	Warehouse 3	GB	WINE A	23988
Warehouse 2	NL	Warehouse 3	GB	BEER B	32640
Warehouse 3	GB	Warehouse 4	IE	WINE A	12480
Warehouse 3	GB	Warehouse 4	IE	WINE B	19600



# Interactive EMRA Visualisations (3)

## Sanitised Brand Movements

- ☒ Open view ☐ Private view
- ☒ Show volumes ☐ Hide Volumes
- ☐ Simulation On ☒ Simulation Off

Warehouse font size: 1.2em  
Volume font size: 1.2em

### Countries

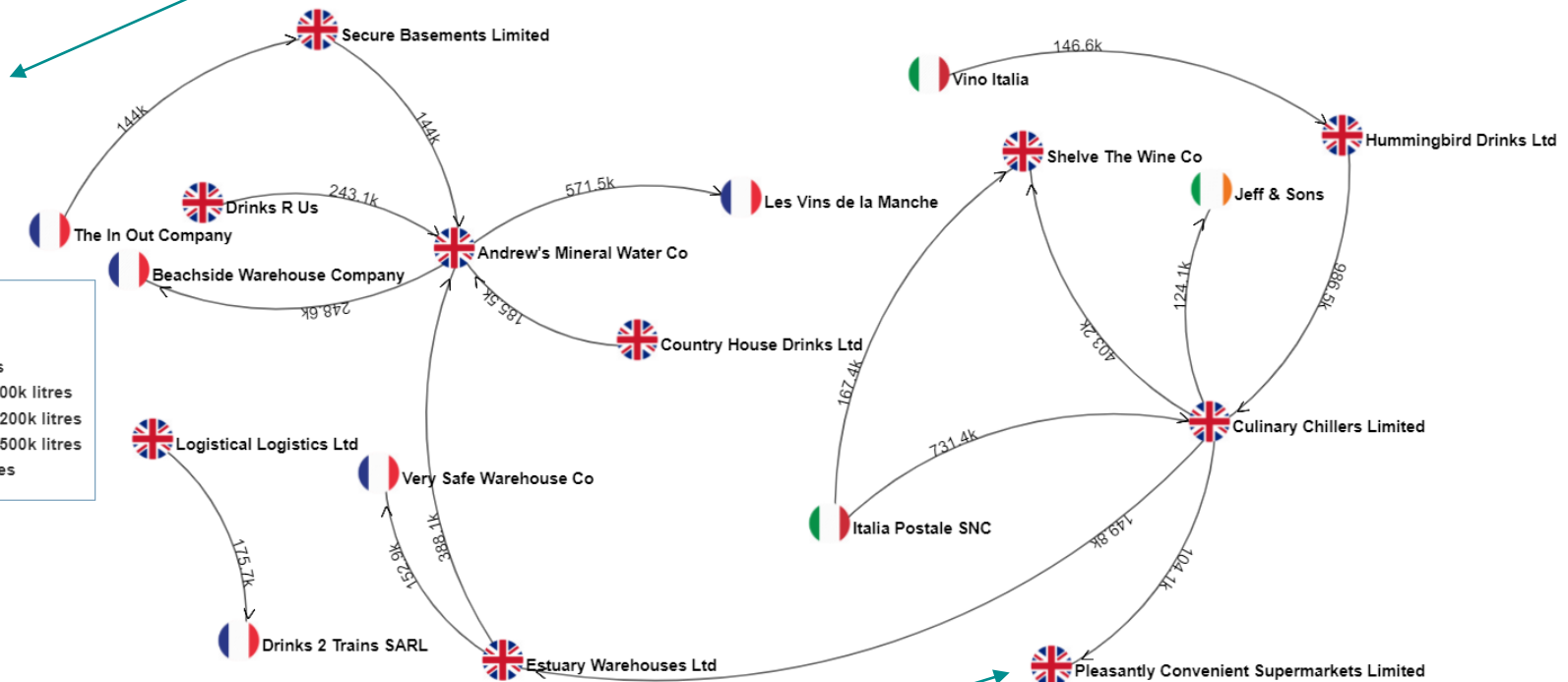
- ☒ BE
- ☒ DK
- ☒ ES
- ☒ FR
- ☒ GB
- ☒ IE
- ☒ IT
- ☒ NL

### Volume threshold

- ☐ Less than 50k litres
- ☐ Between 50k and 100k litres
- ☒ Between 100k and 200k litres
- ☒ Between 200k and 500k litres
- ☒ More than 500k litres

**Filter boxes** – allow the user to focus on specific areas of interest to them.

## Force-Directed Graph



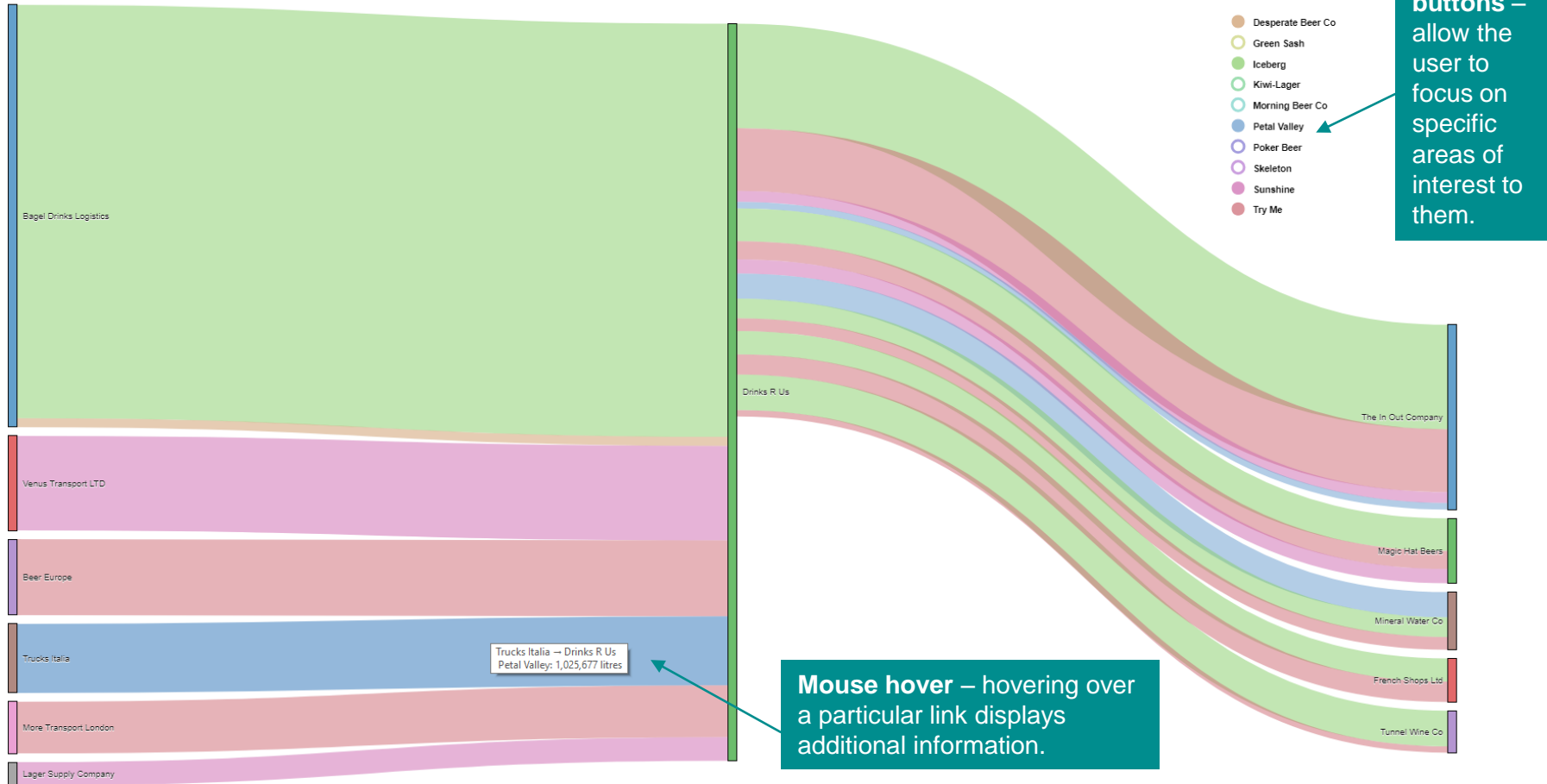
**Nodes and links** – can be moved and positioned by the user to arrange into e.g. supply chains.



HM Revenue  
& Customs

# Interactive EMRA Visualisations (4)

## Anonymised Warehouse Movements



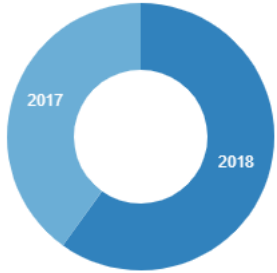
HM Revenue  
& Customs

Sankey Diagram

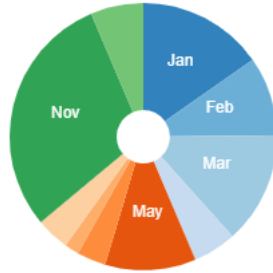
# Interactive EMRA Visualisations (5)

EMRA-Vis Hits **688** selected out of **688** records | [Reset All](#)

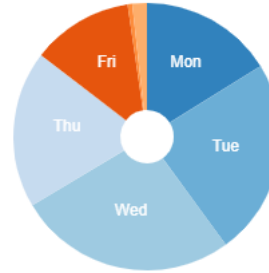
Year reset



Month reset

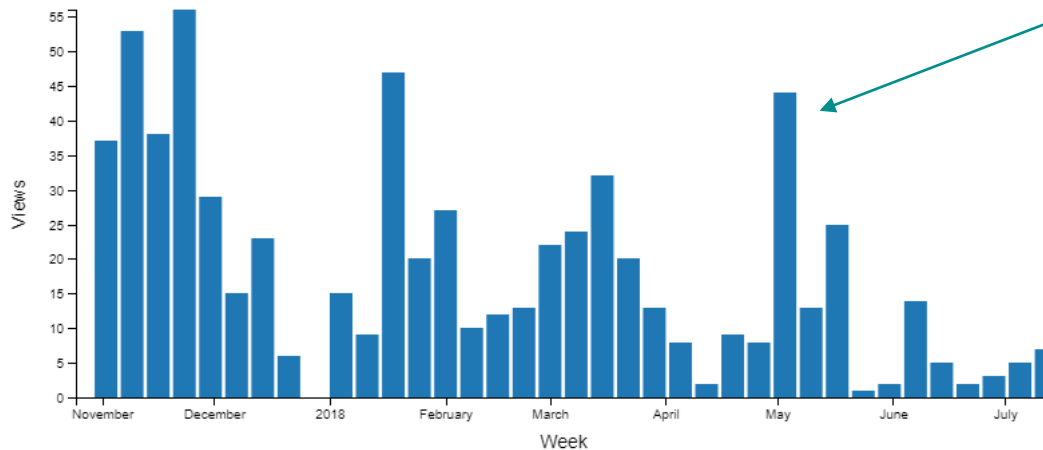


Day reset

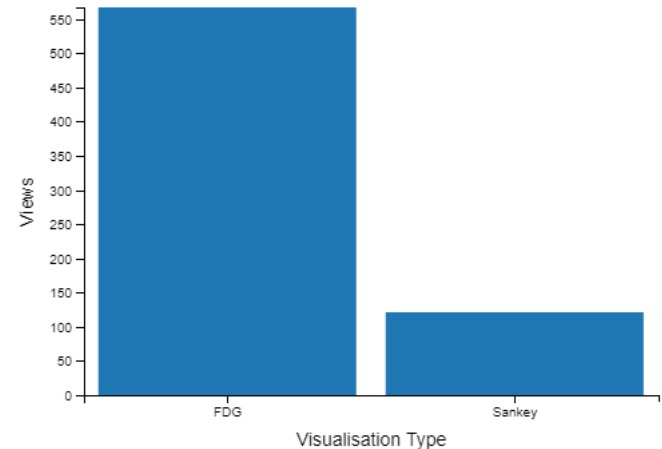


**Filterable charts** – allow the user to focus on specific areas of interest to them.

Views by Week



Type of Visualisation



HM Revenue  
& Customs

Usage Dashboard

**Dashboard** allows us to keep track of which visualisations are being used and when.

# Reproducible analysis for publications



## Issues:

- Large number of national statistics published
- A very manual process (e.g. copy/paste) and time consuming
- Potential errors

## Benefits:

- Collaborate on development and review
- Publish reproducible documents

## Statistical software: R

- Comprehensive set of statistical libraries - KnitR library

## R Studio

- Auto completion, indentation, highlighting
- Linked documentation
- Debugging tools

## Version Control: Gitlab

- Allows collaboration between analysts
- Easily share, amend and reuse code





# KnitR and R Markdown



## KnitR:

- R library to unify statistical analysis and reporting
- Define R functions to produce analysis and outputs
- Documents (text, tables, graphics) are produced dynamically
- Write a mix of markdown and R code (R Markdown)
- Generate reports in .html, .docx, .pdf

## R Markdown:

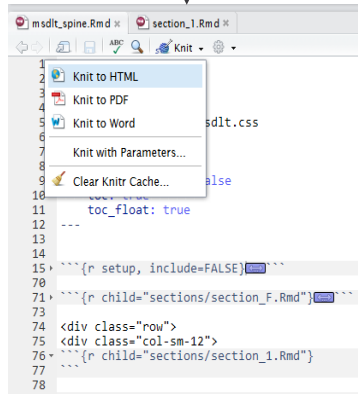
- Independent of R
- Simplified way of typesetting HTML



# Knitting National Statistics



```
88 same = 'similar to'
89 uk.res.no.yr.pct <- get_pct_chg(b=before, a=after, s=same, v=vals, t=1, r=1)
90
91
92
93 The seasonally adjusted estimate of the number of residential property transactions 'r uk.res.sa.mth.pct'
94 between 'r get_mthyr(prev.month)' and 'r get_mthyr(cur.month)'. This month's seasonally adjusted figure is 'r
95 uk.res.sa.yr.pct' the
96 same month last year, 'r uk.res.sa.yr.2.pct' transactions in 'r get_mthyr(prev.2.year)'. Chart 1A shows the historic
97 time series for residential property transactions.
98
99 Caution should be used making comparisons of transactions between 'r get_mthyr(cur.month)' and 'r get_mthyr(prev.year)'
100 as some
101 taxpayers may have changed their behaviour as they considered the result of the June 2017 General Election,
102 and the EU referendum in June 2016.
103
104 For 'r get_mthyr(cur.month)' the number of non-adjusted residential transactions was about 'r uk.res.no.mth.pct'
105 'r get_mthyr(prev.month)'. The number of non-adjusted residential transactions was 'r uk.res.no.yr.pct' 'r
106 get_mthyr(prev.year)'.
107
108 The figures for the three most recent months are provisional and therefore subject to revision.
109
110 [Reviewed 'r get_full_date(release_date)' / Next Review 'r get_full_date(next_release_date)']
111
112 ```{r, echo=FALSE}
113 graph.data %>%
```



## Current Financial Year

### Commentary and Charts

#### Key messages

The provisional seasonally adjusted UK property transaction count for May 2017 was 105,490 residential and 13,110 non-residential transactions.

#### Residential transactions

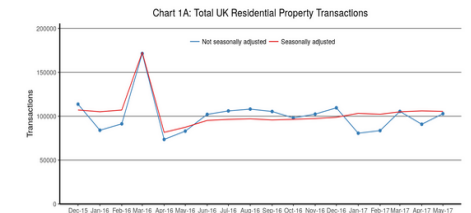
The seasonally adjusted estimate of the number of residential property transactions decreased by 0.5% between April 2017 and May 2017. This month's seasonally adjusted figure is 20.6% higher compared with the same month last year, and has increased by 20.6% compared with transactions in May 2015. Chart 1A shows the historic time series for residential property transactions.

Caution should be used making comparisons of transactions between May 2017 and May 2016 as some taxpayers may have changed their behaviour as they considered the result of the June 2017 General Election, and the EU referendum in June 2016.

For May 2017 the number of non-adjusted residential transactions was about and has increased by 13.3% than in April 2017. The number of non-adjusted residential transactions was 23.9% higher compared with May 2016.

The figures for the three most recent months are provisional and therefore subject to revision.

[Reviewed 21 June 2017 / Next Review 21 July 2017]



HM Revenue  
& Customs

# Useful links and further information

These run best in Google Chrome

- Employment Allowance Publication:  
[https://kai-data-exploitation.github.io/ns-employment-allowance/empall\\_spine.html](https://kai-data-exploitation.github.io/ns-employment-allowance/empall_spine.html)
- Stamp Duty Publication:  
[https://kai-data-exploitation.github.io/monthly-stamp-duty-ns-output/msdlt\\_spine.html](https://kai-data-exploitation.github.io/monthly-stamp-duty-ns-output/msdlt_spine.html)
- D3 applications in JavaScript are located here: <https://d3js.org/>

