

Analysis of Country Mentions in DR, BBC and ARD

Table of contents

1 Set-up	1
2 Numbers	2
2.1 dataframes	2
2.1.1 country_analysis	2
2.1.2 per_day	2
2.1.3 countries_as_placings	2
2.2 Summaries	3
3 Plots	3
3.1 Barplot - top 10	3
3.2 Barplot - unique countries per day	4
3.3 Decay plot	5
3.4 World maps	6
3.4.1 DR	6
3.4.2 BBC	6
3.4.3 ARD	7
3.5 Violin plot	7
3.6 Correlation	8

1 Set-up

Dataen rengøres således at der for hvert medie er samme antal artikler per dag.

```
Rows: 1321 Columns: 6
-- Column specification -----
Delimiter: ","
chr (4): URL, media, type_article, countries
dbl (1): is_nyhed
date (1): date
```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.

I denne analyse medtages 828 artikler

2 Numbers

2.1 dataframes

2.1.1 country_analysis

```
Rows: 274  
Columns: 7  
Groups: media [3]  
# media <chr> "ARD", "ARD", "ARD", "ARD", "ARD", "ARD", "ARD", "A-  
# countries <chr> "Germany", "United States of America", "Russ-  
# mentions <int> 150, 78, 55, 41, 37, 36, 29, 25, 24, 20, 19, ~  
# total_articles <int> 276, 276, 276, 276, 276, 276, 276, 276, 276, ~  
# mentions_per_100_articles <dbl> 54.347826, 28.260870, 19.927536, 14.855072, ~  
# total_country_mentions <int> 855, 855, 855, 855, 855, 855, 855, 855, 855, ~  
# share_of_voice <dbl> 17.5438596, 9.1228070, 6.4327485, 4.7953216, ~
```

Der kigges på 2 ens men også ret forskellige parametre

mentions_per_100_articles: siger sig selv. Hvis man tager 100 artikler, hvor mange gange er omtalte land så nævnt?

share of voice: hvor stor en andel af al land-omtale for givent medie står dette land for

2.1.2 per_day

2.1.3 countries as placings

```
Rows: 274  
Columns: 4  
# media <chr> "ARD", "ARD", "ARD", "ARD", "ARD", "ARD", "ARD", "A~  
# placing <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 1~
```

```
$ countries <chr> "Germany", "United States of America", "Russia", ...
$ mentions_per_100_articles <dbl> 54.347826, 28.260870, 19.927536, 14.855072, ...
```

2.2 Summaries

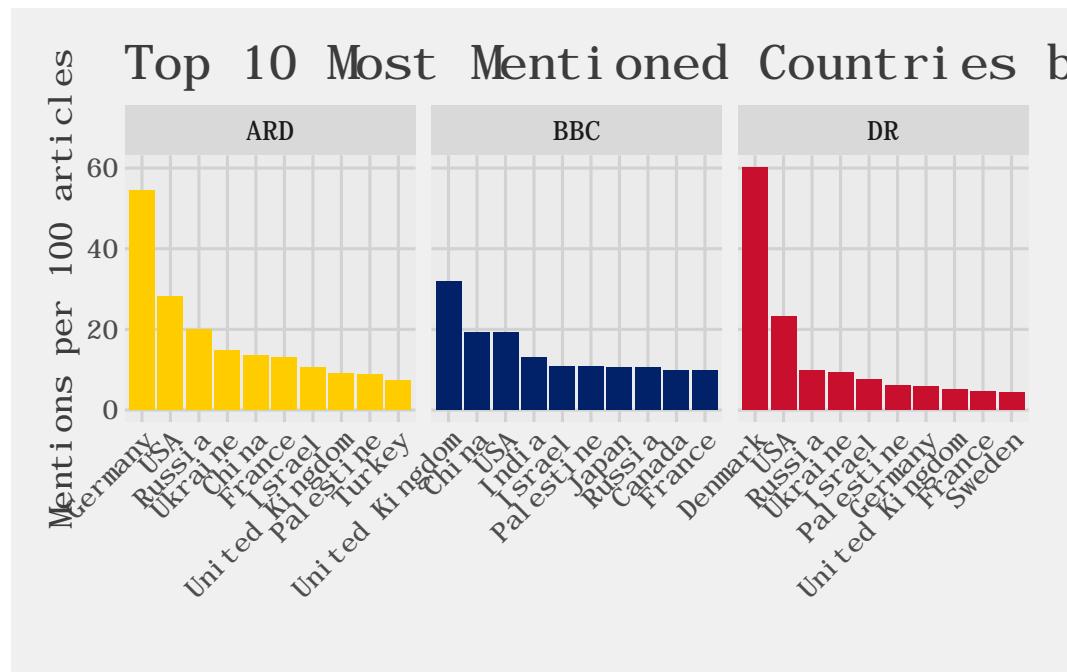
```
# A tibble: 3 x 5
  media n_countries more_than_once gini_coefficient avg_mentions
  <chr>      <int>        <int>          <dbl>           <dbl>
1 ARD         93            52       0.660        2.16
2 BBC        111            53       0.655        1.85
3 DR          70            34       0.708        1.48
```

Generelle stats for hvert medie:

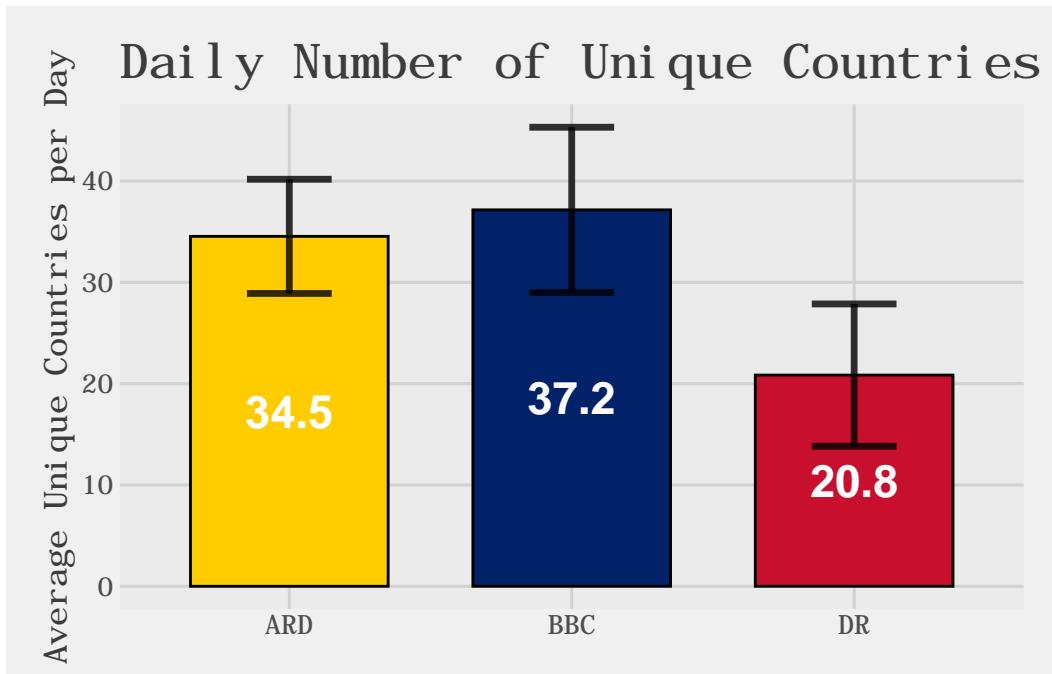
1. n_countries: antal af unikke lande nævnt i alle artiklerne
2. more_than_once: antal af unikke lande nævnt mere end 1 gang per 100 artikler
3. gini_coefficient: parametre fra økonomi, 0 er perfekt lighed og 1 maximum ulighed
4. avg_mentions: gennemsnitligt antal gange en lande bliver nævnt per 100 artikel (4 mest nævnt labde frasorteret)

3 Plots

3.1 Barplot - top 10



3.2 Barplot - unique countries per day



Undersøg om der er signifikant forskel:

ANOVA Table (type II tests)

Effect	DFn	DFd	F	p	p<.05	pes
1 media	2	36	20.276	1.27e-06	*	0.53

ANOVA Table (type II tests)

Effect	DFn	DFd	F	p	p<.05	pes
1 media	1	24	30.061	1.23e-05	*	0.556

ANOVA Table (type II tests)

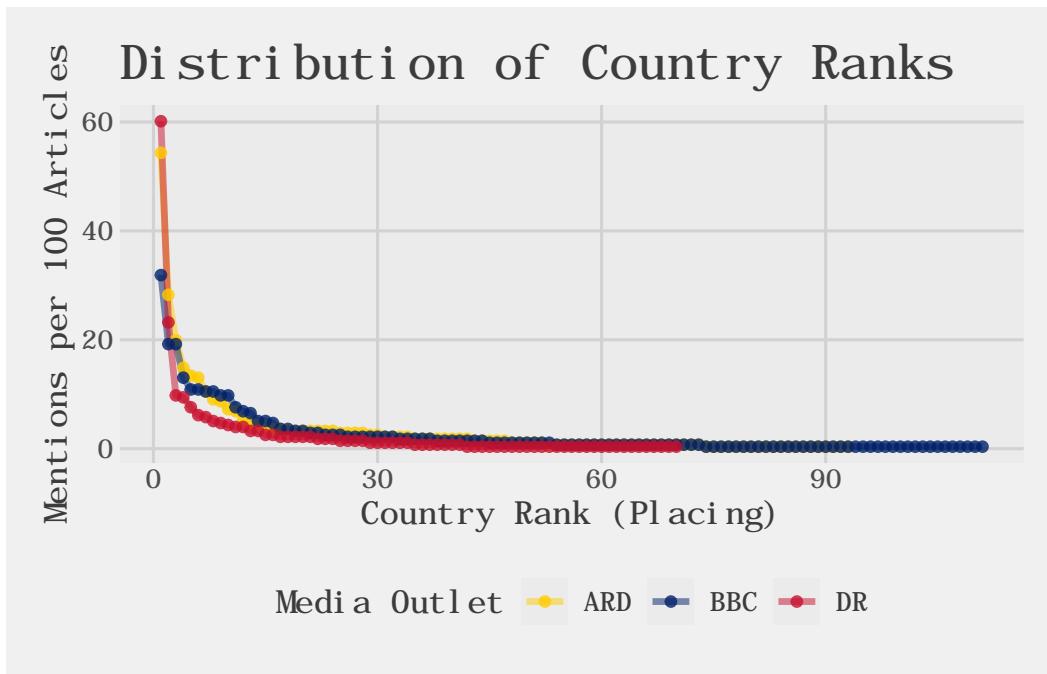
Effect	DFn	DFd	F	p	p<.05	pes
1 media	1	24	29.86	1.29e-05	*	0.554

Her er en indsigt:

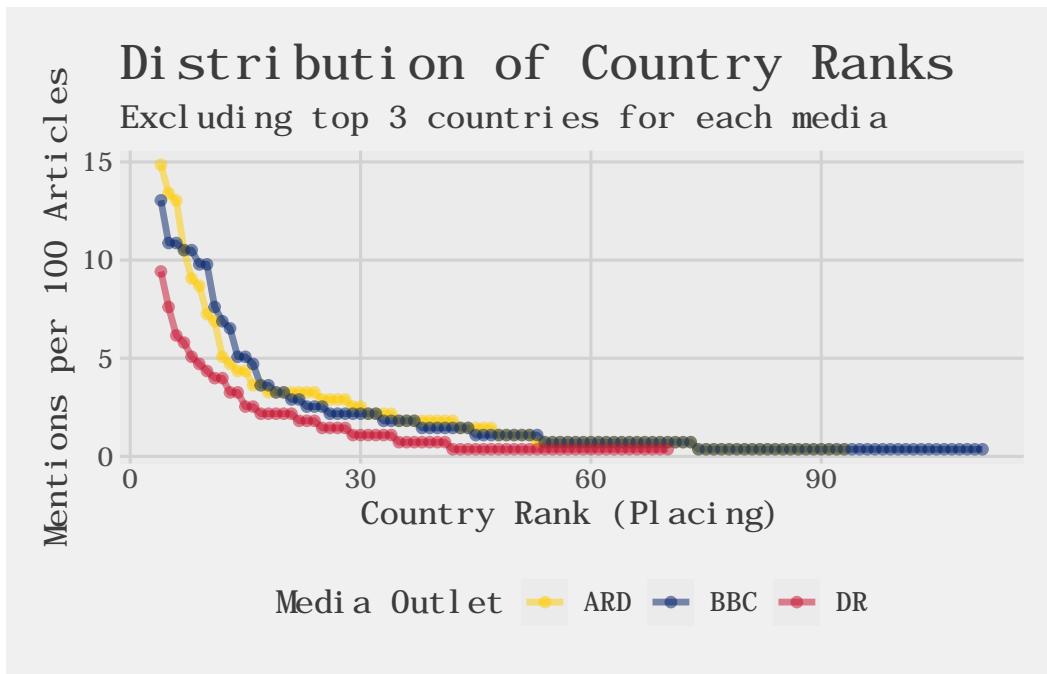
DR skriver om signifikant færre unikke lande end de to andre medier

3.3 Decay plot

Lad os kigge på fordelingen af mentions



Der zoomes ind, ved at frasortere 3 mest nævnte lande fra hvert medie

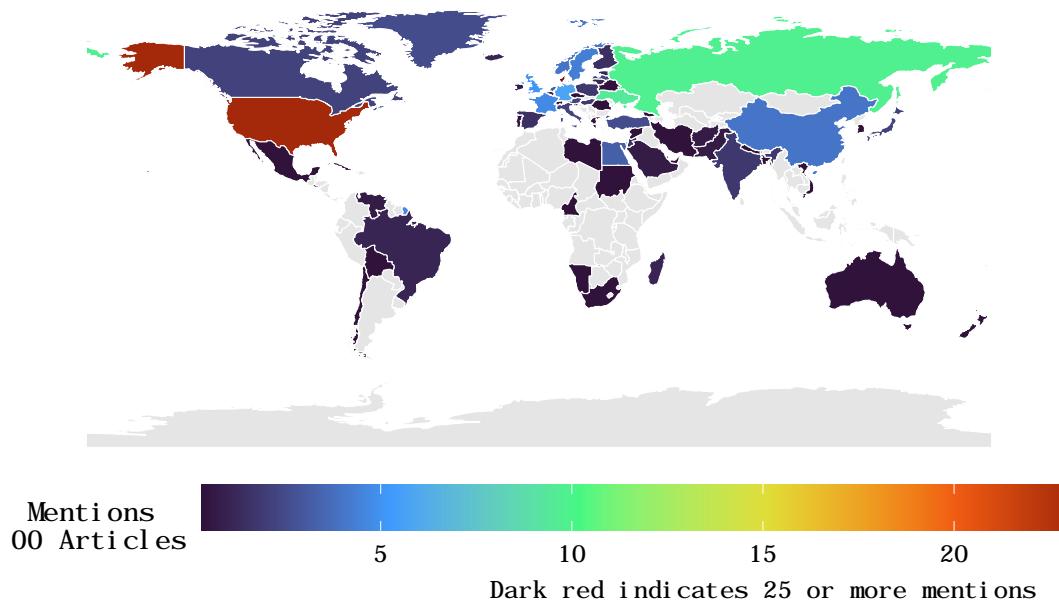


DR omtaler færre lande end de to andre medier i alle dele af fordelingen.

3.4 World maps

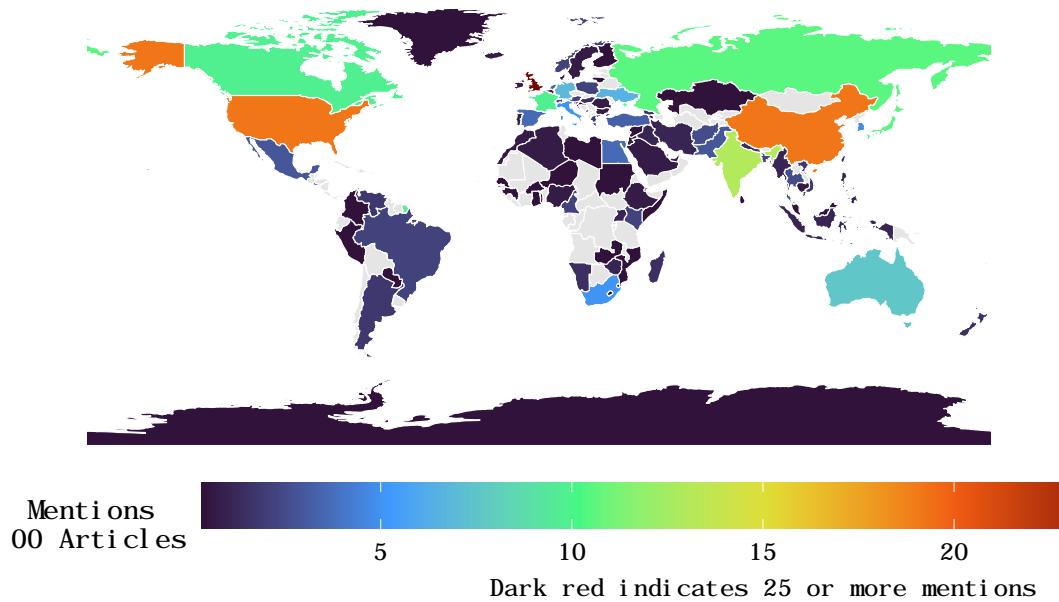
3.4.1 DR

DR: Mentions of Countries



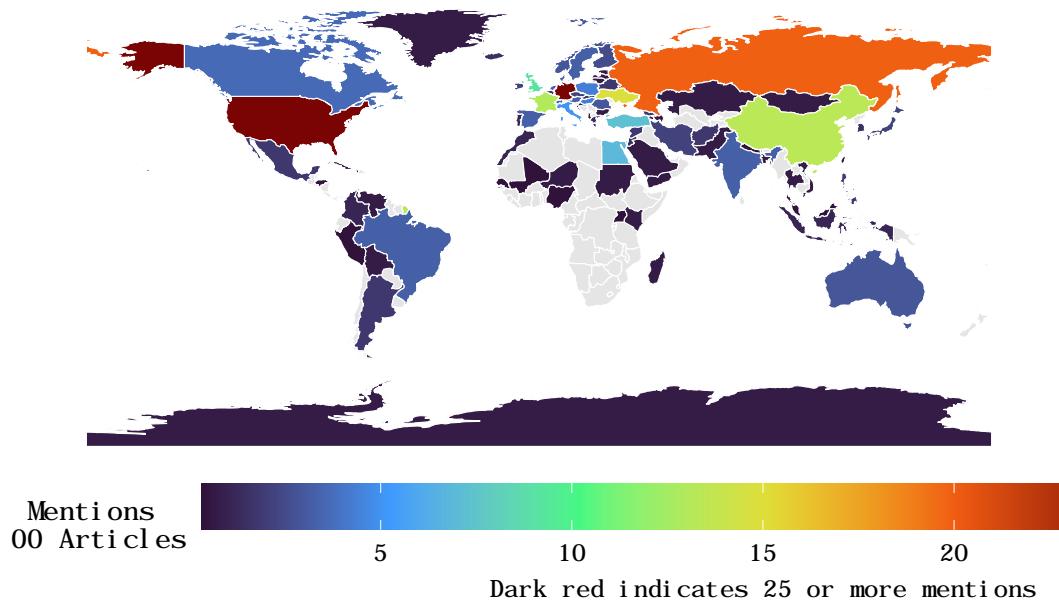
3.4.2 BBC

BBC: Mentions of Countries



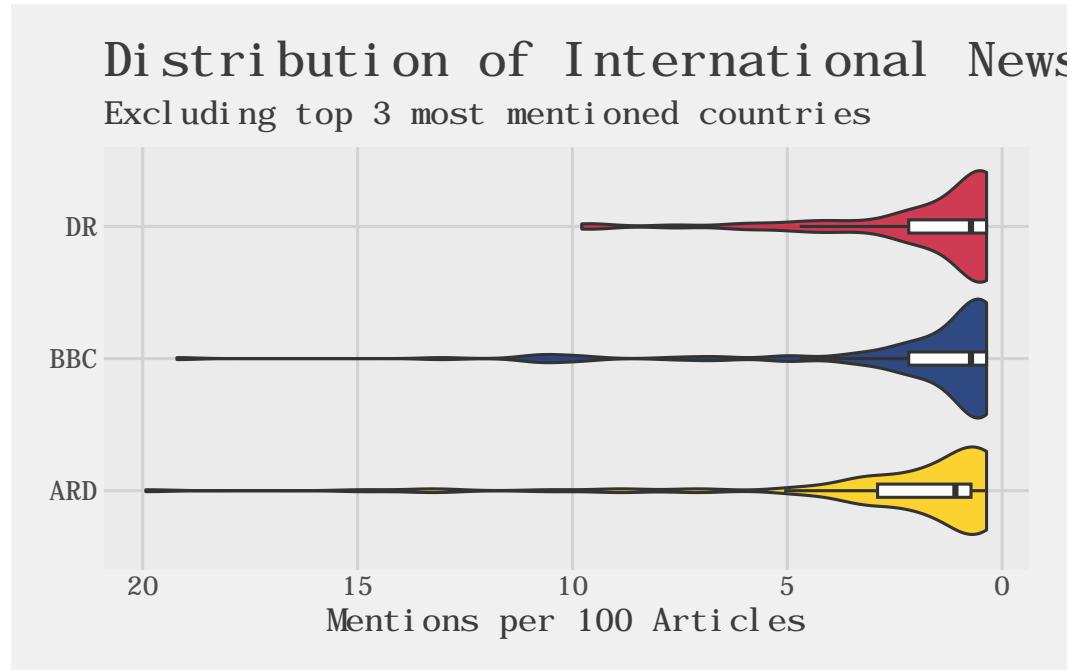
3.4.3 ARD

ARD: Mentions of Countries



3.5 Violin plot

Det her plot siger ikke så meget



3.6 Correlation

```
# A tibble: 3 x 3
  media1 media2 correlation
  <chr>   <chr>      <dbl>
1 BBC     ARD       0.507
2 DR      ARD       0.691
3 DR      BBC       0.539
```

Indsigt: DR og ARD skriver om meget ens lande