

ratio_smallPartD

Zongyan Wang

Thursday, December 03, 2015

This document will load the PartD_newDrug.csv output by PartD_data_drug_pickout.R.

Where those physicians' from?

What is ratio of brand_name / generic_name in this small PartD dataset, is it higher than the average of the rest of the dataset.

Are the physicians appearing in this small PartD use brand_name produced in one medical company very often?

We will also look at what the role those physician play in the referral data.

Our analysis begin here.

Where those physicians' from?

```
source("newDrug_States.R")
```

```
## Loading required package: data.table
```

```
## Warning: package 'data.table' was built under R version 3.1.3
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
## Loading required package: plyr
```

```
## Warning: package 'plyr' was built under R version 3.1.3
```

```
## Loading required package: dplyr
```

```
## Warning: package 'dplyr' was built under R version 3.1.3
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
##
```

```
## The following objects are masked from 'package:plyr':
```

```
##
```

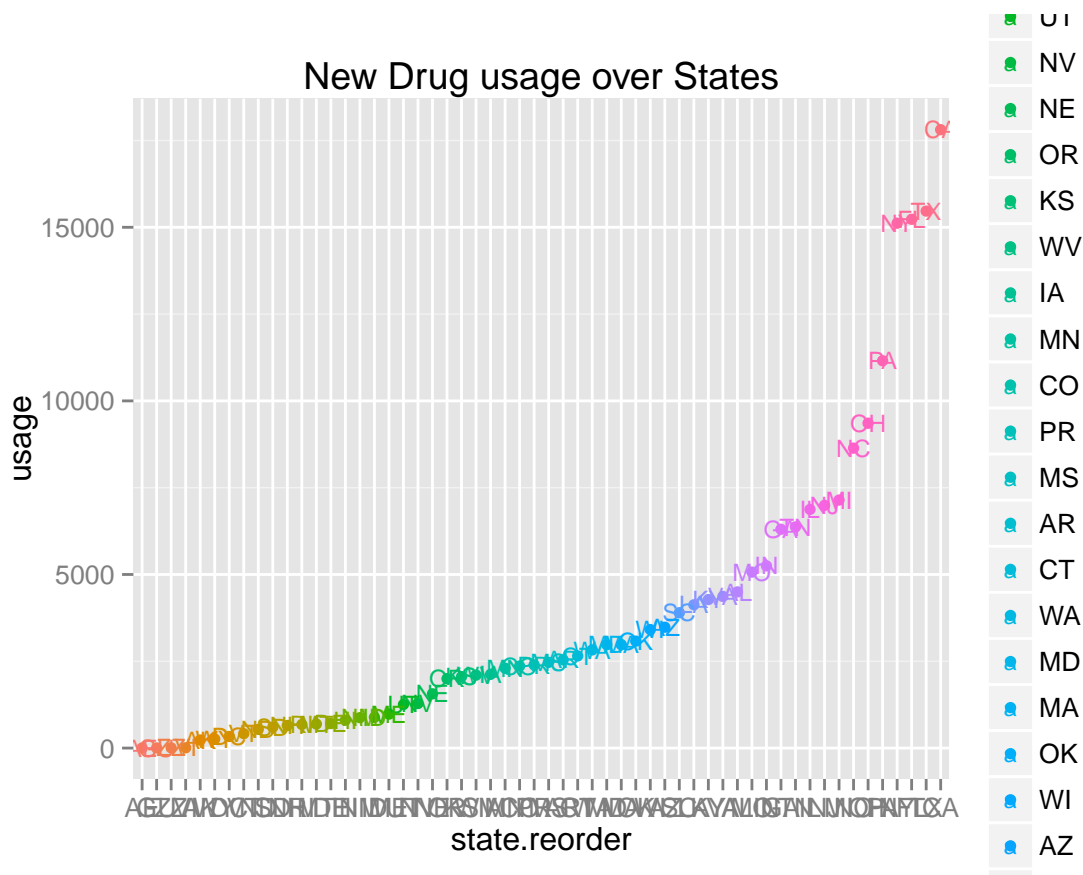
```
##      arrange, count, desc, failwith, id, mutate, rename, summarise,
```

```
##      summarize
```

```
##
```

```
## The following objects are masked from 'package:data.table':
##
##   between, last
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

##
Read 80.0% of 212462 rows
Read 212462 rows and 19 (of 19) columns from 0.027 GB file in 00:00:03
## The total new drug total usage over States
```



```
source("Combine_PartD_drugList.R")
```

```
## Loading required package: tidyr
```

```
## Warning: package 'tidyr' was built under R version 3.1.3
```

```
## check the overlapThere are lots of overlap here. Merge the rowsThe total new drug total usage over S
```

