

Assign2.1

Ambalika

16/04/2021

Statistical Summary for cases of Hepatitis A in different US States from 1966-2011

Research Question : To analyse the trend of Hepatitis A cases in US.

We shall be analysing the rate and case count per state all over US over state distribution and yearly count. This shall help us analyse the year which had the highest count of Hepatitis B and also the state with the highest count. This can help us further help answer as to whether there was a downward trend over the years in Hepatitis due to advancement of technology and which state has the highest count and which state should CDC look into to control this disease.

```
## Warning: package 'dslabs' was built under R version 4.0.5
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
##  
## 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:stats':  
##  
##   filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
summary(us_contagious_diseases)
```

```
##      disease      state      year      weeks_reporting
## Hepatitis A:2346 Alabama : 315 Min. :1928 Min. : 0.00
## Measles :3825 Alaska : 315 1st Qu.:1950 1st Qu.:31.00
## Mumps :1785 Arizona : 315 Median :1975 Median :46.00
## Pertussis :2856 Arkansas : 315 Mean :1971 Mean :37.38
## Polio :2091 California: 315 3rd Qu.:1990 3rd Qu.:50.00
## Rubella :1887 Colorado : 315 Max. :2011 Max. :52.00
## Smallpox :1275 (Other) :14175
##      count      population
## Min. : 0 Min. : 86853
## 1st Qu.: 7 1st Qu.: 1018755
## Median : 69 Median : 2749249
## Mean : 1492 Mean : 4107584
## 3rd Qu.: 525 3rd Qu.: 4996229
## Max. :132342 Max. :37607525
##      NA's :214
```

```
diseases <- head(us_contagious_diseases,1000)
unique(diseases[c("disease")])
```

```
##      disease
## 1 Hepatitis A
```

Figure 1 : Below graph displays the rate of Hepatitis A cases(count vs population) against each year

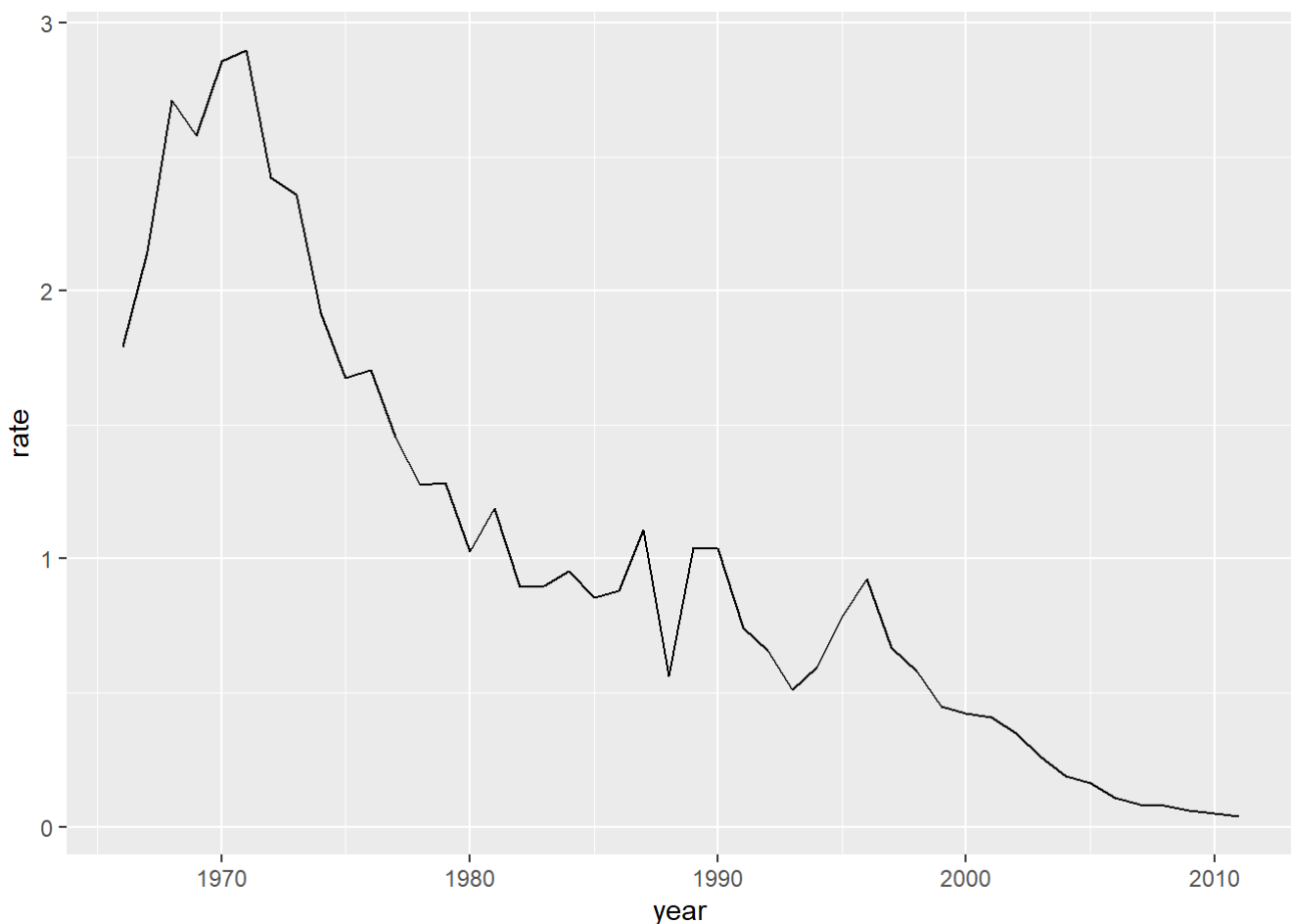


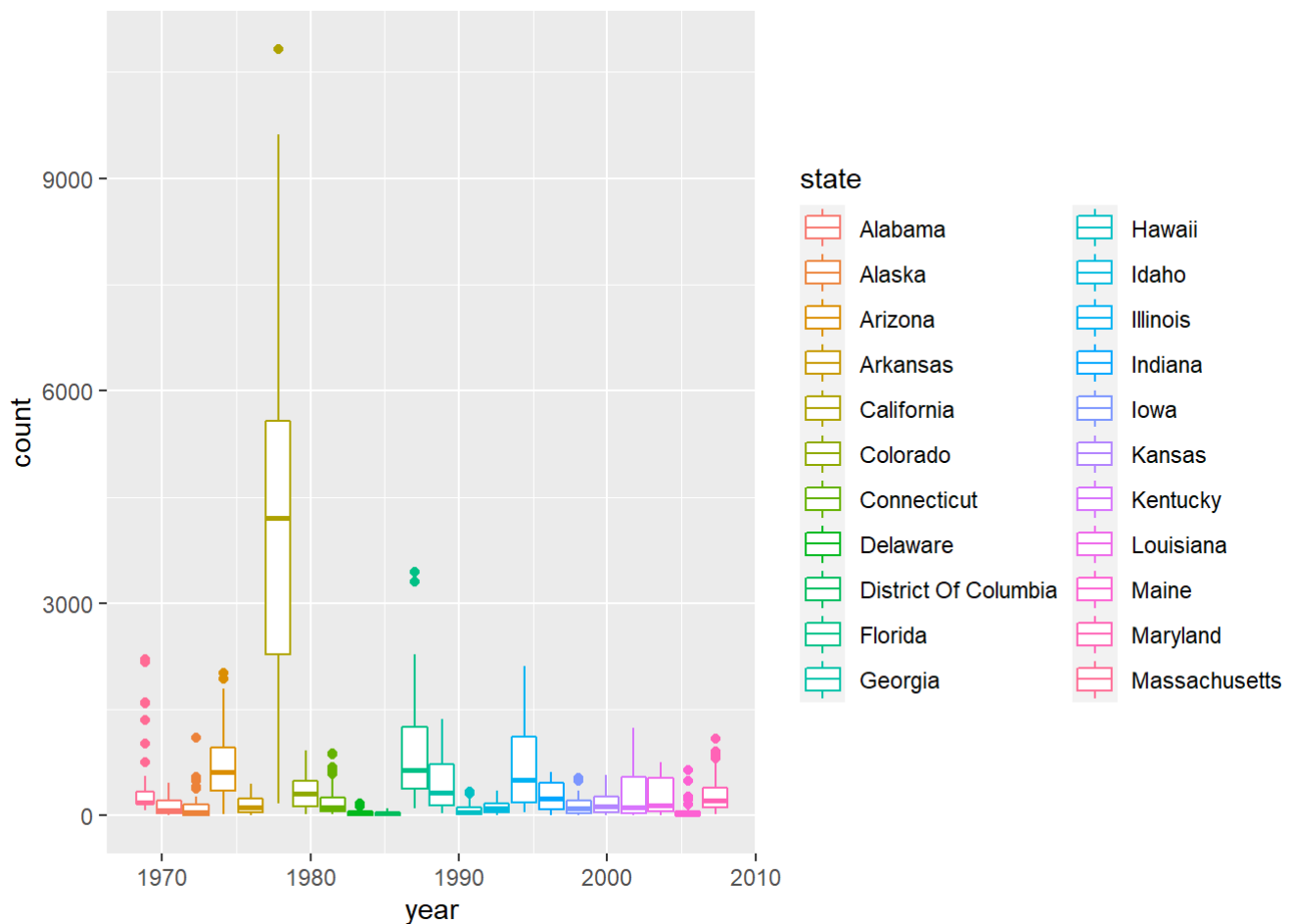
Table 1 : The below is a tabular representation of total case counts of Hepatitis A per US state from the year 1966 - 2011

Cases of Hepatitis A in each US state

state	number_of_cases
Alabama	6298
Alaska	5647
Arizona	31900
Arkansas	7494
California	191314
Colorado	15057
Connecticut	9849
Delaware	1901
District Of Columbia	1537
Florida	42194
Georgia	19832
Hawaii	3585
Idaho	5464
Illinois	32882
Indiana	11833
Iowa	6481
Kansas	8148
Kentucky	15955
Louisiana	12899
Maine	3181
Maryland	14603
Massachusetts	15952

Figure 2 : The below table is a boxplot representing the number of cases over the years recorded for each US state.

```
## `summarise()` has grouped output by 'year'. You can override using the `.groups` argument.
```



Hello

Table 2 : Below is a tabular representation of rate of Hepatitis A cases in the State of California(California has recorded the highest number of cases as compared to other US states).

Rate of Hepatitis A in the State of California for each year

disease	state	year	weeks_reporting	count	population	rate
Hepatitis A	California	1966	52	5933	18429575	3.22
Hepatitis A	California	1967	52	7480	18831882	3.97
Hepatitis A	California	1968	52	10821	19219725	5.63
Hepatitis A	California	1969	50	9051	19593348	4.62
Hepatitis A	California	1970	50	9422	19953134	4.72
Hepatitis A	California	1971	51	9623	20300216	4.74
Hepatitis A	California	1972	47	7190	20638390	3.48
Hepatitis A	California	1973	49	6875	20972407	3.28
Hepatitis A	California	1974	48	5547	21307362	2.60
Hepatitis A	California	1975	51	5134	21648695	2.37

disease	state	year	weeks_reporting	count	population	rate
Hepatitis A	California	1976	51	5590	22002198	2.54
Hepatitis A	California	1977	44	4741	22374038	2.12
Hepatitis A	California	1978	45	4301	22770801	1.89
Hepatitis A	California	1979	44	4101	23199549	1.77
Hepatitis A	California	1980	33	3673	23667902	1.55
Hepatitis A	California	1981	44	5147	24181703	2.13
Hepatitis A	California	1982	44	3726	24737390	1.51
Hepatitis A	California	1983	42	3922	25328529	1.55
Hepatitis A	California	1984	43	6021	25948031	2.32
Hepatitis A	California	1985	39	5371	26588046	2.02
Hepatitis A	California	1986	39	5802	27239883	2.13
Hepatitis A	California	1987	45	5956	27893955	2.14
Hepatitis A	California	1988	23	2717	28539754	0.95
Hepatitis A	California	1989	34	4692	29165865	1.61
Hepatitis A	California	1990	42	4870	29760021	1.64
Hepatitis A	California	1991	36	3791	30311890	1.25
Hepatitis A	California	1992	36	3594	30821629	1.17
Hepatitis A	California	1993	26	2674	31292538	0.85
Hepatitis A	California	1994	27	3367	31728420	1.06
Hepatitis A	California	1995	41	5130	32133526	1.60
Hepatitis A	California	1996	45	5547	32512506	1.71
Hepatitis A	California	1997	28	4438	32870358	1.35
Hepatitis A	California	1998	35	2803	33212395	0.84
Hepatitis A	California	1999	38	2349	33544208	0.70
Hepatitis A	California	2000	39	2261	33871648	0.67
Hepatitis A	California	2001	40	1599	34199784	0.47
Hepatitis A	California	2002	49	1364	34529758	0.40
Hepatitis A	California	2003	46	1045	34861711	0.30
Hepatitis A	California	2004	48	788	35195792	0.22
Hepatitis A	California	2005	49	905	35532154	0.25
Hepatitis A	California	2006	52	688	35870957	0.19
Hepatitis A	California	2007	51	312	36212364	0.09

disease	state	year	weeks_reporting	count	population	rate
Hepatitis A	California	2008	52	337	36556548	0.09
Hepatitis A	California	2009	52	239	36903684	0.06
Hepatitis A	California	2010	49	201	37253956	0.05
Hepatitis A	California	2011	49	176	37607525	0.05

Conclusion

1. From Figure 1 , we can that the highest number of cases were recorded in the year 1970.It further represents a downward trend in the cases, from which we can assume that there has been an advancement in the field of technology and health and more awareness.
2. From Table 1, we can conclude that California has had the highest number of cases from the year 1966-2011, followed by Illinois and Arizona.
3. From Figure 2, we can conclude as above that California has the highest number of cases, with an average of 4000 cases, average of California being significantly higher than the average of other US states. Hence California needed high time intervention at that time.
4. From Table 2, we can conclude that the rate of Hepatitis A cases in the California reduced from around 5 to 0.05. Which marks a significant decrease in the rate of cases, showing true progress in field of health sector. The max recorded rate was 5.63 in the year 1968 and the lowest rate being 0.05 in the year 2011.

Hence we can conclude that the cases of Hepatitis B have gone down significantly in the United States of America from the year 1966 to 2011 especially the highly affected state of California.