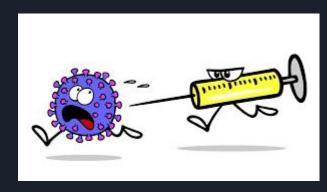
Vaccination
Coverage Among
Children in US
1995-2021

Miraj Patel

Why?

- Vaccination is a "hot" topic
- Firm believer of vaccinations and science
- Current role includes making diagnostic kits
- Roche collaborates with other companies to make vaccines
- Knowledge of trends of vaccination coverage will help decide where to settle down





Data Set 1

- Data set collected from CDC website
- Shows vaccination coverage across the US and its territories
- Data Range: 2011-2021
- ~130,000 observations with 18 variables
- Limitations: may not account for other factors influencing vaccination rates, such as socioeconomic status, access to healthcare, or public health policies

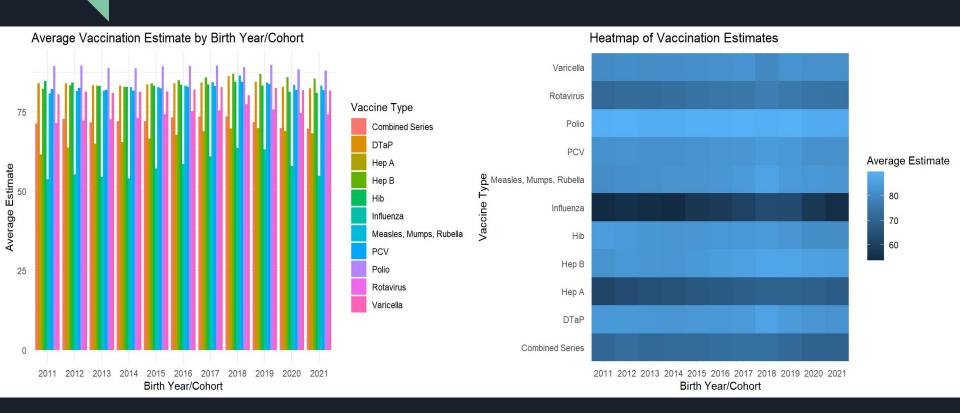
Vaccines in this data set

- Varicella
- o Polio
- PCV
- o MMR (Measles, Mumps, Rubella)
- DTaP (Diphtheria, Pertussis, Tetanus)
- o Hib
- Hep B
- Combined Series (includes all of the above)
- Hep A
- Rotavirus
- o Influenza (flu)

Data Set - Wide

\$	Vaccine Type	Geography	Birth.Year.Birth.Cohort	24m ‡	35m [‡]	13m [‡]	19m [‡]	3m 🔻	5m ‡	7m ‡	01d [‡]	02d [‡]	03d [‡]
2171	Нер В	Delaware	2011	93.90000	93.90000	92.65000	91.00000	99.50000	91.20000	72.90000	78.60000	82.60000	84.10000
2167	Нер В	Connecticut	2018	95.15000	95.40000	94.47500	93.05000	99.45000	93.20000	76.85000	78.30000	83.15000	84.00000
2344	Нер В	Massachusetts	2021	98.50000	99.00000	97.15000	98.50000	99.40000	97.80000	88.45000	85.00000	86.00000	86.10000
2428	Нер В	New Hampshire	2017	97.55000	97.55000	96.87500	97.45000	99.25000	96.35000	91.30000	61.75000	80.95000	83.65000
2406	Нер В	Nebraska	2017	95.90000	96.40000	92.70000	94.75000	99.05000	92.25000	84.35000	80.65000	82.85000	85.10000
2172	Нер В	Delaware	2012	97.80000	97.90000	96.65000	97.10000	98.90000	87.40000	72.80000	73.10000	80.90000	83.90000
2653	Нер В	Rhode Island	2018	98.00000	98.80000	98.37500	97.95000	98.90000	97.15000	94.72500	71.55000	78.05000	79.25000
2655	Нер В	Rhode Island	2020	98.15000	98.15000	98.02500	97.95000	98.70000	97.80000	95.10000	79.90000	81.60000	82.30000
2484	Нер В	North Dakota	2018	95.60000	96.60000	94.85000	93.70000	98.60000	93.10000	88.52500	82.80000	85.20000	85.75000
2646	Нер В	Rhode Island	2011	96.00000	96.20000	96.45000	96.00000	98.60000	91.90000	91.30000	59.40000	69.00000	70.00000
2342	Нер В	Massachusetts	2019	95.40000	96.05000	92.47500	95.25000	98.50000	94.25000	81.02500	83.70000	87.45000	87.60000
2686	Нер В	Tennessee	2018	95.40000	95.65000	94.75000	93.35000	98.45000	92.10000	85.10000	71.65000	74.00000	74.90000
2177	Нер В	Delaware	2017	92.35000	92.95000	92.32500	90.40000	98.40000	93.45000	73.50000	72.10000	79.20000	81.75000
2379	Нер В	Missouri	2012	94.40000	95.30000	92.70000	92.70000	98.30000	92.80000	77.30000	65.90000	72.10000	72.90000
2302	Нер В	Louisiana	2012	95.80000	97.30000	94.85000	95.20000	98.30000	89.30000	82.85000	68.00000	77.50000	79.50000
2654	Нер В	Rhode Island	2019	96.95000	96.95000	97.37500	96.95000	98.30000	94.00000	93.72500	72.80000	78.15000	78.70000
2407	Нер В	Nebraska	2018	93.55000	94.30000	90.95000	90.35000	98.25000	92.65000	84.25000	85.90000	87.40000	87.70000

Data Set 1 Graphs



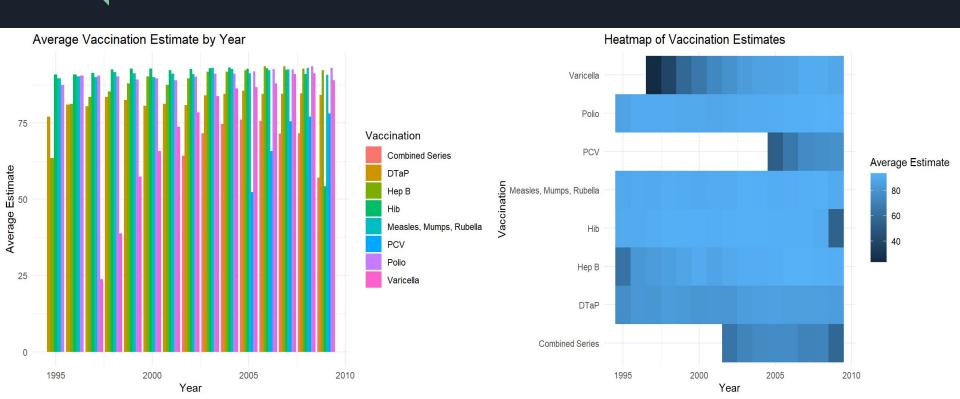
Data Set 2

- Data collected from Kaggle, which was collected from CDC 6 years ago
- Vaccination coverage by race or economic status (above or below poverty level)
- 1116 observations and 9 variables
- Data Range: 1995 2009
- Limitations:
 - Not all vaccines are included.
 - Geographical coverage, hard to tell which states or urban/city/rural

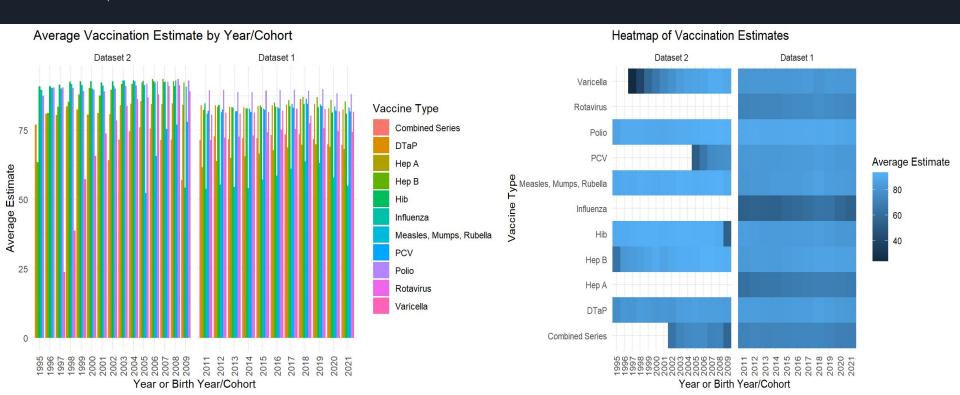
Vaccines in this data set

- Varicella
- o Polio
- PCV
- o MMR (Measles, Mumps, Rubella)
- DTaP (Diphtheria, Pertussis, Tetanus)
- o Hib
- o Hep B
- Combined Series (includes all of the above)

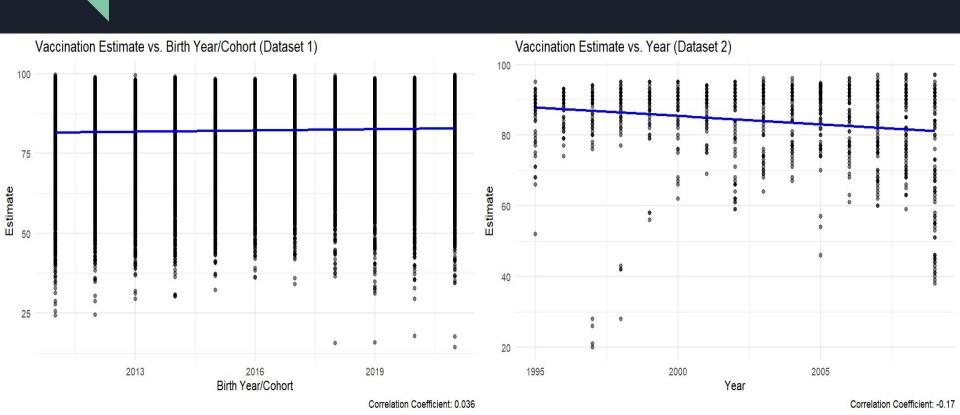
Data Set 2 Graphs



Combined Graph



Statistical Analysis - Correlation



Recommendations and Challenges

Recommendations

- For Data Set 1
 - Further analysis can be done filtering states or regions 1-10 (50 states are grouped into 10 regions)
- For Data Set 2
 - Further analysis filtering races or economic status (above or below poverty level)
- Find Data for 2010!

Challenges

- A lot of data and regex clean
- Hard to find common ground between the two datasets
 - Had to stick with to year vs estimate
 - Data set 1 grouped by states or regions vs Data set 2 grouped by races and economic status