**Outcome Variable EDA**

1. Merge
2. Summary stats
3. Changes over time: abortion\_rate, birth\_rate, pregnancy\_rate all decrease; hiv rate doesn’t change too much; std rate varies, almost all states just begin to decrease these years, but in some states still increase
4. Correlation: these outcome variables are all positive correlated.
5. Clustering: based on time-series change types between 2010 to 2017 to cluster 5 groups

* What does ‘adjusted’ in stds mean? How? Hiv\_data hasn’t been adjusted, so i use both original rate, not adjusted
* The rates for abortion, birth and pregnancy are also per 100,000 people?
* (Pregnancy - abortion) does not equa to birth rate, always larger
* Only 1 record of std rate smaller than hiv rate (std rate=0), maybe some problem

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| State, Year | abortion\_rate | birth\_rate | pregnancy\_rate | HIV Rate per 100,000 | STD Rate per 100,000 People |
| Maryland, 2021 |  |  |  | 16.2 | 0 |