Dataset description

The dataset analysed in this project was collected as part of an abortion surveillance conducted by CDC (Center for Chronic Disease Prevention and Health Promotion) and observes cases of legally induced abortions of certain age/racial/ethnic groups in the USA throughout 2018. The data was voluntarily provided by given states and reporting areas and was further processed by CDC.

Key statistics

Stats about abortions among Hispanic ethnic group:

Mean value: 4387.1875
Median value: 660.5
Standard deviation: 12948.51111490205

Stats about abortions among "15 or less" age group:

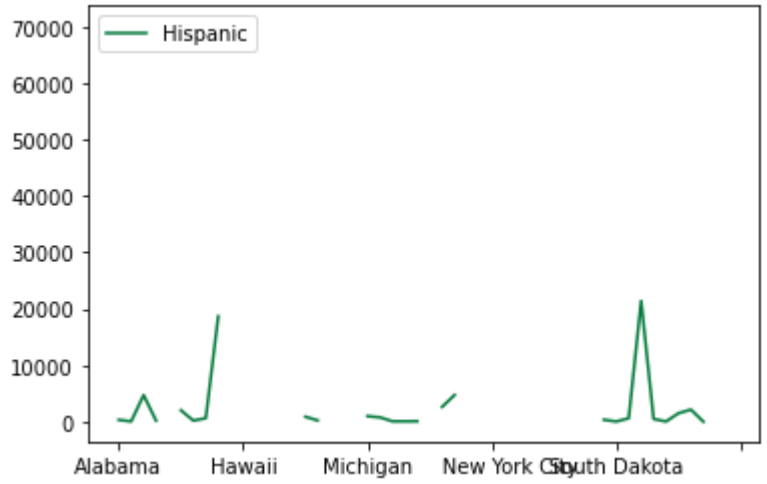
Mean value: 69.85365853658537
Median value: 22.0
Standard deviation: 210.1121796773821

Stats about abortions among "20–24" age group:

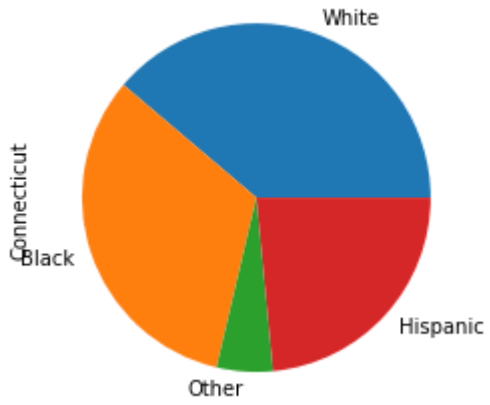
Mean value: 7347.34
Median value: 2288.5
Standard deviation: 24433.840645259992

Data visualisation

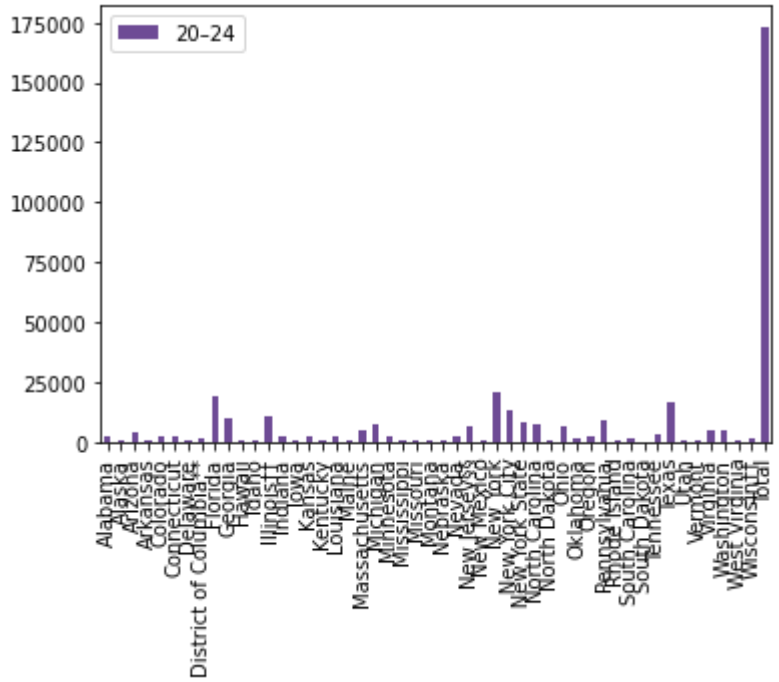
Graph showing abortions among Hispanic ethnic group



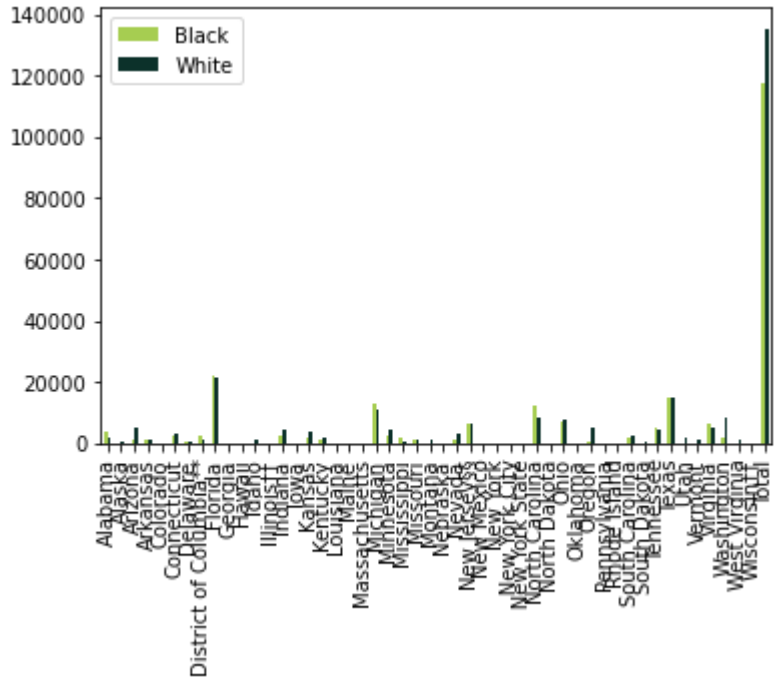
Pie-chart showing proportion of abortions in Connecticut by racial/ethnic group



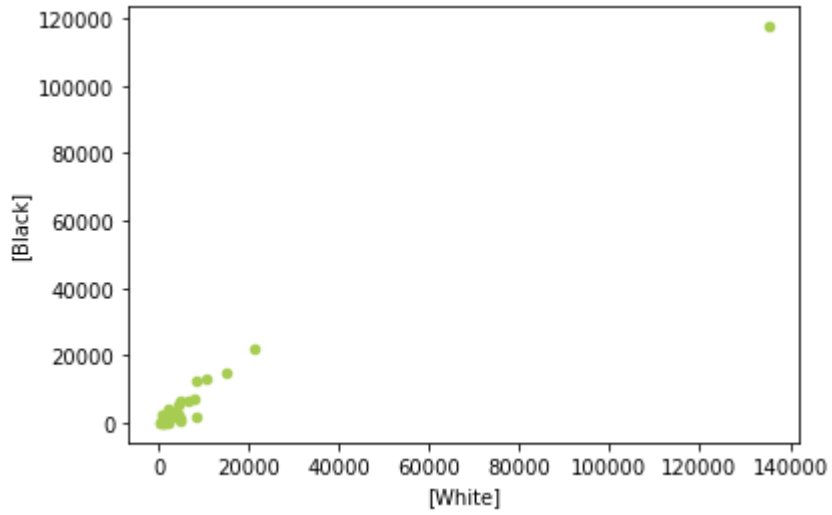
Bar-chart showing abortions among 20–24 age group



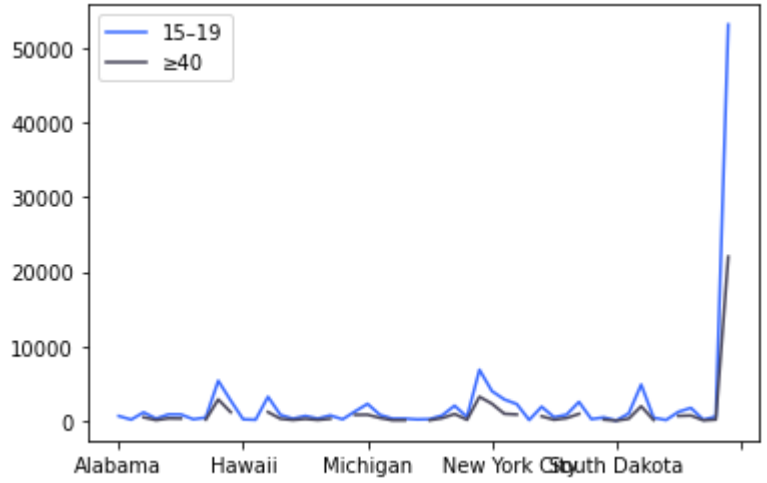
Bar-chart showing comparison of abortions among 2 racial groups



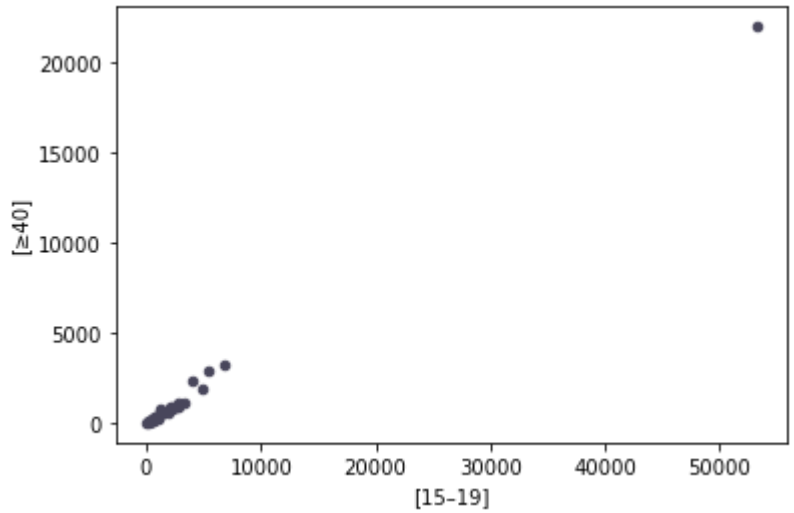
Scatter-plot showing correlation of 2 racial groups



Graph showing Abortions among 2 age groups



Scatter-plot showing correlation between 2 age groups



Conclusions and Analysis

The first part of the analysis provides statistical values (mean value, median value, standard deviation) within three chosen race/age groups. It can be noticed that the mean number of abortions is considerably lower among pre-teenagers and teenagers (approx. 38 cases) than the average number of 20-24 year olds who had an abortion (approx. 3960 cases). From that we can conclude that age is an important factor when observing abortion statistics and that one is less likely to have an abortion at a very early age.

Within the second part of the project there were obtained 3 graphs either presenting data among people of one specific racial/age group (graph 1 and 3) or depicting how a total number of abortions in one state is spread among all groups of one type (graph 2). It can be seen in the latter one (pie-chart 2) that there are two prevailing racial groups, white and black, in the state of Connecticut which people who had an abortion in 2018 belong to. Similarly, one can conclude that most young adults (20-24 y.o.) who had an abortion live in highly populated states (New York, Florida, Texas), as can be seen in bar-chart 3.

The third part of the analysis consists of comparison of values among two racial groups — black and white (graph 4 and 5); and two age groups — teenagers and people over 40 (graph 6 and 7). In both cases the first graphs show the data among two groups by state (bar/line-charts). These graphs allow us to see what states have most abortion cases among both groups simultaneously, as in graph 4, which shows that while in most states the number of cases among black and white racial groups is almost the same, in some state (like Michigan or North Carolina) people belonging to the black race had significantly more abortions than representatives of the white race. The second graphs (scatter-plots) in both comparisons show correlation of abortions among two groups. For instance, one can notice that large figures are spread relatively evenly between two groups in both cases and get sparser as the figures grow. The latter conclusion is much more visible through a scatter-plot rather than a bar-chart.