CS 512

Homework Environment Configuration

Name : Rahul Kumar Nalubandhu

Date : 01-17-2023

Title : Analyzing Birth Rates in the United States

Problem statement : Birth rates in the US

Big Data Referenced from : <https://wonder.cdc.gov/wonder/help/Natality-expanded.html>

OSEMN Process

1. Obtain : Data on birth rates in the United States was obtained from Center for Disease Control (CDC).
2. Scrub : The data capture a range of maternal demographic information, but we shortened to get data for number of births by each state in US also removed null values.
3. Explore : After Scrubbing the data it is ready to use for data visualization. So here the data was analyzed to find the trends in birth rates. Here we utilized Bar chart to visualize the data.
4. Model : [modelling is nothing but reducing the dimensionality of your data set so initially this data capture a range of maternal demographic information, such as state and county of residence, mother's age and race, ethnicity and country of origin, marital status](https://wonder.cdc.gov/wonder/help/Natality-expanded.html) So here we modeled our data large data to just get the birth rates in the USA and using this data we can model it to predict the future rates.
5. Interpret : The results were drawn about the current birth rates in united states and utilizing this data for future developments.

Text

Description automatically generated2)

Chart, histogram

Description automatically generated3)