

REPORT-ASSIGNMENT 1

In order to begin with a template generation I firstly went through a thorough analysis of the dataset and tried identifying those columns in which formatting would be required in order to be easily incorporated in news article. The date column in the dataset is changed from mm/dd/yyyy format to day, month, date, and year format. Initially the data set had abbreviation of states. I changed those abbreviations to their proper state name so that it gives more meaning to the article.

The news article which this code generates talks about a row specific data and also tries to derive insights specific to that row. The article starts with information related to a specific shooting incident and details into demographic information of the victim. Also with this code I have tried to develop insights related to shooting incidents in the given state in form of percentage of male and female victims among the shooting incidents which had happened in the given state, average age of victims in the state and number of shooting incidents in the given state. Police shootings have always been related to racial discrimination; hence the article generated by the code mentions the racial percentage distribution of police shootings in a given state. In order to understand how many such shootings took place every day, I included information in the article about other shooting incidents which had happened on the very same day but in other location. This article also attempts at providing an overview of all shooting incidents which happened in US since 1st Jan 2015. While developing the content and language of this article, I have tried to include maximum degree of freedom in terms of generalization so that inspite of introducing variation in the language of data set through usage of synset of jinja template; the information which it tries to impart remains uncompromised.

This form of automated news production can be really useful for producing news consisting to repetitive information. News articles related to information which is present in form of numbers can be easily generated using such templates. This type of template based news will not only be beneficial for journalists wherein a lot of their time will be saved but also can easily generate numerical insights from the data. Eg: news articles from a company's financial reports, weather report or tax related news articles. However such templates may not prove to be too useful in situations where in journalists use their own cognition in gathering, processing and understanding the information thereby penning down their own insights for e.g.: news articles mentioning implication of new governmental policies or articles mentioning impact of prolonged famine conditions etc. Automated news articles creation is useful in terms of maximizing the outreach of breaking news wherein creating hundreds of news articles in short time span is crucial. Following are the outputs.

OUTPUT-1

On Sunday, January 4, 2015 a 32.0 years old Male named Matthew Hoffman was fatally shot .The shooting was reported in the city of San Francisco located at California. The victim's race was identified as White. According to local police reports Matthew Hoffman possessed a toy weapon. The victim was reported to be not fleeing .There was a certain level of threat associated with the victim .The average age of these shooting victims in California is around 34.8 years. Since 1st January 2015 a total of 327 such incidents have taken place in California. The percentage of male victims among these 327 shootings is 94.8 percent in California. The percentage of victims who are whites is 29.97 percentage in California.3 such incidents have taken place on this date all over USA. Matthew Hoffman was shot in California and Michael Rodriguez was shot in Colorado and Kenneth Joe Brown was shot in Oklahoma. Till now the city of San Francisco has witnessed 9 such acts of violence .States like California, Texas and Florida have very high occurrence of such police shootings. The shooting statistics in California puts this state among very high range of shooting occurrences. Among all the 1954 police shootings in which a police officer, in the line of duty, shot and killed a civilian in US 89.2 of the police officers were not wearing a body camera. The absence of body camera makes it very difficult to ascertain the reason for the shooting and hence doesn't lend proof to the entire situation and adds it as random acts of violence.

OUTPUT 2

On Sunday, February 8, 2015 a 33.0 years old Male named John Martin Whittaker was fatally shot .The shooting took place in the city of Anchorage located at Alaska. The victim's race was identified as White. According to local police reports John Martin Whittaker was armed with a gun. The victim was reported to be fleeing .There was a certain level of threat associated with the victim .The average age of these shooting victims in Alaska is around 34.6 years. Since 1st January 2015 a total of 11 such incidents have taken place in Alaska. The percentage of male victims among these 11 shootings is 90.9 percent in Alaska. The percentage of victims who are whites is 45.45 percentage in Alaska.5 such incidents have taken place on this date all over USA. John Martin Whittaker was shot in Alaska and Sawyer Flache was shot in Texas and Vincent Cordaro was shot in New York and Joseph Paffen was shot in Florida and Larry Hostetter was shot in Texas. The city of Anchorage is tormented by 3 such acts of violence .States like California, Texas and Florida have very high occurrence of such police shootings .The number of shootings in Alaska puts this state among low range of shooting occurrences. Among all the 1954 police shootings in which a police officer, in the line of duty, shot and killed a civilian in US 89.2 of the police officers were not wearing a body camera. .The absence of body camera makes it very difficult to ascertain the reason for the shooting and hence doesn't lend credibility to the entire situation and adds it as random police shootings.