Pokemon with stats

1. Barradas, A. (2016, August 29). Pokemon with stats. [Data file] Retrieved August 29, 2016, from <https://www.kaggle.com/abcsds/pokemon>
2. License: under CC0:Public Domain License <https://creativecommons.org/publicdomain/zero/1.0/>
3. Why these data are interesting:

Nowadays, phone applications have become an indispensable part of people’s daily life. People tend to take advantage of them to lead a pleasant life. For example, with the help of digital map, people nowadays find it easier to reach a place than ever. They can also enjoy life better and entertain themselves with the help of applications. Pokemon is definitely one of them. At the beginning of the release of Pokemon, it did cause a heat. Everyone went everywhere to get a pokemon. Players are indulged in the fun provided by Pokemon. They are likely to collect all pokemons and lead a wonderful battle with other players.

This data set includes 721 Pokemon, including their number, name, first and second type, and basic stats: HP, Attack, Defense, Special Attack, Special Defense, and Speed. In this raw data, people can understand how much damage an attack will do in the games.

In addition, it has been of great use when teaching statistics to kids. With certain types people can also give a geeky introduction to machine learning.

1. Potential users / decision makers: Pokemon players; parents who need to teach some easy statistics to their kids; machine learners
2. Three questions can be answered by the dataset
3. The defenses and attack cannot well describe certain type of pokemon. From the data set, people can find another two variables to define the type of a pokemon. For example, the special attack and special defense; the total mark of the pokemon and the generation.
4. The connection between the speed and type of pokemon

Pokemons with the type of grass or related types tend to have a rather slower speed than others.

1. People always misunderstand that the upper generation tends to have a better defense and attack. After reading the data set, it is not hard to find out the improvement of defenses and attacks cannot simply rely on the generations.

College Scorecard Data

1. U.S. Department of Education (2016, September 13). College Scorecard Data. [Data file] Retrieved September 13, 2016, from https://collegescorecard.ed.gov/data/
2. License: under CC0: Public Domain License <https://creativecommons.org/publicdomain/zero/1.0/>
3. Why these data are interesting:

It is known that students of prestigious school can always be regarded as someone intelligent and get higher salary after graduation. However, students who attend elite colleges tend to be bond to the loan for their studies. The debt repayment obligations sometimes will far outstrip their employment and income prospects. If they had the chance to see the predicted outcome compared with what they should pay, they are more likely to make a right decision. Even more, they are able to plan for their future and get a higher reward. Therefore, I believe this data set is useful and interesting.

1. Potential users / decision makers: students who wants to make educational investments less speculative; parents who care about their children; educators
2. Three questions can be answered by the dataset
3. We can figure out the schools of which part of the U.S.A. have a higher tuition fee than the ones of other religions. E.g. the average tuition fee of NY is higher than 10,000 dollars, and some of them even have more than 40,000 dollars.
4. Which part of the U.S.A have the largest amount of schools. It may have some connections between the local economy.
5. The University of Maryland College Park doesn’t have a very expensive tuition fee. However, the differences between out of state tuition with state tuition is really big.

Emergency-911 calls

1. Chirico, M. (2016, March 02). Emergency (911) Calls: Fire, Traffic, EMS for Montgomery County, PA. [Data file] Retrieved March 02, 2016, from http://montcoalert.org/gettingdata/
2. License: Open Database License(ODbL)v1.0

<http://opendatacommons.org/licenses/odbl/1.0/>

1. Why these data are interesting

By collecting and analyzing these data, the police of Montgomery County can have a better understanding of the most emergency place of the whole county and always keep an eye on this region. In addition, they can make a list of the emergency feature of different area of the county. For example, the officers can summarize the place that most of the traffic accidents happen and then set some signals on that road to warn drivers to pay more attention. It can help prevent the bad accident from happening and save people’s lives. Another approach to better take advantage of these data is to analyze the frequency of emergency happens in different areas and restore the analysis. When local police station needs to expand or have other branches around the county, officers can first consider these highly-emerged emergency place.

From the aspect of inhabitants, they can be aware of those places that criminals often choose to conduct crime. They can spread the awareness to their friends, relatives and neighborhood.

1. Potential users/ decision makers: the governors; the officer of Emergency Center; the local people; the Traffic Department of Montgomery County
2. Three questions can be answered by the dataset
3. Which area of the country is the most dangerous place and residents should pay attention.
4. Which places does fire tend to happen.
5. Which religion of the county is most elite. Because the disadvantaged places are less likely to have a low rate of crime.

Word count:899