Yuheng Zhang (yzhang63@terpmail.umd.edu)

- How much should Stephen Curry to earn in the next contract?

Datasets:

1. lmcmath34. (2016).Stephen Curry 2015-2016 Regular Season [Data file]. Retrieved from https://www.statcrunch.com/app/index.php?dataid=1884311
2. gcbateman22. (2016).stephcurry [Data file]. Retrieved from https://www.statcrunch.com/app/index.php?dataid=1812768
3. lmcmath34. (2016).2015-2016 NBA Salaries [Data file]. Retrieved from https://www.statcrunch.com/app/index.php?dataid=1843341
4. lmcmath34. (2016).Kevin Durant 2015-2016 Regular Season [Data file]. Retrieved from https://www.statcrunch.com/app/index.php?dataid=1884280

Term of Use:

The owner of the website Pearson (2016) indicates that: “Permitted Uses of Materials. You may, on an occasional and irregular basis, include insubstantial portions of Materials from the Website in memoranda, reports and presentations, and then only to the extent that such use is for educational purposes and of a non-commercial nature, does not otherwise diminish the pedagogical or commercial value of the Materials, and is otherwise permissible as "fair use", “fair dealing” or its equivalent under applicable copyright and intellectual property law. In each such instance, You must include proximate to any excerpted Materials all copyright, trademark and other notices from the Website (whether or not included in the particular Materials used), and appropriate source attribution to Pearson and its licensors.” (Pearson Education End User License Agreement and Privacy Policy, 3.2)

Reference:

Pearson. (2016). Pearson Education End User License Agreement and Privacy Policy. Retrieved from https://register.pearsoncmg.com/reg/include/license2.jsp

Stephen Curry, one of the best basketball player in the world, he helped Golden State Warrior to win the NBA champion in 2015. According to comparison between data set #1 and #4, Stephen got 30 scores on average at each match of the 82 matches in 2015-2016 season. He surpassed Kevin Durant,

his new teammate who got 28 scores on average at each match. Moreover, as shown in data set #2, Stephen has made great progress in shooting average of both two-point and three-point since 2009, which indicates that he has indefinite potential in the following few years. However, as shown in dataset #3, this superstar's salary was only $11.3m in 2015, ranking the 57th among all basketball players in NBA! Kevin Durant, one of his teammates, earned $9m more than him.

Data users:

- Boss and Team Manager of Golden State Warriors
- Managers of other teams in NBA
- Stephen Curry
- Stephen's agent

a. How much money will the team pay to Stephen in the next new contract?

Stephen's contract will expire in 2017. Analysis on the aforesaid datasets may help Boss and Team Manager of Golden State Warriors learn how much they will offer in the next contract if they want to sign a contract with this superstar again. They may learn and evaluate the value of Stephen by carrying out a horizontal comparison between the performance of Stephen with that of all players at Point Guard position in NBA, including the data of shooting and assistant, the data of the PG player who gets the most salary or the average scores of all basketball players in the previous season. Meanwhile, they can also carry out a vertical comparison on the salaries of all superstars of NBA 30 teams in order to estimate a number which can satisfy Stephen.

b. How will Stephen perform better?

Maybe salary is not the problem that Stephen cares most. Nevertheless, by comparing his data with the data of other top NBA players, such as his new teammate Kevin Durant, if Stephen wants to be a legend, he has to be able to learn what he needs to improve, so that he can help his team to win in the final champion again.

c. What salary is acceptable and satisfactory to my client?

Similar to the team's manager, Stephen's agent also needs to learn and know the value of his

client through data analysis. Maybe the conditions provided by the team are not acceptable to him. He needs to estimate a reasonable number, and he is responsible for creating the maximum interest and value of his client.

- Gruyter Open Author Survey

Dataset:

1. Gruyter. (2016).de_gruyter_open_author_survey_2016.csv [Data file]. Retrieved from https://s3-eu-west-1.amazonaws.com/pfigshare-u-files/5360981/de_gruyter_open_author_survey_2016.zip

License:

Type of License: CC-0

As the organization names Creative Commons (2016) defines that "CC0 1.0: No copyright. The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission."

Reference:

Creative Commons. (2016). CC0 1.0 Universal. Retrieved from https://creativecommons.org/publicdomain/zero/1.0/

The questionnaire is designed for surveying academic authors' attitudes towards the question whether they are willing to pay APC (Article Processing Charge) for open access publishing. Firstly, the questionnaire is interesting, it is an anonymous survey. The survey is not designed for a specific group, such as professors in universities. Hence, anyone, even those who have never published any academic paper, can give answers despite their age, occupation or educational background. Secondly, the survey scope is not limited to a country or region. Instead, it covers the whole world. We can find out, from the results of users from different cultural backgrounds, the users attitude towards the question whether they are willing to pay APC. Gruyter released 107,296 questionnaires and got 1012 replies. The results of all users of Open author survey are listed in the dataset.

Data users:

- Readers who find academic papers
- Authors of academic papers
- Managers of open access publishing services
- Publishers

a. People of what background are willing to pay APC and choose Open access publishing?

For all readers, the key of an academic paper is quality. Readers can get the answers they need from some good academic papers instead of wasting time in reading some low-grade papers. This dataset can help a reader learn people of what educational background and at what age are willing to pay APC for open access publishing. If most of the authors are very young, have poor education background, or engage in an industry which is very strange to the reader, the reader may search and obtain academic paper in other ways.

b. Authors from which countries and regions are willing to pay APC?

Through dataset analysis, open accessing publishers can know people of which industry, which background and from which country are willing to pay APC. To make the project operate better, they may know which industries they should turn to in order to seek opportunities, or which regions they should focus on to dig the authors' commercial potential.

c. Will the authors choose to pay APC?

Through dataset analysis, authors can learn how the authors of the same industry choose to pay APC or not. In addition to view of their papers, they may need to take into account their own benefits. How should an author make a choice if he wants to publish several academic papers in the next three years? He may check the dataset to see the choice of authors who have published more than two papers in order to determine whether he will choose open access publishing and pay APC.

- Predicting excitement at Donorschoose.org

Datasets:

1. DonorsChoose.org. (2014).project.csv [Data file]. Retrieved from https://www.kaggle.com/c/kdd-cup-2014-predicting-excitement-at-donors-choose/download/projects.csv.zip
2. DonorsChoose.org. (2014).donations.csv [Data file]. Retrieved fromhttps://www.kaggle.com/c/kdd-cup-2014-predicting-excitement-at-donors-choose/download/donations.csv.zip
3. DonorsChoose.org. (2014).outcomes.csv [Data file]. Retrieved from https://www.kaggle.com/c/kdd-cup-2014-predicting-excitement-at-donors-choose/download/outcomes.csv.zip

License:

Type of License: CC-0

As the organization names Creative Commons (2016) defines that "CC0 1.0: No copyright. The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission."

Reference:

Creative Commons. (2016). CC0 1.0 Universal. Retrieved from https://creativecommons.org/publicdomain/zero/1.0/

DonorsChoose is a website on which school teachers can initiate projects and raise fund, and they will obtain the materials they need for an activity once the amount of the fund has been raised. However, not every project can succeed. Dataset #1 provides the data of all projects, and Dataset #2 provides the data of all donors, including where they are from, the amount of their donation and their donation ways; of course the information is anonymous. Dataset #3 provides the results of all projects,

including the results of fund raising, the recorded data of related donors, as well as the reasons why some projects are successful while some are ignored. By analyzing these datasets, we will be able to find out some similarities of very successful projects.

Data users:

- Project initiators or teachers in school
- Project donors
- Website administrators

a. How to make a project a success?

Project initiators can find out the key to success of a project by analyzing some projects which are marked with "exciting". For instance, whether the reason of the success of some projects is recommendation by more teachers; what kind of resource is more inclined to being donated, electronic projects or other materials? People from which places are more willing to aid projects financially, do I need to take account into them in particularly, for instance, seeking cooperation with local schools etc.

b. How to make the money more valuable?

Donors need to ensure that their donations are meaningful in addition to finding out the projects that they are interested in. By analyzing these datasets, they can find out the recognized projects which really need help. For instance, by analyzing poverty level, they can donate to the schools which need help more urgently. Meanwhile, if a project initiator or a school initiates several donations in a period, the donor should consider carefully whether the school really needs money or the initiator has other attempts.

c. How to perfect DonorsChoose?

Website administrators can know whether their website is more and more popular by analyzing the number of newly increased projects every day. If the number is decreasing, then they need to know what goes wrong. By analyzing the initiating places of the projects, they can know the regions where the teachers are not so good at using DonorsChoose. It doesn't mean that they don't need to

raise fund, maybe they just don't know this website. Website administrators can make better plans to facilitate promotion of the website to the potential users.

Word count: 1624