**Individual Peer Evaluation Form**

Your name: **Supraja Rapuru**

Write the name of your classmate you are preparing this review for in the designated column. Using a scale of 1-4 (1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree) answer each question. If you aren’t able to answer the question based on what is posted in the discussion board, reach out to your classmate for more information via the discussion board. Total the numbers in each column. **Make sure to answer the questions on the 2nd page.**

|  |  |
| --- | --- |
| Evaluation Criteria | Peer Name:  **Victoria Hall** |
| Has plan in place to complete course project. | 4 |
| Has found datasets/data sources to support project idea. | 4 |
| Has solidified project idea. | 4 |
| Has identified resources for project. | 4 |
| Topic is related to data science and demonstrates topics learned to date through program. | 4 |
| Risks and potential issues have been identified. | 4 |
| TOTALS | 24 |

Feedback on Individual’s project topic:

1. How clear is the classmate’s project topic? What questions does their topic make you consider?

* The topic Victoria is chosen is very interesting and relevant project now. There always been constant focus on customer retain strategies in marketing departments. It has same importance of customer acquisitions. Her project is clear however would help if the context is provided. I couldn’t get clear understanding if the proposal is to predict customers attrition for a product or a bank(assuming it is bank from the data file).

I have few questions on data and request Victoria to consider these while processing the data.

* + - The data appears to be a snapshot as some point in time e.g. the balance is for a given date which leaves a lot of questions: What date is it and of what relevance is this date, Would it be possible to obtain balances over a period of time as opposed to a single date.
    - There are customers who have exited but still have a balance in their account! What would this mean? Could they have exited from a product and not the bank?

1. What risks or issues should your classmate consider while working on their project?
   1. I would say being that the data is so specific to the company they are using it may be hard to incorporate the code to other data sets from other companies.
   2. The second data set might be hard to incorporate with the initial data set from the employee attrition dataset that was simulated by IBM
   3. There could be lot of factors that would drive customer churn but the dataset here is quite basic so may have to keep that in mind while deriving the conclusions
2. Additional suggestions/comments that might be beneficial to your peer?
   1. The milestone2 you have created is crisp and clear. The topic is interesting and your project plan is great
   2. Usually such simulated datasets would be imbalanced. So I would suggest you to consider over/under sampling methods to increase the effectiveness of dataset (my initial observation, you have more observations with churn 0 than 1)

Adapted from a peer evaluation form developed at Johns Hopkins University (October, 2006)