- Primary Key is UserID since, it is available for all players, including Level 1 players who have not joined a team.
- Selected data type “long integer” for session id since, there could be several session ids and the userID is also a long integer.
- Selected data type “integer” for missionID since, wanted to optimize the storage needs and do not believe there is a need for too many missionIDs.
- Selected data type “binary” for isHit since, it can only take a value of “yes” or “no”
- Added a column for TeamID with datatype “long integer” to accomplish the objectives for this game. Combined with the other 2 tables it should provide the ability to run required queries.
- Intentionally, did not add team rank and score to this table to save space and since, scores and rankings always changing. The Egelence website will provide a link “See Team Scores and Ranking.” A new user will click on this link which will force a query to be run on the below data structure to calculate Team Scores and Rankings. On completion, of the query the score and ranking will be displayed.

userID: long	sessionID: long	timestamp: dateTime	clickedPoint: coordinate	missionID: int	isHit: binary	TeamID: long
100	4356	10/12/2015::14:15:09	(4,8)	13	yes	10000
101	3241	10/23/2015::14:15:19	(20,5)	18	no	10100
102	4537	11/4/2015::14:15:20	(17,43)	21	no	10100

Vertex Table:

Id	Property
User	(User, UserID)
Chat Session	(Chat, Chat Session ID)
Chat Text	(Chat Text, Text)

Edge Table”

SrcID	DstId	Actions	Team ID	TimeStamp
User	Chat Session	Joins		
User	Chat Session	Starts		
User	Chat Session	Leaves		
User	Chat Text	Writes		
Chat Text	User	Mentions		
Chat Session	Chat Text	Contains		

1. Which teams are having more conversations?

- First add Team ID to the data structure. The team that has the most User ID in the Chat Session and Chat Text node will have the most conversations.

1. Do users chat more (or less) before they leave a team?

- First the user must Join or Start a Chat Session. From the above data structure we will determine the timestamp when a user leaves the Chat Session. We will then select a time window before the timestamp when user leaves, and collect how many “Write” Actions

were performed during this time window. Based on such data for several users we can determine if a user chats more or less before leaving the chat session.

2. What are the dominant terms (words) used in a chat session within a specific time period?

- This can be easily done by selecting the time period of interest and creating a “Term Document Matrix (TDM)” for all the chat session. The TDM can provide the dominant terms.

3. Which users are most active in a specific chat session?

- Again this can be answered by running a query and determining which user is performing the write action and gets “mentions” to a specific chat session.

4. How many chat sessions is a user participating in at the same time?

- A Union query on UserID and Chat SessionIDs should answer this question.

Explain how you might extend the tree with at least five specific flamingo properties. For example, beak-color whose values might be “bright-red” or “pink”.

1. Flamingos with beak colors “bright-red” and have number 100 on their chest.
2. Flamingos with beak colors “bright-red” and have number 200 on their chest.
3. Flamingos with beak colors “bright-red” and have number 300 on their chest.
4. Flamingos with beak colors “pink” and have number 100 on their chest.
5. Flamingos with beak colors “pink” and have number 200 on their chest.