Project Report

Jianru Wang/Jing Wang

Part 1

1) Domain: Our project is about baseball.

2) Conceptual models (entities) for a tweet/post, a Social Media user, a person, and a company:
The tweet entities is divided into three parts (consumers, producers and companies). The entities of social media user will contain information like userid, username, tweets, likes and favorites.

3) Conceptual models (entities) that represent consumers, producers and companies in your chosen domain.

Consumers: baseball fans

Producers: baseball players

Companies: baseball teams

4) Conceptual models (entities) for at least two things specific to the domain.

Teams like Boston Red Sox and Los Angeles Dodgers are specific to the domain.

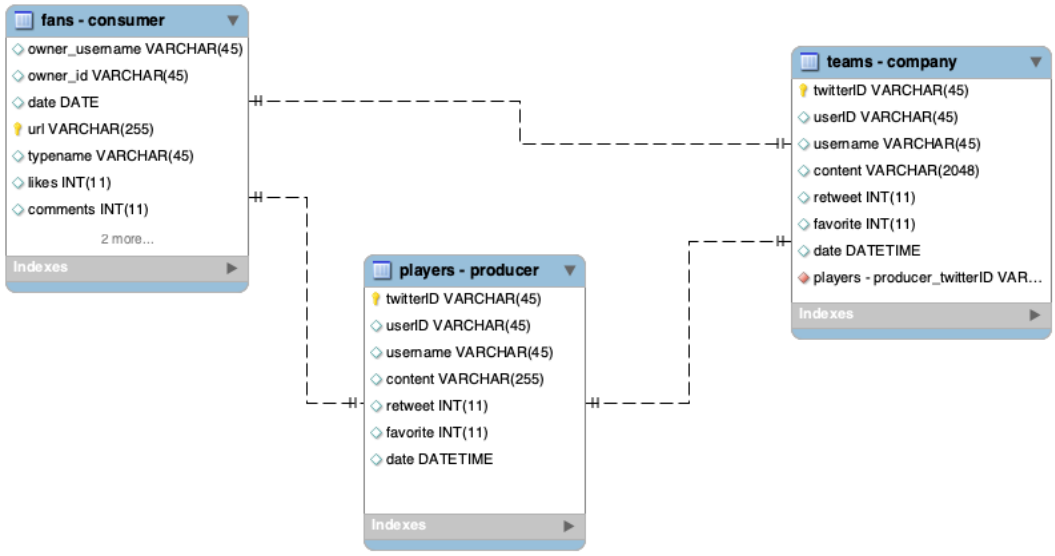
5) Relationships that connect the entities.

Shown in ER diagrams.

6) Appropriate attributes and keys

Shown in ER diagrams.

7) ER diagrams that illustrate the entire conceptual model.



Questions:

1.What are the ranges, data types and format of all of the attributes in your entities?

The attributes of FAN are username, userid, postdate, url, comments and likes. The attributes of PLAYER are userid, username, content, retweet, favorite, date. The attributes of TEAM are twittered, userid, content, retweet, favorite, date and location. They are strings and numerical values.

2.When should you use an entity versus attribute? (Example: address of a person could be modeled as either)

Entities have many attributes. For example, I collected baseball fans' information from Instagram. The fans is entity, and their username, userid, postdate, url, comments and likes are attributes. The entity could be used to connected to the other entities.

3. When should you use an entity or relationship, and placement of attributes? (Example: a manager could be modeled as either)

When the attributes have relationship within them, we should do that.

4. How did you choose your keys? Which are unique?

Primary key is unique. The primary key of FAN is ownerid; the primary key of TEAM is twitterID；the primary key of PLAYER is twitterID.

5. Did you model hierarchies using the "ISA" design element? Why or why not?

No. Since my model in this assignment is a bit simple and actually the relationships among those teams and fans are not so strong. So it's unnecessary to use ISA design element.

6. Were there design alternatives? What are their tradeoffs: entity vs. attribute, entity vs. relationship, binary vs. ternary relationships?

The alternative design is to use user id instead of twitter id as primary key for players and teams. In this way, we need create more charts to build the relationship between them and will consume more space.

7. Where are you going to find real-world data to populate your model?

We collected fans' date from Instagram, players' and teams' data from twitter.

Part 2

1) Updated ER diagrams that illustrate the entire conceptual model in a form that directly maps to SQL. You need to either: incorporate the feedback you are given or respond as to why it isn't needed.

We didn't change the ER diagrams since use the current model and diagrams we can solve the use cases and questions. We think the current ER diagrams are prefect.

2) SQL and diagram for the physical model that represents the entire conceptual/physical model.

The physical model are shown in the ER digrams.

3) SQL that express the queries you are asked to write.

a. What user posted this (e.g. tweet, facebook post, IG post, etc.)?

select username, content from assign2.players

	username	content
1	Christian Yelich	RT @travis_shaw21: Oh yeah... one more thing... retweet and comment a caption and me and @ChristianYelich will pick the best comment and I…
2	Mike Trout	@AugieAguilar 👀
3	Bryce Harper	The Harper 3 Low!🔥 #MemorialDayWeekend https://t.co/m8D7ITI3Rl
4	Mike Trout	Detroit🚀 🚀 🚀 🚀 🚀 !!!
5	Mike Trout	Lebron🔥🔥🔥🔥#wow
6	Christian Yelich	@travis_shaw21 🐻
7	Christian Yelich	RT @travis_shaw21: “On this day, take time to remember those who have fallen. But on every day after, do more; put the freedoms they died f…
8	Mike Trout	RT @LandisOlivia: MY ALBUM HAS FINALLY BEEN RELEASED!! CHECK IT OUT ON ITUNES OR SPOTIFY🎶♡ https://t.co/WP8OJ4sTOK
9	andrew mccutchen	I cant imagine what those who have served/serve have gone thru. What they’ re families go thru.The things they have… https://t.co/y3QbZIMQ60
10	Mike Trout	#MEMORIALDAY us🇺🇸Today we honor those who made the ultimate sacrifice...Never Forgotten & Always Thankful.
11	andrew mccutchen	Here’ s your kid’ s chance to show off their skills! Don’ t miss your opportunity to win a trip to the 2018 All-Star G… https://t.co/31M2nCIjQA
12	Christian Yelich	RT @MLBNetwork: Before today’s win against the Cardinals, @ChristianYelich joined #HighHeat to discuss the @Brewers’ surge to the top of th…
13	Dansby Swanson	RT @NayPlayPod: Very excited to share a new episode of @NayPlayPod. 🗣️Atlanta @Braves shortstop Dansby Swanson (@LieutenantDans7) came by t…
14	Dansby Swanson	RT @SannyDamet: Very excited to share a new episode of @NayPlayPod. 🗣️Atlanta @Braves shortstop Dansby Swanson (@LieutenantDans7) came by t…
15	Dansby Swanson	RT @sportingnews: Cardinals TE Jermaine Gresham helping a person in need ♡ https://t.co/uscygvotGB

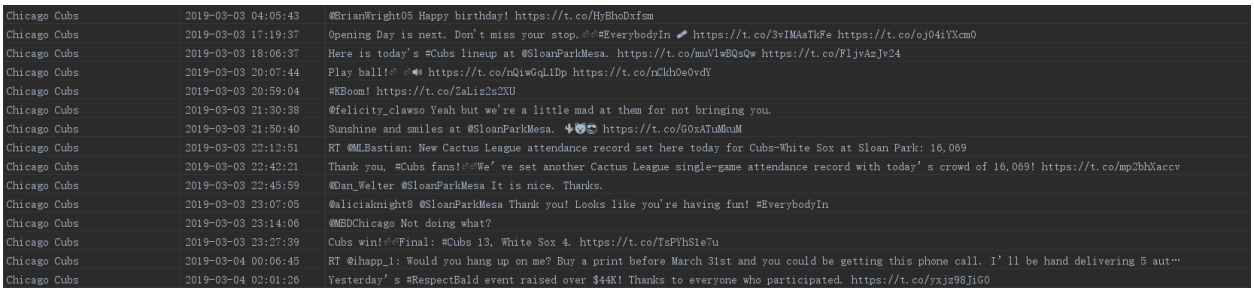
b. When did the user post this (e.g. tweet, facebook post, IG post, etc.)?

select username, date, content from assign2.players

	username	date	content
1	Christian Yelich	2018-05-25 15:26:13	RT @travis_shaw21: Oh yeah... one more thing... retweet and comment a caption and me and @ChristianYelich will pick the best comment and I…
2	Mike Trout	2018-05-26 04:22:52	@AugieAguilar 👀
3	Bryce Harper	2018-05-26 15:05:06	The Harper 3 Low!🔥 #MemorialDayWeekend https://t.co/m8D7ITI3Rl
4	Mike Trout	2018-05-27 22:23:04	Detroit🚀 🚀 🚀 🚀 🚀 !!!
5	Mike Trout	2018-05-28 02:57:27	Lebron🔥🔥🔥🔥#wow
6	Christian Yelich	2018-05-28 03:56:11	@travis_shaw21 🐻
7	Christian Yelich	2018-05-28 15:02:16	RT @travis_shaw21: “On this day, take time to remember those who have fallen. But on every day after, do more; put the freedoms they died f…
8	Mike Trout	2018-05-28 16:03:31	RT @LandisOlivia: MY ALBUM HAS FINALLY BEEN RELEASED!! CHECK IT OUT ON ITUNES OR SPOTIFY🎶♡ https://t.co/WP8OJ4sTOK
9	andrew mccutchen	2018-05-28 18:39:34	I cant imagine what those who have served/serve have gone thru. What they’ re families go thru.The things they have… https://t.co/y3QbZIMQ60
10	Mike Trout	2018-05-28 21:09:15	#MEMORIALDAY us🇺🇸Today we honor those who made the ultimate sacrifice...Never Forgotten & Always Thankful.
11	andrew mccutchen	2018-05-29 17:00:33	Here’ s your kid’ s chance to show off their skills! Don’ t miss your opportunity to win a trip to the 2018 All-Star G… https://t.co/31M2nCIjQA
12	Christian Yelich	2018-05-30 22:46:19	RT @MLBNetwork: Before today’s win against the Cardinals, @ChristianYelich joined #HighHeat to discuss the @Brewers’ surge to the top of th…
13	Dansby Swanson	2018-05-31 15:23:44	RT @NayPlayPod: Very excited to share a new episode of @NayPlayPod. 🗣️Atlanta @Braves shortstop Dansby Swanson (@LieutenantDans7) came by t…
14	Dansby Swanson	2018-05-31 15:23:52	RT @SannyDamet: Very excited to share a new episode of @NayPlayPod. 🗣️Atlanta @Braves shortstop Dansby Swanson (@LieutenantDans7) came by t…
15	Dansby Swanson	2018-05-31 18:00:42	RT @sportingnews: Cardinals TE Jermaine Gresham helping a person in need ♡ https://t.co/uscygvotGB

c. What posts has this user posted in the past 24 hours?

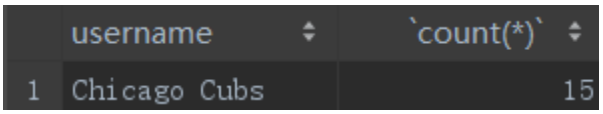
select username, date, content from assign2.teams where username like '%Chicago Cubs%' and date > '2019-03-03'



Chicago Cubs	2019-03-03 04:05:43	@BrianWright05 Happy birthday! https://t.co/HyBhoDxfsm
Chicago Cubs	2019-03-03 17:19:37	Opening Day is next. Don't miss your stop.⚾⚾#EverybodyIn 🔪 https://t.co/3vIMAaTkFe https://t.co/oj04iYXcm0
Chicago Cubs	2019-03-03 18:06:37	Here is today's #Cubs lineup at @SloanParkMesa. https://t.co/muVlwBQsQw https://t.co/FljvAzJv24
Chicago Cubs	2019-03-03 20:07:44	Play ball!⚾ ⚾🔊 https://t.co/nQiwGqL1Dp https://t.co/nCkh0e0vdY
Chicago Cubs	2019-03-03 20:59:04	#KBoom! https://t.co/ZaLiz2s2XU
Chicago Cubs	2019-03-03 21:30:38	@felicity_clawso Yeah but we're a little mad at them for not bringing you.
Chicago Cubs	2019-03-03 21:50:40	Sunshine and smiles at @SloanParkMesa. ☀️🐻😎 https://t.co/G0xATuMkuM
Chicago Cubs	2019-03-03 22:12:51	RT @MLBastian: New Cactus League attendance record set here today for Cubs-White Sox at Sloan Park: 16,069
Chicago Cubs	2019-03-03 22:42:21	Thank you, #Cubs fans!⚾We've set another Cactus League single-game attendance record with today's crowd of 16,069! https://t.co/mp2bhXaccv
Chicago Cubs	2019-03-03 22:45:59	@Dan_Welter @SloanParkMesa It is nice. Thanks.
Chicago Cubs	2019-03-03 23:07:05	@aliciaknight8 @SloanParkMesa Thank you! Looks like you're having fun! #EverybodyIn
Chicago Cubs	2019-03-03 23:14:06	@MBDChicago Not doing what?
Chicago Cubs	2019-03-03 23:27:39	Cubs win!⚾Final: #Cubs 13, White Sox 4. https://t.co/TsPYhS1e7u
Chicago Cubs	2019-03-04 00:06:45	RT @ihapp_1: Would you hang up on me? Buy a print before March 31st and you could be getting this phone call. I'll be hand delivering 5 aut…
Chicago Cubs	2019-03-04 02:01:26	Yesterday's #RespectBald event raised over $44K! Thanks to everyone who participated. https://t.co/yxjz98JiG0

d. How many post has this user posted in the past 24 hours?

select username, count(*)from assign2.teams where username like '%Chicago Cubs%' and date > '2019-03-03'



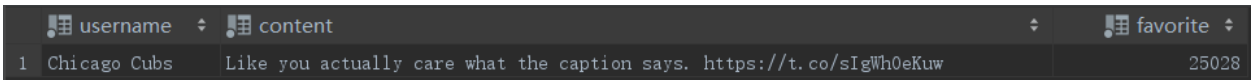
	username	`count(*)`
1	Chicago Cubs	15

e. What keywords/ hashtags are popular?

we have no hashtags in our tables

f. What posts are popular?

select username, content, favorite from assign2.teams where username like '%Chicago Cubs%' order by favorite desc limit 1

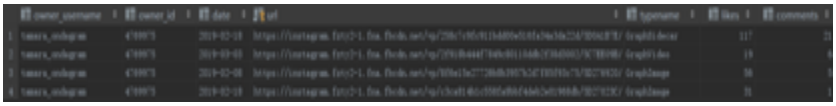


	username	content	favorite
1	Chicago Cubs	Like you actually care what the caption says. https://t.co/sIgWh0eKuw	25028

4) At least 5 (10 if two people) distinct use cases of queries that are particular to your domain. SQL expressions that express the 5 (10 if two people) use cases of queries that you write.

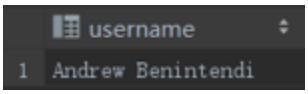
 # What posts has player users posted in 2018?

select * from assign2.fans where year(date) > 2018



 # What's the lastest post of player user?

select username from assign2.players order by date limit 1



	username
1	Andrew Benintendi

 # Which post get the most likes for fans user?

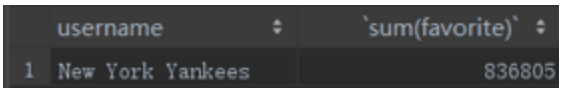
select owner_username, url from assign2.fans order by likes limit 1

 # What posts get more than 5 retweets of player user?

select * from assign2.players where retweet > 5

 # Which team user gets most favorites in total?

select username, sum(favorite) from assign2.teams group by username limit 1



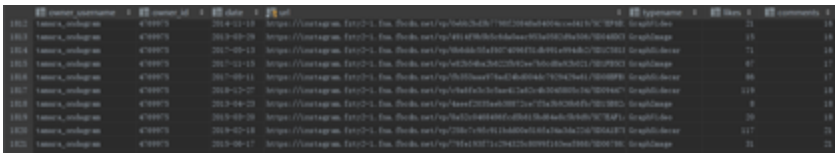
	username	`sum(favorite)`
1	New York Yankees	836805

 # What posts of team user get more than 10 retweets and less than 50 favorites?

select username, content from assign2.teams where retweet > 10 and favorite < 50

 # Rank the posts of fans user according to the number of comments from the least to the most.

select * from assign2.fans order by comments



 # Which team is interested in champion

select * from assign2.teams where content like '%champion%'

 # which player postes least posts this year?

select username, count(*) sum from assign2.players where year(date) > 2018 order by sum limit 1

	username	sum
1	andrew mccutchen	193

 # Rank the posts of fans user according to the number of likes from the most to the least.

select * from assign2.fans order by likes desc