Section III
Description of the Enterprise

The enterprise is a database of CrossFit Athletes, which can be accessed online. This enterprise maintains an up to date list of all CrossFit athletes throughout the season and their statistics. In order to understand the information and the reasoning behind this database, CrossFit must be described. A CrossFit competition will consist of several events or workouts. Similar to traditional sports there is a scoring system. Each event can be broken down into two types: timed and weighted. A timed workout is a race. The objective is to finish the given movements as quickly as possible and the time is your score. A weighted workout works the same as Olympic Weightlifting. One lift is given to the athletes. They each have a set amount of time to complete the movement with the highest weight possible. The weight then becomes their score.

The statistics can be broken down into several sections. The first consists of "Benchmark Workouts" and their "Maxes". The benchmark workouts consist of seven of the iconic CrossFit workouts. Each athlete will have their best to date score for each workout as their statistics. The "Maxes" has four Olympic Lifts and one gymnastics movement where each athlete will have their best score for each, similar to their Benchmark Workouts. The next section is a brief biography of the athlete, which is written by each Athlete on the site, under their profile. All of this information is contained within the main athlete section. The next section is the "Athlete Information" section. This consists of several questions pertaining to the sport and that athlete's training. These questions are also answered personally by each athlete through the website. The next section is "Athlete Coverage." This section consists of any pages on the website that feature this athlete. These can include news articles, photos, and videos.

The final section is the Leaderboard section. This is a very simple leaderboard that displays the top nine athletes in order. The rest of the leaderboards can also be accessed to see any athletes standing. Most of this information is updated by the athletes themselves. Because of this, each section has a different frequency of updating. The only sections that are not updated by the athletes themselves are the Athlete Coverage and the Leaderboard. The section with the most changes is the Leaderboard. However, this section is only updated during the season. The structure of the season regulates how often the leaderboards are updated. I will cover this outline next.

The CrossFit season is broken down into three stages: The Open, Regionals, and the Games.

The Open:

The open is the first stage. This is inclusive to anyone in the world who wants to compete. Any athlete who signs up must create an athlete profile. During this time, there is a massive amount of athletes in the database. (The 2014 Open had 209, 585 participants.) Because of this, the world is divided into 17 "regions". Athletes only compete against other athletes from their region. The Open takes place over five weeks, where one workout is released each week. Every week, every athlete must submit their scores for the workout, which are then stored to their profile and places them accordingly onto the leaderboard. Since the scores have the span of a week, the leaderboards are changing constantly throughout these five weeks. After the five weeks are over and every workout has been completed, the top 48 men and women from each of the 17 regions on the leaderboard are invited to the next phase.

Regionals:

Regionals takes place over four weeks. Each region only competes for a total of three days. The regions are separated over the four weeks. The four new regions compete every weekend and then five regions on the last. The leaderboard is only updated at the end of a day of competition. At the end of this, the top 3 men and women from each region are invited to compete in the final stage, the CrossFit Games.

The Games:

The CrossFit Games consists of six days of competition. Each day consists of multiple workouts. After each workout is over, the scores are updated to the database and the leaderboards are updated accordingly. After the end of the last workout on the sixth day of competition the scores are totaled and the winners are declared.

The leaderboard has several sections. It keeps track of the scores from each stage of the competition, so that the final leaderboard for the Open, Regionals, and the Games can be seen online. The leaderboard is also separated by gender, giving two versions of each. Another feature is being able to compare athletes based on their stats. Some example queries are listed below:

- Who qualified for the Games in 2013 and made it into the top three?
- Who has qualified for regionals every year but has never made it to the Games?
- Which athletes have been in the top three at the Games for the past three years?

- Which athletes have qualified for the Games this year for their first time?

- Which athletes have qualified for the Games this year who have never competed before?

- Which athletes have qualified for the Games this year who have competed every year?

- Which athlete has the best score for Open Workout 14.1?

- Which athletes have qualified for the Games every year except for this year?

- Which athletes compete in the Central East Region?

- Which athletes are affiliated with CrossFit Los Angeles?

Section IV
Definition of the Environment

Section IV.1 Input and Report Forms

IV.1.1
Potential Athlete
- Date of Application
- Name
- Email
- Date of Birth
- Gender
- Height
- Weight
- Region
- Affiliate

IV.1.2
Benchmark Form
- Athlete Name
- Benchmark name
- Score

IV.1.3

Competition Form
- Competition Name
- Region
- Score
- Date of Competition

Section IV.2 Assumptions

1. Potential athletes/competitors fill out form to be reviewed for approval.
2. Athletes get one account/profile
3. An athlete submits their score for an Open workout, which is posted to profile. If the scores place them in the top rankings, they will need to submit a video for proof of performance.
4. Each athlete submits his/her own score.
5. An athlete may have a profile and not submit any scores for workouts.
6. Athletes can redo the workouts and resubmit their score within the give timeframe.

Section IV.3 User-Oriented Data Dictionary

DATUM	INFORMATION DEFINITION
Affiliate	The name of the CrossFit affiliate (gym) that the athlete trains with. Example: CrossFit Los Angeles.
Age	The age of the athlete at the time of application.
Date of Application	The date that the application is filled out.
Date of Birth	The date of birth of the athlete.
Email	The email address of the athlete to link to their account.
Gender	The gender of the athlete.
Height	The height of the athlete in the form of feet, inches. Example: 5, 11
Maxes	The score/weight that the athlete has achieved for each movement in lbs, kg, or reps. Example: Clean and Jerk, 370 lb. Max Pull-ups, 75 reps.
Name	The name of the athlete in the form last, first. Example: Doe, John.
Region	The Region that the athlete lives in. Example: Central East.
Score	The score that the athlete received at the end of a workout in the form of:

	(number, reps) , (weight, lbs/kgs), or (time). Example: Workout 1: 13 reps. Workout 2, 315 lbs. Workout 3, 5 minutes.
userview	A view that conceals private information to other users.
Weight	The weight of the athlete at the time of submitting the form.
Workouts	The best score the athlete has received on this workout throughout their career in the form of a time measurement or repetition count. Example: Fran, 2:13.

Section IV.4 Cross-Reference Table

Datum	Form or screen		
	Potential Athlete	Benchmark Form	Competition Score Submission
Affiliate	X		
Age	X		
Date of Application	X		
Date of Birth	X		
Email	X		
Gender	X		
Height	X		
Maxes		X	
Name	X	X	X
Region	X		X
Score			X
Weight	X		
Workouts		X	

Section V
Enterprise Database Design

Section V.1 Logical Model of the Enterprise

V.1.1
    Entities and Associated Attributes
- Athlete
  - a_ID

- o a_name
- o dob
- o a_gender
- o a_height
- o a_weight
- o a_region
- o affiliate
- o a_email
- o app_date
- Competition
  - o comp_name
  - o comp_type
  - o comp_score
  - o comp_region
- Benchmark
  - o b_name
  - o type
  - o b_score
  - o a_ID
  - o units
- Accomplishment
  - o workout_name
  - o a_score
  - o a_ID
  - o comp_name
- Participation
  - o a_ID
  - o comp_name
- Feat
  - o type
  - o units
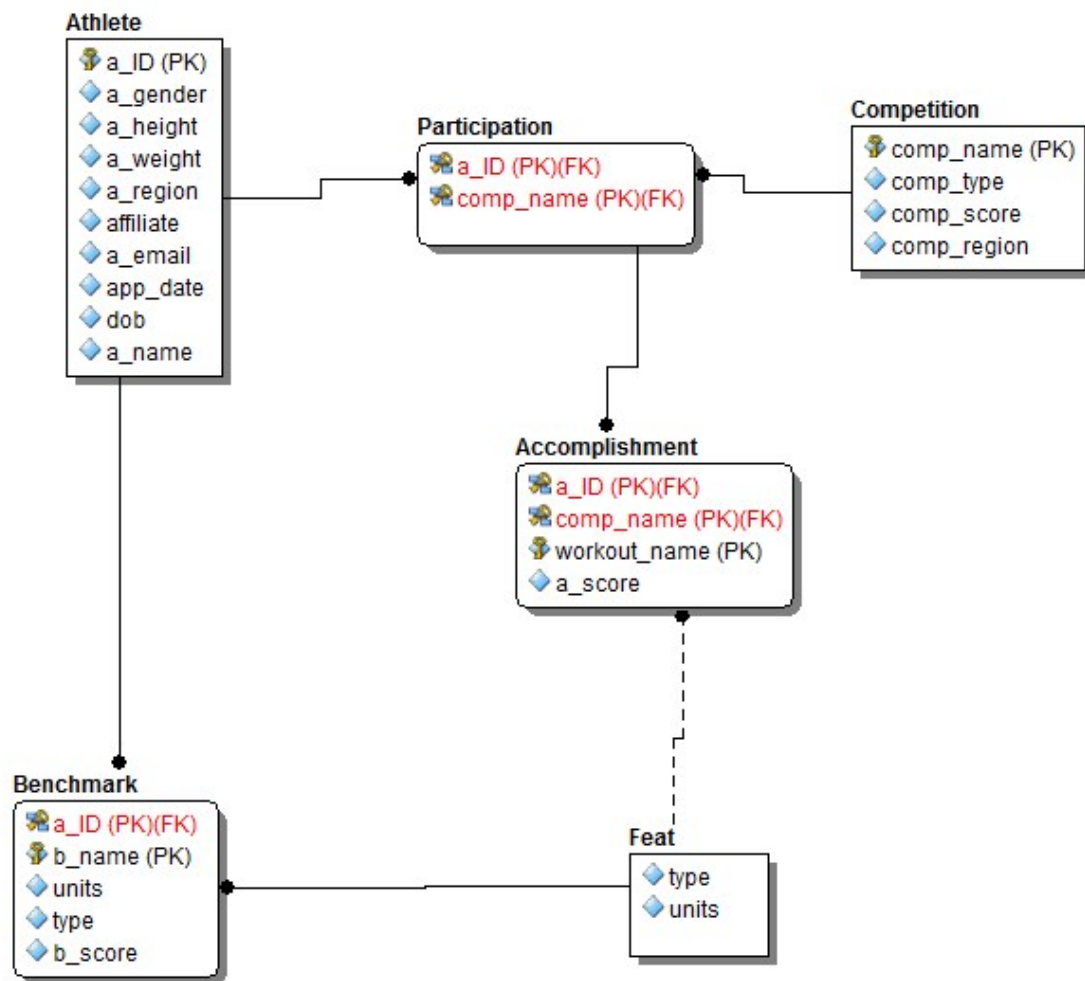
V.1.2

Relationships and Attributes
Participation: a_ID, comp_name

Athlete_Benchmark: a_ID, b_name

Competition_Accomplishment: comp_name, workout_name

V.1.3

Entity-Relationship Diagram

Section V.2 Conceptual Model of the Enterprise

Athlete(a_ID, a_name, dob, a_gender, a_height, a_weight, a_region, affiliate, a_email, app_date)
    PK: a_ID

Competition(comp_name, comp_type, comp_score, comp_region)
    PK: comp_name

Benchmark(a_ID, b_name, type, b_score, units)
    PK: b_name
    FK: a_ID references Athlete.a_ID

Accomplishment(workout_name, a_score, a_ID, comp_name)
    PK: comp_name
    FK: a_ID references Athlete.a_ID

8

comp_name references Competition.comp_name

Participation(a_ID, comp_name)
       PK: a_ID
       FK: a_ID references Athlete.a_ID
           comp_name references Competition.comp_name

Section V.3 Table Dictionary

TABLE	ATTRIBUTES	INFORMAL DEFINITION
Athlete	a_ID, a_name, dob, a_gender, a_height, a_weight, a_region, affiliate, a_email, app_date	Lists each athlete, their unique ID number and other attributes about them
Competition	comp_name, comp_type, comp_score, comp_region	Has every competition, each with a unique name, and stores the scores of each athlete involved
Benchmark	b_name, type, b_score, a_ID, units	A list of the Benchmark workouts with each athletes score
Accomplishment	workout_name, a_score, a_ID, competition_name	Describes a single workout done during a compeition. Includes the scores that each athlete received.

Section V.4 Attribute Dictionary

DATUM	INFORMATION DEFINITION	TABLE
a_email	The email address of the athlete to link to their account.	Athlete
a_gender	The gender of the athlete.	Athlete
a_height	The height of the athlete in the form of feet, inches. Example: 5, 11	Athlete
a_ID	A unique ID for a specific athlete	Athlete
a_name	The name of the athlete in the form last, first. Example: Doe, John.	Athlete
a_region	The Region that the athlete lives in. Example: Central	Athlete

	East.	
a_score	The score that the athlete received at the end of a workout in the form of either: (number, reps) , (weight, lbs/kgs), (time). Example: Workout 1: 13 reps. Workout 2, 315 lbs. Workout 3, 5 minutes.	Accomplishment
a_weight	The weight of the athlete at the time of submitting the form.	Athlete
affiliate	The name of the CrossFit affiliate (gym) that the athlete trains with. Example: CrossFit Los Angeles.	Athlete
age	The age of the athlete at the time of application.	None (Determined later by dob)
app_date	The date that the application is filled out.	Athlete
b_name	The name of a specific benchmark workout	Benchmark
b_score	The score that an athlete received on a specific benchmark workout	Benchmark
comp_name	The unique name of the competition. Example: 2014 Regionals: Central East.	Competition
comp_region	The Region that the athletes are competing in for this competition.	Competition
comp_score	The score that an athlete received at the end of a workout during a competition	Competition
comp_type	The type of competition. Example: Open, Regionals, The Games.	Competition
dob	The date of birth of the athlete.	Athlete
type	The type of workout. Example: Weightlifting, cardio, race.	Benchmark
workout_name	The unique name of a workout. Example: Fran.	Accomplishment

	Open Workout 14.1	

Section VI
Database and Query Definition

Section VI.1 Database Definition
Refer to crossfit_load_data.sql

Section VI.2 SQL Database Queries
 1) What did Froning score on Fran?
```
        SELECT a2.a_score
        FROM Athlete AS a
                , Accomplishment AS a2
        WHERE a.a_ID = a2.a_ID
                AND a.a_name = 'Froning'
                AND a2.workout_name = 'Fran';
```
 2) What is Smith's best score on the benchmark workout Fran?
```
        SELECT b_score
        FROM Benchmark
                , Athlete
        WHERE Athlete.a_ID = Benchmark.a_ID
                AND b_name = 'Fran'
                AND a_name = 'Smith';
```
 3) How many competitions has Bailey participated in?
```
        SELECT COUNT(p.comp_name)
        FROM Participation AS p
                , Athlete AS a
                , Competition AS c
        WHERE a.a_ID = p.a_ID
                AND c.comp_name = p.comp_name
                AND c.comp_region = c.comp_region
                AND a_name = 'Bailey';
```
 4) How many completed benchmarks does Foucher have?
```
        SELECT COUNT(b_name)
        FROM Benchmark AS b
                , Athlete AS a
        WHERE a.a_ID = b.a_ID
                AND a_name = 'Foucher';
```
 5) What athletes competed in the 2014 CrossFit Games?
```
        SELECT Athlete.a_name
        FROM Athlete
                , Participation
        WHERE Athlete.a_ID = Participation.a_ID
                AND Participation.comp_name = '2014 CrossFit Games';
```

6) What did Foucher score at the 2014 Regionals in the Central East?
        SELECT c.comp_score
        FROM Athlete AS a
            , Participation AS p
            , Competition AS c
        WHERE a.a_ID = p.a_ID
            AND c.comp_name = p.comp_name
            AND c.comp_region = 'Central East'
            AND c.comp_name = '2014 Regionals'
            AND a.a_name = 'Foucher';

7) Which Athletes competed in The 2014 Central East Regionals?
        SELECT a_name
        FROM Athlete AS a
            , Participation AS p
            , Competition AS c
        WHERE a.a_ID = p.a_ID
            AND c.comp_name = p.comp_name
            AND c.comp_name = '2014 Regionals'
            AND c.comp_region = 'Central East';

8) What is the best score for Fran during the 2014 CrossFit Games?
        SELECT MIN(a_score)
        FROM Athlete AS a
            , Participation AS p
            , Competition AS c
            , Accomplishment AS accomp
        WHERE a.a_ID = p.a_ID
            AND c.comp_name = p.comp_name
            AND accomp.a_ID = a.a_ID
            AND c.comp_name = '2014 CrossFit Games'
            AND accomp.workout_name = 'Fran';

9) Which athletes compete in the Central East Region?
        SELECT a_name
        FROM Athlete
        WHERE a_region = 'Central East';

10) Which athletes are affiliated with CrossFit Mayhem?
        SELECT a_name
        FROM Athlete
        WHERE affiliate = 'CrossFit Mayhem';

Section VI.3 Review Sign-off Sheet
        Turned in in-class
Section VI.4 Design Limitations

As of now, I cannot see any limitations in my design. It is set up in such a way that each entity can be recombined as needed to answer ad hoc and/or complex questions.

Section VII
Database Integrity and Security

VII.1 Functional Dependencies
a_ID -> a_gender, a_height, a_weight, a_region, affiliate, a_email, app_date, dob, a_name
a_email -> a_ID, a_gender, a_height, a_weight, a_region, affiliate, app_date, dob, a_name
affiliate -> region
b_name -> units, type
comp_name -> comp_type, comp_score, comp_region

VII.2 Adjustments for Normalization
First Normal Form: Every attribute is single-valued for each tuple
Second Normal Form: All nonkey attributes are fully dependent on the key
Third Normal Form: No nonkey attribute is transitively dependent on the key

Conclusion: The database is in third normal form.

VII.3 Integrity and Security
Classes of users: Database Administrator, User

1) Administrator: has access to everything
2) User: has access to all information other than the email address and date of application of users
   a. Created a view called userview which conceals this information. This is in crossfit_load_data.sql.

Section VIII
Implementation Notes

VIII.1 Indices
All of the indices used in my database are to the primary keys of each of my entities. This is because these are the terms/information that will be most likely searched for.

VIII.2 Data
Reference crossfit_load_data.sql

VIII.3 Query Trace

Reference DB-project-Sullivan-trace.rtf and queries.sql

VIII.4 Implementation Assessment

Overall the implementation of data into my database was smooth. One problem that I had was loading and dropping the tables in the proper order. Because of the use of foreign keys, if this is not done properly, certain tables will not be able to be created or dropped. I also added a trigger in my database so that whenever a workout (accomplishment) is logged that is also a benchmark workout, it checks if the athlete beat their previous best time. If they did, it updates their benchmark score; otherwise it leaves it the same. This can be seen in crossfit_load_data.sql.