Brian Taylor

IST 659 M402

Youth Basketball Performance Evaluations

Overview

Youth sports in Northern Virginia have become much more competitive at a much earlier age than when I grew up here 30 years ago. Information on youth player performance has become more important to capture for a variety of reasons, including high-school and college sports recruitment, national competition participation (such as Junior Olympics) and qualification for more competitive developmental programs. I have had the privilege of coaching my children in soccer, lacrosse, swimming and now basketball at the local club level.

In order to create comparable and friendly competition, the local basketball club conducts assessments of all players before the season starts. Youth coaches draft the young players based on combined assessments from coaches assigning a numerical value for each player, with “5” being the highest ranking, “1” being the lowest. Coaches are not allowed to stack their teams with a bunch of “4” and “5” players. As the players participate over multiple seasons, they accumulate more assessments. Additionally, there is a player’s coach assessment at the end of each season. To clarify, a second season player should have three assessments: their initial assessment from all of the youth coaches from their first assessment the previous year (1st), their coach’s assessment at the end of their last season (2nd), and another “all coaches” assessment (3rd) in time for their second season team draft.

With around 90 players per season, and approximately 10 head coaches evaluating 5-10 performance metrics over multiple seasons has created enough of a data collection and analysis problem to support development of a database to store this information. The information can then be stored, updated and used for further data analysis.

However, upon implementation and further consultation with the primary shareholders representing the league, they were extremely limited on the data they would allow me to use. As a result, I was not able to gather multiple coach assessments on each player, only a rough average of the attributes averaged over the players. As a result, the database design changed as will be explained later.

Design Specifications

The database should capture the performance of each player in the evaluated sessions, as well as the player’s coach for the season. The stakeholders for a program like this include the players, the players’ parents, the volunteer coaches, the age group commissioner, the youth basketball program president and Dr. Block. Players and parents can use the information to focus on specified development in the off season, if desired. The coaches and commissioner could use the information to ensure players aren’t gaming the system, evenly distributing the talent throughout the teams and possibly identify coaches that are developing players a bit better or worse than their counterparts. The program president’s stake involves protecting the privacy of the data, so as not to violate any laws, agreements or sensitivities regarding identifiable information about minor children.

Coaches are responsible for providing their assessment sheets to the commissioner, or designated representative, for consolidation and entry into the database. The president and commissioner and can request reports pertaining to coaches and players. Coaches can only request reports on their players and themselves. Players and parents can only request information on themselves. Dr. Block can use the information as necessary to evaluate for educational purposes only.

Business Rules

As with any system that utilizes untrained evaluators, there is a serious question of standardization in the evaluations. As such, the multiple evaluations from the coaches, the ratings will be averaged. In fact, the scores were averaged before I received them, which required a change to the design of the database tables described later. Any missing data in the performance metrics will be assumed to be a value of 3, the mid-point rating on a Likert Scale, after discussion with many coaches who stated that due to the pace of the assessments, they often left values blank if they would have scored them a 3 in the effort to save time and pay better attention to player performance. Allow recorded Likert values to be decimal, not just integer. Coaches can only record first names and possibly last name initials as necessary to protect personal information.

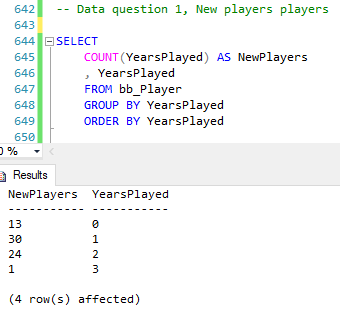
In practice, I was not allowed to average the coaches’ evaluations due to privacy concerns of the club. The commissioner provided me with the rough (rounded to the nearest whole number) averages that he calculated. This resulted in one assessment per evaluation period (pre and post season, each year) rather than multiple assessments on each player from the same evaluation period. Further, I was not privy to previous team assignments for the players from previous seasons.

Data Questions

The top five data questions that this database will support are:

1. How many new players are playing this season?

This output that there are 13 players (about 19%) playing this season that have not played in this league before. Discussion with the stakeholders revealed that is higher than they have seen lately.

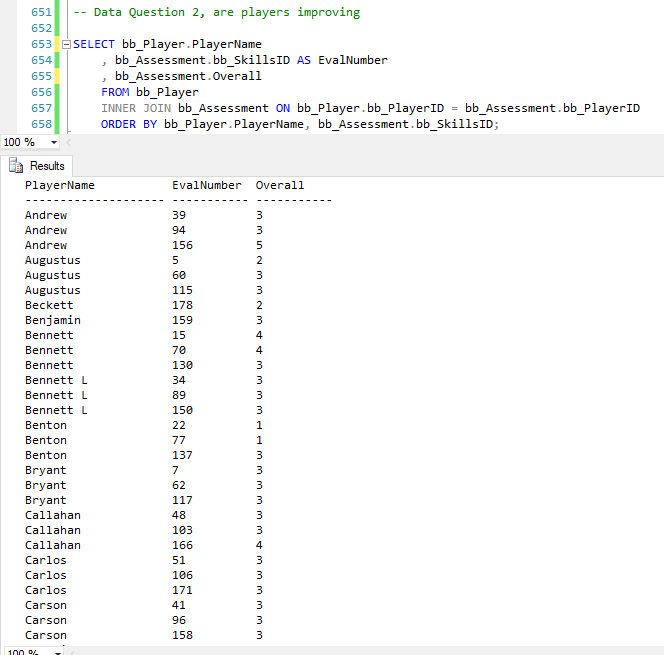


Access report:

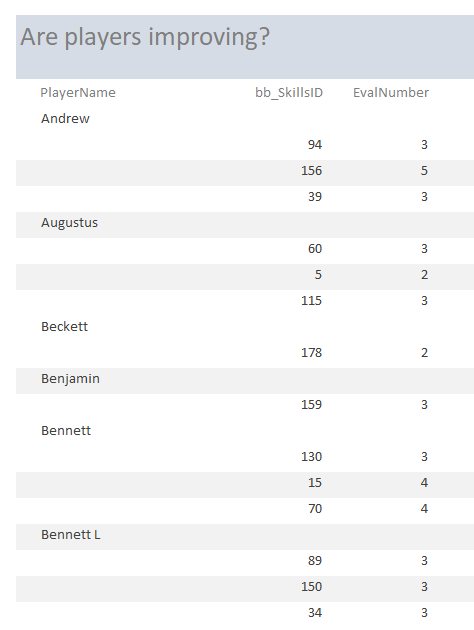


1. Are players improving?

This query shows the grouped results by player, in chronical order of their overall scores from their assessments. By evaluating up to three scores per player, a coach could confirm or question their gut judgement. Factors playing into this score is that as the players age, the rest of the field improves a 3 last year is not indicative of the same objective performance this year. These scores are not independent of cohorts performance.

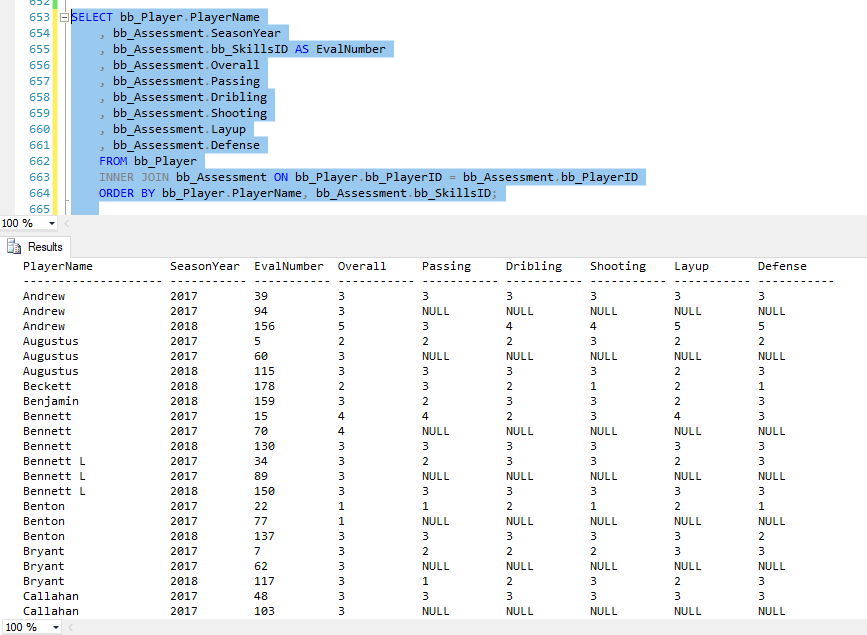


Access report for question 2:



1. In what areas are players improving or regressing?

This review of the players performance trends on a skill-by-skill basis to evaluate if certain areas are improving per player.

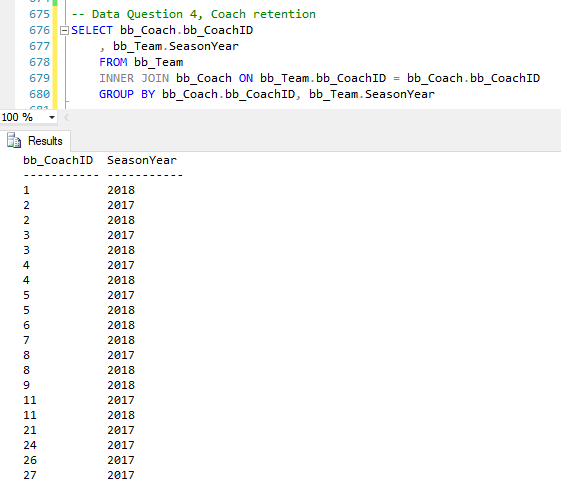


Access report for question 3:



1. What is coach retention, season to season?

There are 4 coaches that did not return either as a head coach or an assistant coach for the 2018 season, however, 4 new coaches volunteered. This is about 20% turnover, a bit higher than the commissioner and club president would like to see.

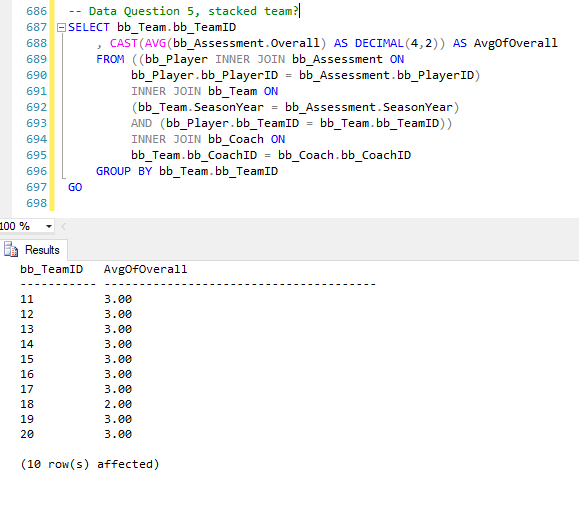


Access report for question 4:



1. Is there a stacked team?

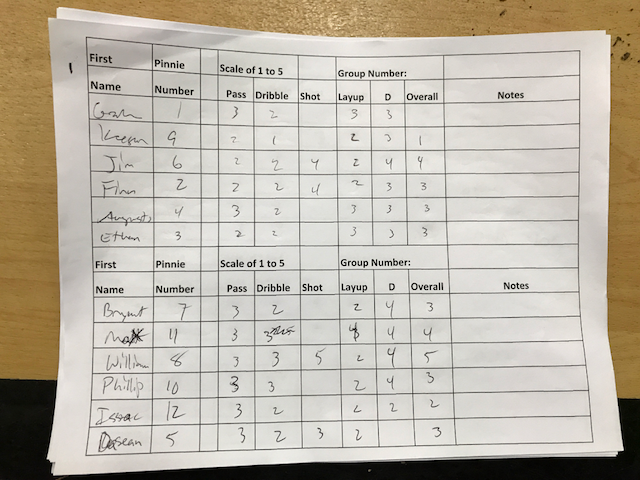
Based on the average overall scores on the assessments, the teams appear fairly well matched with only one team averaging less than a 3.



Access report for question 5:

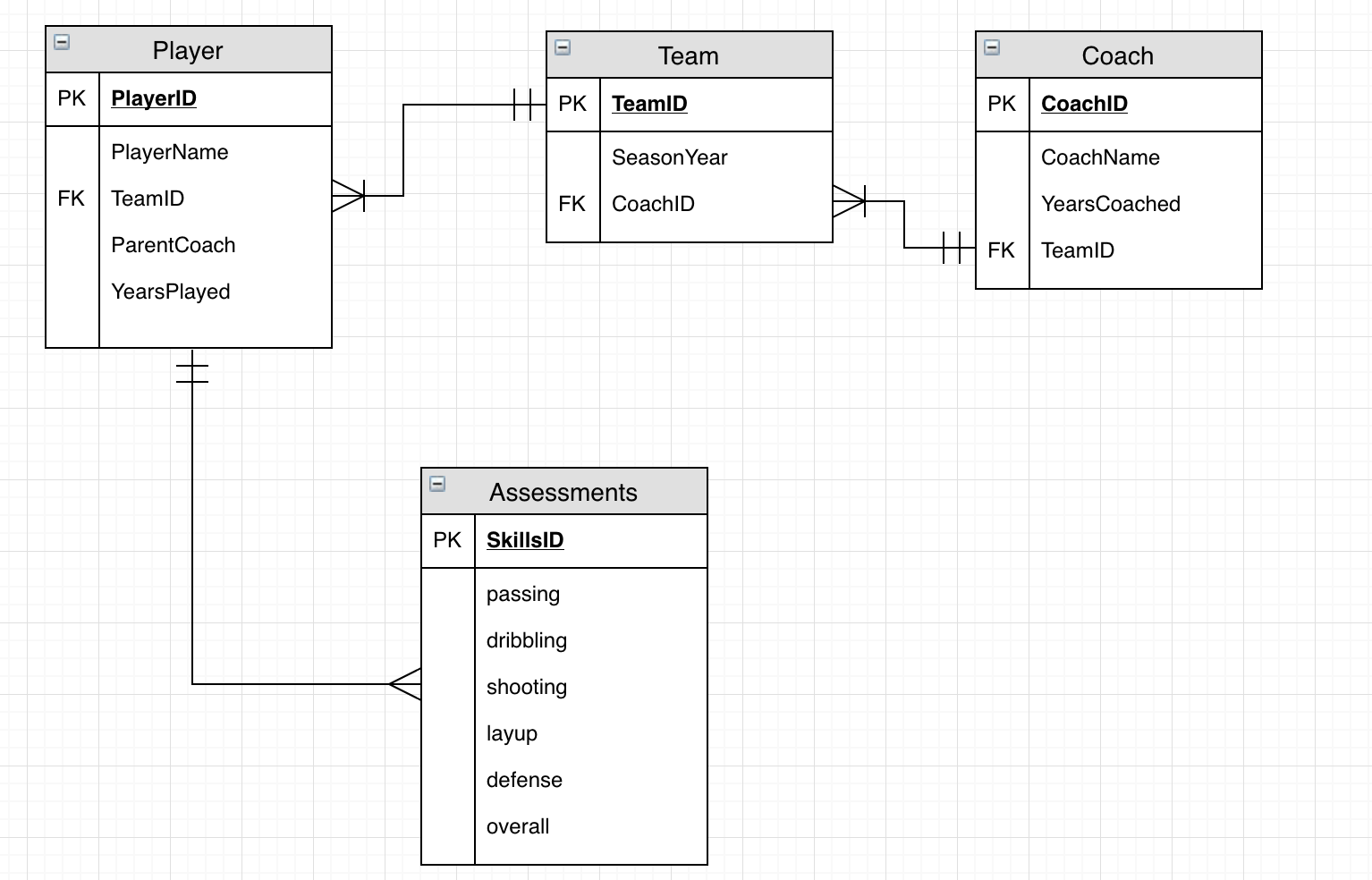


Data Example



The overall column is not an average or summary of the preceding columns. Nature of the movements associated with layup and shot, attributes are likely to be similar.

Entity Relationship Diagram



ERD Glossary

Player – player entity

PlayerID – surrogate key of the player

PlayerName – Name of player

CoachID – surrogate key of the coach of the team

TeamID – surrogate key of the team that multiple players and each coach are assigned to, foreign key to identify teams over time

Parent Coach – attribute is “Y” is the player’s parent is a coach

YearsPlayed – years player played in league

Coach – coach entity

CoachName – Name of coach

YearsCoached – years coach has coached in the league

Assessments – entity capturing the skills evaluated of the player by the coach

SkillsID- surrogate key of the assessment

Passing – passing skill of the player

Dribbling – dribbling skill of the player

Shooting – shooting skill of the player

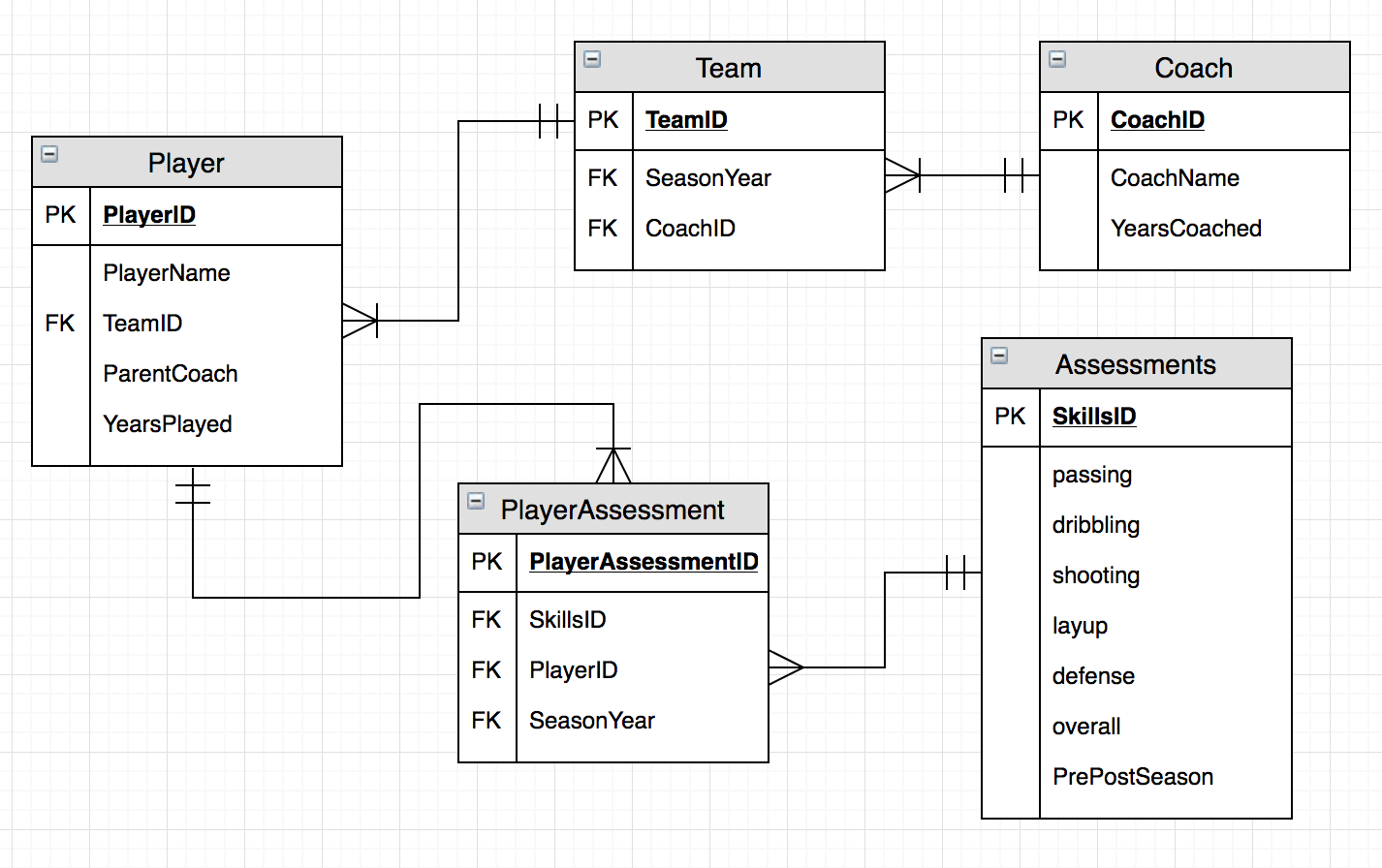
Layup – layup skill of the player

Defense – defensive skill of the player

Overall – overall skill of the player

A player needs a coach, and a coach can coach many players. Players can have many assessments of her skills, and the skills apply to many players. Coaches can make many assessments of the skills of many players and a player’s skills can be assessed by many coaches.

Original Logical Model

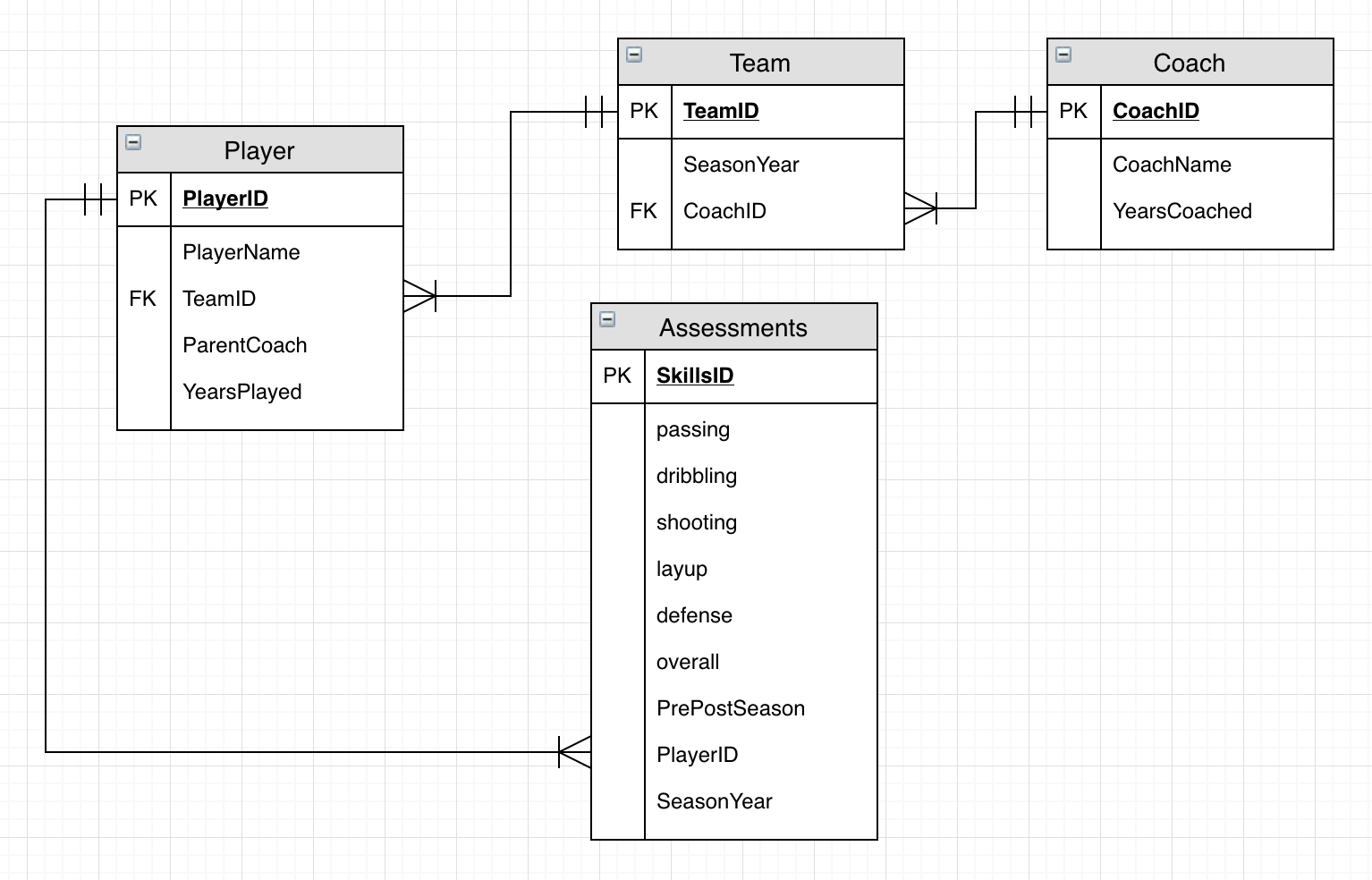


The Team entity developed from the Coach entity to account for coaches coaching multiple teams over multiple years and seasons. Originally, the PlayerAssessment entity represented each, individual assessment. This entity depends on multiple foreign keys to link the time and player to the attributes assessed by coaches. In practice, the Primary key of the PlayerAssessment table simply duplicated the SkillsID key from the Assessment table and was difficult to manage. By simply moving the PlayerID and SeasonYear attributes to the Assessment table, the PlayerAssessment table was removed. Upon further reflection and the previously mentioned change regarding the raw data, the relationship had changed. The data entered was already combined from the multiple coaches by the commissioner and the need for combining many assessments for many players was already computed. Therefore, the relationship also changed to simply a one-to-many between the Player and Assessment table.

Despite the fact that coaches create the assessments, they are not tracked by coach. However, the post season assessment can be attributed to the coach of the season, so PrePostSeason is an additional attribute to the Assessments identity from the ERD. Further, the previous many to many relationship between the Coach and the Assessments is no longer important to track, based on the commissioner and club president consultation.

Each coach can coach multiple teams over the seasons, but must coach one at a time to be considered a coach. There are many players on each team, and players cannot be on multiple teams in the same season, thus the addition of the SeasonYear attribute in the Team entity. Each player must have at least one assessment, but could have many over the seasons and there must be a player to have an assessment. The Assessments entity has attributes that are used on multiple assessments, however, at this point, only two similar format exists and must exist.

The revised logical model is below:



Additional Attributes

SeasonYear- the attribute that helps distinguish when an assessment was conducted

CoachID – surrogate key of the coach of the team and foreign key for the Team entity

PlayerAssessmentID – the primary, surrogate key for each player’s assessment

SkillsID- surrogate key of the assessment

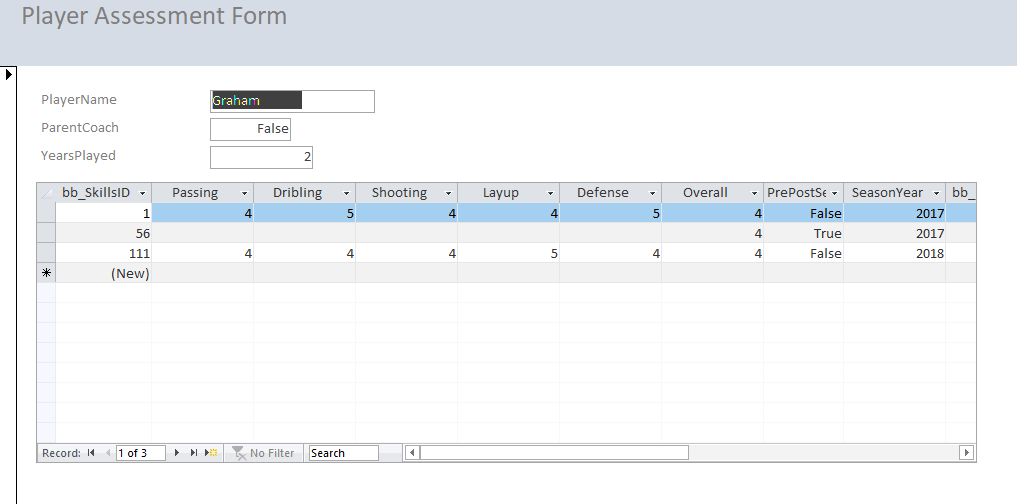
PlayerID – surrogate key of the player and foreign key for the PlayerAssessment

PrePostSeason – an attribute differentiating the mass pre-season assessment and the coach’s post season assessment

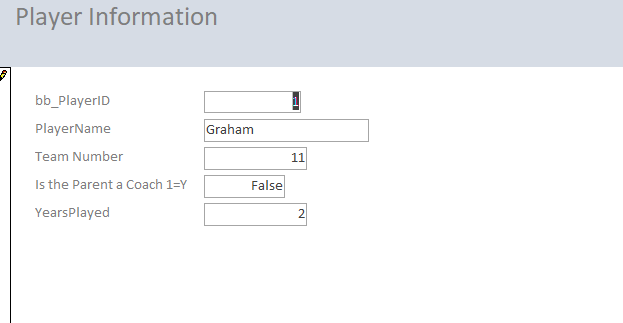
All primary keys are integer and will use the identity property. YearsCoached, SeasonYear, YearsPlayed, and all basketball related skills in the Assessments entity will also be integers as the data is inherently numeric and will be used to make calculations. All others are character fields with ParentCoach and PrePostSeason limited to binary, “Yes”/”No” and “Pre”/”Post” values.

Implementation

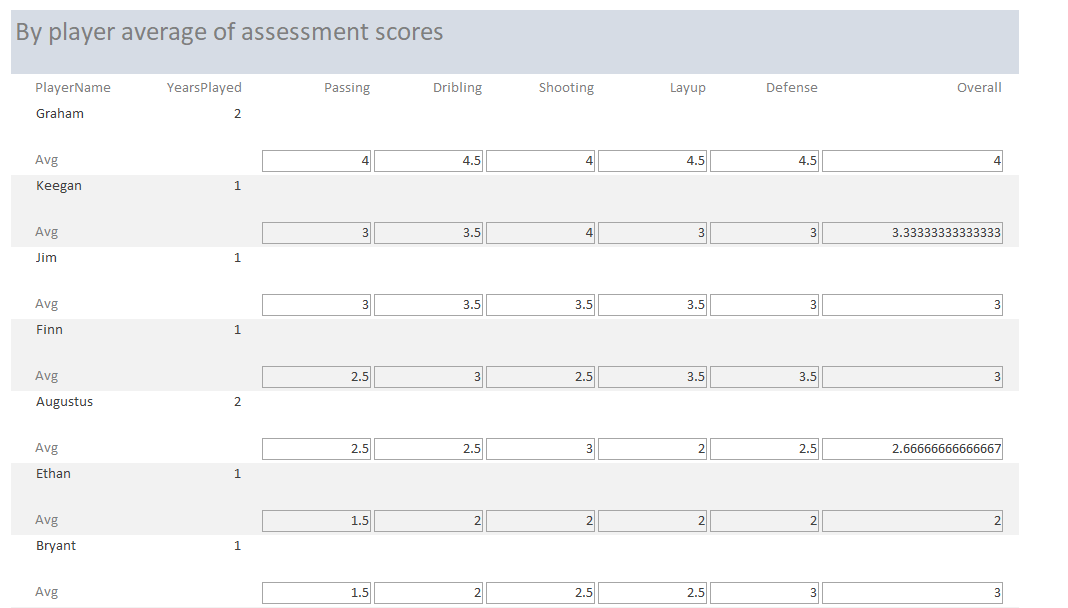
This is a form created with Access to record or review player assessments.



Below is the basic Player Information form used to enter or update player information.



Here is a report showing the average player scores from all of their assessments.



Reflection

Although the task appeared relatively simple on the outset, it became quite daunting. The negotiations with the club president as well as the commissioner was more difficult in expected, and resulted in unexpected design changes. Obviously, my knowledge of database structure and implementation greatly improved from week 4 until know, however, the use of Access as a front end was more difficult than expected. I appreciate the power of a good front end, as once it was connected to the database, entry and updating became much easier.

Next time, and I hope and am pretty sure there will be a next time, I will spend more time on the design to try to flush out the relationships better. I will also search for built in functions and other resources outside the scope of class material to hopefully aid in implementation. I realize I have much more to learn about databases and coding in general.

Summary

This database should allow multiple data questions to be answered through the organization and compilation of multiple assessments for multiple players over many seasons. It can be updated and queries can be constructed to provide standardized reports of interest for multiple stakeholders to find value. The use of Access as an interface is due to the inherent power and flexibility of the program as well as its ability to easily interface with SQL Server.