

**Part 1**

Now you have to make some decisions about if you are more interested in Sports data and being a sports analyst or more interested in music data and being a music analyst.

In the last module, you looked up some statistics websites for sports or music data and saw there is so much data available. Even though computers can analyze lots of data, you don't want to waste energy analyzing things that are not important in decision-making. Today you will look through and decide what data is important. You will focus on either music or sports (but you can do both if you have time) to decide on the factors that are important.

Start with some videos on what people think are important to consider.

**Music**

Where do music producers look for new talent. Watch these videos on how companies discover talent.

<https://www.youtube.com/watch?v=Z_IiBxRPSaQ>

and this video on why Atlantic Records signed Dua Lipa

<https://www.youtube.com/watch?v=BpktkepReno>

What traits, metrics (voice, looks, communication, creativity…) do they look for in a new artist and what places or data sources do they use to discover?

**Sports**

Where do sports agents look for new talent?

<https://www.youtube.com/watch?v=O80XPnxiacU>

How technology aids sports talent discovery

<https://www.youtube.com/watch?v=g-_L4rbMC1c>

What are some of the metrics they think are important?

**Narrowing down the data**

In the last module you looked at data and started writing down some of the information. Below is a snapshot of some data from the WNBA website. There was at least 25 metrics on the website, although only a few were recorded here. How do you know which ones are really important in deciding if a player is really talented or if you wanted to compare a young unsigned player to some current (and highly paid) star?

EX: Website sports - WNBA

| **Name of website** | **URL** | **Data Contained** | **Other information or notes including API access** |
| --- | --- | --- | --- |
| Ex  WNBA Stats | <https://stats.wnba.com/> | Team, age, GP, W, L, Min, PTS, FGM, FGA…. A lot more data | Filters to search by name or years. Can’t API directly, but could paste in a webpage. I searched and the API is on a diff website <https://sportsdata.io/developers/api-documentation/wnba> |

| Statistic/Data | Data Type (primitive, non -primitive…) | Data Type for computer | Example value |
| --- | --- | --- | --- |
| Name | Non-primitive | Character | Chennedy Carter |
| Team | Non-primitive | Character | Chicago Sky |
| Age | primitive | Integer | 26 |
| GP | primitive | Integer | 7 |
| W | primitive | Integer | 3 |
| L | primitive | Integer | 4 |
| PTS | primitive | Float | 12.0 |
| FGM | primitive | Float | 5.1 |
| FGA | primitive | Float | 9.3 |
| Mins | primitive | Float | 18.2 |

Well, it winds up there are only a few metrics that are really important to define talent in the WNBA. It winds up that there are only 4 factors that are really important.  There are a lot of websites that describe the four factors that are important in baseball.

<https://stats.wnba.com/help/faq/>

**TO DO:**

1. Go to the web. Pick either sports or music and try and narrow it down to a particular sport or music genre. For example, the factors required for a talented soccer forward is different than the factors for a midfielder. A country music artist may have different metrics for success than a reggaeton artist. Make sure to make a list of your references.

2. Once you do this, look for sources of this data for the metrics determined to be important. How would you find out data on "creativity" or "communication skills" how do you think those might be determined? Check to see if the site allows for API access.

3. Now that you have sources of data, make a table for that important data and then go look up the information for at least 25 to 30 great artists or players so you have some data you can compare in Matlab. You can start in Word, and then the table can be pasted into an Excel file.

The table might look like

WNBA

| Name | Team | Mins played | PPG | FGM | FGA |
| --- | --- | --- | --- | --- | --- |
| **A'ja****Wilson** | Lva | 33.8 | 27.8 | 10.2 | 19.5 |
| **Arike Ogunbowale** | dallas wings | 38.5 | 23.9 | 7.9 | 22.1 |
| angel reese | chicago sky | 30.5 | 13.2 | 4.5 | 11.5 |
| caitlan clark | indiana fever | 33.7 | 16.3 | 4.9 | 16.3 |
| breanna stewart | ny liberty | 34 | 19.7 | 7.1 | 14.7 |
| brittney griner | PHO | 30.4 | 19.7 | 7.7 | 13.5 |
| alysha clark | lvs | 29.3 | 8.2 | 2.9 | 5.4 |
| **Dearica****Hamby** | LA | 35.5 | 17.4 | 6.8 | 14.3 |
| **Napheesa****Collier** | MIN | 34.8 | 20.9 | 8 | 16.9 |
| **Rhyne****Howard** | ATL | 32.1 | 15.3 | 4.8 | 12.5 |
| **Alanna****Smith** | MIN | 28.2 | 12 | 4.4 | 8.6 |
| **Chennedy****Carter** | CHI | 21.3 | 14.3 | 5.9 | 10.7 |
| **Stefanie****Dolson** | WAS | 25.7 | 9.4 | 3.3 | 7.0 |
| **Kayla****McBride** | MIN | 32.5 | 16.1 | 4.9 | 10.7 |
| **Jonquel****Jones** | NY | 30.1 | 16.8 | 6.4 | 10.8 |
| **Cameron****Brink** | LA | 22 | 7.5 | 2.7 | 6.9 |
| **Cecilia****Zandalasini** | MIN | 12.5 | 4.1 | 1.4 | 3.4 |
| **Skylar****Diggins-Smith** | SEA | 32 | 14.8 | 5.4 | 13.8 |
| **Sabrina****Ionescu** | NYL | 33.2 | 19.3 | 6.7 | 15.3 |
| **Ezi****Magbegor** | SEA | 32.4 | 13.3 | 4.8 | 9.9 |
| **Diana****Taurasi** | PHO | 29.6 | 16.3 | 5.4 | 13.7 |