

**Department of computer science**

**CS 7260 (S01) ADVANCED DATABASE SYSTEMS – (Fall 2019)**

**Project**

**-------------------------------------------------**

**Christopher Regan**

**CREGAN1@students.kennesaw.edu**

**11/30/2019**

1. Website chosen:   
   <https://pastebin.com/raw/Bjsw46PD>
   1. This site was chosen due to simplicity.
2. The schema design for a one table solution.
   1. The design is 2 fielded
      1. There is a column for the player index
      2. There is a column for the player name
3. See code solution here: <https://github.com/GIGA-Money/pyBeautifulSOUPWebScraperProject>
4. Using a localhost MySQL database solution for storing the data.
5. Create table players (
   1. ID VA character(15),
   2. "player\_name varcharacter(255),
   3. primary key (ID));
6. The major query:
7. for vdex in values:
8. for edex in entity:
9. activeCursor.execute("insert into players values ('%s','%s')" % (vdex, edex))
10. Inserts from each list the value and the entry, as they were separated into 2 list for easy insertion.
11. The methods:
    1. # This will format the string, by removing excess text that isn't used.
       1. def text\_format():
    2. # This method extracts the index from the name into 2 separate list.
       1. def splint\_string():
    3. # This will insert the values into the database. Using the existing lists, values and entity.
       1. def insertion():
12. to connect to the web pages, I use the following:
    1. site = <https://pastebin.com/raw/Bjsw46PD>
       1. the web pages, URL.
    2. her = {'User-Agent': 'Mozilla/5.0'}
       1. the web agent.
    3. context = ssl.\_create\_unverified\_context ()
       1. ssl verification indicator
    4. req = Request (site, headers=hdr)
       1. the request header and site
    5. page = urlopen (req, context=context)
       1. the page requests
    6. soup = BeautifulSoup (page, "html.parser")
    7. formattedString = text\_format()