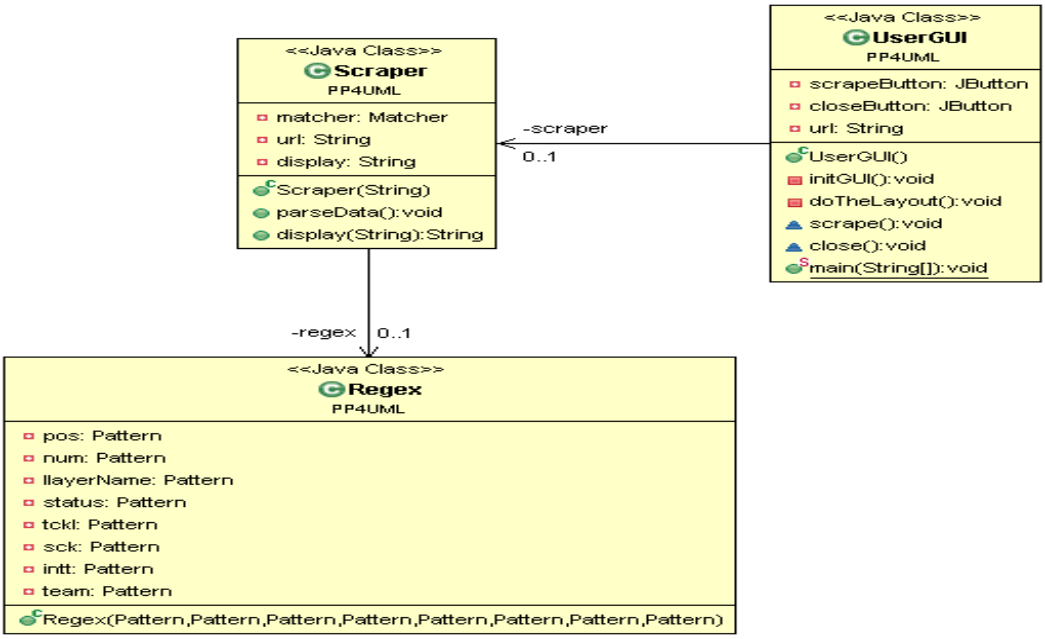
Project Assignment 4
75 Points
References: Referenced Textbook and Week 2, 3, 4, 5, 6, 7, 8, 9, & 10 handouts

Skills Required:

1. Advanced String manipulation techniques
2. Regex, Pattern & Matcher Classes
3. String tokenizer
4. Understanding exceptions, Try and Catch
5. Reading data from the web sources and writing data to files

Description:

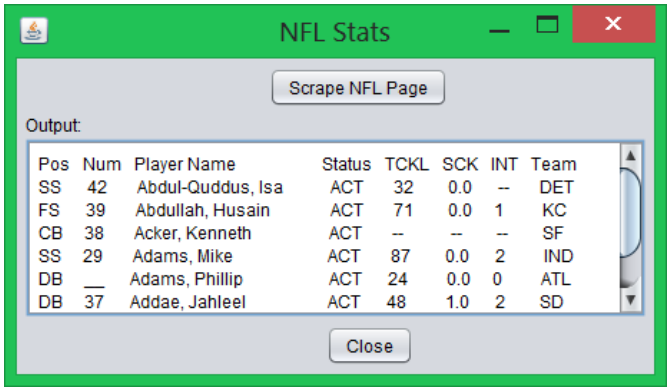
1. The program finds specific text strings of statistics in a website using regex
2. A URL link of the web source HTML file for the NFL statistics is provided



Task Specifications:

1. Create a Java project PP4 in your Eclipse workspace
2. The project has three java classes as shown above, `Regex.java`, `Scraper.java`, and `UserGUI.java`
3. The program must be implemented following the provided class diagram model
4. All class attributes must be private, that means you need public methods to access class attributes
5. `Scraper.java` has two methods to read and display the web source data
6. The *parseData()* method reads the data from a web page and searches for the match with your defined patterns, using the Java `Matcher` and `Pattern` classes

7. The *display()* method shows the output (scraped data) in a text-area component in user GUI
8. The `UserGUI.java` class is the main class that implements the user interface as shown below:



9. When clicking on the Scrape Data button, your program fetches the data and display it in the text-area component
10. A user must be able to scroll up/down within the text-area component to read the displayed data
11. The Close button terminates the program
12. `Regex.java` defines all the patterns (must use special characters) using Pattern class needed to find the required statistics data in the NFL URL web source
13. Your program reads HTML source page content from the NFL.com website using the Java `URL` and `HttpURLConnection` classes
14. You program should open the following URL at NFL.com to find stats of the players classified as Defensive Backs
    a. URL: http://www.nfl.com/players/search?category=position&filter=defensiveback&conferenceAbbr=null&playerType=current&conference=ALL
15. It captures the HTML web source page into a String object reader
16. Using regular expressions, find and extract the following data for all players of the Defensive Backs type:
    a. Pos (position)
    b. Num (jersey number)
    c. Name (last, first)
    d. Status
    e. Quick Stats (TCKL, SCK, and INT)
        i. Numbers of Tackles
        ii. Numbers of Sacks
        iii. Numbers of Interceptions
    f. Team

17. Write the stats into a single file; one row per player, columns delimited by tabs
18. Note that there are several pages of players, so your application must account for this by going to each page until the end (last page of players' stats), use the page numbers to iterate the pages and extract web source data into a String object reader
19. Write scraped data to the text-area component in GUI in multiple rows as shown above
20. Write the output data to a text file, name it NFLStat.txt
21. Format your file properly to so that the data is stored as shown in the original URL page, similar to a table with columns and rows for players' information

Evaluation Criteria:

1. All tasks must be completed to receive credit for this assignment
2. Output should report the correct/expected data values
3. Output file should be formatted for easy viewing
4. Program should not crash while scraping data from the HTML page from NFL.com URL

Grading Rubric - Project 4 (PP04)

Student Names: _____

Evaluation Criteria	Comments	Max	Points
Program compilation and running (Program should run on your machine during the video call/demonstration): Proper use of the directory structure and the class templates provided (6) Successful compilation and running, meaning no unexpected runtime error, or crash (2) The program should not crash if the data were scraped instead from another URL, say nba.com URL (2)		**10**	
Reading and writing data (Demonstrate that it works on your machine during the video call):		**10**	

Reading data from the URL (4) Writing data to the file NFLStat.txt (4) Output file is properly formatted for easy reading (2)			
General code structure (Walk-through of the code during the video call): Proper use of GUI components (3) Proper use of String manipulation, RegEx, Pattern, Matcher, StringTokenizer (8) Proper use of exception handling with Try and catch (3) Proper use of objects and the association e.g., zero to 1 (3) Proper scope (private, public etc.) of class attributes and the methods, proper method signature and return type, proper calling of the methods from the other classes (3)		**20**	
Functionality (Demonstrate that the functions work during the video call): Proper display of the data in the JTextArea component (5) Ability to scroll up/down within the text-area (5) The proper finding of the player data and the quick stats (10) Accounting for several pages of player data (10) Proper functioning of the Close button (5)		**35**	
Total		**75**	

Total = _____/75