**Ciarán Moyne - 123…...**

**Ciaran O’Hagan -** **12317758**

**Simon Markham - 12307233**

**ST3009 - Statistical Analysis- Group Assignment**

**FIFA World Cup**

For this assignment, we acquired two data sets from the internet which contained data which focused on various aspects of the FIFA World Cup:

A dataset compiled by the Guardian containing data concerning individual teams and their performance in the competition from when it began in 1930 to 2002.

The data set can be found here:

[https://docs.google.com/spreadsheet/ccc?key=0AgdO92JOXxAOdGtRLThiUUhYSnhackhXVm9qbm5aQ0E#gid=0](https://docs.google.com/spreadsheet/ccc?key=0AgdO92JOXxAOdGtRLThiUUhYSnhackhXVm9qbm5aQ0E" \l "gid=0)

A German dataset containing detailed data on the players who played in the 2014 Brazilian World Cup.

The data set can be found here:

[https://docs.google.com/spreadsheets/d/1i7aUrjBcle7XtEbRUqQrpSLi6cGxIIE1MpJyPpfdg5Y/edit#gid=1426167629](https://docs.google.com/spreadsheets/d/1i7aUrjBcle7XtEbRUqQrpSLi6cGxIIE1MpJyPpfdg5Y/edit" \l "gid=1426167629)

For our Assignment we plotted several different views of this data.

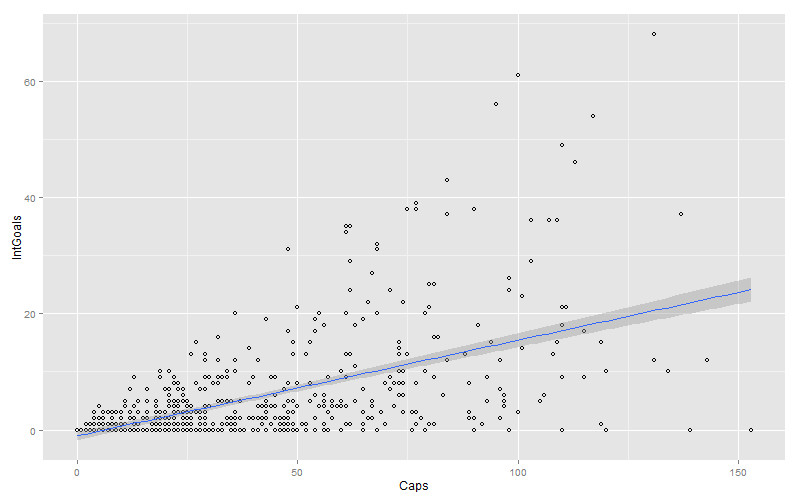
1. The number of goals a player has scored for their country against the number of Appearances, using a scatter plot.
2. The Total number of times a country has been in the World Cup, using a bar plot.
3. The height of a player against the number of goals they have scored, using a scatter plot.
4. The Ratio of Goals per game against the number of appearances, using a scatter plot.
5. The Height of Players against the number of goals scored, investigating whether or not it is better to be a taller player.

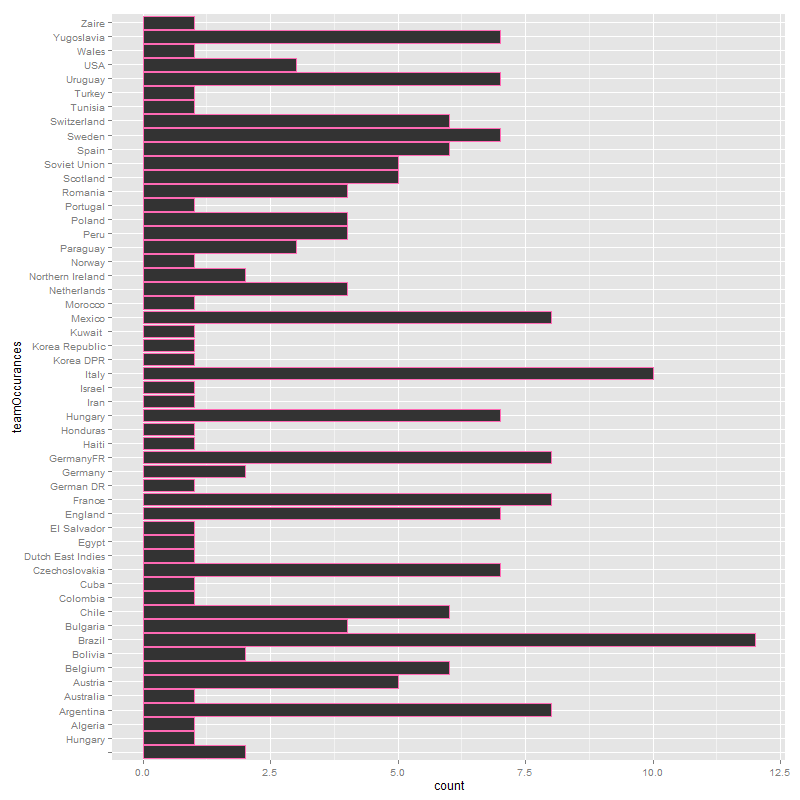
We found there to a correlation between the number of goals scored against the number of appearances a player has, however, there are a few outliers. These being the defenders and the goalkeepers, who don’t normally score that many goals.

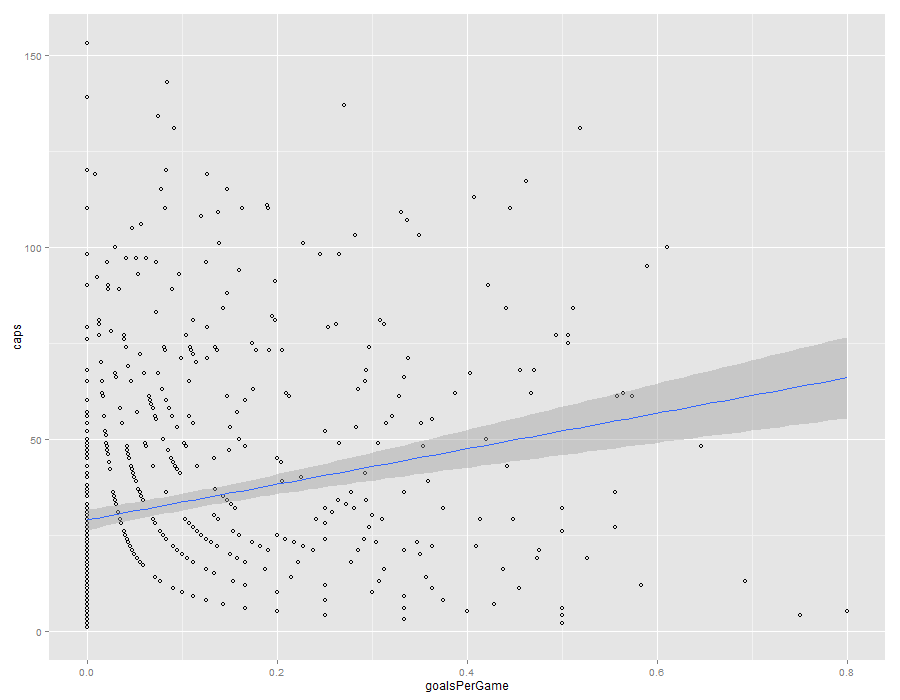
This was also the case for the goals per game ratio too.

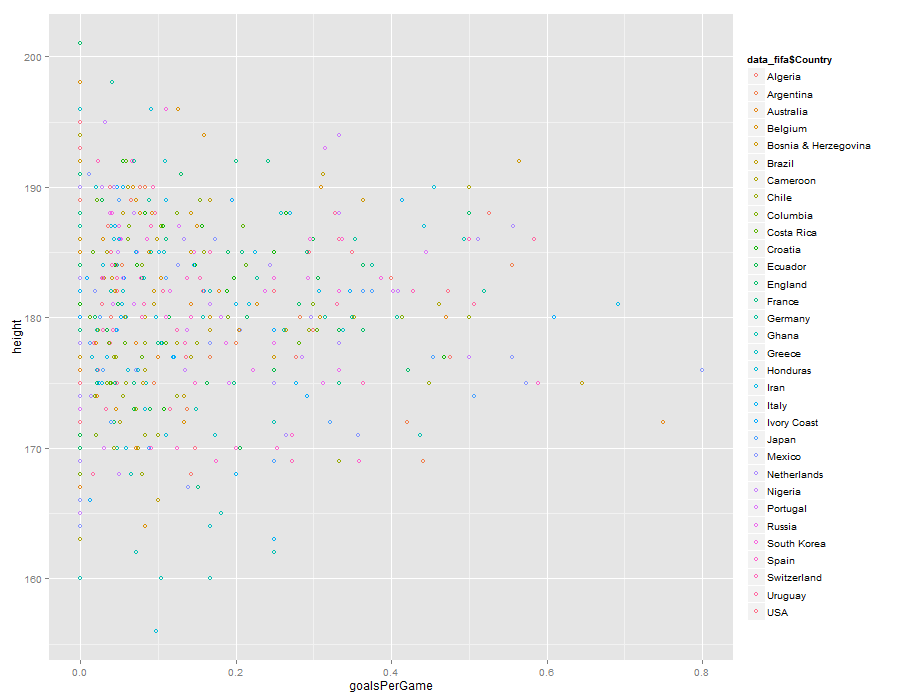
All our plots were used using ggplot.

**Statistical distributions:**





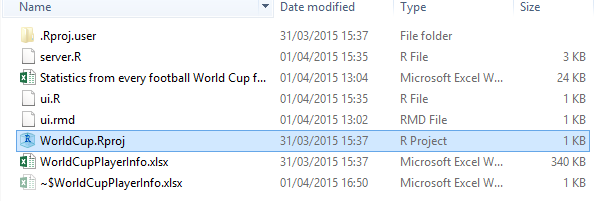




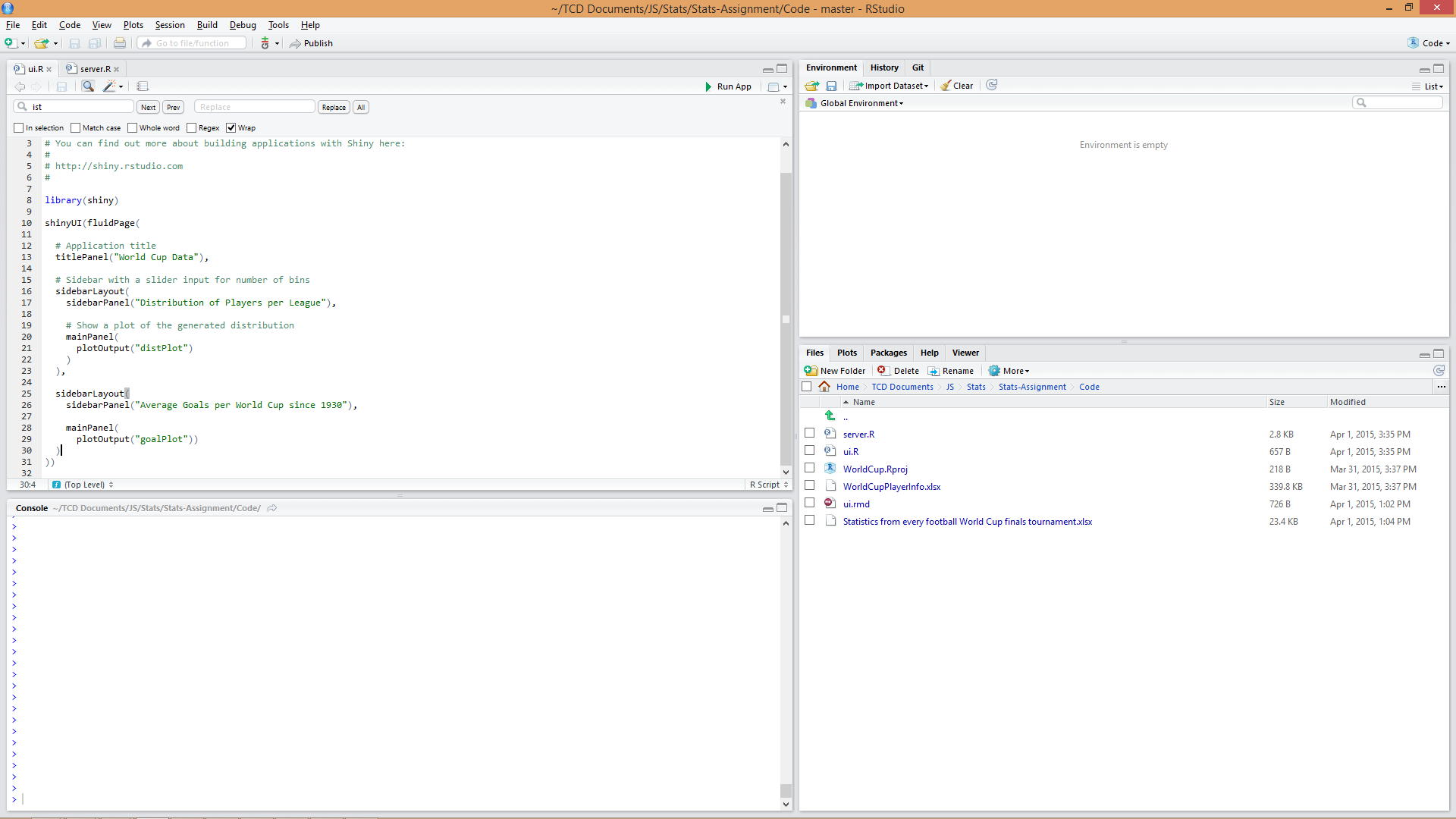
**How to run Application:**

To run any ShinyR application, one must go through RStudio.

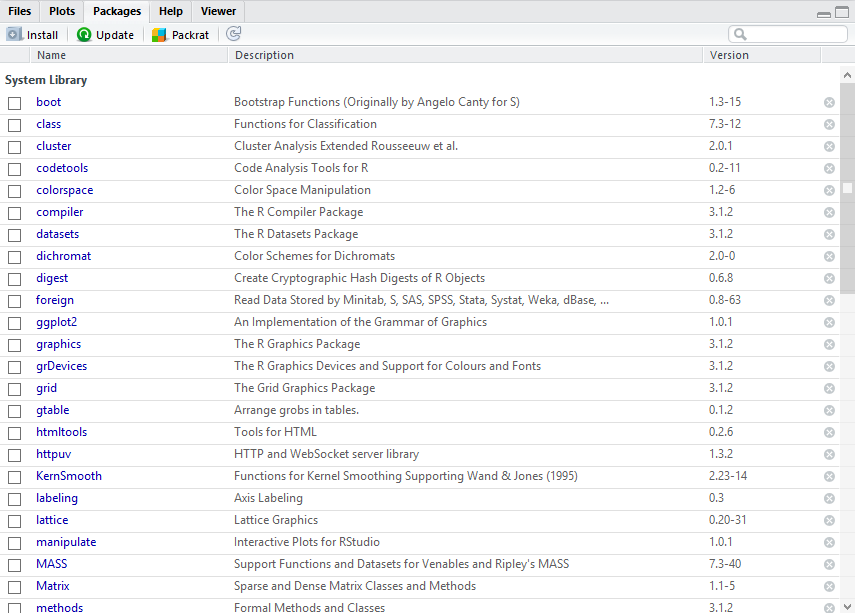
First, open a ShinyR project in RStudio:



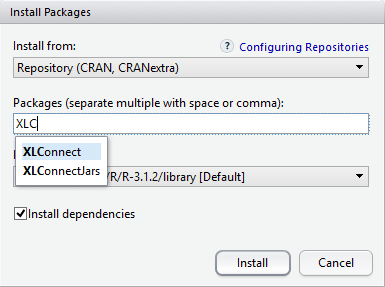
You will then be greeted with the project screen in RStudio:



Before running an application for the first time, one must make sure that all associated packages in RStudio are updated to their latest versions. To do this, click on the “Packages” tab on the bottom right window and click “Update”.



Also, to run the project, you must install two external packages: Shiny & XLConnect. Thankfully, RStudio provides a helpful facility to locate, download and install packages. To start the process, click the Install button in the same window as before. Then, simply search for the required package and click Install:



Finally, to run the ShinyR application, click on the button above the source code window.

**Description of various parameters:**

Our parameters for this project included two spreadsheets, one a .csv file called “prevWorldCup” which describes the results of every world cup since 1930 to 2002.

While our other dataset, describes more information about the players in the tournament. It is called “WorldCupPlayerInfo” and is a .xlsx file. Initially the file had a lot of extra sheets which contained information, but since we didn’t need it we deleted them and only used the sheet we wanted.

To use these data sets we simply imported the XLConnect package in to RStudio and read in the data using the commands located in the global.R file.

Both can be found online at the links on the first page.

**Team member contribution:**

For our project we distributed the work out evenly between the team members so we had the following format:

**Simon Markham**:

* Graphs
* Documentation

**Ciarán Moyne:**

* Documentation
* Graphs

**Ciaran O’Hagan:**

* Documentation,
* Graphs

**Code:**