Consider the `game_goalie_stats.csv` data file posted on Blackboard. This file contains goalie stats from different teams.

1. In **Python**, answer the following:

   (a) (3 points) Using the pandas library, read the csv data file and create a data-frame called `goalie_stats`.

   ```
   import pandas as pd

   ## Reading the csv file
   goalie_stats = pd.read_csv('game_goalie_stats.csv')
   ```

   (b) (3 points) Report the average and the median of the average time on ice of each goalie. What can you conclude about the distribution of the average time on ice?

   ```
   ## the average of the average timeOnIce
   goalie_stats.groupby('player_id')['timeOnIce'].mean().mean()

   ## the median of the average timeOnIce
   goalie_stats.groupby('player_id')['timeOnIce'].mean().median()
   ```

   The mean is 3012.24 and the median is 3257.51. Because the median is greater than the mean, we conclude that the shape of the histogram is left-skewed.

   (c) (3 points) Create a histogram of the average time on ice of each goalie. Does this histogram agree with your conclusions from part (b)?

   ```
   ## Histogram
   goalie_stats.groupby('player_id')['timeOnIce'].mean().hist()
   ```

   Yes, the shape of the histogram agrees with our conclusion from part (b).

2. In **R**, answer the following:

   (a) (3 points) Using the `read.csv` function, read the csv data file and create a data-frame called `goalie_stats`.

   ```
   ## Reading csv file
   goalie_stats = read.csv(file = 'game_goalie_stats.csv')
   ```

   (b) (3 points) Report the number of games of each goalie.

   ```
   ## Loading plyr
   library(plyr)

   ## Counting the number of games of each player
   player_games = ddply(goalie_stats, .(player_id), summarise,
                        numb_of_games = length(unique(game_id)))
   ```

(c) (3 points) Report the IQR of the number of games of each goalie.

```
## Five number summary
summary(player_games$numb_of_games)
```

The IQR is 139.

The End.