Introductory Studies

```
## Loading required package: xml2
```

```
## Warning in split.default(x = seq_len(nrow(x)), f = f, drop = drop, ...):
## data length is not a multiple of split variable
```

We are looking at data from major European football leagues to highlight the importance of our research question. While one might posit that spectators are drawn to attend games because of the intrigue of an uncertain outcome, we can show that this basic analysis alone is not enough to draw any conclusions from. In order to gain a clearer picture of the factors which draw spectators to attend games, we will need more in-depth research.

The data that we have pulled for the European football leagues covers each season from 2008-09 through 2018-19. We have gathered the average attendance for all league games, the average number of points per team for each season (3 for win, 1 for draw, 0 for loss), the standard deviation of points accumulated by teams over the course of the season, the betting odds for the home team winning any given game, the betting odds for the favored team to win the game, the betting odds for the "underdog" team to win the game, and the difference between the odds for the favorite and those for the underdog. Our hypothesis is that attendance will be greater in years with greater uncertainty of outcome (small SD of points, higher number for OddsFavorite, lower number for OddsUnderdog, and lower number for difference in odds). Additionally, we hypothesize that fans will prefer home odds of greater 50% but less than 75% (uncertain outcome but favoring the home team).

One potential mitigating factor is that teams are rewarded for finishing in the top four of the league, meaning that fans will still be drawn in even if there is a team or two that dominate the league (and thus increasing the standard deviation). Additionally, we are looking at league-wide data and thus not looking at the incentive structures for individual fanbases.

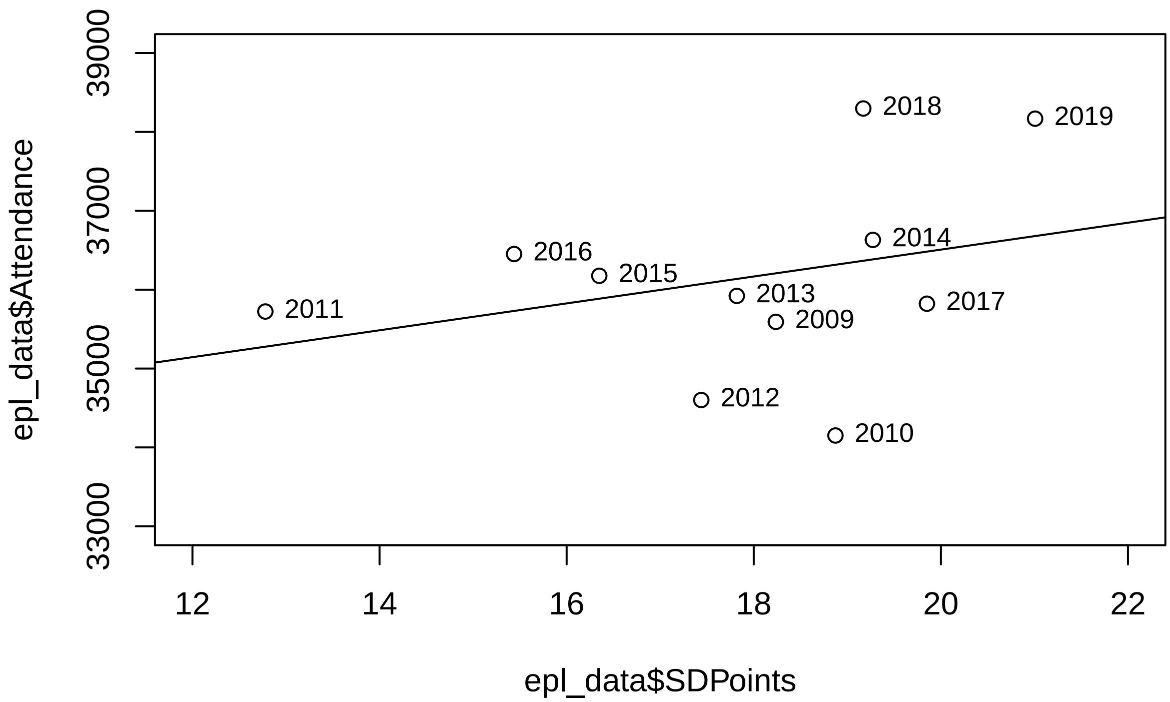
Our methods for the collection and analysis of data are explained below.

The first league we looked at was the English Premier League (EPL). The EPL has 20 teams, and the double round robin format means that each team plays 38 games per season.

The attendance data was pulled manually from worldfootball.net.

The points data was scraped from various sources. The functions `mean' and`sd' were used to generate the data in the dataset. The plot of attendance versus the standard deviation of points is shown below.

```r
fit <- lm(epl_data$Attendance ~ epl_data$SDPoints)
plot(epl_data$SDPoints, epl_data$Attendance, ylim = c(33000, 39000), xlim = c(12,22))
abline(fit)
text(epl_data$Attendance~epl_data$SDPoints, labels=(epl_data$Season), cex=.8, pos=4)
```



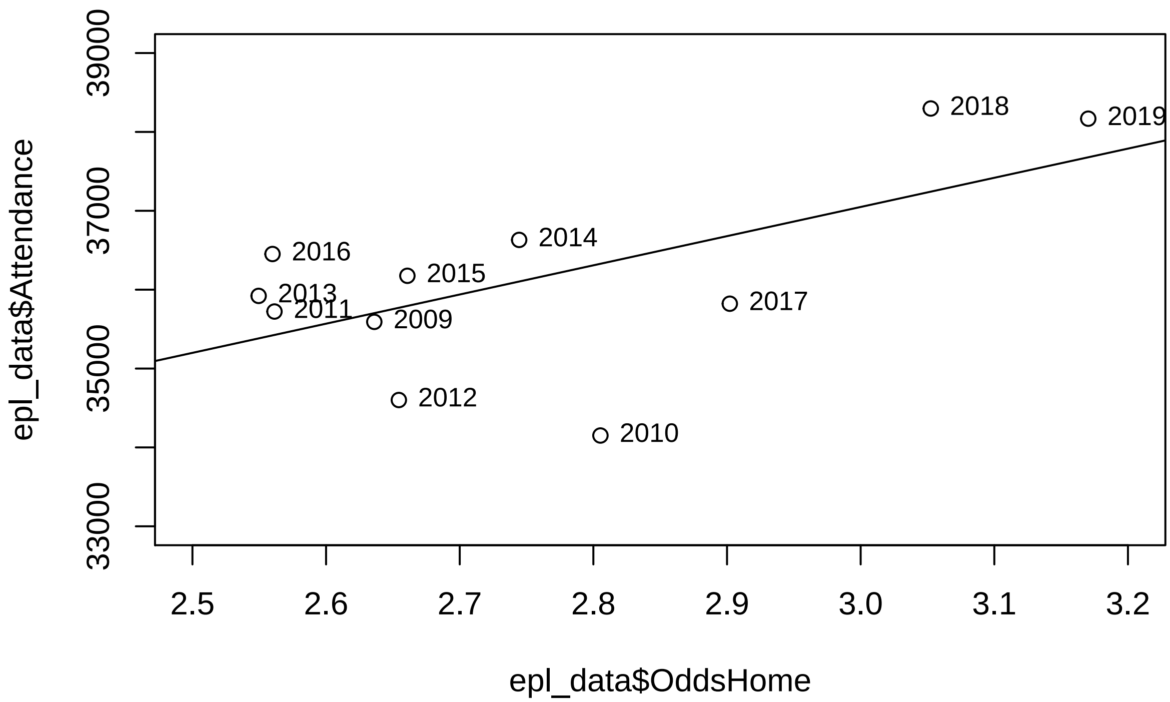

```r
cor(epl_data$Attendance, epl_data$SDPoints)
```

```
## [1] 0.3093882
```

The correlation coefficient of 0.309 indicates that there is a moderate positive correlation between the average attendance and the standard deviation of the points that teams accumulated. This runs contrary to our hypothesis.

The odds data was scraped from football-data.co.uk. The mean of the home odds, favorite odds, and underdog odds for each game was taken for each season. The plot of attendance versus home odds is shown below.

```r
fit <- lm(epl_data$Attendance ~ epl_data$OddsHome)
plot(epl_data$OddsHome, epl_data$Attendance, ylim = c(33000, 39000), xlim = c(2.5, 3.2))
abline(fit)
text(epl_data$Attendance~epl_data$OddsHome, labels=(epl_data$Season), cex=.8, pos=4)
```



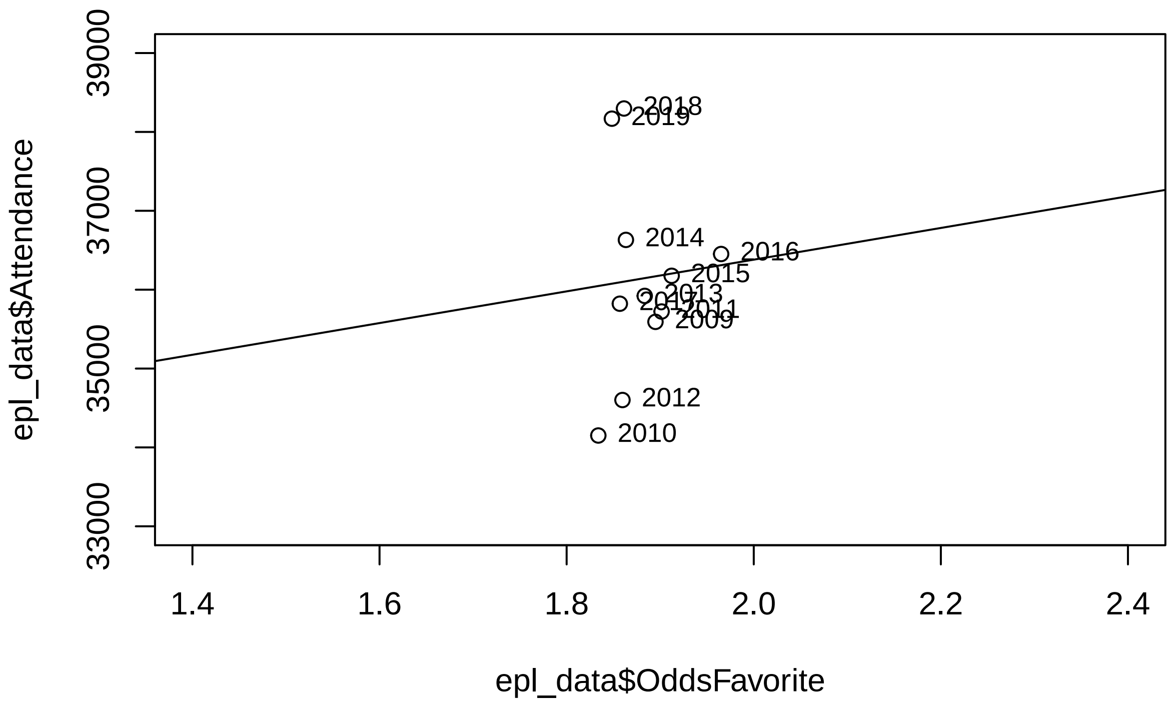

```
cor(epl_data$Attendance, epl_data$OddsHome)
```

```
## [1] 0.6097223
```

The correlation coefficient of 0.610 indicates that there is a strong positive correlation between the average attendance and the odds of the home team winning. In seasons which had more dominant teams and less parity in the league, attendance was generally higher. This also is not in line with our hypothesis.

Next is the plot of favorite odds against attendance. Our hypothesis would expect higher odds to correlate with higher attendance but the data from the prior analyses would indicate the opposite should happen.

```
fit <- lm(epl_data$Attendance ~ epl_data$OddsFavorite)
plot(epl_data$OddsFavorite, epl_data$Attendance, ylim = c(33000, 39000), xlim = c(1.4, 2.4))
abline(fit)
text(epl_data$Attendance~epl_data$OddsFavorite, labels=(epl_data$Season), cex=.8, pos=4)
```



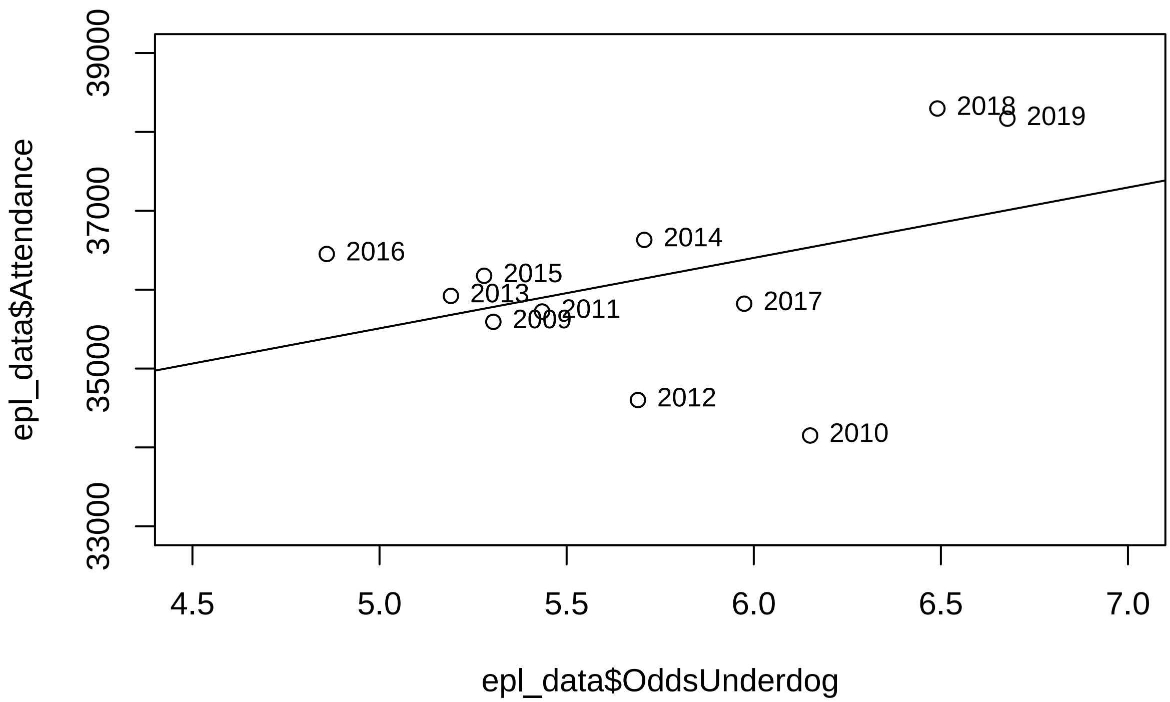

```r
cor(epl_data$Attendance, epl_data$OddsFavorite)
```

```
## [1] 0.05869502
```

The correlation coefficient of 0.059 shows that there is an essentially negligible positive correlation between the favorite being less likely to win and attendance.

Here we have the plot of underdog odds against attendance. Our hypothesis predicts that at lower values of underdog odds, there will be greater attendance numbers. However the prior analyses suggest the opposite relationship.

```r
fit <- lm(epl_data$Attendance ~ epl_data$OddsUnderdog)
plot(epl_data$OddsUnderdog, epl_data$Attendance, ylim = c(33000, 39000), xlim = c(4.5, 7))
abline(fit)
text(epl_data$Attendance~epl_data$OddsUnderdog, labels=(epl_data$Season), cex=.8, pos=4)
```



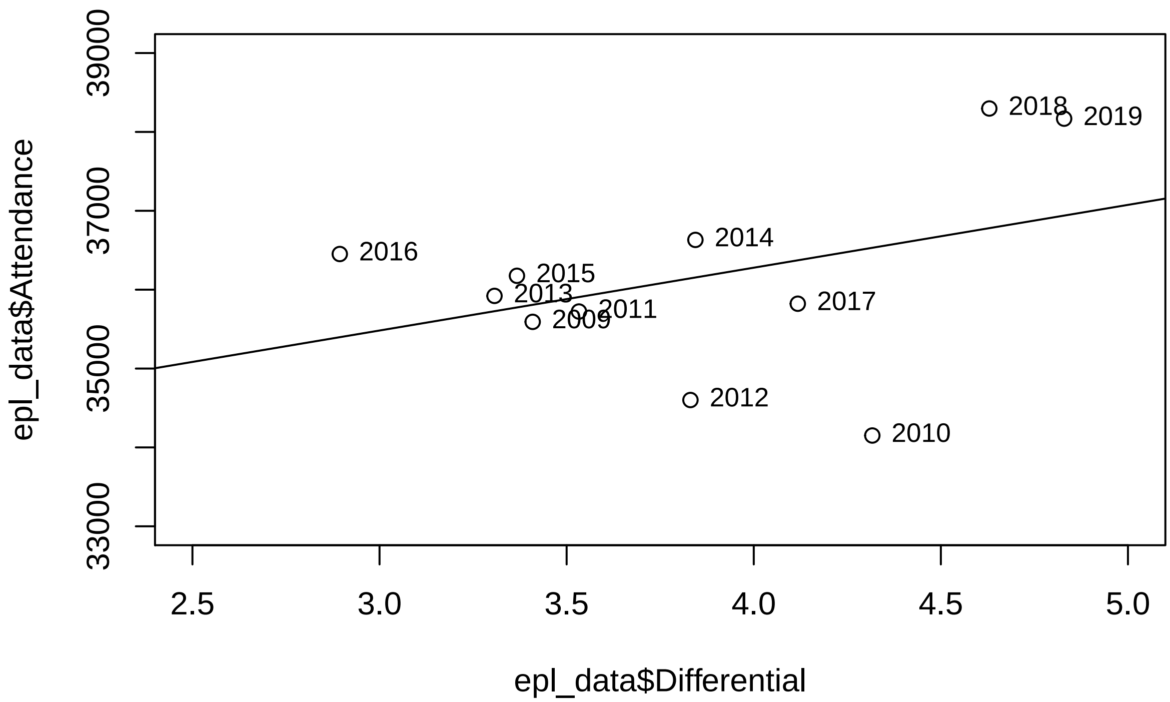

```
cor(epl_data$Attendance, epl_data$OddsUnderdog)
```

```
## [1] 0.400582
```

The correlation coefficient is 0.401, supporting our earlier analyses and providing more evidence against our hypothesis.

Our last plot for the EPL data uses the differential in favorite and underdog odds against attendance.

```
fit <- lm(epl_data$Attendance ~ epl_data$Differential)
plot(epl_data$Differential, epl_data$Attendance, ylim = c(33000, 39000), xlim = c(2.5, 5))
abline(fit)
text(epl_data$Attendance~epl_data$Differential, labels=(epl_data$Season), cex=.8, pos=4)
```




```r
cor(epl_data$Attendance, epl_data$Differential)
```

```
## [1] 0.3765336
```

The correlation coefficient of 0.377 suggests that there is a moderate positive relationship between the difference in favorite vs. underdog odds, therefore showing that with less uncertainty of outcome the attendance is higher. The random scatter of the points suggests that there is not likely a temporal relationship between the variables.

Ultimately what we find is that attendance numbers are greater in years with greater disparity between the skill levels of the teams. This is a surprising result that indicates that fans don't value uncertainty of outcome in the EPL, and may prefer to see higher levels of skill amongst the best teams in the league.

The full dataset is included below.

```r
epl_data
```

```
##    Season Attendance Points SDPoints OddsHome OddsFavorite OddsUnderdog
## 1    2009      35592  52.15 18.23610 2.636079     1.894947     5.304184
## 2    2010      34151  51.75 18.87319 2.805289     1.833868     6.150737
## 3    2011      35723  51.45 12.77940 2.561395     1.901474     5.434368
## 4    2012      34601  52.35 17.43944 2.654447     1.859684     5.690500
## 5    2013      35921  51.60 17.81897 2.549526     1.883474     5.190684
## 6    2014      36631  53.10 19.27338 2.744474     1.863342     5.707421
## 7    2015      36176  52.35 16.34907 2.660842     1.912263     5.279474
## 8    2016      36452  51.65 15.43842 2.559921     1.965132     4.858895
## 9    2017      35822  52.80 19.85102 2.902079     1.856842     5.974526
## 10   2018      38297  52.05 19.17091 3.052447     1.861289     6.490789
## 11   2019      38168  53.45 21.00745 3.170289     1.848368     6.677789
##    Differential
## 1      3.409237
## 2      4.316868
## 3      3.532895
## 4      3.830816
```

```
## 5      3.307211
## 6      3.844079
## 7      3.367211
## 8      2.893763
## 9      4.117684
## 10     4.629500
## 11     4.829421
```

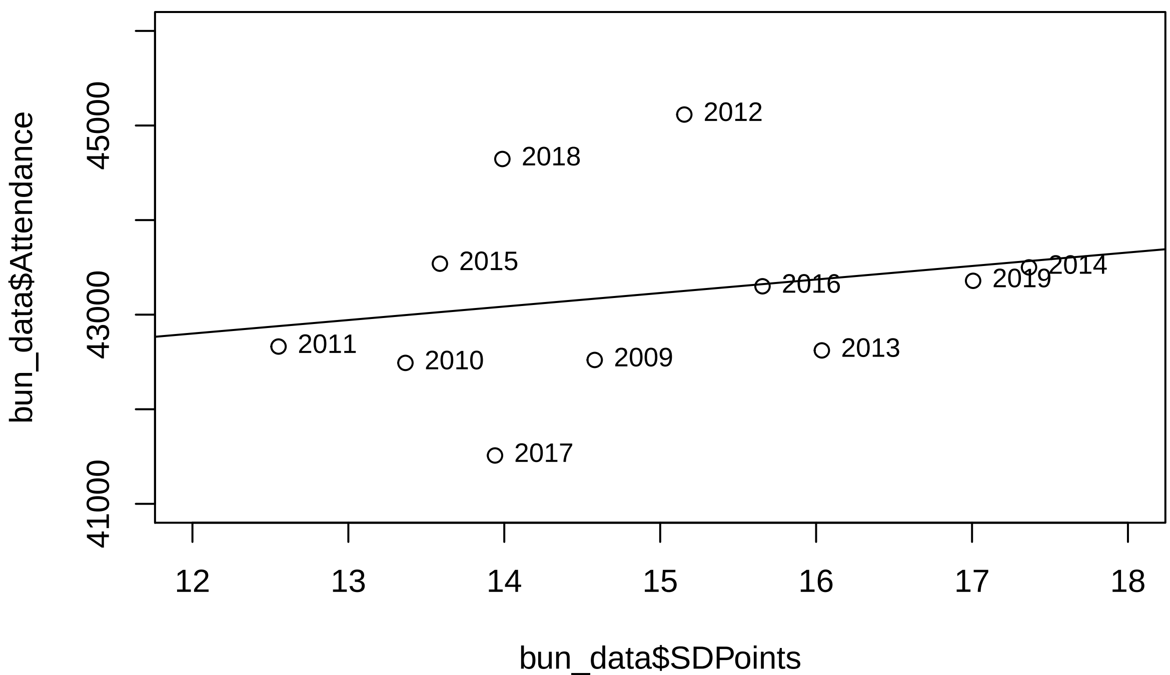
The same methods were repeated for the Bundesliga (Germany).

The Bundesliga has 18 teams, meaning that each team plays 34 games per season. This will likely contribute to a lower standard deviation in points than in the EPL, but that is irrelevant as they are not being compared against one another.

The attendance data was pulled manually from worldfootball.net.

The points data was scraped from kicker.de. The functions `mean' and`sd' were used to generate the data in the dataset. The plot of attendance versus the standard deviation of points is shown below.

```
fit <- lm(bun_data$Attendance ~ bun_data$SDPoints)
plot(bun_data$SDPoints, bun_data$Attendance, ylim = c(41000, 46000), xlim = c(12, 18))
abline(fit)
text(bun_data$Attendance~bun_data$SDPoints, labels=(bun_data$Season), cex=.8, pos=4)
```




```
cor(bun_data$Attendance, bun_data$SDPoints)
```

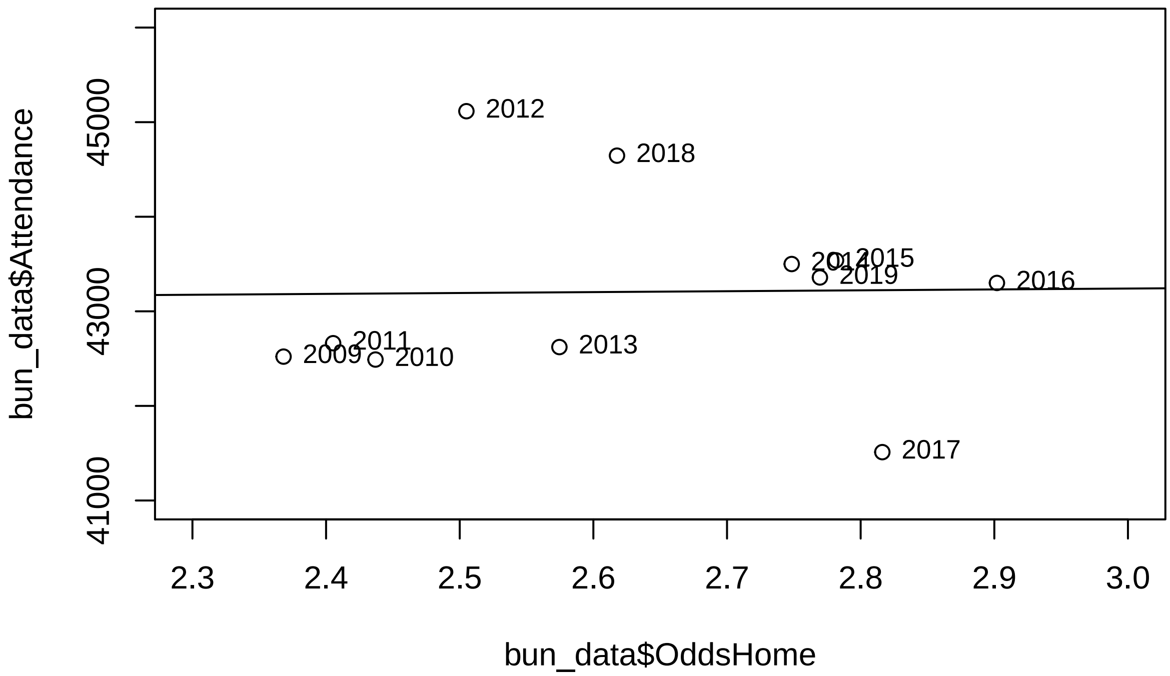
```
## [1] 0.2158874
```

The correlation coefficient of 0.216 indicates that there is a weak positive correlation between the average attendance and the standard deviation of the points that teams accumulated. This information is in line with the EPL data and runs contrary to our hypothesis of uncertainty of outcome drawing fans to attend.

The odds data was scraped from football-data.co.uk. The mean of the home odds, favorite odds, and underdog odds for each game was taken for each season. The plot of attendance versus home odds is shown below.

```
fit <- lm(bun_data$Attendance ~ bun_data$OddsHome)
plot(bun_data$OddsHome, bun_data$Attendance, ylim = c(41000, 46000), xlim = c(2.3, 3.0))
```

```r
abline(fit)
text(bun_data$Attendance~bun_data$OddsHome, labels=(bun_data$Season), cex=.8, pos=4)
```



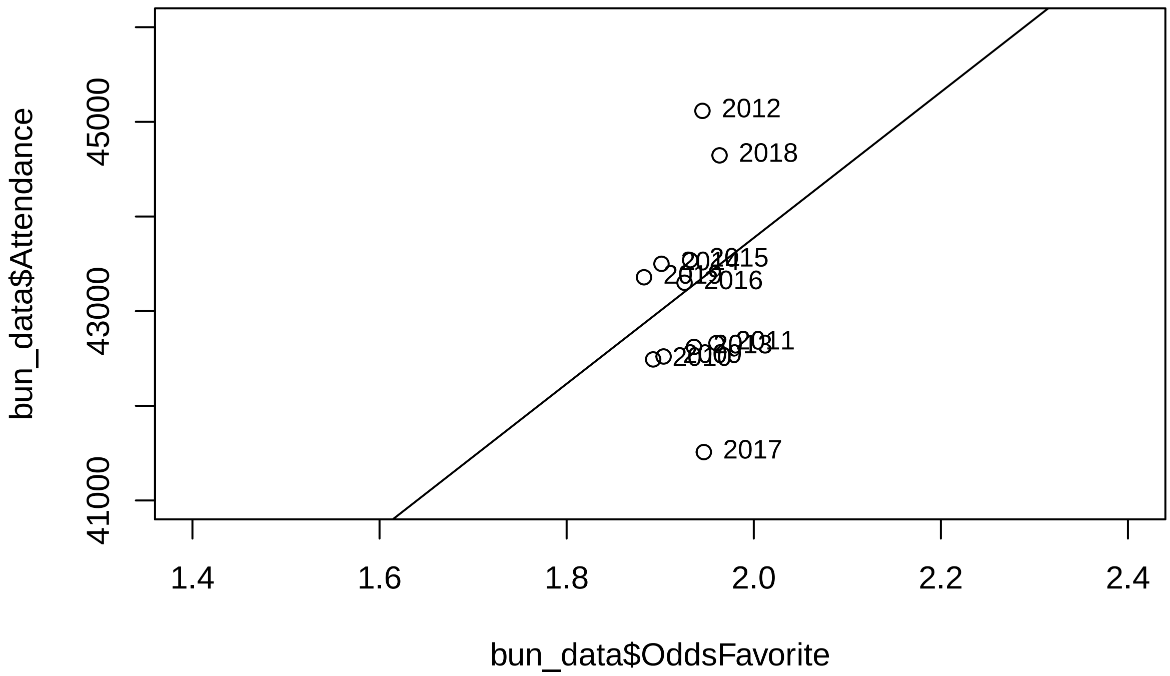

```r
cor(bun_data$Attendance, bun_data$OddsHome)
```

```
## [1] 0.01691811
```

The correlation coefficient of 0.017 indicates that there is essentially no correlation between the average attendance and the odds of the home team winning. Contrary to the result from the EPL, we see that for the Bundesliga the attendance is not affected by the odds of the home team winning.

Next is the plot of favorite odds against attendance. Our hypothesis would expect higher odds to correlate with higher attendance.

```r
fit <- lm(bun_data$Attendance ~ bun_data$OddsFavorite)
plot(bun_data$OddsFavorite, bun_data$Attendance, ylim = c(41000, 46000), xlim = c(1.4, 2.4))
abline(fit)
text(bun_data$Attendance~bun_data$OddsFavorite, labels=(bun_data$Season), cex=.8, pos=4)
```



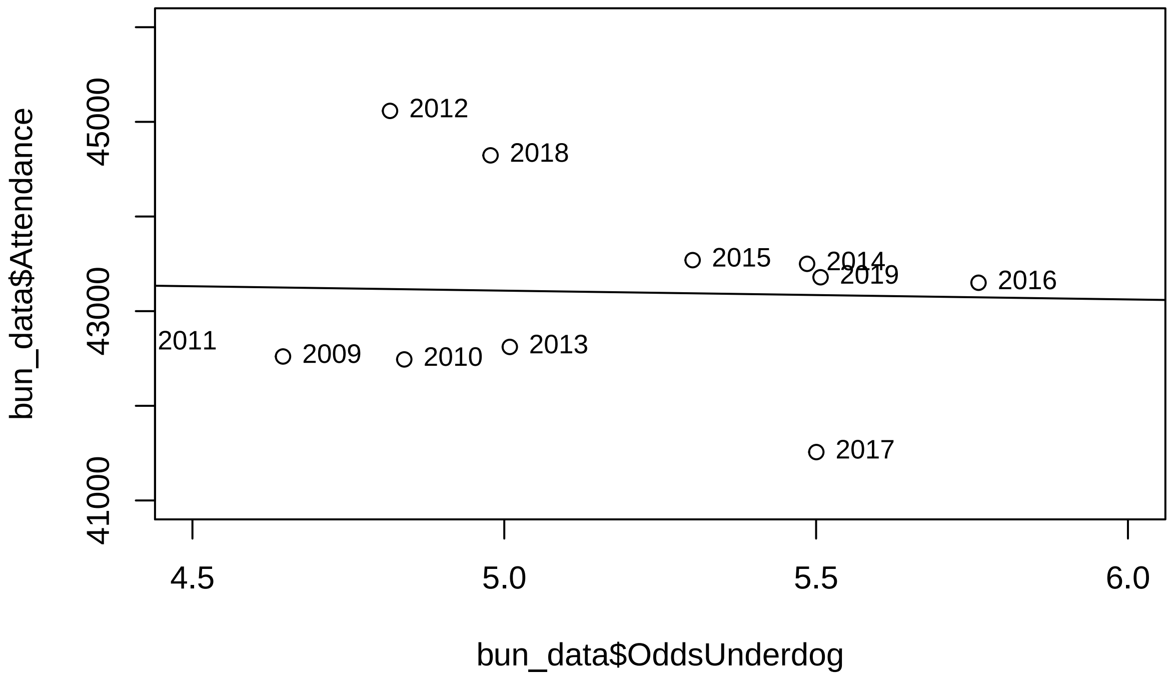

```
cor(bun_data$Attendance, bun_data$OddsFavorite)
```

```
## [1] 0.2080014
```

The correlation coefficient of 0.208 shows that there is a weak positive correlation between the favorite being less likely to win and attendance. This supports our hypothesis of uncertainty of outcome influencing spectators to attend, and shows a relationship opposite of what we found with "SDPoints" against attendance. This is the first result that we have found which supports our hypothesis.

Here we have the plot of underdog odds against attendance. Our hypothesis predicts that at lower values of underdog odds, there will be greater attendance numbers.

```
fit <- lm(bun_data$Attendance ~ bun_data$OddsUnderdog)
plot(bun_data$OddsUnderdog, bun_data$Attendance, ylim = c(41000, 46000), xlim = c(4.5, 6.0))
abline(fit)
text(bun_data$Attendance~bun_data$OddsUnderdog, labels=(bun_data$Season), cex=.8, pos=4)
```



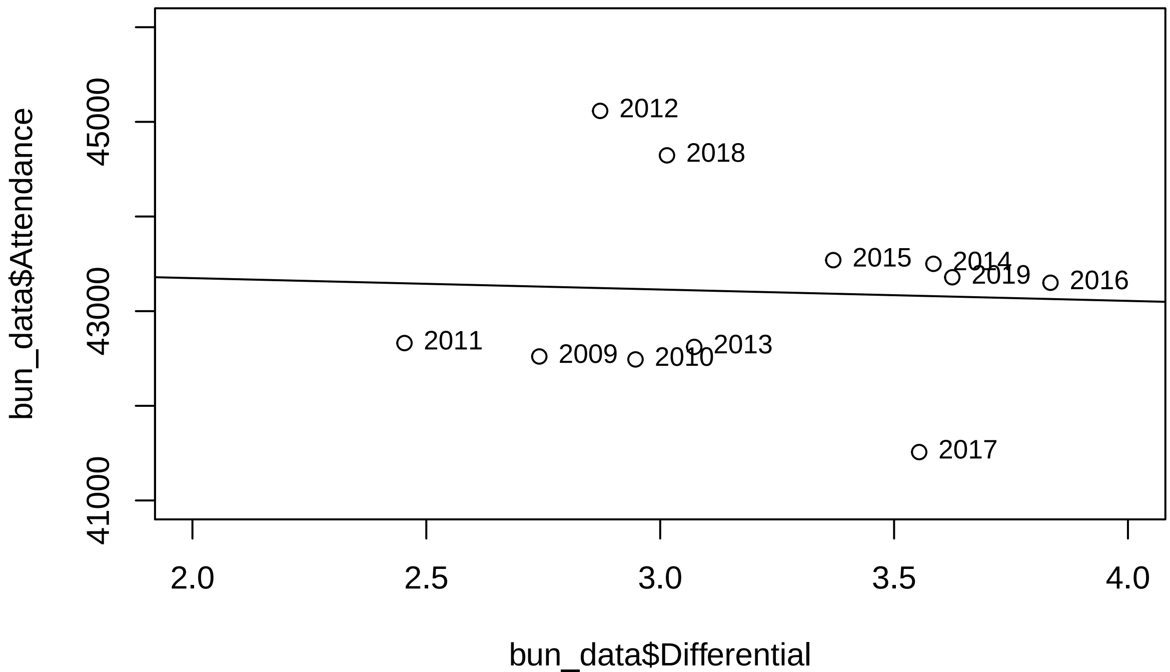

```
cor(bun_data$Attendance, bun_data$OddsUnderdog)
```

```
## [1] -0.03846616
```

The correlation coefficient is -0.038, which is an essentially negligible negative relationship between the decrease in odds of the underdog and the attendance. The direction of this relationship supports our hypothesis but is very minor in degree.

Our last plot for the Bundesliga data uses the differential in favorite and underdog odds against attendance.

```
fit <- lm(bun_data$Attendance ~ bun_data$Differential)
plot(bun_data$Differential, bun_data$Attendance, ylim = c(41000, 46000), xlim = c(2, 4))
abline(fit)
text(bun_data$Attendance~bun_data$Differential, labels=(bun_data$Season), cex=.8, pos=4)
```




```
cor(bun_data$Attendance, bun_data$Differential)
```

```
## [1] -0.05095716
```

The correlation coefficient of -0.051 suggests that there is a very weak negative relationship between the difference in favorite vs. underdog odds, showing that there is an almost negligible negative relationship between uncertainty of outcome and attendance. The random scatter of the points suggests that there is not likely a temporal relationship between the variables.

Ultimately we find contradictory results across the studies as to whether or not fans respond positively to uncertainty of outcome. While we can't form any definitive conclusions, it appears that attendance in the Bundesliga is not changed by parity in the league across seasons, at least not to the extent that it is in the EPL.

The full dataset is included below.

```
bun_data
```

```
##    Season Attendance   Points SDPoints OddsHome OddsFavorite OddsUnderdog
## 1    2009      42521 46.88889 14.57996 2.368170     1.903627     4.645229
## 2    2010      42490 46.22222 13.36614 2.436961     1.892484     4.839673
## 3    2011      42663 47.50000 12.55224 2.405294     1.960163     4.413497
## 4    2012      45116 46.61111 15.15465 2.505000     1.945163     4.816797
## 5    2013      42622 46.66667 16.03672 2.574575     1.936111     5.008856
## 6    2014      43500 47.44444 17.36536 2.748399     1.901373     5.485523
## 7    2015      43539 46.44444 13.58729 2.781569     1.932026     5.302026
## 8    2016      43300 47.05556 15.65613 2.901961     1.926013     5.760392
## 9    2017      41511 46.88889 13.94058 2.816209     1.946634     5.500327
## 10   2018      44646 46.38889 13.98797 2.617680     1.963464     4.977974
## 11   2019      43358 46.94444 17.00682 2.769510     1.882680     5.507124
##    Differential
## 1      2.741601
## 2      2.947190
## 3      2.453333
## 4      2.871634
```

```
## 5       3.072745
## 6       3.584150
## 7       3.370000
## 8       3.834379
## 9       3.553693
## 10      3.014510
## 11      3.624444
```

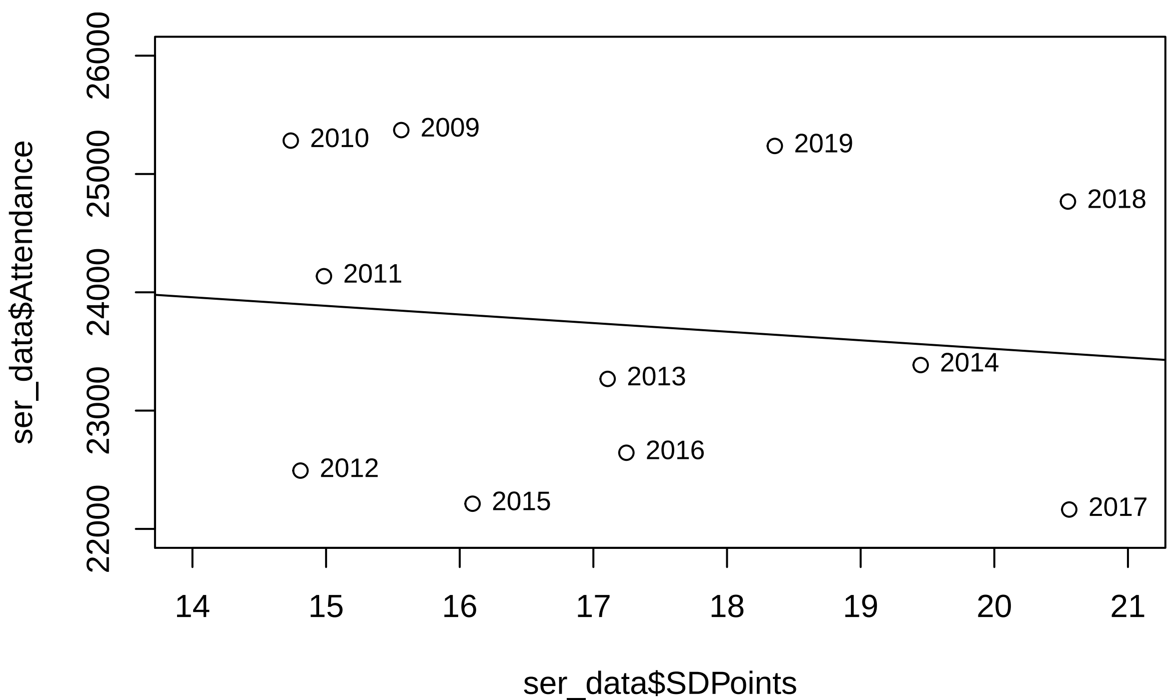
The same methods were repeated for Serie A (Italy).

Serie A has 20 teams, meaning that each team plays 38 games per season. This will likely contribute to a similar standard deviation in points to the EPL, but higher than in the Bundesliga. Again, this discrepancy will not have any effect on our results but may point to slightly less power in this study than in that of the EPL.

The attendance data was pulled manually from worldfootball.net.

The points data was scraped from espn.co.uk. The functions `mean' and`sd' were used to generate the data in the dataset. The plot of attendance versus the standard deviation of points is shown below.

```
fit <- lm(ser_data$Attendance ~ ser_data$SDPoints)
plot(ser_data$SDPoints, ser_data$Attendance, ylim = c(22000, 26000), xlim = c(14,21))
abline(fit)
text(ser_data$Attendance~ser_data$SDPoints, labels=(ser_data$Season), cex=.8, pos=4)
```



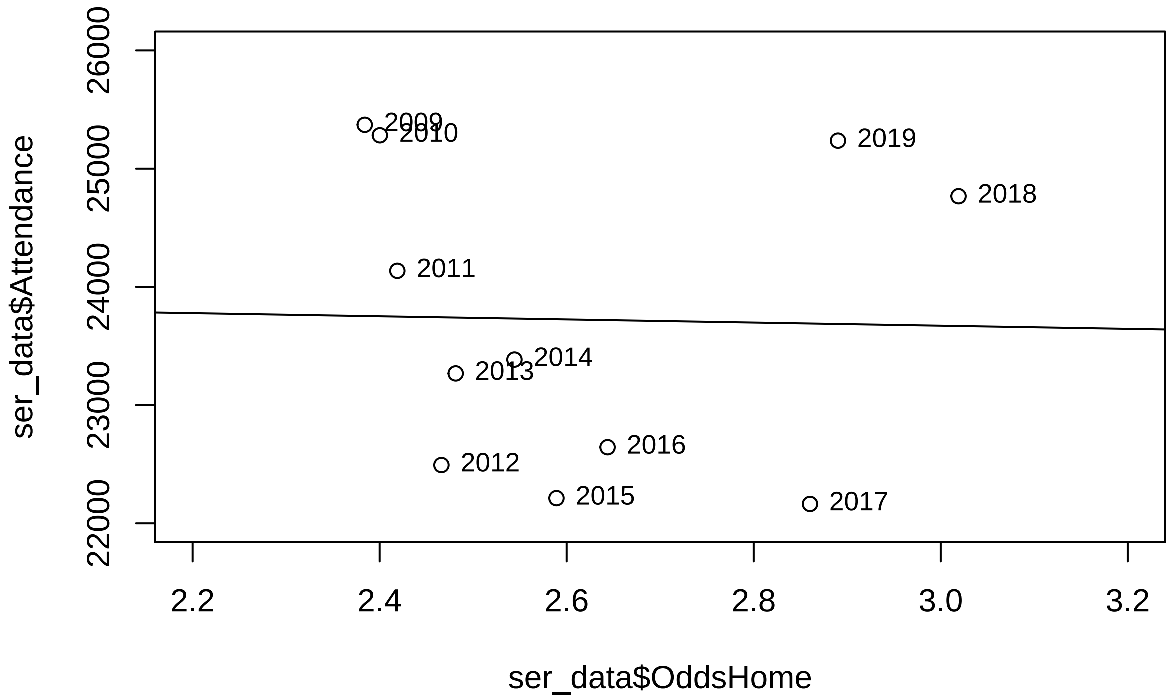

```
cor(ser_data$Attendance, ser_data$SDPoints)
```

```
## [1] -0.1264008
```

The correlation coefficient of -0.126 indicates that there is a weak negative correlation between the average attendance and the standard deviation of the points that teams accumulated. This is the opposite of the result that we saw for the EPL and the Bundesliga and suggests that Serie A fans respond more to parity in the league, supporting our hypothesis.

The odds data was scraped from football-data.co.uk. The mean of the home odds, favorite odds, and underdog odds for each game was taken for each season. The plot of attendance versus home odds is shown below.

```r
fit <- lm(ser_data$Attendance ~ ser_data$OddsHome)
plot(ser_data$OddsHome, ser_data$Attendance, ylim = c(22000, 26000), xlim = c(2.2, 3.2))
abline(fit)
text(ser_data$Attendance~ser_data$OddsHome, labels=(ser_data$Season), cex=.8, pos=4)
```



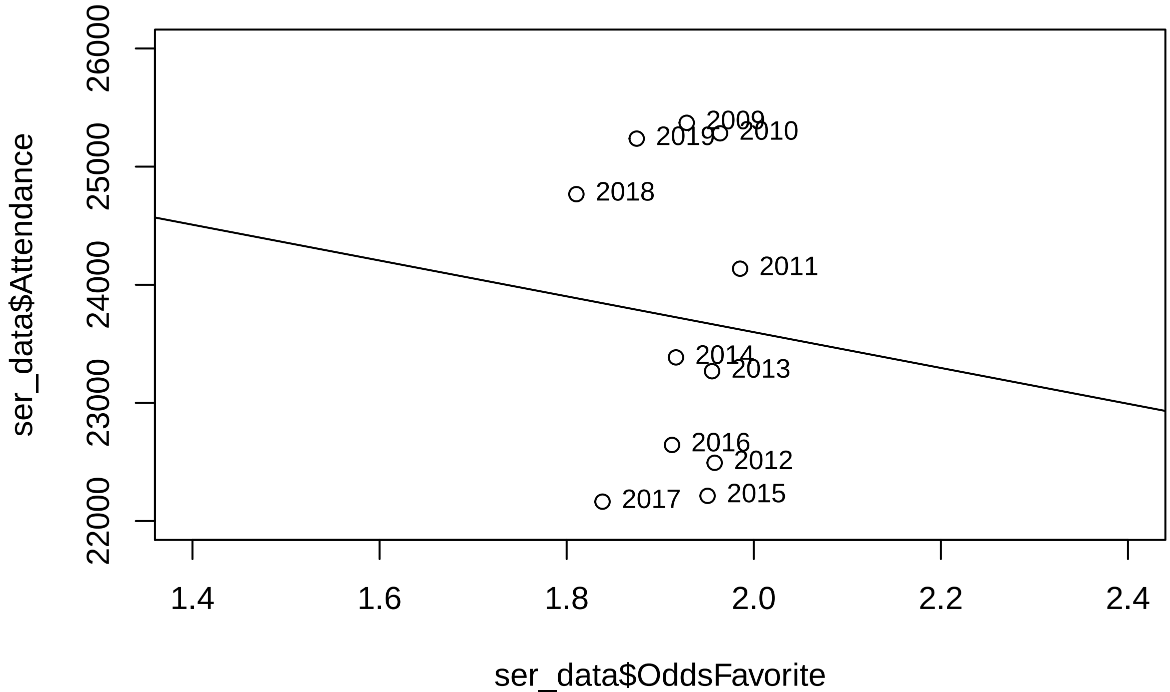

```r
cor(ser_data$Attendance, ser_data$OddsHome)
```

```
## [1] -0.02280357
```

The correlation coefficient of -0.023 indicates that there is essentially no correlation between the average attendance and the odds of the home team winning. Contrary to the result from the EPL but similar to the Bundesliga, we see that for the Bundesliga the attendance is not affected by the odds of the home team winning.

Next is the plot of favorite odds against attendance. Our hypothesis would expect higher odds to correlate with higher attendance but the data from the prior analyses would indicate the opposite should happen.

```r
fit <- lm(ser_data$Attendance ~ ser_data$OddsFavorite)
plot(ser_data$OddsFavorite, ser_data$Attendance, ylim = c(22000, 26000), xlim = c(1.4, 2.4))
abline(fit)
text(ser_data$Attendance~ser_data$OddsFavorite, labels=(ser_data$Season), cex=.8, pos=4)
```



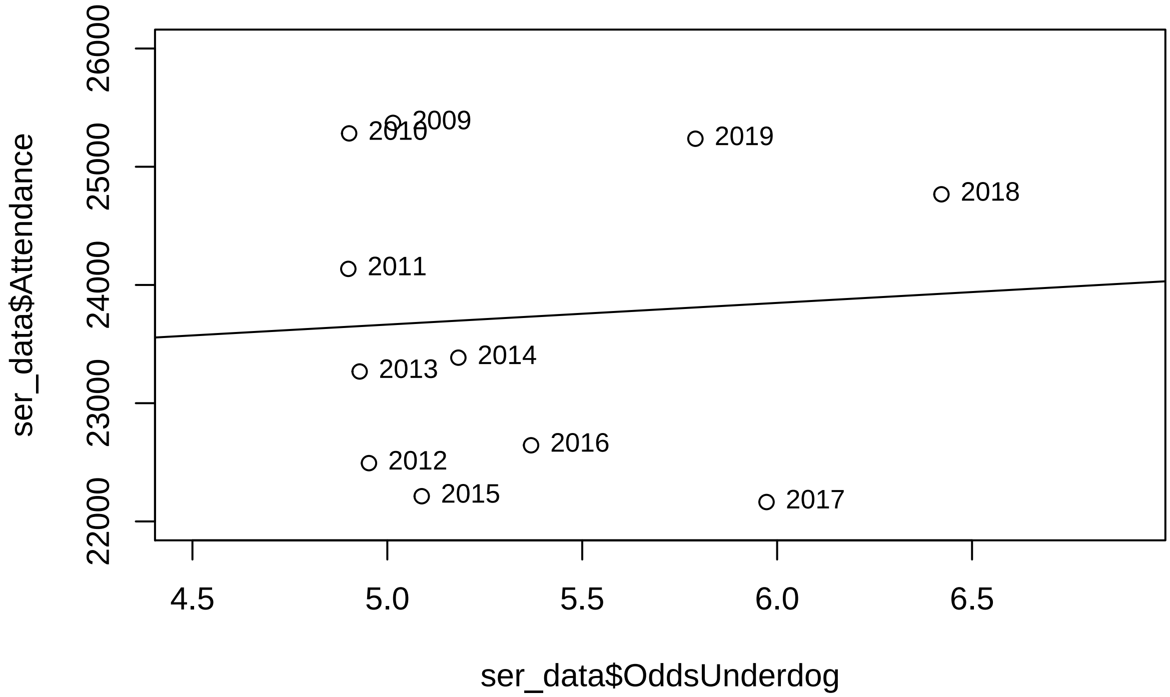

```
cor(ser_data$Attendance, ser_data$OddsFavorite)
```

```
## [1] -0.0655219
```

The correlation coefficient of -0.066 shows that there is an essentially negligible negative correlation between the favorite being less likely to win and attendance, although the negative direction of the relationship is in line with our hypothesis.

Here we have the plot of underdog odds against attendance. Our hypothesis predicts that at lower values of underdog odds, there will be greater attendance numbers. However the prior analyses suggest the opposite relationship.

```
fit <- lm(ser_data$Attendance ~ ser_data$OddsUnderdog)
plot(ser_data$OddsUnderdog, ser_data$Attendance, ylim = c(22000, 26000), xlim = c(4.5, 6.9))
abline(fit)
text(ser_data$Attendance~ser_data$OddsUnderdog, labels=(ser_data$Season), cex=.8, pos=4)
```



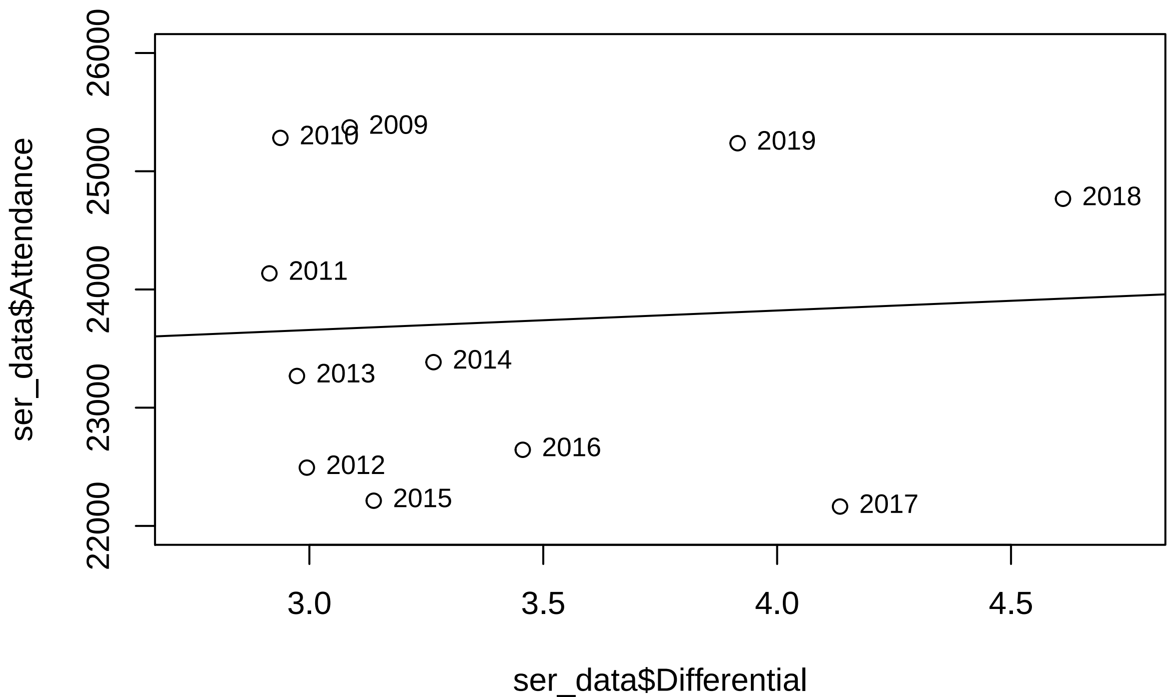

```
cor(ser_data$Attendance, ser_data$OddsUnderdog)
```

```
## [1] 0.07394614
```

The correlation coefficient is 0.074, which is also basically negligible, although the direction of the relationship provides some support of our earlier analyses and more evidence towards our hypothesis although to a very weak degree.

Our last plot for the Serie A data uses the differential in favorite and underdog odds against attendance.

```
fit <- lm(ser_data$Attendance ~ ser_data$Differential)
plot(ser_data$Differential, ser_data$Attendance, ylim = c(22000, 26000), xlim = c(2.75, 4.75))
abline(fit)
text(ser_data$Attendance~ser_data$Differential, labels=(ser_data$Season), cex=.8, pos=4)
```




```r
cor(ser_data$Attendance, ser_data$Differential)
```

```
## [1] 0.07329762
```

The correlation coefficient of 0.073 suggests that there is a minor positive relationship between the difference in favorite vs. underdog odds, therefore showing that with less uncertainty of outcome the attendance is higher. There may be some sort of temporal relationship between the variables given that the past three seasons have seen by far the largest differentials in odds level.

Ultimately what we find is that the relationship between attendance numbers and the disparity between the skill levels of the teams changes when we use different factors.

The full dataset is included below.

```r
ser_data
```

```
##    Season Attendance Points SDPoints OddsHome OddsFavorite OddsUnderdog
## 1    2009      25371  52.25 15.56269 2.384063     1.928338     5.014485
## 2    2010      25282  51.90 14.73592 2.400237     1.963974     4.902053
## 3    2011      24136  52.00 14.98420 2.418947     1.985395     4.899947
## 4    2012      22493  51.15 14.80851 2.466079     1.958211     4.952868
## 5    2013      23268  51.70 17.10679 2.481342     1.955447     4.928947
## 6    2014      23385  52.50 19.44899 2.544158     1.916842     5.182237
## 7    2015      22213  50.65 16.09601 2.589105     1.950658     5.088237
## 8    2016      22644  52.25 17.24704 2.643737     1.912632     5.368763
## 9    2017      22164  53.00 20.56057 2.860211     1.838342     5.972789
## 10   2018      24767  52.85 20.55103 3.019053     1.810421     6.421526
## 11   2019      25237  51.45 18.35749 2.890132     1.874895     5.790342
##    Differential
## 1      3.086148
## 2      2.938079
## 3      2.914553
## 4      2.994658
## 5      2.973500
```

```
## 6      3.265395
## 7      3.137579
## 8      3.456132
## 9      4.134447
## 10     4.611105
## 11     3.915447
```

As we can see from the results across the leagues, we aren't able to learn a lot from these basic studies. While the EPL shows a relatively strong relationship between attendance and less uncertainty of outcome, the Bundesliga and Serie A results are inconclusive. This suggests that either the fans in the various leagues respond to different incentives or that there are confounding variables which we need to explore. We hope that exploring the more in-depth data in the AFL datasets will help us to better learn about the behavioral patterns of sports spectators.

We ran a similar basic analysis with AFL data.

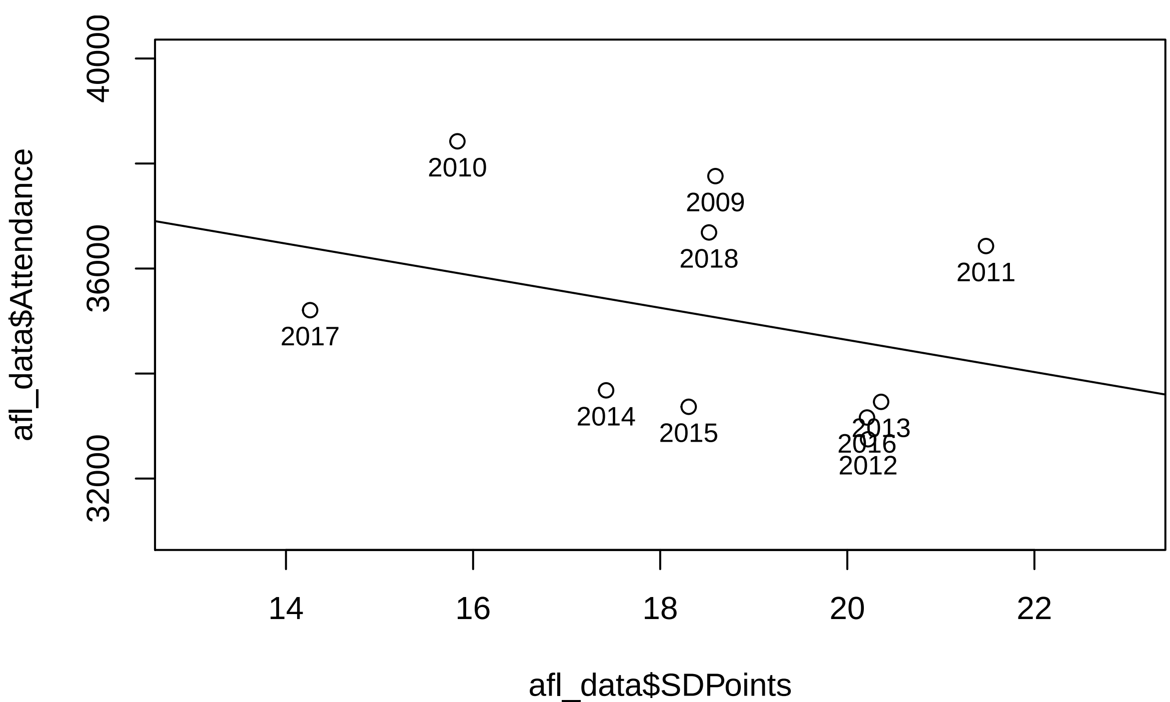
```
## Warning: NAs introduced by coercion
```

The AFL has 18 teams, however the format of the league is different than that of the European football leagues and each team plays 22 games per season. The data encompasses the seasons between 2009 and 2018.

The attendance data was scraped from afltables.com.

The points data was scraped from wikipedia.org. The functions 'mean' and 'sd' were used to generate the data in the dataset. The plot of attendance versus the standard deviation of points is shown below.

```
fit <- lm(afl_data$Attendance~afl_data$SDPoints)
plot(afl_data$SDPoints, afl_data$Attendance, ylim = c(31000, 40000), xlim = c(13,23))
abline(fit)
text(afl_data$Attendance~afl_data$SDPoints, labels=(afl_data$Season),data=afl_data, cex=.8, pos=1)
```



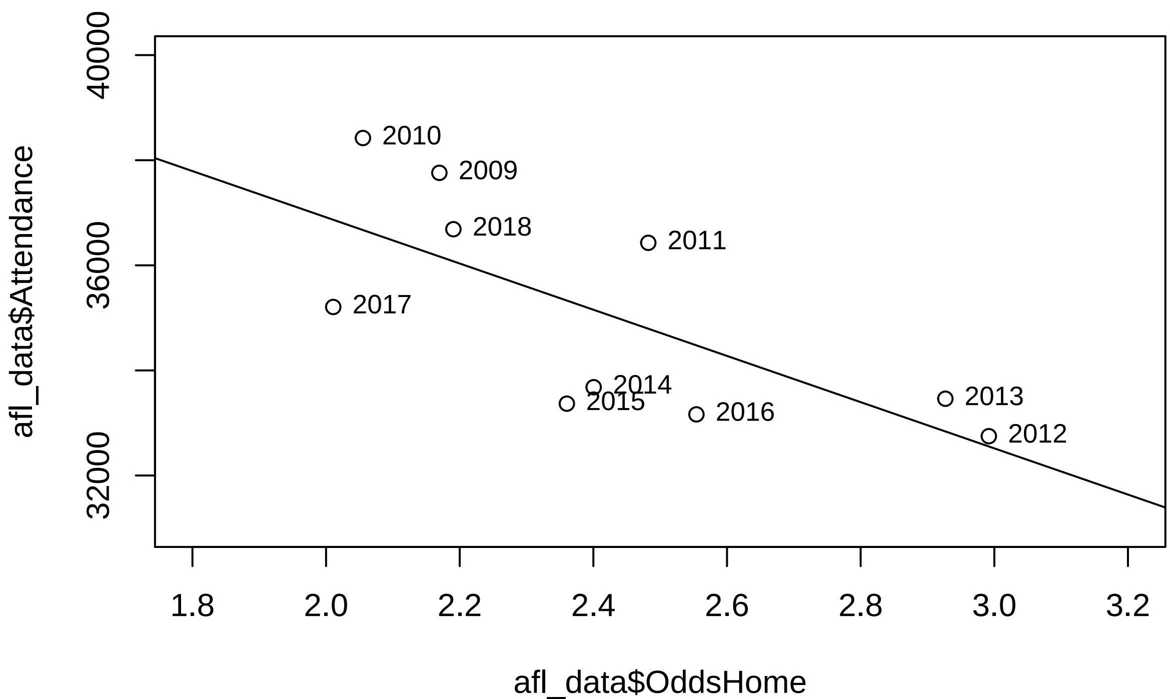

```
cor(afl_data$Attendance, afl_data$SDPoints)
```

```
## [1] -0.3251895
```

The correlation coefficient of -0.325 indicates that there is a moderate negative correlation between the average attendance and the standard deviation of the points that teams accumulated. This points to AFL fans valuing parity more than in the European football leagues.

The odds data was inputted manually from a spreadsheet database. The mean of the home odds for each game was taken for each season. The plot of attendance versus home odds is shown below.

```
fit <- lm(afl_data$Attendance~afl_data$OddsHome)
plot(afl_data$OddsHome, afl_data$Attendance, ylim = c(31000, 40000), xlim = c(1.8,3.2))
abline(fit)
text(afl_data$Attendance~afl_data$OddsHome, labels=(afl_data$Season),data=afl_data, cex=.8, pos=4)
```

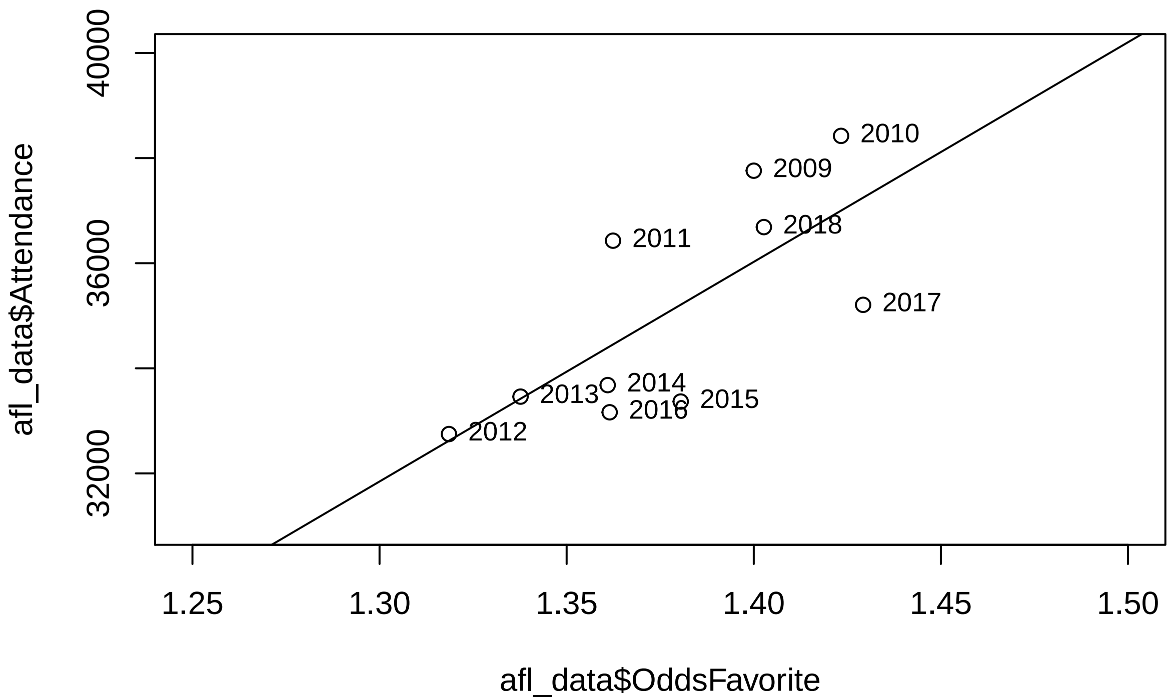



```
cor(afl_data$Attendance, afl_data$OddsHome)
```

```
## [1] -0.7093574
```

The correlation coefficient of -0.709 indicates that there is a strong negative correlation between the average attendance and the odds of the home team winning. Exactly the opposite of the EPL, the AFL gets better attendance when the home team has lower odds of winning. This also points to AFL fans' tendency to value parity.

```
fit <- lm(afl_data$Attendance~afl_data$OddsFavorite)
plot(afl_data$OddsFavorite, afl_data$Attendance, ylim = c(31000, 40000), xlim = c(1.25,1.5))
abline(fit)
text(afl_data$Attendance~afl_data$OddsFavorite, labels=(afl_data$Season),data=afl_data, cex=.8, pos=4)
```

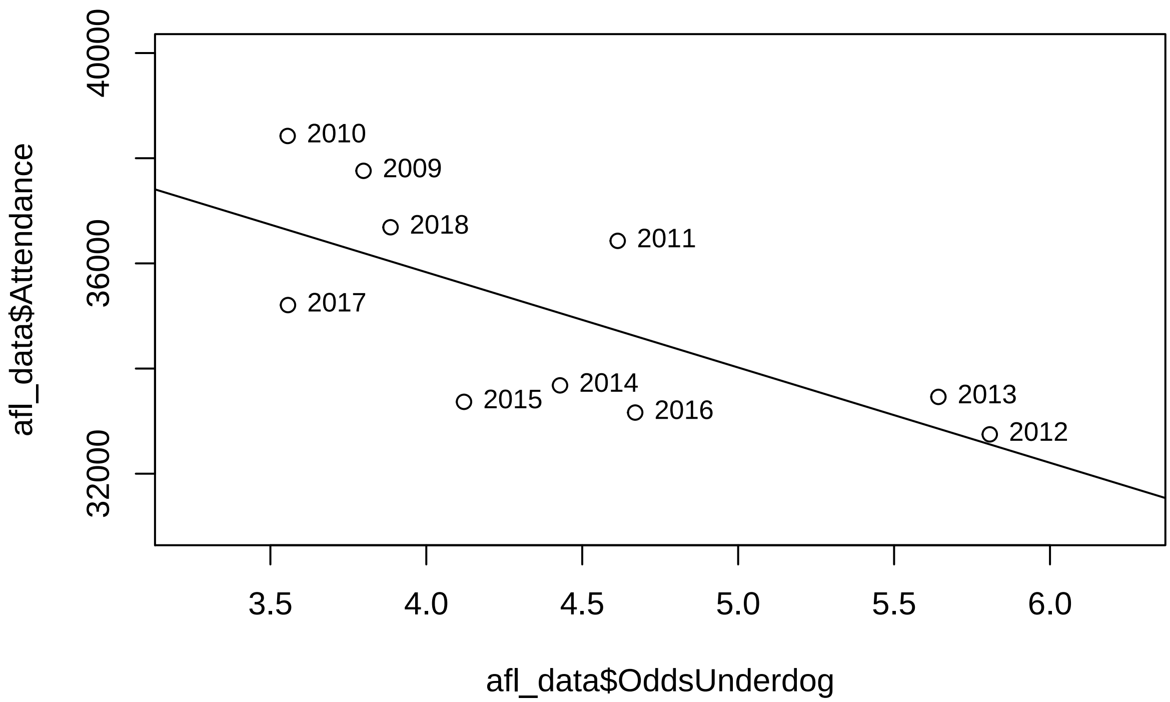



```
cor(afl_data$Attendance, afl_data$OddsFavorite)
```

```
## [1] 0.7215841
```

There is a strong positive correlation (0.722) between a decrease in the favored team's chances of winning and average attendance. This supports our hypothesis that fans respond to uncertainty of outcome.

```
fit <- lm(afl_data$Attendance~afl_data$OddsUnderdog)
plot(afl_data$OddsUnderdog, afl_data$Attendance, ylim = c(31000, 40000), xlim = c(3.25,6.25))
abline(fit)
text(afl_data$Attendance~afl_data$OddsUnderdog, labels=(afl_data$Season),data=afl_data, cex=.8, pos=4)
```

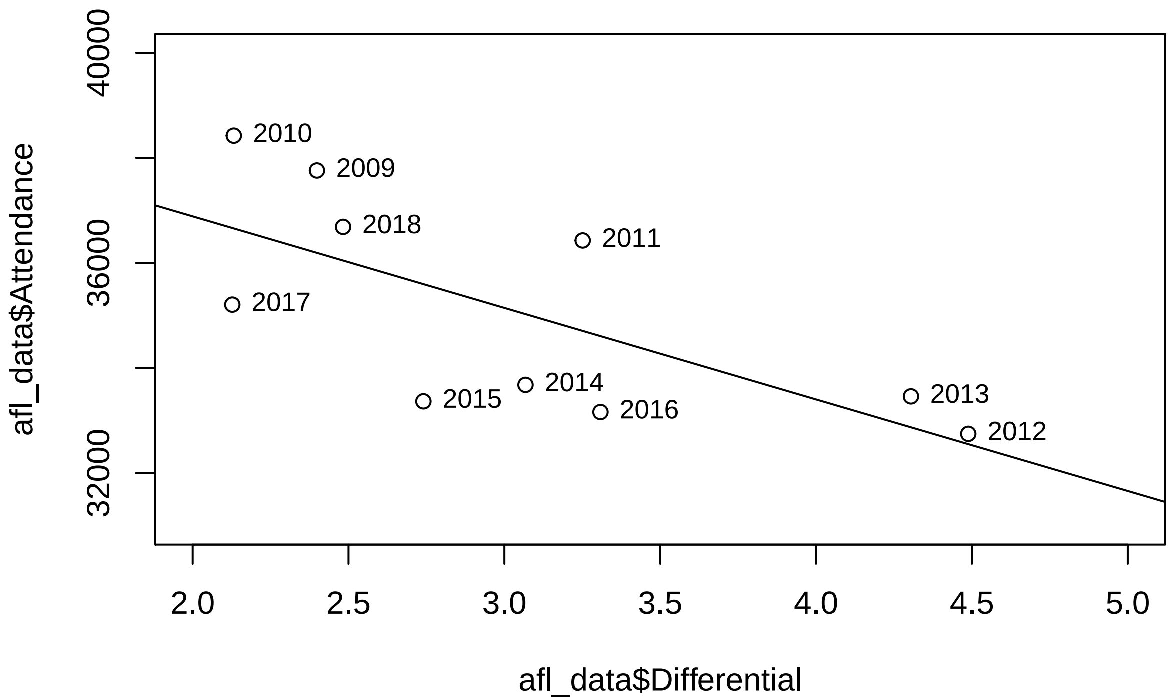



```
cor(afl_data$Attendance, afl_data$OddsUnderdog)
```

```
## [1] -0.6932748
```

The correlation coefficient of -0.693 shows a strong positive relationship between attendance and uncertainty of outcome. This further supports our hypothesis.

```
fit <- lm(afl_data$Attendance~afl_data$Differential)
plot(afl_data$Differential, afl_data$Attendance, ylim = c(31000, 40000), xlim = c(2,5))
abline(fit)
text(afl_data$Attendance~afl_data$Differential, labels=(afl_data$Season),data=afl_data, cex=.8, pos=4)
```




```
cor(afl_data$Attendance, afl_data$Differential)
```

```
## [1] -0.6956248
```

Finally, we see that there is a strong negative correlation of -0.696 between the difference of underdog odds and favorite odds, further reinforcing the fans' propensity towards attending games in seasons which have greater uncertainty of outcome.

The consistent strong relationships through this analysis show that AFL fans respond to uncertainty of outcome in a much stronger manner than fans in the EPL, Bundesliga, or Serie A. The relationship between these factors is opposite for the AFL as opposed to the EPL.

The full dataset is included below.

```
afl_data
```

```
##    Season Attendance Points SDPoints OddsHome OddsFavorite OddsUnderdog
## 1    2018      36687     44 18.52185 2.190531     1.402705     3.885266
## 2    2017      35207     44 14.25812 2.010725     1.429227     3.556377
## 3    2016      33163     44 20.21066 2.554203     1.361498     4.669807
## 4    2015      33367     44 18.30461 2.360291     1.380485     4.120825
## 5    2014      33680     44 17.42210 2.400435     1.360966     4.428792
## 6    2013      33461     44 20.36144 2.926763     1.337681     5.642126
## 7    2012      32748     44 20.22229 2.991787     1.318551     5.806812
## 8    2011      36428     44 21.48255 2.482194     1.362398     4.613724
## 9    2010      38423     44 15.83246 2.055161     1.423333     3.555323
## 10   2009      37760     44 18.59032 2.169565     1.400000     3.798696
##    Differential
## 1      2.482560
## 2      2.127150
## 3      3.308309
## 4      2.740340
## 5      3.067826
## 6      4.304444
```

```
## 7       4.488261
## 8       3.251327
## 9       2.131989
## 10      2.398696
```