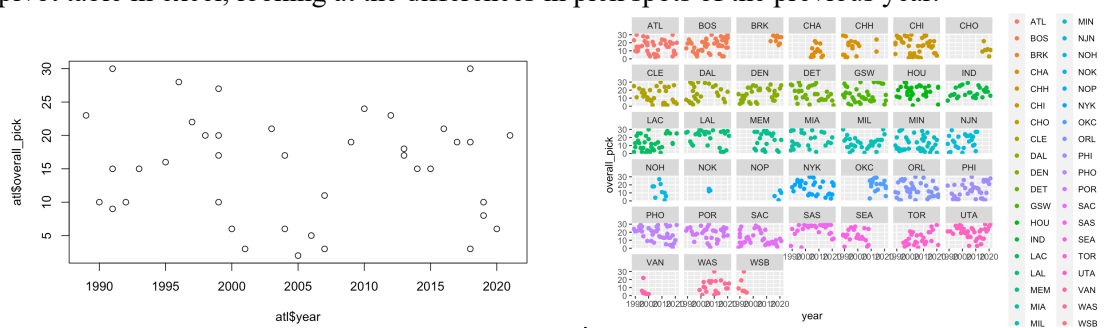


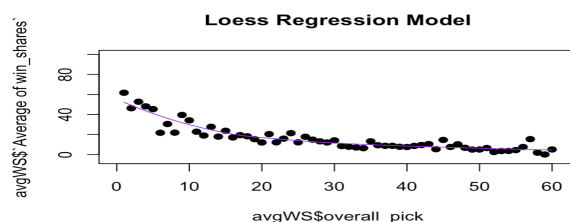
Part 1: Data Comprehension (estimated time: 1.5 hours)

- A. Which NBA team(s) has drafted the most players who...
- Went to Duke and were drafted in or before the 2000 draft?
 - Dallas, Minnesota, Phoenix, all twice.
 - Have a first name that begins with D and were drafted in an even year draft (1990, 1992, 1994, ...)?
 - 110 players.
- B. Describe the relationship between a team's first round pick slot in one year with their first-round pick slot in the subsequent year. When looking at individual team's first round picks (ignoring outliers such as picks acquired via trade etc.), one can notice that the picks tend to be close in range, altering by about 1-4 pick slots from the previous year. For example, when looking at the differences in Atlanta's pick spots, a simple plot suggests that this is true, and when further investigating via a pivot table in excel, looking at the differences in pick spots of the previous year.



Part 2: Analytical Acumen (estimated time: 3.5 hours)

- (A) Prompt: Analyze draft position value and team success/deficiencies compared to expectation.
- Create a method for valuing each draft slot in the NBA Draft (picks 1 through 60 in most drafts).
 - Description of method, description of results. First, I looked at win shares versus draft position for the entire dataset in a simple scatter plot. As expected, the “highest” values tend to be found in the earlier picks. To simplify the graph, I took average WS for each pick number, creating the variable “average Win Share”, which effectively combines the data for each pick number across all years. This was done by using a pivot table of the data in Excel. Next, to optimize, I solved a local regression to the data (so that the squared error is minimized between assigned values and the values in the data and such that each value does not increase from pick to pick), to “smooth” out the values. I used a K-Fold Cross Validation to Find the Best Model/select the best span for the data, then selected the span with the lowest RMSE, which was .75. Then I plotted the values with the predicted values regression lines and used those values as the assigned value for each slot. The results are a smoothed-out curve that decreases as pick increases and is minimized to the required constraints.



b. Conditional on the expected value of the draft positions, which NBA teams have over or underperformed the most when drafting during this time span. Which College Teams have had the players outperform expectations the most after entering the NBA? For expected value of draft positions, I used the calculated values assigned in part a. Then, in excel, I found the difference between an individual's WS value, and the expected value for that pick # overall (the # from part a). After doing this, I then summed the total values of the found differences, using a pivot table, for the individual colleges/teams. Because some schools/teams vary in numbers (for example, Detroit has been around much longer and thus they have more picks than a defunct team like Vancouver), I divided by the # of picks per college/team. The top five scores for "overperformance" (i.e., most positive scores) are NOH, SEA, CHH, SAS, and LAL. The 5 underachievers are: LAC, WAS, NOP, CHO, and NOK. For colleges the top 5 overachievers were: Santa Clara, Trinity Valley, IUPUI, Little Rock, Wake Forest. The 5 underachievers are: San Jose State, San Jose State, Barton County Community College, Loyola Marymount, University of the Pacific.

c. Explain and present your findings with tables and visuals. What additional research areas would you focus on if given the opportunity to expand this study? I found that teams with more picks, tend to have more success in "overachieving", additionally, numerous teams with success in more recent years (i.e., have won a championship in the past ten years), for example, teams like GSW, CLE, LAL, SAS, and TOR are all within the top ten in terms of value/#picks. The other teams within the top team were "perennial" playoff teams. This can indicate that drafting well/over expectation is related to overall team success. For colleges, however, many of the top overperformers have just one draft pick that is very talented, so this is likely skewed. For example, Davidson's one player is Steph Curry. This makes it harder to compare from school to school. I would investigate how the value of each pick changes if you were using metrics other than win shares. Additionally, it would be interesting to investigate these calculations broken down by position, and by round. For example, are forwards "more valuable" than guards? Do they under/overperform? How do those calculations vary when looking at first round picks versus second round picks? Additionally, I would like to investigate finding a way to include undrafted players (you could also do a similar type of research project for how well they sign undrafted players).

