Each year, the National Basketball Association (NBA) holds a draft, where prospective basketball players are able to be chosen to join one of the 30 professional teams across the United States and Canada.

In order to be eligible for the draft, a player must be at least 19 years old and out of high school for at least one year. Prior to 2006, this rule was not in effect, and players could be drafted during/right out of high school.

The draft is comprised of 60 players and takes place over two rounds of 30 selections. Teams pick players in an order based on performance from the previous season, with teams that performed poorly getting earlier picks in order to create a seemingly more level playing field. It's important to note that there weren't always 30 players selected in each round, the number made its way up to 30 as more teams were added into the NBA.

The data you’ll be using focuses on players that were selected in the first round of the NBA draft between the years 1990-2021, and they are divided based on what number in the first round they were selected. For each player, there are a range of different statistics from their careers in the NBA, some of which are ongoing.

The graphs below show a density plot corresponding to the minutes per game based on players who were selected in Round 1 of the NBA draft. The dashed line represents the groups mean. The box plots represent the variance for the minutes played for players who were selected in the first round.

A diagram of a pick in a draft pick

Description automatically generated

A graph of a diagram

Description automatically generated

1. Based on this density plot and the box plots above, can we determine if there is a significant difference between the average number of minutes played per game based on a player’s pick in the first round of the draft? What conclusions can you draw upon first glance of the two data visualizations? Are there any concerns?
2. Now that we've determined ANOVA is most likely appropriate, write and interpret in context the null and alternative hypotheses we'll be using for the ANOVA test.
3. Using the statistics for each category provided below, fill in the ANOVA table and use it to answer the following questions.

A screenshot of a graph

Description automatically generated

k = number of groups

n = total sample size

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **SS** | **MS** | **F-statistic** | **P-value** |
| Pick in Draft |  | 13839 |  |  | 0.000 |
| Error |  |  |  |
| Total |  | 60091 |

1. Is there a significant difference in the mean number of minutes played per game based on when a player was selected in the draft?
2. Between which groups are we *most likely* to see a significant difference. Between which groups are we *least likely* to see a significant difference in minutes played. Provide evidence.