NFL Draft Analysis

Philip Ipe

4/25/2020

Preparation

Load the NFL draft history data, containing the draft data from rounds 1-4 from the 1970-2016 NFL draft.

```
master_df <- read.csv("rd1_4_draft.csv")
head(master_df)
```

```
##   X Rnd Pick  Tm             Player Pos Age   To AP1 PB St CarAV DrAV   G year
## 1 1   1    1 PIT Terry Bradshaw HOF  QB  22 1983   1  3 13   106  106 168 1970
## 2 2   1    2 GNB         Mike McCoy  DT  22 1980   0  0  7    46   40 132 1970
## 3 3   1    3 CLE        Mike Phipps  QB  22 1981   0  0  5    40   32 119 1970
## 4 4   1    4 BOS         Phil Olsen  DT  22 1976   0  0  1    19   NA  79 1970
## 5 5   1    5 BUF        Al Cowlings  DE  23 1979   0  0  5    29   11 101 1970
## 6 6   1    6 PHI        Steve Zabel  LB  22 1979   0  0  7    38   17 124 1970
```

```
master_df <- master_df[complete.cases(master_df[,12:13]),]
master_df$Rnd <- as.factor(master_df$Rnd)
```

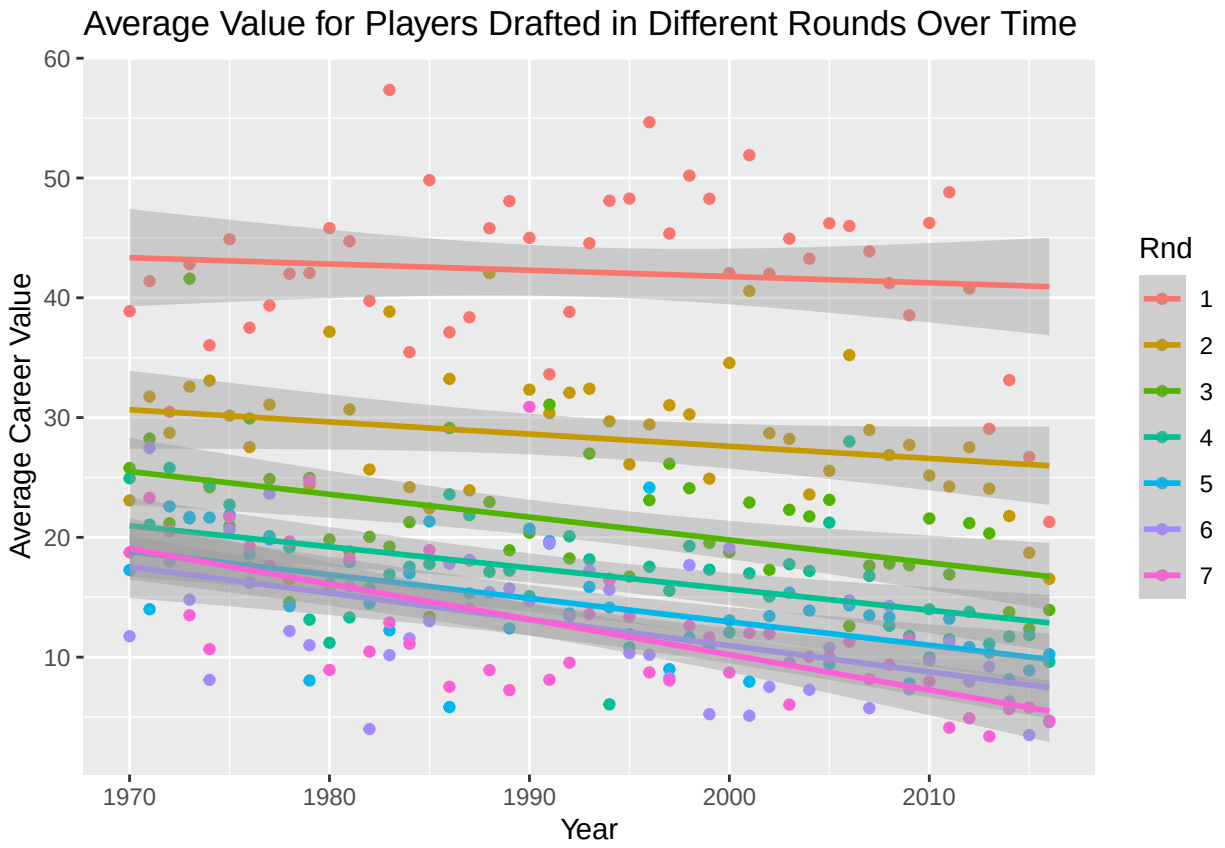
Trends in Average value

Has average value for players selected in different rounds gone up over time? Have GMs gotten better at drafting in the first round, second round, etc. . . ?

```
avg_val_df <- master_df %>%
  group_by(year, Rnd) %>%
  summarise(Avg_AV = mean(CarAV))

avg_val_df %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=year, y=Avg_AV, color=Rnd)) +
    geom_point() +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for Players Drafted in Differen
    geom_smooth(method=lm)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Seems like GMs have not gotten better at drafting in any of the first four rounds. In fact, we see a slight decline in average value for all rounds over time. We do see that GMs are selecting better players in the earlier rounds, as generally the average value for the earlier rounds is consistently higher than the lower rounds throughout the years. We see the trends spread out as time goes on as well.

Analyzing Different Positions

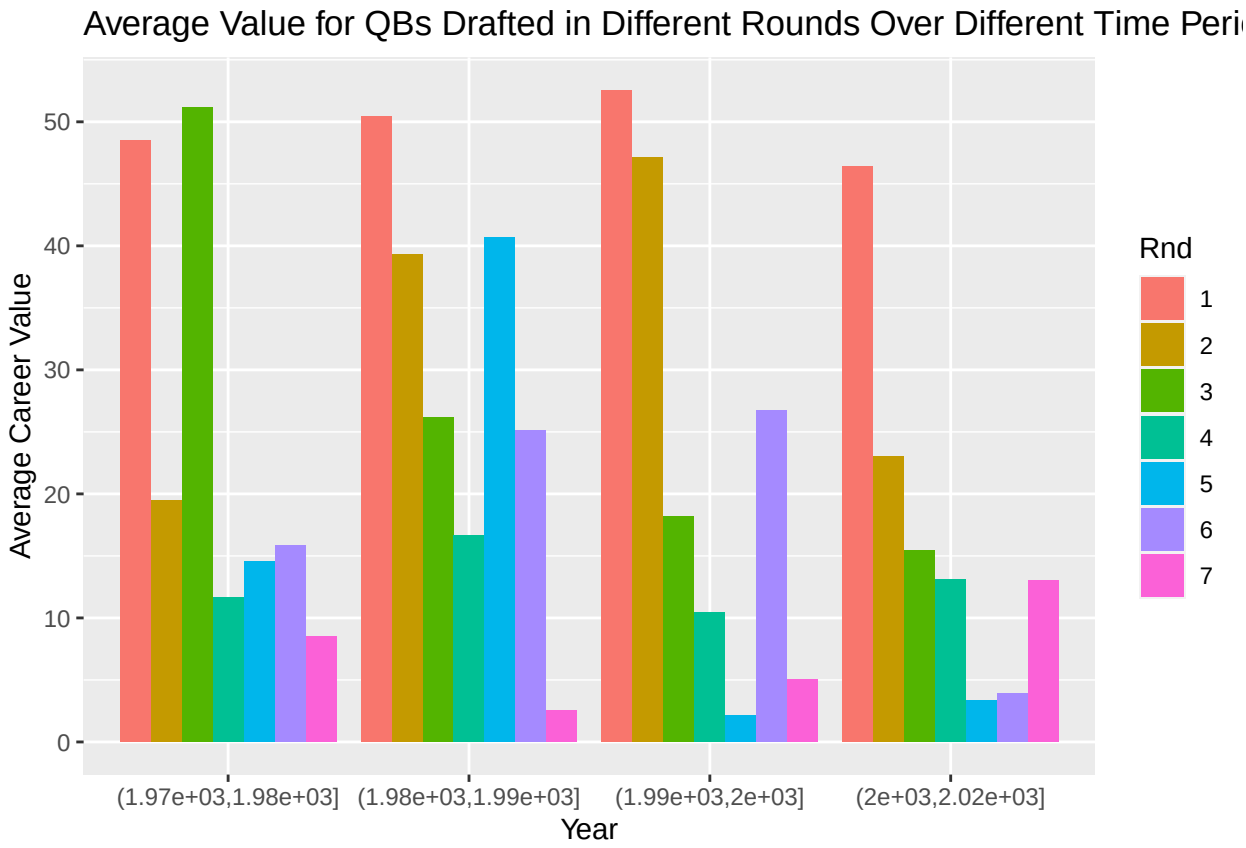
```
pos_df <- master_df %>%
  mutate(period = cut(year, breaks=4)) %>%
  group_by(period, Pos, Rnd) %>%
  summarise(Avg_AV = mean(CarAV))

master_df2 <- master_df

master_df2$Rnd <- as.numeric(master_df2$Rnd)
```

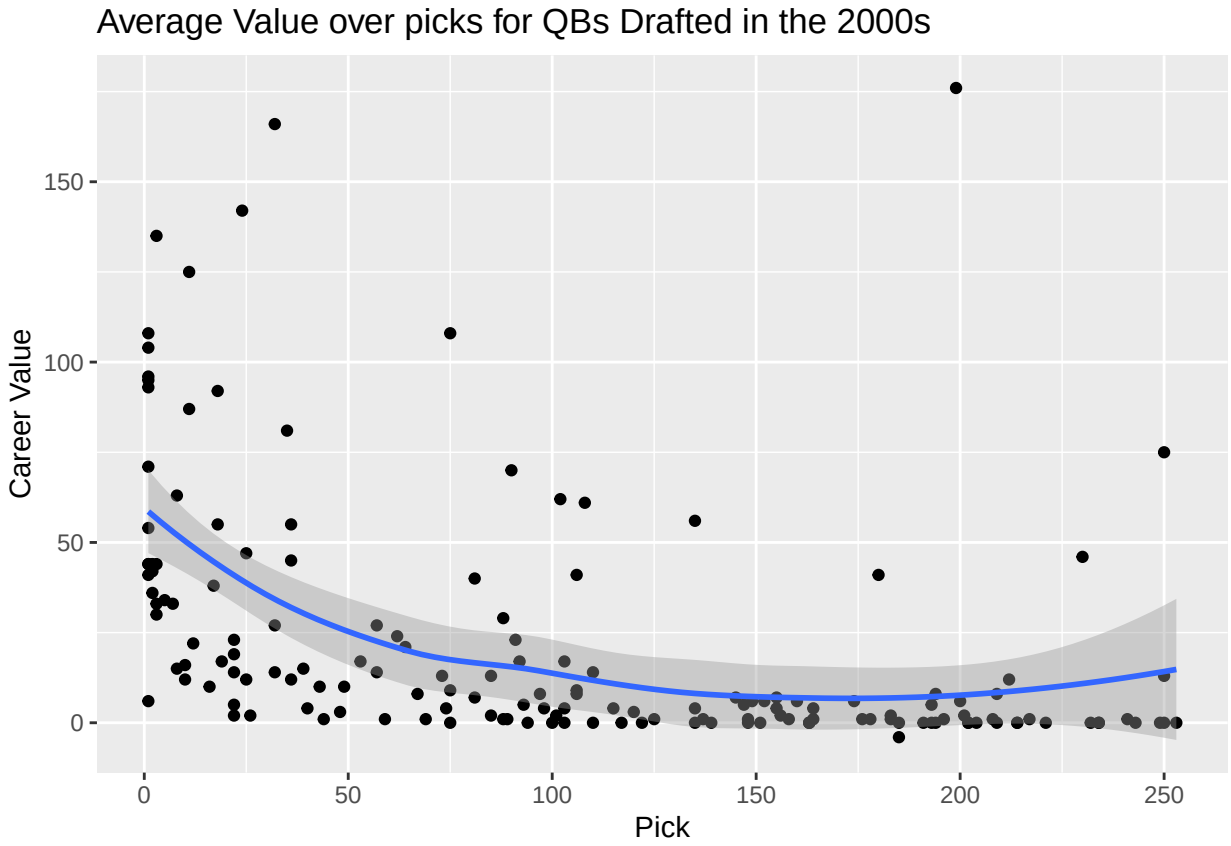
Lets analyze the average values of different positions grouped by round and time period in which they were drafted. Let's start with the most important position in football.

```
pos_df %>%
  filter(Pos == "QB") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for QBs Drafted in Different Ro
```



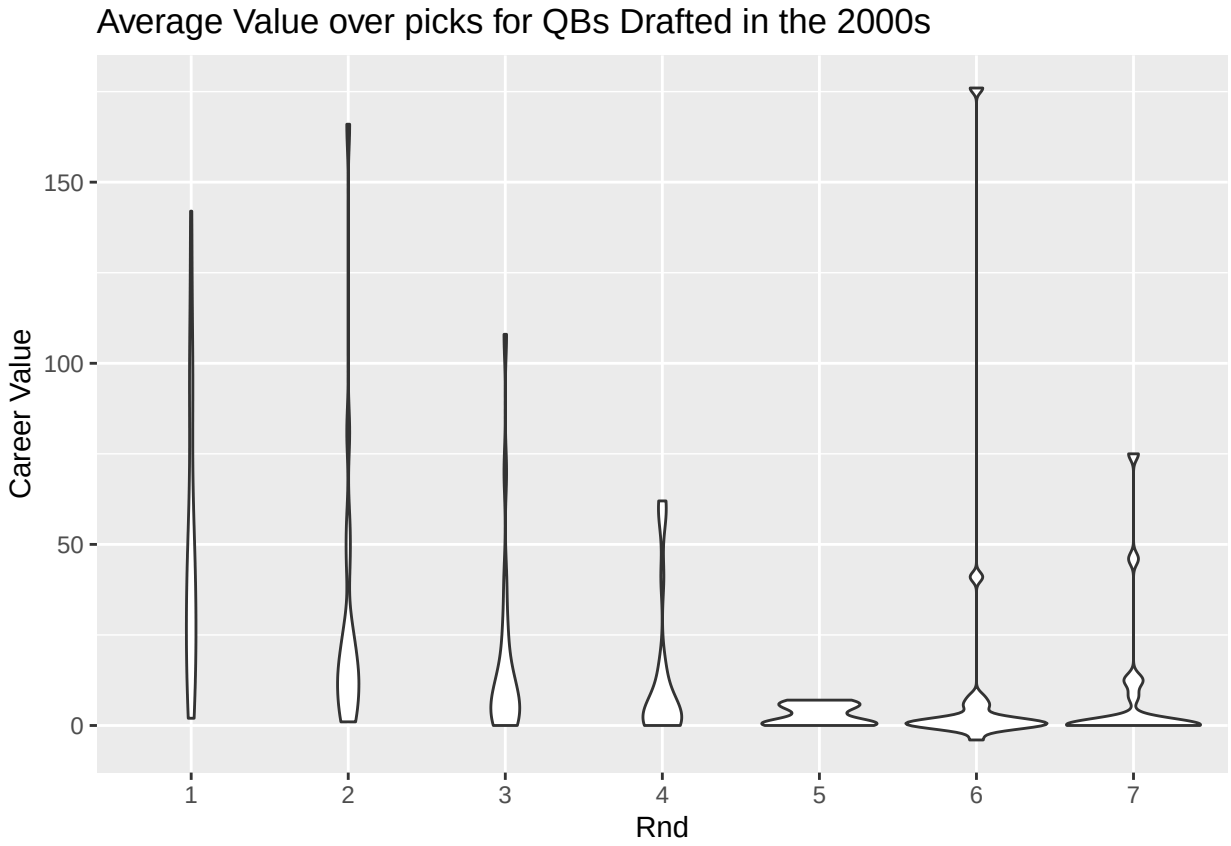

```
master_df2 %>%
  filter(year >= 2000) %>%
  filter(Pos == "QB") %>%
  ggplot(mapping=aes(x=Pick, y=CarAV)) +
    geom_point()  +
    labs(x = "Pick", y = "Career Value", title = "Average Value over picks for QBs Drafted in the 2000s
    geom_smooth(method='loess')
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
master_df2 %>%
  filter(year >= 2000) %>%
  filter(Pos == "QB") %>%
  ggplot(mapping=aes(x=factor(Rnd), y=CarAV)) +
    geom_violin()  +
    labs(x = "Rnd", y = "Career Value", title = "Average Value over picks for QBs Drafted in the 2000s")
    geom_smooth(method=lm)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

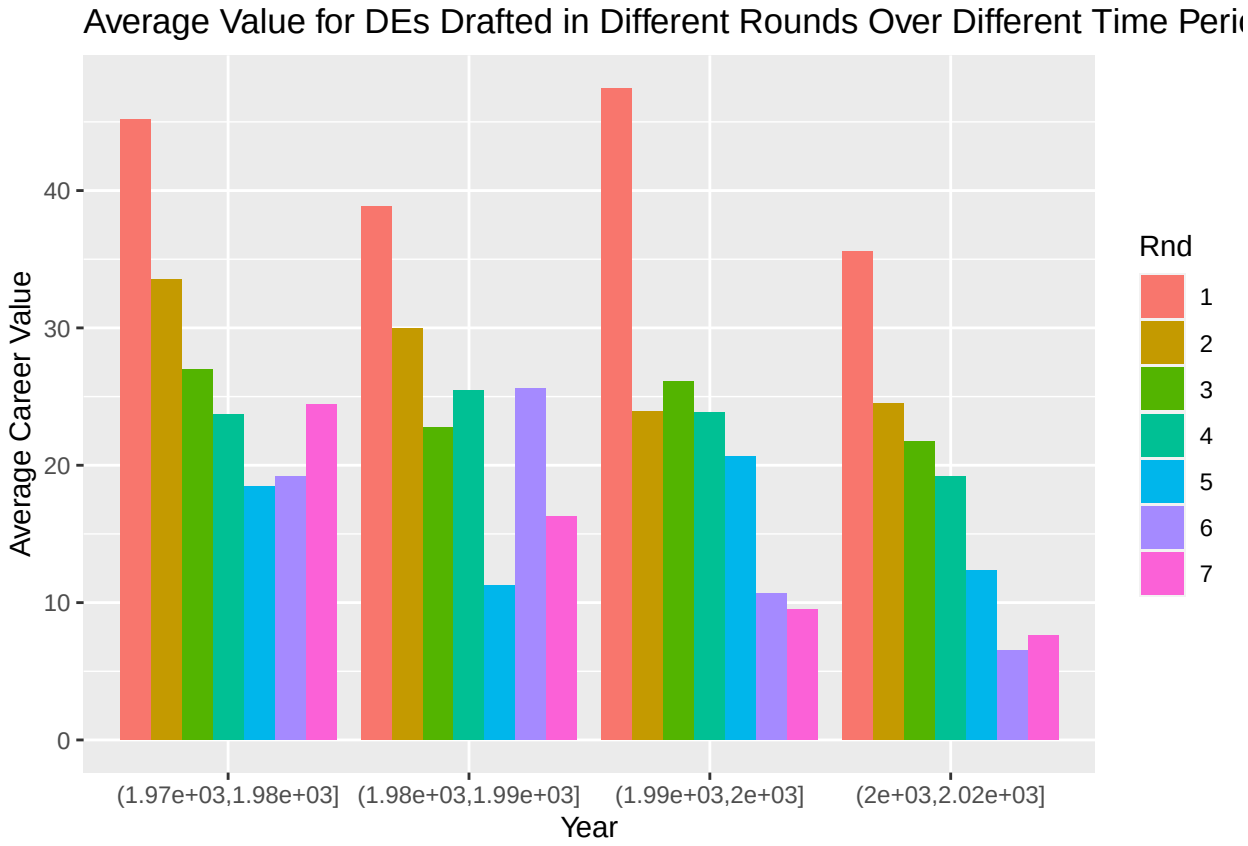


In the past 20 years (2000-2020), it is much more likely to hit on a QB in the first round than in any of the other rounds. We see a huge drop in average value after the first round in the past 20 years. Average value dips in half after that, and continues to decrease until the 7th round. It is interesting to note that the 7th round averge value shoots up again to about the same average values of the 3rd and 4th round in the 2000s.

Before then, we see that the 6th round was a value pick for QB, although Tom Brady might be the outlier that is skewing those numbers.

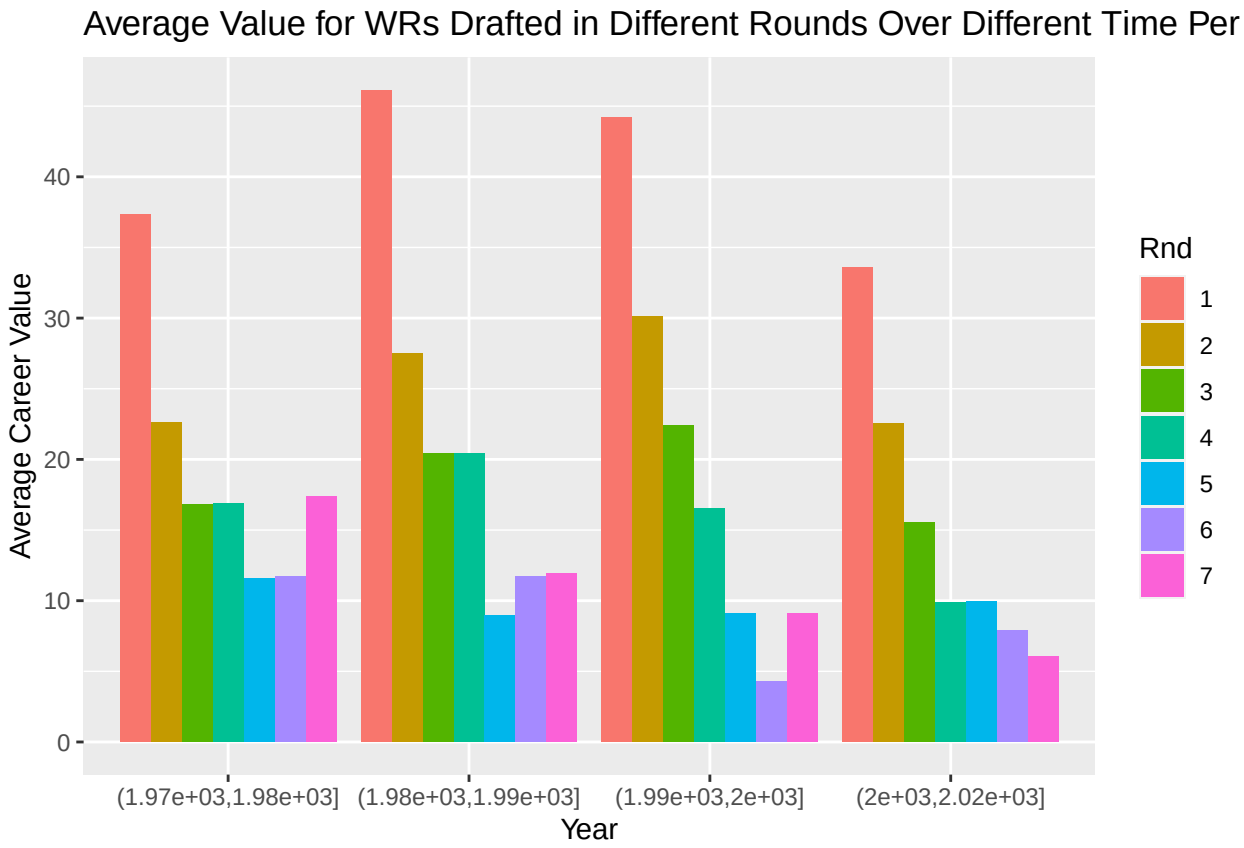
Now let's analyze the second most important position...

```
pos_df %>%
  filter(Pos == "DE") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for DEs Drafted in Different Rou
```



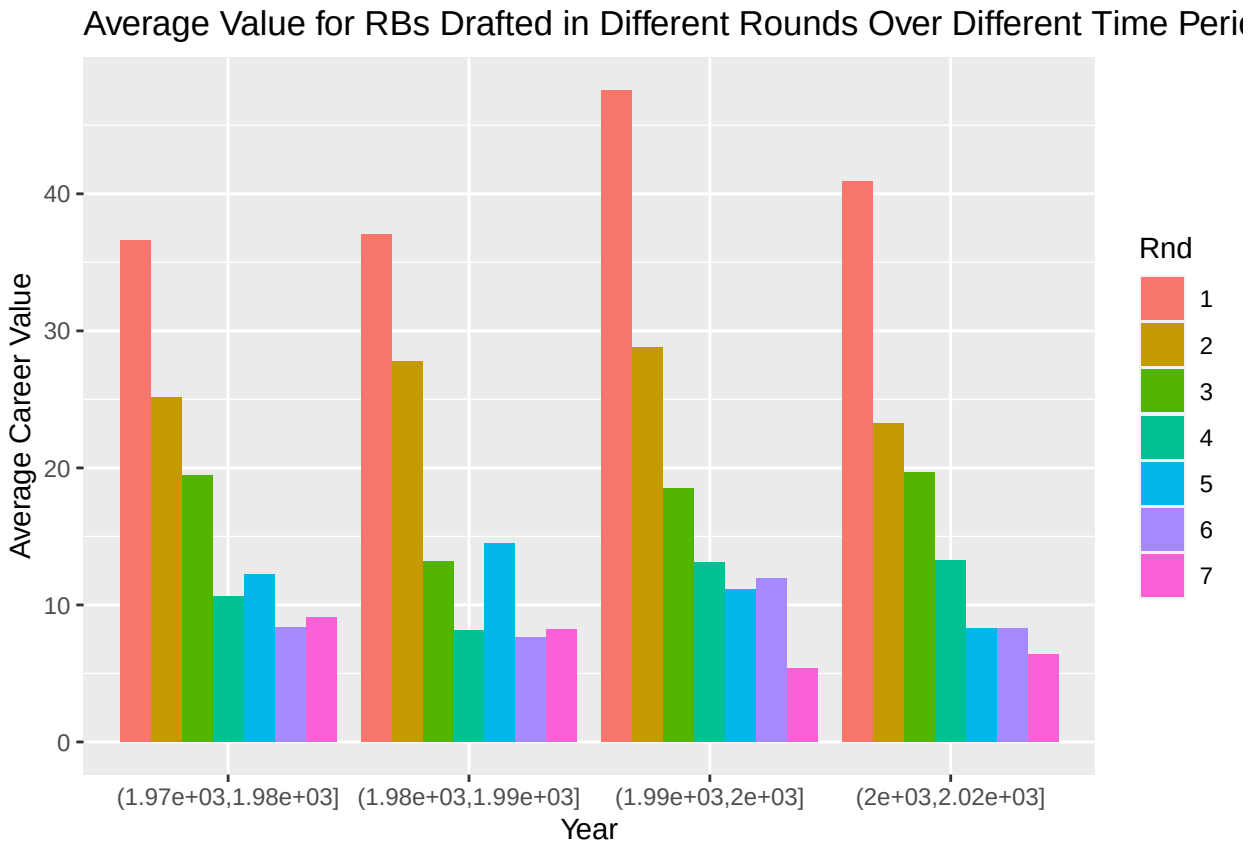
Not as big of a dip in average value from 1-4 round as other positions.

```
pos_df %>%
  filter(Pos == "WR") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for WRs Drafted in Different Rou
```



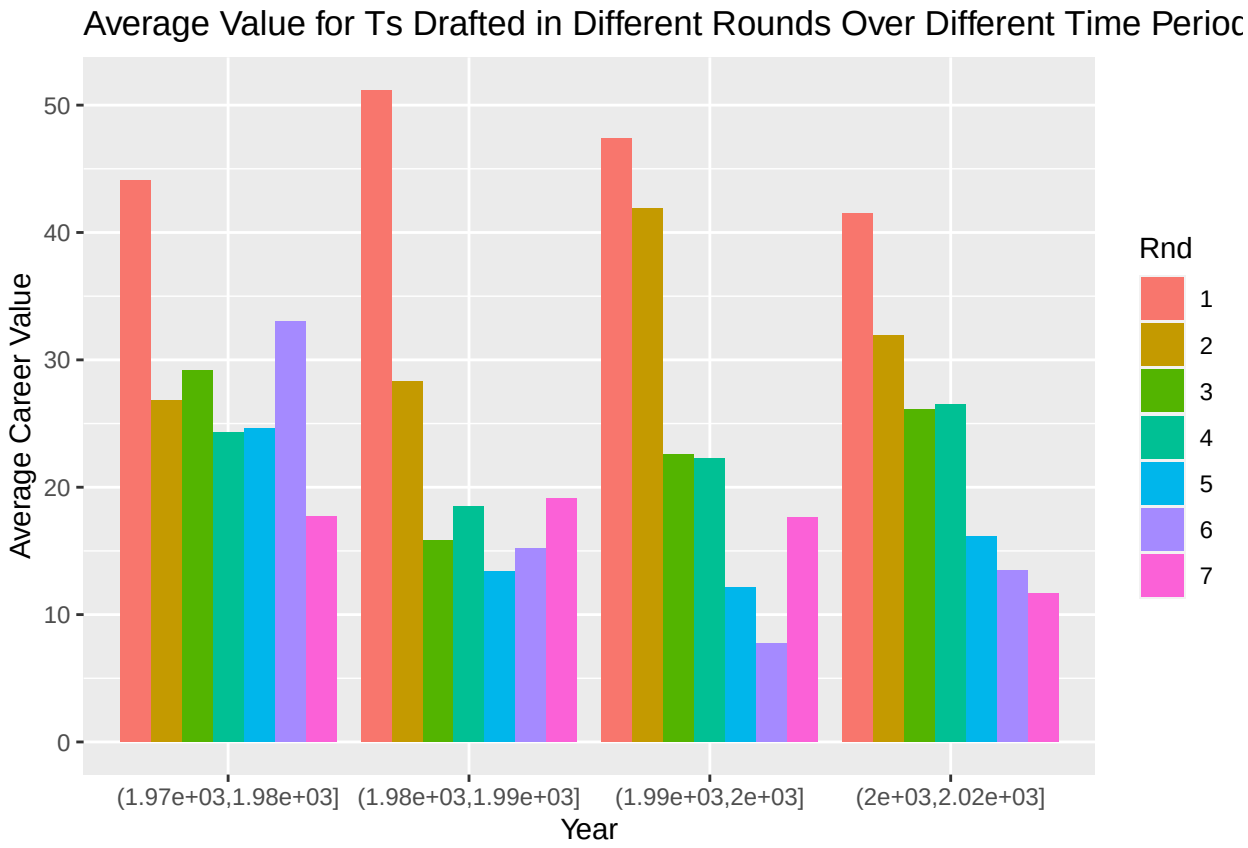
Rounds 4-6 have around the same value, suggests 5th and 6th rounds are value picks for WR. We also don't see as huge of a drop off in value from 1st to 2nd round here (and even from 2nd to 3rd) as we did when we looked at the quarterback position.

```
pos_df %>%
  filter(Pos == "RB") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for RBs Drafted in Different Rou
```



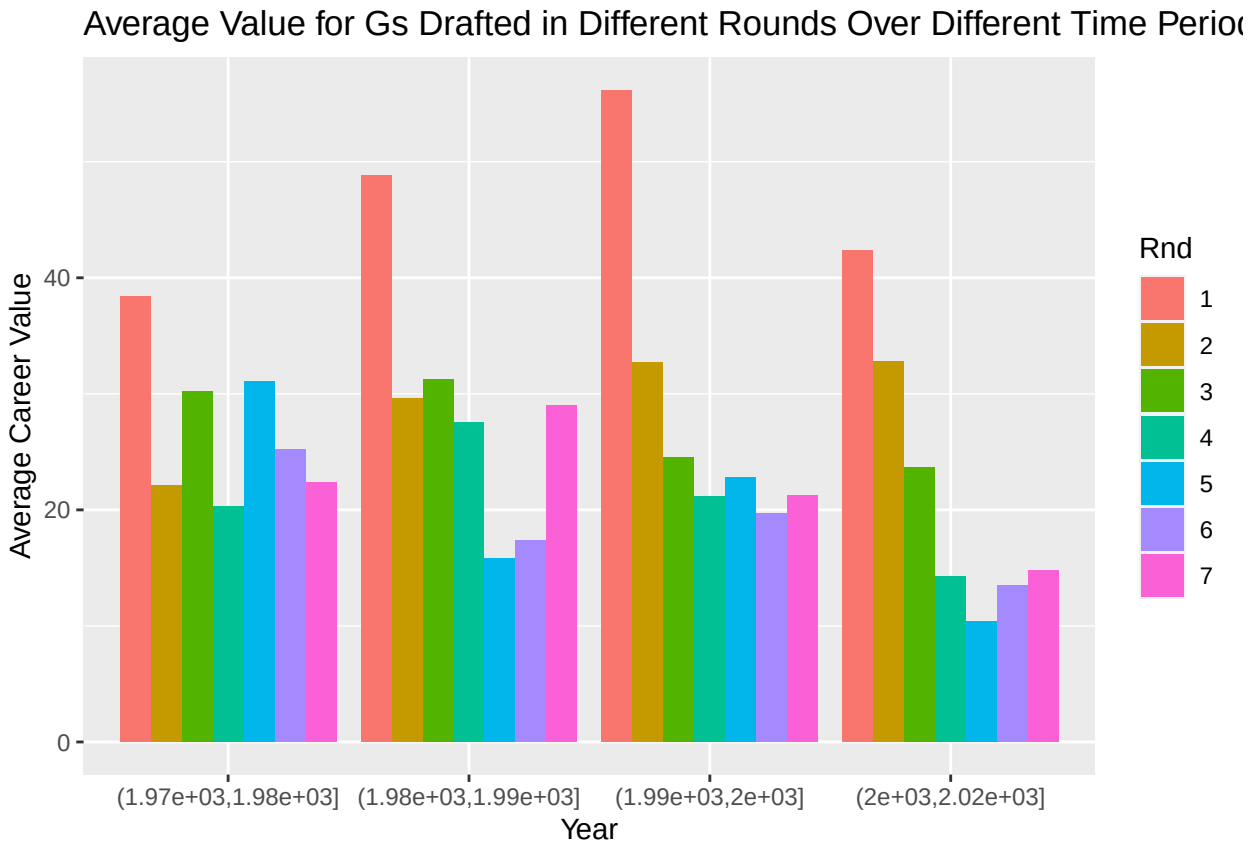
Big drop in average value after first round, almost half (not as much as WR). Pretty surprising given that teams are looking to draft RBs in later rounds nowadays.

```
pos_df %>%
  filter(Pos == "T") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for Ts Drafted in Different Rou
```

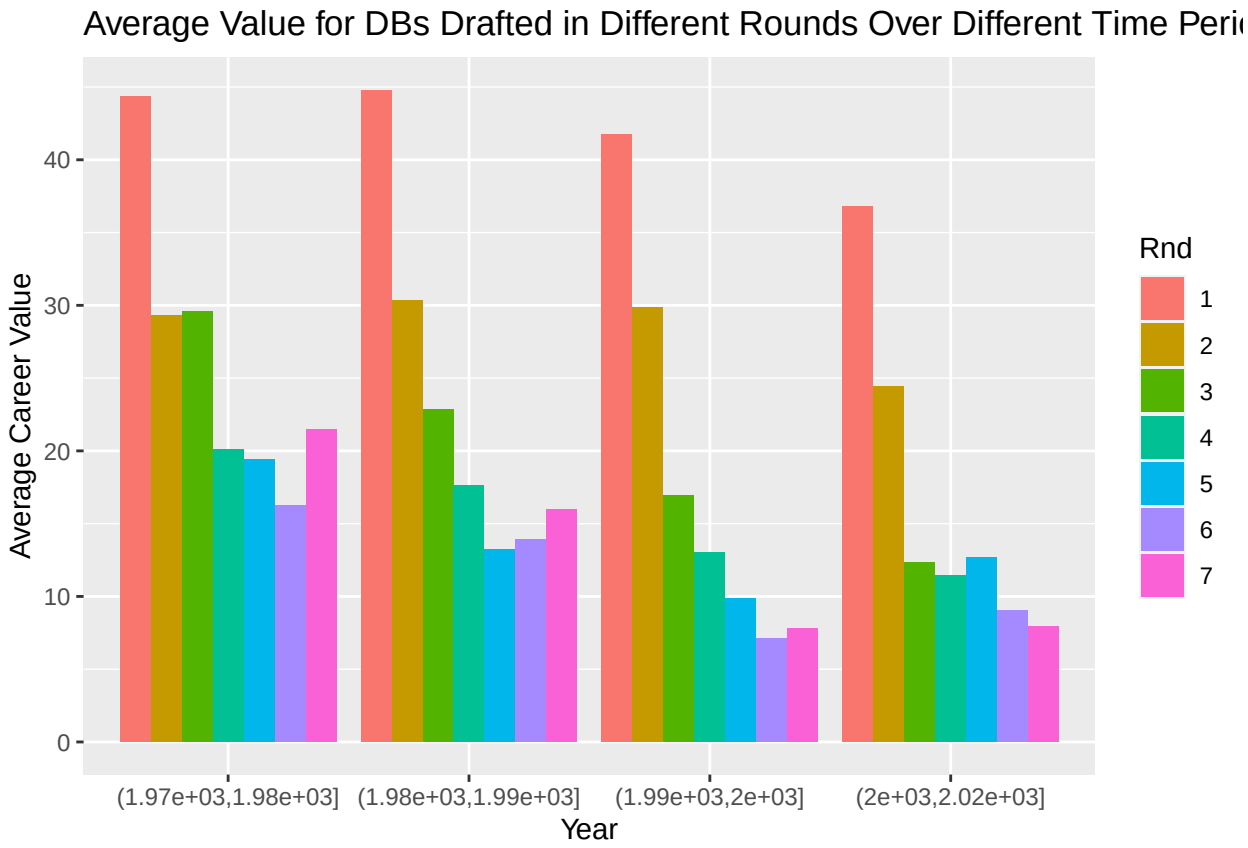


We don't see as big a drop in value from first-fourth rounds as wel do in other positions in the 2000s. Suggests there is typically more value in second round-fourth round T picks than other positions.

```
pos_df %>%
  filter(Pos == "G") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for Gs Drafted in Different Rou
```

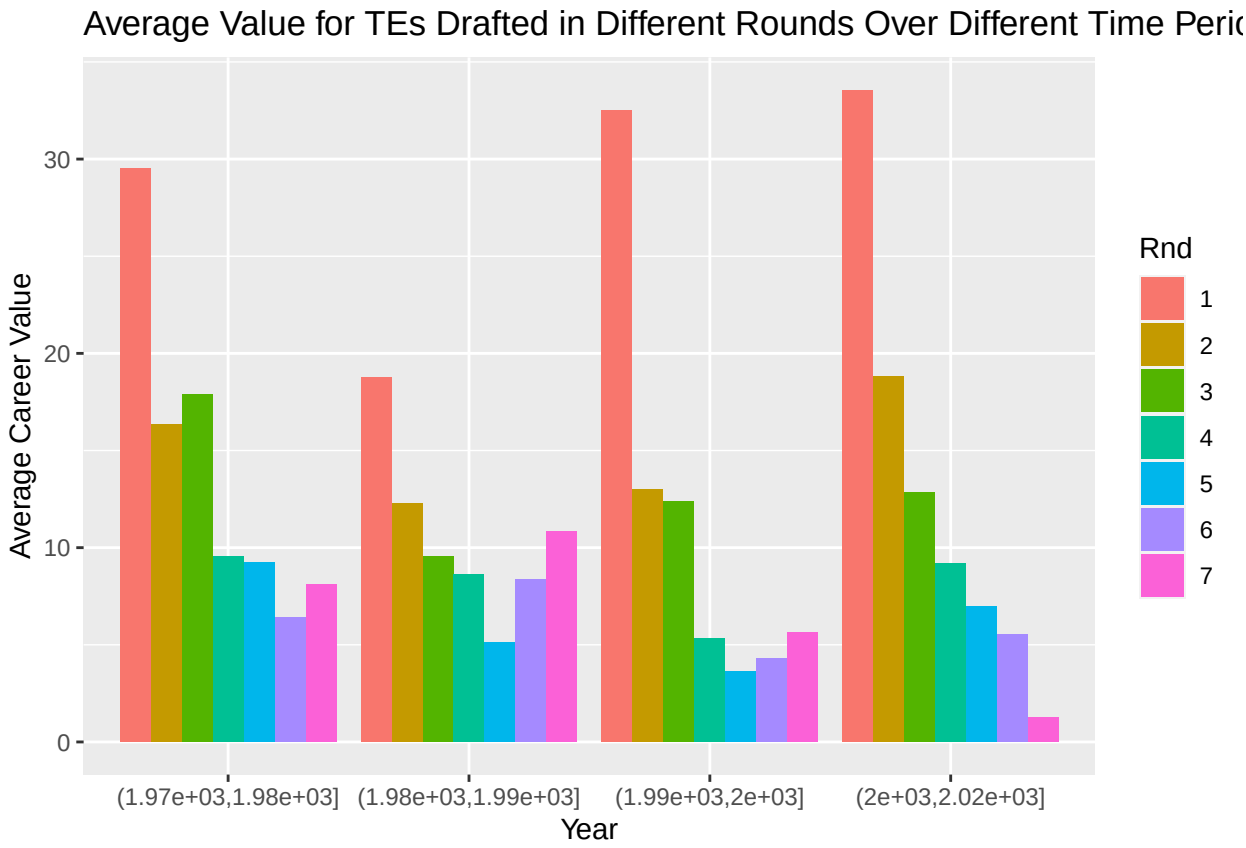



```
pos_df %>%
  filter(Pos == "DB") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for DBs Drafted in Different Rou
```



In the 2000s, value for DBs picked in rounds 3-7 have about the same value.

```
pos_df %>%
  filter(Pos == "TE") %>%
  group_by(Rnd) %>%
  ggplot(mapping=aes(x=period, y=Avg_AV, fill=Rnd)) +
    geom_bar(stat="identity", position = "dodge") +
    labs(x = "Year", y = "Average Career Value", title = "Average Value for TEs Drafted in Different Rou
```



Are there any positions that particularly suck when drafted after a certain round? What position show the least and most correlation between the pick number and average value brought? Which positions have been the best value picks?