THE ANCHORING EFFECT ON NELSON MANDELA'S AGE AT DEATH

Rachel Roggenkemper

CALIFORNIA POLYTECHNIC STATE UNIVERSITY STAT 365: Statistical Communcations

Introduction

The "anchoring" phenomenon is described as the tendency for someone to rely heavily on the piece of information they are given when making decisions or guesses. It is also a type of cognitive bias since someone's decisions can be influenced by some piece of information they receive which can affect their judgement and decision-making. It is also likely to change their decision compared to their decision they would have made without the anchor. We will investigate the anchoring phenomenon by analyzing data collected in my STAT 365: Statistical Communications class. We will be investigating two primary questions:

Does seeing the larger "anchor" value tend to produce larger age guesses, on average than seeing the smaller "anchor" value? If so, how much larger?

Does seeing the larger "anchor" value make a student more likely to provide an age guess of 75 or older than seeing the smaller "anchor" value? If so, how much more likely?

Additionally, we will also investigate two secondary questions:

Did students in the two sections of STAT 365 differ significantly with regard to average age guess (a) or percentage (b) who made a guess of 75 or older?

Did students in STAT 365 differ significantly from students in STAT 130 with regard to average age guess (a) or percentage (b) who made a guess of 75 or older?

We are interested in seeing whether or not the anchor value produced different guesses and whether the guesses differed based on which section or course the students are enrolled in.

Materials and Methods

The data was collected through a Google Form that was taken by the students enrolled in Dr. Allan Rossman's classes. The final data set included four variables and 113 observational units, each observational unit being each student's survey response.

One of the variables is which anchor value they received in their survey, depending on whether their phone number ended in an even or odd digit. This is how students were randomly assigned to which anchor value they were shown. The even digits received the smaller anchor value, where the question being asked was: "Was Nelson Mandela, the first president of South Africa following apartheid, older or younger than age 16 when he died?" and then asking students to make a guess for the age Mandela died. The odd digits received the larger anchor value, where the question being asked was: "Was Nelson Mandela, the first president of South Africa following apartheid, older or younger than age 160 when he died?" and then asking students to make a guess for the age Mandela died.

The second variable was the guess each student made trying to determine what age Nelson Mandela was when he died.

The third variable was which course the student was enrolled in. The students were either enrolled in STAT 365 or STAT 130. From the 113 students in total, 49 students were enrolled in STAT 130 and 64 students were enrolled in STAT 365.

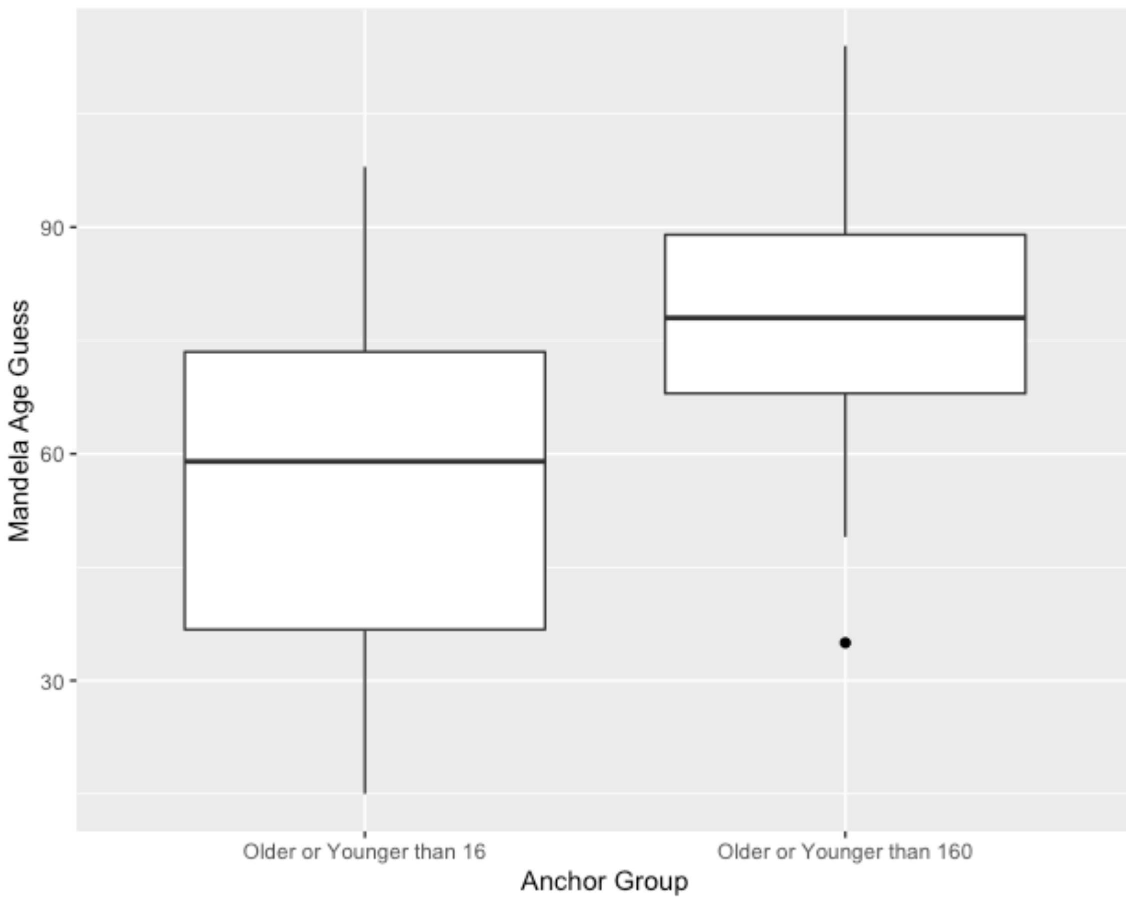
The fourth variable was which section the student was enrolled in, section 1 or 2. This variable only applies to students who are enrolled in STAT 365, because there is only one section of STAT 130.

The goal of the investigation is to see how the student's answer is influenced by the anchor value they were down and do the answers differ depending on which section or course they are enrolled in.

Analysis and Results

Question 1: Does seeing the larger “anchor” value tend to produce larger age guesses, on average than seeing the smaller “anchor” value? If so, how much larger?

Figure 1: Boxplot of Mandela Death Age Guess by Anchor Group



Based on Figure 1, we can see that the guesses of the group who was shown the larger anchor value of 160 tended to have higher guesses compared to the group who was shown the smaller anchor value of 16.

Figure 2: Descriptive Statistics of the Mandela Death Age Guesses based on Anchor Group

	Mean	**Standard Deviation**	**Group Size**
Older/Younger than 16	56.52	15.01	53
Older/Younger than 160	78.30	22.12	60

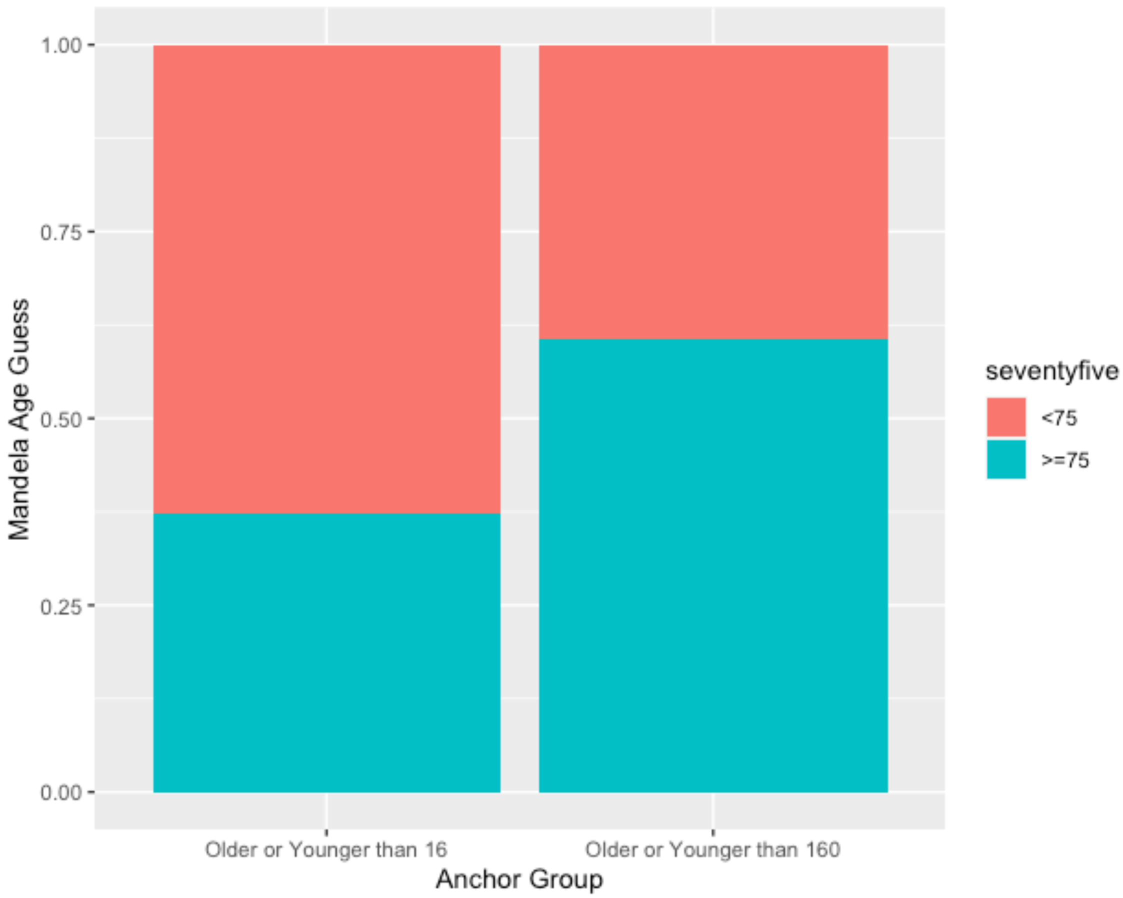
Based on Figure 2, we can support our conclusion from above as we can see that the guesses of the group who was shown the larger anchor value of 160 tended to have higher guesses compared to the group who was shown the smaller anchor value of 16, which can be seen through the difference in mean values of the group. To test to see if what the difference is and whether it is significant, we will conduct a hypothesis test.

After performing a two-sample t-test, with a small p-value of 6.20e-9, we have strong evidence that the larger anchor value tends to produce larger age guesses on average compared to the

group who was shown the smaller anchor value. We are 95% confident that seeing the larger anchor value will produce a guess of death age of between 14.64 to 28.93 larger on average than seeing the smaller anchor value.
This explanation is consistent with the boxplot shown in Figure 1 because we can see that the overall age guesses for the higher anchor value group are higher than the guesses from the lower anchor value.

Question 2: Does seeing the larger "anchor" value make a student more likely to provide an age guess of 75 or older than seeing the smaller "anchor" value? If so, how much more likely?

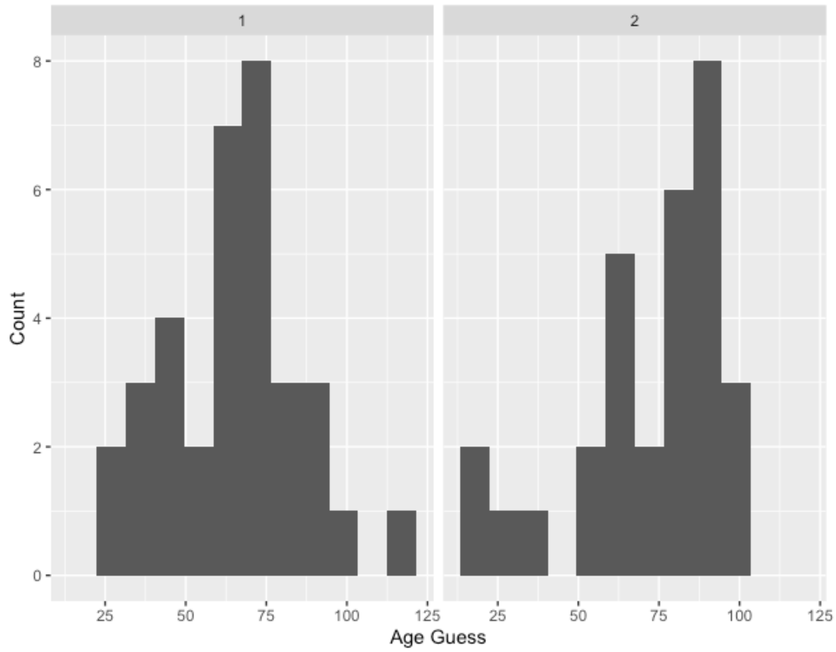
Figure 3: Proportion of Death Age Guesses over 75 by Anchor Group



Based on Figure 3, we can see that the proportion of students who guessed an age of 75 years or higher was larger for the group who received a larger anchor value compared to the group with the smaller anchor smaller. To formally test this, we will be performing a two-sample proportion test. Based on the small p-value of 0.0022, we have strong evidence that seeing the larger anchor value will make a student more likely to provide an age guess of 75 or older compared to seeing the smaller anchor value. We are 95% confident that seeing the larger anchor value makes one between 28.58% to 39.63% more likely on average to provide an age guess of 75 or older compared to seeing the smaller anchor value.

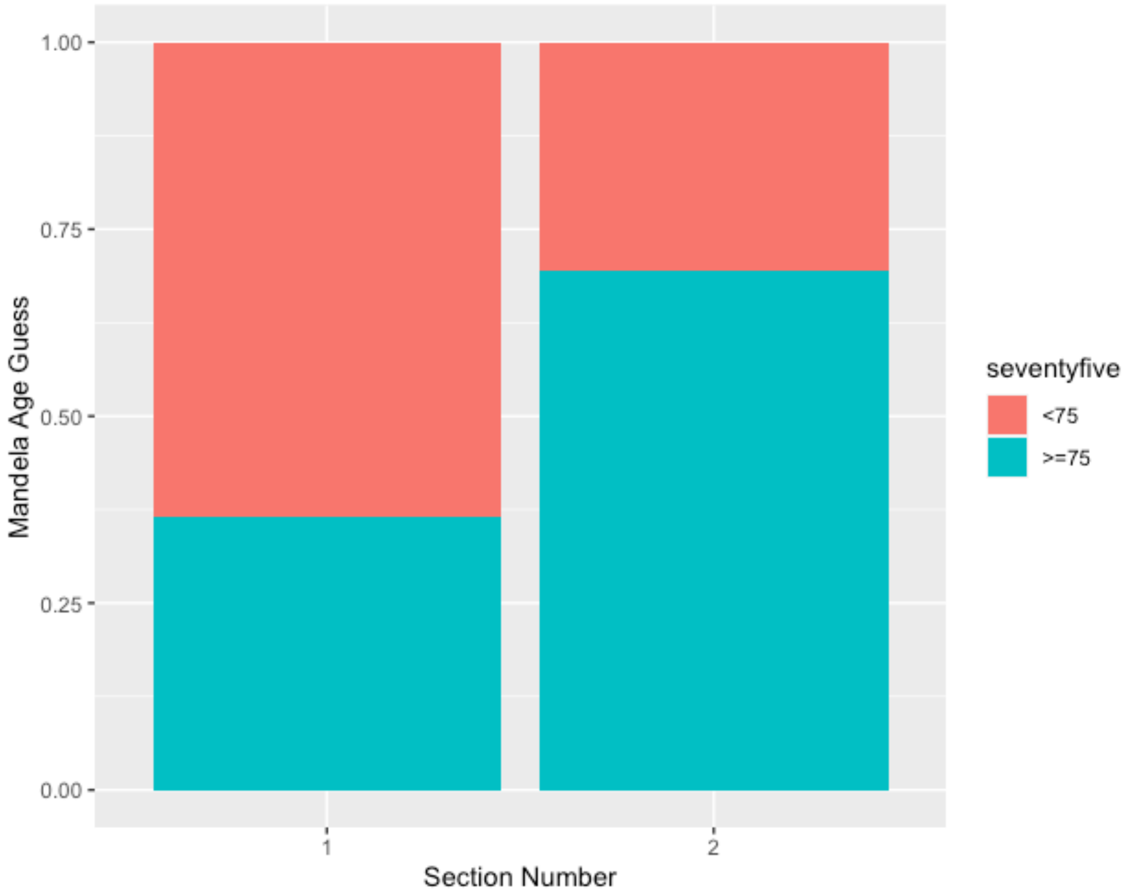
Question 3: Did students in the two sections of STAT 365 differ significantly with regard to average age guess (a) or percentage (b) who made a guess of 75 or older?

Figure 4: Distribution of Mandela Death Age Guess by STAT 365 Section



After running a hypothesis test to investigate whether the students in the two sections of STAT 365 differ significantly with regard to average age guess, based on the large p-value of 0.120, we do not have sufficient evidence that the students in the two sections of STAT 365 differ significantly with regard to average age guess. We are 95% confident that the true mean age guess for students enrolled in section 1 is between 19.316 years younger to 2.222 years older on average than section 2's average age guesses. Notably, we can notice that zero in included in this confidence interval, supporting our hypothesis that there is not significant difference of average age guess between the two sections. This is consistent what Figure 4 shows since both sections appear to have a similar age guess distribution. This makes sense because we used random assignment in each section and expect both sections to have roughly an even mix of students that have phone numbers ending in even and odd digits so the distributions of guesses should be somewhat similar.

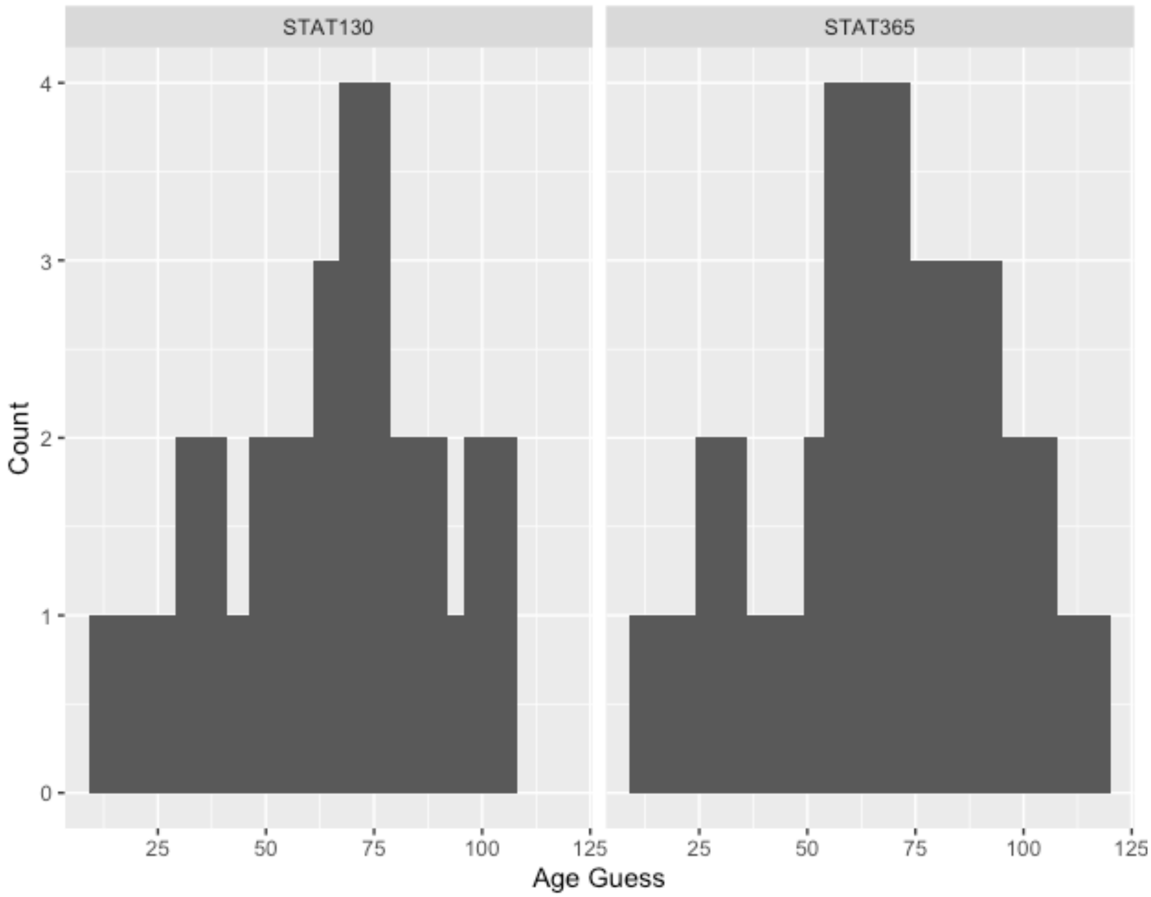
Figure 5: Proportion of Death Age Guesses over 75 by Section Number



Next, to test whether the two sections of STAT 365 differ significantly with regard to percentage who made a guess of 75 or older, we can conduct a two-sample proportion test. Based on the small p-value of 0.028, we have strong evidence that the two sections of STAT 365 differ significantly with regard to percentage who made a guess of 75 or older. We are 95% confident that the true proportion of students that guessed an age of death of 75 or older in the first section was 0.04 to 0.56 lower on average compared to the students that guessed an age of death of 75 or older in the second section.

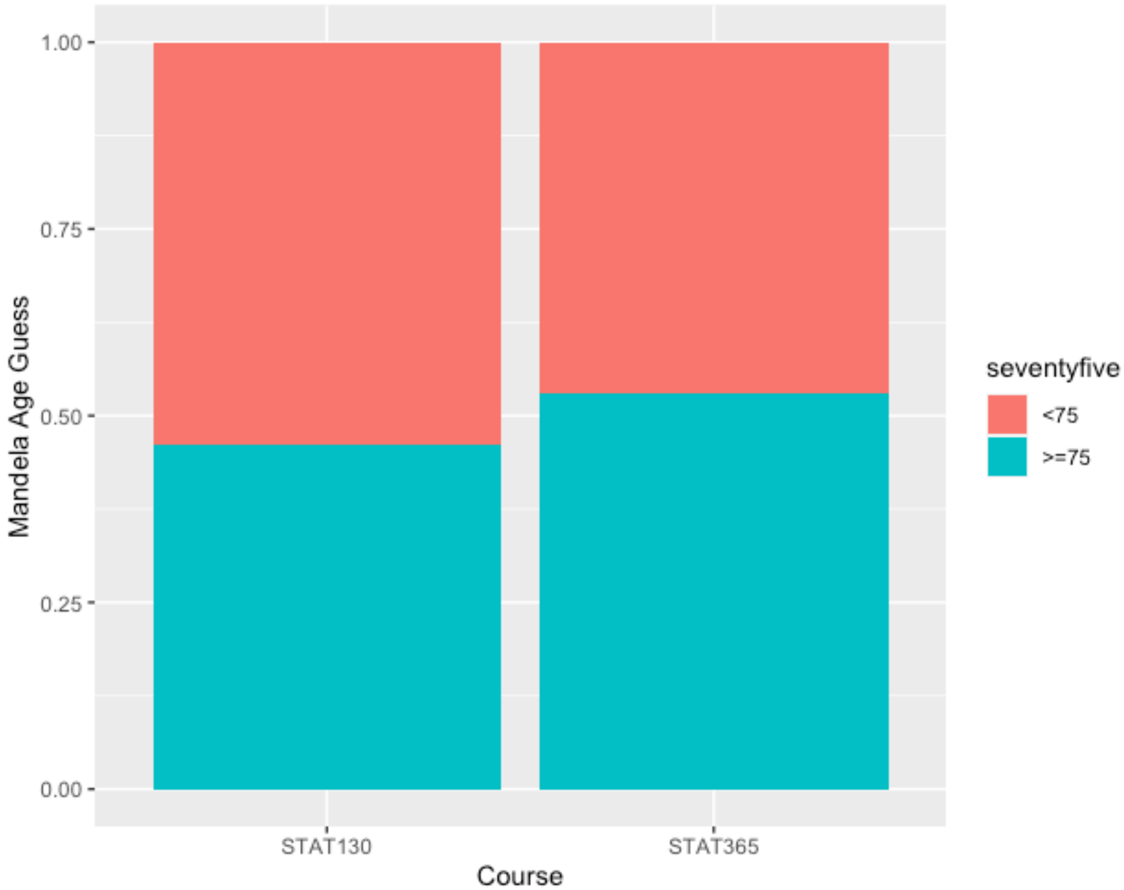
Question 4: Did students in STAT 365 differ significantly from students in STAT 130 with regard to average age guess (a) or percentage (b) who made a guess of 75 or older?

Figure 6: Distribution of Mandela Death Age Guess by Course



Looking at Figure 6, we can see that the distribution of death age guesses is very similar between both courses, STAT 130 and STAT 365. This means that we should not expect a significant different when running our hypothesis test. Based on the large p-value of 0.53, we have do not have sufficient evidence to believe that there is a significant difference between the average age guesses between the students in the two sections. We are 95% confident that the true average age guess is for students enrolled in STAT 365 is between 5.60 years younger to 10.79 years older on average than the true average age guess for students in STAT 130. Notably, zero is in this confidence interval, which further proves that there is not a significant difference between the average age guess between students based on the course they are enrolled in.

Figure 7: Proportion of Death Age Guesses over 75 by Course



Based on Figure 7, we can see that the proportions of students who guessed an age of 75 or older between the two sections are very similar, which means we do not expect to see a significant difference between percentage of students who guessed an age of 75 or older when comparing the two sections. Based on the large p-value of 0.65, we do not have sufficient evidence to conclude that there is a significant difference between percentage of age guesses 75 or older comparing students enrolled in STAT 130 and STAT 365. We are 95% confident that the proportion of age guesses of students enrolled in STAT 365 that were 75 or older is between 0.138 lower and 0.257 higher on average than the proportion of age guesses of students enrolled in STAT 130. Notably, zero is in this confidence interval which further supports our hypothesis that there is not a significant difference in proportion of age guesses based on course.

Conclusion

Ultimately, after analyzing the data collected in class, we can see that the anchoring phenomenon did play a role in the results when asked to guess the age of death of Nelson Mandela. As we investigated in Question 1, students who saw the larger anchor value tended to produce larger age guesses on average than students who saw the smaller anchor value. As seen in the investigation of Question 2, seeing the larger anchor value does make a student more likely to provide an age guess of 75 or older than seeing the smaller anchor value. The investigations of questions 3 and 4 revealed that the average age guesses did not differ significantly based on which section of STAT 365 the student was in, or which course the student was enrolled in. This makes sense because the section or course did not play a role in which anchor value the student was shown, because it was based off a random criterion, in this case whether the last digit of their phone number ends in an even or odd digit. Additionally, since random assignment was used in this experiment, that means we can draw cause and effect conclusions from our tests. So, that means we can conclude that seeing a higher anchor number caused students to produce larger age guesses. However, we cannot further generalize our results to a larger population because we did not utilize random sampling in our experiment.

Appendix

Appendix 1: R Code for Boxplot of Mandela Death Age Guess by Anchor Group (Figure 1)

```
mandela %>%
  ggplot() +
  geom_boxplot(aes(x = group, y = guess), outlier.colour = "black") +
  xlab("Anchor Group") +
  ylab("Mandela Age Guess") +
  scale_x_discrete(labels=c("even" = "Older or Younger than 16", "odd" = "Older or Younger
than 160"))
```

Appendix 3: R Code for Proportion of Death Age Guesses over 75 by Anchor Group (Figure 3)

```
mandela %>%
  mutate(seventyfive = case_when(guess >= 75 ~ '>=75',
                          guess < 75 ~ '<75')) %>%
  ggplot(aes(x = group, y = guess, fill = seventyfive)) +
  geom_col(position = "fill") +
  xlab("Anchor Group") +
  ylab("Mandela Age Guess") +
  scale_x_discrete(labels=c("even" = "Older or Younger than 16", "odd" = "Older or Younger
than 160"))
```

Appendix 4: R Code for Distribution of Mandela Death Age Guess by STAT 365 Section (Figure 4)

```
mandela %>%
  filter(section != "*") %>%
  ggplot(aes(x = guess)) +
  geom_histogram(bins = 12, aes(y = ..count..)) +
  xlab("Age Guess") +
  ylab("Count") +
  facet_wrap(~section)
```

Appendix 5: R Code for Proportion of Death Age Guesses over 75 by Section Number (Figure 5)

```
mandela %>%
  filter(section != "*") %>%
  mutate(seventyfive = case_when(guess >= 75 ~ '>=75',
                             guess < 75 ~ '<75')) %>%
  ggplot(aes(x = section, y = guess, fill = seventyfive)) +
  geom_col(position = "fill") +
  xlab("Section Number") +
  ylab("Mandela Age Guess") +
  scale_x_discrete(labels=c("even" = "Older or Younger than 16", "odd" = "Older or Younger
than 160"))
```

Appendix 6: R Code for Distribution of Mandela Death Age Guess by Course (Figure 6)

```
mandela$course[mandela$course == "C"] = "STAT365"
mandela$course[mandela$course == "D"] = "STAT130"

mandela %>%
  ggplot(aes(x = guess)) +
  geom_bar(width = 12) +
  xlab("Age Guess") +
  ylab("Count") +
  facet_wrap(~course)
```

Appendix 7: R Code for Proportion of Death Age Guesses over 75 by Course (Figure 7)

```
mandela %>%
  mutate(seventyfive = case_when(guess >= 75 ~ '>=75',
                                 guess < 75 ~ '<75')) %>%
  ggplot(aes(x = course, y = guess, fill = seventyfive)) +
  geom_col(position = "fill") +
  xlab("Course") +
  ylab("Mandela Age Guess") +
  scale_x_discrete(labels=c("even" = "Older or Younger than 16", "odd" = "Older or Younger
than 160"))
```