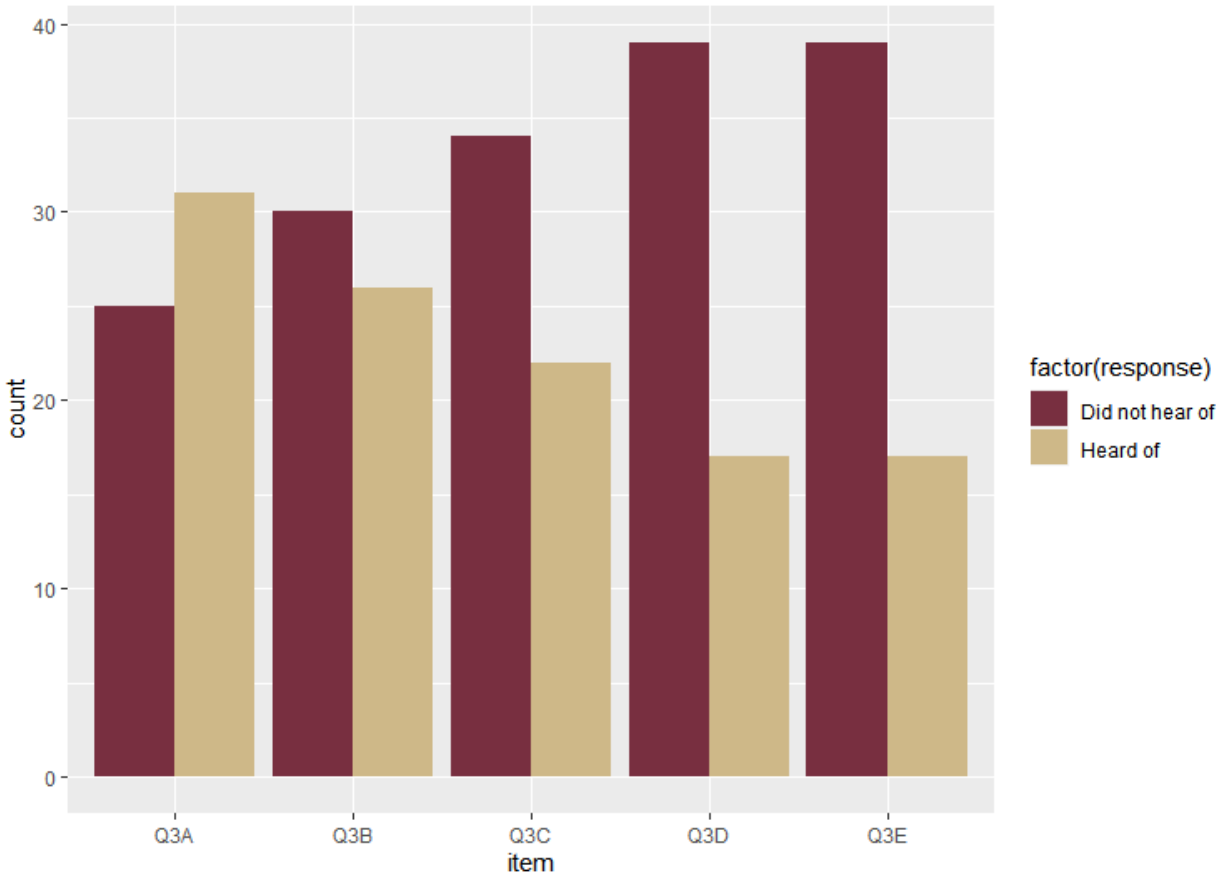
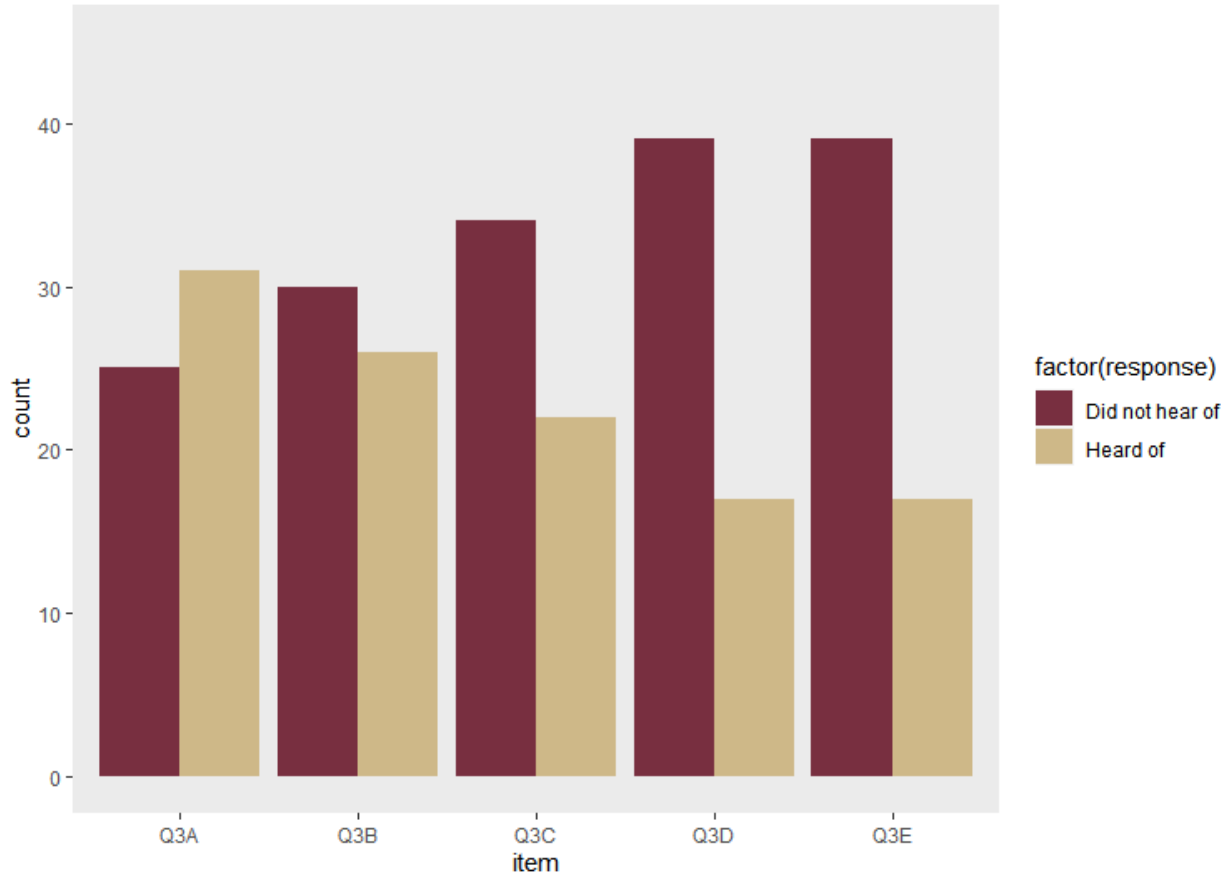
Code for Q3X and Q4X Bar Graphs

```
#Necessary R libraries to execute this code
>library(tidyr)
>library(ggplot2)

#Code for Question 3 graph

>Q3NewDataFrame <- structure(list(Q3A = c(X2022FSULibSurveyDataFinalVersion$Q3A), Q3B
= c(X2022FSULibSurveyDataFinalVersion$Q3B),Q3C =
c(X2022FSULibSurveyDataFinalVersion$Q3C), Q3D =
c(X2022FSULibSurveyDataFinalVersion$Q3D), Q3E =
c(X2022FSULibSurveyDataFinalVersion$Q3E), Q3YN =
c(X2022FSULibSurveyDataFinalVersion$Q3_YN_Num) ), row.names = c(NA,56L), class =
"data.frame")

#Note: 56L in the row.names statement allows us to have all 56 observations from our dataset
present.

> Q3_Long <- Q3NewDataFrame |> pivot_longer(-Q3YN, names_to = "item", values_to =
"response")

#Version 1
>ggplot(Q3_Long, aes(item, fill = factor(response))) + geom_bar(position =
position_dodge(preserve = "single")) + scale_fill_manual(labels = c("0" = "Did not hear of", "1" =
"Heard of"), values = c("0" = "#782F40", "1" = "#CEB888"))
```


40
30
20
10
0
count
Q3A
Q3B
Q3C
Q3D
Q3E
item
factor(response)
Did not hear of
Heard of

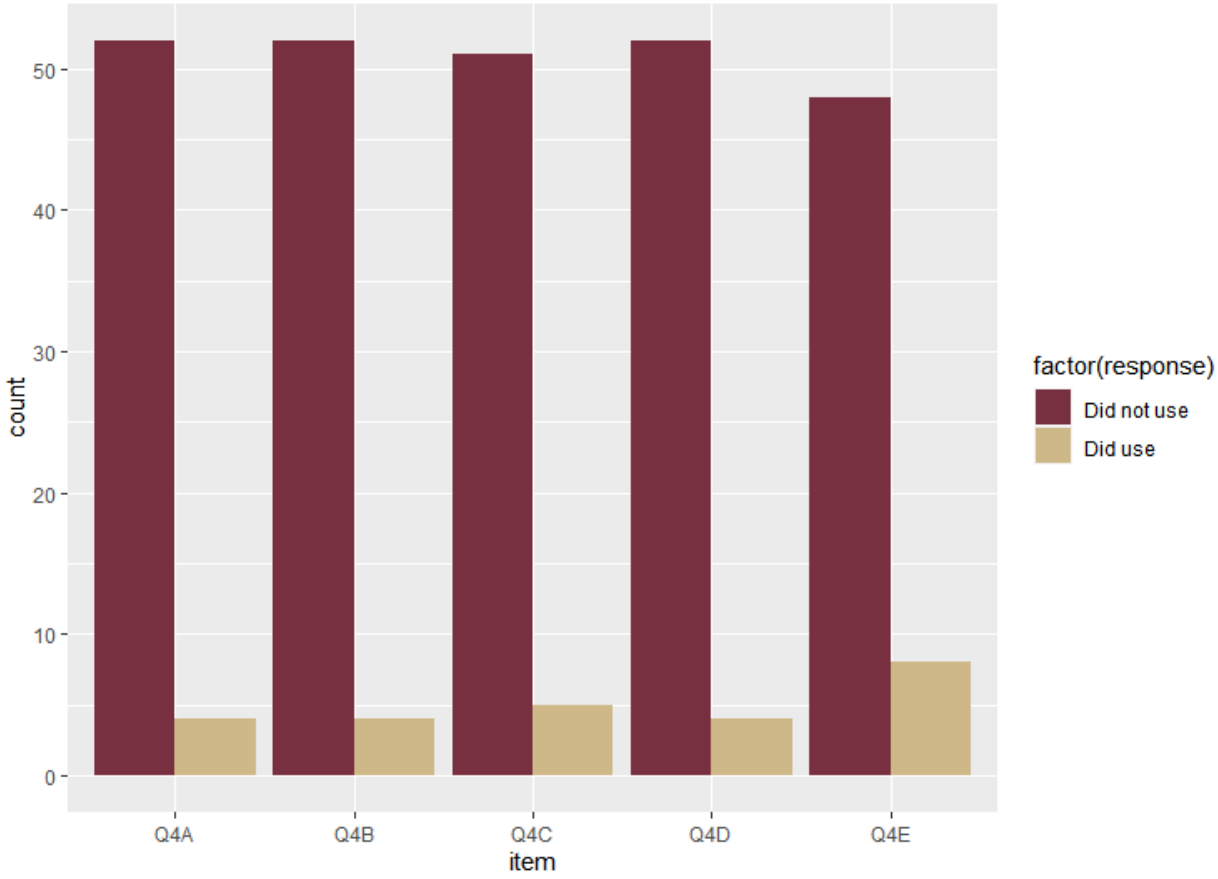
#Version 2 to account for y-axis height
> ggplot(Q3_Long, aes(item, fill = factor(response))) + geom_bar(position = position_dodge(preserve = "single")) + scale_fill_manual(labels = c("0" = "Did not hear of", "1" = "Heard of"), values = c("0" = "#782F40", "1" = "#CEB888")) + theme(panel.grid.major = element_blank(),panel.grid.minor = element_blank())+ylim(0,45)



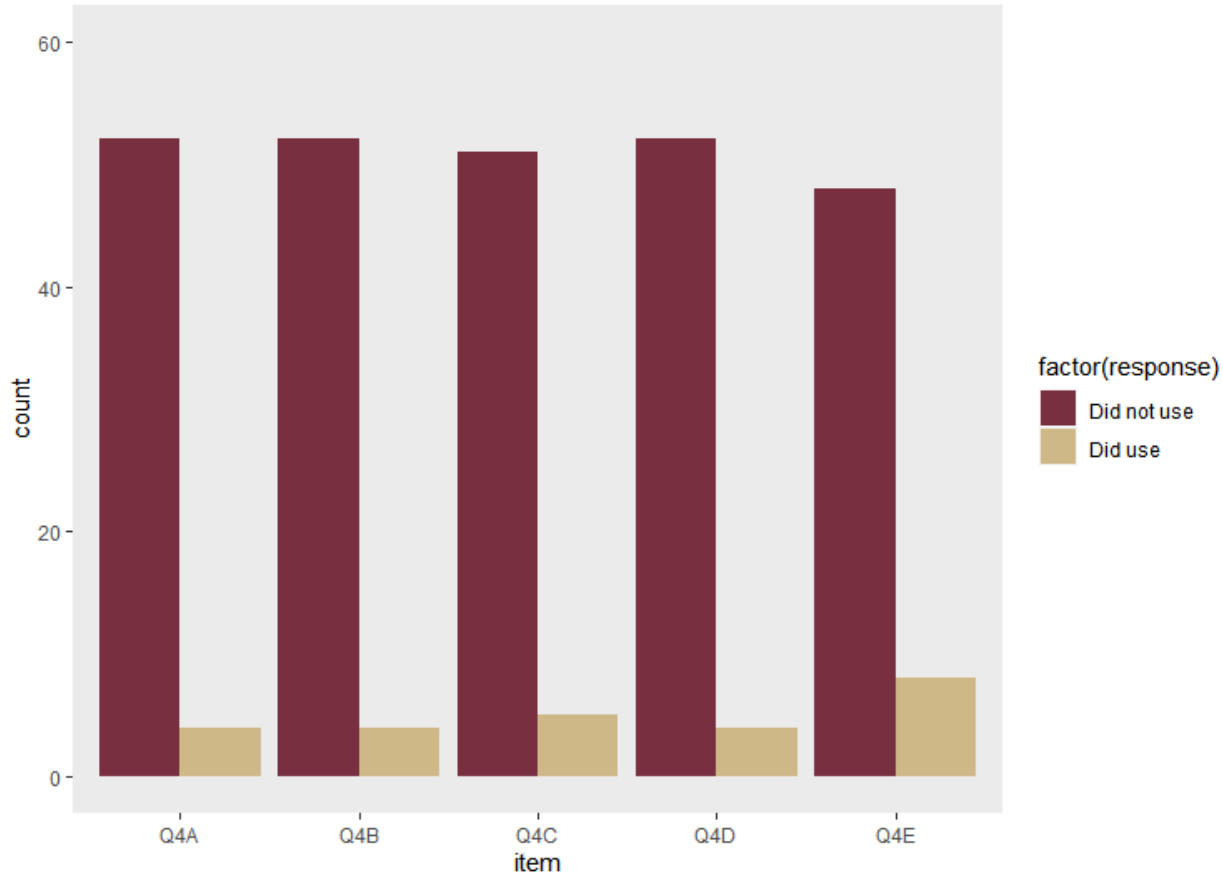
#Code for Q4 Graph

```
> Q4NewDataFrame <- structure(list(Q4A = c(X2022FSULibSurveyDataFinalVersion$Q4A), Q4B = c(X2022FSULibSurveyDataFinalVersion$Q4B),Q4C = c(X2022FSULibSurveyDataFinalVersion$Q4C), Q4D = c(X2022FSULibSurveyDataFinalVersion$Q4D), Q4E = c(X2022FSULibSurveyDataFinalVersion$Q4E), Q4YN = c(X2022FSULibSurveyDataFinalVersion$Q4_Num) ), row.names = c(NA,56L), class = "data.frame")

> Q4_Long <- Q4NewDataFrame |> pivot_longer(-Q4YN, names_to = "item", values_to = "response")
```

#Version 1
>ggplot(Q4_Long, aes(item, fill = factor(response))) + geom_bar(position = position_dodge(preserve = "single")) + scale_fill_manual(labels = c("0" = "Did not use", "1" = "Did use"), values = c("0" = "#782F40", "1" = "#CEB888"))



#Version 2 to account for y-axis
>ggplot(Q4_Long, aes(item, fill = factor(response))) + geom_bar(position = position_dodge(preserve = "single")) + scale_fill_manual(labels = c("0" = "Did not use", "1" = "Did use"), values = c("0" = "#782F40", "1" = "#CEB888")) + theme(panel.grid.major = element_blank(),panel.grid.minor = element_blank())+ylim(0,60)



#The code below allows for further bar graph delineation on which specific services have been heard about or not, as well as some arithmetic observations and properties about the Question 3 and Question 4 responses that allow the following R code to logically work:

- If a respondent did *not* hear about a specific research data service for any given Q3X option in Question 3, the entry for that given data service would equal 0. Likewise, if they *did* hear about a specific research data service for any given Q3X service, the entry for that given data service would equal 1.
- Assuming that a survey respondent did *not* hear about a given Q3X research data service, it would logically follow that they would *not* respond that they *used* that research data service in Question 4. Therefore, if Q3X=0, the corresponding Q4X value should also equal 0. (This is inherently true if the skip logic of the survey design was invoked by a respondent; if a survey respondent did not hear about *any* research data services, the respondent would *not* be asked about which services they used, as it would logically follow that the respondent *probably did not* use any research data services if they did not hear about those services in the first place.)
- If a respondent said they *didn't use* a given research data service in Q4X (if asked), the entry for that given research data service would equal 0. Likewise, if they *did* use a specific research data service for any given Q4X service, the entry for that given data service would equal 1.
- With the above in mind, we can note that if Q3X =1 and Q4X=1, then Q3X+Q4X=2.
- Presumably, Q3X+Q4X=2 can only be true if a survey respondent specifically said that they heard about a data service and then said that they *used* this service. If they did *not* use a service but did hear about it, Q3X=1 and Q4X=0, and 1+0=1.
- ***Therefore***, we can hereby separate responses into three groups with these arithmetic properties, where 0 = Did not hear about or use a specific research data service, 1 = Heard about a research data service, but did not use it, and 2 = Heard about *and used* a research data service
- From here, we can make a combined matrix that combines the variables for both Q3 and Q4 into one group that groups responses into the above outline 0, 1, and 2 response groups.

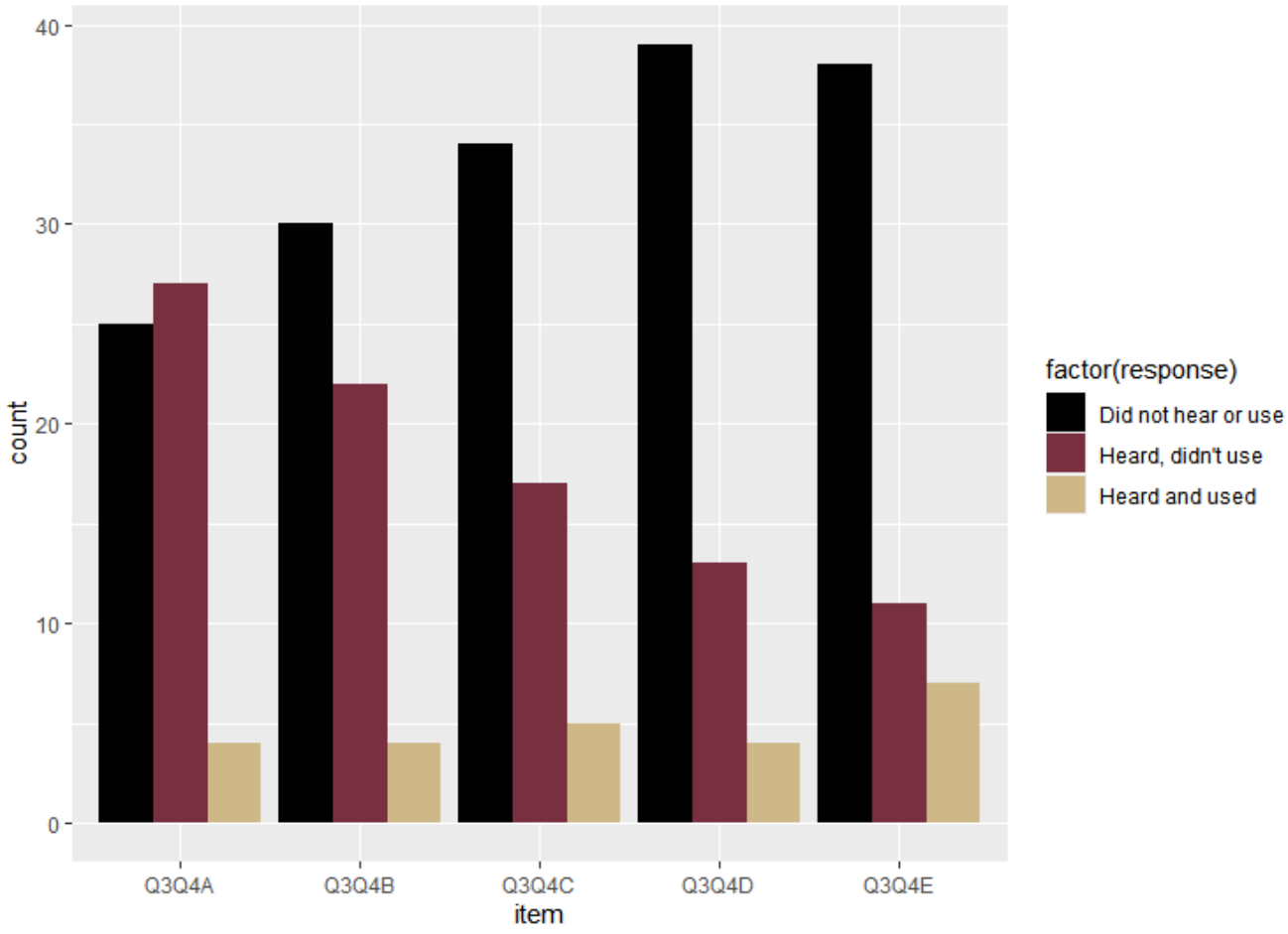
```
Q3Q4NewDataFrame <- structure(list(Q3A = c(X2022FSULibSurveyDataFinalVersion$Q3A),
Q3B = c(X2022FSULibSurveyDataFinalVersion$Q3B),Q3C =
c(X2022FSULibSurveyDataFinalVersion$Q3C), Q3D =
c(X2022FSULibSurveyDataFinalVersion$Q3D), Q3E =
c(X2022FSULibSurveyDataFinalVersion$Q3E), Q3YN =
c(X2022FSULibSurveyDataFinalVersion$Q3_YN_Num), Q4A =
c(X2022FSULibSurveyDataFinalVersion$Q4A), Q4B =
c(X2022FSULibSurveyDataFinalVersion$Q4B), Q4C =
c(X2022FSULibSurveyDataFinalVersion$Q4C), Q4D =
c(X2022FSULibSurveyDataFinalVersion$Q4D), Q4E =
c(X2022FSULibSurveyDataFinalVersion$Q4E), Q4YN =
c(X2022FSULibSurveyDataFinalVersion$Q4_Num)), row.names = c(NA,56L), class =
"data.frame")
```

#Variables to add all QX3 and QX4 responses to create data entries that equal 2

```
> Q3Q4A <- Q3Q4NewDataFrame$Q3A+Q3Q4NewDataFrame$Q4A
> Q3Q4B <- Q3Q4NewDataFrame$Q3B+Q3Q4NewDataFrame$Q4B
> Q3Q4C <- Q3Q4NewDataFrame$Q3C+Q3Q4NewDataFrame$Q4C
> Q3Q4D <- Q3Q4NewDataFrame$Q3D+Q3Q4NewDataFrame$Q4D
> Q3Q4E <- Q3Q4NewDataFrame$Q3E+Q3Q4NewDataFrame$Q4E
```

#Creating a dataset that uses the above variables as well as the Q4YN variable to execute the code necessary to make a bar graph with ggplot.

```
>LinAlgQ3Q4 <- structure(list(Q4YN =
c(X2022FSULibSurveyDataFinalVersion$Q4_Num),Q3Q4A=Q3Q4A, Q3Q4B=Q3Q4B,
Q3Q4C=Q3Q4C, Q3Q4D=Q3Q4D, Q3Q4E=Q3Q4E), row.names = c(NA,56L), class =
"data.frame")

>LinAlgQ3Q4_Long <- LinAlgQ3Q4 |> pivot_longer(-Q4YN, names_to = "item", values_to =
"response")

>ggplot(LinAlgQ3Q4_Long, aes(item, fill = factor(response))) + geom_bar(position =
position_dodge(preserve = "single")) + scale_fill_manual(labels = c("0" = "Did not hear or use",
"1" = "Heard, didn't use", "2" = "Heard and used"), values = c("0" = "#000000", "1" = "#782F40",
"2" = "#CEB888"))
```


40
30
20
10
0
count
Q3Q4A
Q3Q4B
Q3Q4C
Q3Q4D
Q3Q4E
item
factor(response)
Did not hear or use
Heard, didn't use
Heard and used