

```

library('tidyverse')

## Import the data set
lego <- read.csv("/Users/rebeccaswedberg/Downloads/game_data/lego.csv",
header = TRUE)

## Explore the data set
head(lego)

## Which age group submits the most reviews?

##summary of the dataset
summary(lego)

#Glimpse of the dataset
glimpse(lego)

## To plot the number of reviews per ages
qplot(ages, num_reviews, data = lego, main = "Number of reviews per age"
      , xlab = 'Age', ylab = 'Number of Review')

#Showing which age group is most likely to leave a review

lego[which.max(lego$num_reviews),]

#who are at least 25 years old (>25 years)?

#Create a new data frame using the suggested syntax
data_age25 <- lego[lego$ages>=25,]

#exploring the ages 25 data
head(data_age25)
summary(data_age25)
#Plotting a chart with the 25 age data

qplot(ages, list_price, data= data_age25,main = "Price paid by age group"
      , xlab = 'Age', ylab = 'List Price')

#Showing the most expensive product people over 25 bought. As we can see
#it cost 259.87 and was bought by 29 year old
data_age25[which.max(data_age25$list_price),]

# arrange the rows based on department column
#data_age25 %>% arrange(list_price)

#plotting the 20 most expensive lego pieces

# Order data descending based on list price
data_age25New <- data_age25[order(data_age25$list_price, decreasing =
TRUE), ]

# Showing the top most expensive Lego products sorted by list price
head(data_age25New)

#creating a scatterplot with the top prices with age,

```

```
plot(data_age25New$ages, data_age25New$list_price, pch=18,  
      xlab= 'Age', ylab = 'Price')
```