# Data Science & AI Workshop

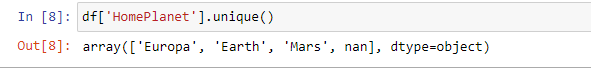
11-21-23

Q10 was omitted. Counting after that.

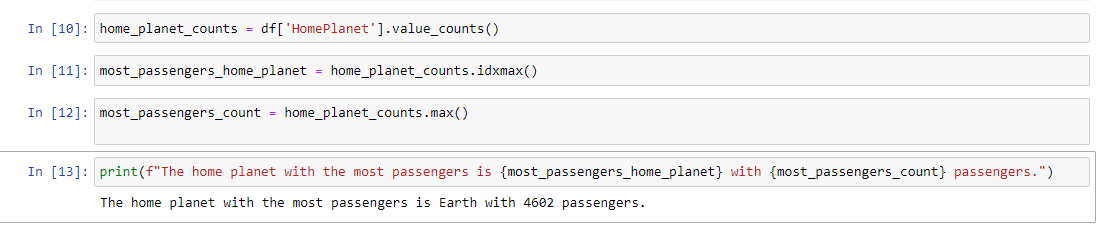
Group Members:

* Jahin Mahbub
* Amir Mohideen Bashir Khan
* Rucha
* Lahari

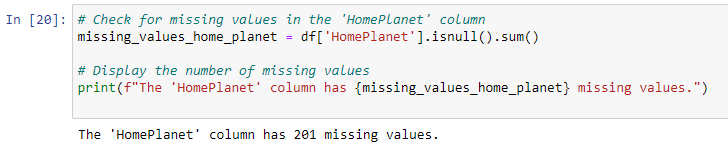
## Name all homeplanets



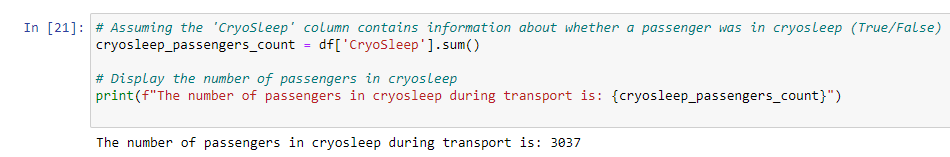
## Which homeplanet has the most passengers?



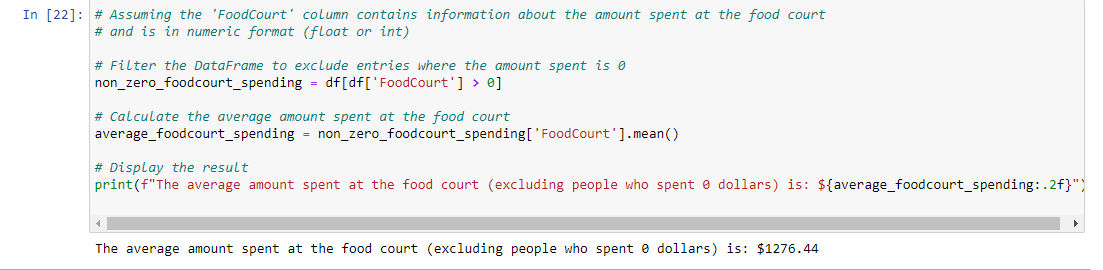
## Are there any missing values from the homeplanet? If so, how many?



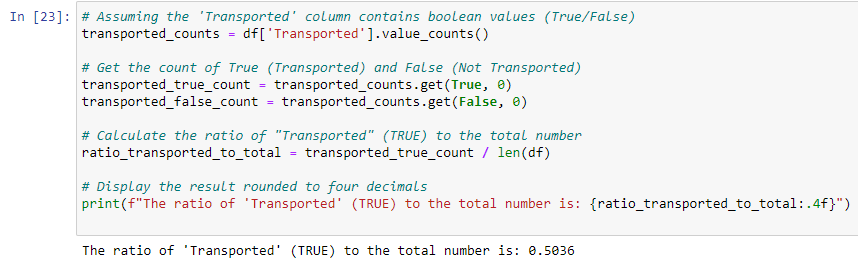
## How many passengers were in cryosleep during the transport?



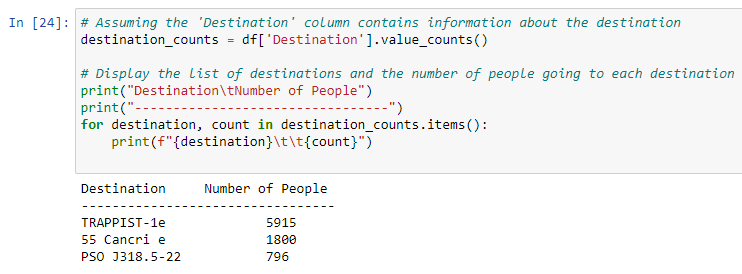
## What is the average amount spent at the foodcourt, ignoring people that spent 0 dollars?



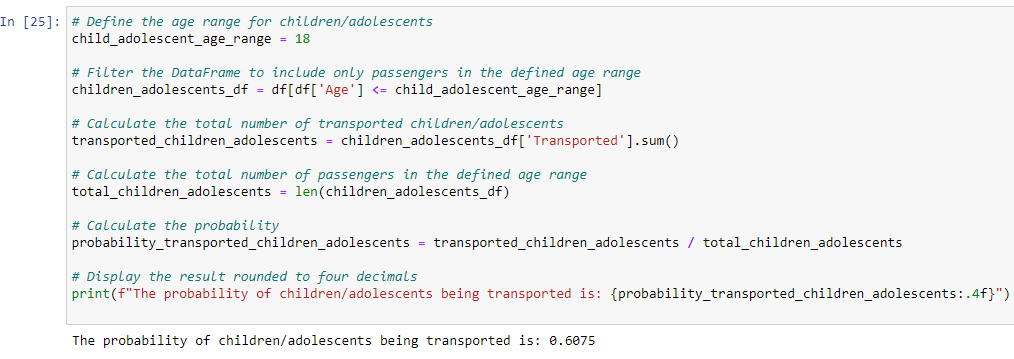
## What's the ratio of "Transported" (TRUE) to the total number? (Round up to four decimals)



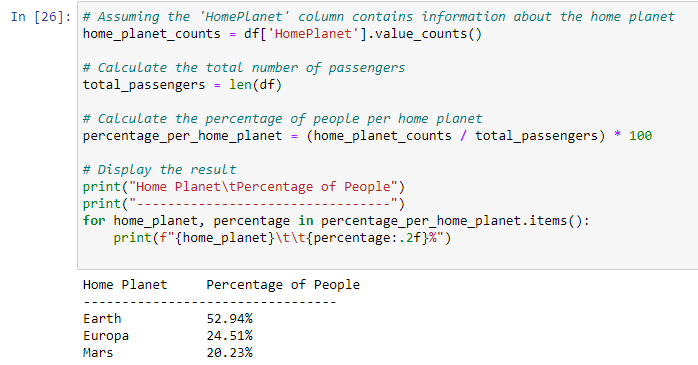
## List every destination and the number of people going.



## What is the probability of children/adolescents being transported?

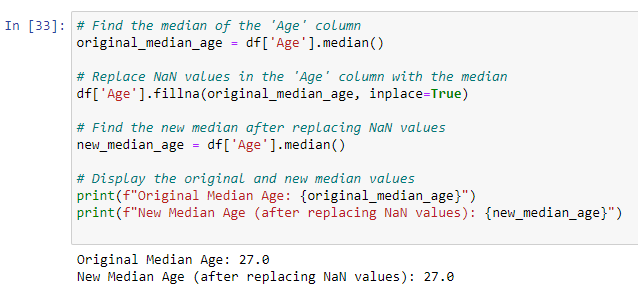


## What is the percentage of people per homeplanet?

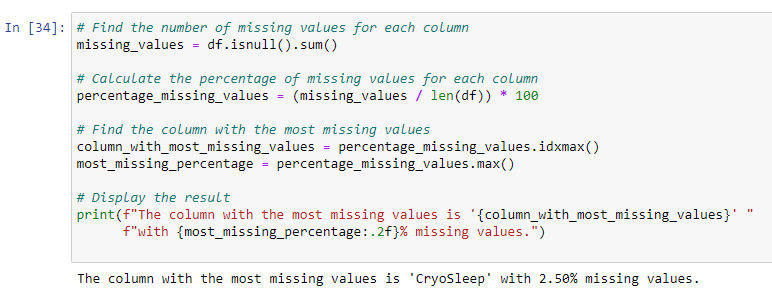


## What is the correlation coefficient (the 'r' value) of those who are "VIP" and if they were "Transported"? What does this value mean? (Round up to four decimals)

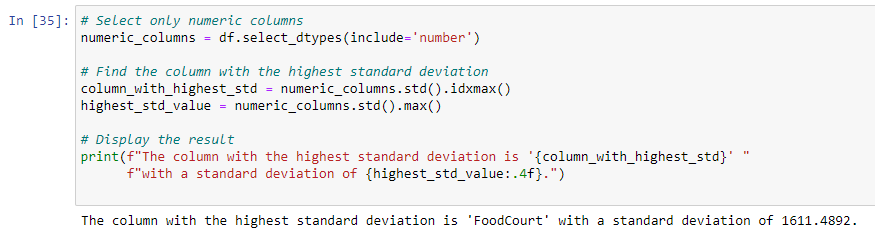
## Find the median of Age and replace all Nan values with the median. After replacing all Nan values in Age, find the new median value.



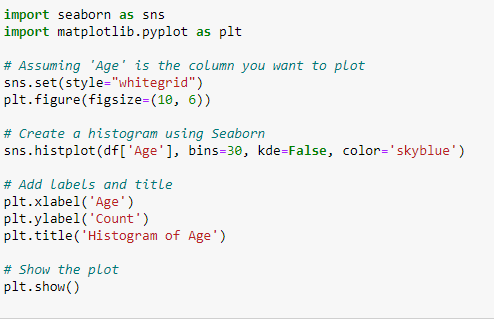
## Find the column with the most missing values and the percentage of its missing values. (Round up to 2 decimal places)

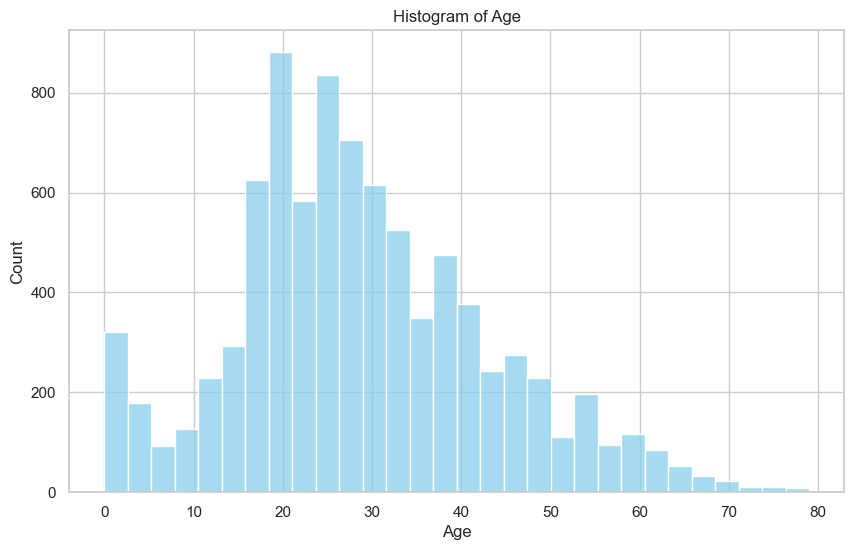


## Of the columns with numeric values, find the column with the highest standard deviation and that value. (Round up to four decimals)

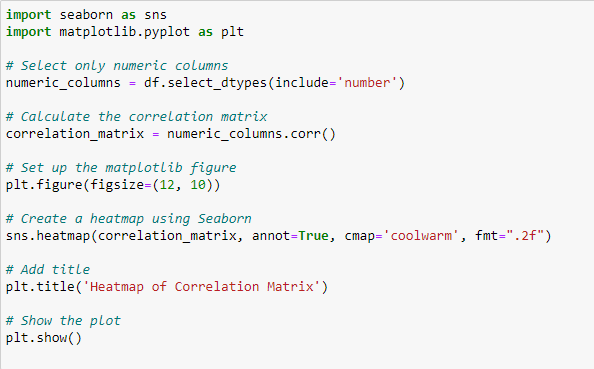


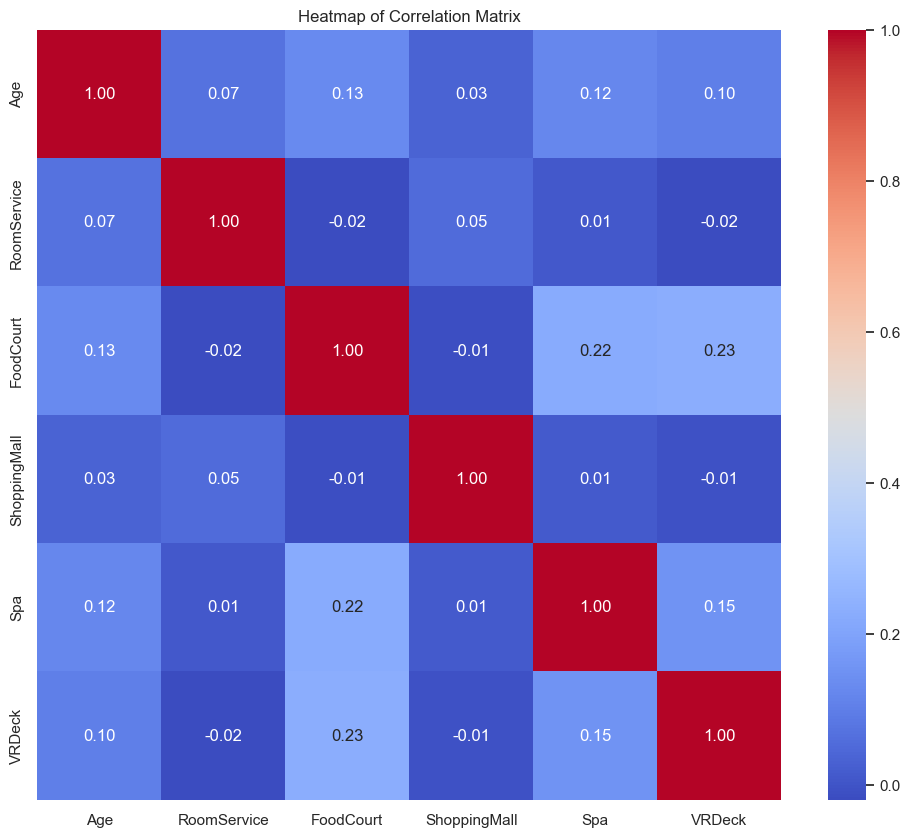
## Create a histogram with Age on the x-axis and Count on the y-axis. (Recommended: Use the Seaborn Library)



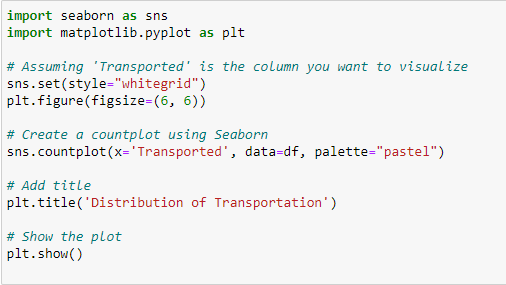


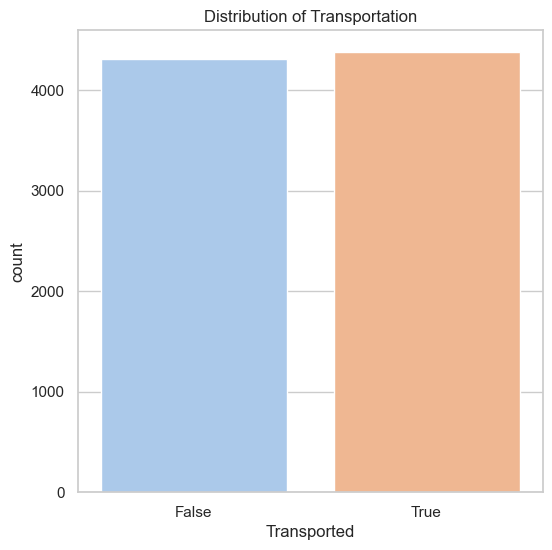
## Create a Heat Map of all the columns with each other. (Recommended: Use the Seaborn Library)



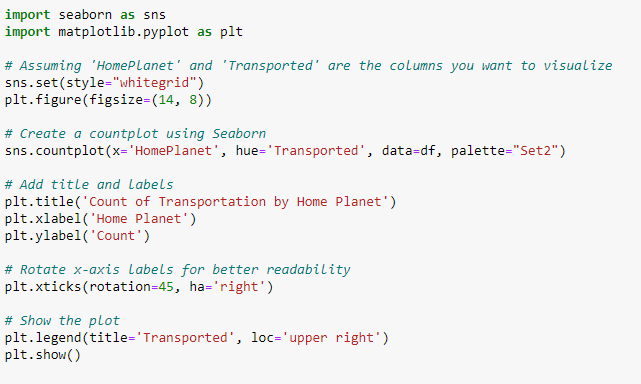


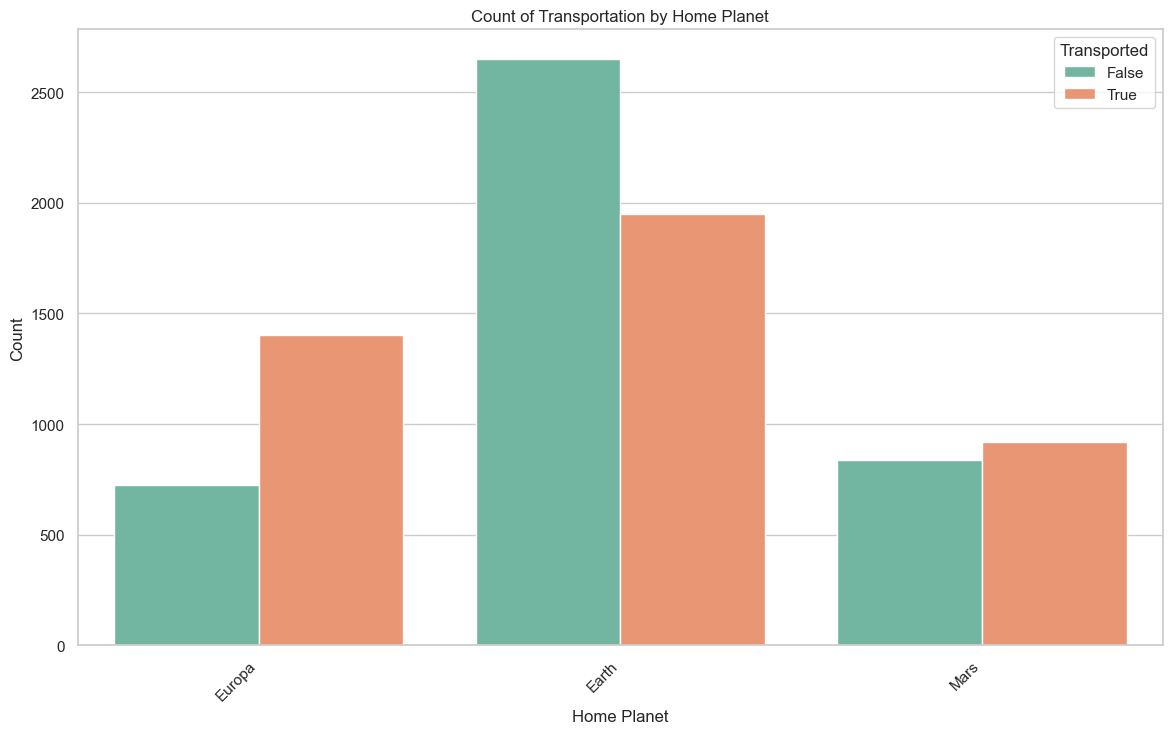
## Create a pie chart to see the percentages of those transported and those not. (Recommended: Use the Seaborn Library)



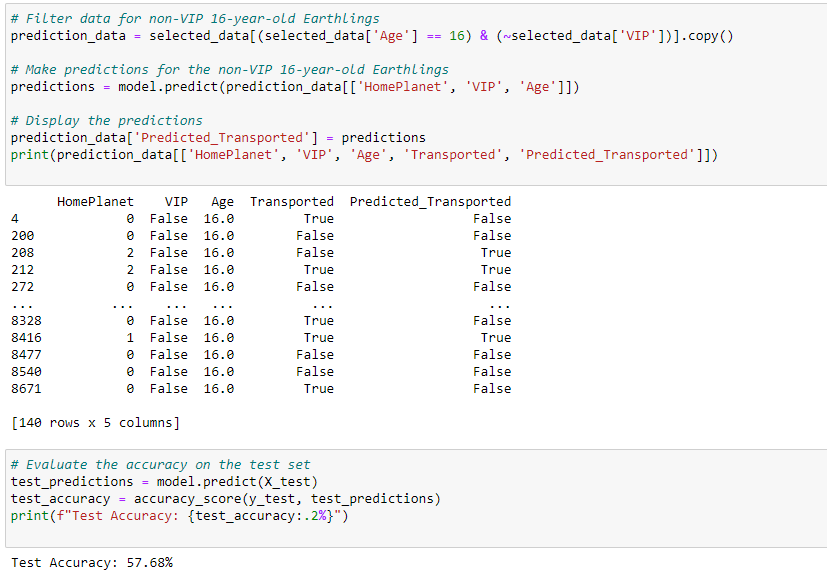


## Create a bar graph of all the planets with the count of those who were and were not transported. (Recommended: Use the Seaborn Library)





## Predict if a non-VIP 16 year old Earthling will be transported (You can use linear regression with features: Home Planet, VIP, Age)



## Predict if a non-VIP 32 year old Europi will be transported. (Use a decision tree classifier with x1 HomePlanet, x2 = VIP, x3 = Age)

