NAME:SUNKESULA SHAHAN
Email:shahan0805@gmail.com
Mobile:9618497495

Asbl assignment report:

Analysis :

Loaded the dataset in to the colab browser and performed eda after data cleansing and column conversion.found out the missing values and replaced them with mean and mode values as per column dtype requirement.plotted box plots to check for outliers.spotted plenty of outliers but cannot apply outlier treatment due to realization of unrealistic values after outlier treatment.instead applied mean and mode to replace the missing values in the rows and columns for which it is required.There are different types of data type columns in the dataset for which I have performed the necessary data conversions .so, we have run the code and derived boxplots,correlation tables and correlation plots for the data set after performing EDA.Necessary graphs have been plotted and error warning should be ignored.Outlier treatment can lead to unrealistic predictions which is not necessary but if needed can be done.Have applied for outlier treatment and removed outliers. we can see the plots after running the codes.converted the data set file from xls format to csv for usage in programming.saved the new processed dataset as newtrademill.csv.

Insights backed up by data and observations from the data:

There are more number of TM1 product users in the data set compared to other product users.There are more number of treadmill usages on the TM3 product than compared to TM1 and TM2.from going through the data we can understand that more fit users are using TM3 and TM2 products than TM1.by observing the data we can see that not relevant correlation exists between the user running miles per week.it can be understood that it is based on user choice.Fitter users run more miles than the less fitter ones according to the data in the dataset.

Recommendations:

We can see there’s a scope for improvement by increasing the number of TM3 and TM2 products in the business which could reduce the number of times the treadmills are used a bit.considering the general mindset of a user he might be transiting to a more advanced treadmill product later after sometime.which raises the need for allocating new TM2 and TM3 products in the business as per user wants.An addition of better numbers of both TM2 and TM3 products in the existing business may be helpful in reducing the rush,demand and could make more users get accommodated to join the business in future.Even though more broader analysis can be done more data is needed than the present data set provides.

Steps To Run Code File:

Open google co-lab and run the file named “asbl.ipynb” that I am attaching you along with this report and run it in google colab.run the code one by one and ignore the attribute errors and check the plots.they are simple warnings which can be ignored.check all the code functions and execute the code til the end and we can see the plots obtained.so, as per the mail instructions I have done the required analysis with simple actionable recommendations that can be implemented for everyone to understand.

Link to open files:

https://drive.google.com/drive/folders/10W0rLK9gjGcvAeOy_oZKXOh8VfTc21mK?usp=sharing