Assignment 4 – Data ViZ.

## Import the attached Indian\_food.csv file in Jupyter notebook and perform following data cleaning operations through operations using Pandas and visualizations using Matplotlib, Seaborn and Plotly:

1. In the state column, replace the -1 with the mode of state column.
2. Whatever the state you got in previous question, fill the corresponding region in the region column.
3. Replace the -1 with mean for cook time variable.

After performing these data cleaning steps, perform the below visualizations. You can use any of the 3 libraries (Matplotlib, seaborn, Plotly):

1. Most preferred flavours by number of customers using pie chart.
2. Most requested course by customers using Bar chart. Print the labels also on the top of each bar. Then sort the bar chart in descending order.
3. Comparison between Vegetarians and Non-Veg orders in each region. *Hint*: You can use stacked/grouped bar chart for comparison.
4. Create the box plot for Order cooking time to see the distribution. Also check if there are any outliers in this column.



## Import the attached Billionaires\_Stats.csv file in Jupyter notebook and perform following operations using Pandas and Data visualization:

1. Create a dataframe which will have the billionaires whose age is greater than or equal to 30. Then create a scatter plot to show the relationship between age and net worth.
2. Create a histogram to see the distribution of ages.
3. Create a heatmap and see the correlation between all the numeric variables. Print the correlation value also in the heatmap.
4. Save the figure in your pwd generated through previous question.



1. Name any 5 types of charts you are familiar with (using any of 3 libraries).

Write the sample code also to create those charts using random data.