# First Data Science Boot Camp

# Feb 27, 2021

**Instructions:**

We have designed three data Science projects for you. These are very simple projects and these projects would give you enough hands on experience on the basics of data science.

First two projects are provided with the instructions and tell you what you have to do.

The third project is without any instructions and we want you to think about it that what information can you get out of this data. For this project imagine that some company hired you as a data scientist and gave you this data set and asked you do perform some analysis and let them know what information is hidden in this data.

**Project Submissions**

* You should have your GitHub account
* Submit your code on GitHub
* All the plots and graphs should be submitted on GitHub
* Submit a 2 paragraph summary for each project
* You have 1 week to finish the projects
* The completion certificate would be given to those who finishes the projects and submit on the GitHub
* Finally, you have to provide your GitHub link

**You can use Google search or any other resources to complete the projects.**

# Codes for Basic Concepts:

Read and repeat work on this code for practice:

1-ProjectBasics.ipynb

2-ProjectStats.ipynb

3-ProjectPlots.ipynb

# Project-1:

## What is the best-selling category?

• Visualize payment method and order status frequency

• Find a correlation between payment method and order status

• Find a correlation between order date and item category

• Find any hidden patterns that are counterintuitive for a layman

• Can we predict the number of orders, or item category or number of customers/amount in advance?

## Project1-DataSet-Pakistan Largest Ecommerce Dataset.csv

# Project-2: *Healthcare*

## What is the most prominent and relevant factor causing the stroke?

* Visualize Age and BMI frequency
* Find a correlation between Age and Gender
* Find a correlation between hypertension and Avg Glucose level
* Find a correlation between Marital status and Smoking status
* Find if there is any relation between work type and heart disease
* Find the frequency of Smoking status in the form of Histogram
* What is Min, Max and Avg glucose level
* Find the frequency of Males and females in the data
* Find if marital status is related to smoking status

For those who have done ML, DL course

Can you make a predictive Model using ML methods? … Stroke chances.

## Project2-DataSet.csv

# Project-3: Real Estate

Work on it to dig out the hidden information and facts to prove it that how can we use this data for business purpose like evaluation or prediction.

## Project3-DataSet-Real estate.csv