****

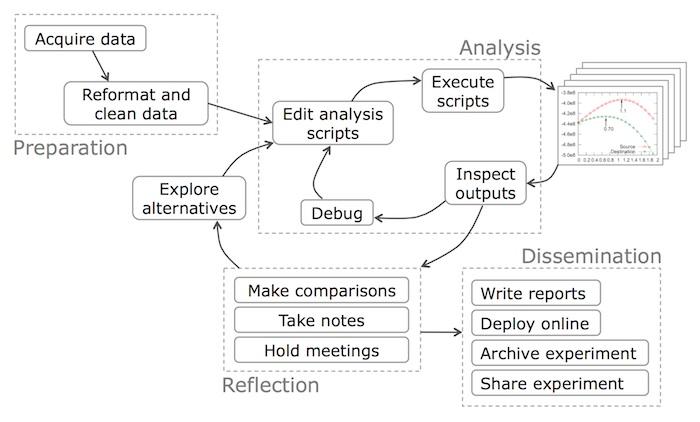
**Data Science Capstone Project Plan**

Created by - Nitin Kumar

Each student in the data science program will have to work on a 3 weeks capstone project. The project will be split into following phases

**Project Week - 1**

1. **Kickoff meeting** - This will be the first meeting between all the concerned parties (students, propulsion, company (if any)). After the meeting the student will have to focus on drawing a work plan (something similar to the image below). This should result in a plan for next week/s. This should include what tools and techniques you are planning to use.



Taken from https://cacm.acm.org/blogs/blog-cacm/169199-data-science-workflow-overview-and-challenges/fulltext

1. **Data exploration** - The first few days should be focused on understanding the data with the help of visualisation, define every row and column in your data, what do they mean, how do the distributions across features look like, what kind of techniques you are gonna try, how would the output look like e.g. numbers (accuracy, precision), interactive web application.
2. **Data cleaning** - Perform all the necessary data cleaning learnt in the data exploration step.
3. **First results (involving all concerned parties)** - This should include results from data exploration and cleaning, what has been done so far, what issues are you encountering, how are you planning to solve them, what is your outlook for next 1-2 weeks.

**Project Week - 2**

1. **More Analysis** - Create all necessary scripts for all the analysis, sequentially, separate different parts of analysis in different scripts. Always keep the your project questions in mind. Ideally you should be done with all the necessary analysis by end of this week.
2. **Second results** **(probably involving all concerned parties)** - Final results, your conclusions, how you are gonna present it, what is gonna happen with your work afterwards, what are you planning to do in week 3.

**Project Week - 3**

1. **Final testing** - Optimise all your code, clean it, organise it, finish your analysis and work on deployment.
2. **Deployment** - The deployment in your case might be a presentation, an interactive dashboard, a jupyter notebook report etc., etc. This will vary from project to project.