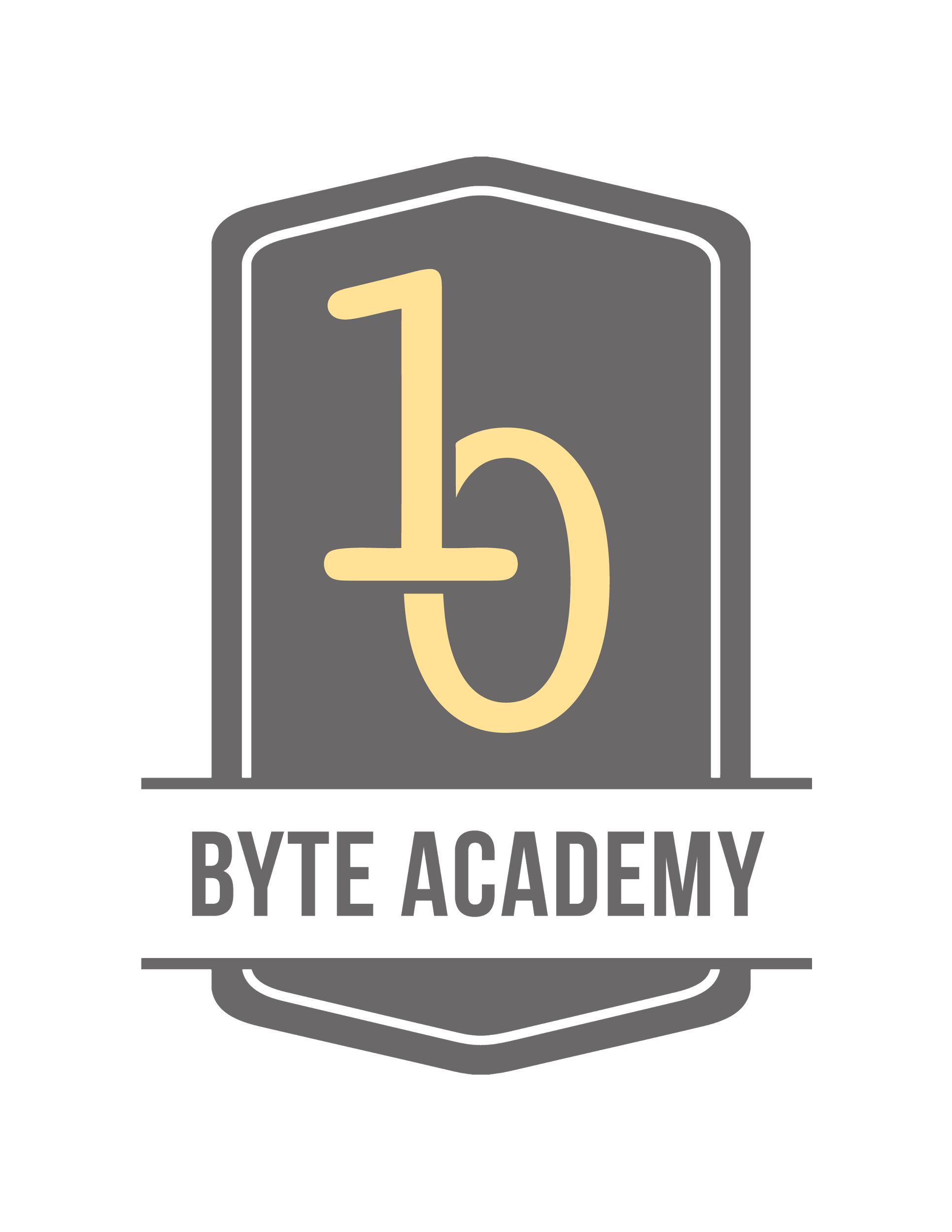
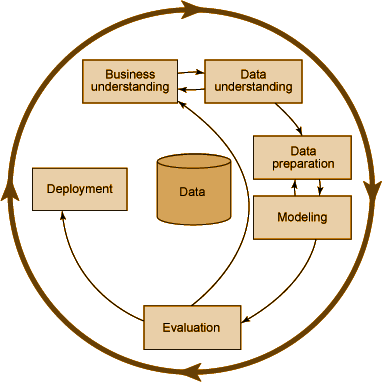
***Final Project Guidelines***

***Methodological Approach***

A common and classic approach to a Data Science project is the CRISP-DM model conceptualized by IBM. Most Data Science/mining projects follow a similar trajectory which can be summarized with the image below:

We recommend that students use the above methodological approach in the project. For further details: <https://www.sv-europe.com/crisp-dm-methodology/> 

***Presentation Guidelines***

*Golden Rules for any Data Science presentation:*

1. Every project should be presented in a PPT format. The idea is for the project to be presentable to individuals with both technical and non-technical backgrounds as if it were an interview or a Meetup.
2. No code snippets! : Your presentation is less an opportunity to demonstrate your ability to code and more a chance to show your *holistic thinking.* Code snippets should only be included if they further clarify your approach or results.
3. The presentation should last a maximum of **15 minutes.** Don’t be repetitive, be concise!
4. Avoid looking at the screen or your notes for too long, look straight and engage with audience!

*The powerpoint should contain the following sections:*

1. **Problem statement**: The problem statement should concisely describe the technical nature of your project.

Example: If your project is to predict crime rates across Bangalore, you may state: “To estimate the percentage of reported, Class I and II crimes for December 2017.”

2. **Hypothesis:** The hypothesis should describe your educated guess on various parameters of the project prior to model building.

Example: “I think that crime rates will be highest in East Bangalore” or “I think the most important predictor will be % of functioning street lamps”

3. **Assumptions:** State any assumptions made during data cleaning, feature engineering, or modelling.

Example: “I assumed that the data followed a normal distribution”

4. **Outline of approach**: Here, you should detail how you approached your project. This can be in the form of a timeline or can be simple bullet points.

Example:



5. **Results**: While the results must include standard reporting like, for example, the precision % for a classification problem. The results must be richened by contextualization and visualizations. This can be accomplished with comparing your results with other attempts at similar problems. Fundamentally, the audience must be able to understand your results in a wider context.

Example: “While my best RMSE for predicting crime rates was 23, similar studies were only able to achieve a minimum of 30!” or “While a study completed by the IIM Bangalore achieved better results, they applied deep learning with the help of high powered GPU

*Optional sections:*

1. **Actionable results**- If applicable, contextualize your results in a business setting.

Example: If your prediction for crime rates is that they will increase in December 2017, and this result is most correlated with the historical shortage of electricity and policemen in the month, you may suggest and calculate how hiring policemen may reduce crime by X%.

***Submission Rules***

1. All final projects should be uploaded in the Student Projects repository in the ByteAcademyIn organization on Github. Link: <https://github.com/ByteAcademyIn/Students-Projects>
2. While submitting, create a folder in the repository with the following name structure: FIRST NAME\_PROJECT NAME\_COHORT NAME
3. All files should be uploaded in the folder, including requirements.txt and setup.py files.