**Final Project Guidelines (indicative only)**

Step 1: Get your data approved

Step 2: Frame a business case (as strong as possible)

Step 3: List all assumptions

Step 4: Do EDA, Cleaning and Feature engineering

Step 5: Find one or more proxy for PII and treat it as PII in your analysis

Step 6: Apply DP on one or more column of your choice. Produce reasoning. Also build one DP Linear and non-linear model (two on dp data and two on non dp data)

Step 7: For classification: On non-DP data apply reweighing on one feature or one composite feature. Run it through a tree and linear model. Report before and after on multiple accuracy, Fairness metrics and cost.

Step 8: For classification: On DP data apply reweighing on one feature or one composite feature. Run it through a tree and linear model. Report before and after on multiple accuracy, Fairness metrics and cost.

Step 9. For Regression and classification: Take the above model (if exists) and apply ACF on DP data model and non DP data model on both Tree and Linear.

Step 10: Repeat step 9 with another residual formula. Be careful of your choice as it depends on the type of protected feature

Step 11. Report before and after on accuracy and fairness

Step 12. Choose the best two models and perform XAI on model, data and counterfactual (if business case makes sense)

Step 13. If there exists any more bias, try Reject Option Classifier

1. Create a 10-slider power point presentation as minimum text and maximum plots/graphs.
2. You will get 8 min to present as a group. Followed by questions.
3. You need to get part 1 and part two of your project verified on 6th (upto reweighting) and 8th (ACF) lecture

**Any kind of plagiarism (that was found in assignments) from internet or among your mates will lead to a heavy penalty. Please keep your work secret, original and exciting.**

PS: when in doubt, please ask. Let’s not Assume

All the best