FAQ HACKATHON RULES (TO JUMP TO GITHUB RELATED FAQ'S. CLICK HERE)

1. **What is the final deliverable of the python hackathon? What are we supposed to submit at the end of 1 week?**
   At the end of a week, we expect at least 1 Jupyter notebook containing great questions that help us understand the dataset and great answers to those questions(In Python). Any extra notebooks or links are at your discretion. But if you choose to submit more than 1 notebook, please give us a word document to help us decode what each of it is for.
   SDET's: Same rules as above, please submit at least 1 Jupyter notebook with your 2 categories. You can submit up to 20 questions in the descriptive stat category.

2. **How should the presentation be? Does it have to be in python?**
   There are 100 extra points available for being able to set up all your insights in a python dashboard(like streamlit or plotly grids, no need to deploy it or have it go live).But you don't have to do it this way, you can also choose to screenshot your best insights into PPT and present it.
   The hackathon awards maximum points for reasoning. You can lose 100 points and still win if you maximize other categories.
   SDET's: The rules of presentation apply to you, but you will not need to consider a dashboard. That category is not applicable to you.

3. **What are these categories? Do I fail if I can't complete every category?**
   Categories are just buckets for different types of questions. Some buckets have more value/more points than other buckets, but every bucket is optional.
   That said, category 3 is the one with the highest points and while you will not be disqualified for not attempting that category, it would be hard for you to place if you don't.
   Cat3 is also the most important one to practice at some level if your final goal is to be an analyst, so we strongly encourage attempting cat3 in whatever capacity you're able to.

4. **What if there is nothing in a category? For example, there is no cleaning to do?**
   It's possible that the right thing to do for a dataset, sometimes, is to leave it alone. If that's the case and you 'correctly' identified that there is no cleaning needed, write down your reasoning for why you think that is.

If you're right, you will get all 100 points from that category.
Knowing what to do is as important as knowing what not to.
SDET's: If there is no cleaning. Display your technical skills with data wrangling and manipulation skills and then proceed to descriptive statistics

5. **What if I don't know any ML and can't do category 4?**
   Category 4 is only worth 90 points. While it's true that you can create as many models as you like, it's also true that a model that solves nothing and is created for no reason except points, is going to be marked at 0. So it's more than likely that each team will only be able to build one good prediction. This amounts to 90/1160 points. It is not what will help you win/lose the hackathon. It exists to allow our more experienced participants a chance to practice this skill.
   If you can't do ML, move on and find some great insights in other categories. You can still win without it.

6. **What If we are able to come up with more questions than the recommended maximum?**
   Choose the best questions of all the ones you have and follow the rules. For example, if you have identified 25 descriptive statistics measures and the rules allow a maximum of 10, choose the best 10 from your 25 and upload only those for evaluation.
   We will only evaluate the first 10 in cat2/ the first 30 in cat3.
   The ability to decide what insights are important is a trait of a good analyst, so choose wisely.

7. **What If we can't come up with as many questions as allowed?**
   Remember, you're all working with the same dataset and it's likely there are only so many good insights that can receive points.
   So a team that presents some good insights that *can* be graded will receive the same points as a team that chooses to "fill" the category with questions that have no analysis or reasoning behind it.
   It's the reasoning that's being graded, not the number of questions.

8. **Does the final dashboard need to be live on Streamlit or Heroku or an external web app? What if I don't know these technologies?**
   Out of the 100 points available per dashboard, only 30 is for design and layout. We recommend keeping this simple.

A plotly/local streamlit dashboard within the Jupyter notebook has the same value as a very complicated streamlit deployment.
What we're grading is the intelligence/insights and the analysis behind the dashboard. Give us a great analysis and we'll give you all the points.

9. **Does the dashboard need to contain the packaged predictive model if we create one?**
The dashboard and predictive model can be separate. You may deploy to a dashboard if you have the time. But the dashboard only needs to contain a collection of your best solutions/insights. It doesn't have to be any more complicated than that. The models' code and reasoning can be separate. The model will be scored on the basis of parameters chosen and accurate hypothesis, not on your ability to deploy it under a time constraint.

GITHUB FAQ

1. **Is cloning mandatory?**
   Yes. Cloning is mandatory. Each team member must clone the repository and push their contributions. Direct uploads are not allowed.
2. **Are branches required?**
   No. Branching is encouraged but not required. Teams can work directly on the main branch.
3. **How can I ensure my commits are recognized?**
   - Commit regularly with descriptive messages
   - Push your commits to the team repository
4. **Can one team member upload all files?**
   No. All members must commit their own work to demonstrate contribution.
5. **Should the entire analysis be in one notebook?**
   Not necessary. Organize into multiple clean notebooks or scripts for clarity.
6. **What should the README include?**
   - Description about the dataset
   - Problem statement
   - Methods and analysis
   - Key findings
   - Team member names