SHOPEE CODE LEAGUE 2020

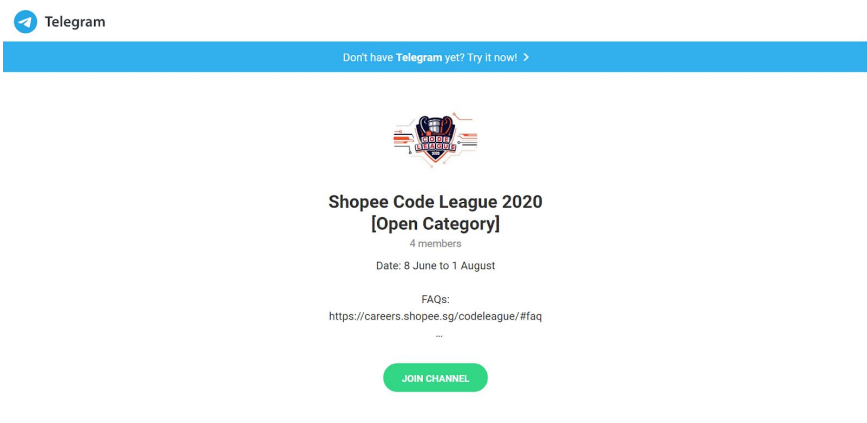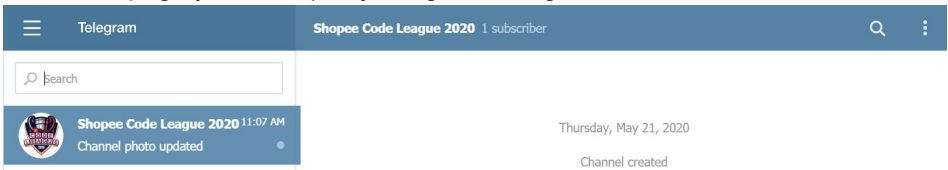**KAGGLE COMPETITION GUIDE**

Getting Started

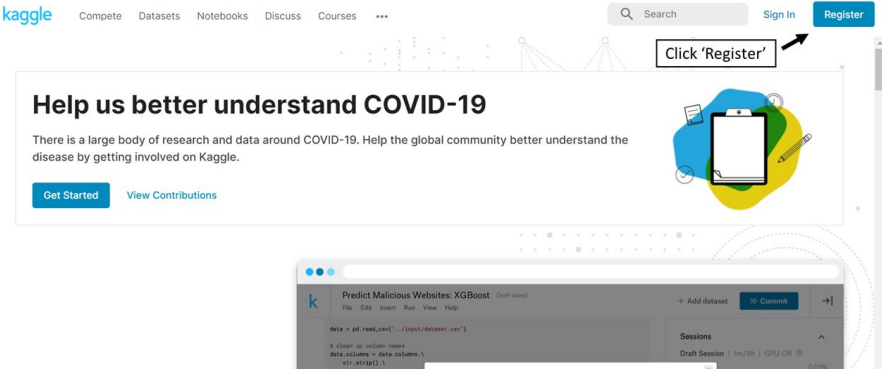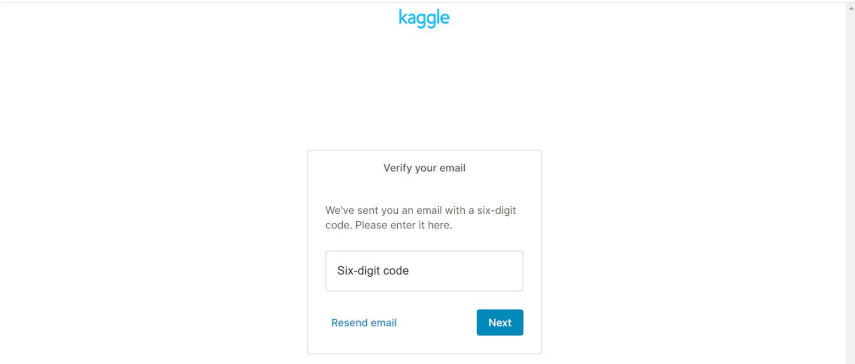To get started, please register on Telegram & Kaggle prior to the start of the competition.

The competition link will be released on the Zoom Webinar Chat real-time. (There is a briefing session on Zoom 1 hour before the start of each competition. The link to access the Zoom webinar is released on our Weekly Announcement Newsletter.) The competition link will also be released on the Telegram channel thereafter on the day of the challenge.

(1) **[Compulsory]** Set up your Telegram Channel to receive important updates and announcements.

Steps	Instructions
1	**Download the app on your mobile** For iOS users: https://apps.apple.com/us/app/telegram-messenger/id686449807 For Android users: https://play.google.com/store/apps/details?id=org.telegram.messenger&hl=en_SG
2	**Enter verification code sent to your mobile**
3	**Access Telegram Web on your Laptop:** https://web.telegram.org/#/login
4	**Join the Shopee Code League 2020 Telegram Channel** **Open Category:** https://t.me/joinchat/AAAAAE1a4Q3cv3ECS3CHhg **Students Category:** https://t.me/joinchat/AAAAAFZlqDaDyG9tgmIuOg *Please ensure you join the channel of your category (i.e. Open **OR** Student Category)

*Note: The invite link would bring you to this page. Click "Join Channel". |
| **5** | **Successful Joining**
This should be the page you see upon joining the Telegram Channel.
 | |

(2) **[Compulsory]** Set up and register for your Kaggle account.

Steps	Instructions
1	**Register for a Kaggle Account** Go to Kaggle homepage: https://www.kaggle.com/ and click register.
2	**Once you've registered, you will see this page. Once you do, check your email for the verification code.**
3	**Explore the Dummy Competition Page** https://www.kaggle.com/t/18b1475a9f3a4673affb8c73b7fa2757 After joining the Dummy Competition, follow Annex C to form a team and try submitting your answers as a team.

Competition Rules

Shopee reserves the right to disqualify any Participants and/or teams at its sole discretion if such Participants and/or teams, where applicable is/are in breach of these Competition Terms and Conditions.

- Participants may use any language they prefer to solve the problem statements - Python, Java, C++, etc
- Participants are allowed to use Internet search engines such as Google during the competition
- Participants are not allowed to enlist remote support from individuals outside of their teams
- Teams are not allowed to collaborate with each other during the course of the competition, this includes any form of discussion, sharing of notebooks between teams, etc.
- Only submissions made within the time limit for each round will be considered valid, late submissions will not be taken into consideration
- Participants may only compete in the category that they have registered for
- Team names on Kaggle must be the same as the ones used to register for the competition. Only submissions under registered team names will be considered valid

Important: The top 100 ranking teams based on the private leaderboard, must submit their complete scripts within 30 mins after each challenge via email to techsg@shopee.com. Requirements for the submission are detailed in Annex E of the Competition Guide.

Note that these rules are non-exhaustive, for more information please refer to the Terms and Conditions on the Competition website.

Further Information

The following Annexes are enclosed for your information:
- Annex A - Joining Telegram Channel
- Annex B - Setting up a Kaggle account
- Annex C - Team Formation
- Annex D - Discussion Board Guidelines
- Annex E - Kaggle Guide & Submission Requirements
- Annex F - Common Issues during Submission
- Annex G - Determining your Scores
- Annex H - Common Kaggle Issues

JOINING TELEGRAM CHANNEL

Download Telegram for iOS or Android
Register for an account
Join the Shopee Code League 2020 Telegram Channel
Click Open in Web to join the channel

◢ Telegram

Don't have **Telegram** yet? Try it now! ❯
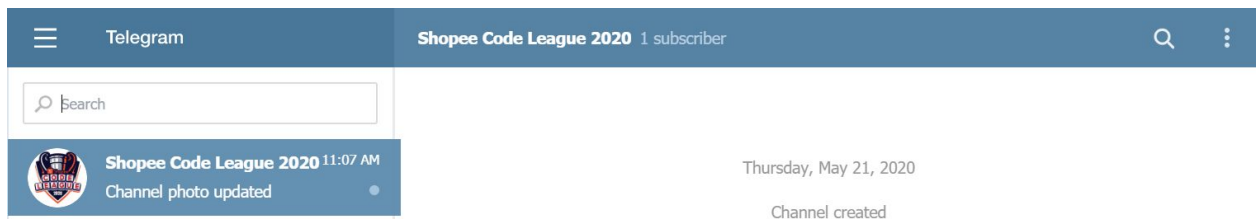
Shopee Code League 2020
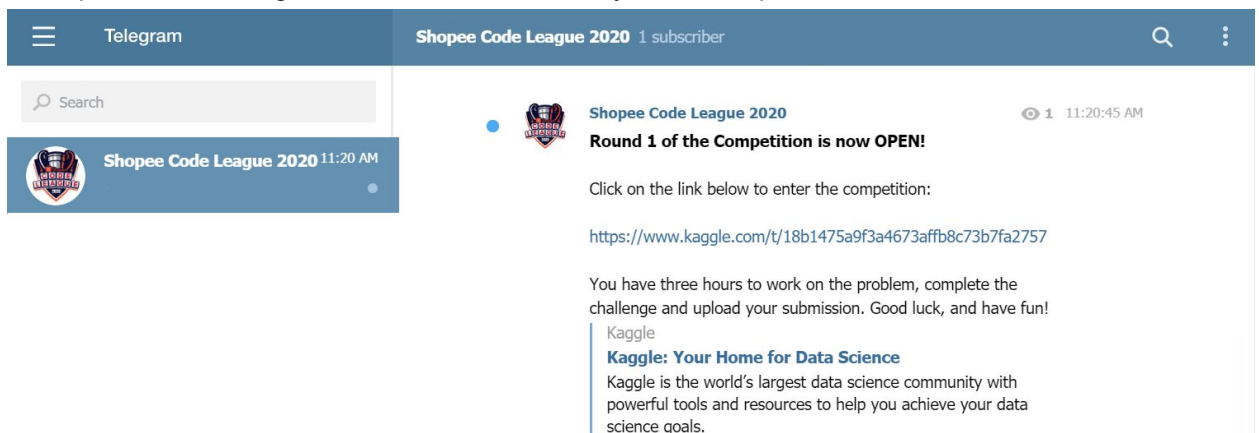1 member
Stay tuned for updates on our challenges!

JOIN CHANNEL

OPEN IN WEB

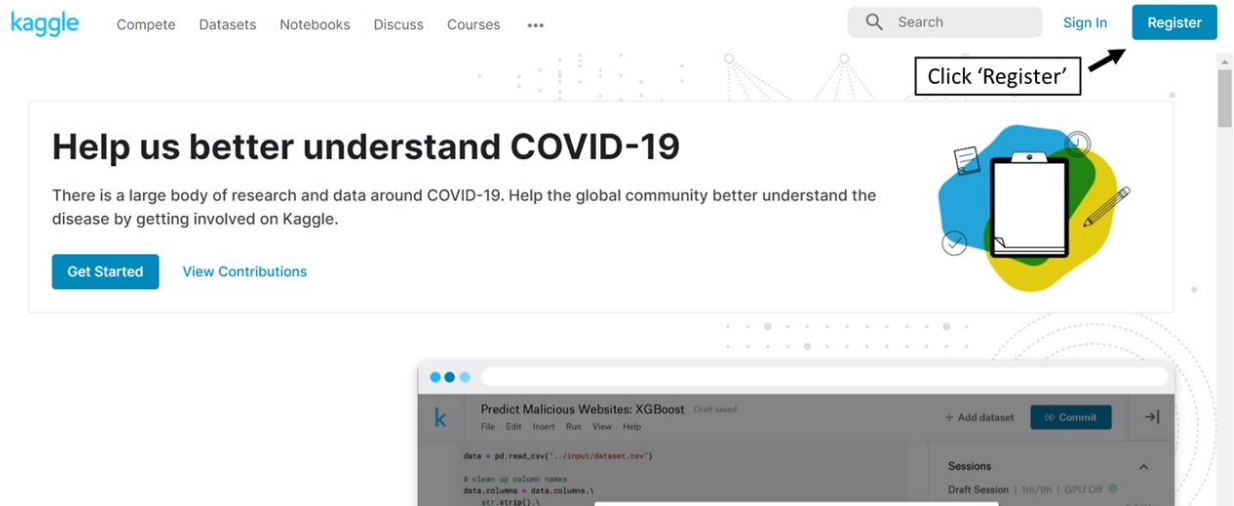This should be the page you see on successfully joining the Telegram Channel.

≡   Telegram   Shopee Code League 2020 1 subscriber   🔍   ⋮

🔍 Search

Shopee Code League 2020 11:07 AM
Channel photo updated   Thursday, May 21, 2020

                                                        Channel created

A sample of the message that will be sent on the day of the competition.

≡   Telegram   Shopee Code League 2020 1 subscriber   🔍   ⋮

🔍 Search

Shopee Code League 2020 11:20 AM   Shopee Code League 2020   👁 1 11:20:45 AM
                                         Round 1 of the Competition is now OPEN!

                                         Click on the link below to enter the competition:

                                         https://www.kaggle.com/t/18b1475a9f3a4673affb8c73b7fa2757

                                         You have three hours to work on the problem, complete the
                                         challenge and upload your submission. Good luck, and have fun!
                                         Kaggle
                                         Kaggle: Your Home for Data Science
                                         Kaggle is the world's largest data science community with
                                         powerful tools and resources to help you achieve your data
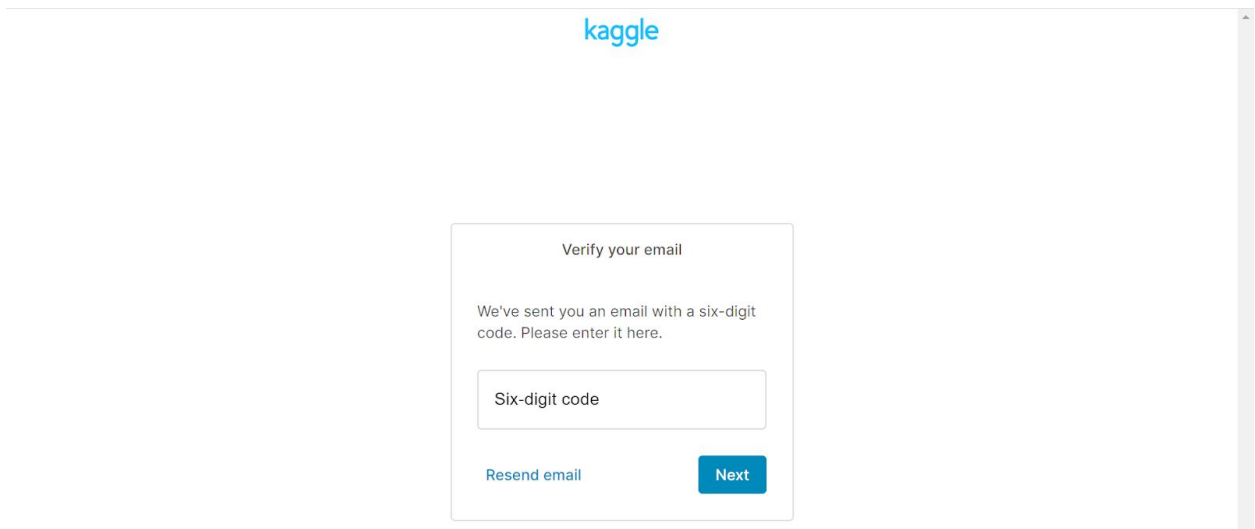   science goals.

SETTING UP A KAGGLE ACCOUNT
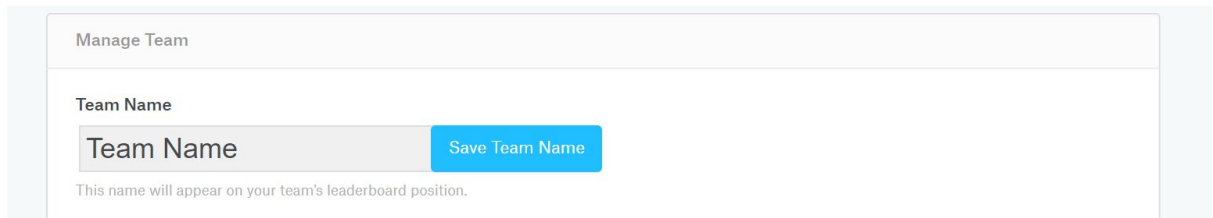
1) Go to the Kaggle homepage
2) Click Register

3) Register with your preferred method and fill in the relevant details
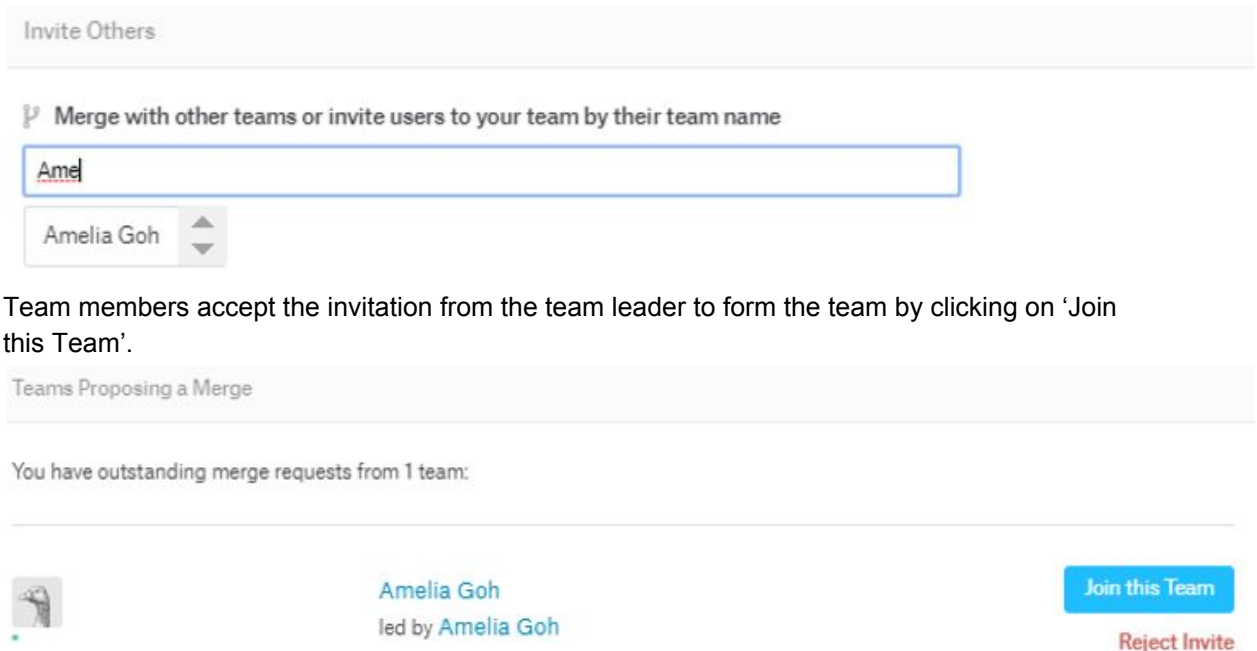4) Once you see this page, check the email you've signed up with, for the verification code

TEAM FORMATION

.

1) The team should form their team on the "Team" tab (In the main page of the competition, the tabs are: Overview, Data, Kernels, Discussion, Leaderboard, Rules, Team).

2) Team leader to enter their registered team name in the 'Team Name' field. **Please ensure that the team name on Kaggle is the same as the team name you registered for via the Shopee Code League website. Failure to comply will result in disqualification from the respective Competition.**

3) Under "Merge with other teams", team leader enter the name of your team members. Select the name from the drop-down list and click on 'Request Merge'. (Note: The team leader can only send a Merge Request to one member at a time. Each member must accept the merge request before the leader can add another member to the team)

4) It is recommended that team members choose names that are specific and unique (e.g "Amelia Goh", instead of "Amelia" or "A") to ensure that Team Leaders add the correct members to their teams.

5) **Please ensure that all team members have registered and joined the competition prior to setting up the teams.** Otherwise, the member's name will not appear in the drop-down list.


6) Team members accept the invitation from the team leader to form the team by clicking on 'Join this Team'.

7) Team leaders are to check that they have added the right members and team members are to verify the invitation before accepting. Do note that **teams cannot be changed once formed.**

DISCUSSION BOARD GUIDELINES

All technical questions related to the problem statement are to be posted on the Discussion Board. Please DO NOT email techsg@shopee.com on any queries related to the problem statement.

1) Click on the Discussion tab to post problem statement related questions on the Discussion Board.
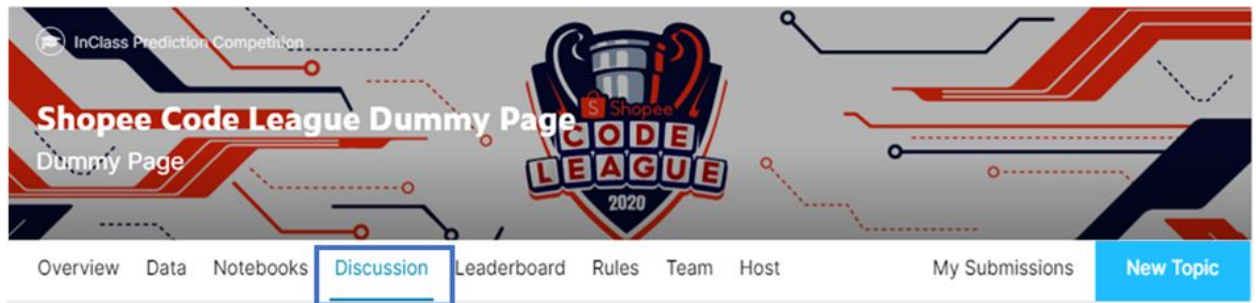
2) Include the Topic Title and Content in your post
   a) Write a title that summarizes the specific problem, do not use generic titles like "Question" or "I need help".
   b) See this link for reference on posting a good question

3) Do ensure you have read through the other topics already posted on the page to see if the question has already been answered before posting a new question.
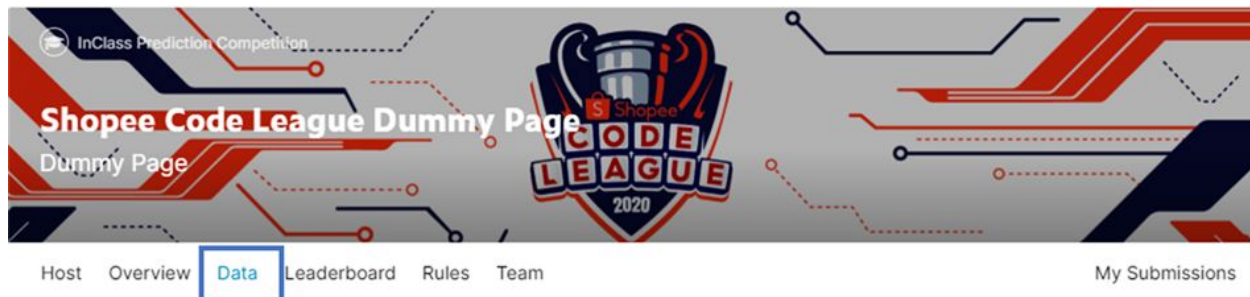
KAGGLE GUIDE & SUBMISSION REQUIREMENTS

Teams are allowed to upload a maximum **combined total of 20 submissions** for each competition. Additional submissions beyond this number will be automatically denied by Kaggle.

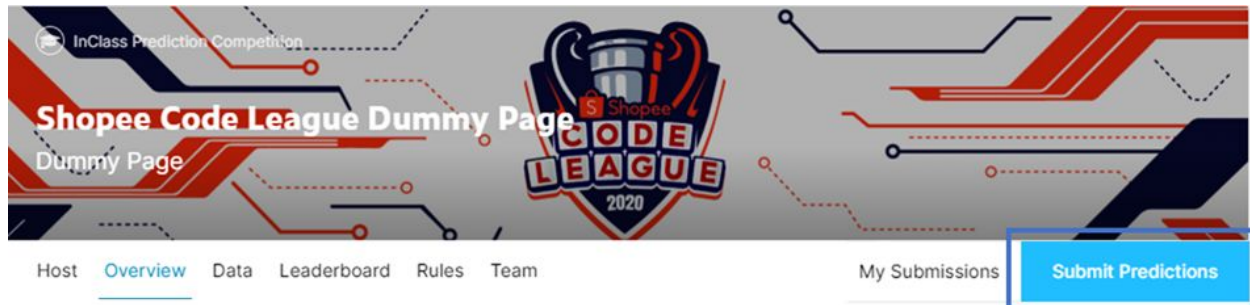1) Read and understand the problem statement and submission requirements in the Overview Tab
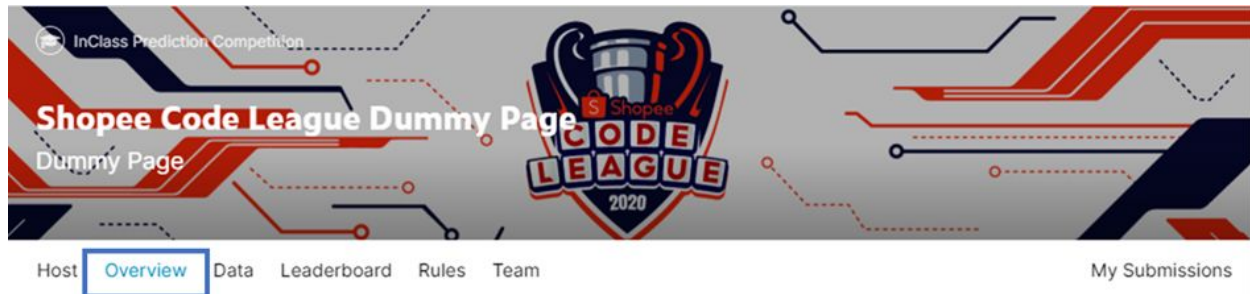
2) Download the data required in the Data tab

3) Submit Predictions
   Please upload in .csv file on Kaggle.

Step 1: Click on Submit Predictions

Step 2: Upload the required csv file and click 'Make Submission'. **Ensure that your submission meets the requirements before you submit your file.**

Step 1
Upload submission file

File Format
Your submission should be in CSV format. You can upload this in a zip/gz/rar/7z archive, if you prefer.

Number of Predictions
We expect the solution file to have 3 prediction rows. This file should have a header row. Please see sample submission file on the data page.

Note: The Top 100 teams based on the Private Leaderboard will be required to send a copy of their script within 30 minutes at the end of each challenge. Any libraries/modules/packages required to run the script must be clearly specified in the script. Email the script to techsg@shopee.com with the subject line: *Category - Team Name (Rank) - Challenge Date* The name of the script should be the same as the subject line. A sample is shown below:
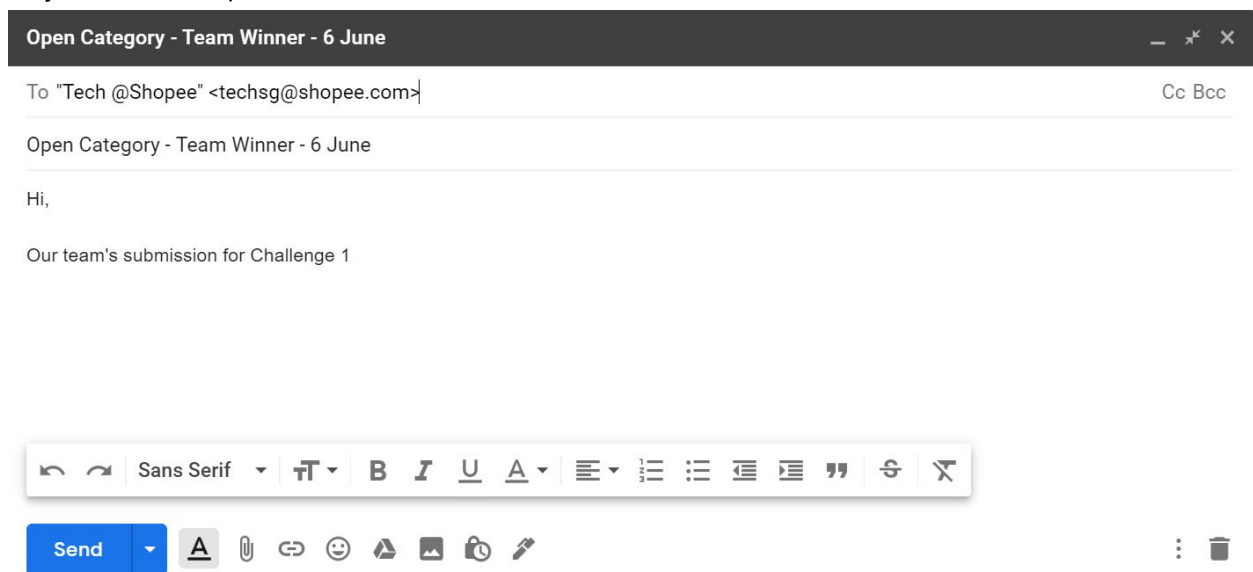
Only one submission per team is required.

Scripts sent have to be **complete** - it should run from start to end without errors and output the required .csv file of predictions for the competition.

COMMON ISSUES DURING SUBMISSION

These are a list of common issues during submission:

1) Wrong Headers

In this example, the required column header was 'new_number'. However, this submission labelled the header as 'newnumber'. Kaggle would show this error when trying to submit a solution with wrongly labelled headers.

2) Missing Rows

In this example, the required number of rows is 101, excluding the header row. However this submission only contained 100 rows instead of 101 rows (excluding the header row). Kaggle would show this error when trying to submit a solution with the wrong number of rows.

3) Additional Columns

In this example, the required columns are 'id' and 'new_number'. However this submission contains an additional column labelled 'another_column'. Kaggle would show this error when trying to submit a solution with an additional column.

Your most recent submission

Name	Submitted	Wait time	Execution time	Score
Additional Column.csv	just now	0 seconds	0 seconds	

Failed

ERROR: Column 'another_column' was not expected (Line 1, Column 15)

Wrong

index	newnumber	another_column
0	1	4543
1	2	34323
2	3	2432

Correct

index	newnumber
0	1
1	2
2	3

4) Wrong Index

In this example, the required index (or 'id' column) is a series of numbers from 0-100. However this submission labels the index as 1-100. Kaggle would show this error when trying to submit a solution with the wrong index specified.

Your most recent submission

Name	Submitted	Wait time	Execution time	Score
Wrong Index.csv	just now	0 seconds	0 seconds	

Failed

ERROR: Unable to find the required key value '0' in the 'id' column

Wrong

id	newnumber
1	2
2	3
3	4

Correct

id	newnumber
0	1
1	2
2	3

5) Duplicated Index Value

In this example, all values in the 1st column need to be unique. However, this submission has duplicate values in the 1st column. Kaggle would show this error when trying to submit a solution with any duplicate values in the 1st column.

Failed

ERROR: The value '1' in the key column 'stmt_id' has already been defined (Line 3, Column 1)
ERROR: The value '1' in the key column 'stmt_id' has already been defined (Line 4, Column 1)

Wrong

| stmt_id	ckt_id
1	1234
1	4567
1	7890
4	9012

Correct

stmt_id	ckt_id
1	1234
2	4567
3	7890
4	9012

Participants are reminded to read the problem overview and submission requirements carefully before attempting the question. No additional time will be given to teams who face problems in submitting their solution.

Important note: While Kaggle provides the Notebook feature for participants to use in order to collaborate on the problems, we **WILL NOT** be providing any support to participants who may face issues in its use. Participants are also reminded it is their responsibility to set their sharing settings on these notebooks to private to prevent others from copying their code.

DETERMINING YOUR SCORES

1) Overall Rankings for each challenge will be determined by your score under the **Private Leaderboard** page.
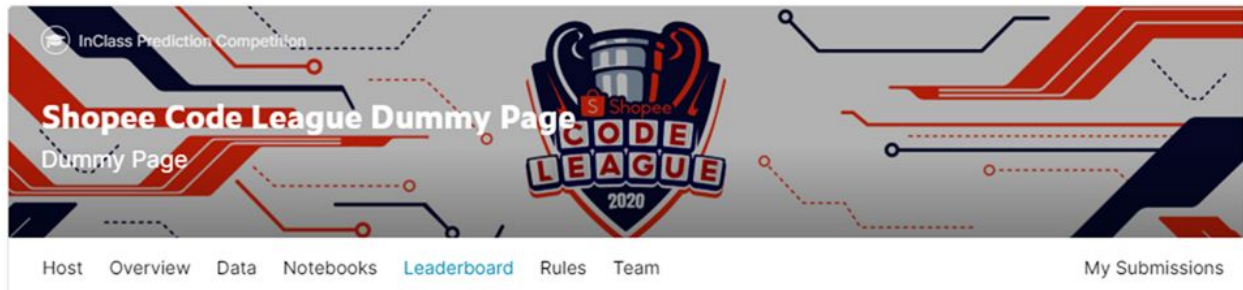
2) Regardless of the split between the Public and Private Leaderboards, only the scores on the **Private Leaderboard** will be taken into consideration for calculation of points.
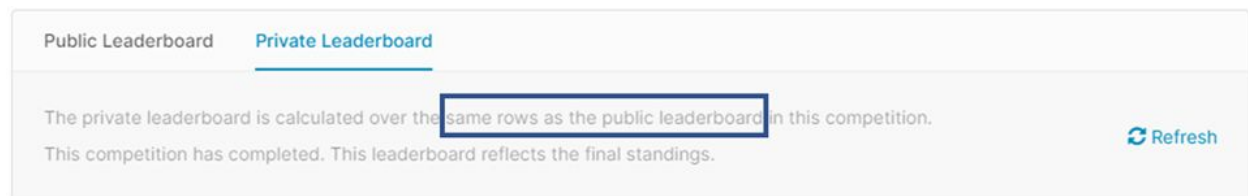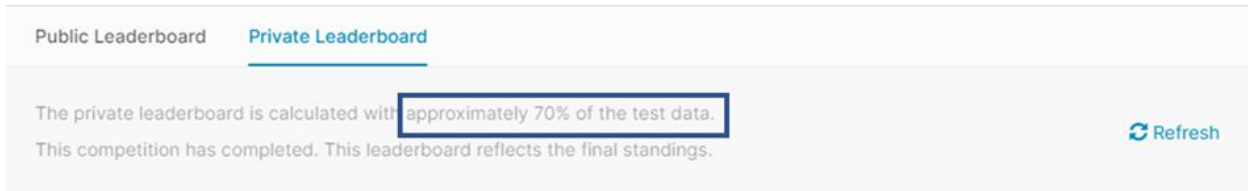
The split of the private and public leaderboard differs for each competition. In general, competitions which involve the creation of models will utilise **70%** of the test data to prevent overfitting. While competitions which do not require creation of models will utilise **100%** of the test data. Please find the information below:

#	Competitions	Type	Scoring on Leaderboard
1	Order Brushing	Data Analytics	Private leaderboard is calculated with **100%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.
2	Product Detection	Data Science	Private leaderboard is calculated with approximately **70%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.
3	Title Translation	Data Science	Private leaderboard is calculated with approximately **70%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.
4	Logistics	Data Analytics	Private leaderboard is calculated with **100%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.
5	Sentiment Analysis	Data Science	Private leaderboard is calculated with approximately **70%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.
6	Marketing Analytics	Data Analytics	Private leaderboard is calculated with approximately **70%** of the test data. Only the Scores on the Kaggle private leaderboard will be taken into consideration for the calculation of points for the Shopee Code League leaderboard.

Example

Each of the competitions has a different amount of points awarded for different tiers of scores on the Private Leaderboard based on challenge difficulty.

In this example the top 5 teams in this challenge have a score of between 0.91577 to 0.91802.

#	Team Name	Notebook	Team Members	Score ❓	Entries	Last
1	HengCherKeng		🐸	0.91802	105	6h
2	vladimirsydor[ods.ai][AfterPa...			0.91735	72	20h
3	Как закалялась сталь			0.91655	28	1h
4	MPWARE			0.91580	101	15h
5	[ods.ai] resnet34 is all you need			0.91577	84	4d

Assume the following point system for this competition is:

Tiers	Points
0 to 0.25	5
>0.25 to 0.5	10
>0.5 to 0.75	15
>0.75 to <1	20
1	25

As the scores for the top 5 teams in this example are all approximately 0.91 (2 decimal places), they will all be awarded 20 points for this competition.

COMMON KAGGLE ISSUES

Should you come across any of these issues while attempting the competition, please refer to the guide below on the necessary actions to take.

#	Issue	Solution
1	Accessing the Page ● Too many requests error ● Page not loading	1. Try refreshing the page again at a later time, Kaggle may be experiencing issues with multiple participants trying to enter the page at the same time 2. Check if your teammate is able to access the page, ask your teammate to send you a copy of the problem statement and data
2	Formation of Teams ● Team leader not able to invite team members ● Unexpected token e in json	1. Trying clearing your browser cookies and cache and try again 2. Get a different team member to form the team, at least one person has to set the registered team name. 3. Email techsg@shopee.com to alert us of this issue with the subject: Team Formation Issue (Your Team Name). a. Include your members details and your current registered team name **BEFORE** the end of the competition
3	Downloading Dataset ● Cannot download dataset	1. Check if your teammate is able to access the page, ask your teammate to send you a copy of the problem statement and data
4	Uploading Solutions ● Scoring takes a long time after making a submission on Kaggle	1. Please be reminded to **make a submission early**, at least 15 mins before the end of competition 2. If your team is unable to upload the solutions directly on Kaggle, please email techsg@shopee.com with the email subject: Submission Issue (Your Team Name): a. Include your members details and your current registered team name b. Only emails received **BEFORE** the end of the competition deadline will be taken into consideration