



SHOPEE CODE LEAGUE 2020

KAGGLE COMPETITION GUIDE (PRODUCT TITLE TRANSLATION)

Submission Opens: 11 July 2020, 12pm (GMT+8) onwards

This challenge utilises “Bleu score” a metric currently not supported on Kaggle. As such, submission via Kaggle’s In-Class platform is unavailable for this challenge.

There will be no submission from 4 July to 10 July 2020.

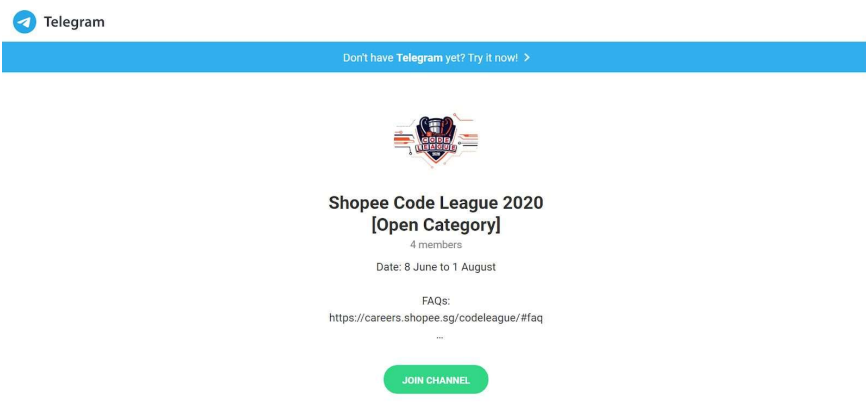
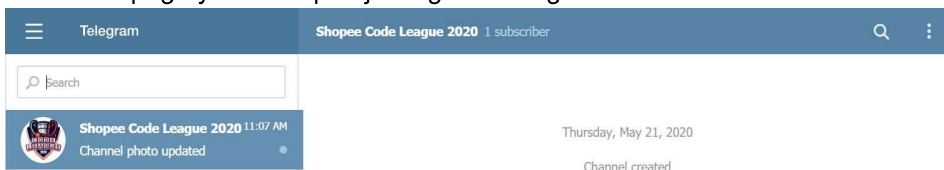
Important: Please kindly note the changes in the submission requirements and leaderboard in Annex C, E, F & G. The rest of the guide remains the same.

Getting Started

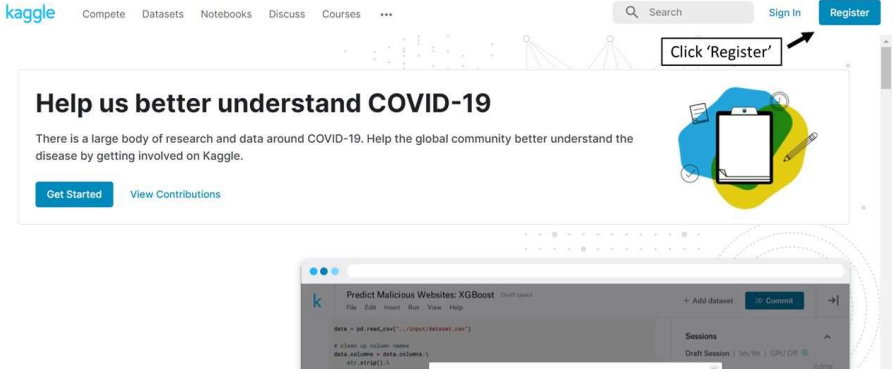
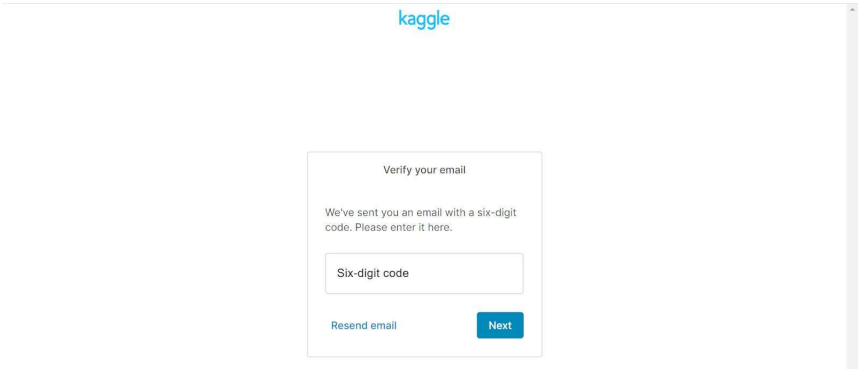
To get started, please register on Telegram & Kaggle prior to the start of the competition.

- (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	<p>Join the Shopee Code League 2020 Telegram Channel</p> <p>Open Category: https://t.me/joinchat/AAAAAE1a4Q3cv3ECS3CHhg Students Category: https://t.me/joinchat/AAAAAFZlqDaDyG9tgmluOg</p> <p>*Please ensure you join the channel of your category (i.e. Open OR Student Category)</p>  <p>*Note: The invite link would bring you to this page. Click “Join Channel”.</p>
5	<p>Successful Joining</p> <p>This should be the page you see upon joining the Telegram Channel.</p> 

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

Steps	Instructions
1	<p>Register for a Kaggle Account Go to Kaggle homepage: https://www.kaggle.com/ and click register.</p>  <p>The screenshot shows the Kaggle homepage with a navigation bar at the top containing links for 'kaggle', 'Compete', 'Datasets', 'Notebooks', 'Discuss', and 'Courses'. On the right side of the navigation bar are 'Sign In' and 'Register' buttons. A callout box with an arrow points to the 'Register' button, labeled 'Click Register'. Below the navigation bar is a large banner titled 'Help us better understand COVID-19' with a subtext 'There is a large body of research and data around COVID-19. Help the global community better understand the disease by getting involved on Kaggle.' and two buttons: 'Get Started' and 'View Contributions'. Below the banner is a preview of a Kaggle notebook titled 'Predict Malicious Websites: XGBoost'.</p>
2	<p>Once you've registered, you will see this page. Once you do, check your email for the verification code.</p>  <p>The screenshot shows the Kaggle email verification page. It has the Kaggle logo at the top. In the center, there is a box titled 'Verify your email' with the text 'We've sent you an email with a six-digit code. Please enter it here.' Below this text is a text input field labeled 'Six-digit code'. At the bottom of the box are two links: 'Resend email' and a 'Next' button.</p>

Competition Rules (Data Analytics)

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 after each challenge via email to techsg@shopee.com. Requirements for the submission are detailed in [Annex E](#) of the Competition Guide.

- Data Analytics Competition submission, 30 minutes after the end of competition.
 - [13 Jun] Order Brushing
 - [11 Jul] Logistics
 - [25 Jul] Marketing Analytics
- Data Science Competition submission, by 11am (GMT+7) / 12pm (GMT+8) the next day.
 - [20 Jun] Product Detection
 - [4 Jul] Title Translation
 - [18 Jul] Sentiment Analysis

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

Competition Rules (Data Science)

One account per participant

Each participant shall be limited to only one Kaggle account for the purpose of submission.

No sharing outside of teams

Privately sharing algorithms, codes and/or data for the purposes of the competition, outside of the participant's team is not permitted.

No hand labelling of data

Participants are not allowed to hand label Shopee Data (defined herein) provided to the participants during the competition period. Shopee reserves the right to test the teams' source codes in their respective entries to verify the authenticity of such entry. Shopee reserves the right to revoke the teams' ranking if hand labelling is detected in any entries of such team.

Team size limits

The team size consists of a minimum of 2 to a maximum of 4 Participants per team. Individuals who register for the competition and do not belong to a team shall be allocated to a team at Shopee's sole discretion.

Teams with less than 2 members will be disqualified.

Team mergers

Mergers between teams are not allowed.

External data

Usage of publicly, freely available external data in the teams' entry is permitted.

Eligibility

Employees of Sea Limited are not eligible to participate in the challenge in the interest of fairness.

Entry Requirements

The Competition is open only to registered Participants and Participants can register for the Competition via the Competition Website.

Shopee shall announce the rules and judging criteria of the Competition and any such rules and judging criteria shall be deemed to be incorporated into these Competition Terms and Conditions.

Teams may submit a maximum of five (5) Entries per day and teams can select two (2) final Entries from such teams' past submissions during the Competition Period for judging.

Shopee reserves the final right to reject any Entries in its sole discretion.

Obligations of the Participants

Participant grants Shopee the right to use the Participant's name in connection with such Participant's submitted Entry as well as in connection with any marketing, advertising or promotional material related to the Competition without requiring any further consent or compensation from Shopee.

Participant represents and warrants (where applicable) that it is duly organized, validly existing, and in good standing in the jurisdiction it is formed, that it will comply with all applicable laws in performing under these Competition Terms and Conditions and that it has full capacity, authority and legal right to agree to these Competition Terms and Conditions.

Participants agree that they shall not:

a) use or access datasets provided by Shopee for the purpose of this Competition ("Shopee Data") for any commercial, gambling, or illegal purpose; b) copy, modify, reproduce, license, sublicense, distribute, transmit, republish, exhibit, or display the Shopee Data (except to the extent copying, modifying, reproducing or displaying is necessary and directly related to the Participants' participation in the Competition); c) remove, obscure, or modify any copyright notices or other notices, designations, or attributions within the Shopee Data; or d) use any Shopee Marks or any copyright or other proprietary rights associated with the Shopee Data for any purpose without the express written consent of Shopee. "Shopee Marks" means all names, trademarks, logos, symbols, designations, emblems, designs, and trade dress of the Shopee, together with all identifications, labels, insignia or indicia thereof.

The winning team of the Competition shall license the winning Entry and the source code used to generate the Entry under an Open Source Initiative-approved license that in no event limits commercial use of such code or model containing or depending on such code.

Intellectual Property Rights

All intellectual property and proprietary rights in each Participant's Entry shall belong to Participant, save for Shopee's intellectual property embedded in such Entry which shall remain the sole property of Shopee.

Participant grants to Shopee, its and its subsidiaries' and affiliates' agents, contractor and partners, an irrevocable, non-exclusive, royalty-free, worldwide, perpetual and sub-licensable license to reproduce, distribute, transmit, create derivative works of, publicly display and publicly perform the Entry.

Shopee grants to Participants for the Competition Period a non-exclusive, non-transferrable and revocable license to use Shopee's Data solely in connection with the Competition. Participants shall not, without the prior written consent of Shopee, alter or modify or create derivative works of the Shopee Data or any of Shopee's intellectual property except as expressly set forth in these Competition Terms and Conditions.

Nothing in these Competition Terms and Conditions is intended to grant the Participant any rights to use any of Shopee's intellectual property, except as specified in these Competition Terms and Conditions.

Participant warrants that (a) its Entry is their original work and (b) Participant has all rights and ownership or is a licensed user of all algorithms, code and/or data used within the Entry.

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](#) **[Amended]**
- Annex D - [Discussion Board Guidelines](#)
- Annex E - [Kaggle Guide & Submission Requirements](#) **[Amended]**
- Annex F - [Common Issues during Submission](#) **[Removed]**
- Annex G - [Determining your Scores](#) **[Amended]**
- Annex H - [Common Kaggle Issues](#)
- Annex I - [Kaggle Frequently-Asked Questions \(FAQs\)](#)

Important: For the Product Title Translation (Data Science) challenge, the problem statement and dataset will be made available to participants on Kaggle. However, submissions will not be done on Kaggle.

Submission will only start from 11 July 2020, 12pm (GMT+8) onwards.

There will be no submission from 4 July to 10 July.

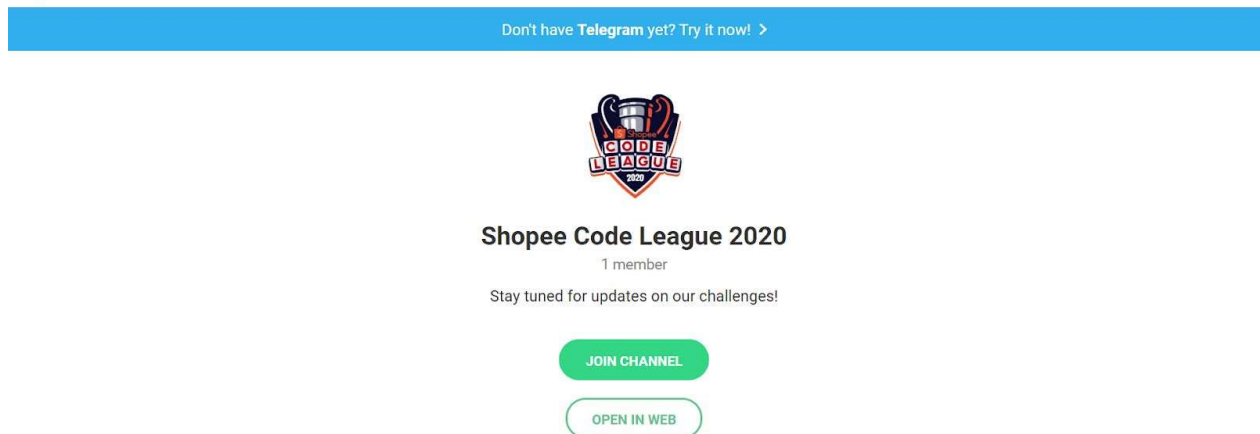
JOINING TELEGRAM CHANNEL

Download Telegram for [iOS](#) or [Android](#)

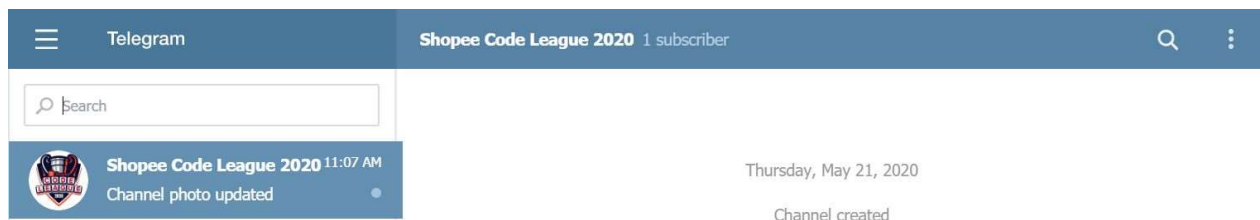
Register for an account

Join the Shopee Code League 2020 Telegram Channel

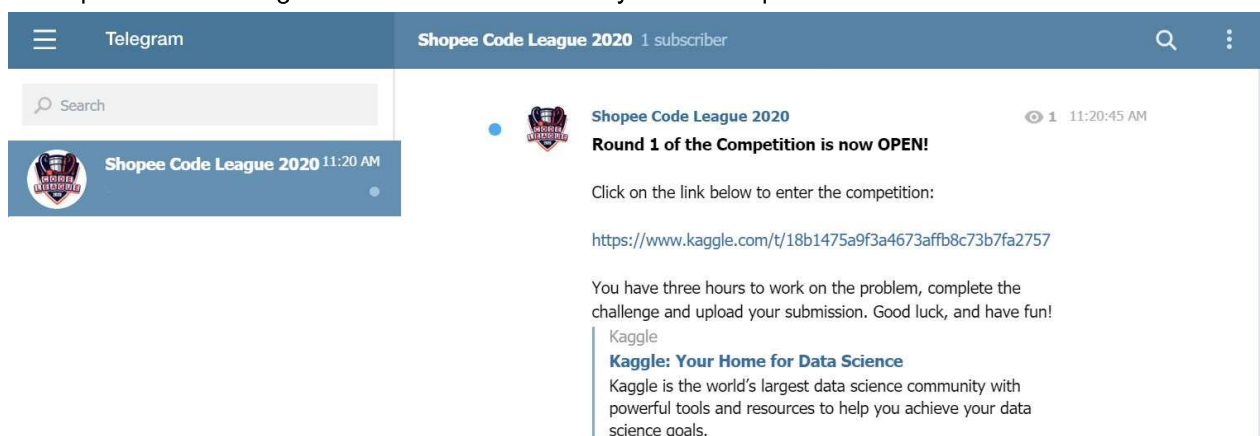
Click Open in Web to join the channel



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

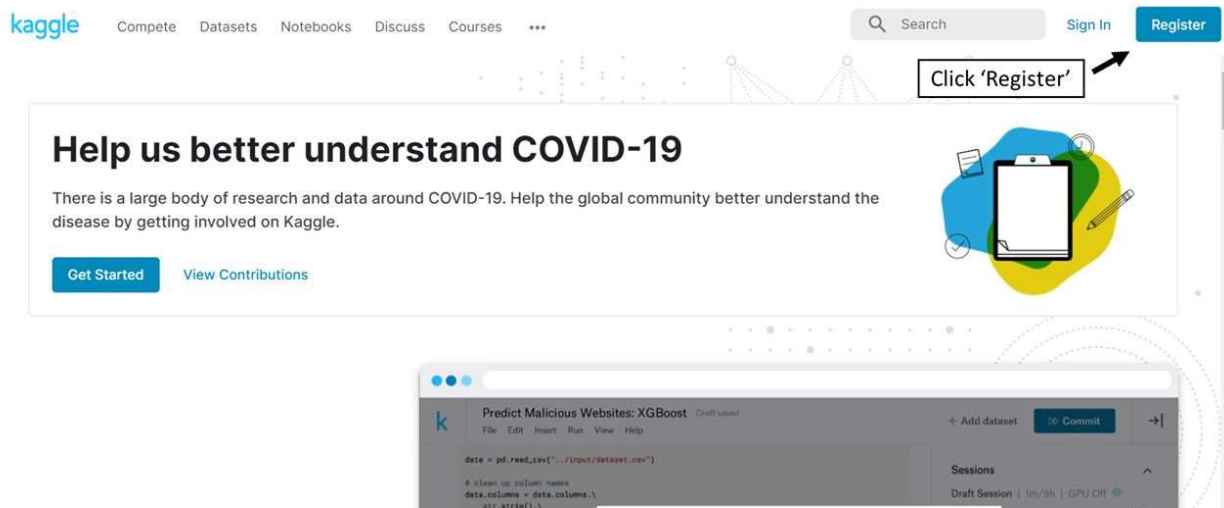


A sample of the message that will be sent on the day of the competition.



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

kaggle

Verify your email

We've sent you an email with a six-digit code. Please enter it here.

Six-digit code

Resend email
Next

TEAM FORMATION

Team formation on Kaggle will not be required for the Product Title Translation challenge as submissions will not be done on Kaggle.

Submission will only start from 11 July 2020, 12pm (GMT+8) onwards.

There will be no submission from 4 July to 10 July.

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

 A screenshot of the 'Discussion' tab interface. It shows a form for creating a new topic. At the top is a text input field labeled 'Topic Title...'. Below it is a rich text editor with various formatting icons (bold, italic, underline, link, unlink, quote, code, list, table, image, emoji, etc.). Below the editor is a text input field labeled 'Topic Content...'. At the bottom right are two buttons: 'Preview' and 'Publish Topic'.

- 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

You may submit up to **20 times as a team in a day for the Data Analytics questions** and **5 times a day for the Data Science questions**.

Additional submissions beyond this number will be automatically denied.

- 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

This challenge utilises **“Bleu score”** a metric currently not supported on Kaggle. As such, submission via Kaggle’s In-Class platform is unavailable for this challenge.

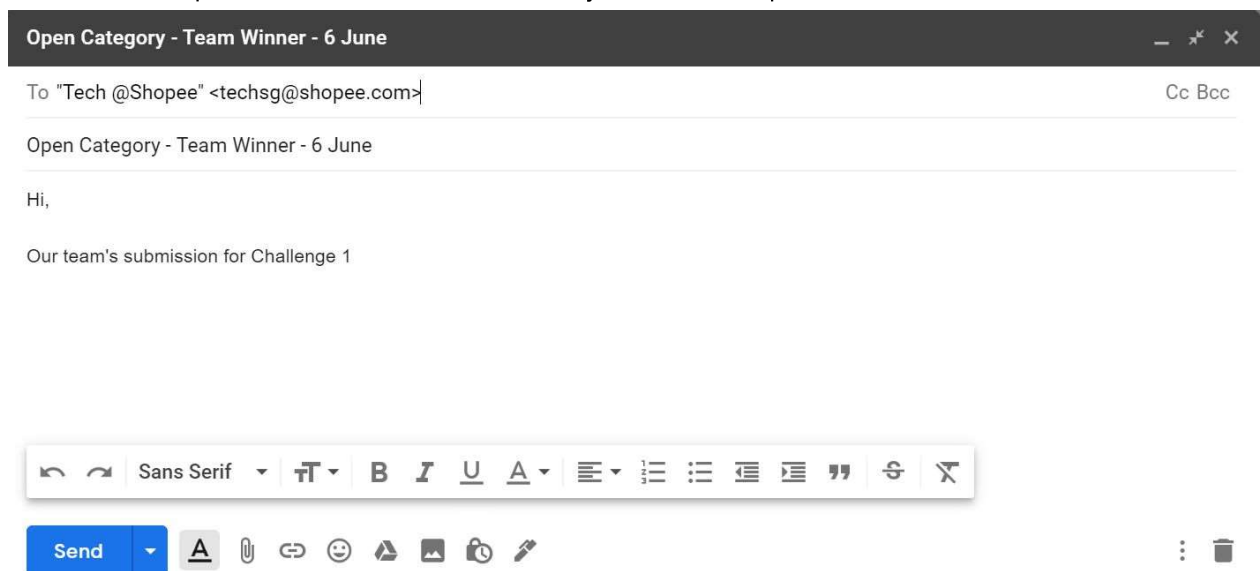
Submission will only start from 11 July 2020, 12pm (GMT+8) onwards.

There will be no submission from 4 July to 10 July.

Note: The **Top 100 teams** based on the Private Leaderboard will be required to send in a copy of their script after the end of the competition.

- Data Analytics Competition submission, 30 minutes after the end of competition.
 - [13 Jun] Order Brushing
 - [11 Jul] Logistics
 - [25 Jul] Marketing Analytics
- Data Science Competition submission, by 11am (GMT+7) / 12pm (GMT+8) the next day.
 - [20 Jun] Product Detection
 - [4 Jul] Title Translation
 - [18 Jul] Sentiment Analysis

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.

DETERMINING YOUR SCORES

While the Product Title Translation (Data Science) challenge is active, teams may view their scores on the Leaderboard after 11 July 2020 once they have uploaded a submission successfully.

Submission will only start from 11 July 2020, 12pm (GMT+8) onwards.

There will be no submission from 4 July to 10 July.

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 <ul style="list-style-type: none"> Too many requests error Page not loading 	<ol style="list-style-type: none"> 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 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 <ul style="list-style-type: none"> Team leader not able to invite team members Unexpected token e in json 	<ol style="list-style-type: none"> Trying clearing your browser cookies and cache and try again Get a different team member to form the team, at least one person has to set the registered team name. Email techsg@shopee.com to alert us of this issue with the subject: Team Formation Issue (Your Team Name). <ol style="list-style-type: none"> Include your members details and your current registered team name BEFORE the end of the competition
3	Downloading Dataset <ul style="list-style-type: none"> Cannot download dataset 	<ol style="list-style-type: none"> 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 <ul style="list-style-type: none"> Scoring takes a long time after making a submission on Kaggle 	<ol style="list-style-type: none"> Please be reminded to make a submission early in case of technical difficulties. 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): <ol style="list-style-type: none"> Include your members details and your current registered team name Only emails received BEFORE the end of the competition deadline will be taken into consideration

KAGGLE FREQUENTLY-ASKED QUESTIONS (FAQs)

This section outlines the Frequently-Asked Questions about the Kaggle Platform. This section has been separated into the following categories:

1. [Joining Competition](#)
2. [Registration](#)
3. [Team Formation](#)
4. [Submission](#)
5. [Scoring](#)
6. [Kaggle Kernels](#)
7. [Others](#)

1. Joining Competition

Questions	Answers
How do we attend competitions on Kaggle? Is it automatically assigned?	On the day of each competition, there will be a briefing on zoom 1 hour before the competition where the competition link will be released real-time on the zoom chat.
How to join the competition?	The competition link will be released on the zoom chat real-time upon the launch of the competition by our guest-of-honour. Please stay till the end of the opening!
Why do I see “This is a limited participation competition. Only invited users may participate”?	The competition can only be accessed via the <u>UNIQUE link</u> sent to each participant’s email.

2. Registration

Questions	Answers
Must the kaggle account be linked to the email account we signed up with the code league?	<p>For Kaggle, you may use your own individual username to sign up for an account in Kaggle. However, when you are in the competition page, please ensure that you have created the same team name which you have registered for the Shopee Code League.</p> <p>If you need to make edits to your Shopee Code League username, please submit an amendment request via this form here.</p> <p>Please send in all requests by 26 June.</p> <p>Any requests sent after the deadline will not be considered.</p>
My name in Kaggle was automatically set to a username and not my real name. Does it affect the competition?	
Must the email address for the Kaggle account be the same as the email address used for Code League registration?	
What if I have a Kaggle account under a different email address from the one I used for this Shopee Code League competition? Do I need to sign up for a new Kaggle account?	
Are kaggle accounts created using a single sign on (e.g. google) usable for the contest?	
Do we need to create new kaggle accounts if we already have our own accounts?	
Can the team name be in mandarin?	Yes, your team name may be in any language you choose. If you can type it, you can use it.
Does every team member need to sign up for their own Kaggle account but for Hackerearth only 1 team account is needed, am I right?	<p>Yes for Kaggle, every team member needs to sign up for their own Kaggle account.</p> <p>For Hackerearth only 1 team account is needed.</p>

3. Team Formation

Questions	Answers
How do I join the kaggle groups?	Once you have clicked 'Join Competition' and accepted the rules. You may click on the Teams tab in order to form teams. Please refer to Annex C of this Kaggle Competition Guide for more details on Team Formation.
How to make a team in kaggle?	
How do I pull out of a team, if I was added into the wrong group on Kaggle?	Please check that it is the right team before accepting the merger. You are unable to pull out of the team once you accept the merger.

4. Submission

Questions	Answers
For Kaggle competitions, we are required to send the codes written or only submit the result generated?	The scripts submitted for the competition must be complete - it should run from start to end without errors and output the required .csv file of predictions for the competition Any libraries or dependencies should be clearly specified within the script. Please refer to Annex E of this Kaggle Competition Guide for more details.
Can we use tableau, attach screenshot as markdown in kaggle for data analytics competition?	You may use data visualisations to aid in Exploratory Data Analysis, however they will not be required for submission
How many times can we submit each day?	You may submit up to 20 times as a team in a day for the Data Analytics questions and 5 times a day for the Data Science questions. Data Analytics Questions: [13 Jun] Order Brushing [11 Jul] Logistics [25 Jul] Marketing Analytics Data Science Questions: [20 Jun] Product Detection [4 Jul] Title Translation [18 Jul] Sentiment Analysis

Which kaggle account will be used to count points if each member has an individual account?	<p>All team members should create an account on Kaggle and form teams prior to submission.</p> <p>The highest individual submission score would automatically be chosen as the team's overall score.</p>
How will you determine the score for a team on kaggle if different team members submitted the solutions?	
Do we all need to create a personal kaggle account or just one team account on kaggle ?	
How to submit the code in this competition, especially in kaggle platform?	<p>Once you have joined the competition. You may click the Submit Predictions button to submit your answers.</p> <p>Please refer to Annex E of this Kaggle Competition Guide for more details on submission.</p>
I am unable to make my submission on Kaggle. What do I do?	<p>Please email techsg@shopee.com with the email subject: Submission Issue (Your Team Name)</p> <p>Include your Team Name and members details.</p> <p>Only emails received BEFORE the deadline of the competition will be taken into consideration.</p>

5. Scoring

Questions	Answers
What's the rule of calculating team leaderboard score?	Scoring methods on the Private Leaderboard on Kaggle can be found on the Evaluation page for each of the Kaggle challenges.
What is the difference between the public and private leaderboard?	<p>For the Data Analytics problem statements, there is no difference between the two leaderboards i.e. the public leaderboard is the same as the private leaderboard.</p> <p>For the Data Science problem statements, the Public Leaderboard is based on ~30% of the dataset while the Private Leaderboard is based on ~70% of the dataset.</p>

6. Kaggle Kernels

Questions	Answers
Is the Kaggle Data Science competition Kaggle kernel only? Or can we use our own machine?	You may run your scripts on your own local machine instead of using the Kaggle kernels.
Will the teams be provided by additional GPU hours on kaggle during the weeks when the competition is hosted on kaggle platform?	No. Additional GPU on Kaggle kernels will not be provided during the competition.
Are both data analysis and data science competitions held as in-kernel competitions in Kaggle?	No, you may run your scripts on your local machine. Only the .csv output generated is required for submission on Kaggle. The top 100 teams will need to submit their complete scripts via email to techsg@shopee.com according to guidelines specified in Annex E of this Kaggle Competition Guide.

7. Others

Questions	Answers
I'm new to data science. I've been using kaggle for a month now, but I never heard about data analytics. Do we need to make a model as well?	Model creation will only be required for the Data Science questions. Data Analytic questions will be focused towards processing of data to find answers to the question posed.
My team member will not have a computer until he gets it repaired - is it ok that he only makes a kaggle account and joins our team later?	Yes, it is alright since submission is done as a team.