SHOPEE CODE LEAGUE 2021
ADMINISTRATIVE GUIDE V2 FOR PARTICIPANTS

Dates of Competition:	6 to 20 March 2021
Dates of Training Workshops:	Training workshops will be held from 19 February to 18 March 2021. For more details, please check our website **here** and refer to our weekly newsletter announcements.
Content:	

EVENT BACKGROUND

Shopee Code League is a 3-week coding challenge consisting of 3 coding competitions open to all students and professionals across the region. The competitions, specially designed by the Shopee tech teams, cover data analytics, data science and algorithmic problems. Participants must analyse the datasets, draw insightful conclusions and solve the problems in a specified amount of time.

Through Shopee Code League, we aim to:

- Equip undergraduates and professionals across the region with essential tech skills and expertise to prepare them for the new digital economy
- Bring the tech communities closer through problem solving and knowledge sharing together across the region
- Provide participants with the opportunity to work on real datasets and challenges in the Internet industry

SHOPEE CODE LEAGUE DETAILS: HOW IT WORKS

Shopee Code League will operate online entirely from 6 to 20 March 2021.

- Competitions are held every weekend of the league (6 March - Data Analytics Challenge, 13 March - Data Science Challenge, 20 March - Programming Contest)
- Participation in all competitions is not compulsory. Your team is free to choose which competitions to participate in.
    - However, participating in more competitions will allow your team to gather more points and rise up the leaderboard.

You can look forward to attending online sharings and training workshops organized by Shopee tech teams and our training partners, on the following topics:

- Programming Languages such as Python
- Data Analytics
- Data Science
- Software Engineering

LIST OF RESOURCES

We have curated the following non-exhaustive resources that you and your team can look into, in preparation for Shopee Code League 2021.

I. Online Learning Platforms

Section I offers four online learning platforms that may be useful in your preparation for Shopee Code League. These platforms offer video tutorials, courses and other resources outlining various data-driven concepts and programming language skills.

Disclaimer: We are not promoting any of these platforms, these are just suggestions you can look into.

Platforms	Overview
Coursera	Coursera provides one of the longest-established online data science educations through John Hopkins University. Learners can access thousands of courses for free, without access to features like graded homework assignments or certificates of completion. ● Data Science Courses ● Data Science Specialisation Course ● C for Everyone: Programming Fundamentals ● Python Programming: A Concise Introduction ● Programming Languages, Part A ● Programming Languages, Part B ● Programming Languages, Part C ● Code Yourself! An Introduction to Programming ● Introduction to Programming with MATLAB ● C++ For C Programmers, Part A ● Computer Vision Basics ● Algorithms, Part I ● Algorithms, Part II ● Analysis of Algorithms ● Computer Science: Algorithms, Theory, and Machines ● Computer Science: Programming with a Purpose
RealPython	At Real Python, you'll learn real-world programming skills from a community of professional Pythonistas from all around the world. Find courses and articles according to your level of expertise.

	• I'm new to Python and to programming in general • I'm an intermediate Python developer—How do I take my skills and my productivity to the next level? • I've got experience with other programming languages—How do I get up to speed with Python as quickly as possible?
Udemy	Udemy offers highly-rated data science courses that will help you learn how to visualize and respond to new data, as well as develop innovative new technologies. The free courses offer online video content, whereas the paid courses include certificate of completion, instructor Q&A and instructor direct message. • Introduction to Python Programming • Python Core and Advanced • Python For Data Science • Java Tutorial for Complete Beginners • Java Multithreading • Java Programming Basics • Practice Java by Building Projects • Java 8 Functional Programming: Lambda Expressions Quickly • Learn how to use all Java keywords • Intermediate Level C++
edX	edX is the trusted platform for education and learning. Founded by Harvard and MIT, edX is home to more than 20 million learners, the majority of top-ranked universities in the world and industry-leading companies. • Data Science Courses • Software Engineering Courses • Java Programming Fundamentals • CS50's Web Programming with Python and JavaScript • Visualizing Data with Python • Deep Learning with Python and PyTorch • Software Engineering Essentials
Udacity	Udacity partners with leading technology companies to learn how technology is transforming industries, and teach the critical tech skills that companies are looking for in their workforce.

	With their powerful and flexible digital education platform, even the busiest learners can prepare themselves to take on the most in-demand tech roles. ● Business Analytics ● Predictive Analytics for Business ● Programming for Data Science with Python ● Data Structures and Algorithms

II. Understanding Data Science Models

Section II offers a series of data science models/topics that would be helpful when solving the Shopee Code League problem statements. You would find links to articles and other online resources, categorised by model type.

Image Classification Model

1) InceptionV3
    - A Simple Guide to the Versions of the Inception Network
    - Review: Inception-v3 — 1st Runner Up (Image Classification) in ILSVRC 2015

2) MobileNet
    - Transfer Learning using Mobilenet and Keras
    - Review: MobileNetV1 — Depthwise Separable Convolution (Light Weight Model)
    - Creating insanely fast image classifiers with MobileNet in TensorFlow

3) SqueezeNEt
    - Review: SqueezeNet (Image Classification)

4) VGGNet
    - What is the VGG neural network?
    - Review: VGGNet — 1st Runner-Up (Image Classification), Winner (Localization) in ILSVRC 2014
    - Difference between AlexNet, VGGNet, ResNet, and Inception

5) ResNet
    - Introduction to ResNets
    - An Overview of ResNet and its Variants
      Intuition For: ResNet — Deep Residual Learning for Image Recognition

6) ResNeXt
    - Understanding and Implementing Architectures of ResNet and ResNeXt for state-of-the-art Image Classification: From Microsoft to Facebook [Part 1]
    - Understanding and Implementing Architectures of ResNet and ResNeXt for state-of-the-art Image Classification: From Microsoft to Facebook [Part 2]
    - Review: ResNeXt — 1st Runner Up in ILSVRC 2016 (Image Classification)
    - Enhancing ResNet to ResNeXt for image classification

Image Segmentation Model

1) SegNet
    - Summary of — SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation
    - Review: SegNet (Semantic Segmentation)
    - Understanding of Semantic Segmentation & How Segnet Model work to perform Semantic Segmentation

2) Deeplab
    - The Evolution of Deeplab for Semantic Segmentation
    - Semantic Image Segmentation with DeepLab in TensorFlow
    - How to use DeepLab in TensorFlow for object segmentation using Deep Learning

3) MaskRCNN
    - Simple Understanding of Mask RCNN
    - Computer Vision: Instance Segmentation with Mask R-CNN
    - Mask R-CNN for Ship Detection & Segmentation

Text Model

1) RNN
2) LSTM
    - Understanding RNN and LSTM
    - Recurrent Neural Networks and LSTM explained
    - Recurrent Neural Networks
    - Report on Text Classification using CNN, RNN & HAN
    - Generating text using a Recurrent Neural Network
    - Multi-Class Text Classification with LSTM
    - Illustrated Guide to LSTM's and GRU's: A step by step explanation
    - The magic of LSTM neural networks
    - Video/Course: Long Short Term Memory (LSTM)

Image Processing

- Getting Started with Image Processing using Python
- Exploring Image Processing Techniques — OpenCV
- Image processing with Python & Open-CV part-1
- Image processing using Python & Open-CV part-2
- Image processing using Python & Open-CV part-3
- Image Processing using Python basic -I

Convolution Neutral Network

- Understanding of Convolutional Neural Network (CNN) — Deep Learning
- A Beginner Intro to Convolutional Neural Networks
- A Brief Guide to Convolutional Neural Network(CNN)
- Convolutional Neural Networks — CNN
- The best explanation of Convolutional Neural Networks on the Internet!

Object Detection

- Understanding Object Detection
- Object Detection with 10 lines of code
- Beginner's Guide to Object Detection Algorithms

Natural Language Processing

- Natural Language Processing is Fun!
- A Practitioner's Guide to Natural Language Processing (Part I) — Processing & Understanding Text

TF-IDF

- TF-IDF from scratch in python on real world dataset.
- What is TF-IDF in Feature Engineering?
- TF IDF | TFIDF Python Example
- How to process textual data using TF-IDF in Python
- TF-IDF/Term Frequency Technique: Easiest explanation for Text classification in NLP using Python (Chatbot training on words)

Word Embeddings

- Introduction to Word Embedding and Word2Vec
- Word embeddings in NLP
- Video: Using Word Embeddings

BERT - Language Modeling

- BERT Explained: State of the art language model for NLP
- Understanding BERT: Is it a Game Changer in NLP?
- Google BERT — Pre Training and Fine Tuning for NLP Tasks
- Building State-of-the-Art Language Models with BERT
- Google Scholar Papers/Articles on BERT

Classification Models

- Machine Learning Classifiers
- A Lesson on Modern Classification Models
- Classification Algorithms Used in Data Science
- Intro to types of classification algorithms in Machine Learning (2017)

Transformer

- What is a Transformer?
- How Transformers Work
- Transformer: A Novel Neural Network Architecture for Language Understanding (2017)

Sequence to Sequence Model

- Understanding Encoder-Decoder Sequence to Sequence Model (2019)
- Sequence To Sequence Models (2018)
- Sequence to sequence model: Introduction and concepts (2017)
- NLP | Sequence to Sequence Networks| Part 1| Processing text data
- NLP | Sequence to Sequence Networks| Part 2| Seq2seq Model (EncoderDecoder Model)
- Sequence Modeling with Deep Learning

Attention Model

- Brief Introduction to Attention Models
- An introduction to Attention
- Intuitive Understanding of Attention Mechanism in Deep Learning
- Attention and its Different Forms
- Attention Mechanisms in Deep Learning — Not So Special
- Coursera Video: Attention Model

III. Engineering Algorithm Questions

Section III offers tips and tricks on answering algorithm questions.

- Software Engineering Must Know: Algorithms
- Top Algorithms and Data Structures You Really Need To Know
- 8 Common Data Structures every Programmer must know
- Data Structures
- Fundamentals of Algorithms

[UPDATED] HEAR FROM US

The Shopee Team will keep you updated via email throughout the entire Shopee Code League. You would expect two types of emails:

I. (Bi)Weekly Announcements Newsletter

This newsletter contains important updates on Shopee Code League 2021. These updates include details on upcoming competitions, registration and detailed access to training workshops, and a weekly leaderboard (Refer to Section D for more information).

II. Weekly Scores E-mail

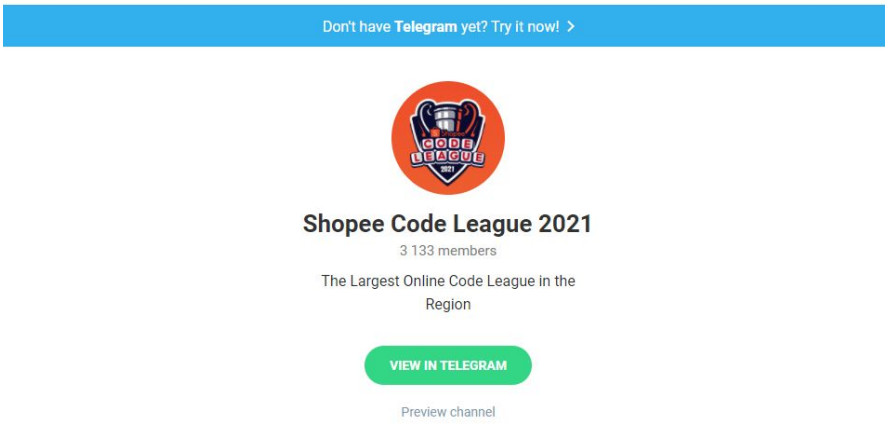
You would receive a weekly scores email that contains your team's tabulated scores after the first weekend (the week of 8th Mar). Refer to Section C for more information on scores tabulation.

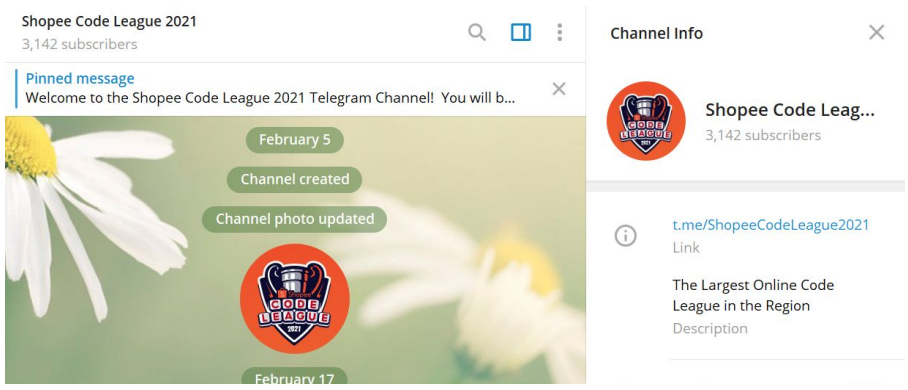
Do keep a lookout for our emailers to be updated with the latest information. **Do not unsubscribe** to any of these emails and ensure that your email account will accept emails from techsg@shopee.com. If you have not been receiving any emails from techsg@shopee.com, please check your spam inbox, or send us an email to check.

A. [NEW] GETTING STARTED

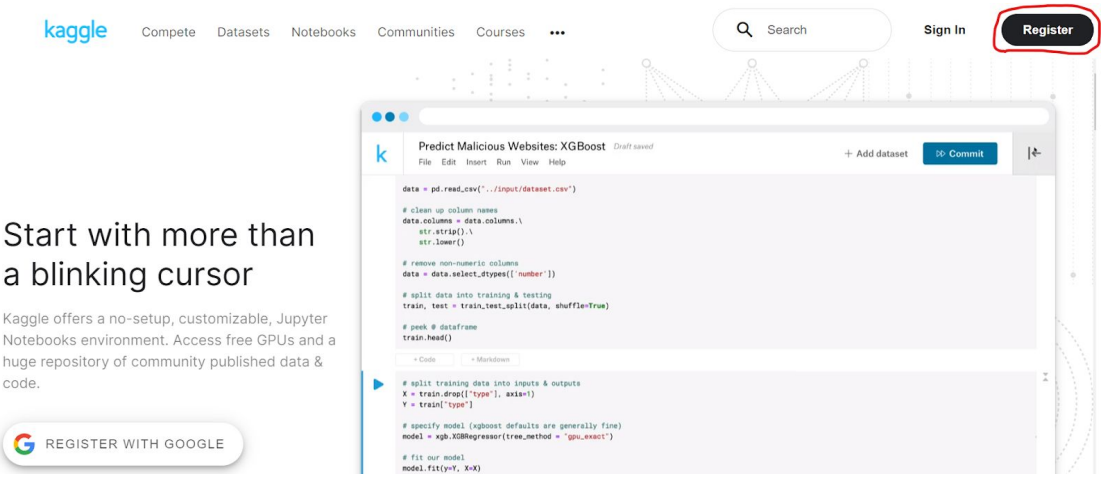
[For your action] Participants will be required to have the following platforms ready by 6 March, Sunday.

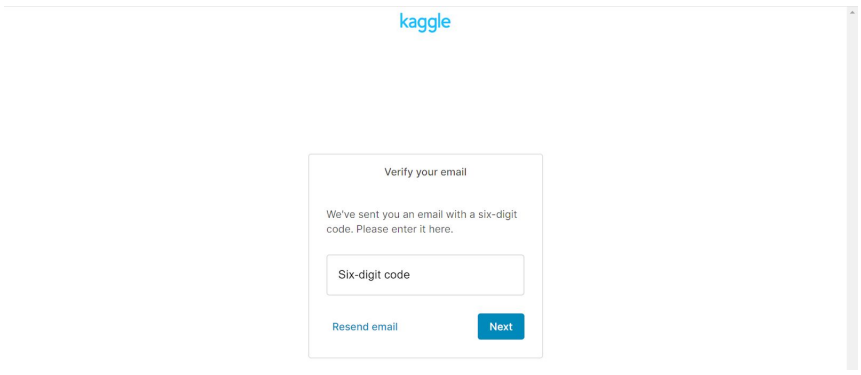
I. **[Compulsory]** Set up your Telegram Channel to receive important updates and announcements.

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 2021 Telegram Channel** https://t.me/ShopeeCodeLeague2021  *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.

II. **[Compulsory]** Set up and register for your Individual 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.**

B. [NEW] POINTS SYSTEM

- Each online competition carries a maximum of 100 points. Each team's entry in respect of the online competitions will be evaluated and points are awarded accordingly.
- Participation in the various online competitions can increase accumulation of points for each team.
- Each team's total accumulation of points will determine its team's ranking on the Competition's leaderboard.

Points for this competition will be normalised using **Min/Max Normalisation and multiplied by 100**, the formula is as such:

$$\textit{Formula}: \left(\frac{value - min}{max - min}\right) x\ 100$$

This will be applied uniformly across the 3 competitions in order to determine the points earned from each contest.

For the Data Analytics and Data Science problem statements, the full score on Kaggle is 1.0. The points awarded will range from 0 to 100. Please refer to Example 1 and 2.

Example 1 (Data Analytics & Data Science) - Range of 0 to 1:

Team	A	B	C	D
Score	0	0.1	0.2	1.0

The normalised score for team C would be:

$$\frac{0.2-0}{1.0-0} \times 100 = 20$$

Therefore for this competition team C scored 20 point(s).

Example 2 (Data Analytics & Data Science) - Range of 0.4 to 0.8:

Team	A	B	C	D
Score	0.4	0.5	0.8	0.8

The normalised score for team C would be:

$$\frac{0.8-0.4}{0.8-0.4} \times 100 = 100$$

Therefore for this competition team C scored 100 point(s).

For the Programming problem statements, the full score on HackerEarth is 100. Team scores will range from 0 to 100. Please refer to Example 3.

Example 3 (Programming Contest) - Range of 50 to 100:

Team	A	B	C	D
Score	50	60	60	100

The normalised score for C would be:

$$\frac{60-50}{100-50} \times 100 = 20$$

Therefore for this competition team C scored 20 point(s).

The total points earned by Team C would be 20+100+20 = 140 points out of a maximum of 300 points.

[Bonus Points] For each competition, teams who are able to attain a perfect score (100 points) in less than one (1) hour will be awarded a total of 5 bonus points. These bonus points will be added to the normalised score, thereby making the maximum attainable score of each competition to be potentially 105 points.

C. [NEW] COMPETITION LEADERBOARD

- Each competition platform will have a designated leaderboard.
- The points from each leaderboard will be accumulated and displayed on the Competition's leaderboard.
- The Competition's leaderboard will be reflected and updated on a weekly basis on the following platforms:
    - [Website] Shopee Code League Competition Website: https://careers.shopee.sg/codeleague/#leaderboard
    - [Email]: Shopee Code League Weekly Announcements Newsletter
- The Competition's Leaderboard will display the top 15 teams and top all-female team with the highest accumulated points per participation category.
    - Terms and Conditions apply.

D. [NEW] PAST COMPETITION RESOURCES

This section contains access to our past problem statements and data files that could aid in your preparation for Shopee Code League 2021. You would be able to access the respective links after registering for your Kaggle Account.

You would also be able to upload your submissions however it would not be recorded on the leaderboard as the competitions have closed.

Note: The file sizes of the datasets are large. We do recommend that you download/install a software that can extract from zip files such as WinRar.

I. I'm the Best Coder 2019! Challenge

I'm the Best Coder 2019 is a one-day data analytics competition where coders and programmers alike collaborate to analyse complex problem statements and examine real datasets to solve real industry problems.

- Pre-Tertiary Round 1
- Pre-Tertiary Round 2
- Undergraduate Round 1
- Undergraduate Round 2
- Open Round 1
- Open Round 2

II. National Data Science Challenge 2019

National Data Science Challenge 2019 is a 4-week long competition where participants are required to extract insightful knowledge from large volumes of textual and visual data using Machine Learning Analytics.

- Beginner: Product Category Classification
- Advanced: Product Information Extraction

III. Shopee Code League 2020

Data Analytics Problem Statements

Participants are required to come up with a script which can output the correct answer.

- Order Brushing
- Logistics
- Marketing Analytics

Data Science Problem Statements

Participants will have to build a model which can accurately predict the labels for each of the rows in the given datasets.

- Product Detection
- Title Translation
- Sentiment Analysis

Programming Contest Questions

Participants will need to come up with the correct solutions to each of the test cases provided by the Shopee Engineering teams.

- Shopee Programming Contest #1
- Shopee Programming Contest #2

FURTHER INFORMATION

The following Annexes are enclosed for your information.

- Annex A - Shopee Code League FAQs
- Annex B - Terms and Conditions
- Annex C - Privacy Policy

ANNEX A

SHOPEE CODE LEAGUE FAQs

You would be able to find answers to Frequently Asked Questions (FAQs) regarding Eligibility, the Competition and Registration here.

ANNEX B

TERMS AND CONDITIONS

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 of Shopee Code League 2021.

ANNEX C

PRIVACY POLICY

By consenting and submitting the registration form, you have acknowledged and agreed that you accept the practices, requirements, and/or policies outlined in our Competition Privacy Policy of Shopee Code League 2021 here.