The War of the Wizards just got bigger and better!

**WNS Analytics Wizard**, a one-of-its-kind online analytics hackathon, is back with a second edition. Here’s an exciting opportunity for young and aspiring analytics professionals to experience challenging, real-life business scenarios and showcase your analytical acumen and problem-solving skills.

The stage is set for all you budding **data wizards** to showcase your acumen and creativity and win with insights.

So, are you ready for the **War of the Wizards**?

Registernow to participate in the largest online hackathon yet, at WNS.

Why should you participate?

WNS Analytics Wizard 2019 is the ultimate challenge for all the young aspiring analytics professionals. Participate in the online hackathon to:

* Get a taste of real-life business scenarios and data challenges, and improve your analytical skills
* Compete with data and analytics masterminds across the globe
* Boost your CV with the title of being the ultimate WNS Analytics Wizard and gain recognition among peers
* Take home cash prizes up to INR 4 Lacs

Who should participate?

Open for all data lovers – statisticians, data scientists, business analysts and students.

Prizes

Cash prizes up to INR 4 lacs up for grabs

* 1st Prize – INR 2,00,000
* 2nd Prize – INR 1,25,000
* 3rd Prize –INR 75,000

Along with prizes, top participants will also get interview opportunities with WNS

About WNS

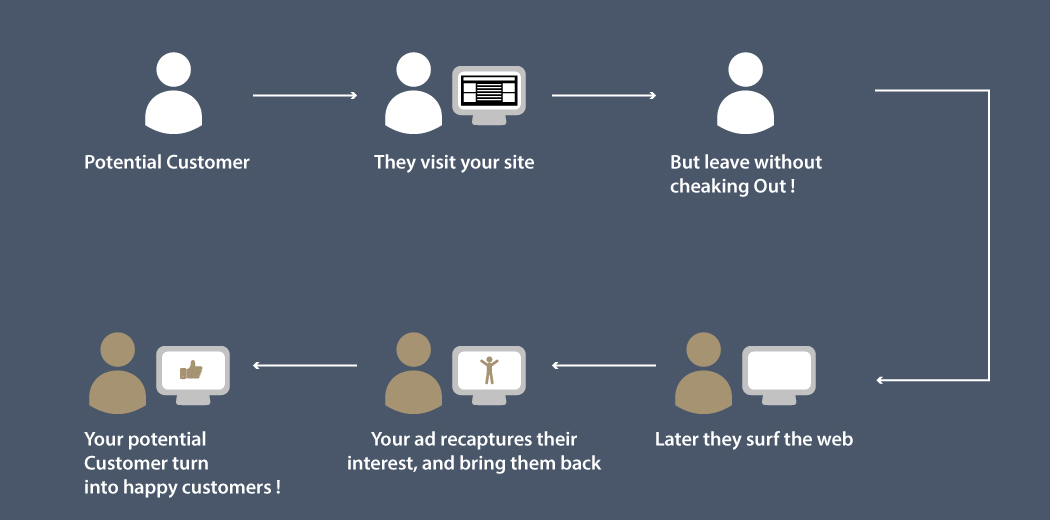
WNS (Holdings) Limited(NYSE: WNS) is a leading Business Process Management (BPM) company. WNS combines deep industry knowledge with technology, analytics and process expertise to co-create innovative, digitally led transformational solutions with over 350 clients across various industries. It is a trusted analytics partner to over 90+ clients from banking, insurance, hospitality, retail, CPG, logistics, healthcare providers, airlines and many more. WNS leverages its analytics expertise in combination with rich industry domain knowledge and custom-built proprietary frameworks to help clients make informed decisions at the right time. It has more than a decade of proven, core analytics experience with over 2,700+ analytics professionals spread across 16+ delivery centers. For more information, visit<https://www.wns.com/solutions/functional-solutions/analytics>

Rules of Participation

1. The contest starts on Saturday, 24th August 2019, 00:00:00 (GMT).The data set and the problem statement will be released at the same time.
2. Entries submitted after the contest is closed, will not be considered.
3. Individual participation is allowed in the hackathon, and the participant can either be a part of a team or can participate individually.
4. Use of external dataset is not allowed.
5. Participation is free-of-charge.
6. Throughout the hackathon, you are expected to respect fellow hackers and act with high integrity.
7. Slack Live Chat admins hold the right to block any participant found to use foul/disrespectful language.
8. Current WNS employees are also eligible to participate.
9. Analytics Vidhya and WNS hold the right to disqualify any participant at any stage of the competition if the participant(s) are deemed to be acting fraudulently.
10. In case of any dispute, decision of Analytics Vidhya shall be final.
11. Prize money is subject to tax deduction as per Income Tax Rules.

Team Formation

* **Click**[**here**](https://docs.google.com/document/d/1nOT8zfI7FMkhwuEiJR_4U5xhX05swotTL8YGOiETIBo/edit)**to view process flow for Team Creation**
* Maximum of 2 people can form a team.
* One person can be a part of one team only.
* In case a team wins, prize would be distributed equally among team members
* Team once created can't be dissolved.
* Teams can't be merged.

Zbay is an E-commerce website which sells a variety of products at its online platform. Zbay records user behaviour of its customers and stores it as a log. However, most of the times, users do not buy the products instantly and there is a time gap during which the customer might surf the internet and maybe visit competitor websites.   Now, to improve sales of products, Zbay has hired Adiza, an Adtech company which built a system such that ads are being shown for Zbay’s products on its partner websites.   If a user comes to Zbay’s website and searches for a product, and then visits these partner websites or apps, his/her previously viewed items or their similar items are shown on as an ad. If the user clicks this ad, he/she will be redirected to the Zbay’s website and might buy the product.      In this problem, the task is to predict click probability i.e. probability of user clicking the ad which is shown to them on the partner websites for the next 7 days on the basis of historical view log data, ad impression data and user data.   You are provided with the view log of users (2018/10/15 - 2018/12/11) and the product description collected from the Zbay website. We also provide the training data and test data containing details for ad impressions at the partner websites(Train + Test).   Train data contains the impression logs during 2018/11/15 – 2018/12/13 along with the label which specifies whether the ad is clicked or not. Your model will be evaluated on the test data which have impression logs during 2018/12/12 – 2018/12/18 without the labels.

## **Dataset Description**

You are provided with the following files: **train.zip**: This contains 3 files and description of each is given below: train.csv :

|  |  |
| --- | --- |
| Variable | Definition |
| impression\_id | AD impression id |
| impression\_time | Time of the impression at partner website |
| user\_id | user id |
| app\_code | Application Code for a partner website where the ad was shown |
| os\_version | Version of operating system |
| is\_4G | 1-Using 4G, 0-No 4G |
| is\_click | (target) Whether user clicked the AD (1-click, 0-no click) |

  view\_log.csv :

|  |  |
| --- | --- |
| Variable | Definition |
| server\_time | Timestamp of the log |
| device\_type | Device type of the user |
| session\_id | Browser session id |
| user\_id | user id |
| item\_id | Item id |

  item\_data.csv:

|  |  |
| --- | --- |
| Variable | Definition |
| item\_id | Item id |
| item\_price | Price of the item |
| category\_1 | Category depth 1 |
| category\_2 | Category depth 2 |
| category\_3 | Category depth 3 |
| product\_type | anonymized item type |

**test.csv:** test file contains the impressions for which the participants need to predict the click rate **sample\_submission.csv:** This file contains the format in which you have to submit your predictions.

## **Evaluation Metric**

Submissions are evaluated on [area under the ROC curve](http://en.wikipedia.org/wiki/Receiver_operating_characteristic) between the predicted probability and the observed target.

## **Public and Private Split**

Test data is further randomly divided into Public and Private data. Public leaderboard is based on first 3 days of data from the test dataset (2018/12/12 – 2018/12/14), while private leaderboard will be evaluated on remaining 4 days of data of the test dataset (2018/12/15 – 2018/12/18).

* Your initial responses will be checked and scored on the Public data.
* The final rankings would be based on your private score which will be published once the competition is over.

## **Hackathon Rules**

1. Setting the final submission is mandatory. Without a final submission, the submission corresponding to best public score will be taken as final submission
2. Use of external datasets is not allowed
3. You can only make 15 submissions per day
4. Code file is mandatory while setting final submission. For GUI based tools, please upload a zip file of snapshots of steps taken by you, else upload code file.
5. The code file uploaded should be pertaining to your final submission.
6. No submission will be accepted after the contest deadline