

CaseQuest: An Epic Journey in Problem-Solving

DATATHON OVERVIEW

Introduction

In an era where data is the new currency, the ability to harness its potential has become a paramount skill. Recognizing the need to nurture the future leaders in the field of Data Science and Machine Learning, [Pickl.AI](#) by [TransOrg Analytics](#) is proud to present the TransOrg Analytics Datathon “**CaseQuest**” at LPU. This exciting competition aims to empower budding Data Scientists and Machine Learning enthusiasts by offering them a platform to showcase their skills and creativity while solving real-world challenges.

A Legacy of Excellence, **TransOrg Analytics** is a renowned leader in the field of data analytics and machine learning solutions. With years of experience spanning various sectors and domains, TransOrg has established itself as a pioneer in delivering cutting-edge data-driven solutions to clients worldwide. Their commitment to innovation, coupled with a deep understanding of data, has enabled them to drive transformative change across industries.

TransOrg Analytics takes pride in its ability to leverage its vast experience across diverse sectors. From aviation to finance, retail to hospitality, and beyond, TransOrg has successfully addressed complex problems with data-driven insights and predictive analytics.

The TransOrg Analytics CaseQuest at LPU is a testament to their dedication to nurturing the next generation of data science and machine learning leaders. Participants in this datathon will have the opportunity to take on challenging real-world problems, much like the ones TransOrg routinely tackles for its clients. Through this experience, participants will not only sharpen their analytical skills but also learn to think critically and creatively in the world of data-driven decision-making.

Why Participate?

1. **Real-world Experience:** Contestants will gain hands-on experience in solving complex data science and machine learning challenges, enhancing their problem-solving skills.
2. **Mentorship:** Participants will have the chance to interact with TransOrg's seasoned experts, receiving valuable guidance and insights to enhance their understanding of data analytics.

3. Recognition: Outstanding solutions will be recognized and rewarded, offering participants the opportunity to showcase their talent and enhance their resumes.

The TransOrg Analytics' Datathon at LPU is an exciting opportunity for aspiring data scientists and machine learning enthusiasts to test their mettle, learn from industry experts, and shape themselves into the leaders of tomorrow. TransOrg Analytics' commitment to excellence and innovation promises a challenging and rewarding experience for all participants. Join us on this journey to unlock your potential and become a future leader in the dynamic world of Data Science and Machine Learning.

The contest has 2 rounds,

1. Round 1 – Qualifier Round, an aptitude, and a technical quiz to test your grasp over the concepts.
2. Round 2 – CaseQuest Round where you all will be working in teams and submitting the solutions/insights/building a model. Solution(s) submission link would be shared with all the participating teams.

Start Date	End Date	Team Size
<p>You can add your team details here: https://docs.google.com/spreadsheets/d/1s0mq_fJt6wtJ_JUSf6Lx4m1xey41elqJ_5gJwKwulmk/edit?usp=sharing</p> <p>5th October 2023 – Quiz will be available till 9th October 2023</p>	5 th November 2023	5-6 (please try for the mixed sections)

Round 1 - Qualifier Round (Suggested Date – 5th – 9th Oct 2023)

This is a quiz based on Aptitude and Data Analysis. It has 20 questions and teams will be given 30 minutes.

Rules:

1. Team participants should be preferably from different sections with a maximum capping of 6 per team.
2. Only 1 submission per team is allowed. Team will be disqualified if multiple submissions from the same team are found.
3. There will be an active window of 5 days available for the teams to attempt the quiz, post that quiz can't be attempted.
4. Once the team picks up the quiz, it has to be completed within 30 minutes at one go, after the completion of 30 minutes, the quiz will be auto submitted, and the quiz link will be disabled.

5. The top 6 teams would be selected to go into the next CaseQuest round on the basis of the scores in the quiz.
6. Qualifying teams would be notified through mail from the TransOrg Analytics team.

Round 2 – Case Quest Round (Suggested Date – 9th Oct – 5th Nov)

This is a Data Science Case where teams will be given a problem statement, dataset as well as a few questions and the team need to build a solution e.g., a dashboard/ML model. The final solution should incorporate the insights/answers to the questions asked in the problem statement. Along with that, the team also needs to write a blog and record a video (make a presentation first & record it).

SOLUTIONS/DELIVERABLES

Teams are required to submit the following documents:

1. A dashboard/Model in an appropriate format to evaluate.
2. A blog (in .pdf format) - The blog should additionally contain the approach to solving the problem, how they tackled challenges, improved the dashboard/model, and how this contest helped them towards their data science/ analytics journey.
3. A video (in .mp4 format) - Introduce your team and demonstrate your work.

Note: Video will have to be shared in .mp4 format, and the duration of the video has to be strictly less than 5 minutes.

JUDGING RUBRICS FOR ROUND 2

Evaluation criteria on the basis of which the participants will be evaluated are as follows:

Interpretation of Data (20 points)

This category examines the more technical aspects of visualization and the submission's treatment of the data.

- Did the chosen method of visualization make sense for the data the submission was working with?
- Are interesting insights or patterns from the data clearly highlighted through the visualization?
- Do the conclusions or implications made (ex. In the title) match the patterns shown in the visualization?

- Has the submission demonstrated good use of available visualization tools or platforms?

EDA & Modeling (20 points)

This category examines the EDA and model accuracy of the submission.

- How good is feature importance and the EDA process?
- Is the model working good or not?

Presentation and Visual Appeal of Dashboard (20 points)

This category examines the design choices and visual appeal of the submission.

- Is the overall dashboard looking interactive or not?
- Is the formatting of the visualization good or not (Select the correct chart /graph or not)?

Blog (15 points)

This category evaluates the writing skills and the organization of the content.

- Is the information correctly organized or not?
- Is the flow maintained throughout the blog?
- Is the presentation well written and understandable?

(Bonus) Creativity (25 points)

This is a bonus category to reward especially innovative submissions and presentations.

- Did the submission do something especially innovative, interesting, or creative? (ex. In their chosen method of visualization or their overall presentation)
- Are particularly interesting insights or conclusions made?
- Did the team present the solution well?
- Were they able to satisfactorily answer the questions of the judges?
- Did all members of the team participate in the final presentation?

For more details of how the marks are divided, please go through the document [Judging Rubrics](#)

RULES

Below are a few rules that every participant needs to follow:

- One Submission Per Team
 - a. The contest will allow only one account per participant.
 - b. Submissions from multiple accounts will lead to disqualification.
- No Private Sharing

- a. Privately sharing code or data outside of teams is not permitted.
 - b. It's okay to share code if made available to all participants on the forums.
- Disqualification
 - a. If any of the details entered are found incorrect, TransOrg Analytics reserves the right to disqualify any participant.
 - b. Any plagiarism is strictly prohibited.

PRIZES

1. The best blog will be published on Pickl.AI's website(www.pickl.ai) with due credits given to the winning team.
2. **Free "Data Science Bootcamp" worth INR 25,000 for the winning team.**
3. Discounted membership for Pickl.AI industry-ready courses for the participating teams.
4. Goodies for the Top 3 teams.

PICKL.AI's RECOMMENDED RESOURCES

These resources are here as a starting point for the participants. There are so many platforms out there for visualization. Creating a successful visualization doesn't mean that one has to learn all of them! We recommend checking out a few of the general resources, then skimming through the different platforms and picking 1-2 to focus on.

General Resources

[How to Start Data Science](#) - This article will introduce what data science is, what one can do with the power of numbers and how to learn it.

[How to Choose the Right Data Visualization](#) - There are many ways data could be communicated, so this article explains what some of the standard visualizations of data are and when they should be used.

[5 Data Visualization Best Practices: The Secrets Behind Easily Digestible Visualizations](#) - Quick tips to improve your visualizations, with some great examples thrown in!

[academic/awesome-datascience: An awesome Data Science repository to learn and apply for real world problems.](#) - This GitHub repository is a huge resource list of data science links, courses, and information for those who are ready to dive in deep.

[FreeCodeCamp Data Science Tutorial](#) - A full video course on data science for those who learn best with videos.

[Kaggle Course on Data Visualization](#) - A course on data visualization produced by Kaggle.

[Nightingale – Medium](#) - Nightingale is the journal of the Data Visualization Society, an organization dedicated to featuring and educating others as they learn the principles and best practices of visualization. Their medium has various articles on visualizing certain types of data, amazing examples of visualization, and information on best practices. Some particularly good articles for beginners include:

- [Is Your Data Story Actually A Story? | by Joshua Smith | Nightingale](#)
- [I Learned Data Viz in a Year, and You Can Too](#)
- [Endless River: An Overview of DataViz for Categorical Data](#)
- [How to Choose Colors for Data Visualizations | Tutorial by Chartio](#) (originally published on Nightingale)

Tableau

[Tableau](#) is a popular visualization platform great for creating shareable dashboards and analyzing data. Students can access [Tableau for free](#)!

Intro resources to Tableau:

- [Tableau - What is data visualization?](#)
- [Tableau Public Resources](#)
- [Tableau Tutorial For Beginners | Data Visualization Tableau](#)

Python

Intro resources to Python:

- [Python Data Visualization Tutorials](#)
- [Introduction to Data Visualization in Python](#)
- [Python Data Visualization Tutorial: Matplotlib & Seaborn Examples](#)
- [Data Analysis w/ Python](#)
- [Gentle Introduction to Data Visualization Using Python](#)

Numpy & Pandas

- To get started, check out: [NumPy's Official Documentation and Tutorials](#)
- For a more detailed guide: [Guide to NumPy by Travis E. Oliphant](#)
- For beginners: [Introduction to Pandas Data Structures](#)
- Machine Learning Cheat Sheet (for code snippets): <https://chrisalbon.com/>

R

[R](#) is a language specifically designed for statistics and graphical representations of data.

- Start with this Medium article from Towards Data Science: [A Comprehensive Guide to Data Visualisation in R for Beginners](#)
- Learn more about what R can do (with the addition of ggplot) here: [Data visualization with R and ggplot2 | the R Graph Gallery](#)
- Dive into things with this book on data visualization in R: [Data Visualization with R](#)
- [Introduction to R & Data Visualization - IRB Chairman Vito Leonardo](#)

ArcGIS

[ArcGIS](#) is a great tool if one is thinking of creating a map or doing more geographical analysis. Learn more about the platform and gain access to various lessons/tutorials [here](#).

Other ArcGIS resources:

- [ArcGIS Online Resources | Tutorials, Documentation, Videos & More](#)

If you don't have access to ArcGIS, QGIS is a great open-source option!

PowerBI

[PowerBI](#) is a great way to create interactive dashboards and reports. Check out a quick guide [here](#) and go through this tutorial [here](#).

Google Data Studio

[Google Data Studio Overview](#) - Google Data Studio allows one to compile their visualizations into an interactive report (much like Tableau).

OBS Studio for recording videos

Here are a few videos that can help record videos using OBS Studio

1. <https://www.youtube.com/watch?v=QKmrDUJFRkM>
2. <https://www.youtube.com/watch?v=RtDQDbPTd9E>
3. <https://www.youtube.com/watch?v=BMXQ4yD0KMc>

QUERIES

If you encounter any problem or have any queries, you can write to us at care@pickl.ai

For more details on the problem statement, dataset, and other relevant information, please go through the document [Bank Case Study](#)