Waze User Churn Project

Executive Summary

This PACE (Plan, Analyze, Construct, Execute) strategy document outlines the approach for developing a machine learning model to predict user churn for Waze. The project is initiated by leadership to help prevent user churn, improve user retention, and increase growth. This document will serve as a living record of decisions and reflections throughout the data analytics process.

Background

Waze is a community-driven navigation map app acquired by Google in 2013, now operating as a subsidiary of Alphabet, Google's parent company. As the newest member of the Waze data analytics team, I have been tasked to complete a project proposal for predicting user churn based on sample Waze user data.

Project Overview

Primary Objective: Build a machine learning model that accurately predicts user churn based on Waze user data.

Business Impact

An effective predictive model will enable the company to:
- Proactively identify users at risk of churning
- Implement targeted retention strategies
- Increase overall user retention rates
- Support sustainable growth objectives

Project Questions & Considerations

PACE: **Plan Stage**

- Who is your audience for this project?

  The audience will be the Waze data team and cross-functional team members.

- What are you trying to solve or accomplish? And, what do you anticipate the impact of this work will be on the larger needs of the client?

  *W*orking to develop a machine learning model to predict user churn.

- What questions need to be asked or answered?

  - What is the condition of the provided dataset?
  - What variables will be the most useful?
  - Are there trends within the data that can provide insight?
  - What steps can I take to reduce the impact of bias?

- What resources are required to complete this project?

  Python notebook, and input from stakeholders.

- What are the deliverables that will need to be created over the course of this project?

  - *t*he deliverables include *a* dataset scrubbed for exploratory data analysis
  - Visualizations
  - statistical model
  - regression analysis and/or machine learning model.
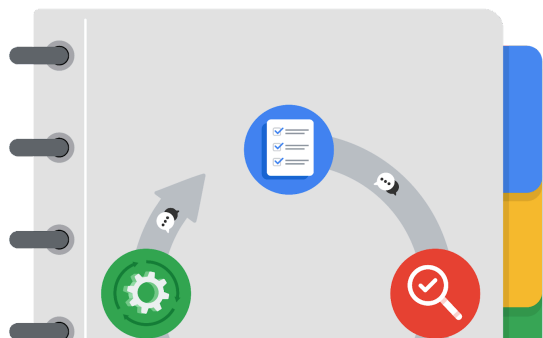
THE PACE WORKFLOW

Project tasks

Following are a group of tasks your company's data team has determined need to be completed within this project. The data analysis manager has asked you to organize these tasks in preparation for the project proposal document. First, identify which stage of the PACE workflow each task would best fit under using the drop down menu. Next, give an explanation of why you selected the stage for each task. Review the following readings to help guide your selections and explanation: The PACE stages and Communicate objectives with a project proposal. You will later reorder these tasks within a project proposal.

1. **Evaluating the model:** Execute ▾

   Why did you select this stage for this task?

   > After the model has been constructed, data is run through to evaluate whether it meets the project's expectations and goals.

2. **Conduct hypothesis testing:** Analyze ▾ and Construct ▾

   Why did you select these stages for this task?

   > During the analyzing stage, it is determined that a statistical test will be used. During the construction phase, the test is carried out.

3. **Begin exploring the data:** Analyze ▾

   Why did you select this stage for this task?

   > During the analysis phase, you will gain a deeper understanding of the dataset and the information within it.

4. **Data exploration and cleaning:** Plan ▾ and Analyze ▾

   Why did you select these stages for this task?

   > Planning takes place when you first make choices about the methods needed. The cleaning process then takes place in the analyzing stage.

5. **Establish structure for project workflow (PACE):** Plan ▾

Why did you select this stage for this task?

Planning stage. Creating an initial project PACE document outlines the workflow and helps to plan how to best approach a project.

6. **Communicate final insights with stakeholders:** Execute ▾

Why did you select this stage for this task?

Communication is necessary at various points throughout a project. Final insights are shared with stakeholders in the execute phase of the data project workflow.

7. **Compute descriptive statistics:** Analyze ▾

Why did you select this stage for this task?

Investigating the statistics within data takes place during analysis.

8. **Visualization building:** Analyze ▾ and Construct ▾

Why did you select these stages for this task?

Visualization begins with data assessment and is created during the construction stage.

9. **Write a project proposal:** Plan ▾

Why did you select this stage for this task?

Planning stage. A project proposal is the initial document used to define a project.

10. **Build a regression model:** Analyze ▾ and Construct ▾

Why did you select this stage for this task?

During the analyzing stage, the model is examined in detail to be sure it will meet the needs of the task. The building of the regression model will take place in the construction phase.

11. **Compile summary information about the data:** `Analyze ▾`

Why did you select this stage for this task?

> Inspecting a dataset to compile information would take place in the analysis phase.

12. **Build machine learning model:** `Construct ▾`

Why did you select this stage for this task?

> The building of a data model would take place in the construct stage.

Questions to Always keep in mind

- As a new member of a data analytics team, what steps could you take to get 'up to speed' with a current project? What steps would you take? Who would you like to meet with?

- How would you plan an analytics project?

- What steps would you take to translate a business question to an analytical solution?

- Why is actively managing data an important part of a data analytics team's responsibilities?

- What are some considerations you might need to be mindful of when reporting results?