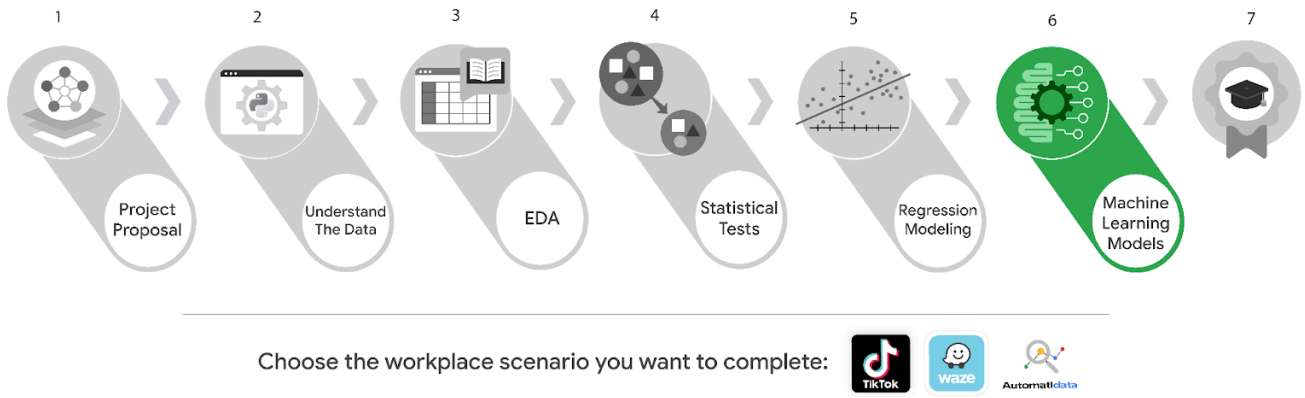
# Reading: Explore your Course 6 workplace scenarios

## Overview

This certificate offers you a choice of several different workplace scenarios to use when completing each end-of-course project:

- Automatidata, featuring a fictional data consulting firm
- TikTok, created in partnership with the short-form video hosting company
- Waze, created in partnership with the realtime driving directions app

Each scenario offers you an opportunity to apply your skills and create work samples to share when applying for jobs; so, you will be practicing similar skills regardless of the workplace scenario. It is recommended that you work with the same scenario for each end-of-course project to have a cohesive experience. However, you are welcome to investigate any of the workplace scenarios you are interested in as you progress through the program.



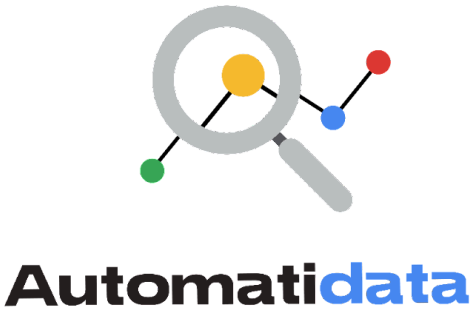

***Reminder:*** *We recommend that you choose one workplace scenario to follow for all end-of-course projects to ensure end-to-end project development.*

The minimum requirement to earn your Advanced Data Analytics Certificate is to complete the end-of-course project, using one workplace scenario, for each course. You may complete the project for as many of the workplace scenarios as you wish. Completing the project for more than one workplace scenario in a single course offers you additional practice and work examples you can add to your portfolio and share with prospective employers during your job search.

This reading offers an overview of all available workplace scenarios. Before moving on, identify the scenario you would like to complete for the Course 6 end-of-course project.

# Course 6 workplace scenarios

## Automatidata




**Project goal:**
In this fictional scenario, the New York City Taxi and Limousine Commission (TLC) has approached the data consulting firm Automatidata to develop an app that enables TLC riders to estimate the taxi fares in advance of their ride.

**Background:**
Since 1971, TLC has been regulating and overseeing the licensing of New York City's taxi cabs, for-hire vehicles, commuter vans, and paratransit vehicles.

**Scenario:**
New York City TLC stakeholders have been impressed with the data analytical work completed by the Automatidata team in this project. As a result, they have reached out once again for assistance in creating a machine learning model that can help predict whether or not a rider will be a generous tipper.

**Course 6 tasks:**
To complete this task, you'll need to:

- Build a random forest model from the New York City TLC dataset
- Create an executive summary for the Automatidata data team before sharing the results with the client

***Note:*** *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. And, the data shared in this project has been created for pedagogical purposes.*

## TikTok

**Project goal:**
The TikTok data team is developing a machine learning model for classifying claims made in videos submitted to the platform.

**Background:**
TikTok is the leading destination for short-form mobile video. The platform is built to help imaginations thrive. TikTok's mission is to create a place for inclusive, joyful, and authentic content–where people can safely discover, create, and connect.

**Scenario:**
The data team at TikTok is nearing the end of the claims classification project. The final milestone left for the team: creating the machine learning model. You will be responsible for leading these final tasks, which include feature engineering, model development, and evaluation.

**Course 6 tasks:**

- Import relevant packages and TikTok data
- Exploratory data analysis
- Feature engineering
- Check model assumptions
- Model building
- Model evaluation
- Summarize findings for cross-departmental stakeholders within TikTok

***Note:*** *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. And, the data shared in this project has been created for pedagogical purposes.*

## Waze

**Project goal:**
Waze leadership has asked your data team to develop a machine learning model to predict user churn. Churn quantifies the number of users who have uninstalled the Waze app or stopped using the app. This project focuses on monthly user churn. An accurate model will help prevent churn, improve user retention, and grow Waze's business.

**Background:**
Waze's free navigation app makes it easier for drivers around the world to get to where they want to go. Waze's community of map editors, beta testers, translators, partners, and users helps make each drive better and safer.

**Scenario:**
Your team is close to completing their user churn project. Previously, you completed a project proposal, used Python to analyze and visualize Waze's user data, conducted a hypothesis test, and built a binomial logistic regression model. The final step is to build and test different machine learning models for predicting user churn.

**Course 6 tasks:**

- Perform feature engineering
- Build the following machine learning models: random forest and XGBoost
- Evaluate the models
- Share an executive summary with the Waze leadership team

***Note:*** *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. And, the data shared in this project has been created for pedagogical purposes.*

# Key Takeaways

In Course 6, The Nuts and Bolts of Machine Learning you were introduced to supervised and unsupervised machine learning techniques. Additionally, you developed a working familiarity with Naive Bayes, decision trees, random forests, gradient boosting machines, and K-means clustering.

## Course 6 skills:

- Apply feature engineering techniques using Python
- Construct a Naive Bayes model
- Articulate how unsupervised learning differs from supervised learning
- Code a K-means algorithm in Python
- Explore decision tree and forest models
- Understand bagging and boosting
- Identify tuning model parameters and how they affect performance and evaluation metrics
- Share Insights and Ideas with Stakeholders

## Course 6 end-of-course project deliverables:

- Machine learning  model with Python
- Executive summary with recommendations and insights based on results

The end-of-course portfolio projects are designed for you to apply your data analytical skills within a workplace scenario. No matter which scenario you work with, you will practice your ability to discuss data analytic topics with coworkers, internal team members, and external clients.

As a reminder, you are required to complete one project for each course. To gain additional practice, or to add more samples to your portfolio, you may complete as many of the scenarios as you wish.

---