Course 1 workplace scenarios

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.

TikTok users have the ability to report videos and comments that contain user claims. These reports identify content that needs to be reviewed by moderators. This process generates a large number of user reports that are difficult to address quickly.

TikTok is working on the development of a predictive model that can determine whether a video contains a claim or offers an opinion. With a successful prediction model, TikTok can reduce the backlog of user reports and prioritize them more efficiently

Course 1 tasks:

- Gather information from stakeholder notes from within TikTok
- Assign PACE stages to the requested tasks for the classification project
- Organize tasks into milestones
- Create a project proposal for the TikTok data team

Course 2 workplace scenarios

Scenario:

As a data analyst on TikTok's data team, you'll help by preparing the data needed for the claims classification project. You'll build a dataframe, organize the claims data for the process of exploratory data analysis, and update the team on your progress and insights.

Course 2 tasks:

- Build a dataframe for the TikTok dataset
- Read in data from TikTok csv file
- Display rows within dataframe
- Examine data type of each column
- Gather descriptive statistics
- Visualize the TikTok data in Python
- Report to TikTok's data team through an executive summary

Course 3 workplace scenarios

Scenario:

It is now time to begin the process of exploratory data analysis (EDA). As a data analyst on TikTok's data team, you will complete the EDA process for the claims classification project. You'll also use Tableau to create visuals for an executive summary to help non-technical stakeholders engage and interact with the data.

Course 3 tasks:

- Imports of relevant packages and TikTok data into Python
- EDA and cleaning

- Assess Tableau measures and dimensions
- Select and build visualization(s) type
    - Create plots to visualize variables and relationships between variables
- Share your results with the TikTok team

Course 4 workplace scenarios

Scenario:

The TikTok data team has successfully completed exploratory data analysis on the data for the claims classification project. The team is ready to begin the process of hypothesis testing. You've been asked to investigate TikTok's user claim dataset to determine which hypothesis testing method best serves the data and the claims classification project.

Course 4 tasks:

- Import relevant packages and TikTok data
- Explore the project data
- Implement a hypothesis test
- Communicate insights with stakeholders within TikTok

Course 5 workplace scenarios.

Scenario:

The data team at TikTok is close to their goal of building a model to assist in the classification of claims in videos. The next step is to use the project data to create a regression model. As a member of TikTok's data team, you'll determine the type of regression model

that is needed and develop one using TikTok's claim classification data.

Course 5 tasks:

- Import relevant packages and TikTok data
- Exploratory data analysis and check model assumptions
- Determine the correct modelling approach
- Build the regression model
- Finish checking model assumptions
- Evaluate the model
- Interpret model results and summarize findings for cross-departmental stakeholders within TikTok

Course 5 workplace scenarios.

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