

ASA DATAFEST 2024

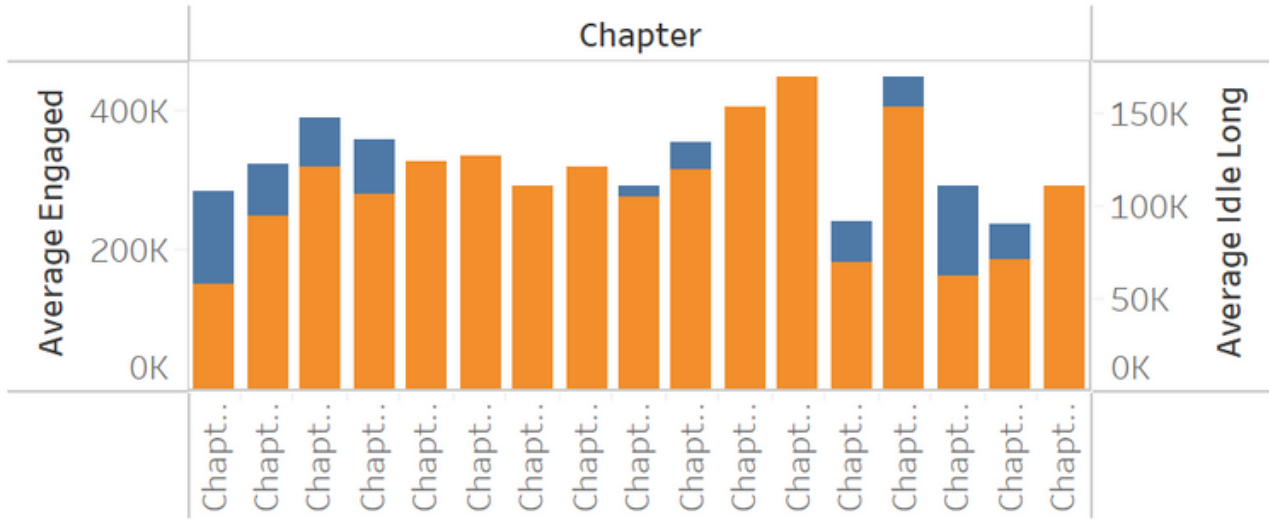
PRESENTED BY: TEAM 9

MEMEBER: BO XUN SHAO, JOSEPH WILSON, JACK
OEBKER, VIJAY SITHAMBARAM

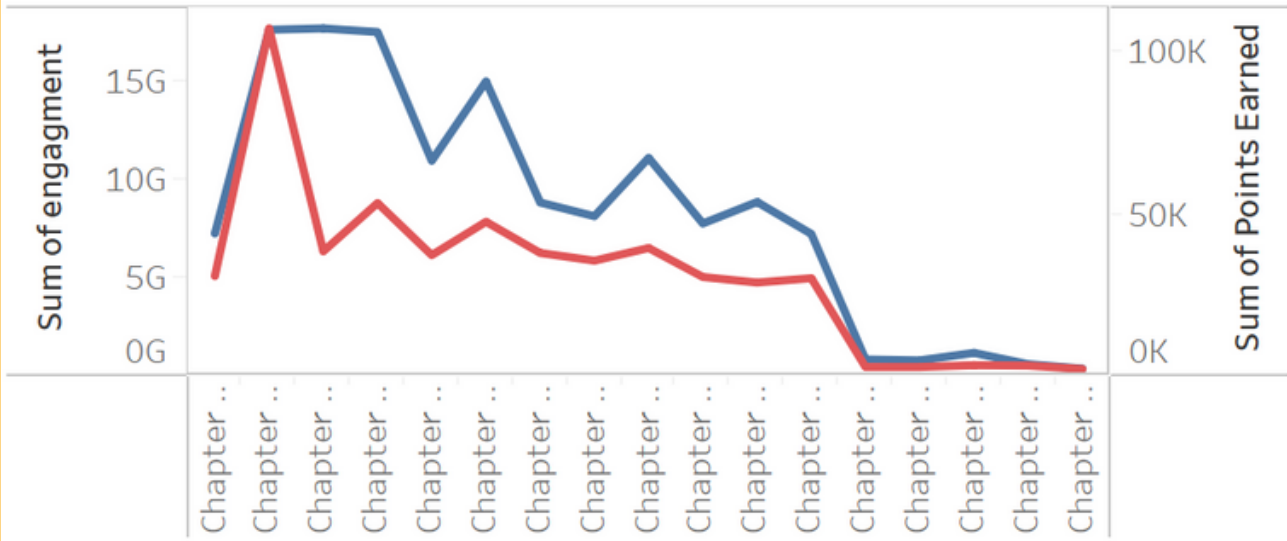
SUGGESTION

- Pop out some multiple choice questions between videos. When the question comes out, it will pasue the video, students need to answer it in order to let video keep running. There are already some platfroms using this method, such as coursera.

Average Engaged v.s Average Idle Long

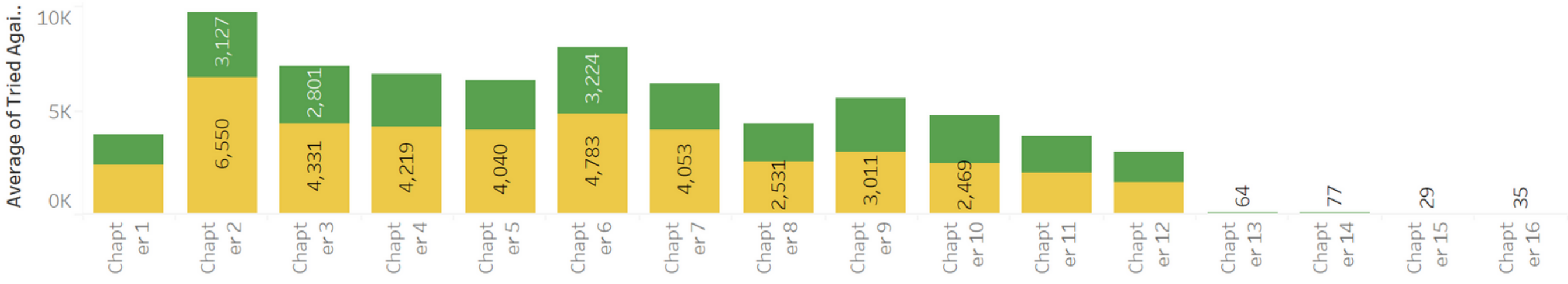


Engaged v.s Points Earned

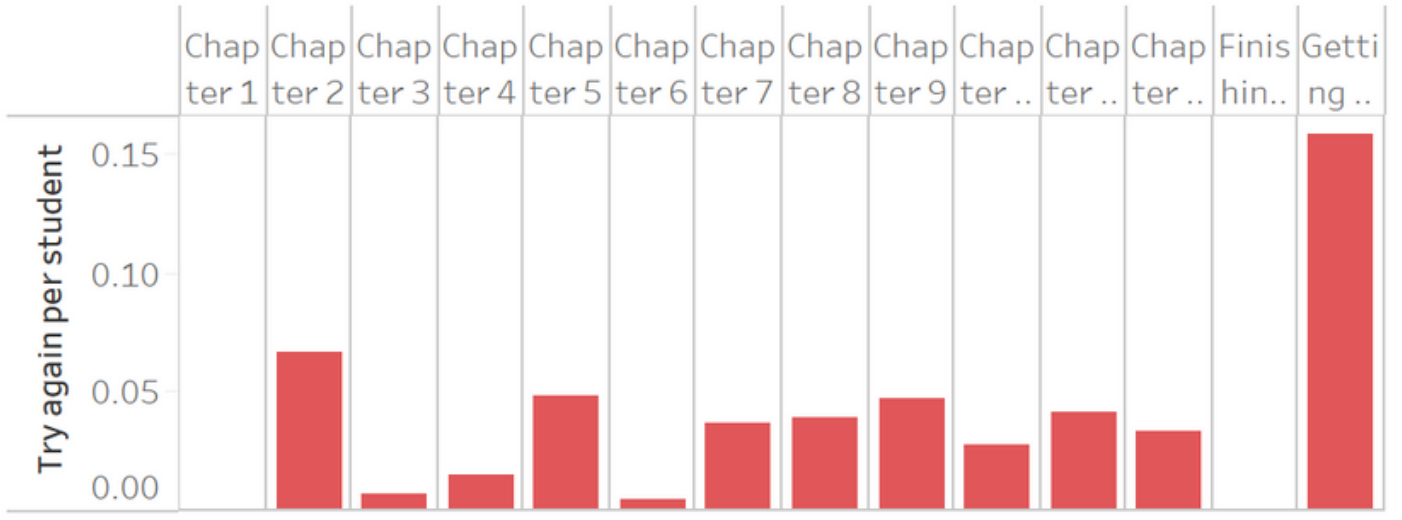


Number of Try Again over chapters

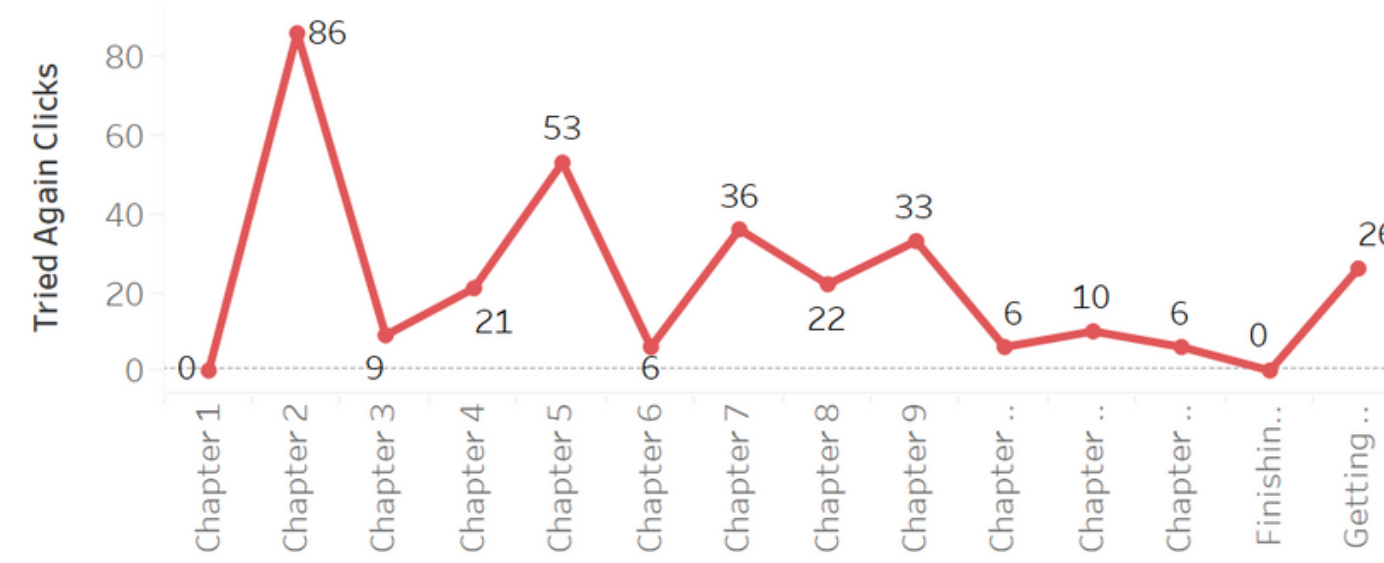
Advance Standaed



Average of Try Again over chapters and Text Book



Number of Try Again over chapters in High School



PREDICTING MODEL

- Logistic regression is a statistical model used for binary classification tasks, which can be extended to multiclass classification, making it suitable for categorizing students into High, Medium, and Low Achievers based on our requirement. We would then compare our prediction against students' actual score to verify our prediction model.

```
# Fit the GridSearchCV on the training data
grid_search.fit(X_train, y_train)

# Get the best parameters from GridSearchCV
best_params = grid_search.best_params_
print("Best Hyperparameters:", best_params)

# Get the best model from GridSearchCV
best_model = grid_search.best_estimator_

# Predict using the best model on training and testing sets
y_train_pred_gb = best_model.predict(X_train)
y_test_pred_gb = best_model.predict(X_test)

# Print classification report for training set
print("Classification Report for Training (Logistic Regrssiog):")
print(classification_report(y_train, y_train_pred_gb, target_names=category_labels))

# Print classification report for training set
print("Classification Report for Training (Logistic regression):")
print(classification_report(y_train, y_train_pred_gb))

# Print classification report for testing set
```

