Analysis of User Adoption Feature Importance

This task analyzes the provided data to identify factors predicting user adoption, defined as a user logging into the product on three separate days in at least one seven-day period. Full notebook here.

Methodology:

1. **Data Preparation:**
    - Merged 'takehome_users.csv' and 'takehome_user_engagement.csv' based on user ID.
    - Identified ‘adopted’ users (target column)
    - Created numeric dataframe for classification using dummies.
    - Imputed missing values in 'invited_by_user_id' with -1, representing no invitation.
2. **Feature Importance Analysis:**
    - Employed permutation importance from an XGBClassifier to evaluate the impact of each feature on model performance.

Findings:

The analysis revealed the following factors as the most significant predictors of user adoption:

1. **'last_session_creation_time':** The timestamp of a user's most recent login
2. **'creation_time':** The time of account creation.
3. **'org_id':** The organization a user belongs to might suggest shared usage patterns or organizational policies affecting adoption.
4. **'invited_by_user_id':** Being invited by an existing user could imply social influence or collaborative use cases, potentially boosting adoption. However, this factor's impact may vary as the user base expands.
5. **'time_stamp':** The timestamp of each login event provides granular information about user activity and contributes to identifying adoption patterns.