User Behavior Analytics (UBA)

What It Does

- Loads and preprocesses user activity data from a CSV file
- Performs exploratory data analysis (EDA) to visualize behavior patterns
- Applies K-Means clustering to group similar user behaviors
- Flags users with behavior far from normal clusters as anomalies
- Outputs detected anomalies to a CSV file for further investigation

Key Features

- Easy setup using Python and popular libraries (pandas, scikit-learn, matplotlib)
- Visualizes clusters and anomalies using intuitive plots
- Saves clean reports for anomaly review
- Customizable to new datasets or additional behavioral metrics

Use Case Examples

- Detecting logins at unusual hours
- Identifying users who perform excessive or minimal actions
- Monitoring employees logging in from unexpected locations
- Alerting analysts to deviations from organizational behavior norms

Dataset Format

The dataset should be a CSV file named `user_behavior.csv` with the following columns:

user_id - Unique identifier for each user

login_time - Numeric value representing login hour

location - Location of login (e.g., Office, Home)

action_count - Number of actions performed by user

How to Use

1. Place your CSV file in the project directory as `user_behavior.csv`
2. Run the script:

```
python user_behavior_analytics.py
```

3. Anomalies will be saved in `anomalous_user_behavior.csv`

Requirements

Python 3.6+

- pandas
- matplotlib
- seaborn
- scikit-learn

Install with:

```
pip install -r requirements.txt
```

Contributing

Have ideas or improvements? Feel free to fork this repo, submit pull requests, or open issues.

Disclaimer

This is a basic prototype intended for educational or early-stage security monitoring. For production environments, further enhancements such as real-time streaming, richer features, and advanced anomaly detection models (e.g., Isolation Forest, DBSCAN, Autoencoders) are recommended.