



# Unsupervised Discovery of Mobile User Personas for UX Strategy

**I'm a Director of User Experience at Google,  
and an aspiring student of Artificial  
Intelligence at University of Colorado  
Boulder. I wanted to connect these two world  
and work on a problem related to User  
Experience Design.**

# I found on Kaggle a fascinating data set - Mobile Device Usage and User Behavior

*The dataset consists of: 700 records of individual user interaction patterns, including key features relevant to mobile behavior:*

- *App Usage Time (min/day): Total daily minutes spent actively using applications.*
- *Screen On Time (hours/day): Total daily time the screen is active.*
- *Battery Drain (mAh/day): Daily battery consumption.*
- *Data Usage (MB/day): Daily mobile data consumption.*
- *Age and Gender: Demographic features used to profile the final clusters.*

This screenshot shows a Kaggle dataset page for 'Mobile Device Usage and User Behavior Data'. At the top right, there's a user icon and the name 'VALA KHORASANI - UPDATED A YEAR AGO'. Below the title, it says 'Mobile Device Usage and User Behavior Data' and 'Analyzing Mobile Usage Patterns and User Behavior Classification Across Devices'. There are tabs for 'Data Card' (which is selected), 'Code (140)', 'Discussion (1)', and 'Suggestions (0)'. Under the 'About Dataset' section, it says: 'This dataset provides a comprehensive analysis of mobile device usage patterns and user behavior classification. It includes user data, such as app usage time, screen-on time, battery drain, and data consumption. Each entry represents five user behavior classes, ranging from light to extreme usage, allowing for insightful analysis and modeling.' It lists 'Key Features': User ID, Device Model, Operating System, App Usage Time, Screen On Time, Battery Drain, Number of Apps Installed, Data Usage, Age, Gender, and User Behavior Class. Below this is a 'View more' button. At the bottom, there's a file download section for 'user\_behavior\_dataset.csv (38.88 kB)' with options for 'Detail', 'Compact', and 'Column'. It also includes sections for 'About this file', 'Usage Instructions', and a note about the dataset being used for 'Analyzing mobile user behavior patterns'.

# The problem

While personas are often based on demographics, unsupervised learning can help us find them based on user behavior. This helps mitigate biases and find actionable clusters of users.

I'm going to use unsupervised learning to find behavioral clusters!

# Goal

Build an unsupervised model capable of finding personas based on user behavior.

# Research Plan

1. **EAD.** After loading the data set I'm going to thoroughly evaluate it. Checking data distribution and correlations between features.
2. **Data processing.** I'll have to standardize features and run a PCA analysis.
3. **K-means training and optimization.**
4. **Hierarchical clustering training and optimization.**
5. **Summary of results.**