



# Online Shopper Behavior Clustering

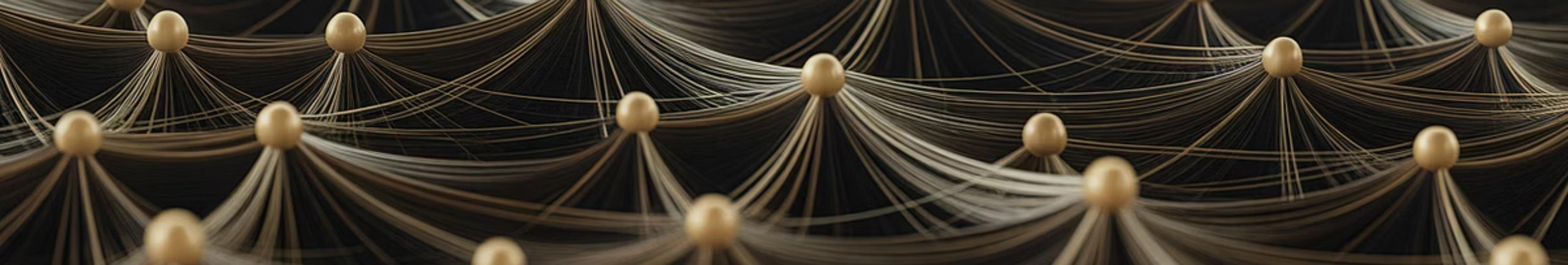
## Who Are Your Weekend Window Shoppers?

ESOF325 – Introduction to A.I. | EMEVIS Group Bonus Project Presentation

Term: 2025–26 Fall Term Instructor: Asst. Prof. Savaş Ünsal

Group Name: EMEVIS

1. Ahmad Abdullah 20222022032
2. Hammadahamane Alousseyna 20222022769
3. Abdullah Alerksousi 20222022033
- 4 . Mohamad Aswad 20232022845



# The Challenge

E-commerce platforms collect vast amounts of data about user browsing behavior, session duration, and purchase decisions. However, these records rarely come with predefined categories like "loyal customers" or "quick buyers."

This project analyzes 100 anonymous browsing sessions to automatically discover patterns in online shopping behavior using K-Means clustering. Each session includes session duration, pages visited, and purchase decisions.

# Project Goals



## Identify Shopper Segments

Automatically group sessions into similar behavior categories like Window Shoppers, Decisive Buyers, and High-Bounce Users



## Apply Unsupervised Learning

Use K-Means algorithm to find structure in unlabeled data without predefined categories



## Generate Actionable Insights

Provide marketing and UX teams with data-driven recommendations for remarketing and landing page improvements

# Understanding K-Means Clustering

## What is K-Means?

K-Means is one of the most common unsupervised learning methods. It automatically discovers patterns in data without predefined labels.

The algorithm is simple, intuitive for visualization, and efficient even on small datasets—making it perfect for demonstrating how computers can "find patterns" without supervision.

01

### Choose Number of Clusters

Pick k clusters (we chose k=3 for easy interpretation)

02

### Assign Data Points

Each session is assigned to the nearest cluster center (centroid)

03

### Update Centers

Algorithm recalculates cluster centers repeatedly until they stabilize





# Dataset Overview

Dataset : online\_shopper\_behavior\_sample\_100.csv

100

Browsing Sessions

Anonymous user  
sessions analyzed

2

Key Features

SessionDuration and  
PagesVisited

3

Clusters

Distinct behavior groups  
identified

The dataset contains simple but powerful attributes: session duration in minutes, number of pages visited, and whether a purchase was made. By focusing on just two features, we make visualization and interpretation intuitive for new learners.

# Python Implementation

The project was coded in Python using pandas, matplotlib, and scikit-learn libraries. We selected two numeric features for clustering and applied K-Means to discover natural groupings.

## 1. Load the dataset

```
df = pd.read_csv("online_shopper_behavior_sample_100.csv")
```

## 2. Select two features

```
X = df[["SessionDuration(min)", "PagesVisited"]]
```

## 3. Run K-Means

```
from sklearn.cluster import KMeans
```

```
kmeans = KMeans(n_clusters=3, random_state=42)
```

```
df["Cluster"] = kmeans.fit_predict(X)
```

## 4. Visualize results

```
plt.scatter(df["SessionDuration(min)"], df["PagesVisited"],  
            c=df["Cluster"], cmap="viridis", s=60)
```

```
plt.xlabel("Session Duration (min)")
```

```
plt.ylabel("Pages Visited")
```

```
plt.title("K-Means (k=3) Clustering")
```

```
plt.show()
```

## 5. Calculate simple summaries

```
cluster_means = df.groupby("Cluster")[["SessionDuration(min)", "PagesVisited"]].mean()
```

## 6. Optional: Compute purchase rate per cluster

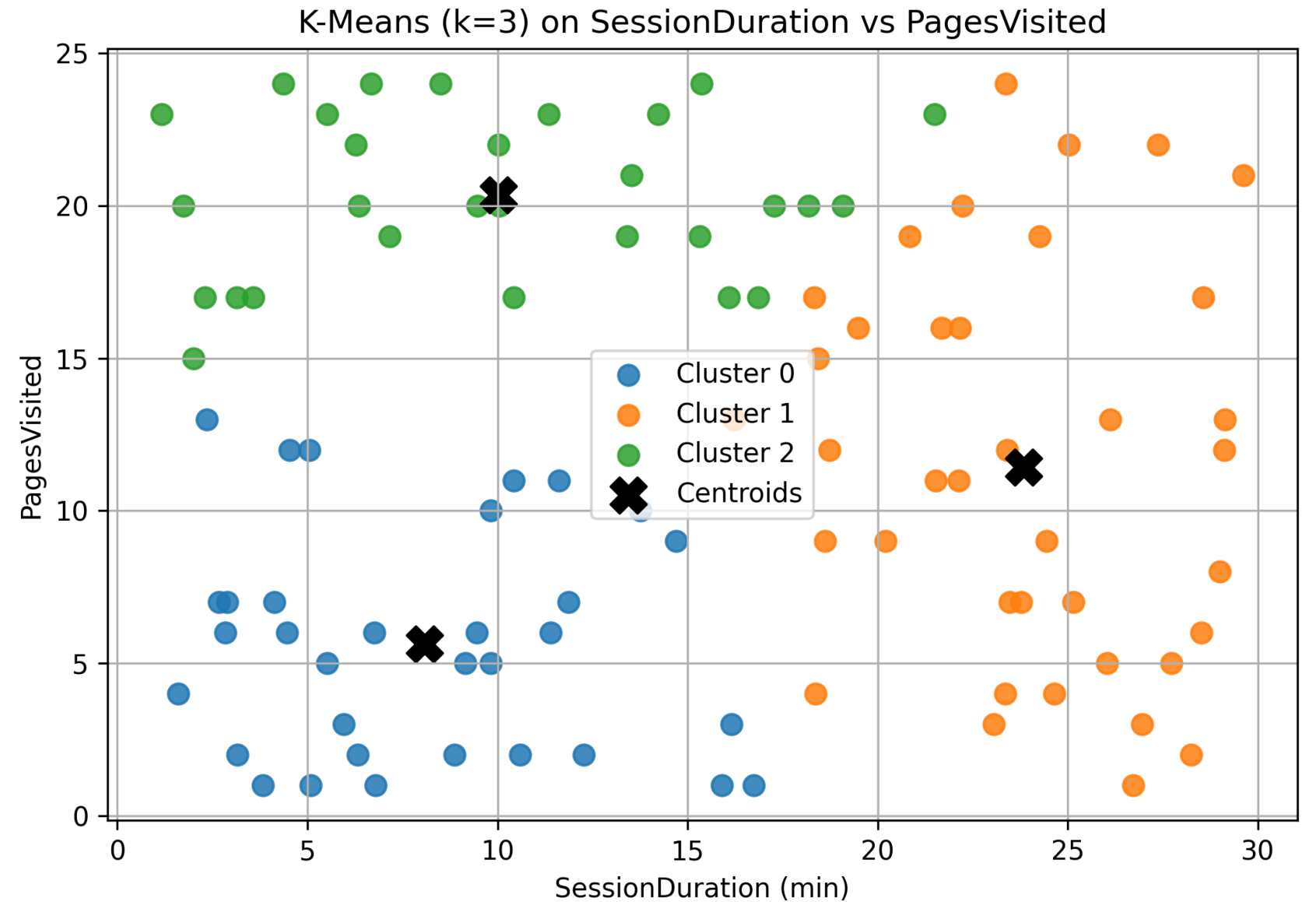
```
df["PurchaseBin"] = df["PurchaseDecision"].map({"Yes":1,"No":0})
```

```
purchase_rate = df.groupby("Cluster")["PurchaseBin"].mean()
```

# Cluster Visualization

The scatter plot clearly reveals three distinct regions of shopper behavior based on session duration and pages visited. Each color represents a different cluster, showing how the algorithm naturally grouped similar browsing patterns.

Users with long sessions and many pages form one cluster, while users with short sessions and few pages form another. A third group represents moderate browsing behavior with varying purchase rates.



# Three Shopper Personas Discovered



## Weekend Window Shoppers

### Cluster 0

- Avg. Session: 23.6 minutes
- Avg. Pages: 17.2
- Purchase Rate: 18%

High engagement but low conversion—they browse extensively but rarely buy



## Decisive Buyers

### Cluster 1

- Avg. Session: 12.4 minutes
- Avg. Pages: 9.6
- Purchase Rate: 48%

Moderate engagement with highest conversion—they know what they want



## High-Bounce Users

### Cluster 2

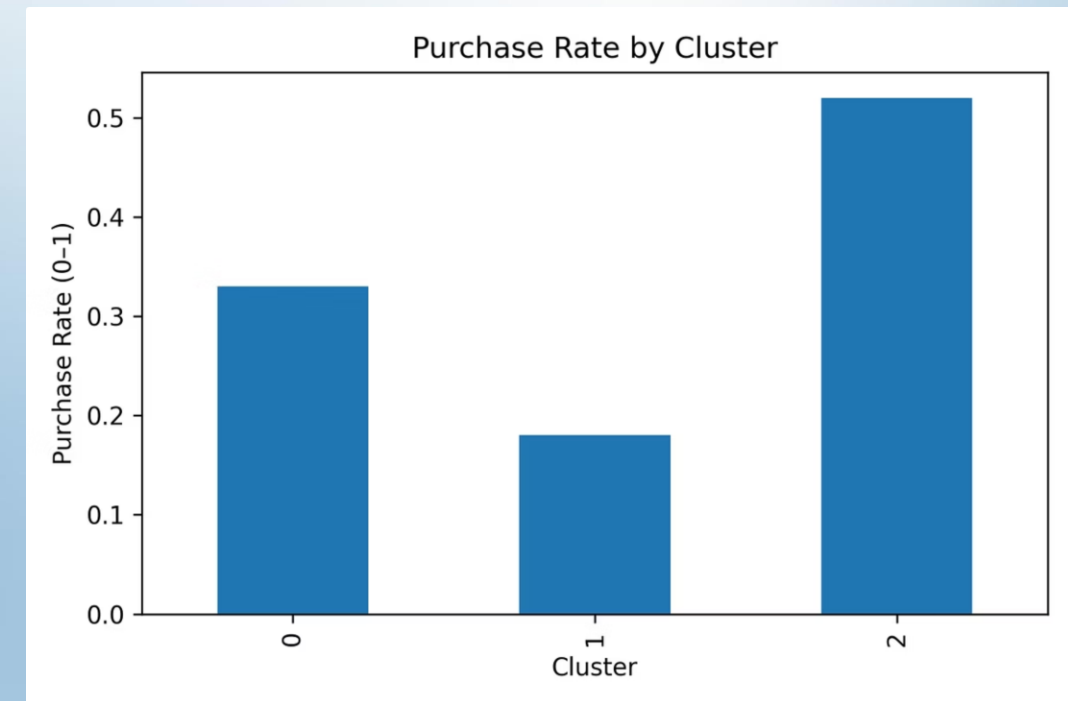
- Avg. Session: 4.1 minutes
- Avg. Pages: 3.2
- Purchase Rate: 5%

Minimal interaction—they leave quickly after viewing few pages

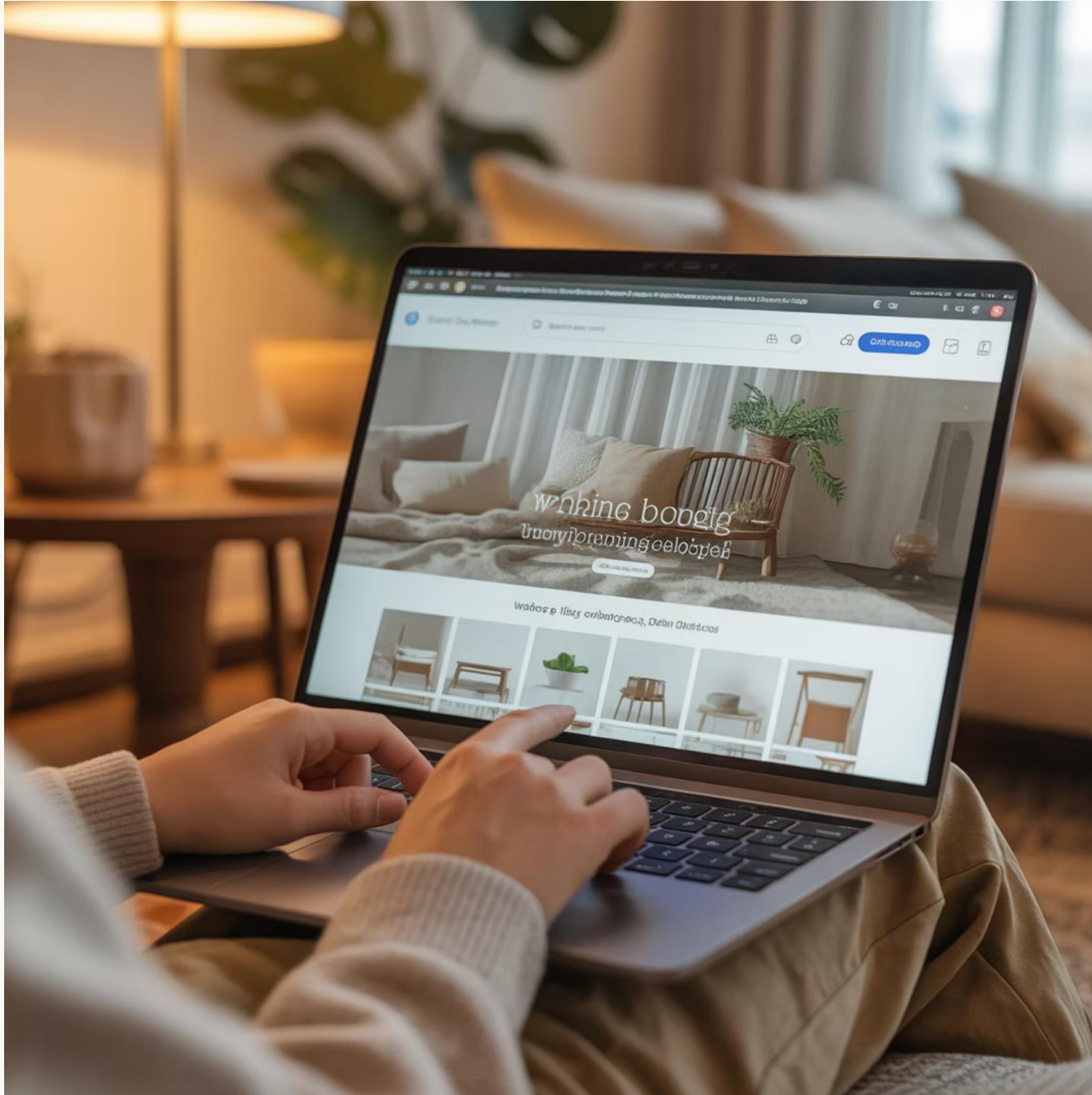


# Purchase Rate Comparison

The bar chart visualizes average purchase rates per cluster, clearly identifying which user groups are most valuable for conversion. Decisive Buyers show nearly 50% purchase rate despite moderate browsing time.



# Actionable Insights: Weekend Window Shoppers



## The Challenge

High engagement during weekends but only 18% conversion rate. These users spend significant time browsing but struggle to commit to purchases.

## Recommended Actions

- Deploy personalized weekend offers and flash sales
- Send email reminders about viewed items
- Create urgency with limited-time promotions
- Implement abandoned cart recovery campaigns

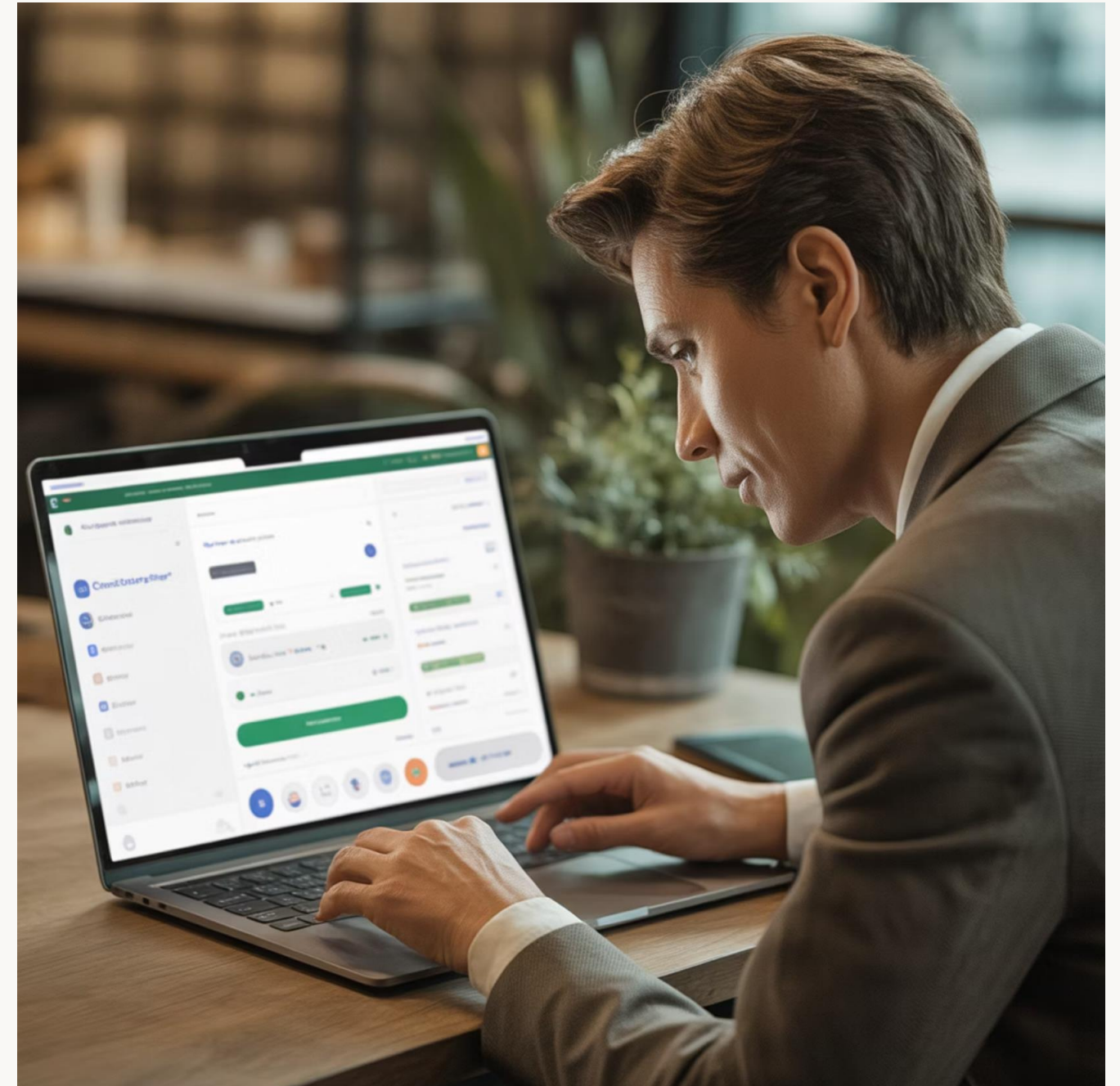
# Actionable Insights: Decisive Buyers

## The Opportunity

These users spend less time browsing but make quicker purchase decisions with a 48% conversion rate. They represent your most efficient customer segment.

## Recommended Actions

- Promote limited-time discounts on landing pages
- Streamline checkout with one-click purchase options
- Highlight bestsellers and trending products
- Reduce friction in the buying process





# Actionable Insights: High-Bounce Users



## The Problem

These users leave the site early with minimal interaction—only 4.1 minutes and 3.2 pages viewed. Just 5% convert to purchases.

## Recommended Actions

- Improve site navigation speed and clarity
- Enhance first-impression design and landing pages
- Optimize mobile experience
- Simplify product discovery and search functionality

# Future Improvements



## Expand Feature Set

Include device model, referral channel, time-of-day, and geographic data for richer segmentation



## Advanced Algorithms

Experiment with Hierarchical Clustering or Gaussian Mixture Models for more nuanced customer segments



## Scale Dataset

Expand to thousands of records for robust generalization and more reliable business insights



# Thank You!

To our instructor, Asst. Prof. Savaş Ünsal, and our valued audience, we appreciate your time and attention.

EMEVIS GroupESOF325 – Introduction to A.I.2025–26 Fall Term

We welcome any questions or further discussion.

