Customer Persona Generator

Chi-Yuan (Tina) Peng

2022 . 12 . 15

Table of Contents

1 **Introduction**

customer segmentation and persona

2 **Objectives**

my plan and thoughts

3 **Preprocessing**

data cleaning and preprocessing

4 **Modeling**

standard scaler, KMeans, PCA

5 **Segment Analysis**

segment reports analysis

6 **Persona Generator**

how I make the persona generator

Introduction 1

What is **customer segmentation**?

- a process of analyzing customers' buying behavior and clustering customers into groups.

What is **persona**?

- a character that represents the behavior and features of a larger group of customers.

2 Objectives

Collect Data → Data Preprocessing → Modeling → Segmentation → Segment Analysis → Make a Report → DONE !

Preprocessing 3

Data collection

Data cleaning and preprocessing

Collect Data

from Kaggle (customer segmentation data)

Customer Info

- Annual Income
- Year Birth
- Marital Status
- Number of Kids
- Education Level

Buying Behavior

- Amount Spend on Wines
- Amount Spend on Fruits
- Amount Spend on Meats
- Amount Spend on Seafoods
- Amount Spend on Sweets
- Amount Spend on Gold Products
- Amount Spend on Deal Products

Data Preprocessing

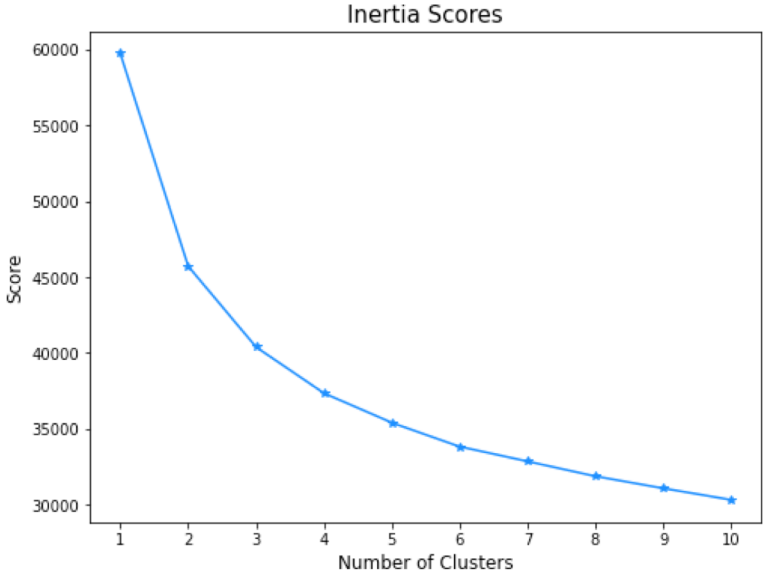
New Columns

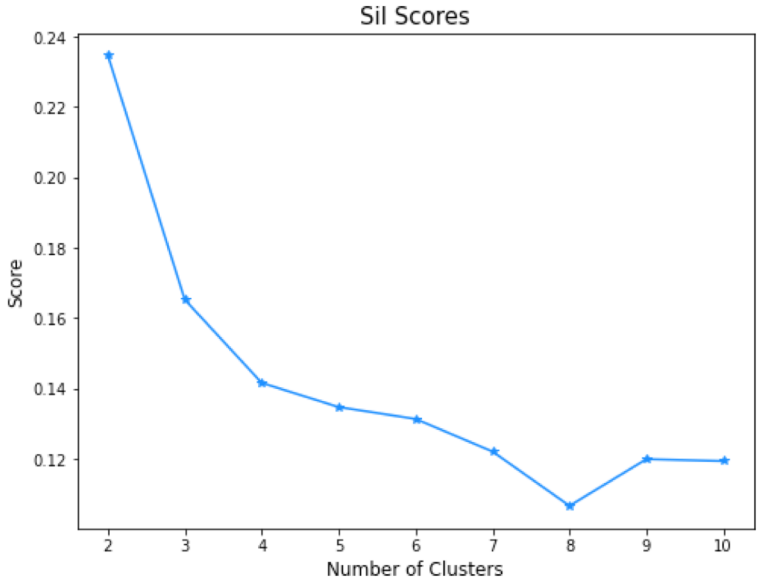
- Year Birth → Age
- Marital Status → Not Live Alone
- Number of Kids & Not Live Alone → Number of Family Members
- Marital Status & Number of Kids → Single Parent
- Amount Spent on Products → Total Spent
- Amount Spent on XX Products / Total Spent → Pr. Spent on XX Products

Modeling 4

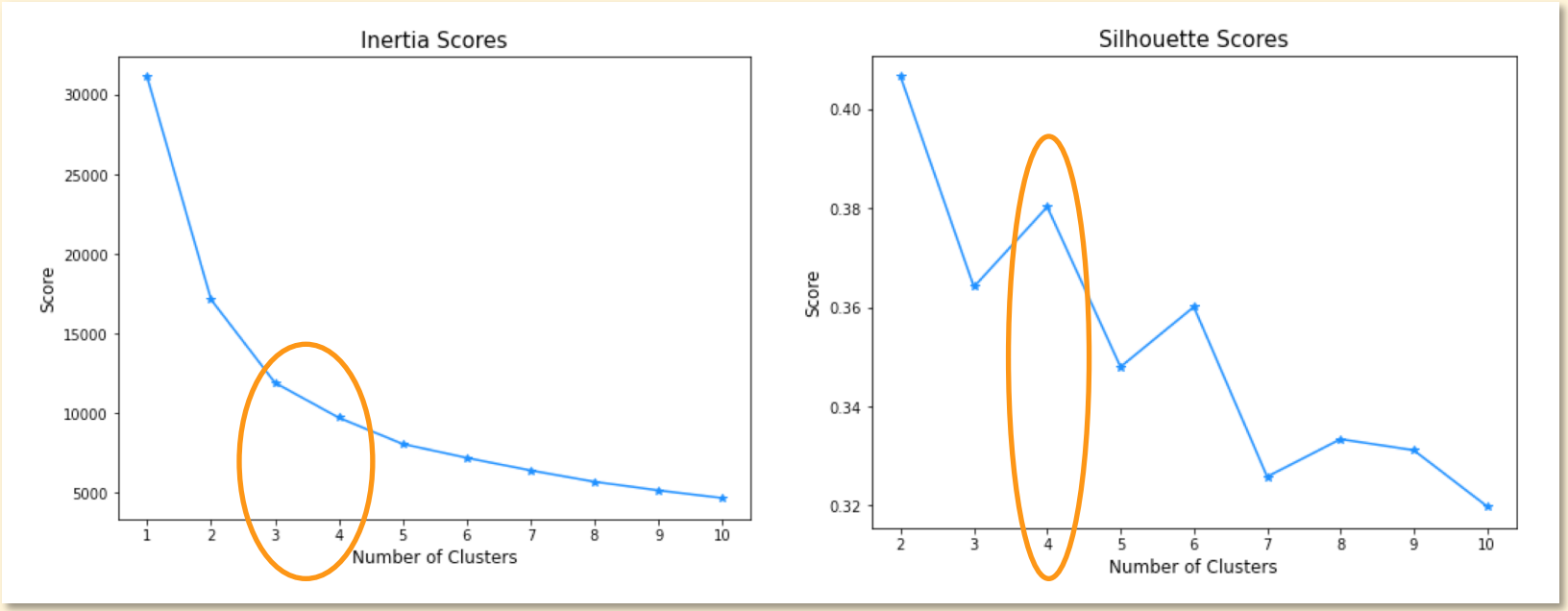
Unsupervised learning: **KMeans**

KMeans





PCA + KMeans



Segment Analysis 5

Traditional Segment Analysis

Group A – 29%

- **All** of them are **married**
- **Largest family** size
- Prefer **wine** and **on-sale** products

Group B – 18%

- **Single parents**
- Most of them has at least one kid
- Prefer **wine** and **on-sale** products

Group C – 30%

- **Richest** group with highest spending
- Most of them has **no kids**
- Prefer **meat** products

Group D – 23%

- **Lowest** income and spending
- **Youngest** group
- Prefer **seafood** and **gold** products

Automatic Segment Analysis

Group A – 29%

🛒 **All** of them are **married**

🛒 **Largest family** size

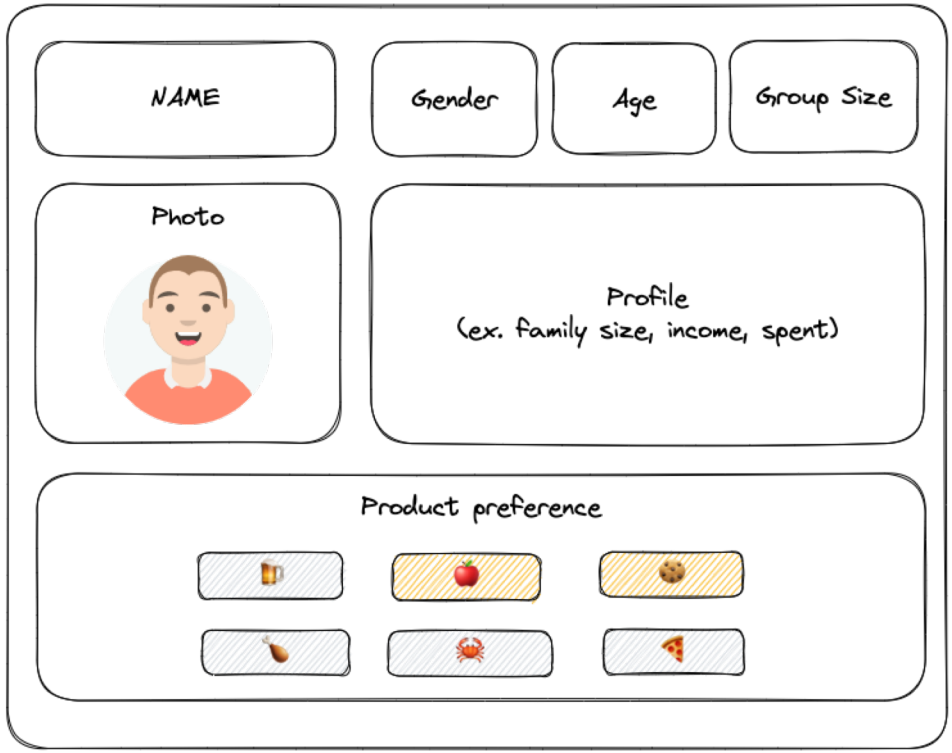
🛒 Prefer **wine** and **on-sale** products

Group A

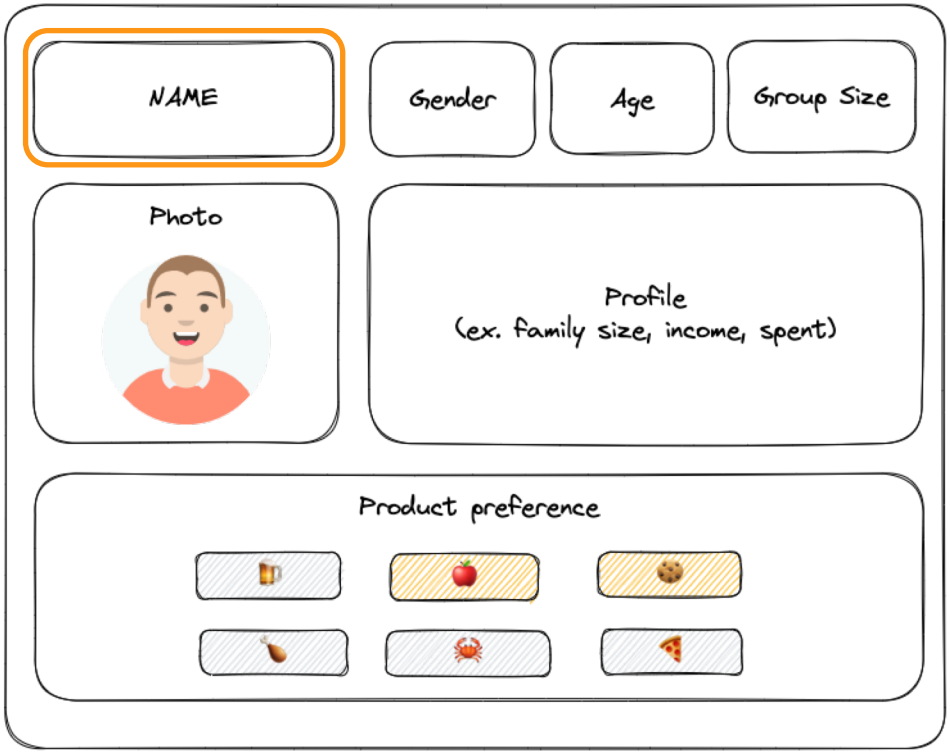
- Group size: **29%**
- Annual income: **$41,000 ~ 59,000**
- Spending amount: **$92 ~ 642**
- Age: **48 ~ 65**
- Marital status: **99% are married**
- Family size: **2 ~ 3**
- Product preference: **Wine, On-sale**

Persona Generator 6

Customer Persona Draft

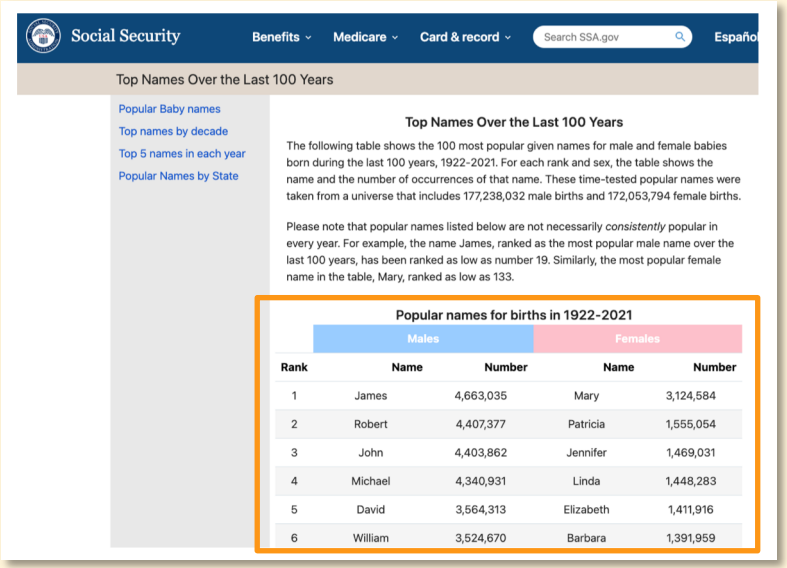


Customer Persona Draft

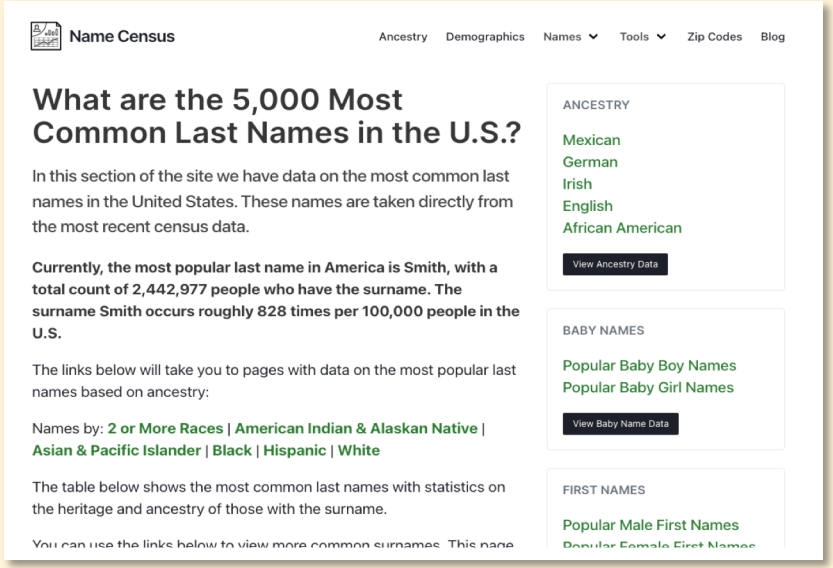


To randomly generate a name ...

Step 1:
web scraping the **first name** from SSA .

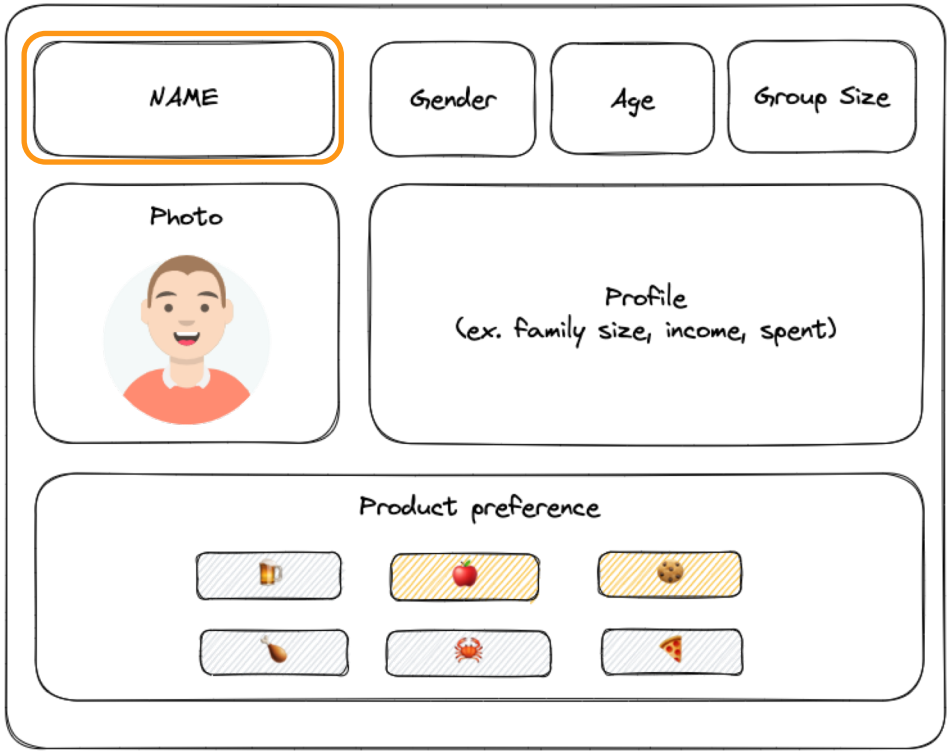


Step 2:
web scraping the **last name** from Name Census.



Customer Persona Draft

- Male first name x 100
- Female first name x 100
- Last name x 100



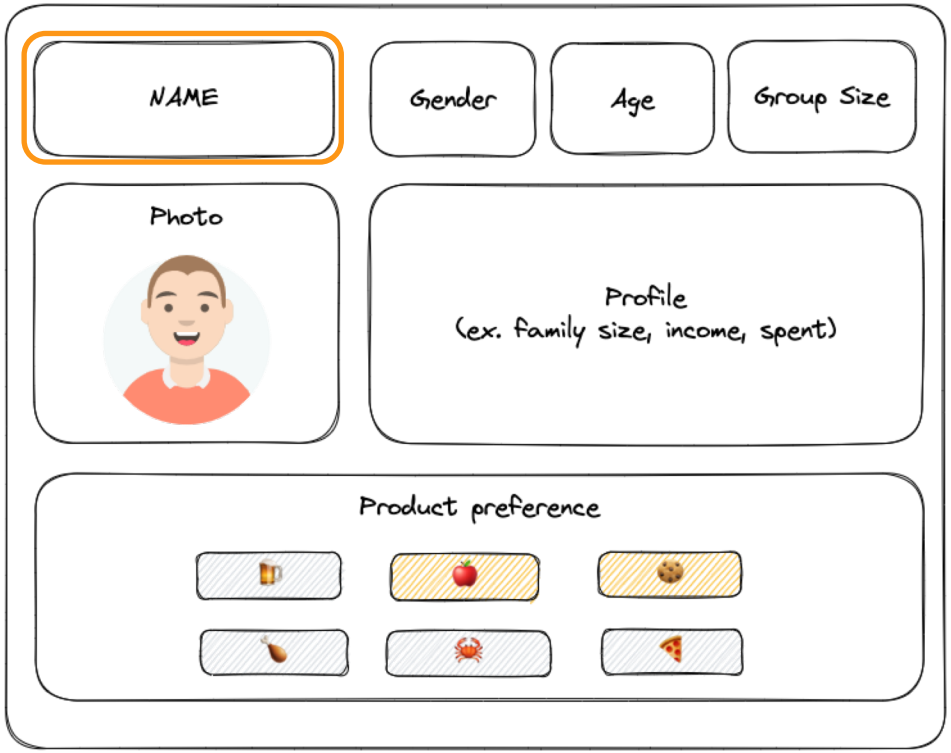
Customer Persona Draft

Male first name x 100

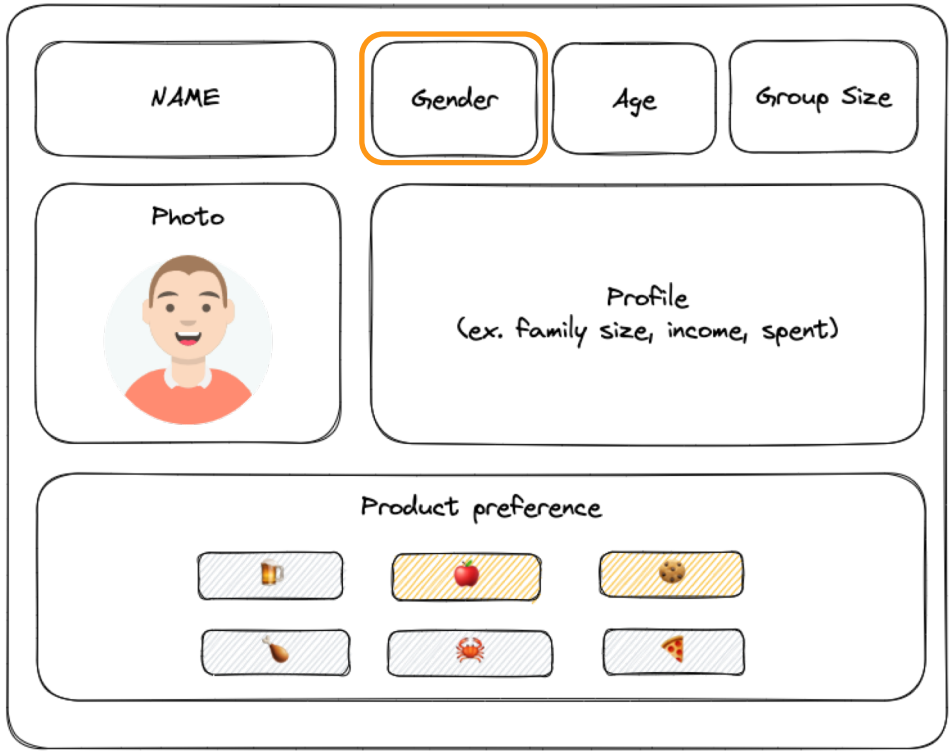
Female first name x 100

Last name x 100

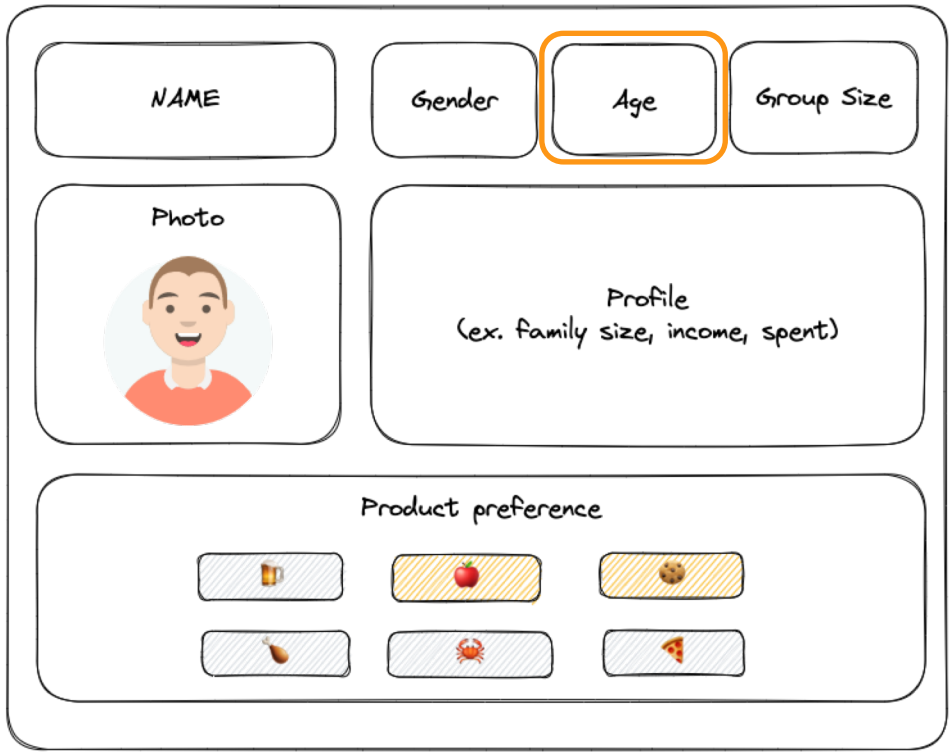
Jennifer Morris



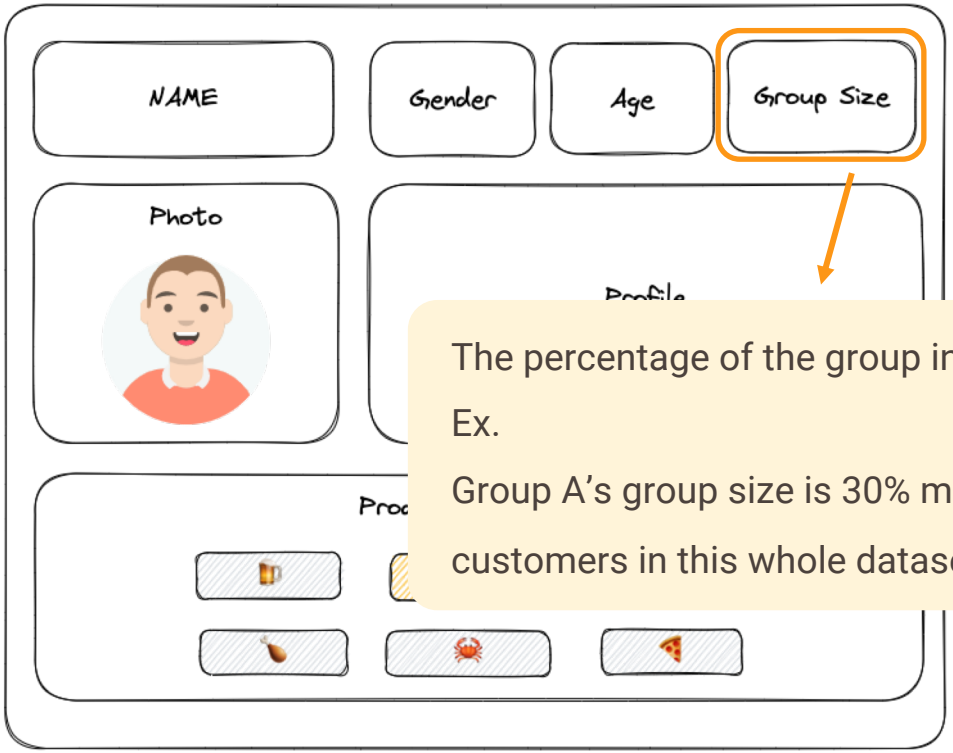
Customer Persona Draft



Customer Persona Draft



Customer Persona Draft

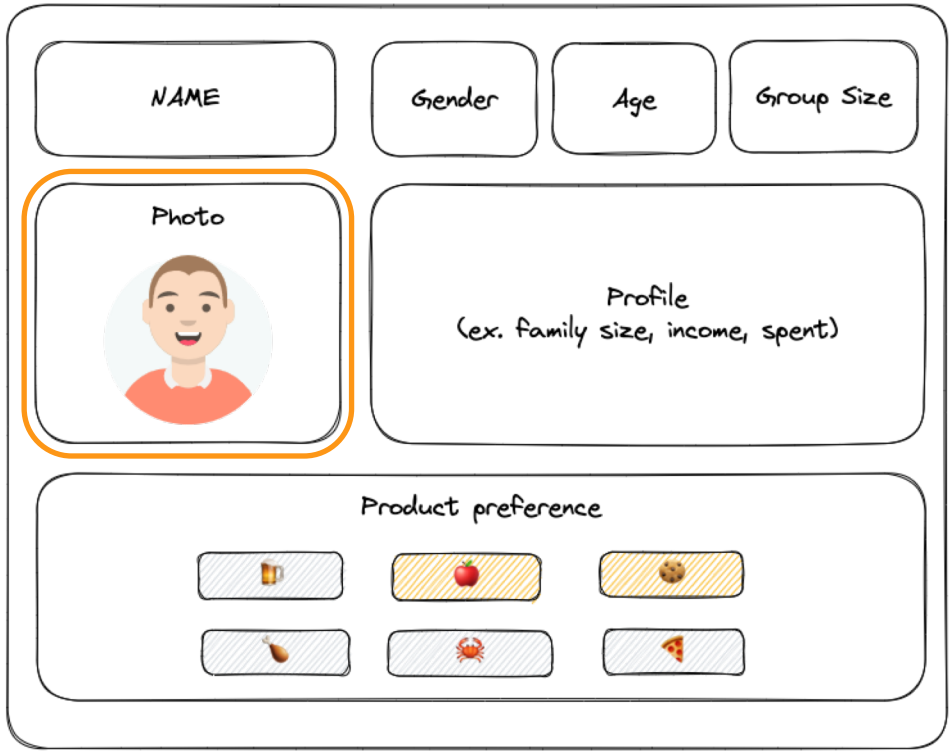


The percentage of the group in the whole dataset.

Ex.

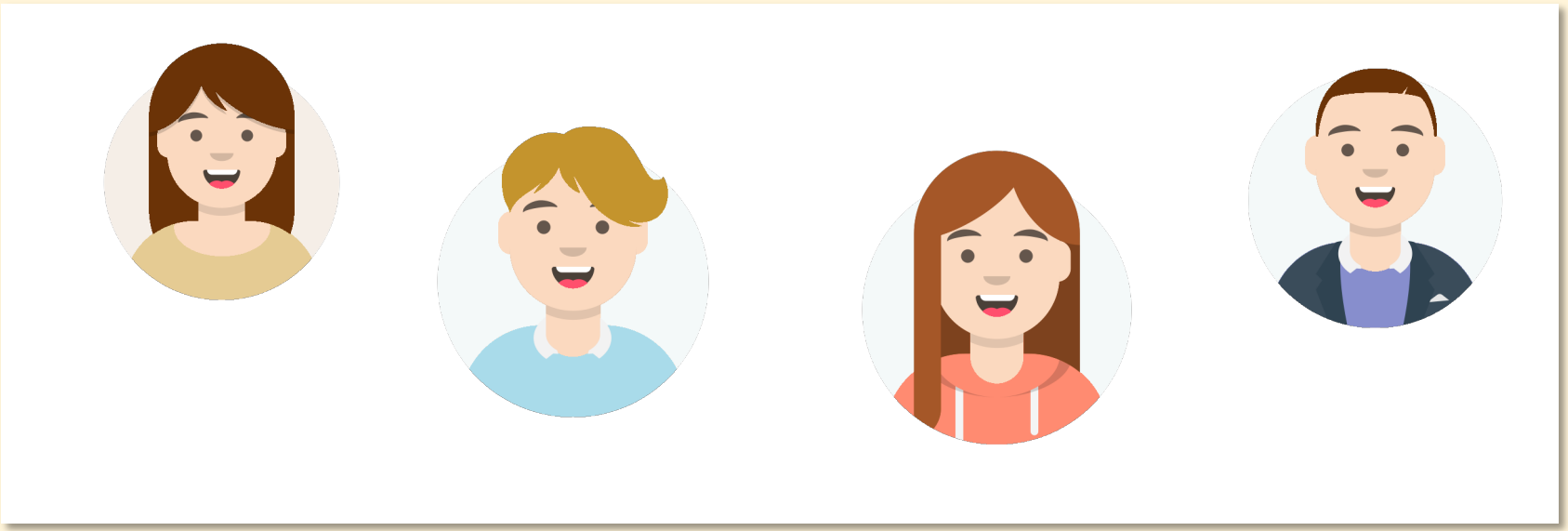
Group A's group size is 30% means that 30% of the customers in this whole dataset are labeled group A.

Customer Persona Draft

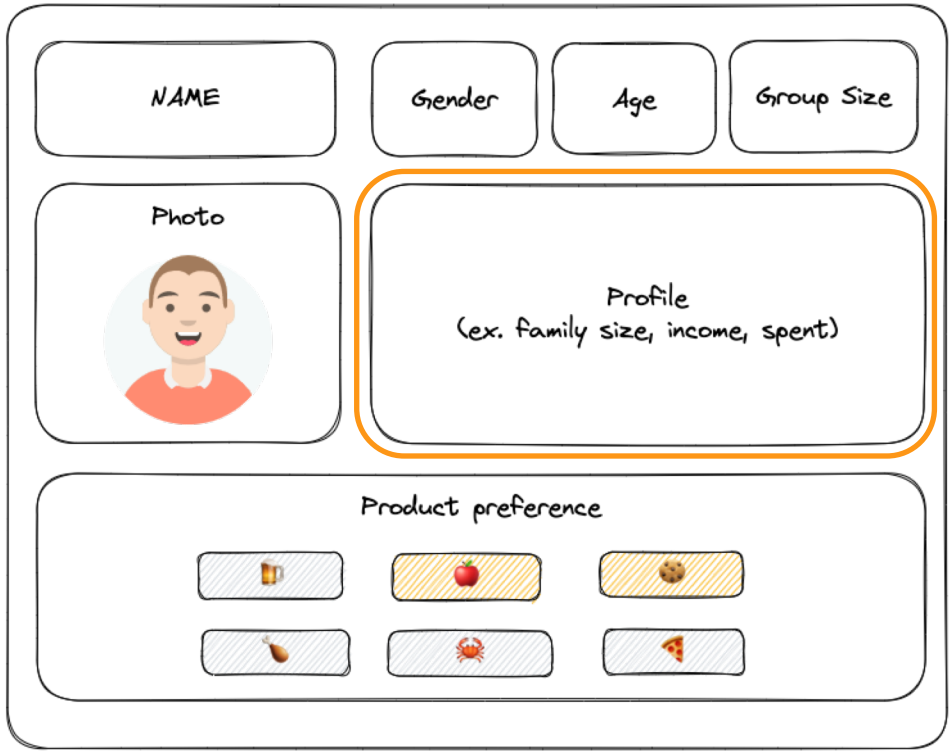


To randomly generate a persona photo ...

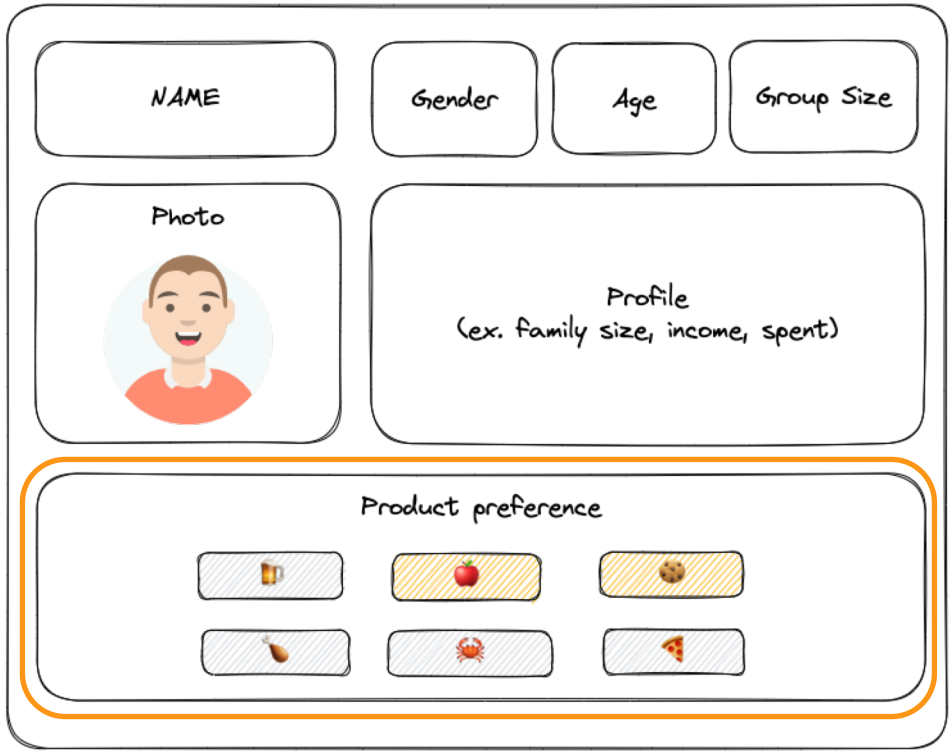
Python package: python-avatars



Customer Persona Draft



Customer Persona Draft



Persona Generator

Display through Streamlit

°○ Customer Persona ○°

Group A | Group B | Group C | Group D | Summary

NAME	GENDER	AGE	GROUP SIZE
Raymond Chavez	Male	52	29%

CUSTOMER PROFILE

- Annual income: **$50,000**
- Spending level: **Medium**
- Marital status: **Married**
- Number of family member: **3**
- Number of kids: **1**

PRODUCT PREFERENCE

Wine | Fruits | Meats | Seafood | Sweets | Gold | On-sale Products

Citations

- https://www.kaggle.com/code/karnikakapoor/customer-segmentation-clustering/data
- https://pypi.org/project/python-avatars/
- https://github.com/ibonn/python_avatars
- https://streamlit.io/
- https://www.ssa.gov/oact/babynames/decades/century.html
- https://namecensus.com/last-names/
- https://www.gartner.com/en/sales/glossary/customer-segmentation
- https://xd.adobe.com/ideas/process/user-research/putting-personas-to-work-in-ux-design/
- https://excalidraw.com/
- https://slidesgo.com/theme/flat-style-buyer-persona-infographics#search-persona&position-5&results-12&rs=search

Thank you

Any questions?