Insights:

- In the given data, 75% users are male and 25% are female
- There are 20 different product categories and 21 different types of occupations
- 40% users are between the age of 26-35 years and 28% between the age of 36-50 years
- 42% users are living in city category B, 31% in C and 27% in city category A
- 35% users are staying in current city from 1 year and 19% from 2 years
- 59% users are single and 41% are married
- Most male users are in 0, 4, 7 occupations while most female users are in 0,1,4 occupations
- Product category 1,5 and 8 are more popular in both male and female users
- Average amount per customer spend by males is more than females
- Users between the age of 26-35 years spend more than other age groups
- Customers living in city category A spend more money than others and second closer is city category B
- Customers who are living in their current city for 3 years spend more than others
- Single customers spend more than married customers
- Customers with occupation 20 spend more than others
- Married customers having occupation 8, spend higher amount than others while Single customers having occupation 16 spend highest amount
- **Average** purchase amount by **Male** customers: 925344.4
- **Average** purchase amount by **Female** customers: 712024.39

- Confidence interval for Gender:
  Using the CLT we can infer that t**he population mean 90% of the times:**
  Average amount spend by **male** customers will lie in between: **(900471.15, 950217.65)**
  Average amount spend by **female** customers will lie in between: **(679584.51, 744464.28)**

  Using the CLT we can infer that t**he population mean 95% of the times:**
  Average amount spend by **male** customers will lie in between: **(895617.83, 955070.97)**
  Average amount spend by **female** customers will lie in between: **(673254.77, 750794.02)**

  Using the CLT we can infer that t**he population mean 99% of the times:**
  Average amount spend by **male** customers will lie in between: **(886214.53, 964474.27)**
  Average amount spend by **female** customers will lie in between: **(660990.91, 763057.88)**

- **Average** purchase amount by **Single** customers: 880575.78
- **Average** purchase amount by **Married** customers: 843526.8

- Confidence interval for Marital Status:
  Using the CLT we can infer that t**he population mean 90% of the times:**
  Average amount spend by **Single** customers will lie in between: **(853938.67, 907212.9)**
  Average amount spend by **Married** customers will lie in between: **(812686.46, 874367.13)**

  Using the CLT we can infer that t**he population mean 95% of the times:**
  Average amount spend by **Single** customers will lie in between: **(848741.18, 912410.38)**
  Average amount spend by **Married** customers will lie in between: **(806668.83, 880384.76)**

Using the CLT we can infer that t**he population mean 99% of the times:**
Average amount spend by **Single** customers will lie in between: **(838671.05, 922480.51)**
Average amount spend by **Married** customers will lie in between: **(795009.68, 892043.91)**

- **Average** purchase amount by customers between the **age 0-17** years is: 618867.81
- **Average** purchase amount by customers between the **age 18-25** years is: 854863.12
- **Average** purchase amount by customers between the **age 26-35** years is: 989659.32
- **Average** purchase amount by customers between the **age 36-50** years is: 852422.43
- **Average** purchase amount by customers between the **age 51+** years is: 665729.21

- Confidence interval for different age groups:
  Using the CLT we can infer that t**he population mean 90% of the times:**
  Confidence interval for **age 0-17 years: (542553.13, 695182.5)**
  Confidence interval for **age 18-25 years: (810323.44, 899402.8)**
  Confidence interval for **age 26-35 years: (952320.12, 1026998.51)**
  Confidence interval for **age 36-50 years: (813972.08, 890872.78)**
  Confidence interval for **age 51+ years: (624772.43, 706686.0)**

  Using the CLT we can infer that t**he population mean 95% of the times:**
  Confidence interval for **age 0-17 years: (527662.46, 710073.17)**
  Confidence interval for **age 18-25 years: (801632.78, 908093.46)**
  Confidence interval for **age 26-35 years: (945034.42, 1034284.21)**
  Confidence interval for **age 36-50 years: (806469.58, 898375.29)**
  Confidence interval for **age 51+ years: (616780.86, 714677.56)**

  Using the CLT we can infer that t**he population mean 99% of the times:**
  Confidence interval for **age 0-17 years: (498811.78, 738923.84)**
  Confidence interval for **age 18-25 years: (784794.6, 924931.63)**
  Confidence interval for **age 26-35 years: (930918.39, 1048400.25)**
  Confidence interval for **age 36-50 years: (791933.47, 912911.39)**
  Confidence interval for **age 51+ years: (601297.2, 730161.23)**

Recommendations:

- Confidence intervals for male and female are not overlapping with each other and male customers spend more than female customers so Walmart should focus on retaining and getting more male customers
- Single customers spend more than married customers so Walmart should focus on acquiring more single customers
- Customers living in city categories A and B spend higher so company should aim these cities to acquire more customers
- Product category 1,5 and 8 are more popular among customers so company can focus on selling more products from these categories
- Customers with occupation 5, 16, 19 and 20, spend more than others so Walmart can aim more customers who are working in these occupations
- Male customers with occupation 9 spend more than others and female customers with occupation 8 and 19 spend more than others so Walmart can target these occupations by keeping the gender in mind
- Customers between the age of 18-50 years spend more than other age groups so company should target this age group
- There are product categories like 2, 3, 4 and 14 where female customers spend more than males so company can target these product categories for female customers