# Grocery Basket Analysis

instacart

Email: danielnuort@gmail.com Github: https://github.com/DanielNuOrt/ Phone: +4915906192998


## Objective:

Instacart, an online grocery store that operates through an app.
We need to uncover more information about their sales patterns.
Perform data and exploratory analysis of some of their data in order to derive insights and suggest strategies for better segmentation based on the provided criteria.

### Data:

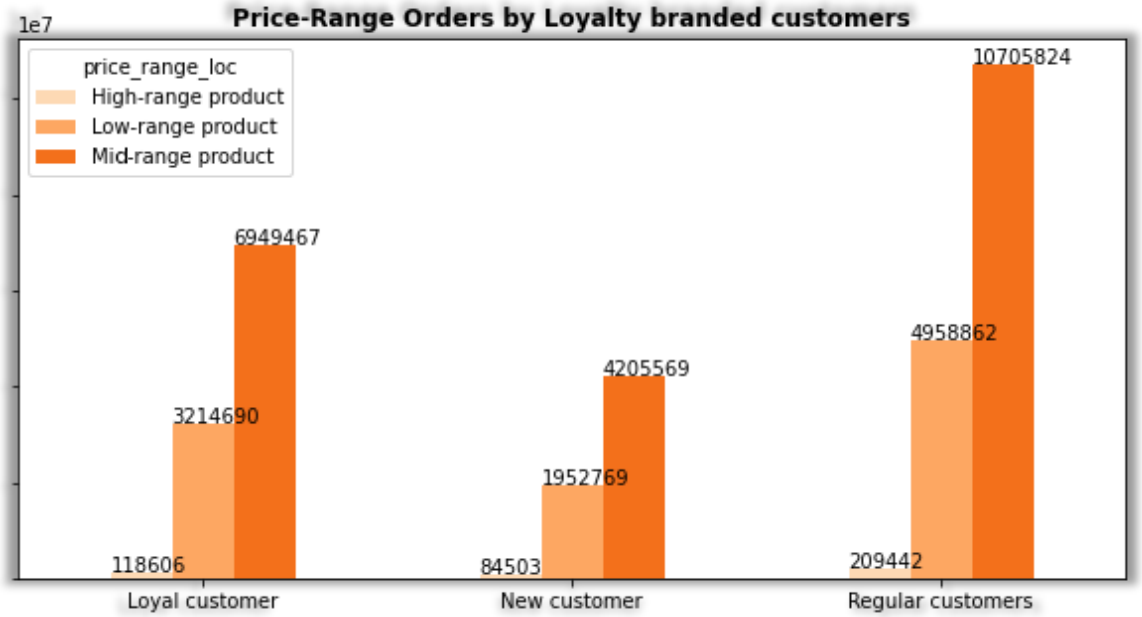
- Orders & Product Data Set1 and Data Set 2, plus from Instacart
- Customer Data Set from CareerFoundry

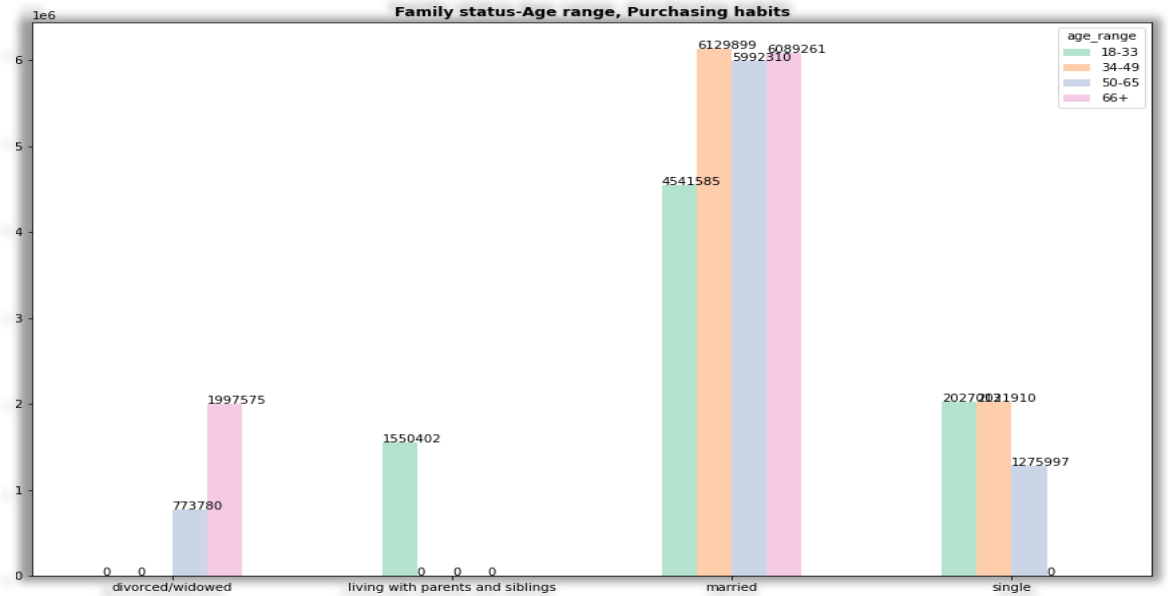
### Skills:

- Python
- Data wrangling
- Data merging
- Deriving variables
- Grouping data
- Aggregating data
- Reporting in Excel
- Population flows

### Tools:

# Profiling & Exploratory Analysis




## Findings:

Customers categorized as "regular customers" had bought higher num of product in all three price-range categories

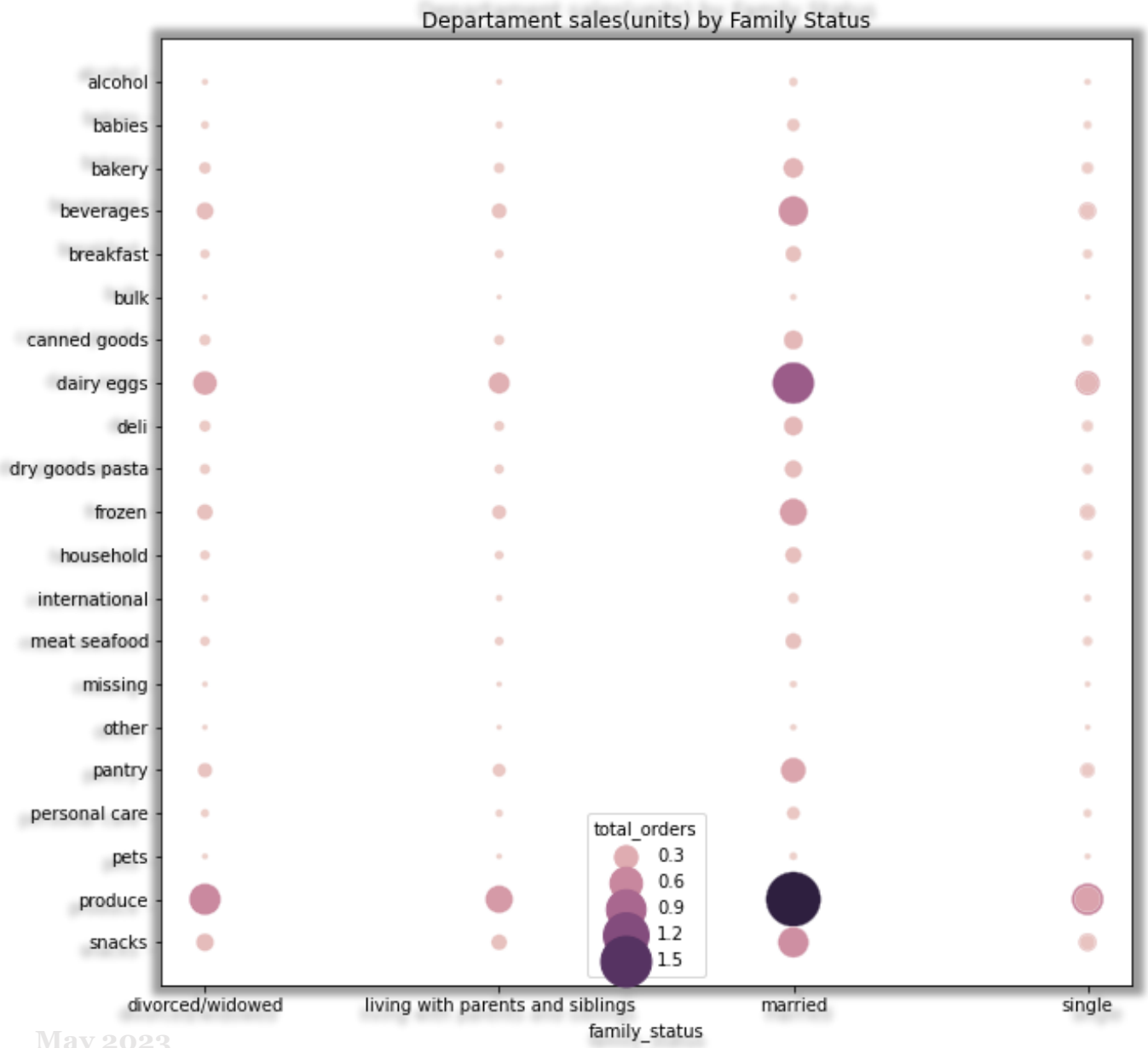
'Low-range product', prices equal or lower than 5$
'Mid-range product' prices above 5$, and lower or equal to 15$
'High-range product' prices above 15$
'Loyal customer', more than 40 orders.
'Regular customers' between 11 – 40 orders
'New customers' 10 orders or fewer

Age-range orders by family status. The bars show that the highest number of orders is being placed by married customers.

However, we observe that there are subgroups that are not reflected in the sales data (values = 0), such as divorced customers between 18 and 49 years old.

This indicates potential errors in data collection or a need for improvement in gathering data to enhance the reliability of the analysis.

# Profiles base in Number of Orders




## Findings:

The customer profile that **order the most** :

**Married**
**34-49 years old**
**Interested on 'Produce' department**

The customer profile that **order the least** :

**Divorce/widowed**
**50-65 years old**
**Interested on 'bulk' department**

# Profiles base in Number of Orders

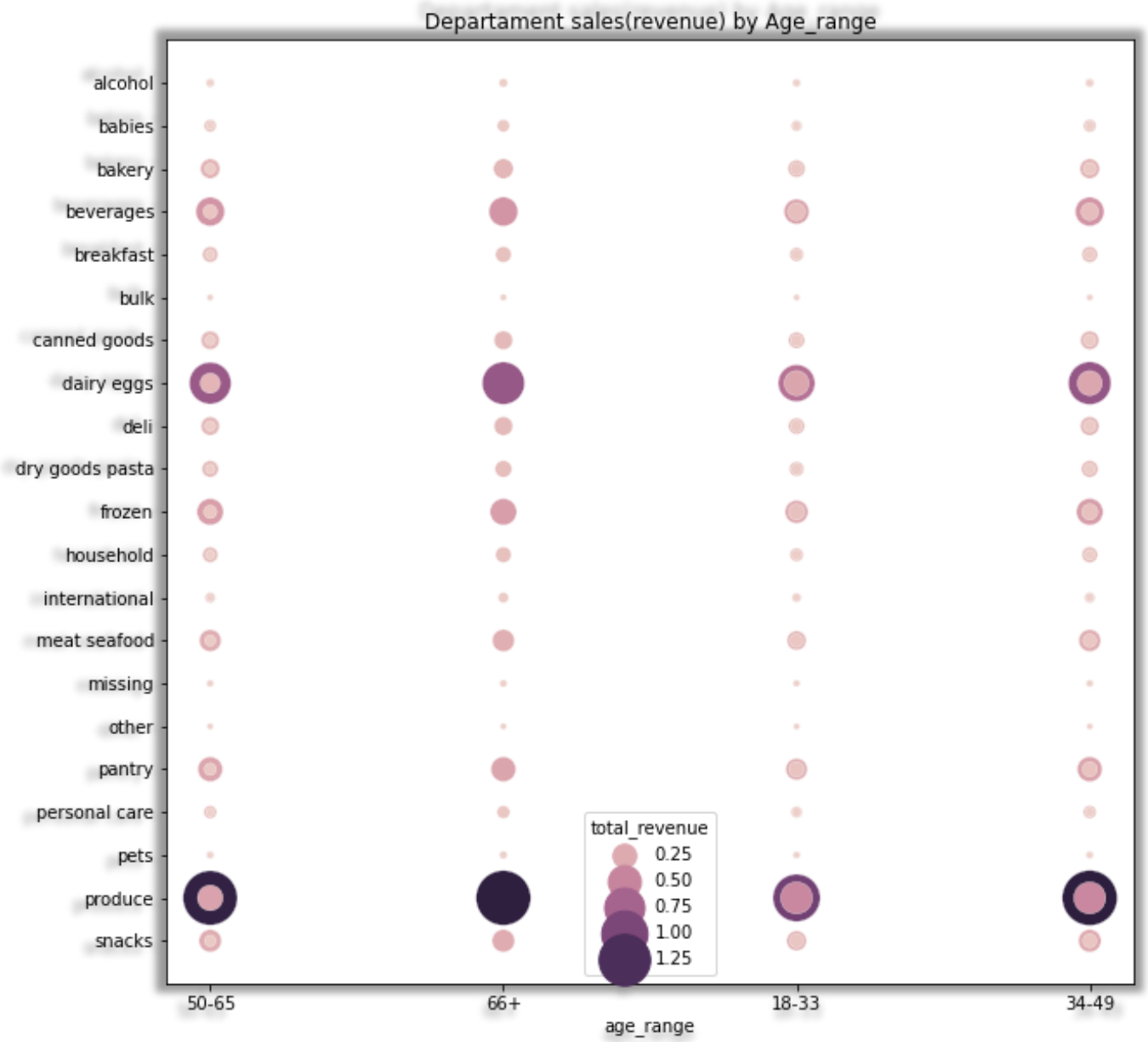
## Findings:

The most profitable customer profile is:

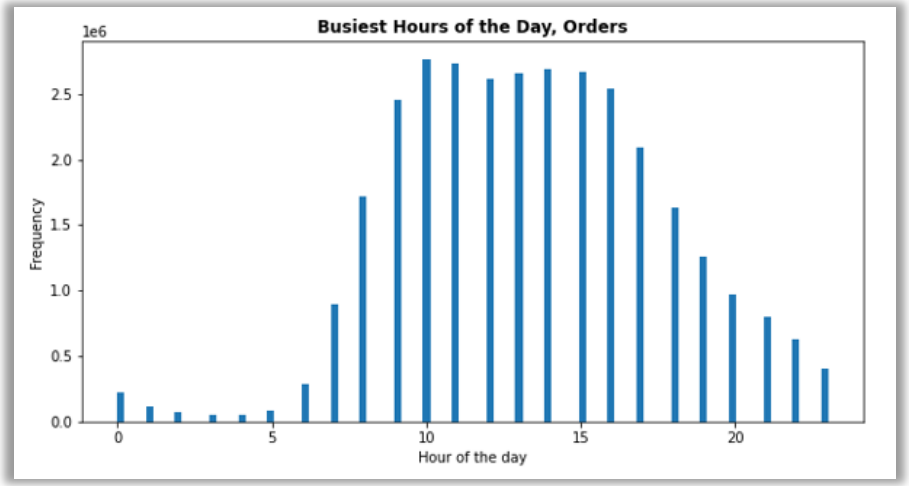
Married
34-49 years old
Interested on 'Produce' department

The least profitable customer profile is:

Divorce/widowed
50-65 years old
Interested on 'bulk' department



# Exploratory Analysis





# Findings:

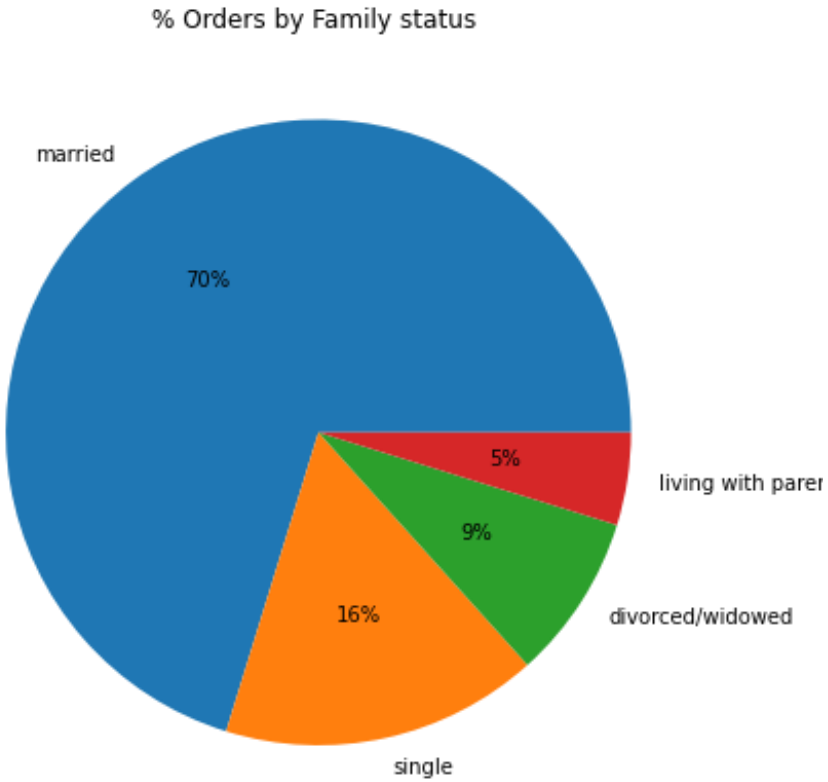
The busiest time of the day is between 8 am and 6 pm

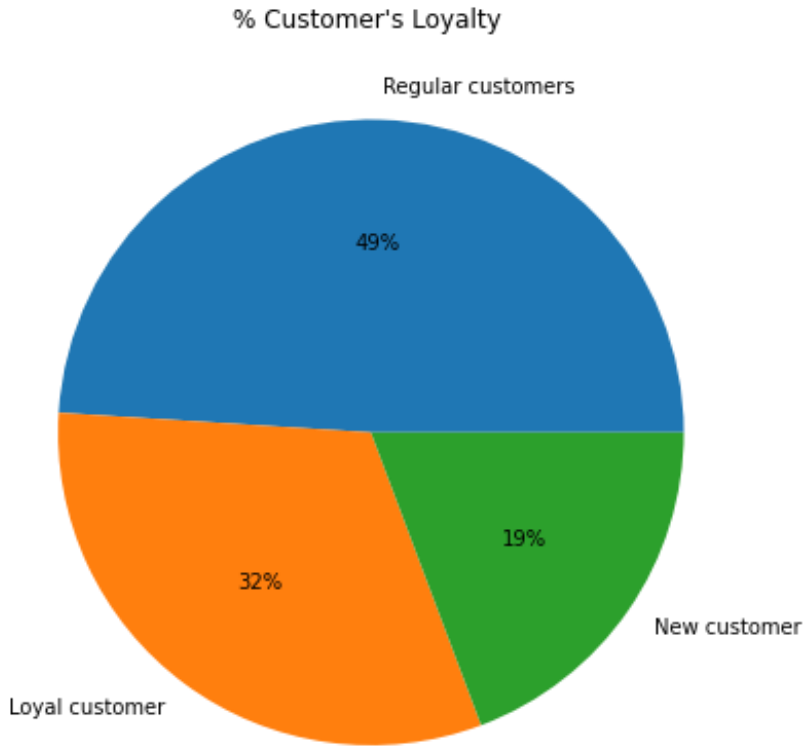
The quietest time of the day is between 23 pm and 6 am

The line chart shows high fluctuations prices of the goods purchased during the first half of the day, and steady average around 7.80 $ during the second half.

By looking at both chart we see two different purchase habits depending the time of the day.

# Profiles



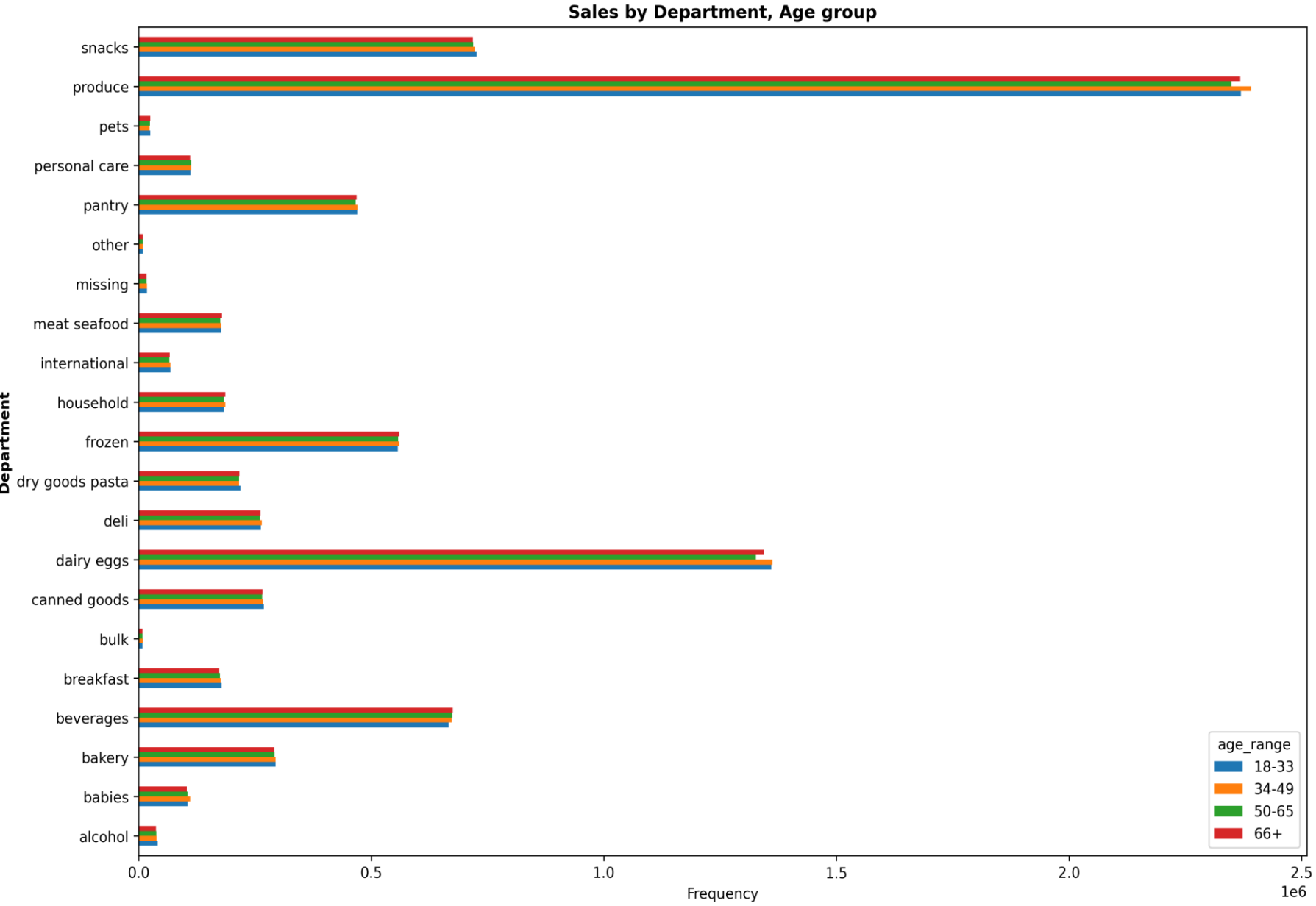

## Findings:

**70 %** of the purchases are done by '**married**' customers.

Only **5%** are categorized as **' living with parents and siblings'**

**New customers**
Less than **10** orders. 19% of the population.
**Regular customers**
From 10 to 40 orders . 49% of the population.
**Loyal customers**
More than 40 orders. 32% of the population.

# Profiles




In percentile terms the Department **"produce“** (29%), **"dairy eggs“** (17%), and **"snacks“** (9%) represent **58**% of **the total orders**.

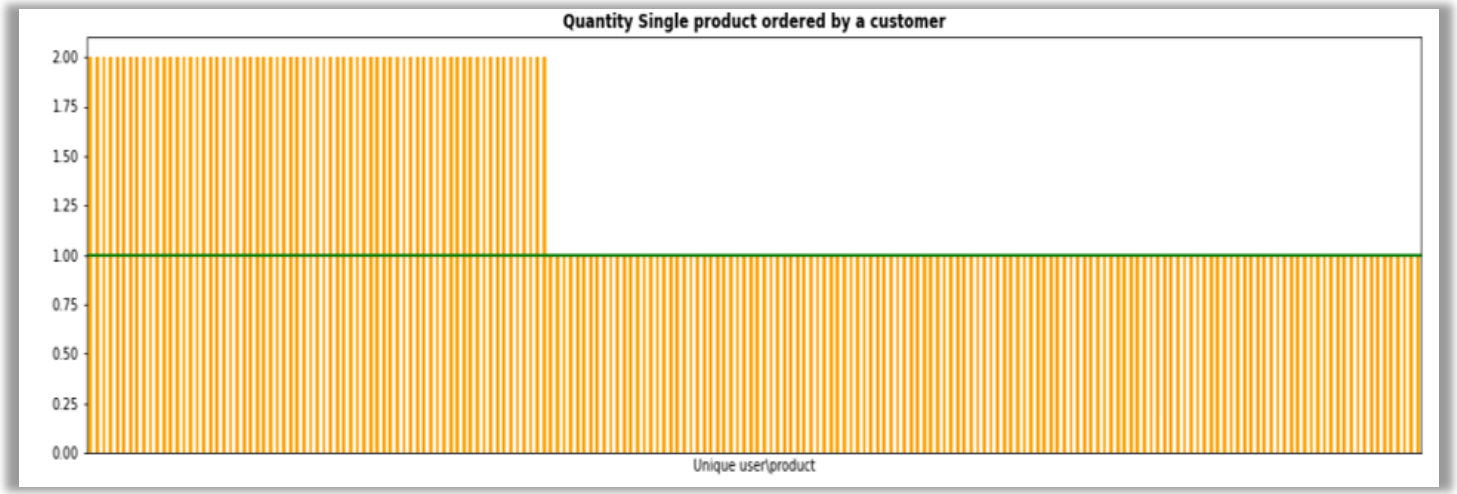
Without mayor differences between age groups.

# Conclusions

| **The busiest Days of the week Saturday & Sunday** | **The quietest Days of the week Tuesday $ Wednesday** |
|---|---|
| The **busiest** time of the day is **between 8 am & 6 pm** | The **quietest time** of the day is **between 23 pm and 6 am** |
| The **highest** average price in sales is **between 4-6am** | |
| Department **"produce“** (29%), **"dairy eggs“** (17%), and **"snacks“** (9%) represent **58**% of **the total orders**. | **“Alcohol”, “Pets”, “Bulk”, and “other”** department sells the **least** |
| **70 %** of the purchases are done by ‘**married’** customers. | Only **5%** are categorized as **' living with parents and siblings'** |
| The **customer profile that order the most** :<br><br>Married<br>34-49 years old<br>Interested on 'Produce' department | The customer profile that order **the least** :<br><br>Divorce/widowed<br>50-65 years old<br>Interested on 'bulk' department |
| The **most profitable customer** profile is:<br><br>Married<br>34-49 years old<br>Interested on 'Produce' department | The **least profitable** customer profile is:<br><br>Divorce/widowed<br>50-65 years old<br>Interested on 'bulk' department |

# Exploratory Analysis

There is a 25% of unique entries "Customer, Product, Product count" where a single customer bought more than 260 units of the same Product.


There are 131 instances "Customer, Product, Product count" where a single customer bought only once a specific product.

## Other findings

| | |
|---|---|
| There are **131 instances** "Customer, Product, Product count" where a single customer bought only once a specific product. | Further studies is recommended to determine: if the prices are correct according to the current market, or questioning the quality of the given items. |
| There is a **25% of unique entries "Customer, Product, Product count"** where a single customer bought **more than 260 units** of the same Product. | Exist the possibility to diversify and expand the business by exploring a B2B business model. |

# Findings & Recommendations

| Base on busyness: | Base on Department: | Base on Customer profiles: | Base on unique product sold to a single customer : |
|---|---|---|---|
| Lowest days of the week:<br>• Tuesday and Wednesday<br><br>Lowest hours based on Sales:<br>• From 23 pm to 6<br><br>To increase sales:<br>Locate ads between 22 pm till 5 am.<br><br>To promote new products, in-app ads between 8 am and 18 pm, from Thursday to Sunday. | In-app ads when "checking-out" promoting products from departments with lower sales, and tailored to the customer profile, could encourage the customer to explore different departments and products that it may be interested. | "Loyalty points card" to retain valuable customers depending on their profitability profile. | Some products have been bought once by a single customer.<br><br>Verify the price and perceived value.<br><br>Explore for existing quality issue with these product.<br><br>Some product was bought more than 260 times by a single customers.<br><br>Explore for B2B unattended market and an opportunity for business growth. |