

# Visualization Project Component 1

Bhumika Dilip Chawla | MDS202215

## Abstract:

The purpose of this project is to analyse various characteristics of small businesses spread across India. This component of the project contains some key insights from the data used. Further insights will be available in the second component of this project.

## Introduction:

The topic for this project is: Analysis of small business owners

The data used for this project had been collected with the help of a Google form link without collecting any information that would lead to respondent's identification.

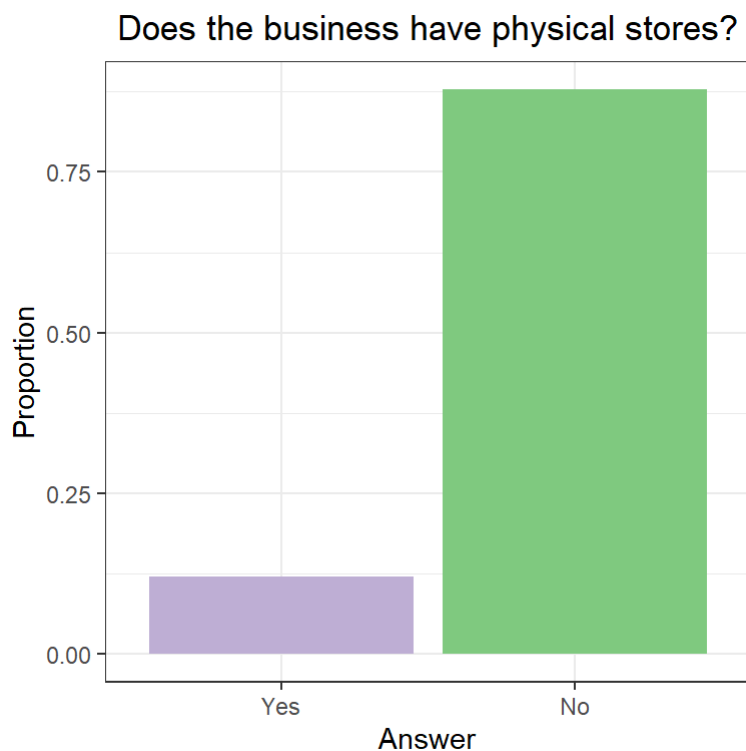
## Data definition:

The variables present in our data are as follows:

- Age of the business owner
- Gender
- City: The city to which the business owner belongs.
- Product Category : Decor, Clothing, Home Kitchen etc.
- Whether the business was established during Covid-19 or prior to it.
- Number of months since the business has been running.
- A ranking of Instagram, Facebook, Youtube and WhatsApp as a mode of preference for business by the owner.
- Target audience
- Shipping range : Whether within state, pan india, international etc.
- Average Net Profit/month? :  $(\text{Revenue} - \text{Cost}) / \text{Revenue} * 100 \%$
- Monthly orders
- Content interaction
- Marketing strategy
- Whether the business owner faces the following constraints:
  - Budget constraints
  - Lack of clients
  - Lack of inventory
  - Shortage of labour
  - Family barrier
  - Time Constraint

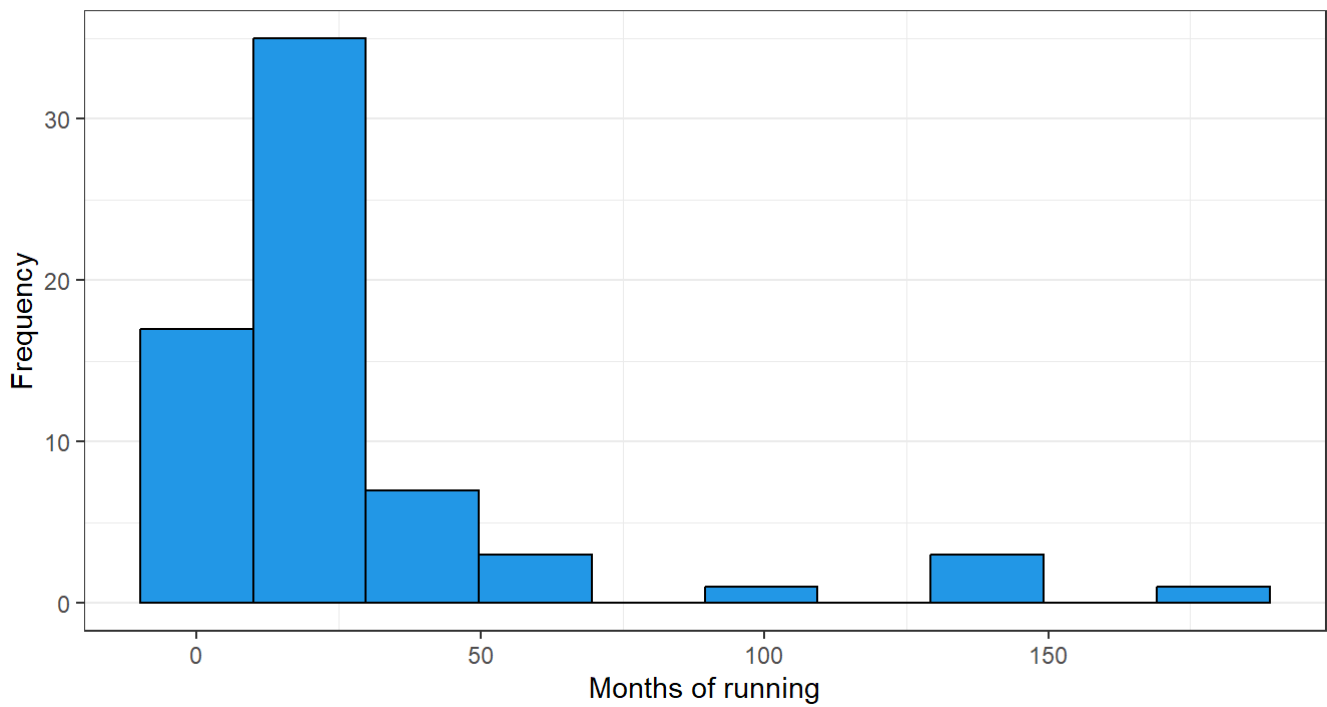
- Place Constraint
- Lack Of Management
- Capital Investment
- Mode of delivery
- Mode of payment
- Whether registered on other established websites.
- Whether the business has a physical store.
- Whether there is a difference between the profitability from your physical and online store.

## Visualizations:



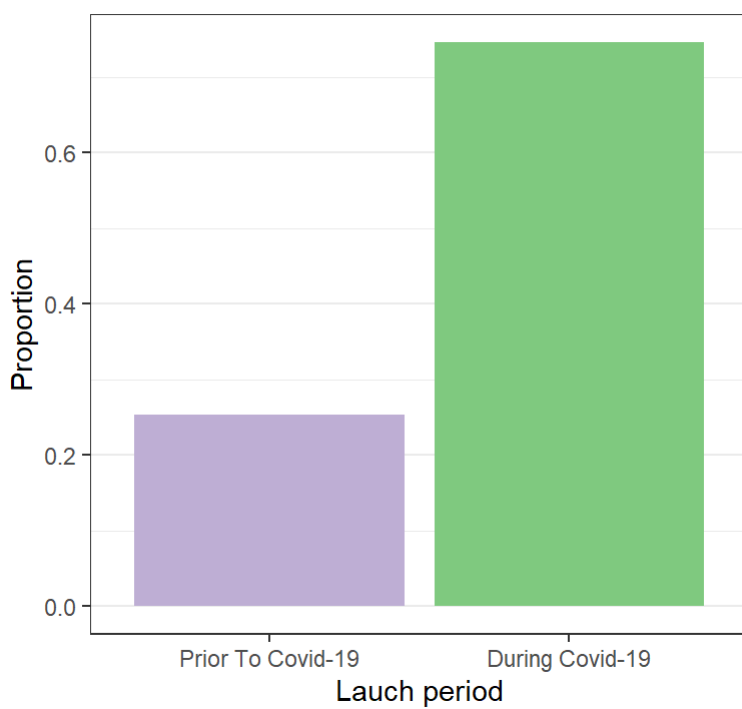
About **88%** owners in our sample don't have a physical store and run their business through the internet, basically they are **E-commerce** businesses.

Distribution of months of running for businesses in our sample.

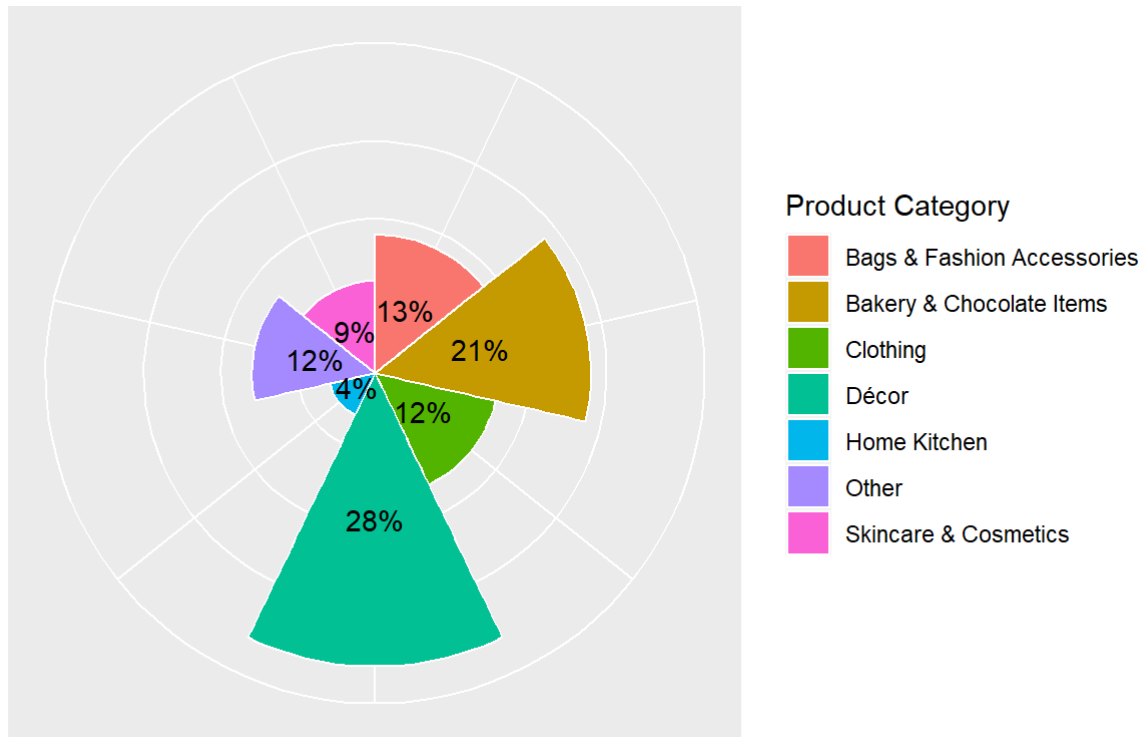


Most of the businesses in our sample have **less** than **50 months** of running/ **less** than **5 years** of running.

When was the business launched?

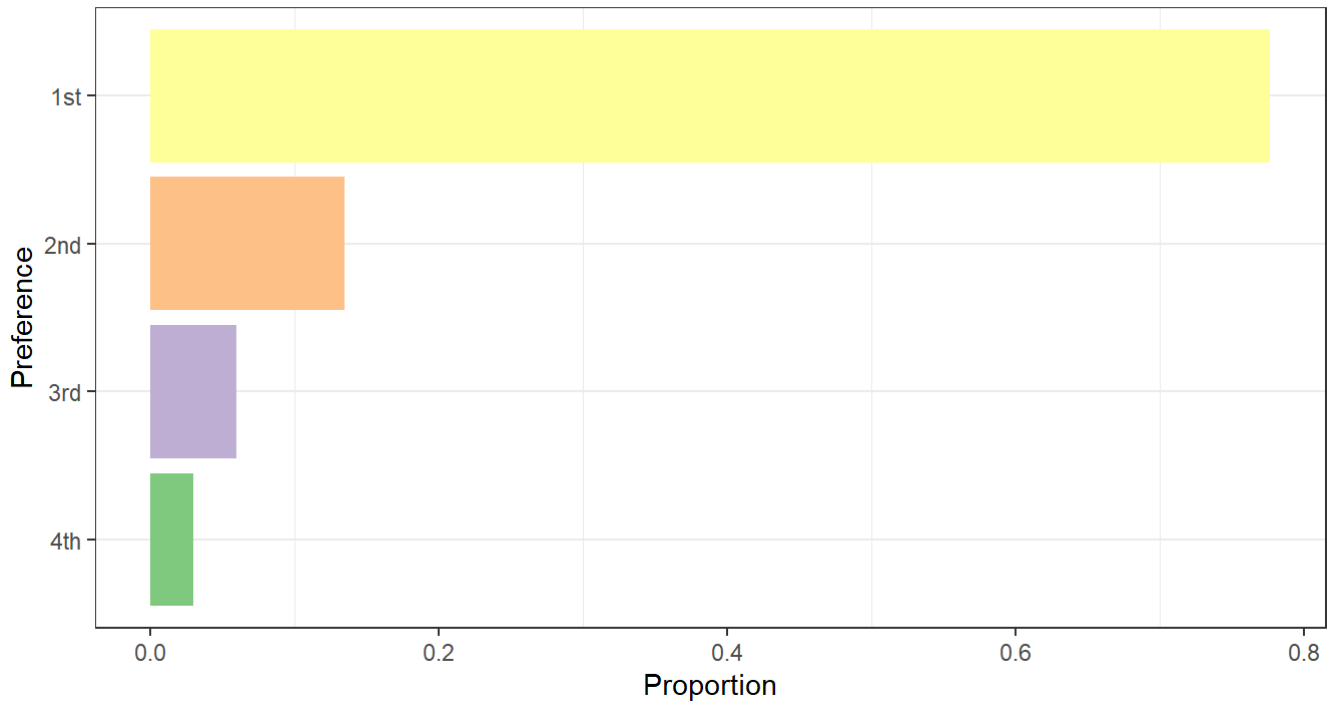


Most businesses in our data have started their operation **during** the pandemic of **COVID 19**.



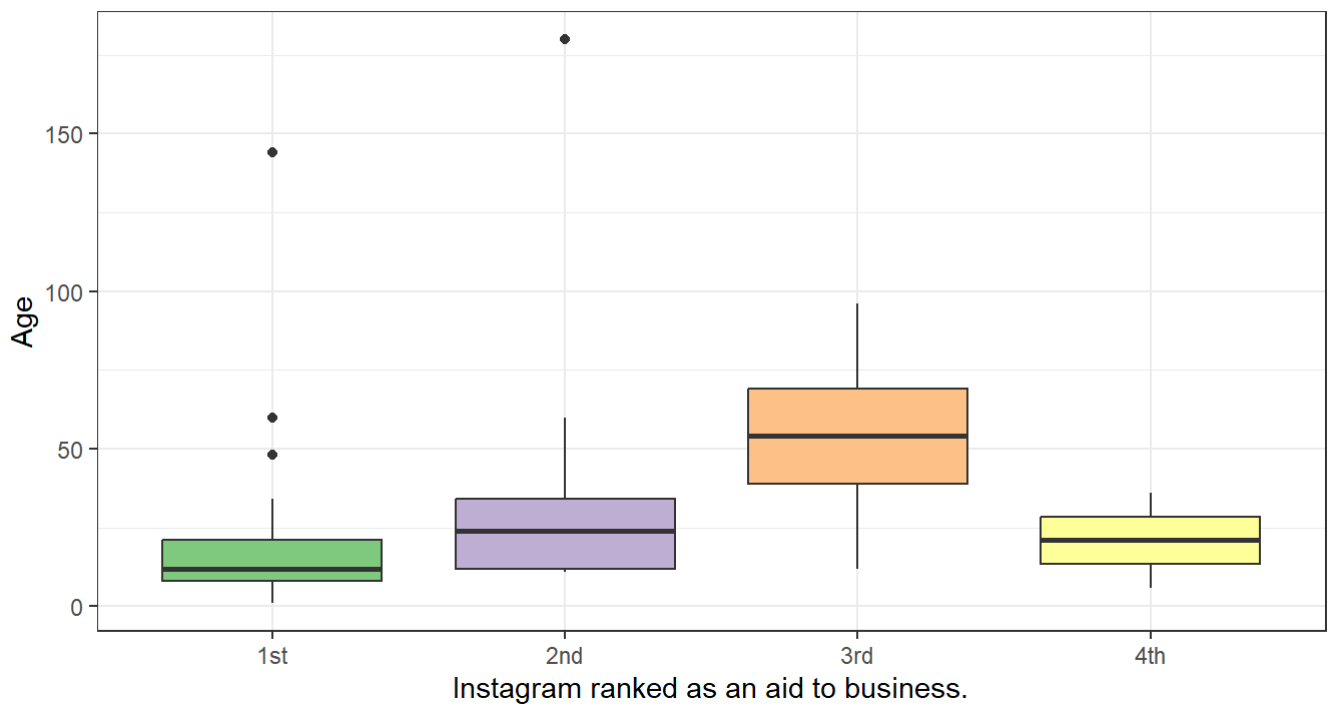
- Top 4 product categories in our sample:
  - Décor
  - Bakery and chocolate items
  - Bags and accessories
  - Clothing

Instagram as a preference for business by the owners.

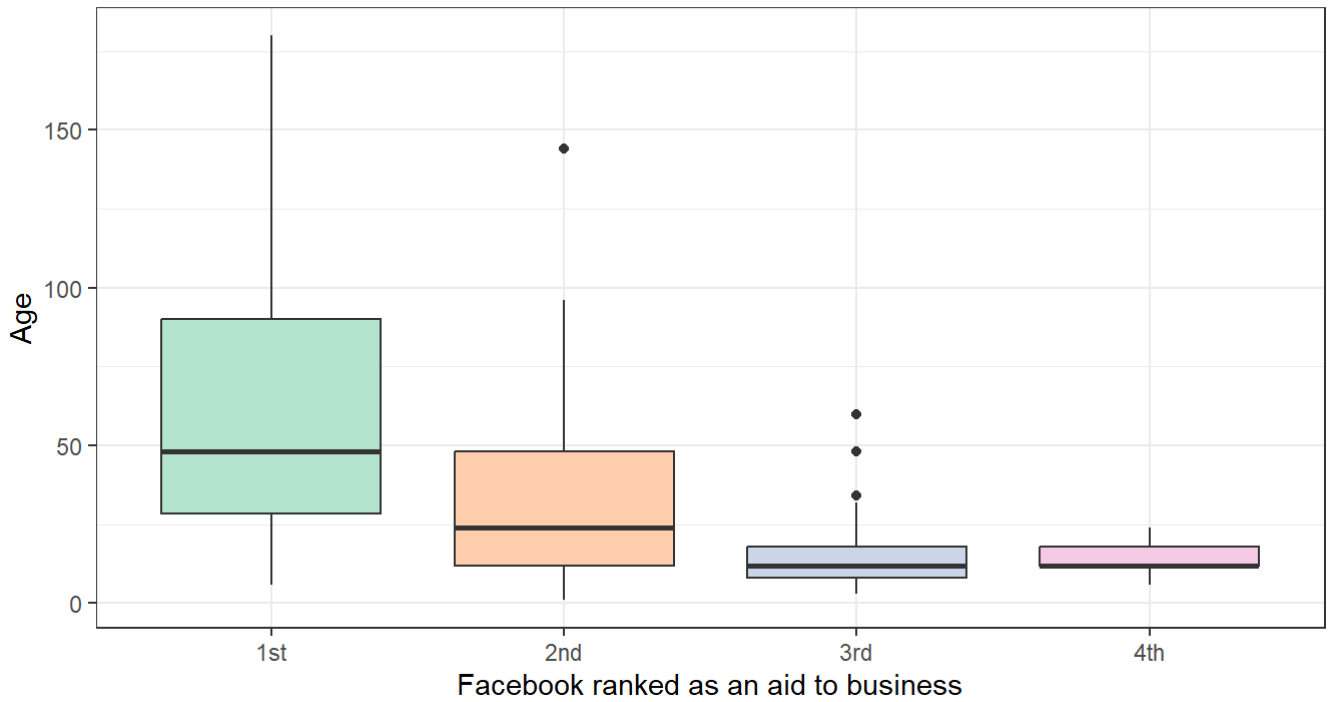


**Most** owners in our sample consider **Instagram** to be their **1st** choice of preference for their business.

Age vs Instagram as a preference

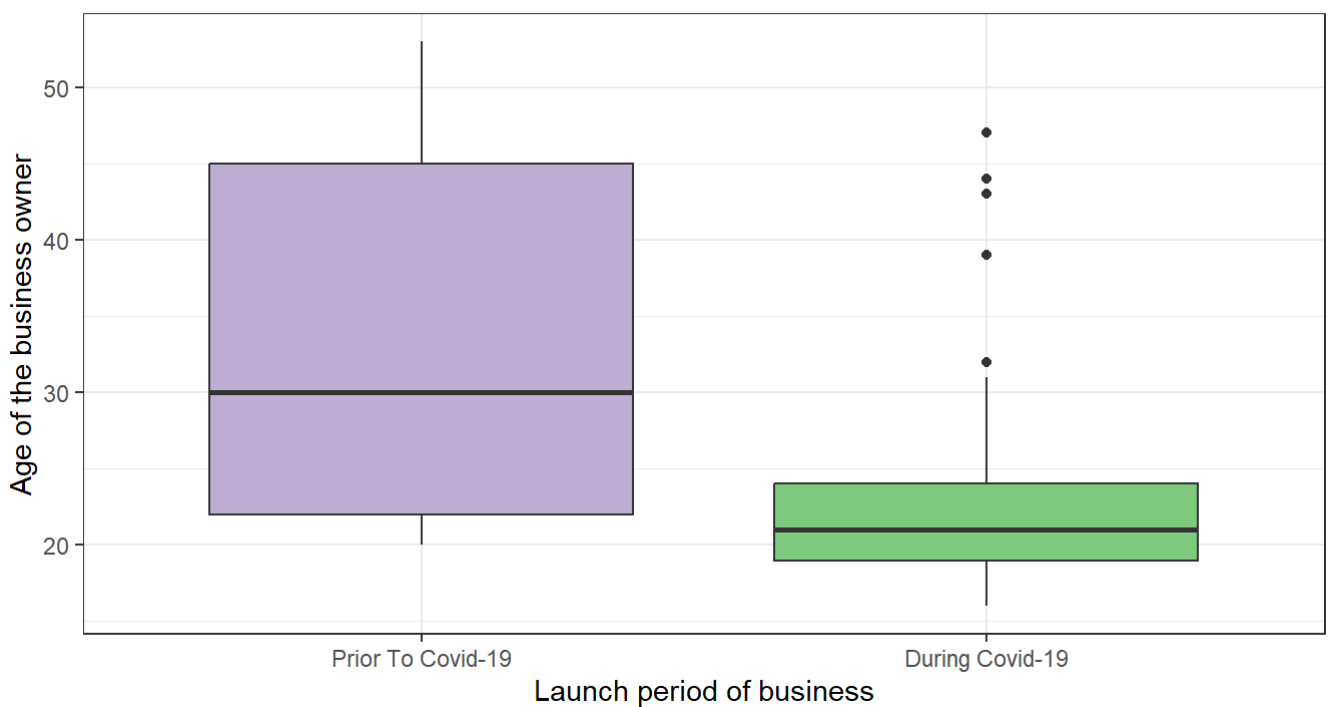


Age vs Facebook as a preference



We can observe from the 2 graphs above that the businesses which have **Facebook** as their **1st** platform of preference are **older** businesses as compared to the businesses which have **Instagram** as their **1st** platform of preference. Because Instagram has **recently** started as compared to Facebook and hence the businesses which were present **long back** use **Facebook** for their business and businesses which have started **recently** use **Instagram**.

Age of business owner vs Launch period of business

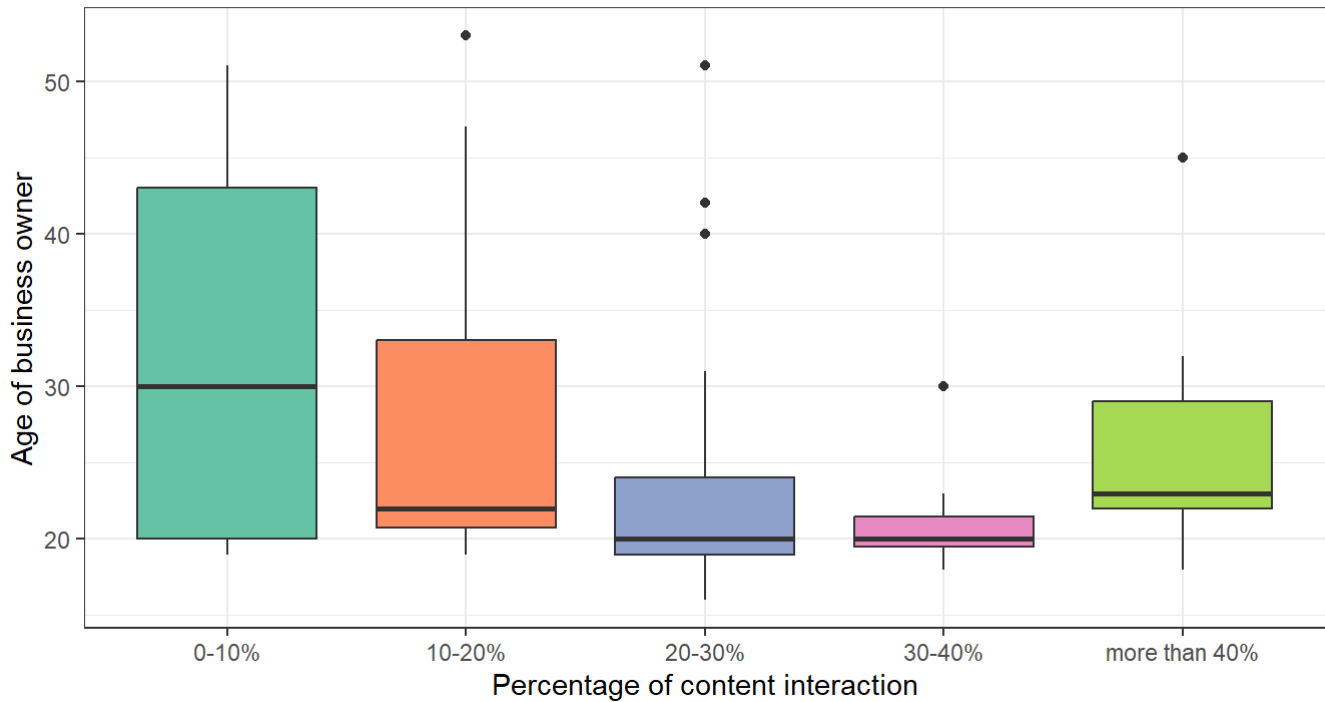


Businesses which were established prior to Covid-19 have business owners with **median** age of **30 years**. While businesses which were established during Covid-19 have business owners with a median age of **20 years**, with a **50%** concentration in the age group of **18-24 years**.



As expected, owners which are **aged**, own businesses which have been running for a **longer duration** as compared to businesses owned by **younger** business owners.

## Age of owner vs the content interaction on various social media websites



The **median** age of the business owners who have **0-10%** content interaction is **higher** as compared to others who have **more than 10%** content interaction.

## Caution:

From the above analysis, we understood that most businesses in our sample are those which have been launched during the period of Covid-19, hence there is bias in our sample. Thus any insights drawn from our sample should not be generalised to the entire population of small business owners in India.

## Conclusion:

We explored distributions of various variables and their interrelationships with other variables using univariate and bivariate plots here. Further exploration of data would be performed in the second component i.e. the R Shiny Dashboard.

The link for dashboard.

([https://bhumikac.shinyapps.io/Analysis\\_Of\\_Small\\_Business\\_Owners/](https://bhumikac.shinyapps.io/Analysis_Of_Small_Business_Owners/))

The link for Youtube video explaining the dashboard.

([https://www.youtube.com/watch?v=phXTv\\_9e4CE](https://www.youtube.com/watch?v=phXTv_9e4CE))