Fundamentals of Data Analytics Project

Kintali Preeti

21BCE9185

I am going to use the **R programming language** in this project.

Project topic- Ordering food online

```
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```
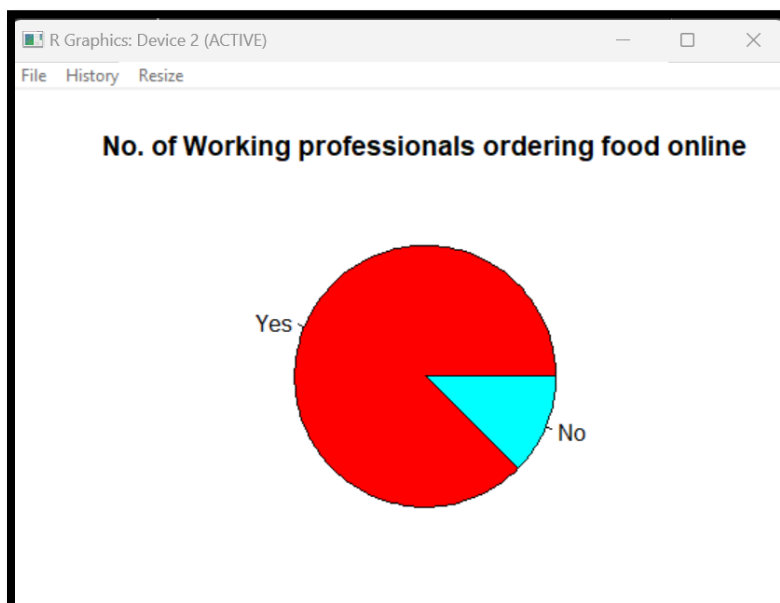
1. How many working professionals prefer to order food online?

Code:

```
> mydata1=filter(mydata, X4..What.do.you.do.=='Working Professional')
> mydata2=filter(mydata1,X3..Do.you.order.food.online.=="Yes")
> yes=nrow(mydata2)
> mydata3=filter(mydata1,X3..Do.you.order.food.online.=="No")
> no=nrow(mydata3)
> x<-c(yes,no)
> x
[1] 21  3
> labels <- c('Yes','No')
> pie(x,labels,main = "No. of Working professionals ordering food online",col = rainbow(length(x)))
> paste('The no. of professionals who chose to order food online: ',yes)
[1] "The no. of professionals who chose to order food online:  21"
```

Output:

The **majority** of the working professionals **prefer to order food online**.

2. How many people living in city outskirts say food reaches on time?

Code:

```
> mydata3=filter(mydata,X11..Does.the.food.ordered.online.reaches.on.time..=='Yes')
> d1=select(mydata3,X3..Where.do.you.live...,X11..Does.the.food.ordered.online.reaches.o
n.time..)
> d2=mydata3 %>% group_by(X3..Where.do.you.live...) %>% summarise(n())
> d2
> labels<-d2$X3..Where.do.you.live...
> values<-d2$`n()`
> barplot(names.arg = labels, values, main = "No of people who says food ordered online r
eaches on time", xlab = "Location", ylab = "No. of people", col = "blue", border = "Green
")
> paste('The number of people living in city outskirts who says food reaches on time are:
',nrow(outskirts))
[1] "The number of people living in city outskirts who says food reaches on time are:  33
"
```

Output:
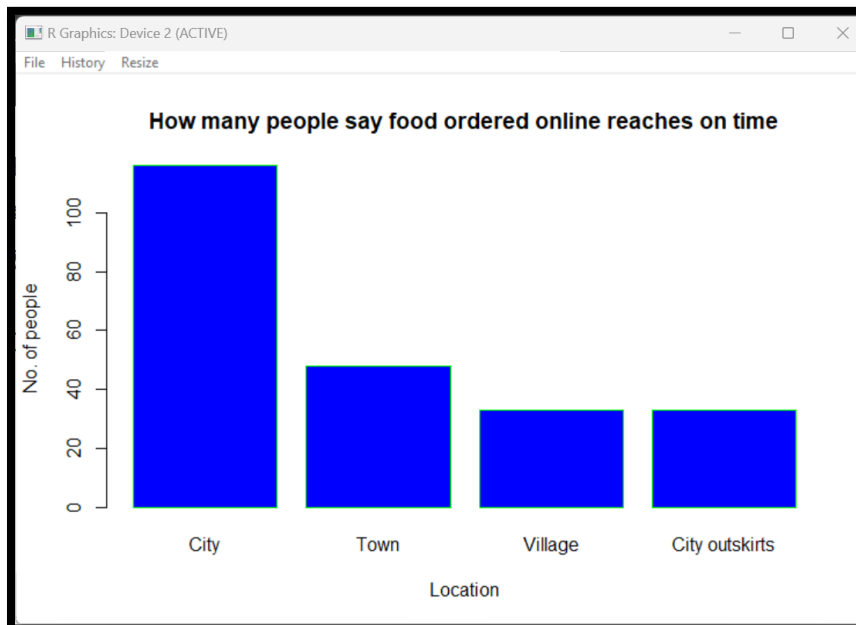
Majority of people living in city says that ordered food reaches on time.

3. How many people of age 18-25 are ordering food from swiggy?

Code:

```
> mydata4 = filter(mydata,X1..What.is.your.age.=='18-25',X6..Which.app.do.you.use.the.mos
t.to.order.food.online.=='Swiggy')
> ques3=select(mydata4,X1..What.is.your.age.,X6..Which.app.do.you.use.the.most.to.order.f
ood.online.)
> nrow(ques3)
[1] 96
> paste('No. of people aged 18-25 using swiggy are: ',nrow(ques3))
[1] "No. of people aged 18-25 using swiggy are:  96"
> mydata4 = filter(mydata,X6..Which.app.do.you.use.the.most.to.order.food.online.=='Swigg
y')
> d3=mydata4 %>% group_by(X1..What.is.your.age.) %>% summarise(n())
> age<-c(d3$X1..What.is.your.age.)
> noofpeople<-c(d3$`n()`)
> pie(noofpeople,age,main='People who use Swiggy',col=rainbow(length(noofpeople)))
> piepercent<- round(100*noofpeople/sum(noofpeople), 1)
```

```
> pie(noofpeople, labels = piepercent, main = "People who use Swiggy",col = rainbow(lengt
h(noofpeople)))
> legend("topright",age, cex = 0.8,fill = rainbow(length(noofpeople)))
```

Output:
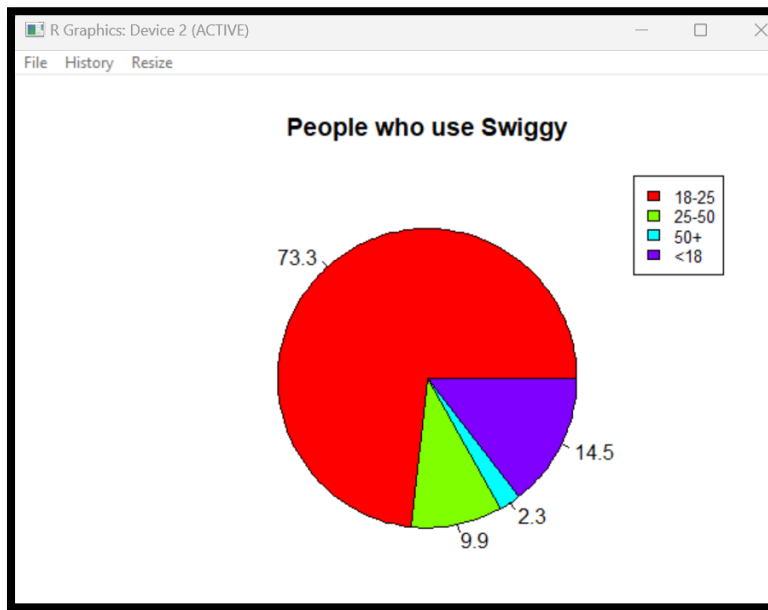
73.3% are people who age **18-25** who use swiggy to order food online. They are **96 people** in number which was retrieved from using R language.

4.How much percentage of people living in city are ordering food?

Code:

```
> mydata5= filter(mydata,X3..Where.do.you.live...=='City' & X3..Do.you.order.food.online.
=='Yes')
> ques4=select(mydata5,X3..Where.do.you.live...,X3..Do.you.order.food.online.)
> n<- nrow(ques4)
> paste('The no. of city living people who order food online are: ',n)
[1] "The no. of city living people who order food online are:  133"
> mydata5= filter(mydata,X3..Do.you.order.food.online.=='Yes')
> d4 =mydata5 %>% group_by(X3..Where.do.you.live...) %>% summarise(n())
> d4
# A tibble: 4 x 2
  X3..Where.do.you.live...  `n()`
  <chr>                     <int>
1 City                        133
2 City outskirts               38
3 Town                         49
4 Village                      33
> location<-c(d4$X3..Where.do.you.live...)
> values<-c(d4$`n()`)
> pie(values,location,main='City living people ordering food online',col=rainbow(length(v
alues)))
> piepercent<- round(100*values/sum(values), 1)
> pie(values, labels = piepercent, main = "Percentage of city living people ordering food
 online",col = rainbow(length(values)))
> legend("topright", c('City','City outskirts','Town','Village'), cex = 0.8,fill = rainbo
w(length(values))
+ )
```

Output:
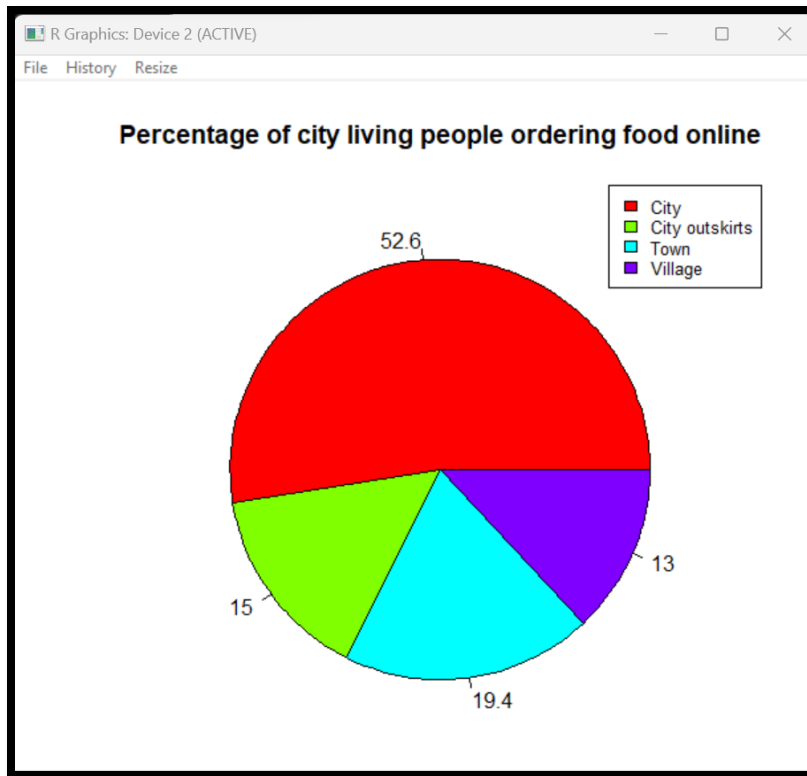
Hence, majority of people who ordering food online live in **city**.

5. How many working professionals think obesity is impact of ordering food online?

Code:

```
> mydata5<- filter(mydata,grepl('Obesity',X13..What.impacts.do.online.food.ordering.ha
s.))
> mydata6<- filter(mydata,grepl('obesity',X13..What.impacts.do.online.food.ordering.ha
s.))
> d5<-rbind(mydata5,mydata6)
> d5<-filter(d5,X4..What.do.you.do.=='Working Professional')
> paste('The number of working professsionals who think obesity is main impact of orderin
g food online: ',nrow(d5))
[1] "The number of working professsionals who think obesity is main impact of ordering fo
od online:  10"
> d<-d5$X4..what.do.you.do.
> table(d)
> d<-as.data.frame(table(d))
> barplot(names.arg=x,d$Freq,main='Distribution of people who say obesity is main impact
',col='cyan',border='green')
```

Output:

Distribution of people who say obesity is main impact
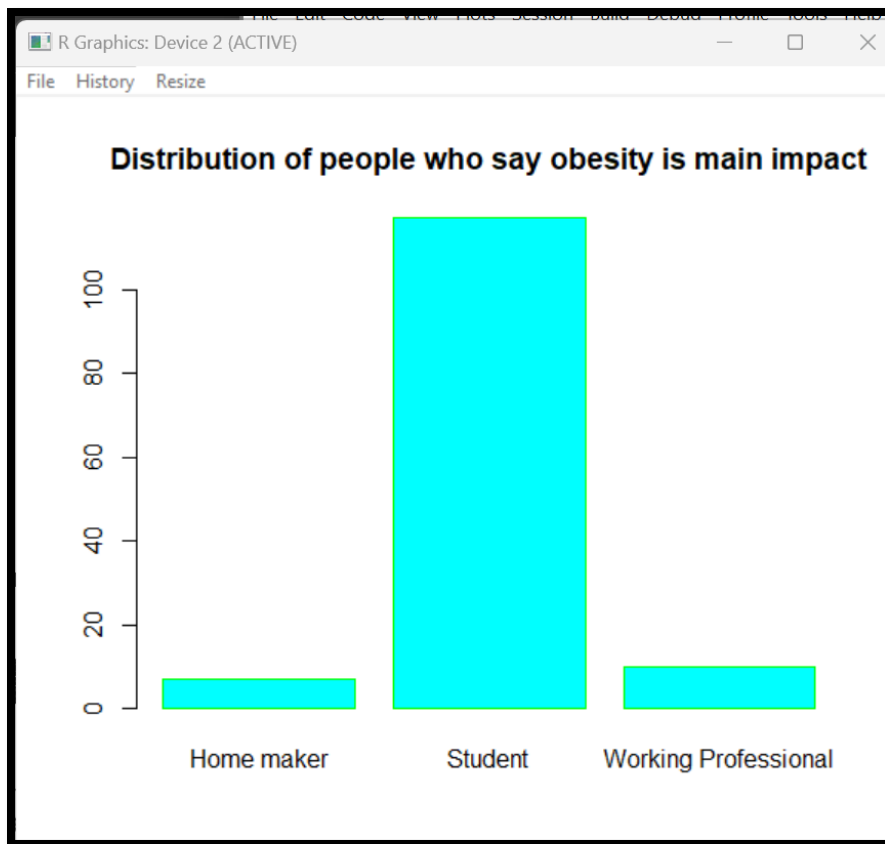
6.What is the difficulty level of ordering food online for the people living in village and town?

Code:

```
> d6<-filter(mydata,grepl('Village',X3..Where.do.you.live...))
> d7<-filter(mydata,grepl('Town',X3..Where.do.you.live...))
> d<-rbind(d6,d7)
> d$X17..How.difficult.do.you.think.it.is.to.order.online.food..
  [1]  2 3 3 2 2 2 2 1 1 2 3 1 3 1 3 1 2 2 2 3 1 3 2 5 1 2 2 3 3 2 2 2 1 1 3 2
 [37]  3 3 3 1 1 2 1 4 1 2 3 2 1 1 1 2 1 2 4 1 3 3 3 2 1 2 3 3 3 2 3 3 3 1 3 1
 [73]  1 1 1 1 1 4 3 3 4 3 3 1 2 3 4 2 2 2 1 3 2 1 1 2 3 1 1 2 1 1 3 2 2 1 4 3
[109]  3 3 3 5 2 3 3 3 2 3 1 3 3 1 3 2 3 3 1 2 5 1 4 2 1
> level<-d$X17..How.difficult.do.you.think.it.is.to.order.online.food..
> table(level)
level
 1  2  3  4  5
40 38 45  7  3
> as.data.frame(table(level))
  level Freq
1     1   40
2     2   38
3     3   45
4     4    7
5     5    3
> x<-as.data.frame(table(level))
> pie(x$Freq,x$level,main='Difficulty level to order for village/town people',col=rainbow
(length(x$Freq)))
```

Output:
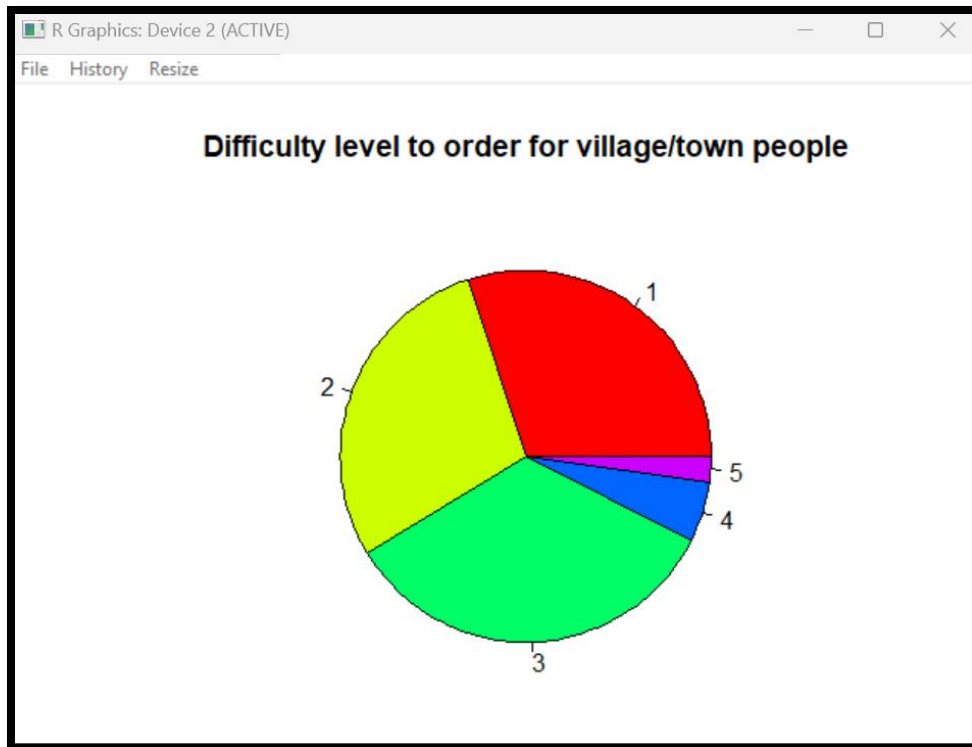
Most people find it not that hard, because almost of people chose 1,2,3 which means easy. And very less people chose 4,5.

This says that most of people have now got habituated to order food online even in towns and villages.

7.Through which medium did most people come to know about online food ordering?

Code:

```
> d7<-mydata$X15..How.did.you.know.about.the.online.food.ordering.
> table(d7)
> d<-as.data.frame(table(d7))
> piepercent<- round(100*d$Freq/sum(d$Freq), 1)
> pie(d$Freq, labels = piepercent, main = "How did they get to know?",col = rainbow(length(d$Freq)))
> v<-c('Ad came in zoro.to ','Adverstisement on TV','Advertisement','Friends/ Family','I dont remember','Myself','Neibours', 'Newspaper','Social media','Television')
> legend("topright",v, cex = 0.7,fill = rainbow(length(d$Freq)))
```

Output:

How did they get to know?

Legend:
- Ad came in zoro.to
- Adverstisement on TV
- Advertisement
- Friends/ Family
- I dont remember
- Myself
- Neibours
- Newspaper
- Social media
- Television

Values shown: 35.8, 0.3, 2.3, 0.3, 0.3, 15.6, 0.3, 44.3
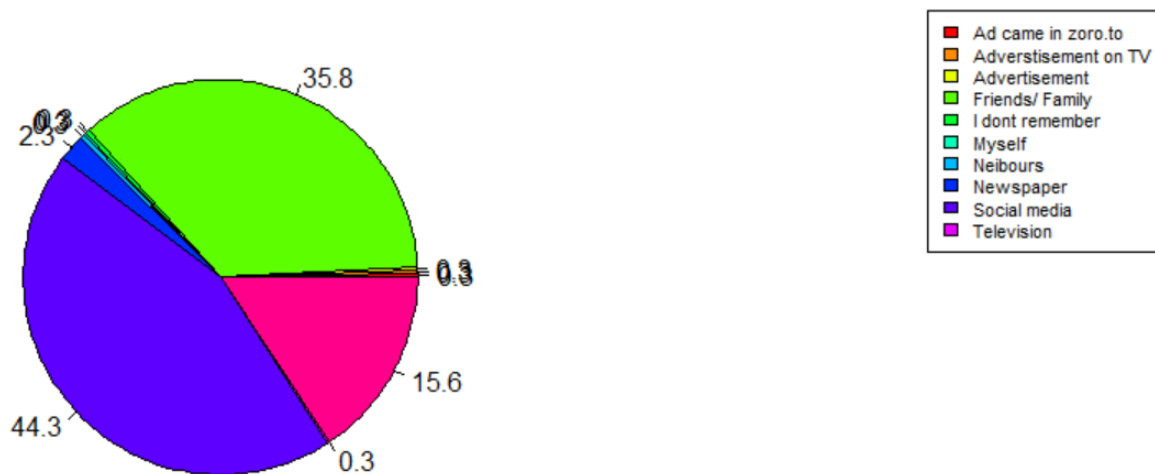
## 8.Where do the people who order food the most live?

Code:

```
> d8<- filter(mydata,X4..How.often.do.you.order.food.online.=='Daily')
> d8$X4..How.often.do.you.order.food.online.
 [1] "Daily" "Daily" "Daily" "Daily" "Daily" "Daily" "Daily" "Daily" "Daily" "Daily"
> d8$X3..Where.do.you.live...
 [1] "City"           "City"           "Town"     "Town"     "Village"
 [6] "City"           "City"           "Town"     "Town"     "City outskirts"
> table(d8$X3..Where.do.you.live...)

         City City outskirts           Town        Village
            4              1              4              1
> as.data.frame(table(d8$X3..Where.do.you.live...))
            Var1 Freq
1           City    4
2 City outskirts    1
3           Town    4
4        Village    1
> d=as.data.frame(table(d8$X3..Where.do.you.live...))
> pie(d$Freq,d$Var1,main='Where do people who order food most live?',col=rainbow(length(d
$Freq)))
```

Output:
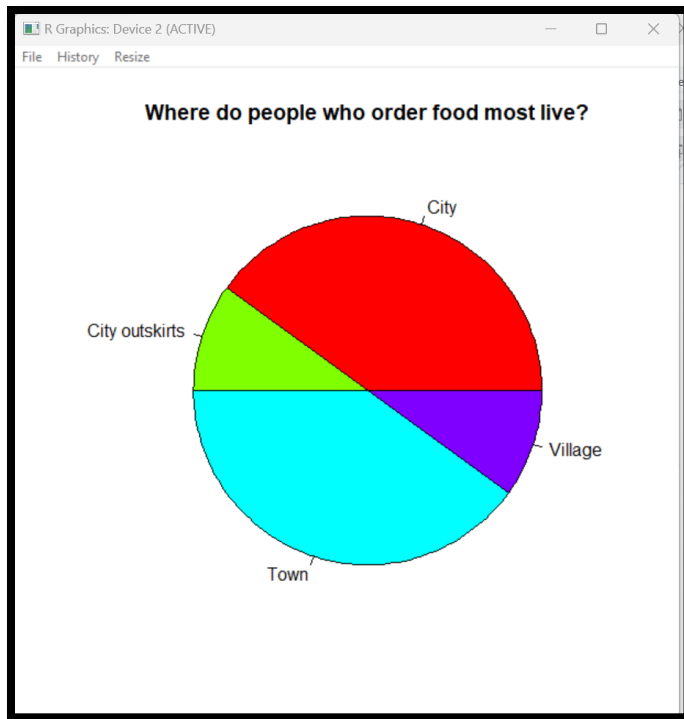
It is clear that people in both **cities and towns** order the food the most.

9.What is the majority (parent/self/chef/outside food) which wins for the question 'Food cooked by whom do you like the most?'.

Code:

```
> d9<-select(mydata,X1..Food.cooked.by.whom.do.you.like.the.most.)
> table(d9)
> x<-as.data.frame(table(d9))
> pie(x$Freq,x$X1..Food.cooked.by.whom.do.you.like.the.most.,main='Whose cooked food is l
iked the most?',col=rainbow(length(x$Freq)))
```

Output:
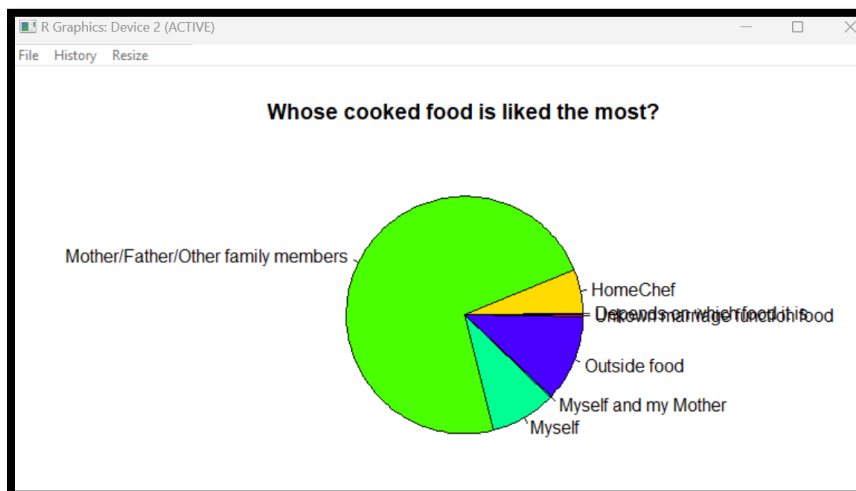
Most people like the food cooked by their mother/father/other family members as it covers the maximum portion of the pie chart.

10.How many people who live in hostel/pg's order food rarely/occassionally?

Code:

```
> d10<-filter(mydata,grepl("Hostel",X5..From.where.do.you.order.the.online.food.mostl
y..))
> d<-filter(d10,grepl('Rarely',X4..How.often.do.you.order.food.online.))
> paste('The no. of people who live in hostel/pgs who order food rarely or occasionally a
re: ',nrow(d))
[1] "The no. of people who live in hostel/pgs who order food rarely or occasionally are:
 68"
```

Output:

The no. of people who live in hostel/pg's who order food rarely/occasionally is **68**.

11. Which platform is the most used to order food online?

Code:

```
> d11<-mydata$X6..Which.app.do.you.use.the.most.to.order.food.online.
> table(d11)
d11
              Domino's                    Honest  I don't order food online
                     1                         1                         37
 I dont order food online      I never ordered food                 Playstore
                    12                         1                          1
             Restaurant                Restaurants                    Swiggy
                     1                         1                        131
                Ubereats                    Zomato
                    10                       155
> as.data.frame(table(d11))
> app<-as.data.frame(table(d11))
> pie(app$Freq,app$d11,main='Which app wins the race?',col=rainbow(length(app$Freq)))
```
**Output:**
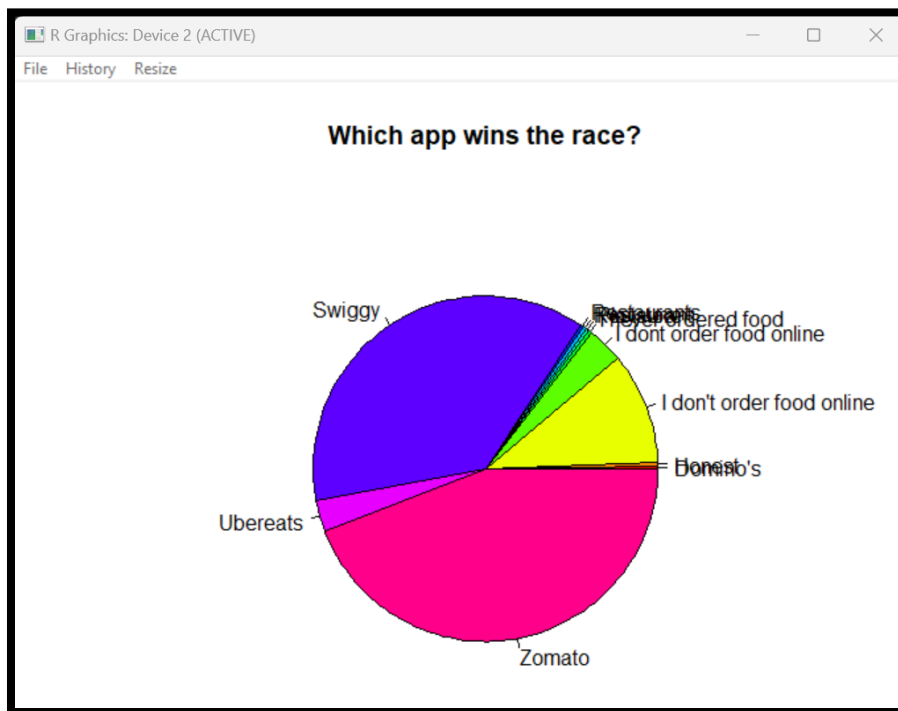
We can see that the majority portion is covered by **Zomato**. Hence, we can declare that it is the most used platform.

12. People from which location order food online the most to satisfy their taste buds?

Code:

```
> d12<-filter(mydata,grepl("To satisfy my taste buds",X7..Why.do.you.order.food.onlin
e..))
> d<-select(d12,X3..Where.do.you.live...)
> grouped<-group_by(d,X3..Where.do.you.live...)
> table(d)
> as.data.frame(table(d))
> data<-as.data.frame(table(d))
> pie(data$Freq,data$X3..Where.do.you.live...,main='Where do people who order food the mo
st live?',col=rainbow(length(data$Freq))
```

Output:
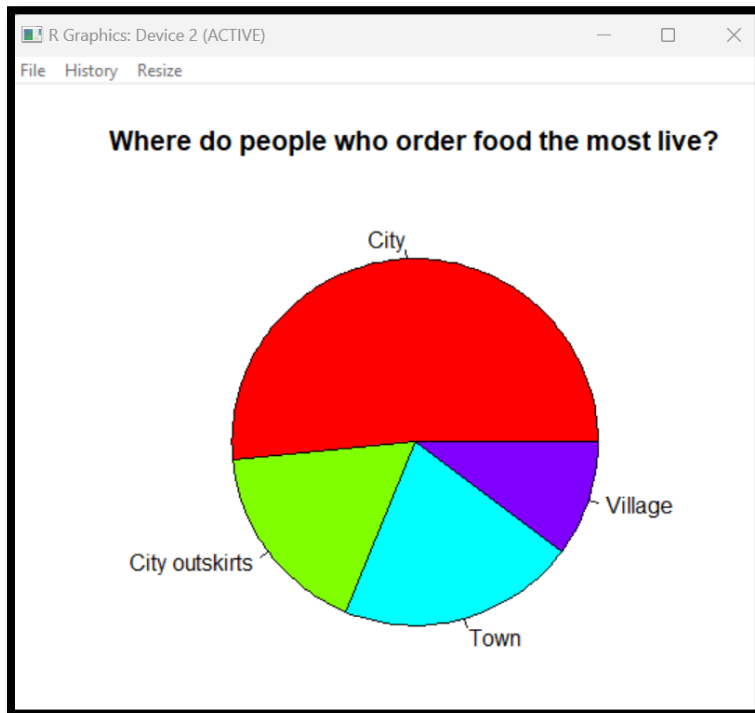
From the pie chart it is clear that most people who order food to satisfy their taste buds live in **cities**.

13.At what time do majority of working professionals order food?

Code:

```
> d13<-filter(mydata,X4..What.do.you.do.=='Working Professional')
> d13$X14..Which.meal.do.you.typically.order.from.online..
> d1<-filter(d13,grepl("Snacks",X14..Which.meal.do.you.typically.order.from.online..))
> snacks<- nrow(d1)
> d2<-filter(d13,grepl("Breakfast",X14..Which.meal.do.you.typically.order.from.online..))
> breakfast<-nrow(d2)
> d3<-filter(d13,grepl("Lunch",X14..Which.meal.do.you.typically.order.from.online..))
> lunch<-nrow(d3)
> d4<-filter(d13,grepl("Dinner",X14..Which.meal.do.you.typically.order.from.online..))
> dinner<-nrow(d4)
> values<-c(breakfast,lunch,snacks,dinner)
> labels<-c('Breakfast','Lunch','Snacks','Dinner')
> barplot(names.arg = labels, values, main = "Which food do working professionals prefer
most?",xlab = "Type of food", ylab = "Freq", col = "green", border = "blue")
```

Output:
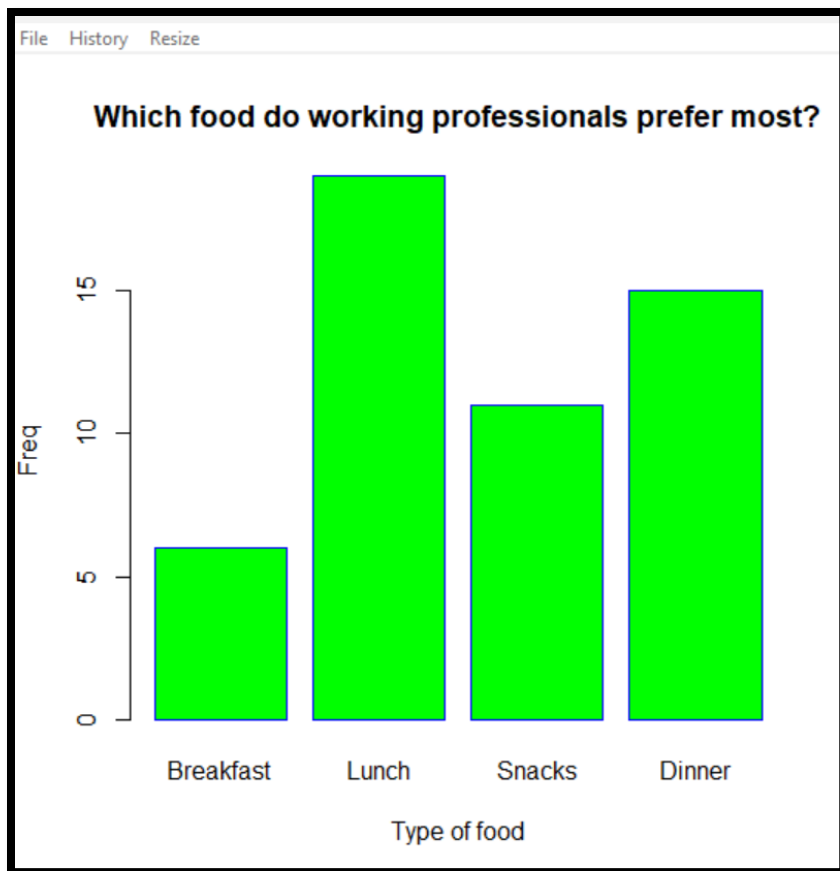
From the bar plot, it is clear that the majority of the working professionals prefer to order food for **lunch**.

14. Who are the majority audience who wanted faster delivery from the online food ordering?

Code:

```
> d14<-filter(mydata,grepl('Faster delivery',X9..Why.do.you.order.food.online.rather.tha
n.going.and.eating.at.a.restaurant.))
> d<-group_by(d14,X4..What.do.you.do.)
> table(d$X4..What.do.you.do.)

            Breath            Home maker              Intern             Student Worki
ng Professional
                 1                     3                   1                 163
            13
> as.data.frame(table(d$X4..What.do.you.do.))
                  Var1 Freq
1              Breath    1
2          Home maker    3
3              Intern    1
4             Student  163
5 Working Professional   13
> d2<-as.data.frame(table(d$X4..What.do.you.do.))
 > pie(d2$Freq,d2$Var1,main='What majority of profession wants food to be de
 livered fast?',col=rainbow(length(d2$Freq)))
```
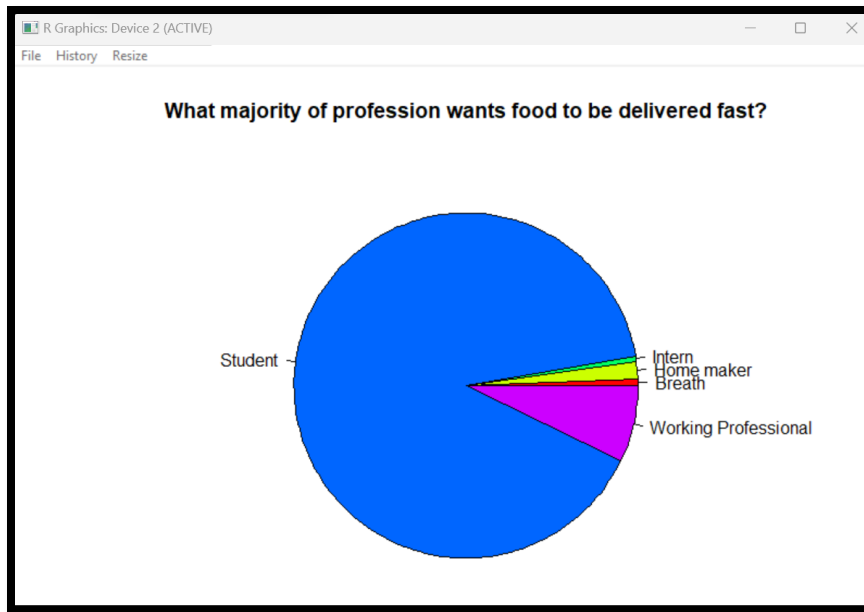
Output:

The majority of people who wanted fast delivery are **students** which can be concluded from this pie chart.

15.What is the percentage distribution for the satisfaction levels of the customers with online food ordering services?

Code:

```
> d15<-mydata$X10..Are.you.satisfied.with.the.quality.of.the.food.ordered.online.
> table(d15)
d15
  1   2   3   4   5
 17  29 128 142  33
> d<-as.data.frame(table(d15))
> labels<- d$d15
> values <-d$Freq
> piepercent <- round (100 * values/sum(values),1)
> pie(values,labels = piepercent,main='Ratings of people',col=rainbow(length(values)))
> legend("topright", c("1","2","3","4","5"), cex = 0.7,fill = rainbow(length(values)))
```

Output:

Ratings of people
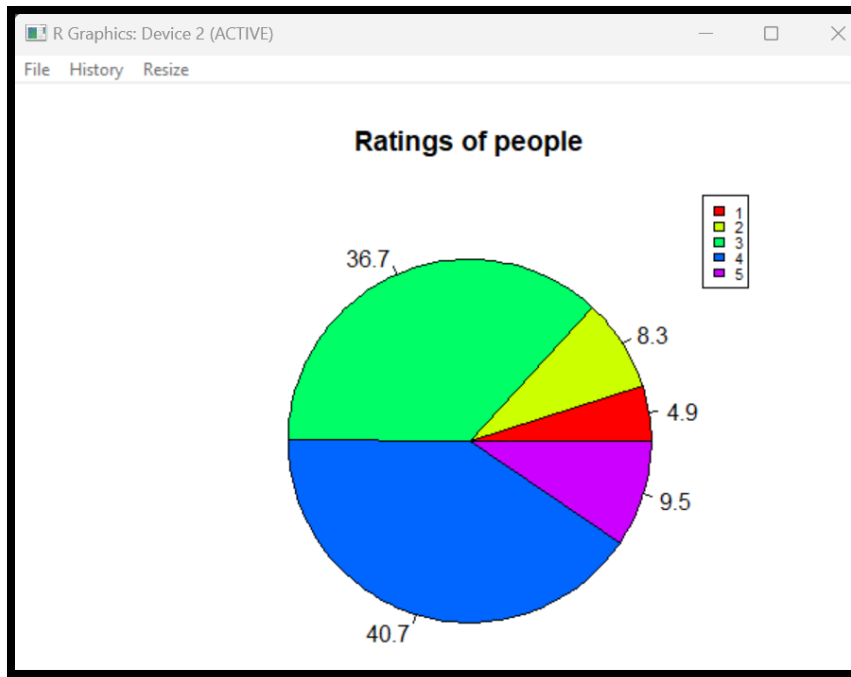
From the pie chart, we can observe that maximum people have rated the ordering food online as **4** have the highest percentage distribution, which is true as we got the same result from retrieval of data using command, which is 142 people who voted for 4.

Hence we can say most of the customers are **satisfied** with online food ordering service.