Italian Resteurant analysis

2025-01-07

My analysis starts here

- Dataset consists of 10000 observations and 9 variables namely : Date ,Time ,Menu item , Category , Quantity purchased ,Price per item , Revenue , Payment method , Customer type

```
library(readr)
library(tidyverse)
library(readxl)
library(sqldf)
library(stringr)
library(knitr)


data <-read_excel('Italian_resteurant_dataset.xlsx')

# Filtering data

data$Date<-as.Date(data$Date)

data <-data %>%mutate(Morning_afternoon=str_extract(Time,'AM|PM'),Month=months.Date(Date),
                      Week_day=weekdays(Date))%>%select(-c(Date,Time))


mat <-matrix(0,nrow = 3,ncol = 2)
mat[,1]<-c('Price','Revenue','Quantity purchased')
mat[,2]<-c(mean(data$`Price (per item)`),mean(data$Revenue),mean(data$Quantity))
colnames(mat)<- c('-','Average')
kable(mat)
```

-	Average
Price	8.32485
Revenue	24.77305
Quantity purchased	2.978

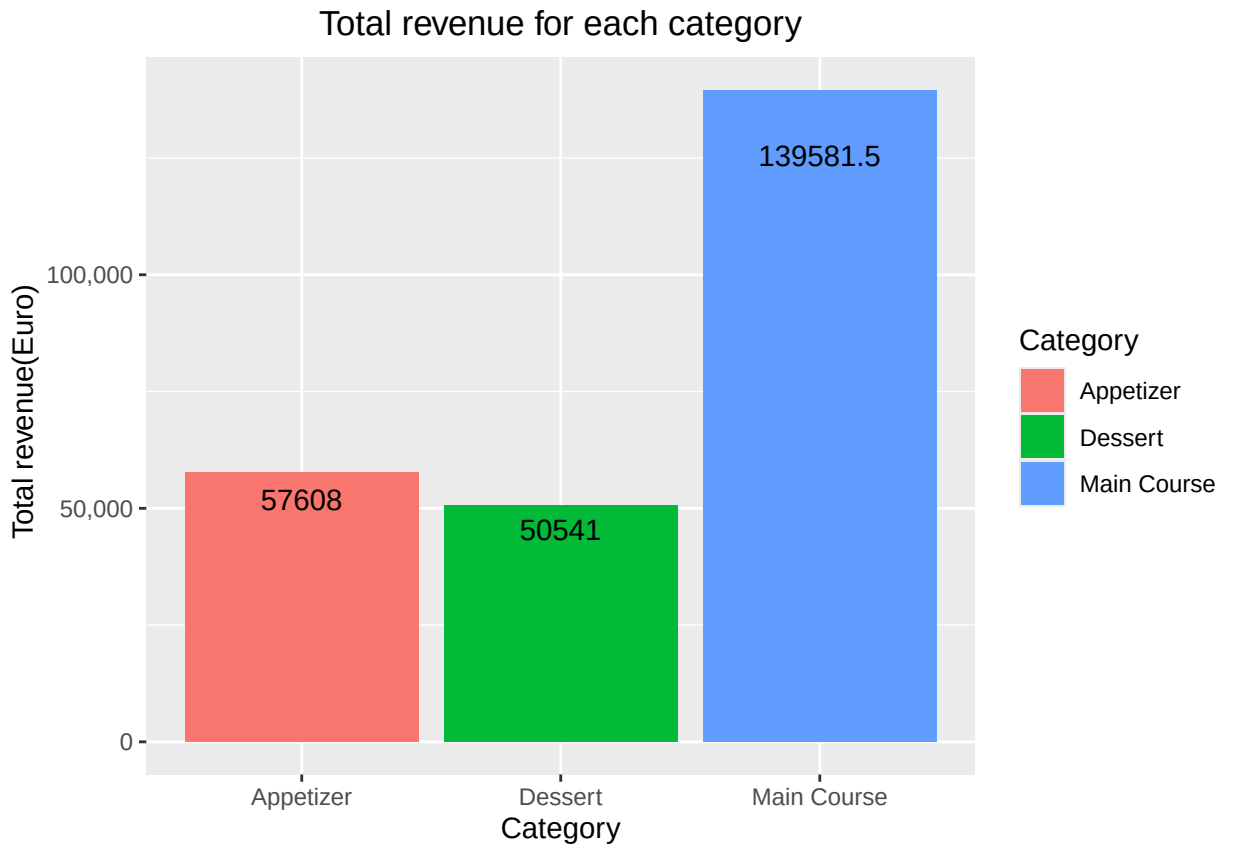
```
####
print('Total revenue for each Category')
```

```
## [1] "Total revenue for each Category"
```

```
revenue_by_category <- data %>% select(Category,`Price (per item)`,Revenue,Quantity) %>%
  group_by(Category) %>% summarise(Average_price=mean(`Price (per item)`),Total_customers=length(Catego
  Total_quantity=sum(Quantity) ,Total_revenue=sum(Revenue)) %>% arrange(desc(Total_revenue))
kable(revenue_by_category)
```

Category	Average_price	Total_customers	Total_quantity	Total_revenue
Main Course	11.646592	4035	11985	139581.5
Appetizer	6.665403	2902	8633	57608.0
Dessert	5.521221	3063	9162	50541.0

```
revenue_by_category %>% ggplot(aes(x=Category,y=Total_revenue)) + geom_bar(stat='Identity',
aes(fill = Category)) + scale_y_continuous(labels =scales::comma ) +
  geom_text(aes(label = Total_revenue),position = position_stack(vjust=0.9)) +
  labs(title = 'Total revenue for each category', y=' Total revenue(Euro)') +
  theme(plot.title = element_text(hjust = 0.5))
```




```
####
```

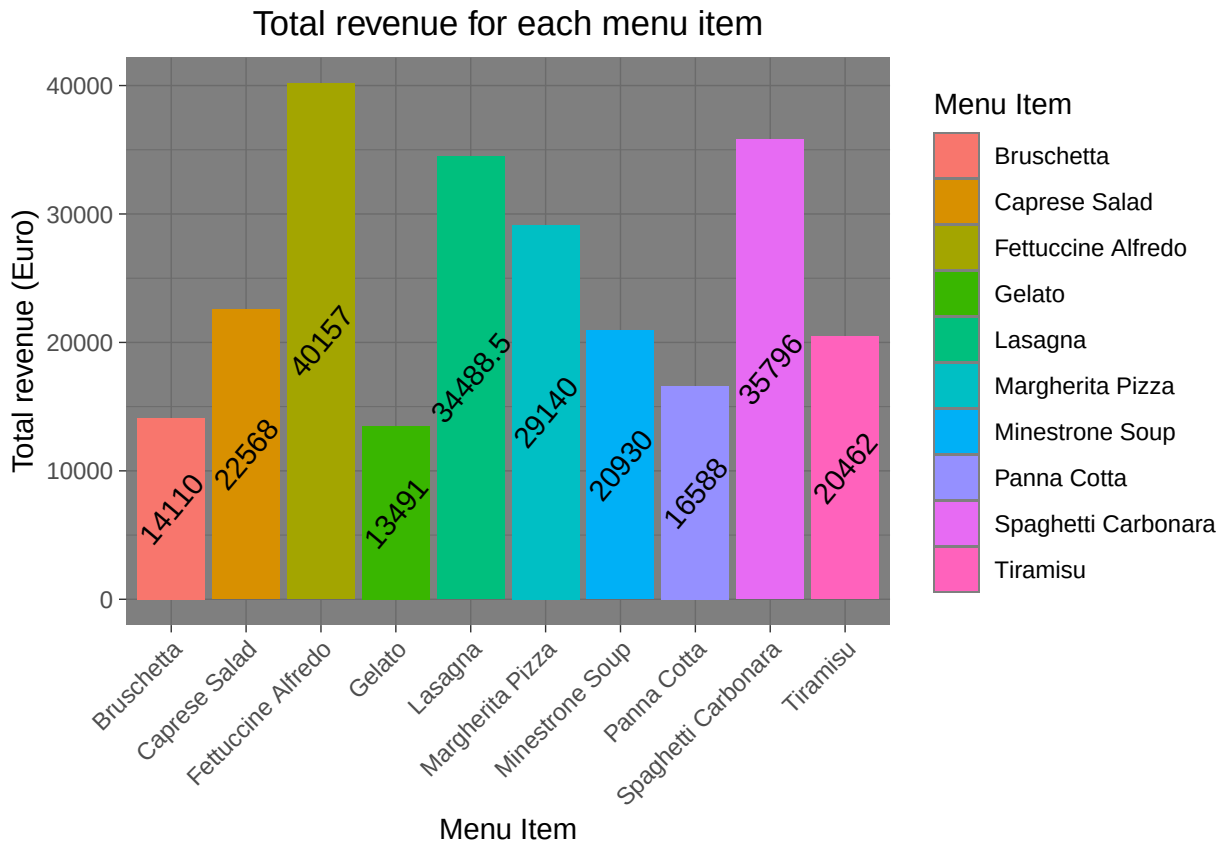
Total revenue per Category for each menu item'

```
Revenue_per_category_by_menu_item <- data %>% select(Category,Quantity,`Menu Item`,Revenue)%>%
  group_by(Category,`Menu Item`) %>% summarise(Total_customers=length(`Menu Item`),

kable(Revenue_per_category_by_menu_item)
```

Category	Menu Item	Total_customers	Total_quantity	Total_Revenue
Appetizer	Bruschetta	958	2822	14110.0
Appetizer	Caprese Salad	945	2821	22568.0
Appetizer	Minestrone Soup	999	2990	20930.0
Dessert	Gelato	987	2998	13491.0
Dessert	Panna Cotta	1024	3016	16588.0
Dessert	Tiramisu	1052	3148	20462.0
Main Course	Fettuccine Alfredo	1036	3089	40157.0
Main Course	Lasagna	1012	2999	34488.5
Main Course	Margherita Pizza	978	2914	29140.0
Main Course	Spaghetti Carbonara	1009	2983	35796.0

```
Revenue_per_category_by_menu_item %>% ggplot(aes(x=`Menu Item`,y=Total_Revenue,fill = `Menu Item`)) +
  geom_bar(stat='Identity') +labs(title = 'Total revenue for each menu item',y= 'Total revenue (Euro)')
  theme_dark()+theme(axis.text.x = element_text(angle = 45,hjust=1),
                     plot.title = element_text(hjust = 0.5))
```




```
####
```

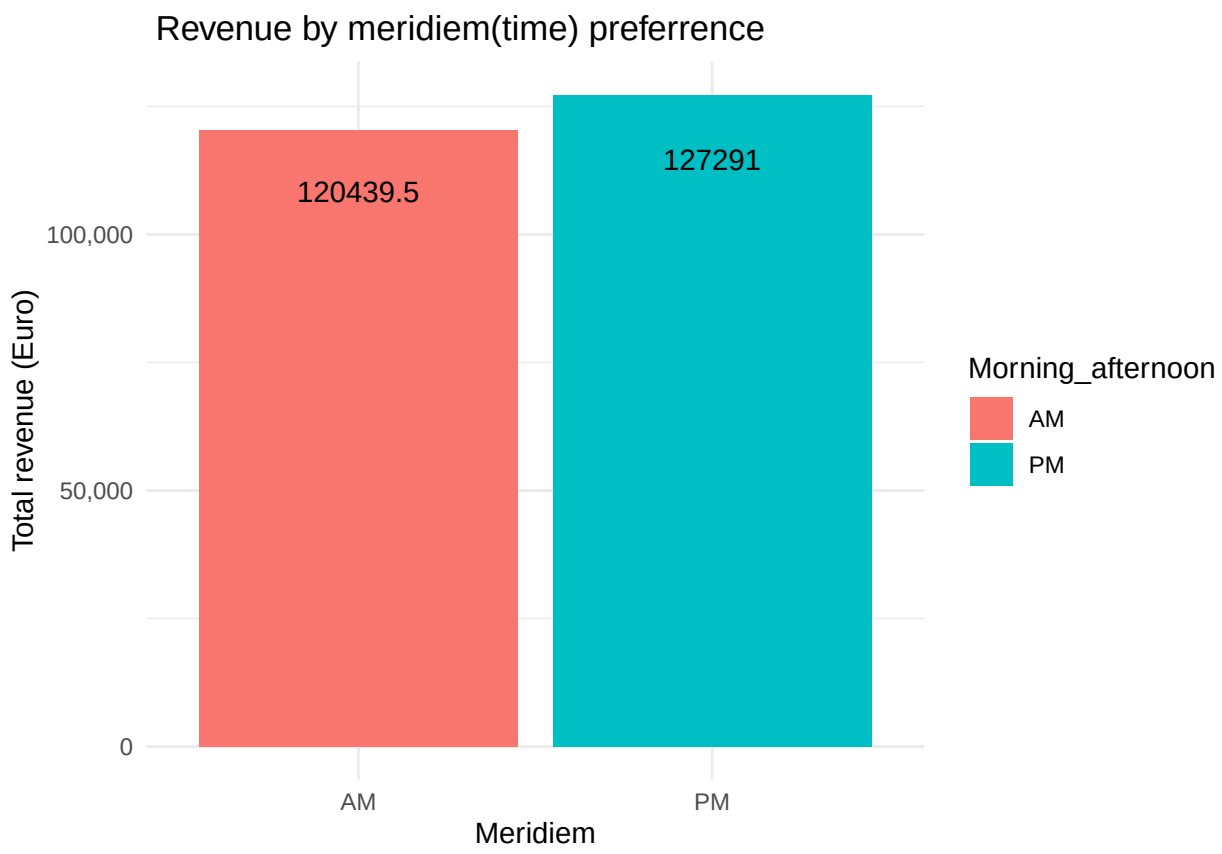
Customer time preference (AM or PM)

```
customer_preferred_time <- data%>% select(Morning_afternoon,Revenue) %>%
  group_by(Morning_afternoon)%>%
summarise(Total_customers=length(Morning_afternoon),Revenue=sum(Revenue))
```

```
kable(customer_preferred_time)
```

Morning_afternoon	Total_customers	Revenue
AM	4887	120439.5
PM	5113	127291.0

```
customer_preferred_time %>% ggplot(aes(x=Morning_afternoon,y=Revenue,fill = Morning_afternoon)) +
  geom_bar(stat='Identity') + labs(title = ' Revenue by meridiem(time) preferrence ',
  y='Total revenue (Euro)',x='Meridiem') + geom_text(aes(label = Revenue),
 position = position_stack(vjust=0.9)) + scale_y_continuous(labels = scales::comma) +theme_minimal()
```



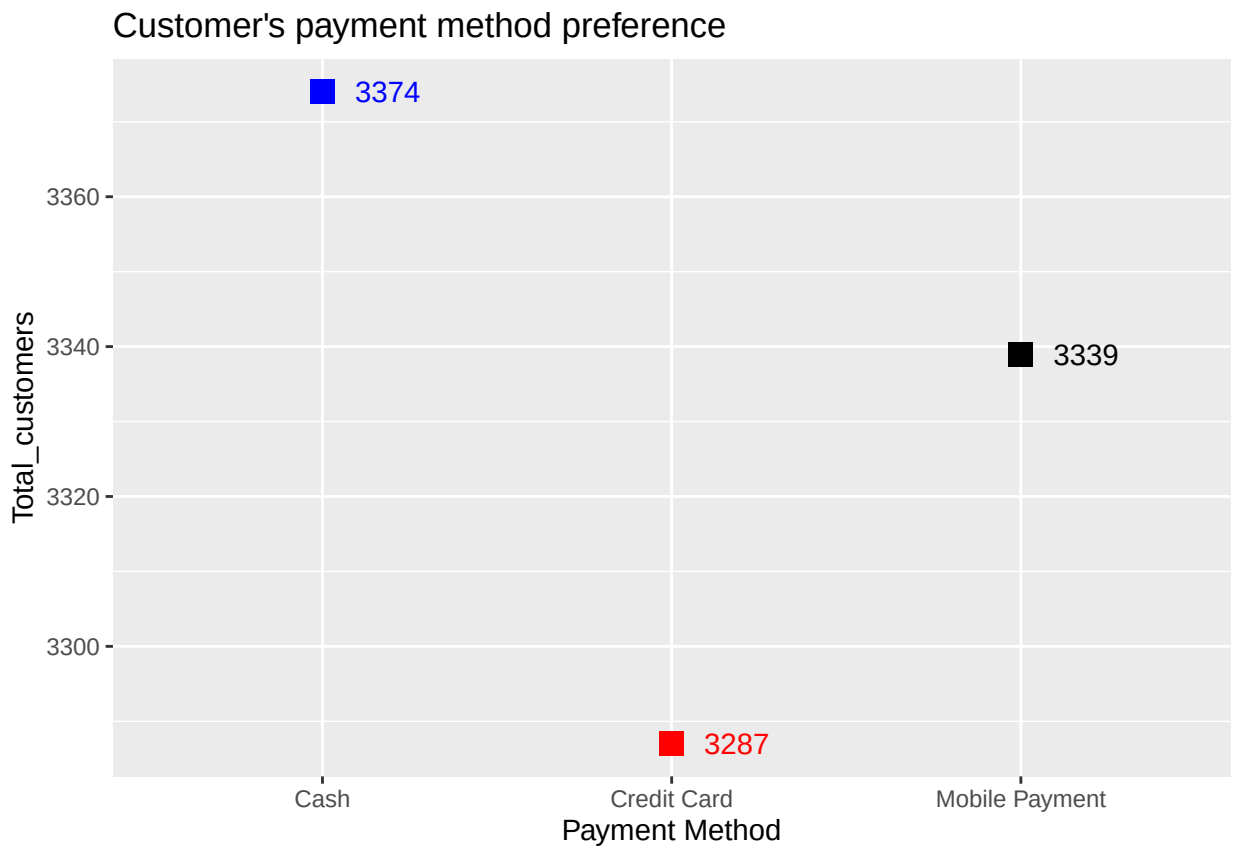

```
####
```

Customer Payment Preference

```
payment_preference<- data%>% select(`Payment Method`) %>%group_by(`Payment Method`) %>%
  summarise(Total_customers=length(`Payment Method`)) %>% arrange(Total_customers)
kable(payment_preference)
```

Payment Method	Total_customers
Credit Card	3287
Mobile Payment	3339
Cash	3374

```
payment_preference %>% ggplot(aes(x=`Payment Method`,y=Total_customers))+
  geom_point(col=c('red','black','blue'),size=4,shape=15)+
  labs( title = "Customer's payment method preference") +
  theme(plot.title = element_text(hjust=0.4))+theme_classic() +
  geom_text(aes(label = Total_customers),
            hjust=-0.5,col=c('red','black','blue'))+theme_gray()
```




```
####
```

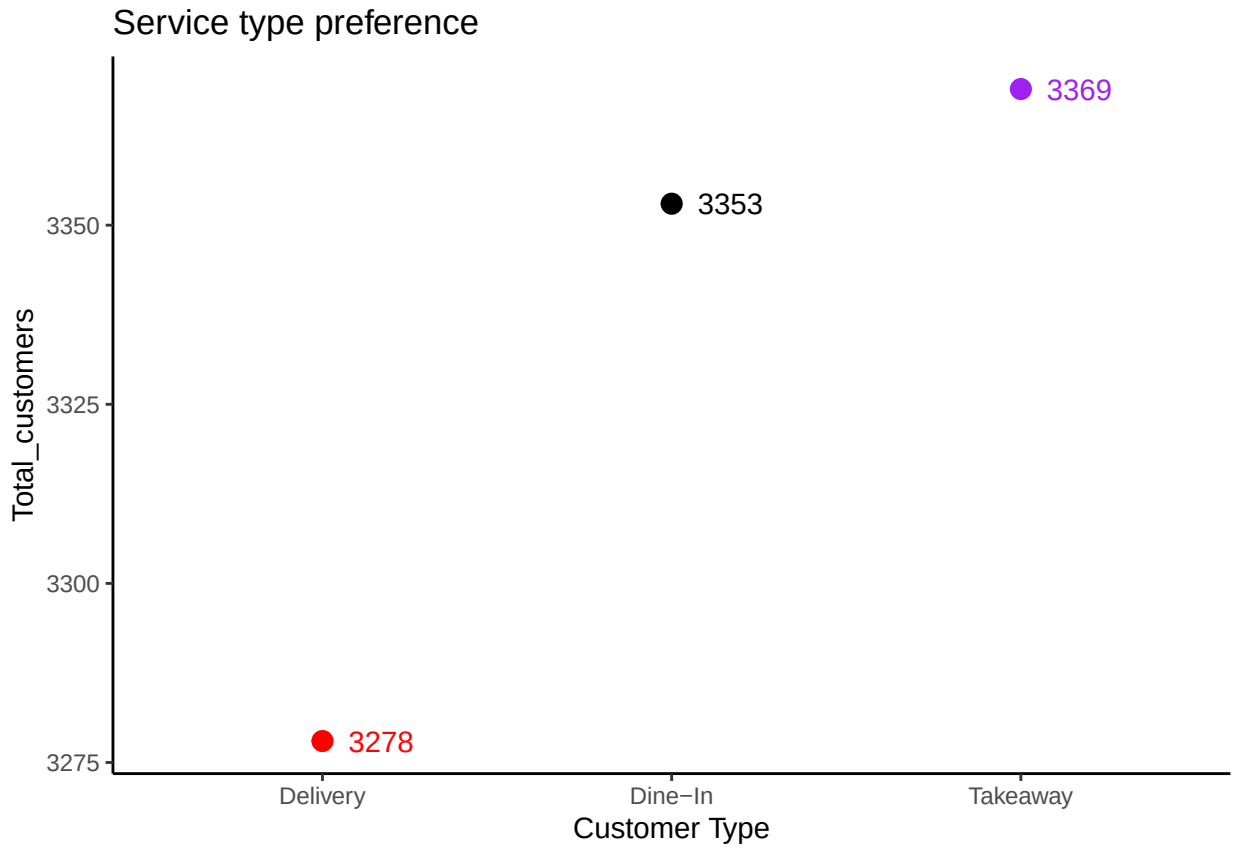
customer type in terms of food distribution

```
customer_type<-data %>% select(`Customer Type`)%>%group_by(`Customer Type`) %>%
  summarise(Total_customers=length(`Customer Type`)) %>%
  arrange(Total_customers)

kable(customer_type)
```

Customer Type	Total_customers
Delivery	3278
Dine-In	3353
Takeaway	3369

```
customer_type%>%ggplot(aes(x=`Customer Type`,y=Total_customers)) +
  geom_point(col=c('red','black','purple'),size=5,shape=20) +theme_classic() +
  geom_text(aes(label = Total_customers),hjust=-0.4,col=c('red','black','purple'))+
  labs(title = "Service type preference")+theme(element_text(hjust = 0.5))
```




```
####
```

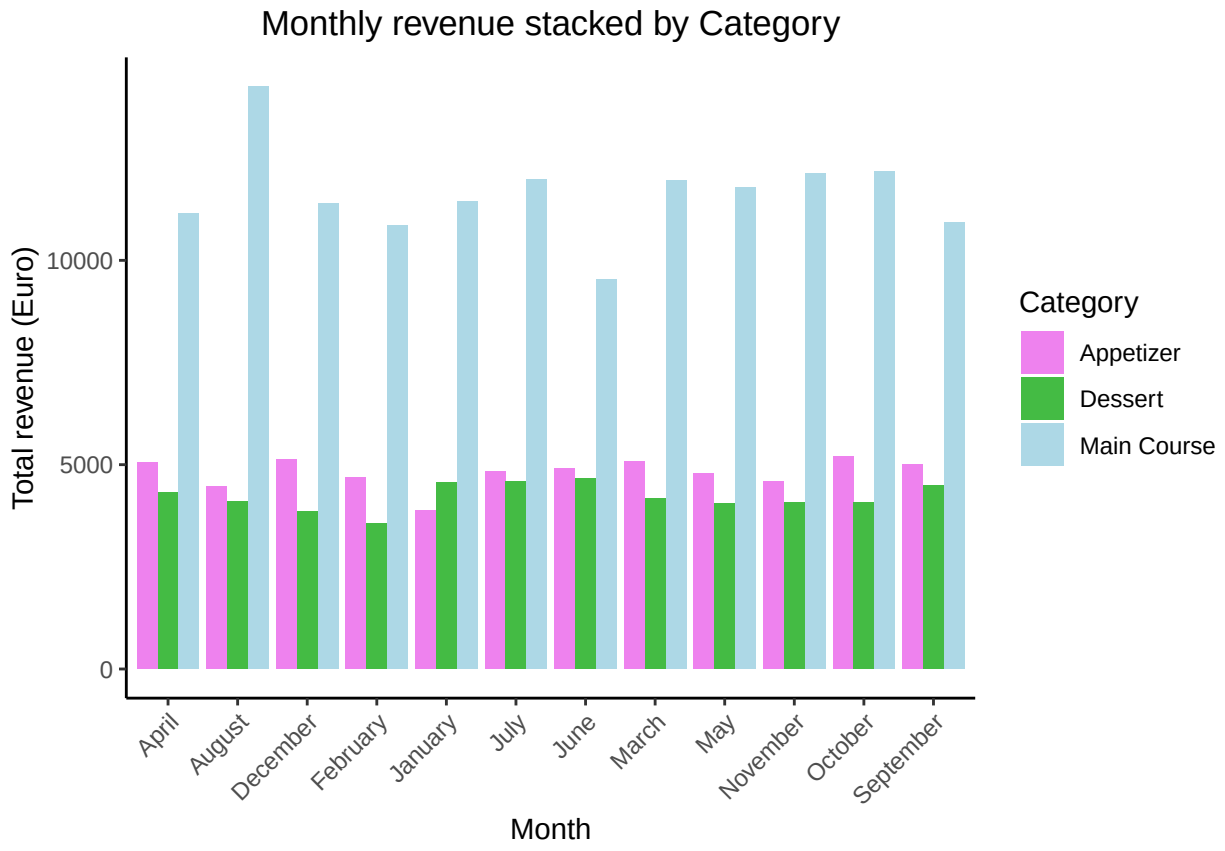
Monthly Revenue

```
monthly_revenue <- data %>% select(Month,Revenue,Category) %>% group_by(Month,Category)%>%
  summarise(Total_revenue=sum(Revenue)) %>%
  arrange(factor(Month,levels = c("January","February","March","April", "May" ,"June",
   "July", "August","September" ,"October","November","December" )))

kable(monthly_revenue%>%select(Month,Total_revenue) %>% group_by(Month) %>%
        summarise(Total_revenue=sum(Total_revenue)))
```

Month	Total_revenue
April	20521.5
August	22827.5
December	20382.0
February	19113.0
January	19871.5
July	21410.5
June	19110.5
March	21217.5
May	20614.5
November	20804.0
October	21448.5
September	20409.5

```
monthly_revenue %>% ggplot(aes(x=Month,y=Total_revenue,fill = Category))+
  geom_bar(stat = 'Identity',position = 'dodge') + labs(y='Total revenue (Euro)',
 title = 'Monthly revenue stacked by Category')+
  scale_fill_manual(values = c('Violet','#4B4','lightblue')) +
  theme_classic() +theme(axis.text.x=element_text(angle = 45,hjust = 1),
                         plot.title = element_text(hjust=0.5))
```




```
####
```

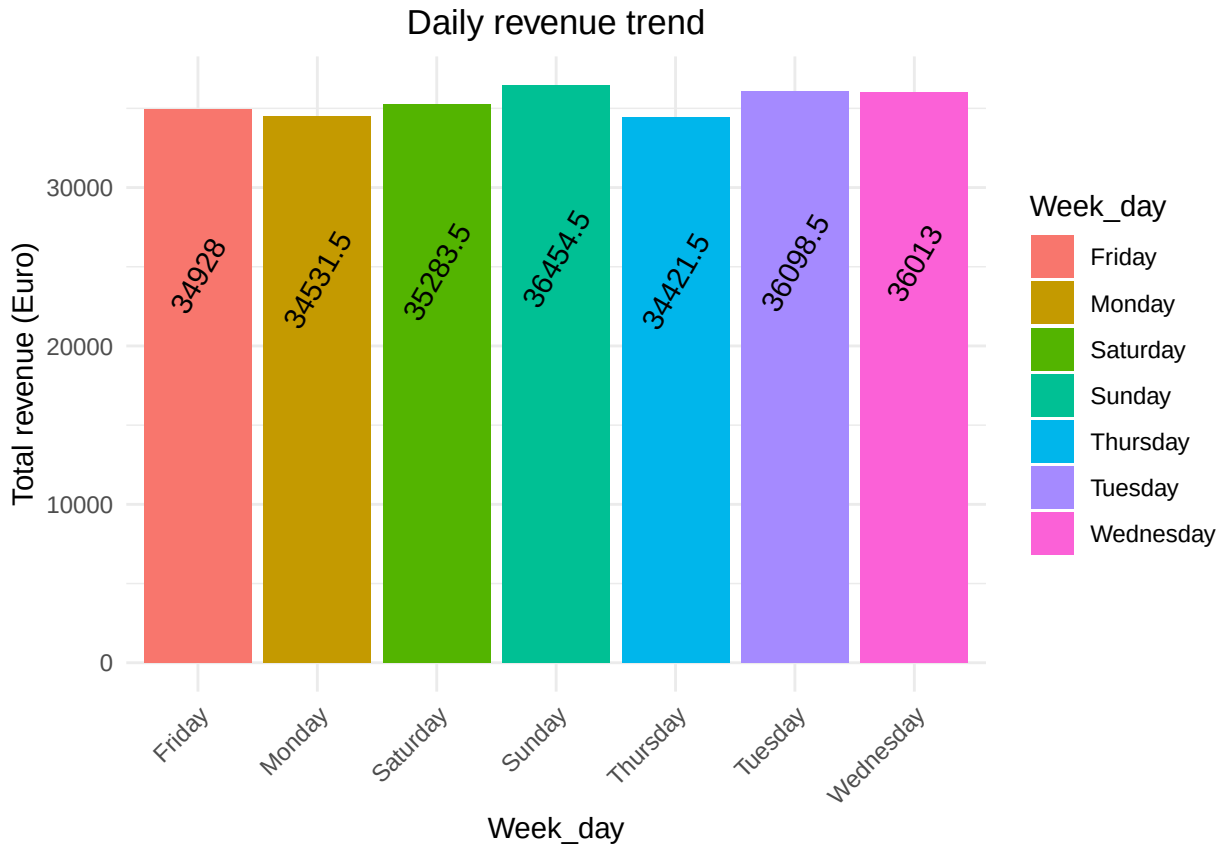
Daily revenue

```
weekly_revenue <- data %>% select(Week_day,Revenue) %>% group_by(Week_day) %>%
  summarise(Total_revenue=sum(Revenue)) %>% arrange(factor(Week_day,
  levels = c("Monday" ,"Tuesday","Wednesday" ,"Thursday","Friday","Saturday","Sunday" )))

kable(weekly_revenue)
```

Week_day	Total_revenue
Monday	34531.5
Tuesday	36098.5
Wednesday	36013.0
Thursday	34421.5
Friday	34928.0
Saturday	35283.5
Sunday	36454.5

```
weekly_revenue %>% ggplot(aes(x=Week_day,y=Total_revenue,fill = Week_day))+
  geom_bar(stat = 'Identity') + theme_minimal() +labs(title = 'Daily revenue trend',
  y='Total revenue (Euro)')+ theme(plot.title = element_text(hjust = 0.5),
 axis.text.x=element_text(angle = 45,hjust = 1)) +
  geom_text(aes(label = Total_revenue),angle=60,position = position_stack(vjust=0.7))
```



```
####
```

Insights

- Food items from the *Main course* category generates more revenue(139581.5 Euros) than the *Appetizer* and *Dessert* categories, with Dessert generating the least revenue(50541 Euros) among the three categories
- *Fettuccine Alfredo* and *Toramisu* are the most liked food items by customers, *Caprese Salad* is the least favourite
- *Fettuccine Alfredo* ,*Spaghetti Carbonara* and *Lasagna* are the top three revenue generating food items, *Gelato* and *Panna Cotta* generate way less revenue
- Most customers buy more in the Afternoon or Evening (PM) than they do in the Morning (AM), hence the Restaurant generates more revenue in post meridiem (PM)
- Most customer prefer paying *cash* compared to the likes of *Credit card* and *Mobile payment*
- Most customers prefer *Takeaways* than *Delivery* and *eating in-store*
- *August* and *October* generated the most revenue with *June* and *February* generating lowest revenue
- The revenue of the weekdays is quite negligible, the difference is barely visible, but *Sunday* seems to generate more revenue than the other weekdays, with *Monday* generating the least revenue