Final Exam – Business Intelligence

Saba Fatima -18556

May 21, 2022

Data Set: https://archive.ics.uci.edu/ml/datasets/in-vehicle+coupon+recommendation

Tool: Power BI

Wrangling

1. Destination **(OK)**
2. Passenger **(OK)**
3. Weather **(OK)**
4. Temperature **(OK)**
5. Time **(OK)**
6. Coupon **(OK)**
7. Expiration **(OK)**
8. Gender **(OK)**
9. Age – The age column was set as data type number initially, yet it had some values for which it can be considered as a categorical column, which was an error to Power BI file. So, editing the data type of this column to **Text,** to handle the categorical values, then replaced **'50plus' with 50** and **'below21' with 20.** After that changed the data type of the column to **Number.**
10. maritalStatus **(OK)**
11. has_children **(OK)**
12. education **(OK)**
13. occupation **(OK)**
14. income **(OK)**
15. car – This column had more missing values than what could be filled in the respective empty columns, so ideally deleting the column was the best approach in wrangling the column or else any method could lead to wrong analysis and wrong insights from the data.
16. Bar – This column had blank cells and replaced the empty cells with **'never',** to analyze it with the never category.
17. CoffeeHouse - This column had blank cells and replaced the empty cells with **'never',** to analyze it with the never category.
18. CarryAway - This column had blank cells and replaced the empty cells with **'never',** to analyze it with the never category.
19. RestaurantLessThan20 - This column had blank cells and replaced the empty cells with **'never',** to analyze it with the never category.
20. Restaurant20To50 - This column had blank cells and replaced the empty cells with **'never',** to analyze it with the never category.
21. toCoupon_GEQ5min – This column is removed since every entry had same values for all the rows.
22. toCoupon_GEQ15min **(OK)**
23. toCoupon_GEQ25min **(OK)**
24. direction_same **(OK)**
25. direction_opp **–** This column was removed, because if the direction_same column had 0's this column would have 1's, so the 0's in direction_same column will imply that the venue is opposite to the driver's residence.
26. Y **(OK)**

KPIs and Dimensions

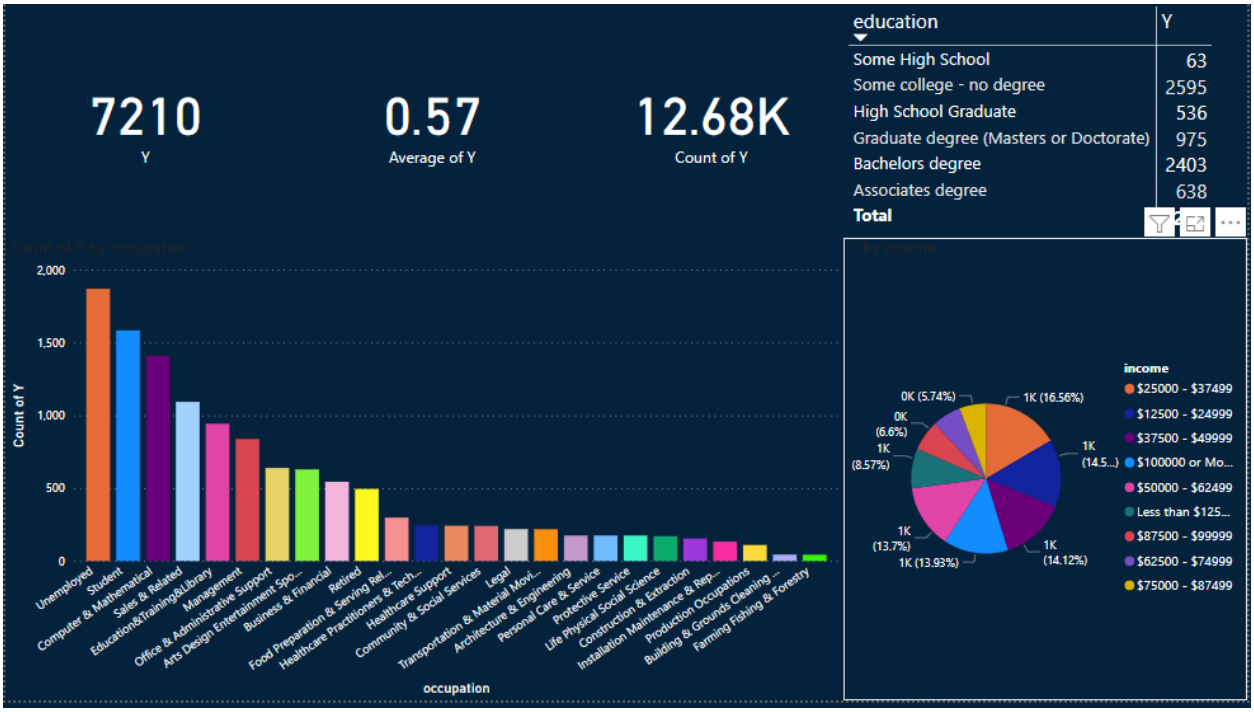
KPIs	Dimensions
Y	Destination Passenger Weather Temperature Time Coupon Expiration Gender Age maritalStatus has_children education occupation income Bar CoffeeHouse CarryAway RestaurantLessThan20 Restaurant20To50 toCoupon_GEQ15min toCoupon_GEQ25min direction_same

Problem Statement

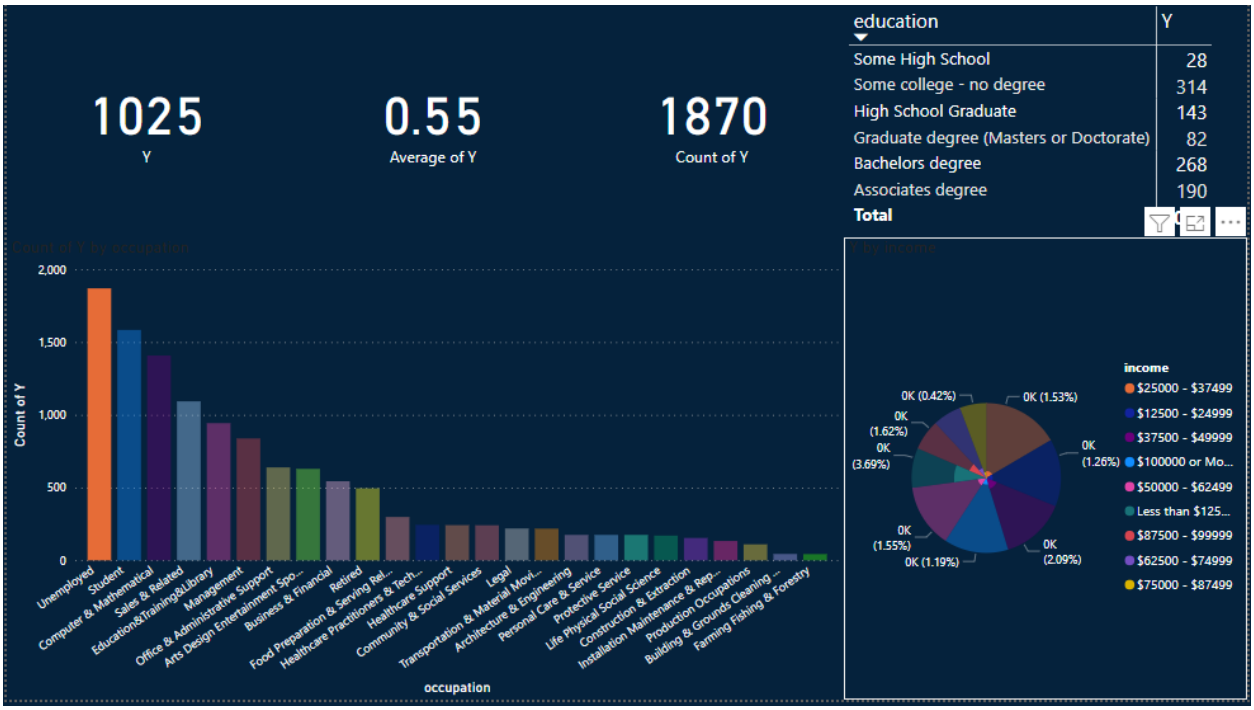
Factors affecting in making a drivers' decision most or least likely to be 'yes'.

Strategy

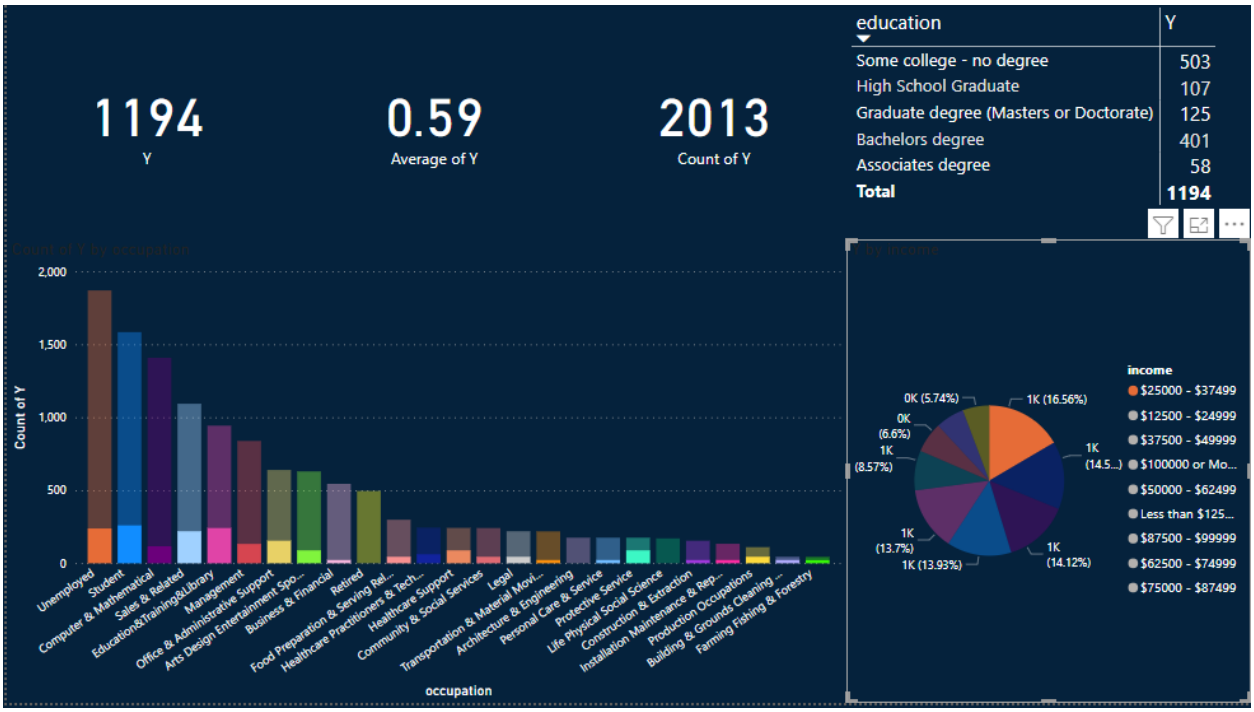
Decision of a driver i.e., Y against; education, occupation, and income



We can see that in terms of occupation, the highest rate for accepting in-vehicle coupons is of Unemployed category, it consists of more than half of all the occupations, which is then followed by student's category and the least contributing occupation is Farming Fishing and Forestry.

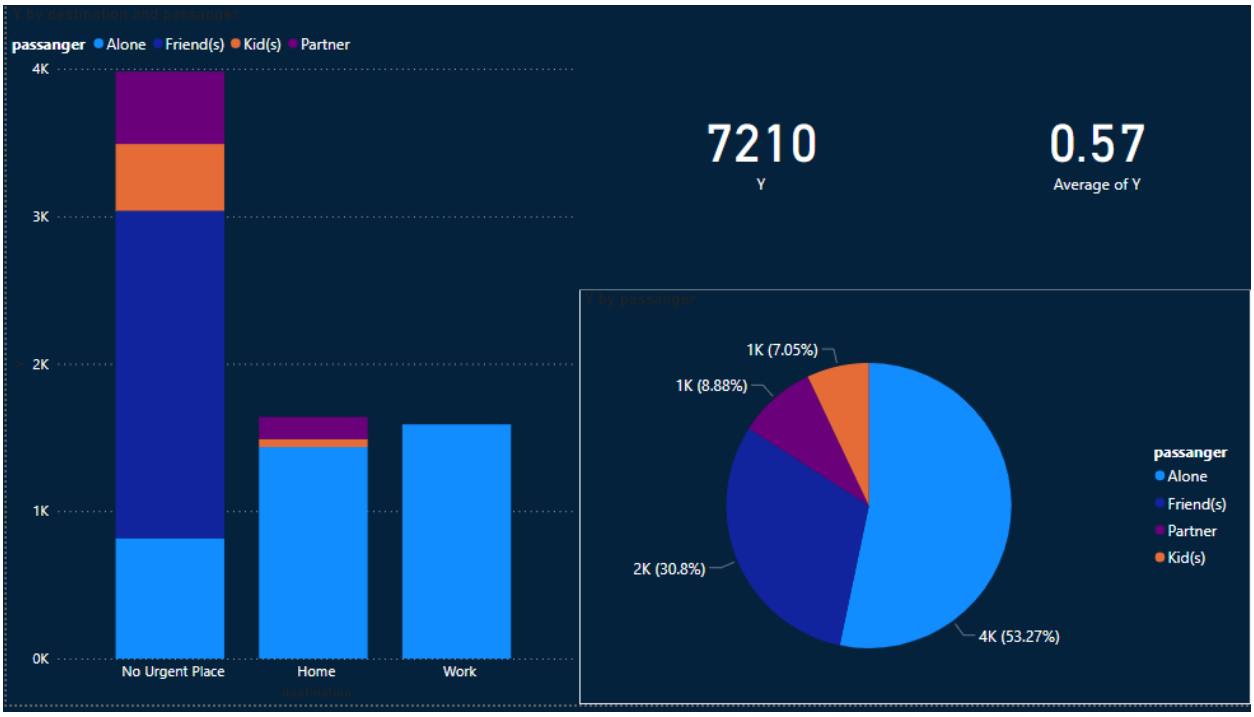


We can also see that category of Unemployed consists of most of the educational categories like, bachelor's degree, Master's degree, etc.

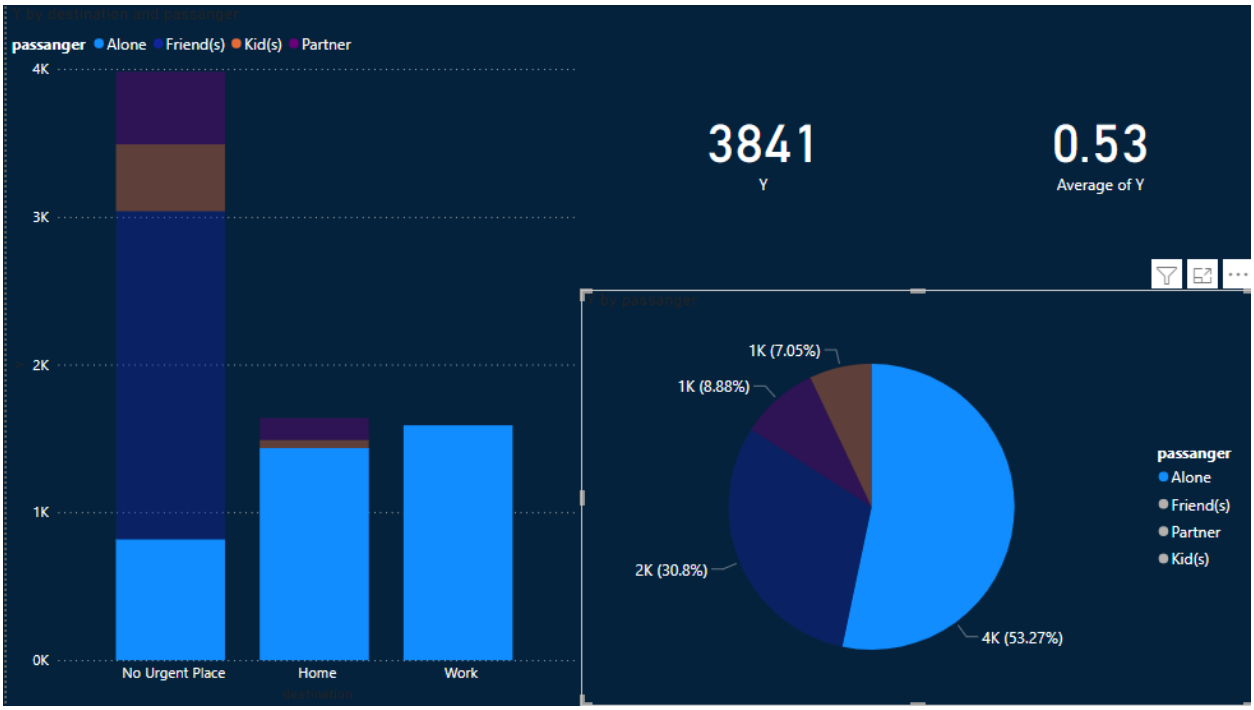


Most of the people belong to the income range of 25000 – 37499, which contributes to most of the occupations and educational categories as well. Yet, the people which has an income ranging from 75000-87499 prefers the coupons least of all categories.

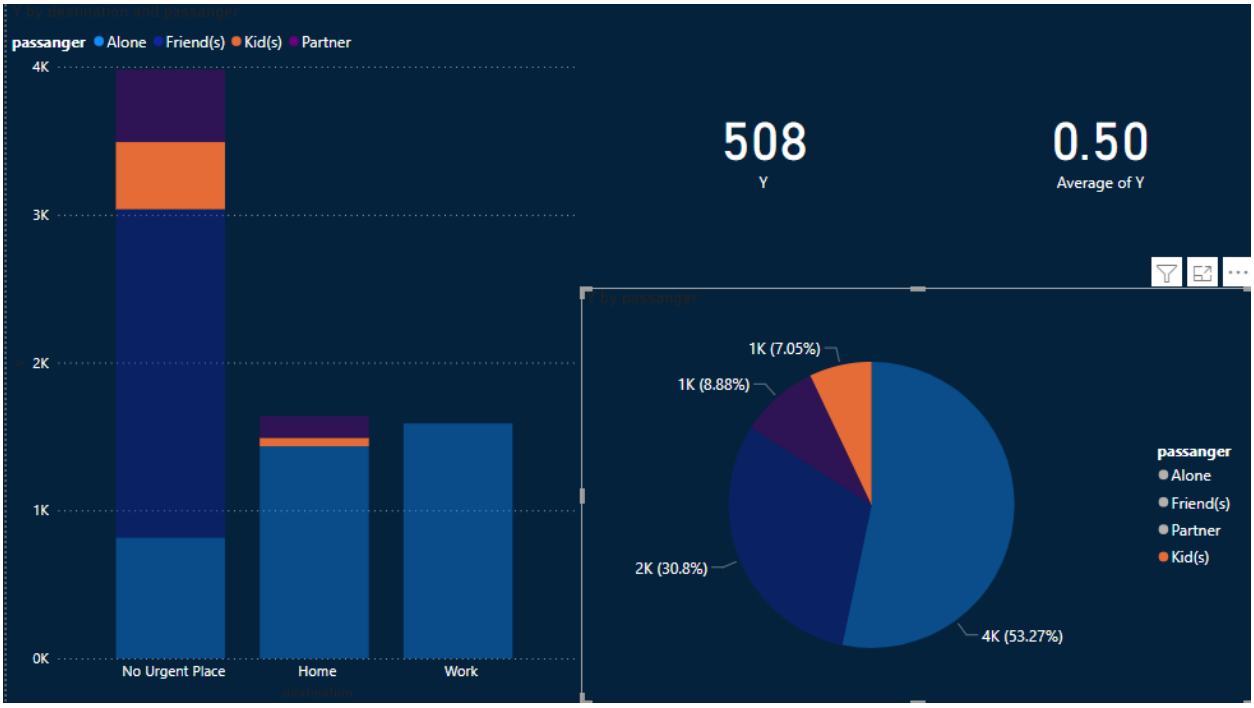
Decision of the driver with respect to Passenger type and destination



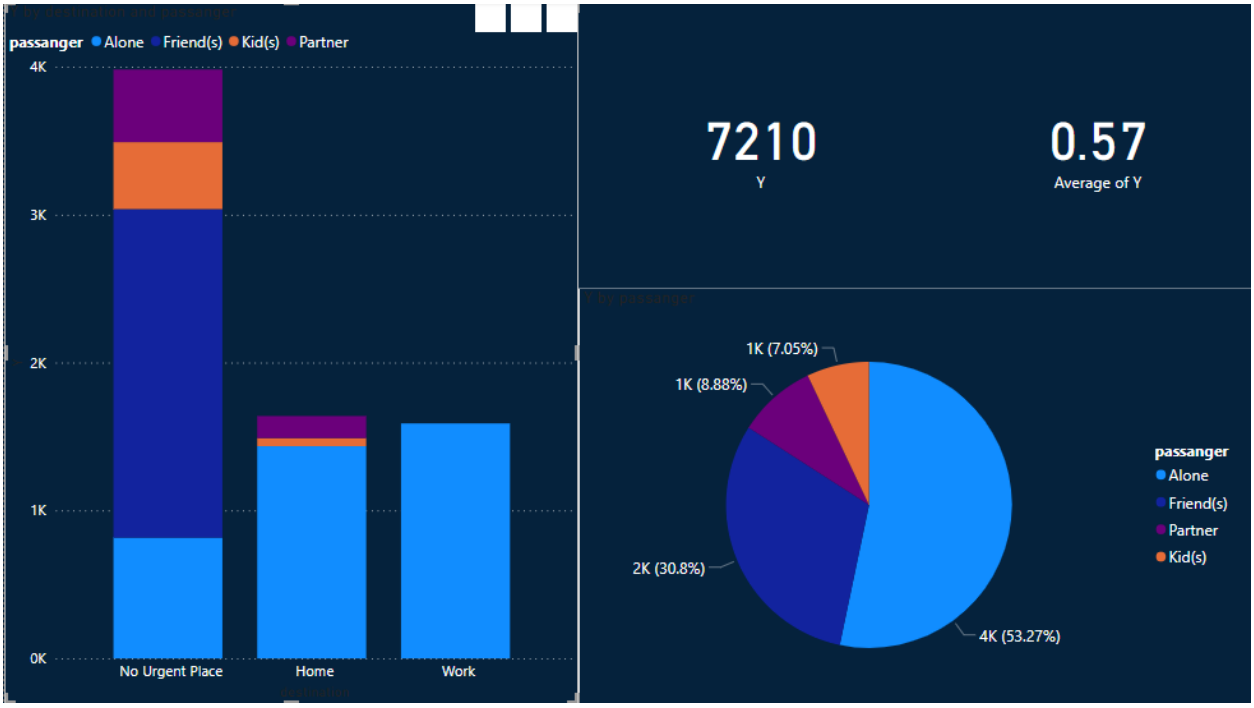
We can analyze from the data that most of the drivers that accept in-vehicle coupons, are most of the time alone.



And when passengers are with their kids, there is a lower chance they might accept a coupon.

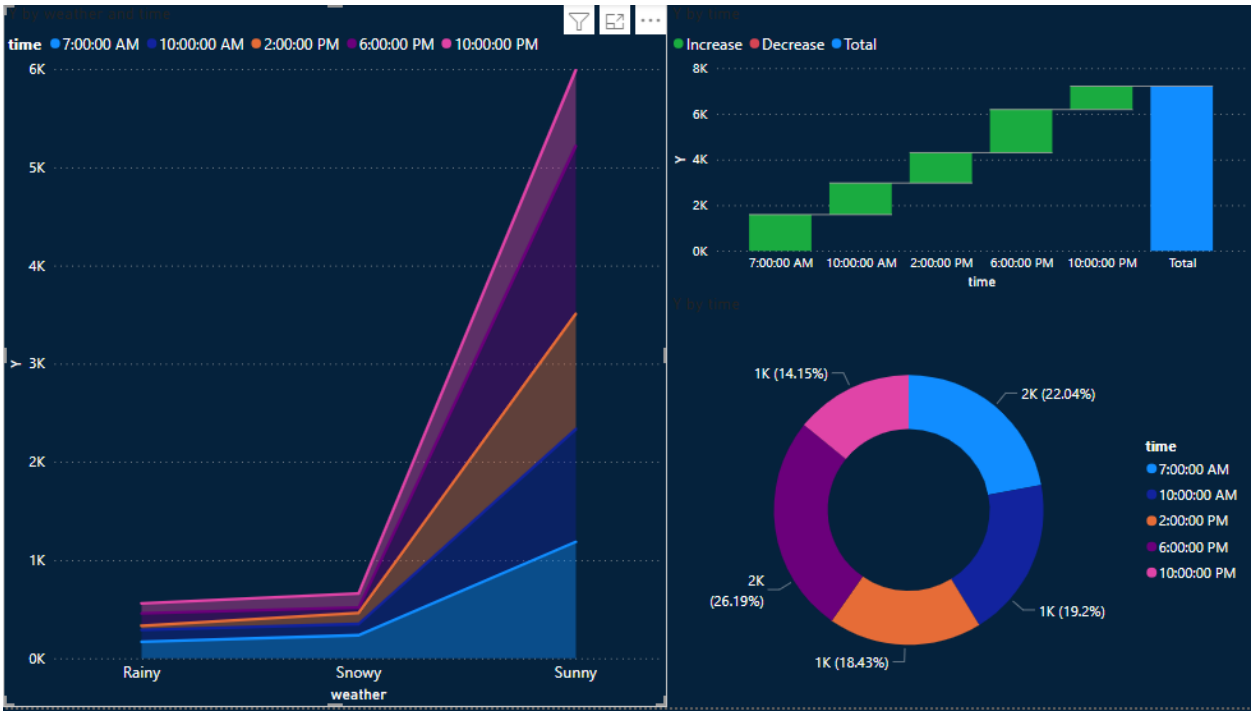


We can also see that, in terms of destination, when a driver is not in urgency the chance of accepting a coupon tends to be more rather than, going for home or work.

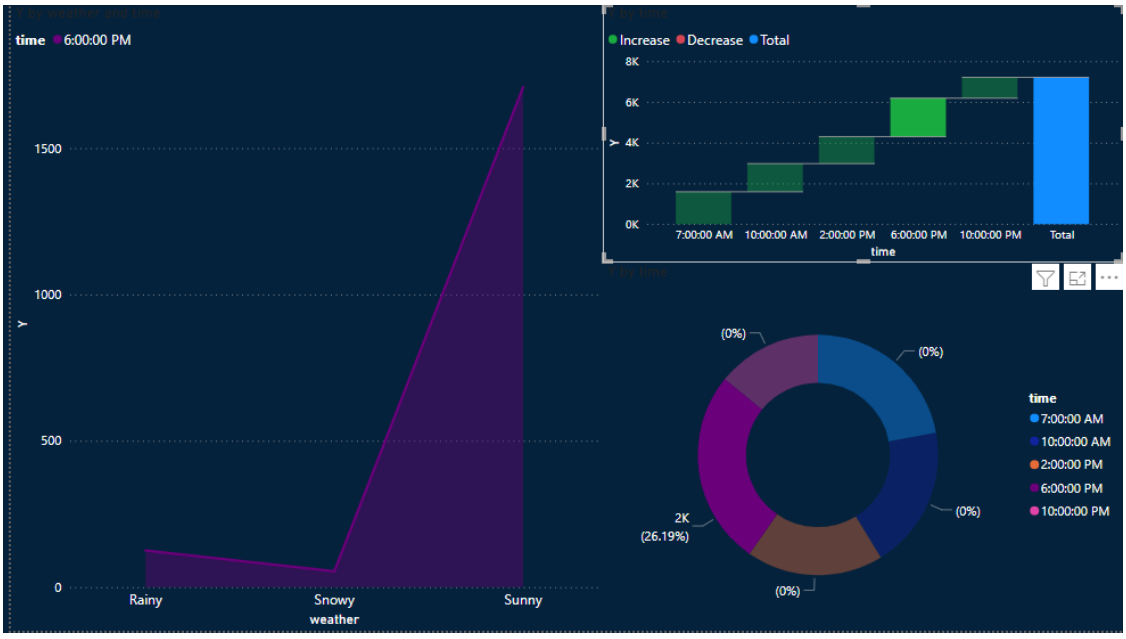


And when the driver is not in urgency, he may or may not be alone, but in most cases drivers who are not in a hurry and are with their friends, may be more willing to accept a coupon, yet in the other cases when a driver is in hurry either for home or work, he will prefer having a coupon when he is alone.

Decision of a driver with respect to time and weather



We can see from the data that most of the times the driver accepted a coupon is when the weather was Sunny, it's pretty explanatory that in Rainy weathers driving is more of a hassle and during monsoon business can witness a loss. But from Snowy to Sunny weathers, we can see a steep which means that the acceptance of coupons grew at a faster pace leading to many profits during the months of May, June, July.

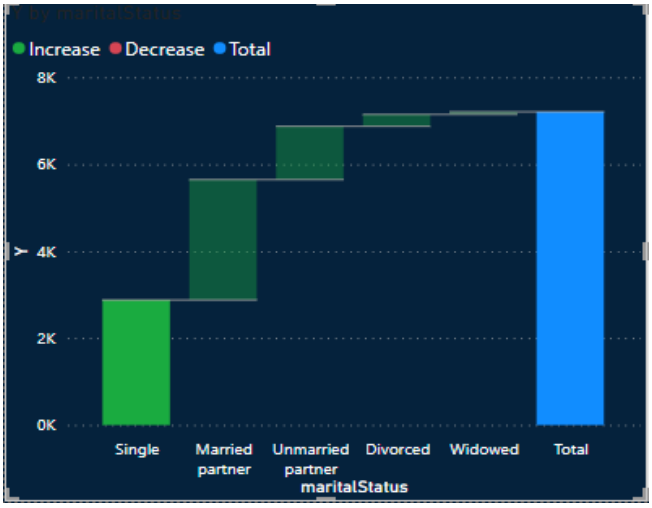


We can analyze from the data, that drivers prefer to accept coupons in evening or around 6:00 PM, maybe this is the time when people get off from their work and would prefer to hang out with their peers.

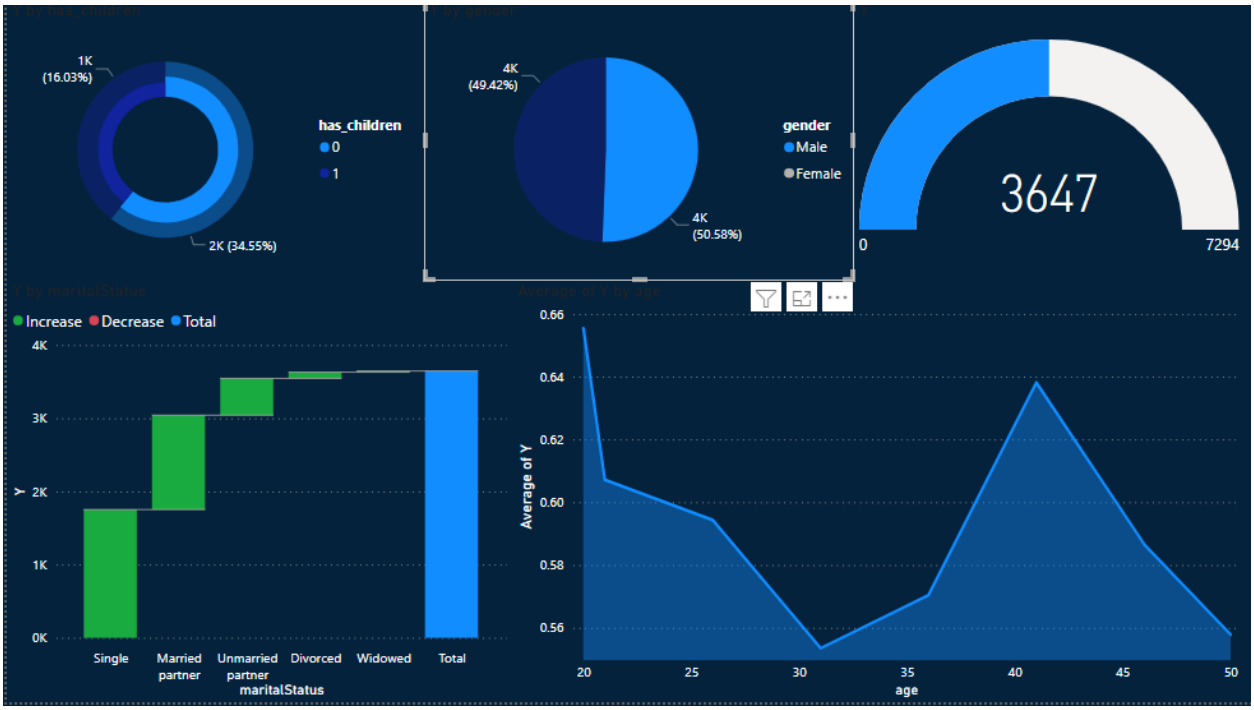
Decision of a driver with respect to its personal data (age, marital status, children, gender)

A driver's decision for accepting a coupon can also be impacted by his personal data, likewise people with age between 20 to 25 years, prefer to accept coupons much more often than people from other age groups. And we can see that old people tend to accept it less.

We can also see that there is a dip in acceptance of coupons between the ages of 30 and 40 years (mainly 35). This is an interesting pattern that have come to notice.

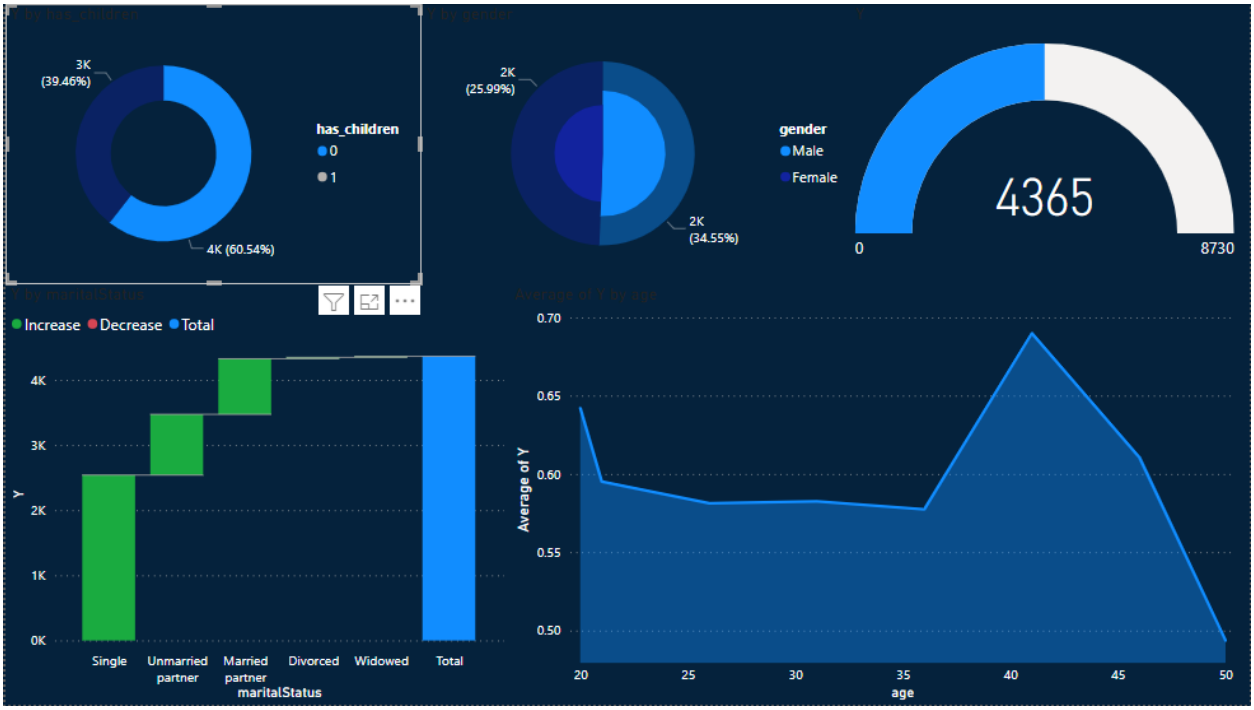


From the marital status of the drivers, we can see that most of the times Single drivers are willing to accept a coupon and after those, married partners consider accepting coupons. Out of all the groups, widows consider less to accept a coupon.



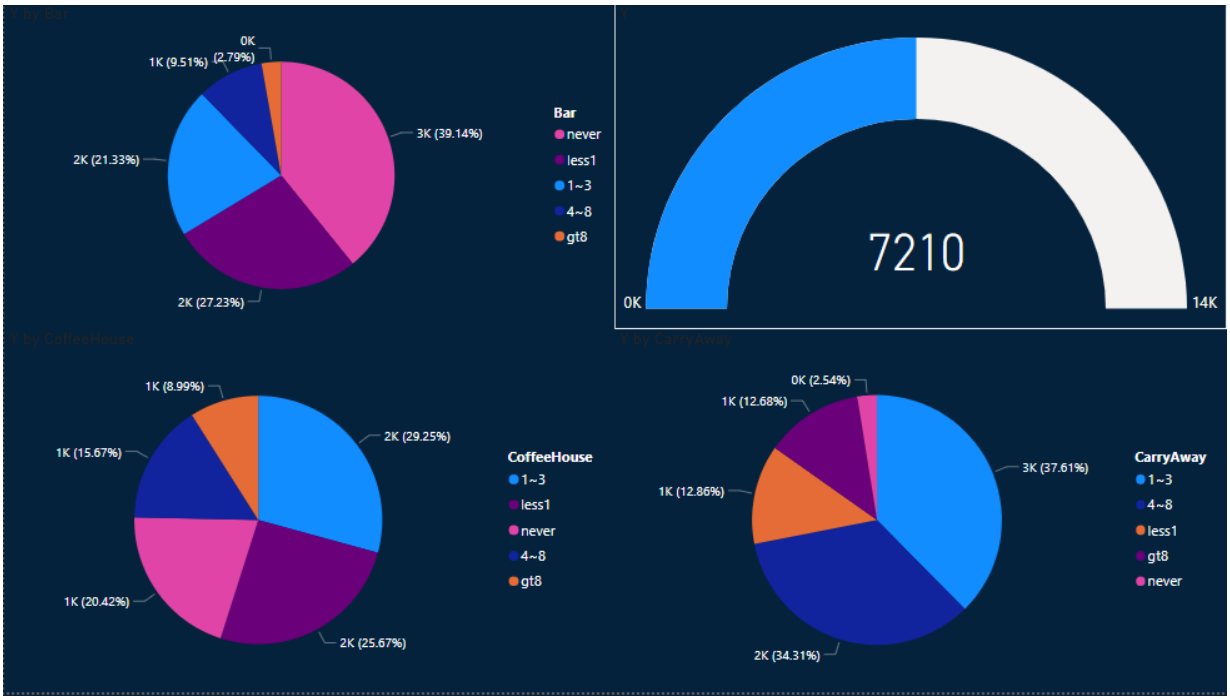
We can see that most of the drivers who accept the coupons are male and most of them are 20 or below, after that we see a decline in pattern of accepting coupons by male between the age group of 21 to 30 and after that the acceptance increases steadily while it reached the peak between the ages of 40 and 45, which is again an interesting pattern.

We can see that, females have a lower rate of acceptance towards coupons than males, yet most women lie between the age group of above 20 to 30 and after 30 we see a drop in acceptance which increases again after the age of 35.

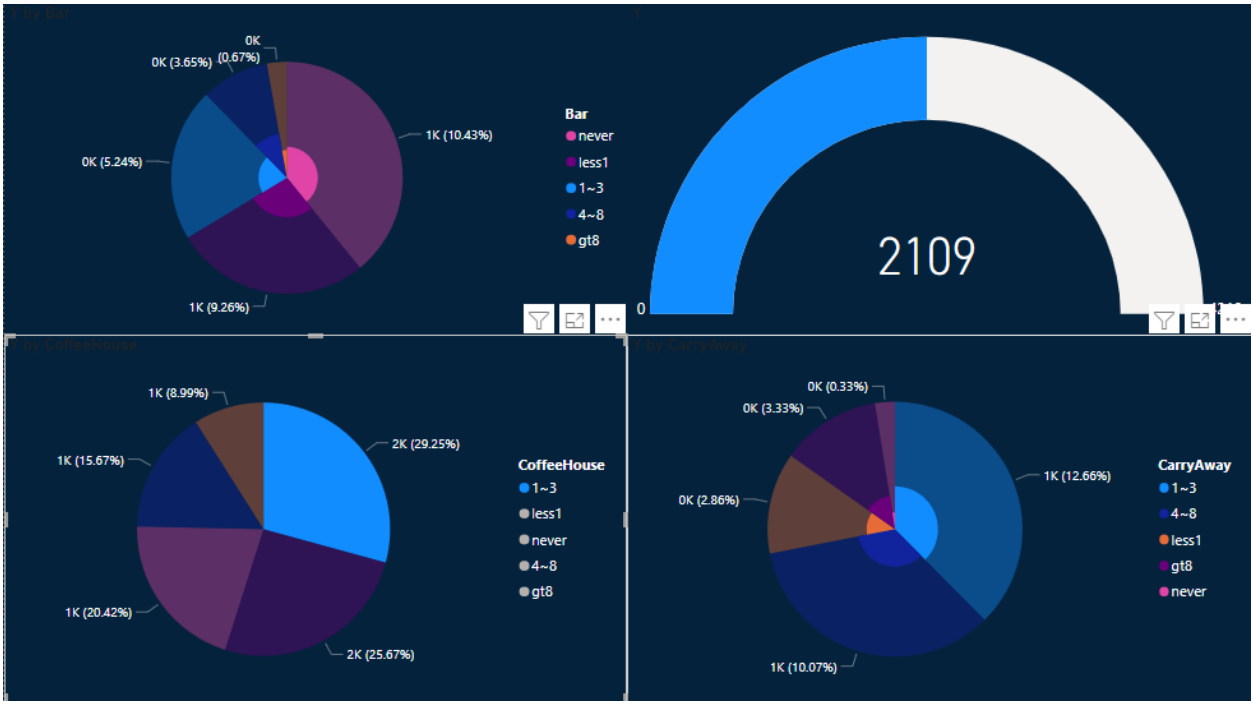


The drivers, who don't have children are either single or belong to the age group between 35 to 50 years and mostly consists of males, tend to accept coupons more than other groups.

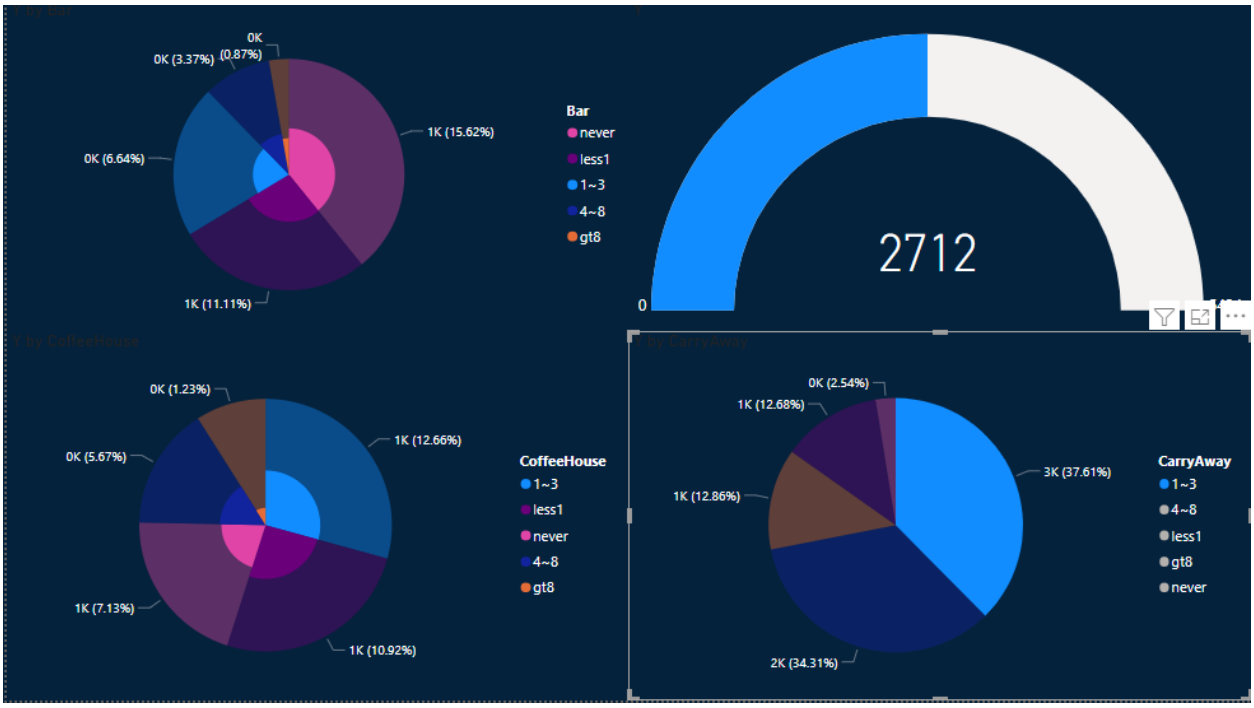
Driver's decision against visiting Bar, coffee house and take away restaurants



We can see that, most of the people who doesn't visits the bar are more inclined to accept a coupon, followed by people who visits the bar once a month.

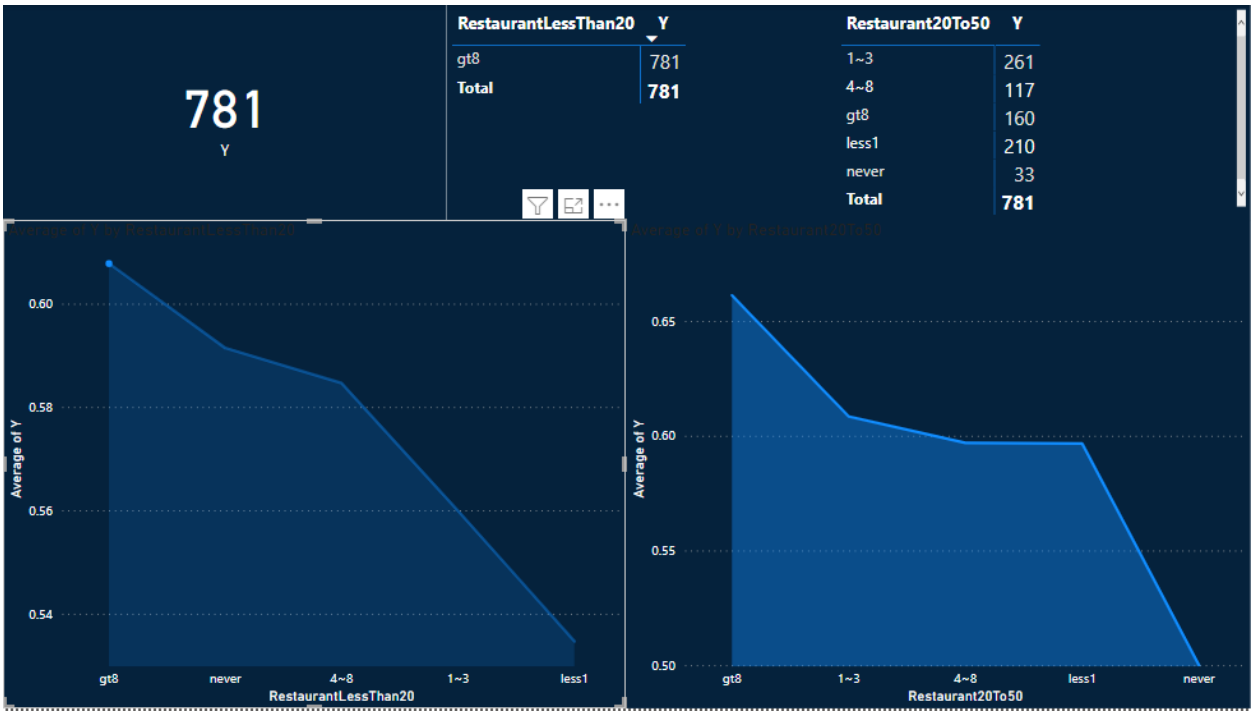


People who visit the coffee house at least for 1 to 3 times a month tend to accept coupons much more often than other groups of people.

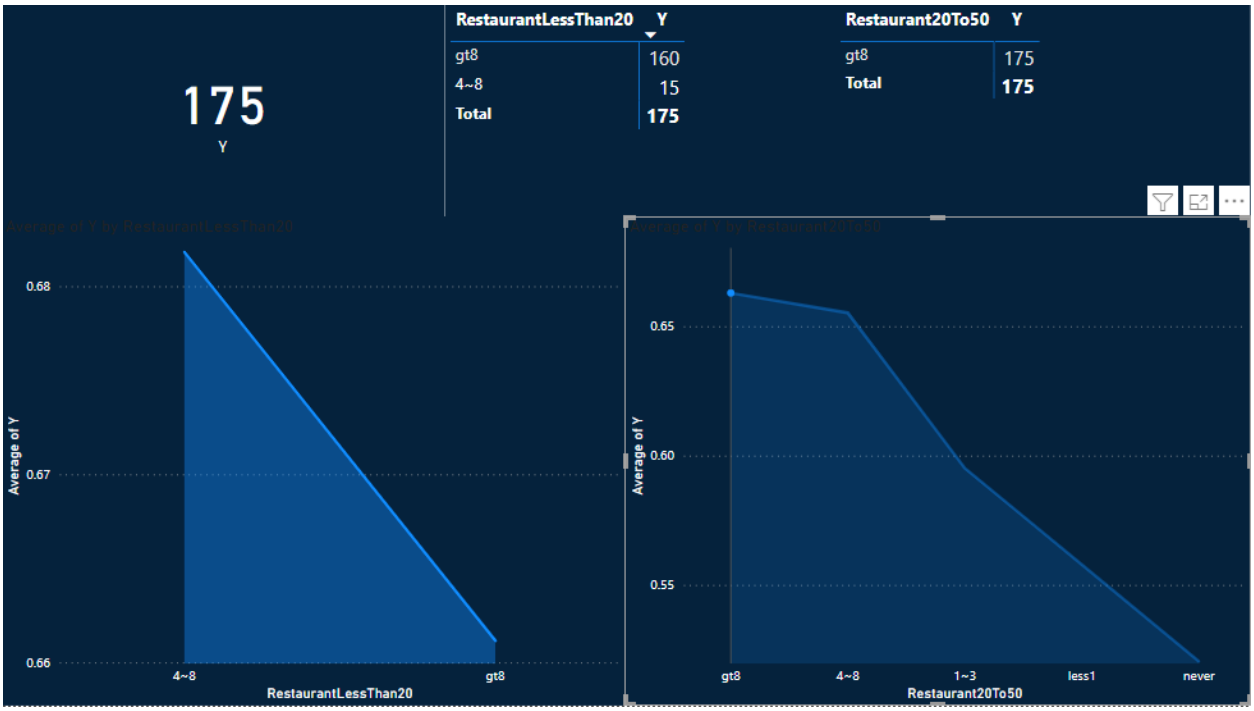


People who prefer take-away 1 to 3 times a month also prefers to accept coupons more often.

Driver's decision with respect to Restaurant visits with less than $20 and more than $20

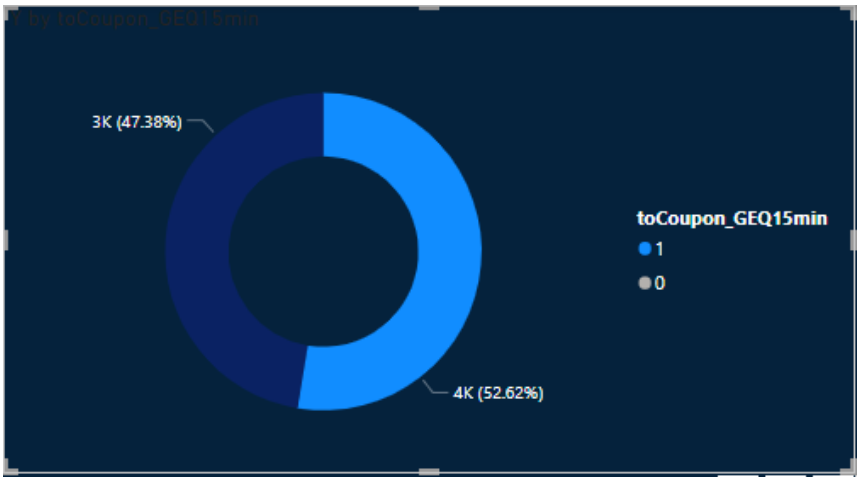


People who have visits to restaurants where the average expense is less than $20 tend to accept a coupon more often. In fact, those people who have these visits more than 8 times a month are more likely to accept the next coupon.

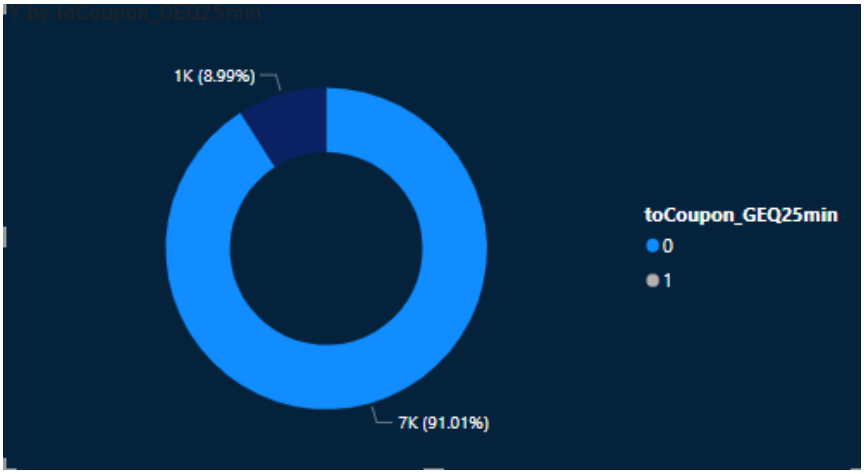


Relatively, people don’t prefer to visit restaurants with an expense per person of more than $20. Yet, among those people who visit the restaurant more than 8 times a month, are more likely to accept a coupon.

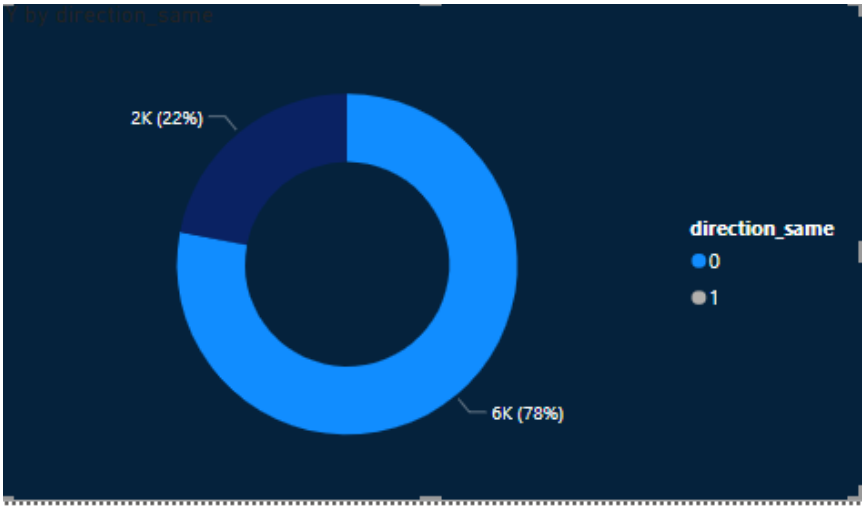
Driver’s decision with respect to commute time to restaurants



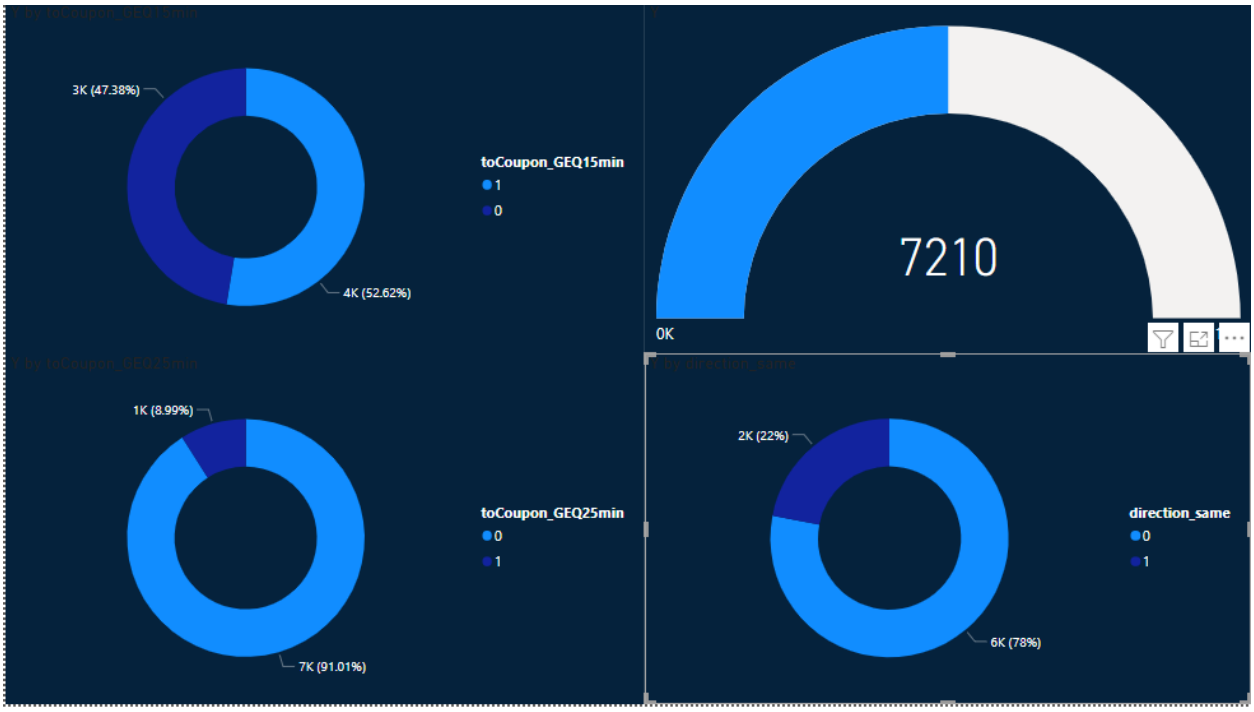
People who have a commute time to restaurants is 15 minutes max are more likely to accept a coupon than those who have more commute time than 15 minutes.



People who have a commute time to restaurants is 25 minutes max, are less likely to prefer a coupon.



And we can see a strange pattern that those people who tend to live in the opposite direction of the restaurants are more likely to accept a coupon than those who live in the same direction.



Key Insights

- **Drivers who are more likely to accept a coupon:**
  Unemployed (mostly students)
  Income 25000-37499
  Not Urgent (with friends)
  If Urgent (Alone)
  During Sunny weathers
  In evening (around 6 PM)
  Between the age group 20-25
  Single
  Has no children
  Males (between 40 - 45)
  Females (between 20 – 35)
  Bar Visit (less1 to never)
  Coffee House visit (1 – 3 times)
  Carry Away (1 – 3 times)
  Restaurant visits that expense less than $20
  Commute to restaurant (Max 15 minutes)
  Restaurant in opposite direction

- **Drivers who are less likely to accept a coupon:**
  Employed in Farming, fishing, and forestry.
  Income 75000 – 87499
  Not Urgent (with Kids)
  During rainy and snowy weathers
  At night (10 PM)
  Between the age group 30- 40
  Widows
  Has children
  Males (21 – 30)
  Females (30 – 40)
  Bar Visit (more than 8 times)
  Coffee House visit (more than 8 times)
  Carry Away (never)
  Restaurant visits that expense more than $20
  Commute to restaurant (max 25 minutes)
  Restaurant in same direction

Conclusion

From the above analysis and key observations, we can conclude that, we need to initiate such campaigns for the coupons, that the target customers will be satisfied, likewise drivers that are from farming or forestry or drivers with income ranging from 75000 – 87499.

We can also conclude that, some campaigns should be stopped such as the one during rainy or snowy weathers as they might be a overhead to the budget, we can also take initiatives to tackle this problem by initiating incentives at this season of the year which can also attract many new customers.

The End 😊