Segmentation_3865

Mokeli Pelamoko

June 23, 2016

1. About the Project

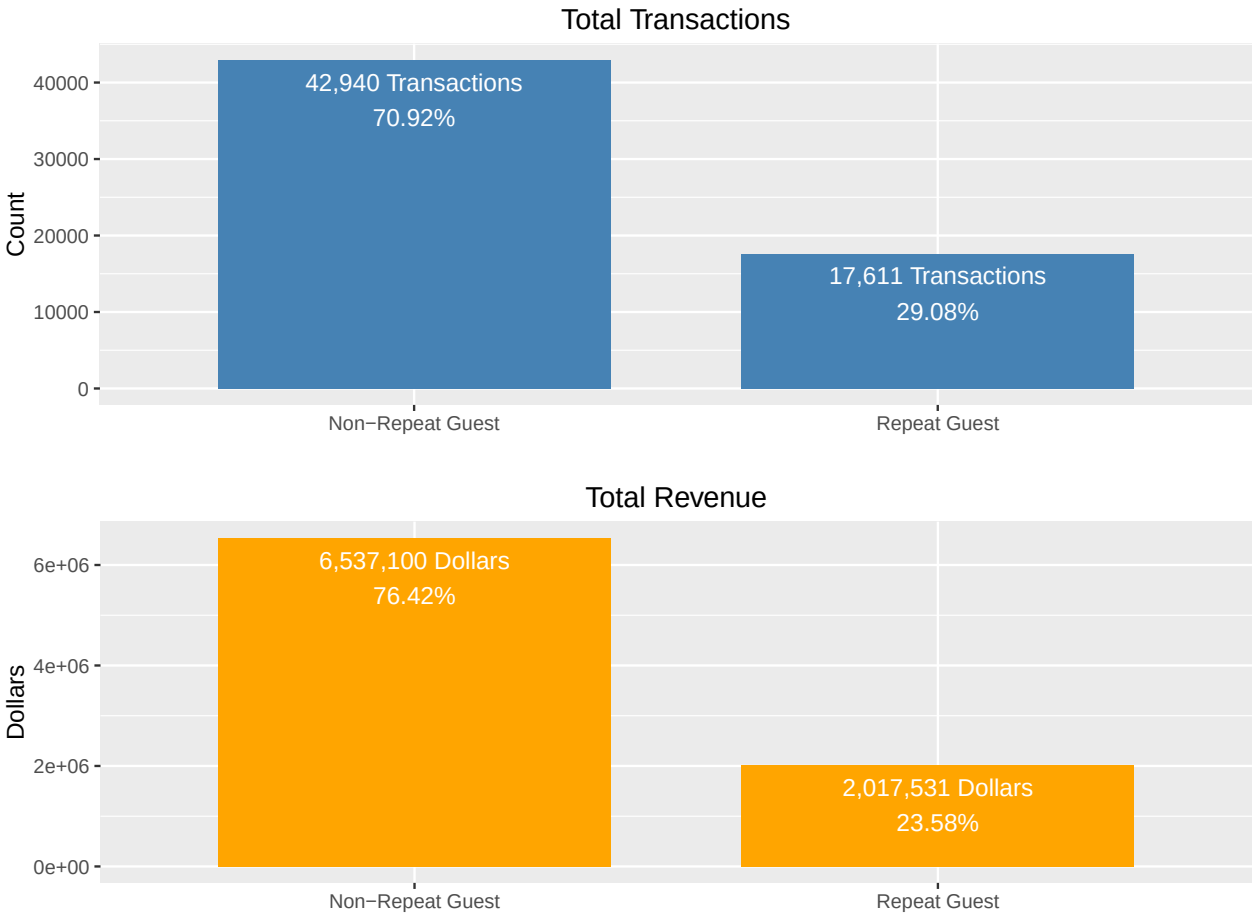
1.1 About the Dataset

Segmentation for restaurant 3865. Insights are derived from the summary of each visit by merging POS, POS Item, and Reservation by loyalty reservation id number. The POS Item was summarized by loyalty reservation before the merging with POS and Reservation. In many of the visualizations, I removed the top 5% of feature in order to reduce the noise from outliers.

1.2 Loading Packages

1.2 Loading Dataset

1.3 Total Transactions

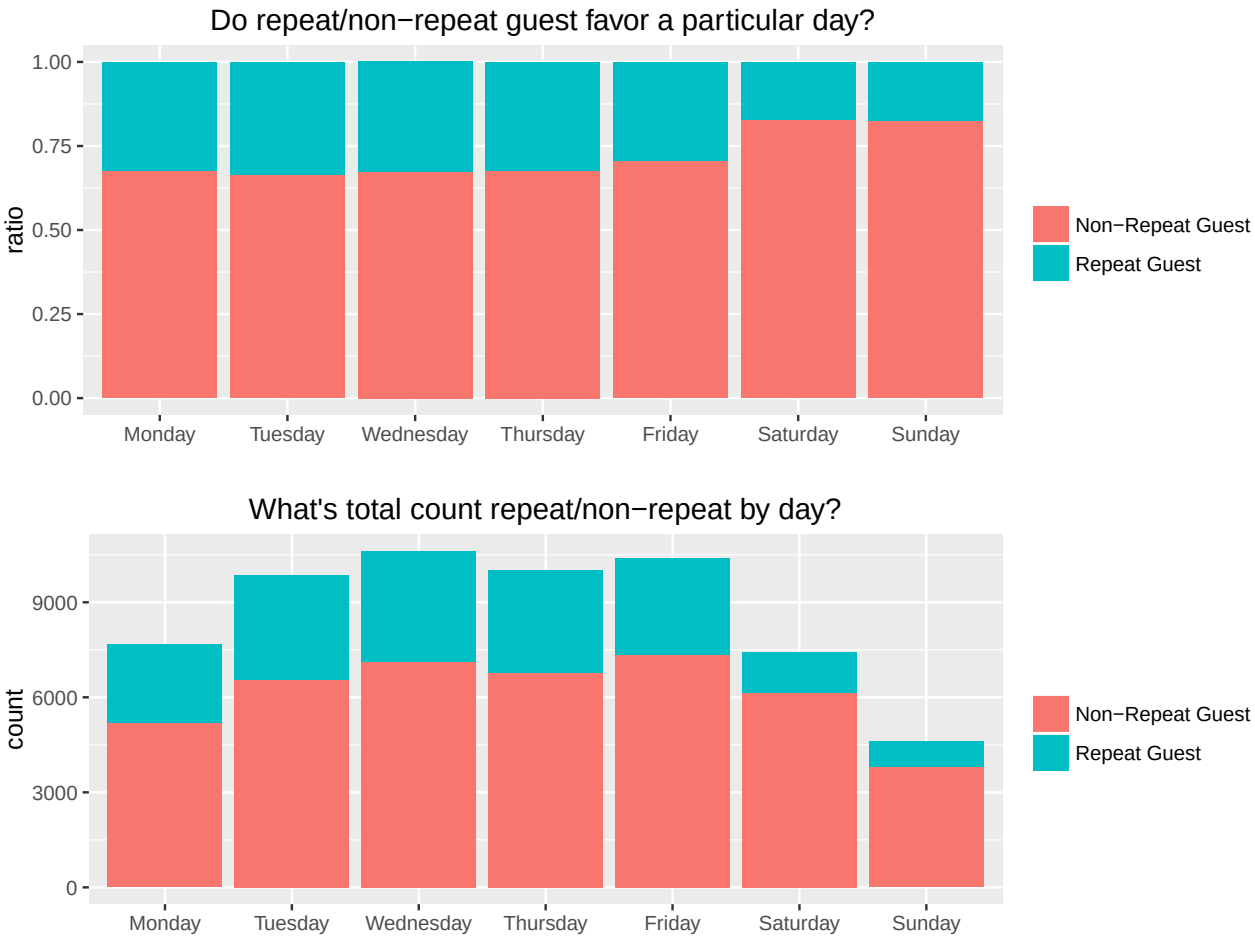


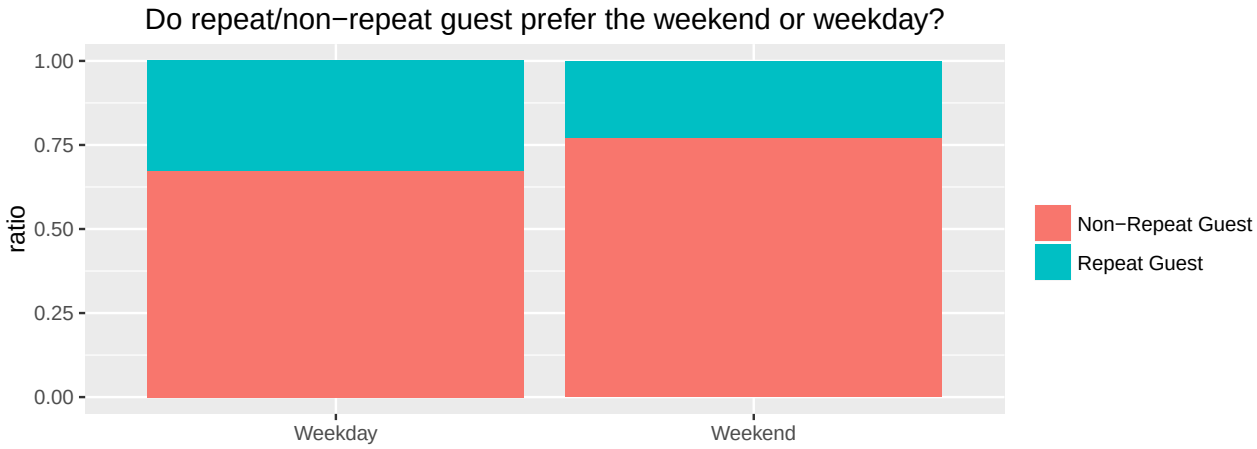
Repeat guest account for about 30% of the transactions yet only 24% of the revenue. Why do they spend less?

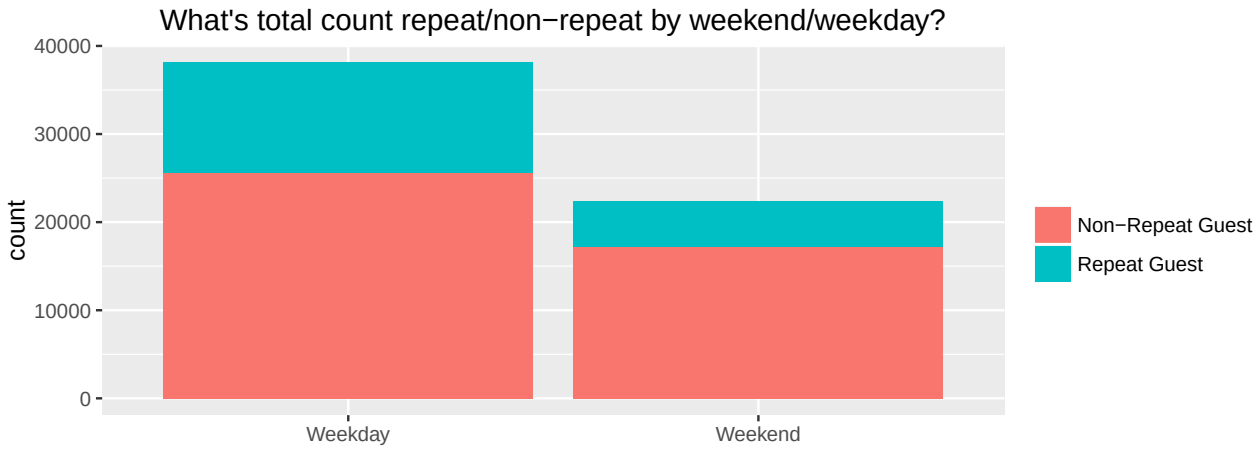
1.4 Visit Time vs Guest Retention

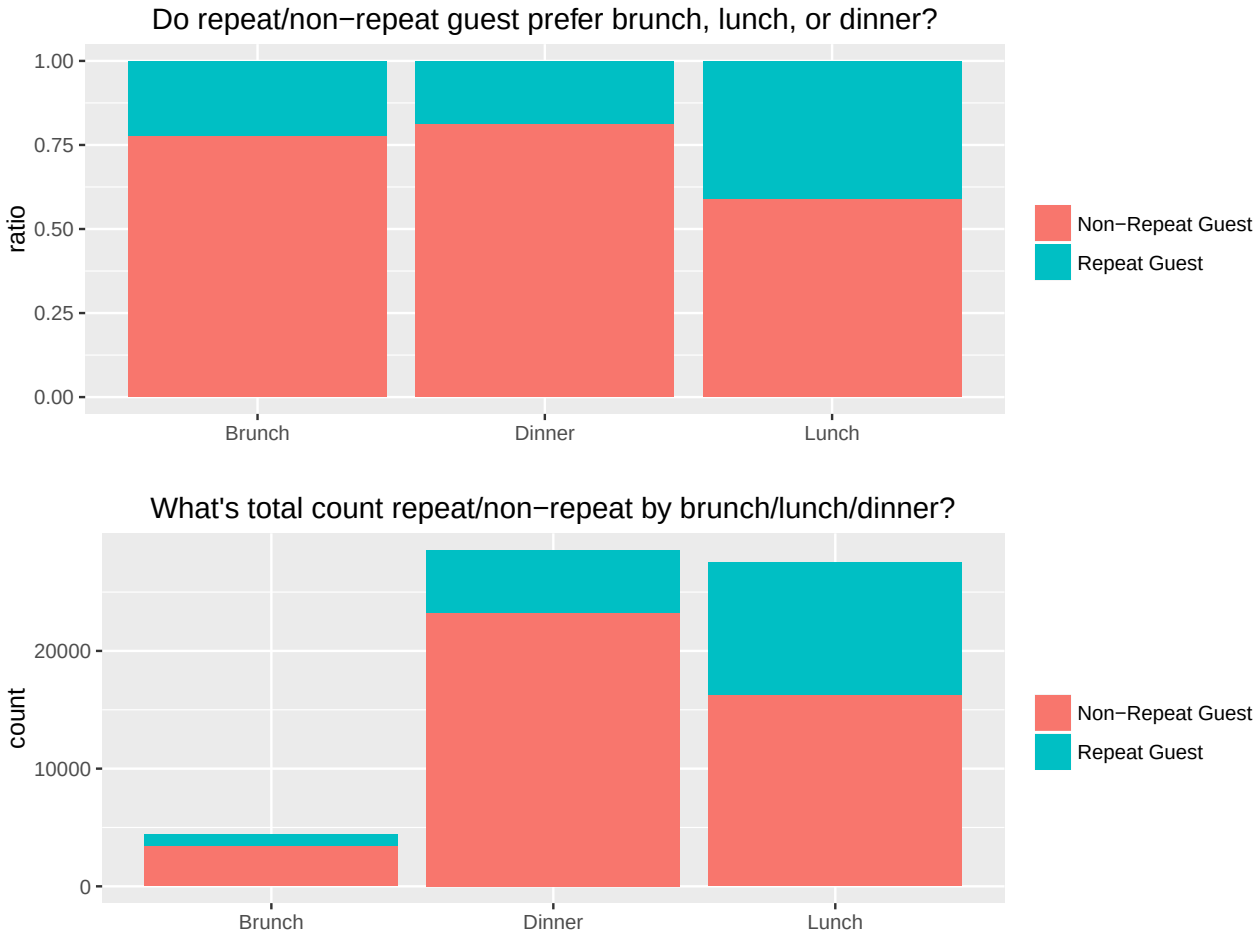
Holidays are defined as:

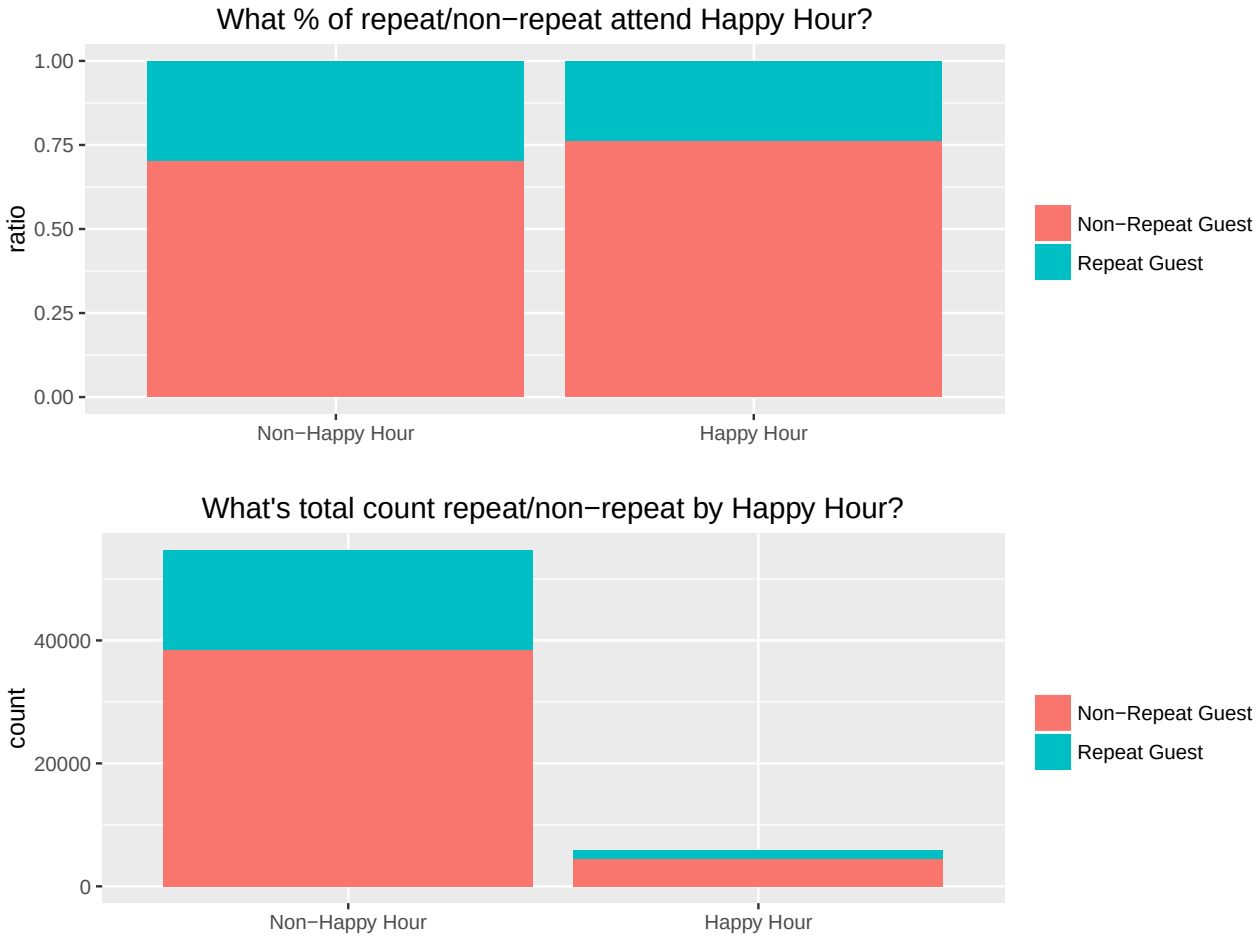
- Christmas 24th, 25th, and 26th
- Mother's Day 2nd Saturday and Sunday in May
- New Year's December 31st and January 1st

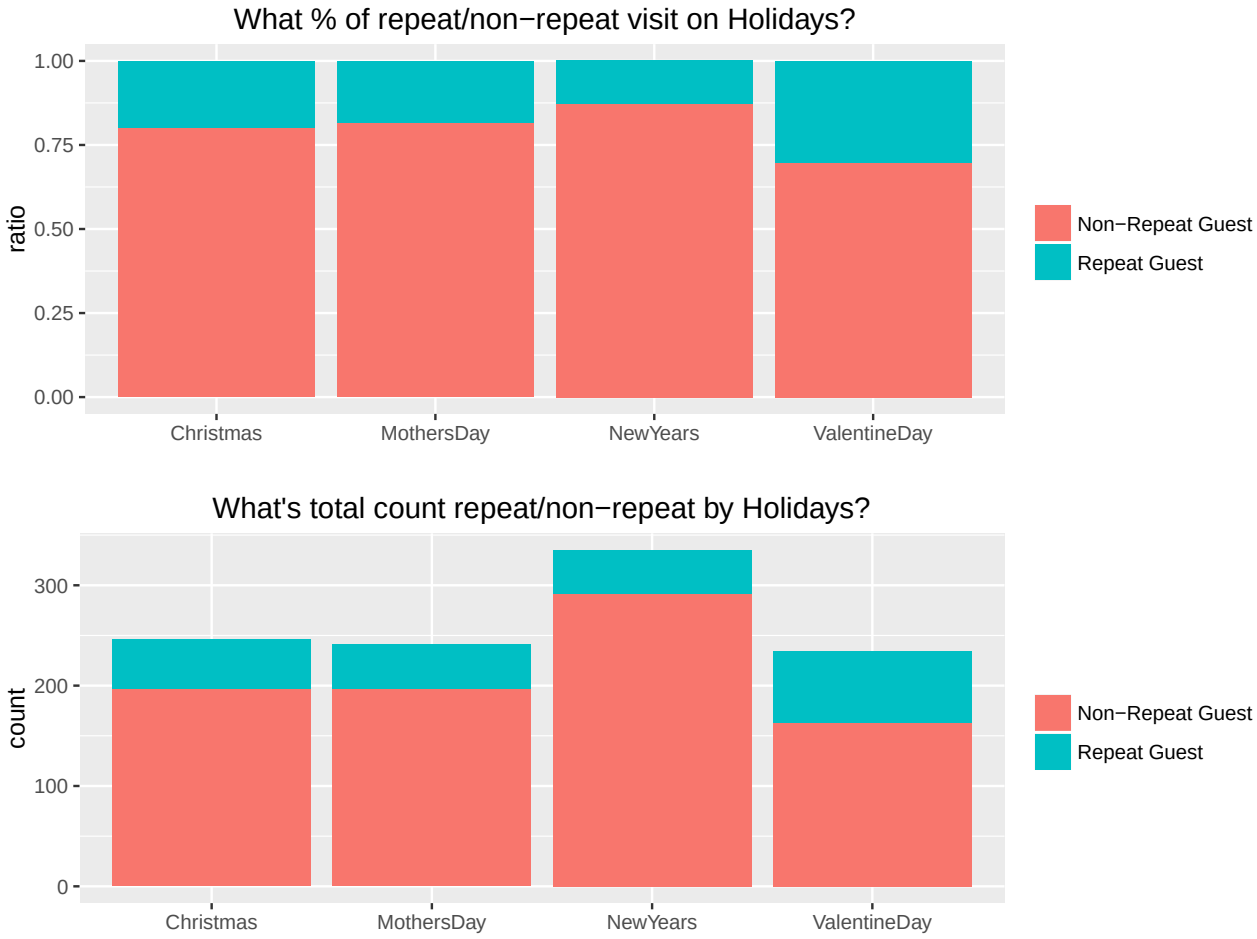


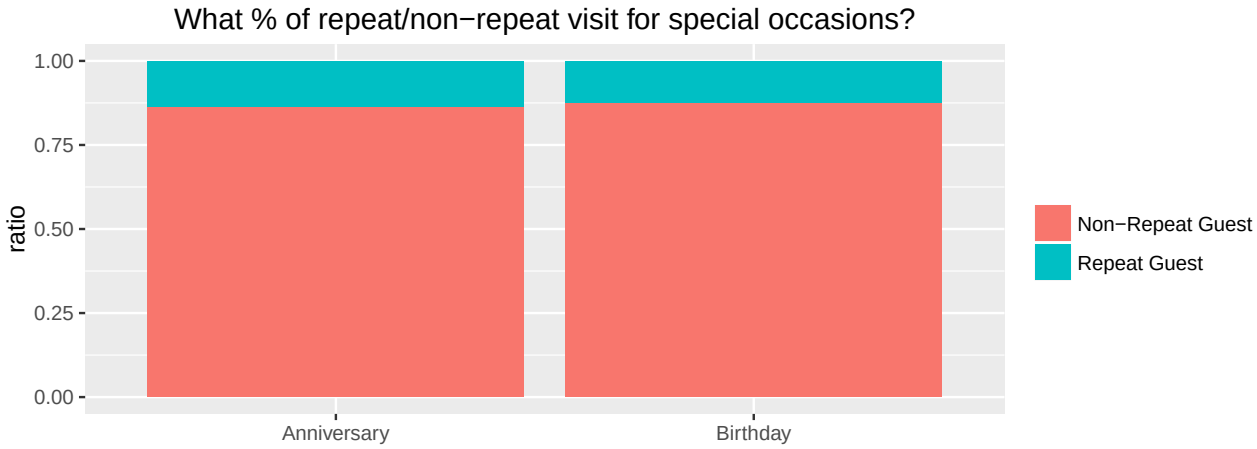

Do repeat/non−repeat guest prefer the weekend or weekday?
ratio
1.00
0.75
0.50
0.25
0.00
Weekday
Weekend
Non−Repeat Guest
Repeat Guest

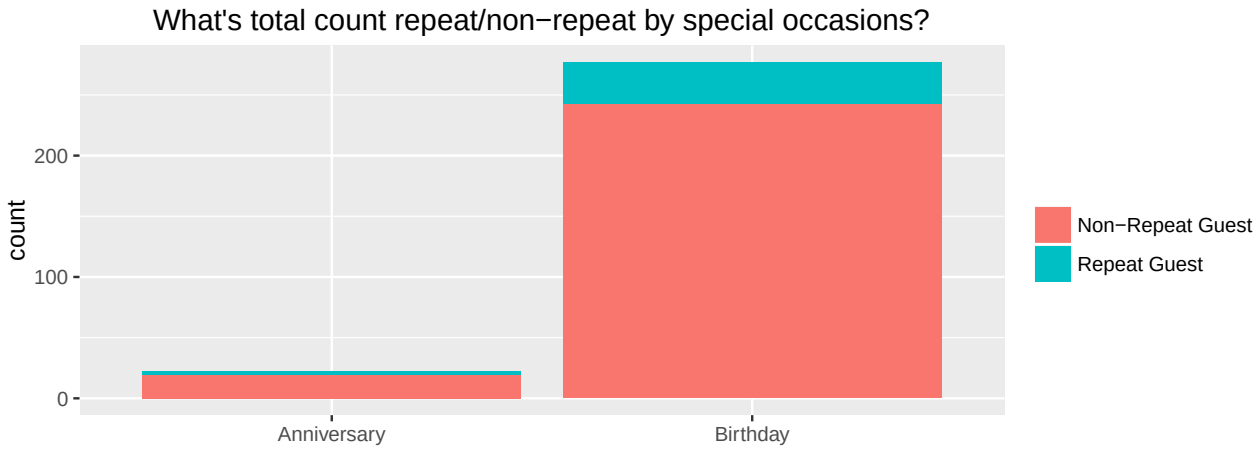

What's total count repeat/non−repeat by weekend/weekday?
count
40000
30000
20000
10000
0
Weekday
Weekend
Non−Repeat Guest
Repeat Guest


Do repeat/non−repeat guest prefer brunch, lunch, or dinner?
ratio
1.00
0.75
0.50
0.25
0.00
Brunch
Dinner
Lunch
Non−Repeat Guest
Repeat Guest
What's total count repeat/non−repeat by brunch/lunch/dinner?
count
20000
10000
0
Brunch
Dinner
Lunch
Non−Repeat Guest
Repeat Guest


What % of repeat/non−repeat attend Happy Hour?
ratio
1.00
0.75
0.50
0.25
0.00
Non−Happy Hour
Happy Hour
Non−Repeat Guest
Repeat Guest
What's total count repeat/non−repeat by Happy Hour?
count
40000
20000
0
Non−Happy Hour
Happy Hour
Non−Repeat Guest
Repeat Guest

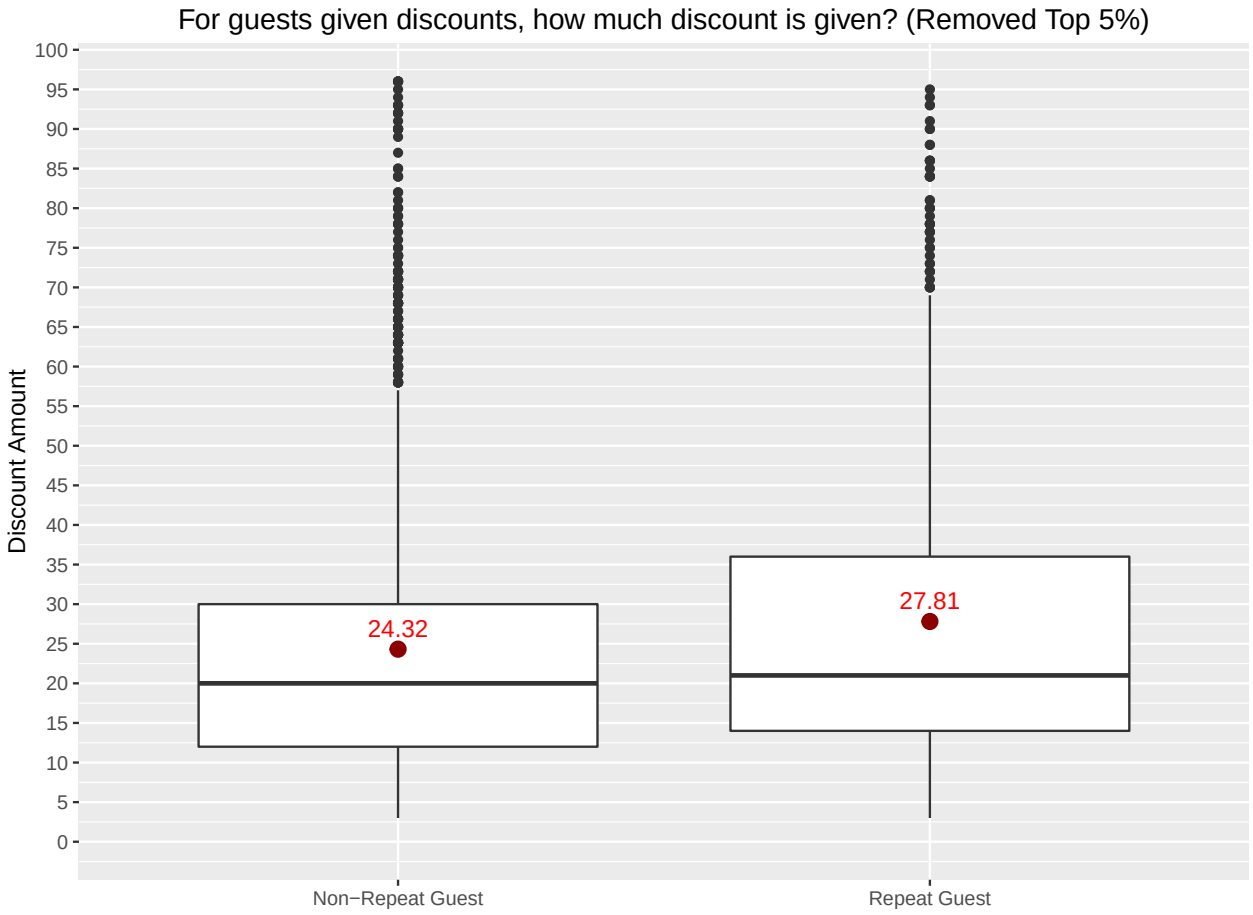

What % of repeat/non−repeat visit on Holidays?
ratio
1.00
0.75
0.50
0.25
0.00
Christmas
MothersDay
NewYears
ValentineDay
Non−Repeat Guest
Repeat Guest
What's total count repeat/non−repeat by Holidays?
count
300
200
100
0
Christmas
MothersDay
NewYears
ValentineDay
Non−Repeat Guest
Repeat Guest

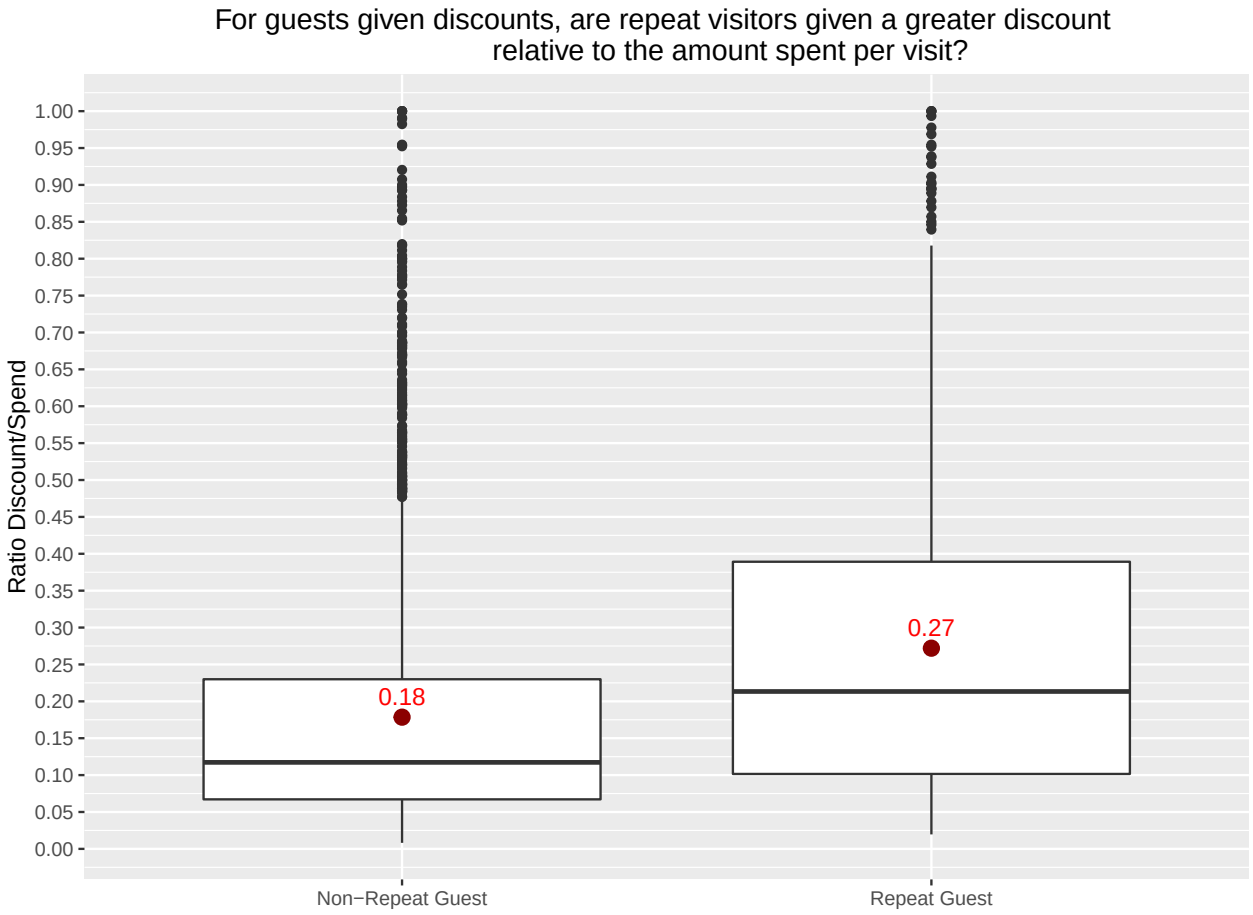




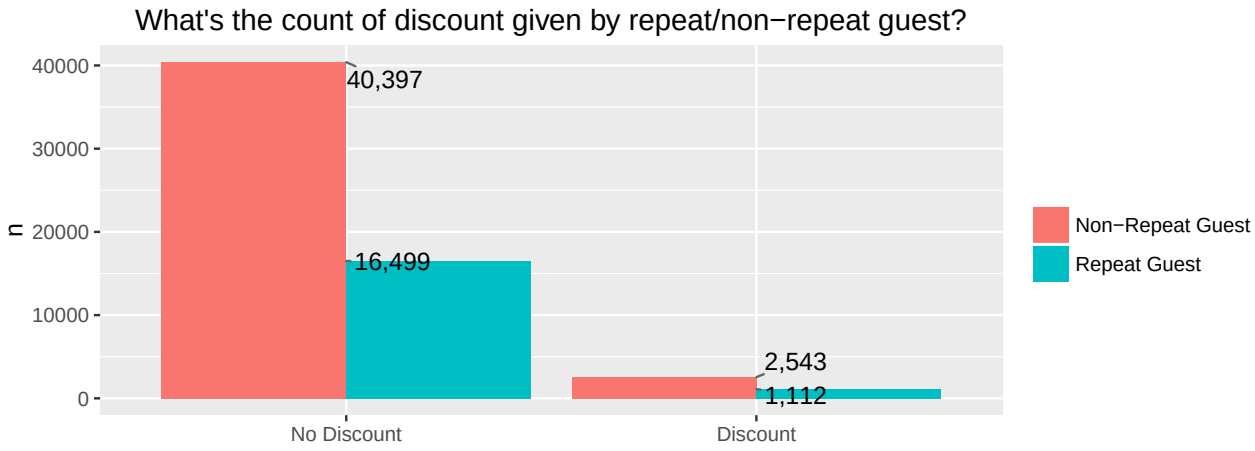
Repeat guest have a preference for weekdays and lunch.

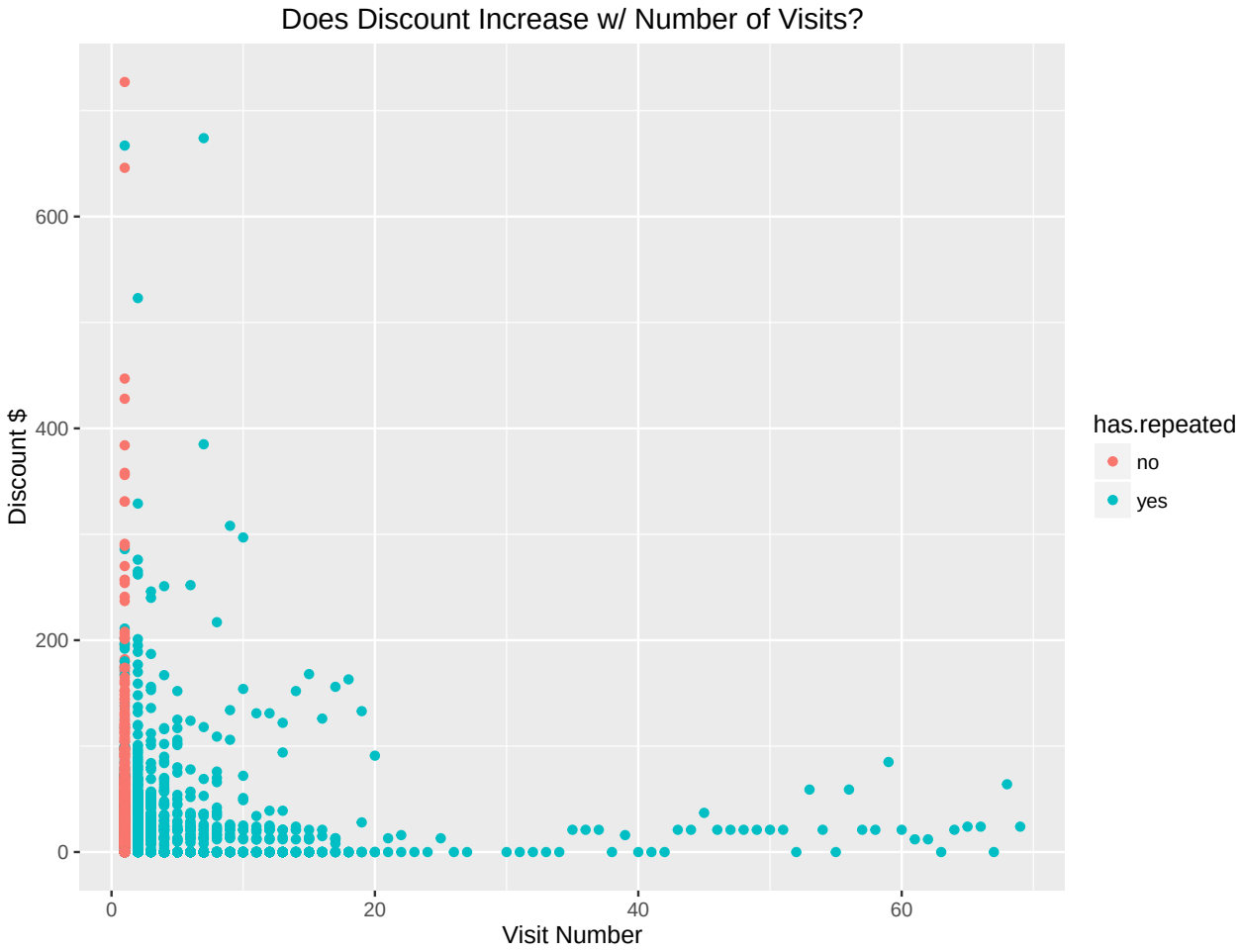
1.5 Discount


For guests given discounts, how much discount is given? (Removed Top 5%)
Discount Amount
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
24.32
27.81
Non−Repeat Guest
Repeat Guest


For guests given discounts, are repeat visitors given a greater discount relative to the amount spent per visit?
Ratio Discount/Spend
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.18
0.27
Non−Repeat Guest
Repeat Guest

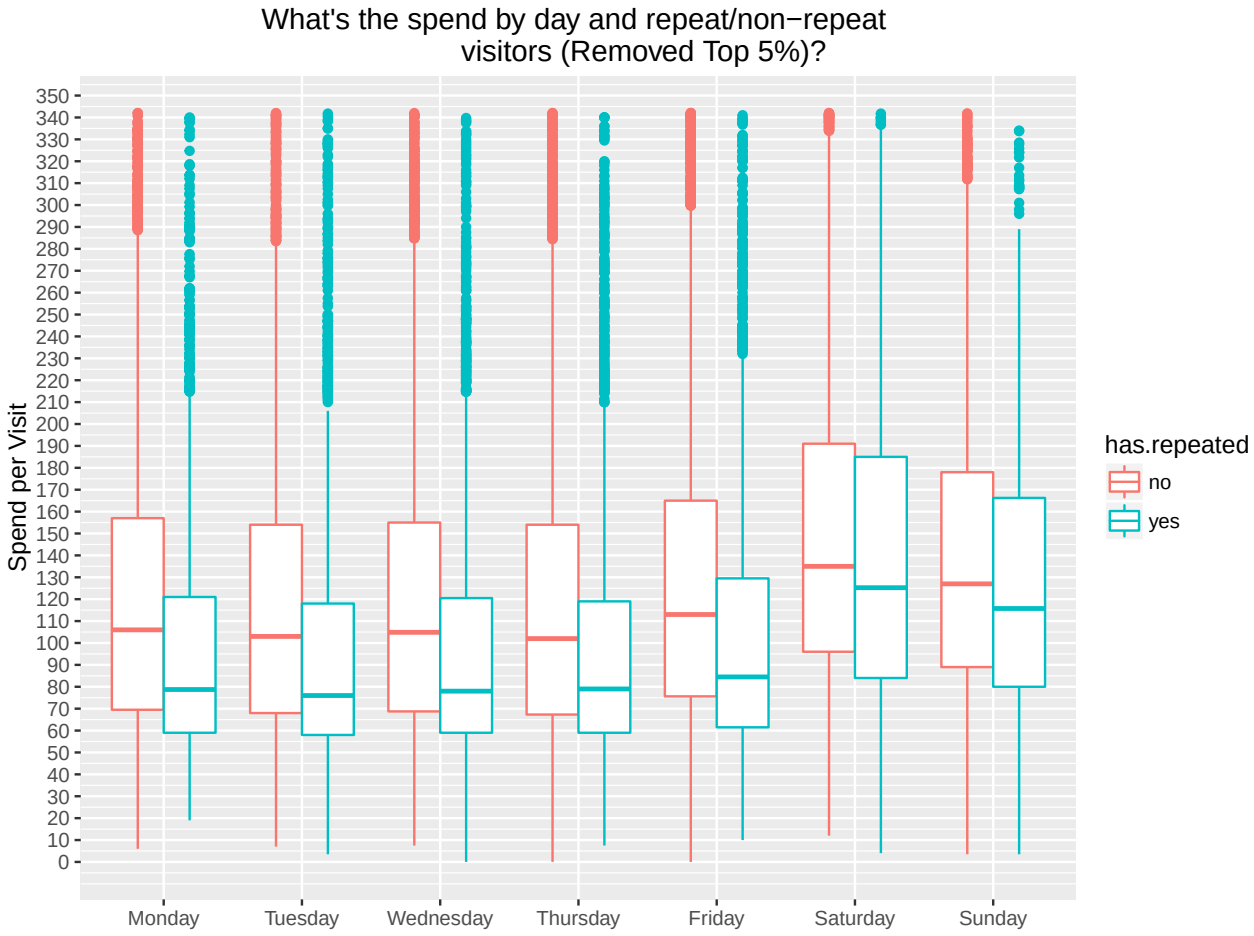

Are repeat guests given more discounts than 1st time visitors?
ratio
1.00
0.75
0.50
0.25
0.00
No Discount
Discount
Non−Repeat Guest
Repeat Guest

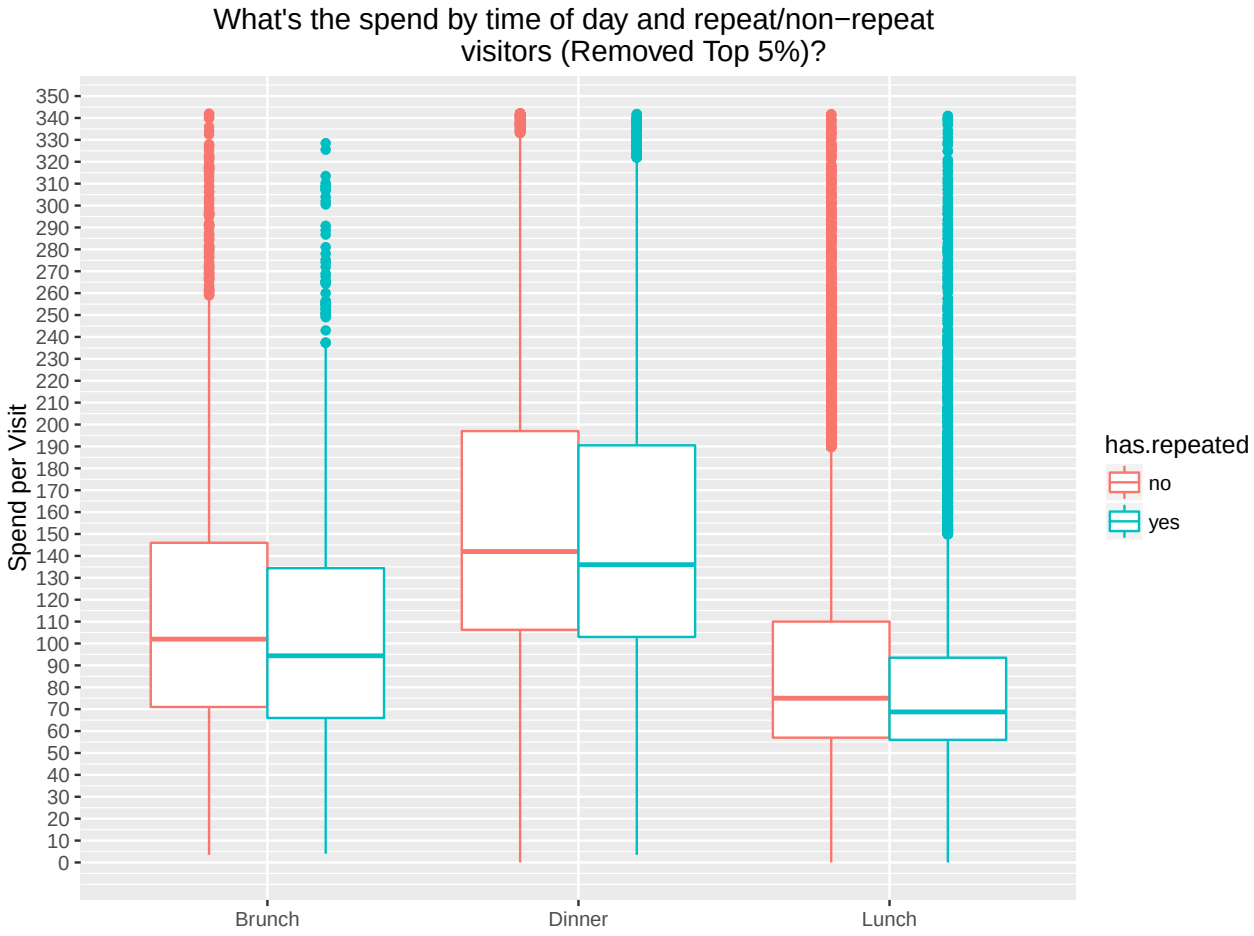

What's the count of discount given by repeat/non−repeat guest?
n
40000
30000
20000
10000
0
40,397
16,499
2,543
1,112
No Discount
Discount
Non−Repeat Guest
Repeat Guest

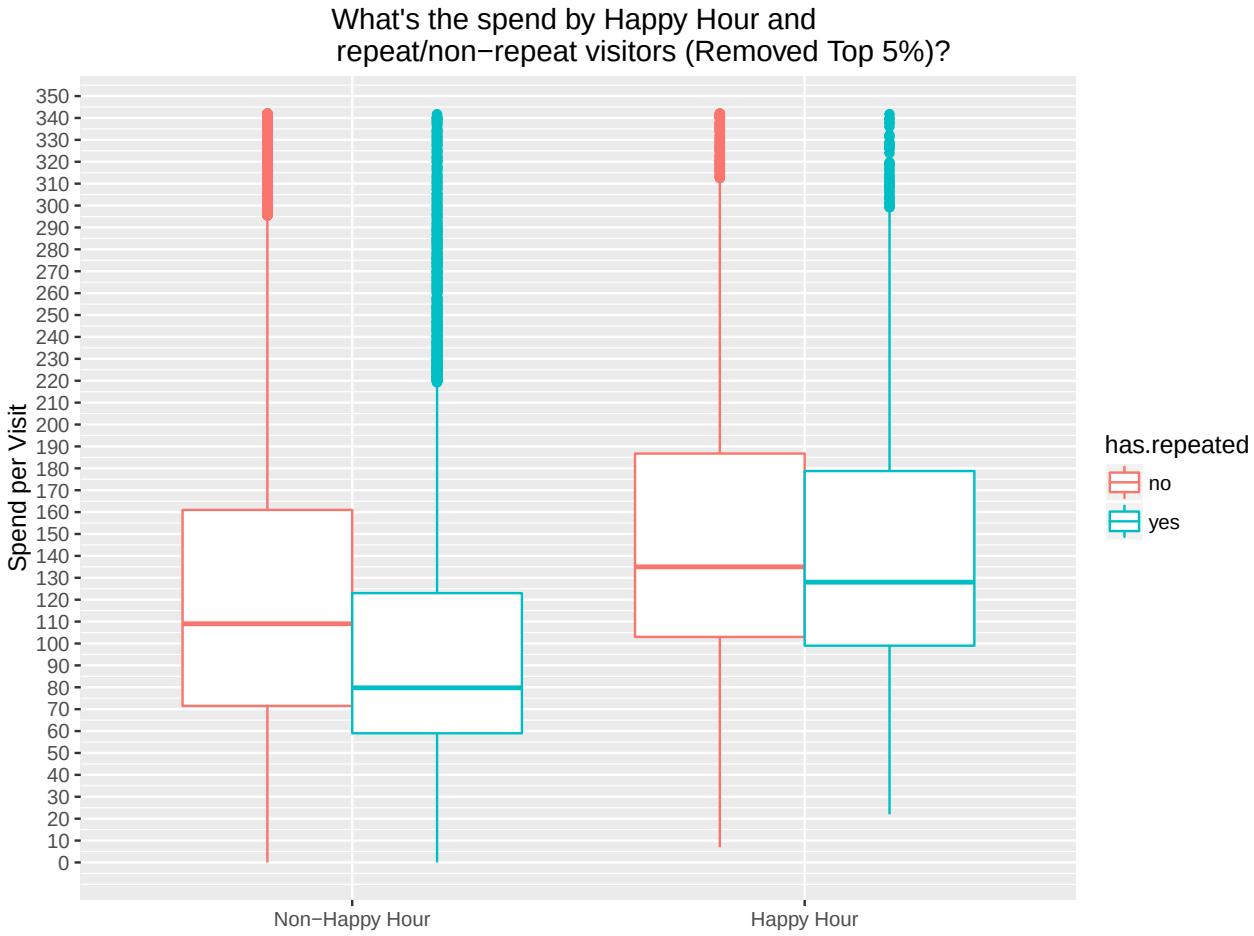

Does Discount Increase w/ Number of Visits?
Discount $
600
400
200
0
0
20
40
60
Visit Number
has.repeated
no
yes

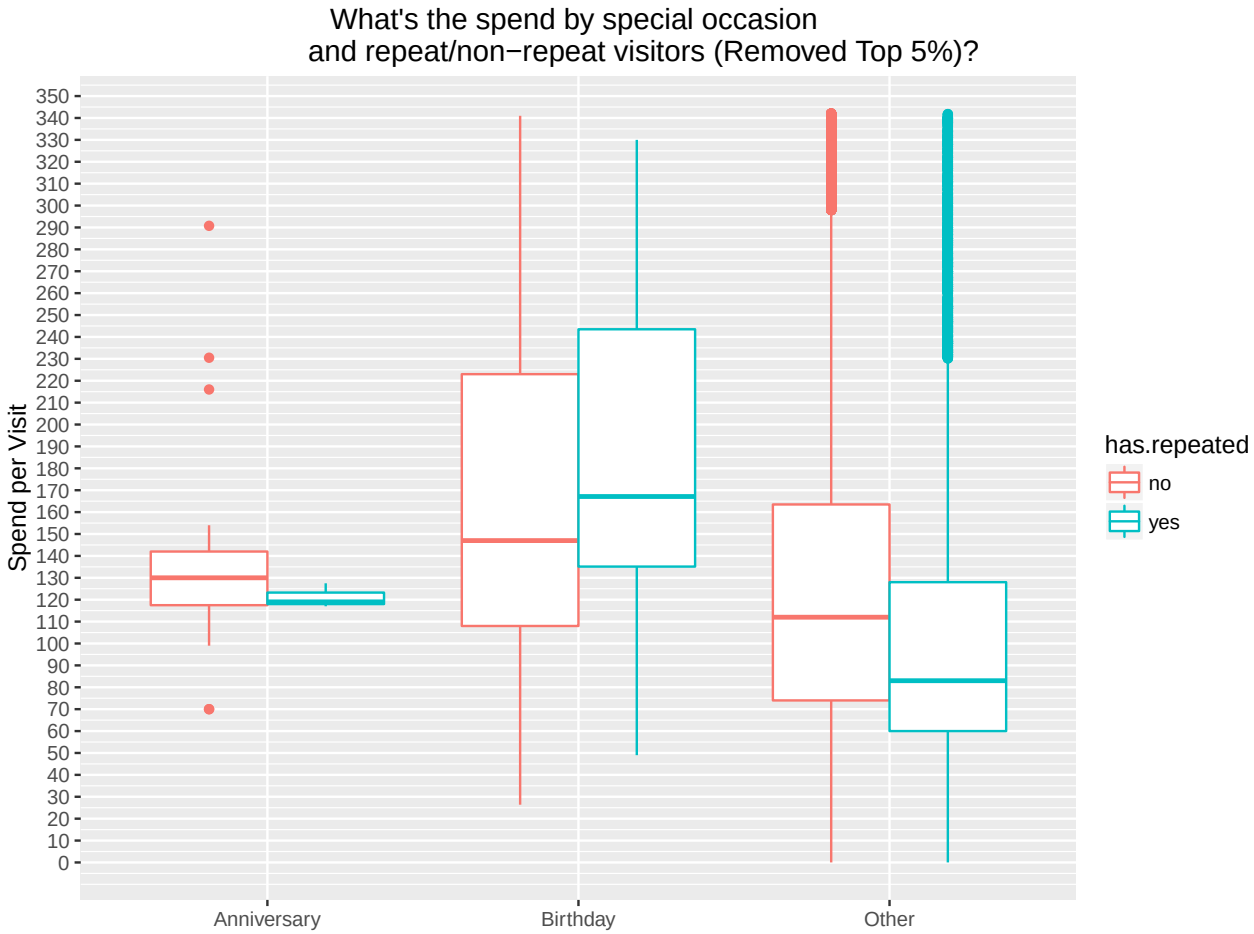
Both repeat and non-repeat guests are equally given discounts, but repeat guests are discounted at higher dollar amounts and higher percentage relative to their visit spend amount. 6% and 10% of all discount transactions for both groups are given even though the total visit spend amount was 0. Why is that? Is this how the POS system handles voids? Are these employee meals? These discounts account for about 13% and 30% of discount spending for non-repeat and repeat guest, respectively

1.6 Spend per Visit



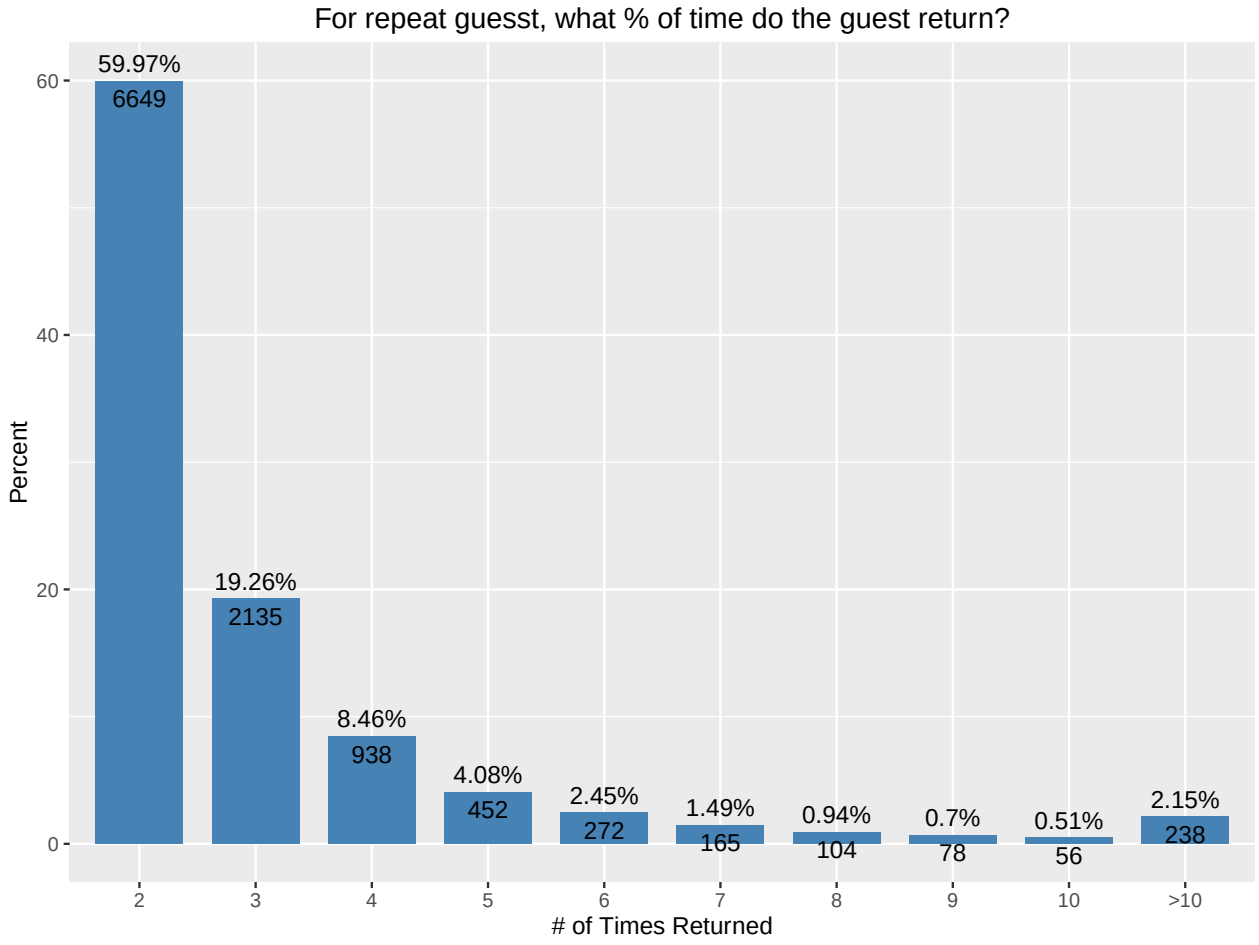

What's the spend by time of day and repeat/non−repeat visitors (Removed Top 5%)?
Spend per Visit
350
340
330
320
310
300
290
280
270
260
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Brunch
Dinner
Lunch
has.repeated
no
yes

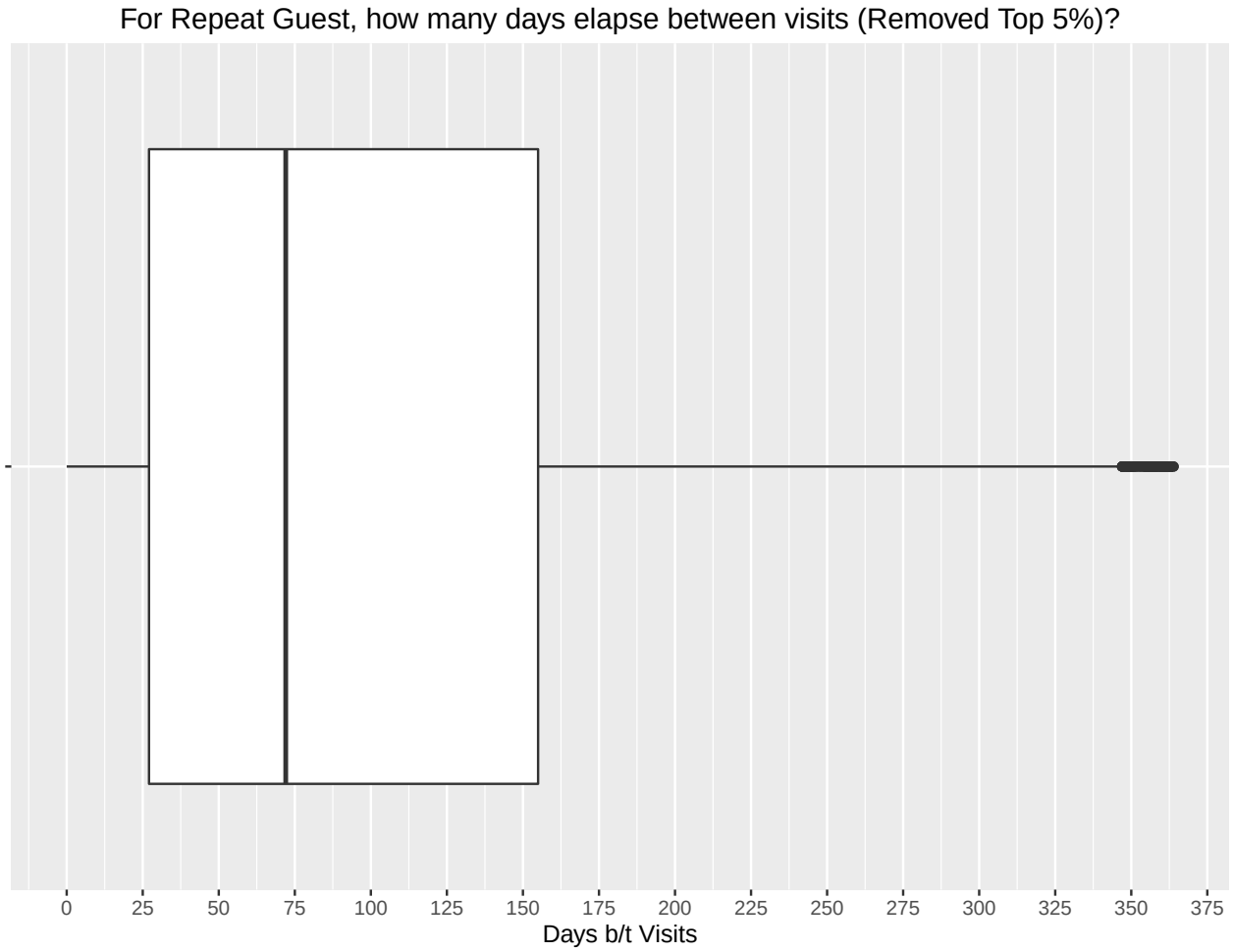

What's the spend by Happy Hour and
repeat/non−repeat visitors (Removed Top 5%)?
Spend per Visit
350
340
330
320
310
300
290
280
270
260
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Non−Happy Hour
Happy Hour
has.repeated
no
yes

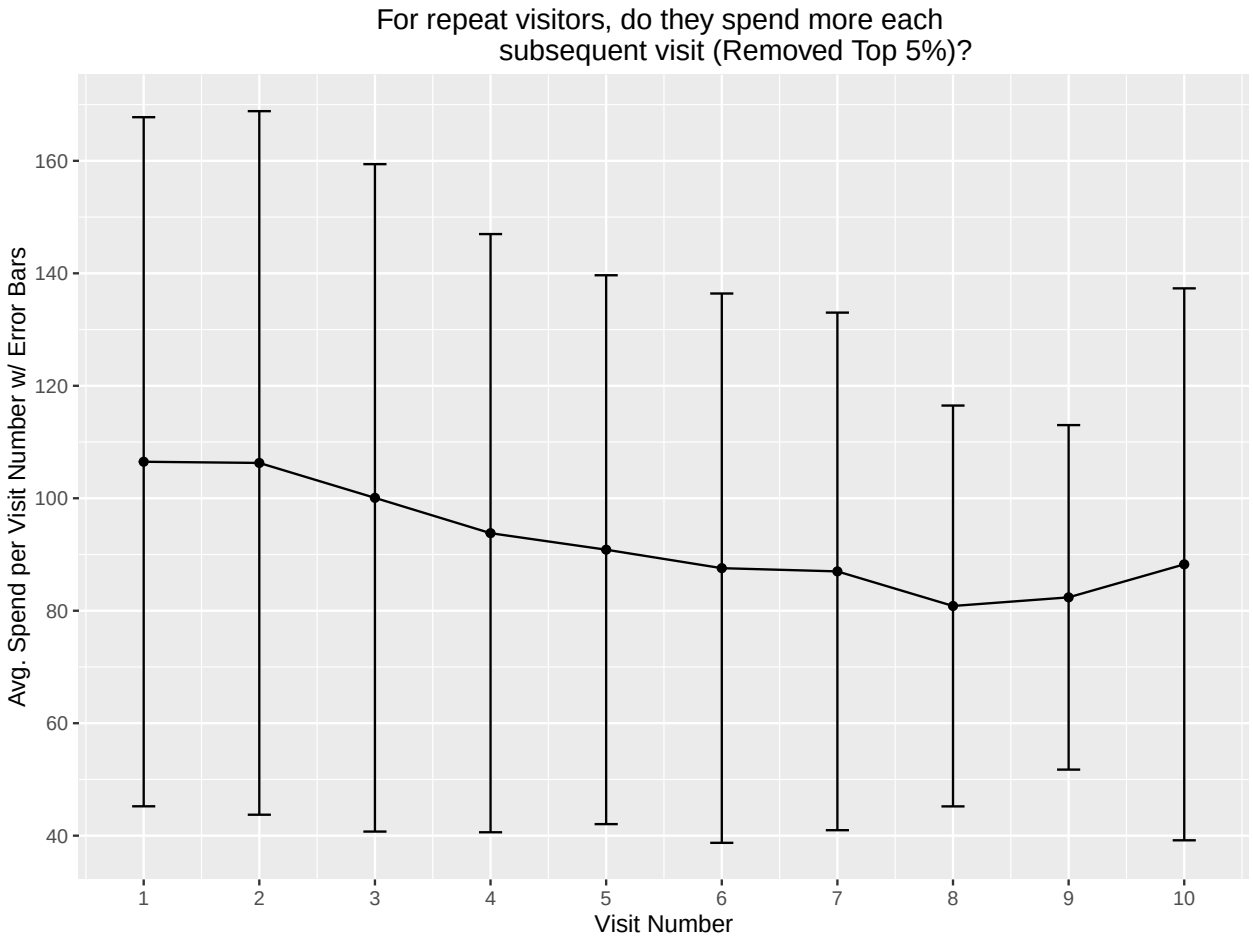


Repeat guest spend less per visit than new customers.

1.7 Repeat Guests

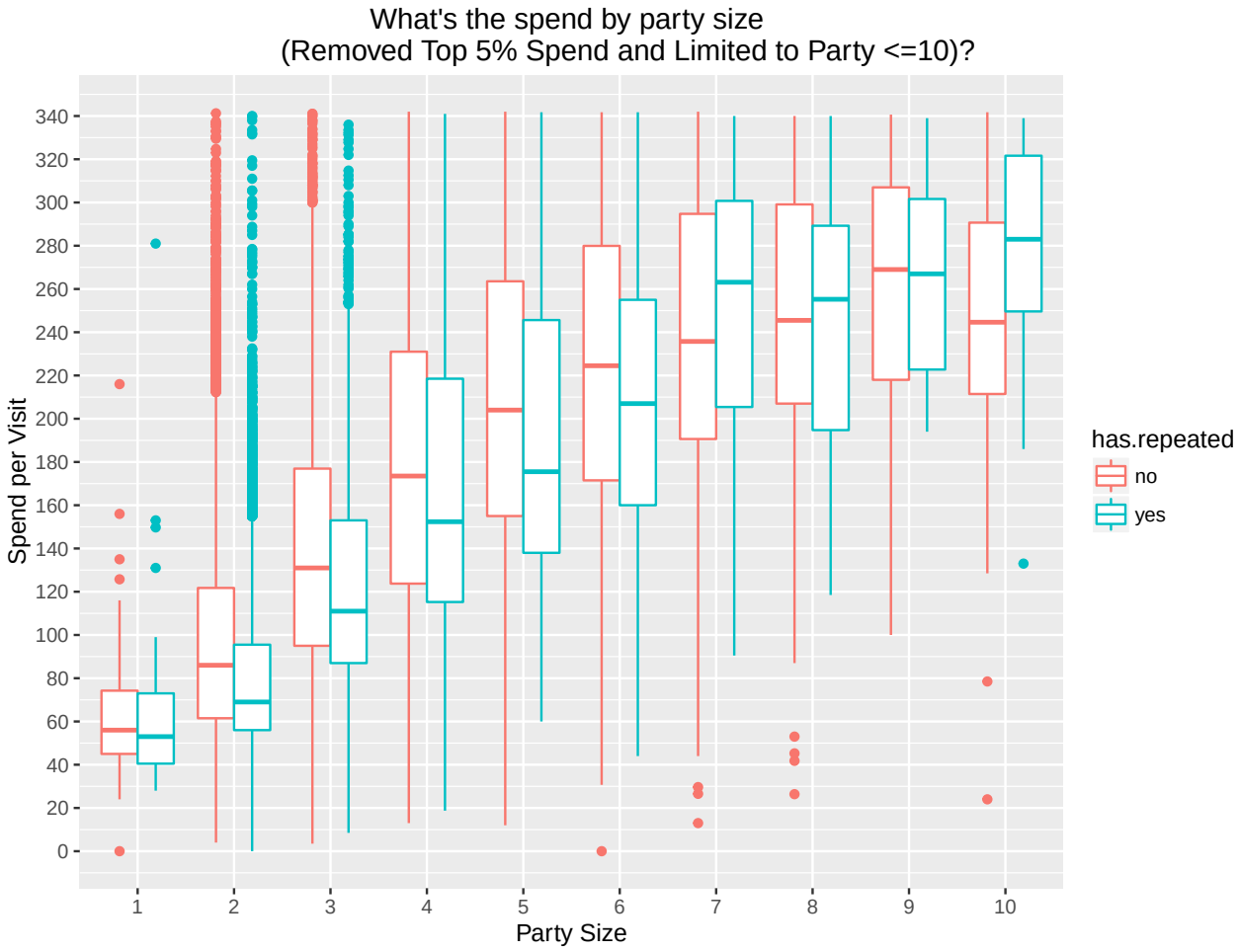


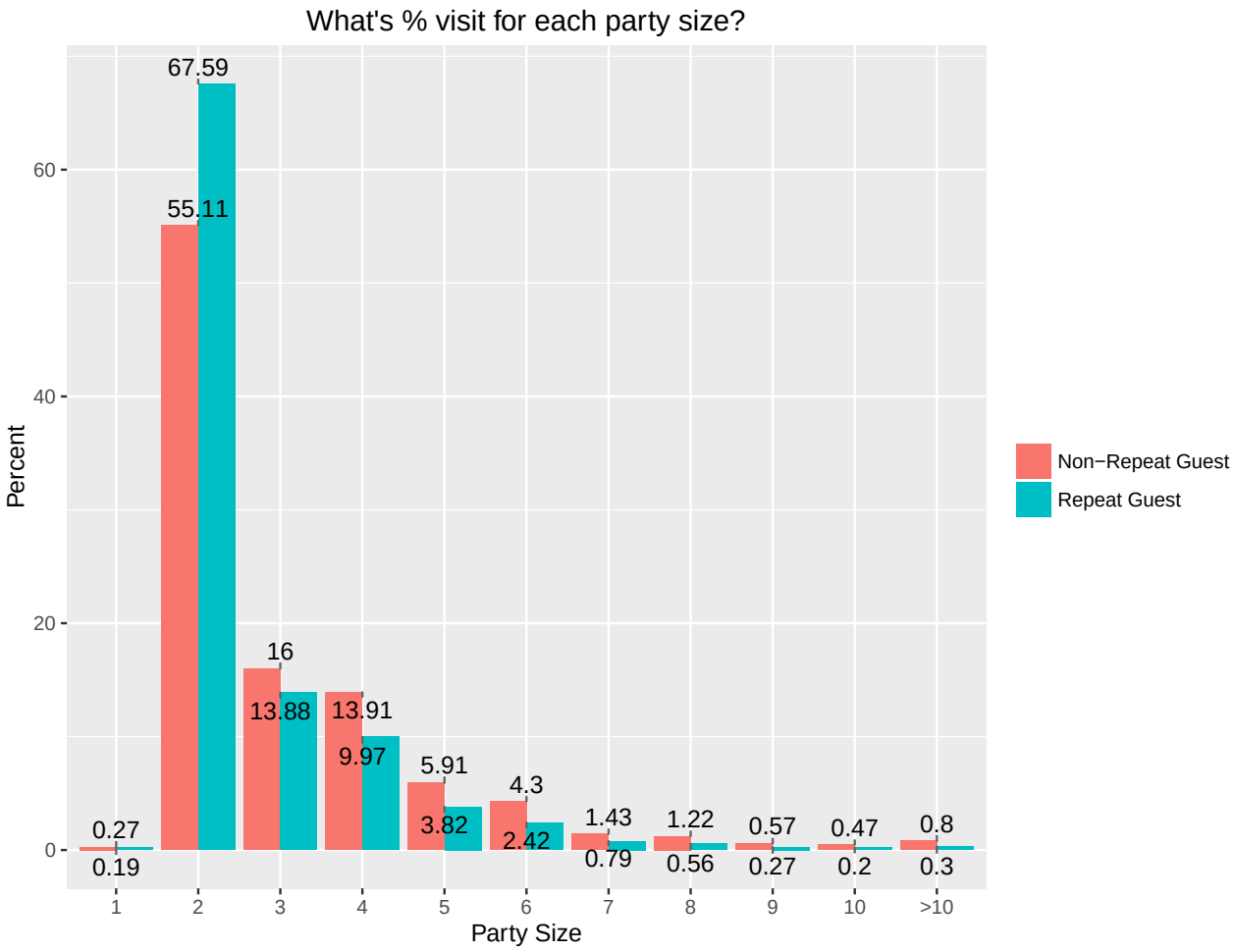

For Repeat Guest, how many days elapse between visits (Removed Top 5%)?
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
Days b/t Visits

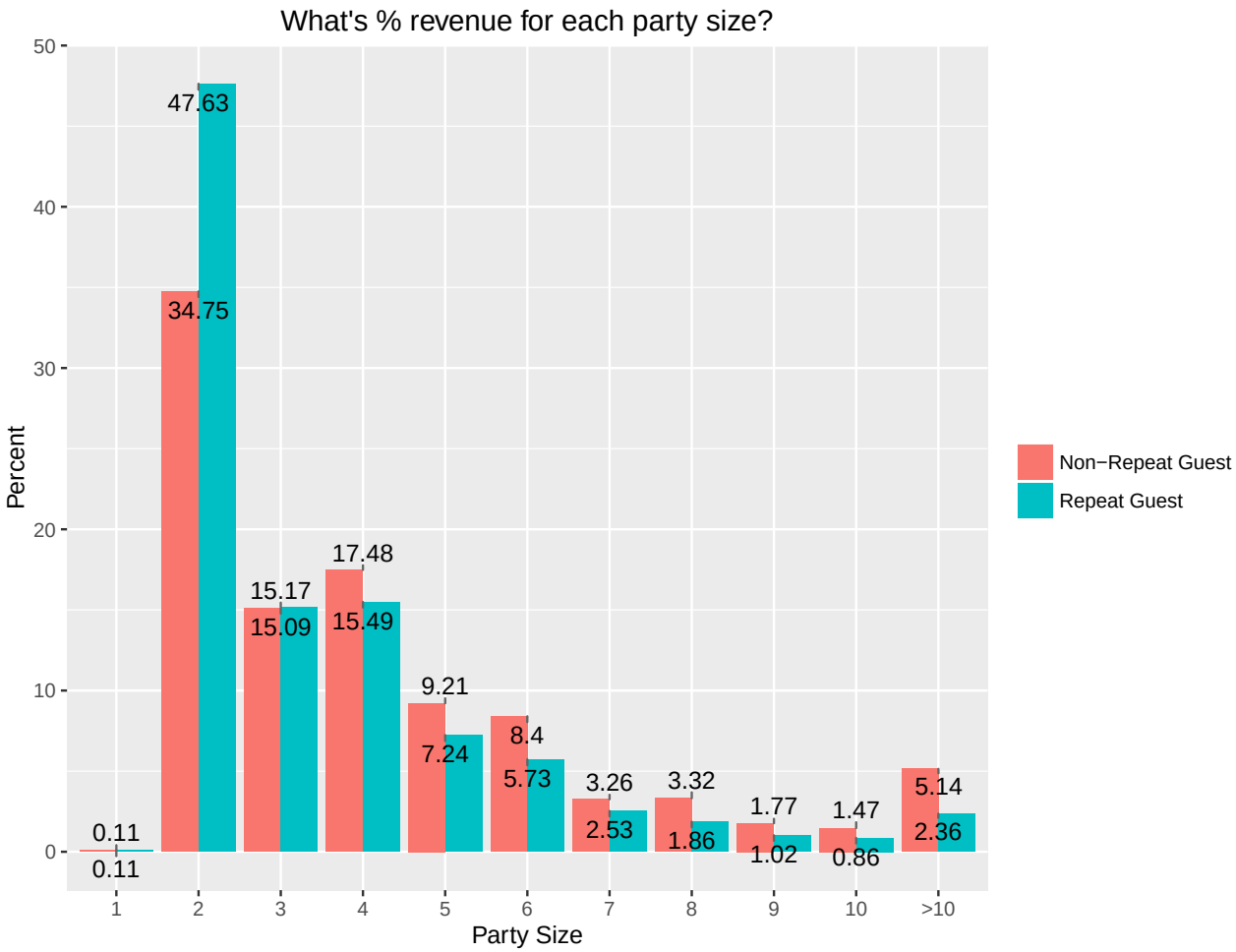


About 75% of the repeat guest only come for 2 or 3 visits. The median days between visits is about 70 days. Finally, on average, repeat guest tend to spend less the more they visit.

1.8 Party Size



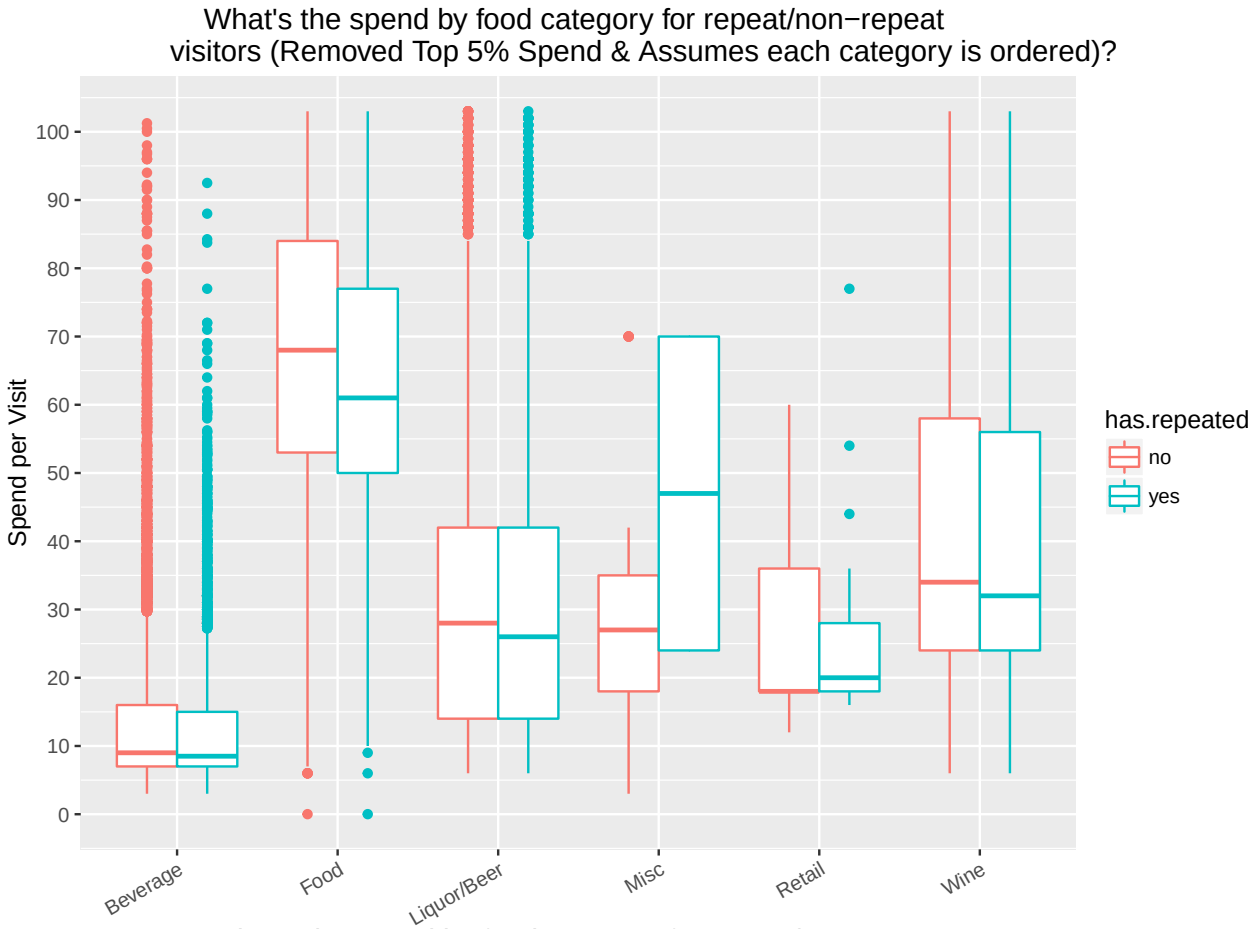

What's % visit for each party size?
Percent
Party Size
Non−Repeat Guest
Repeat Guest
0
20
40
60
1
2
3
4
5
6
7
8
9
10
>10
0.27
0.19
67.59
55.11
16
13.88
13.91
9.97
5.91
3.82
4.3
2.42
1.43
0.79
1.22
0.56
0.57
0.27
0.47
0.2
0.8
0.3



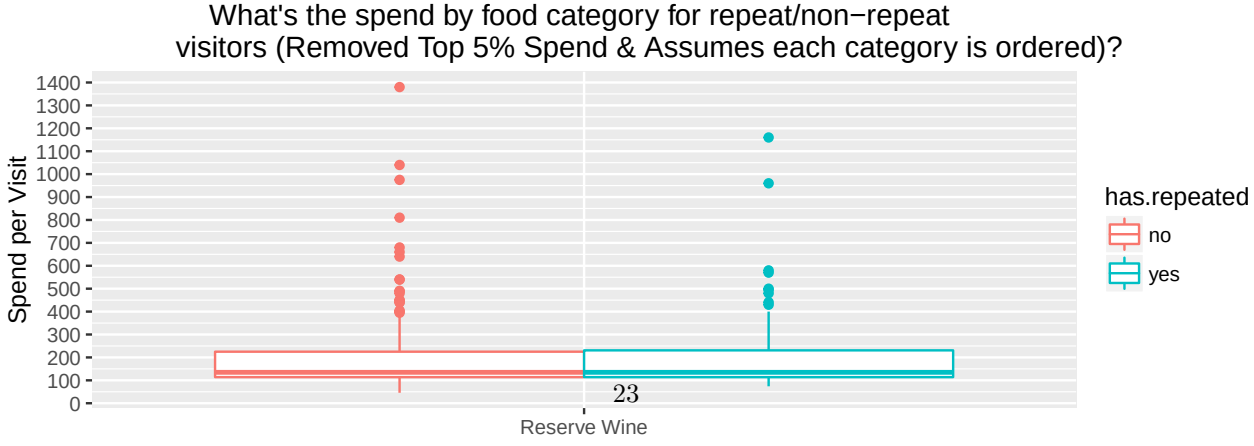
As expected, larger parties do spend more, but again repeat guest spend less regardless of the party size. For both groups, guests tend to come in parties of 2; for repeat guest, it's about 70% of the time. Finally, large parties only account for 0.8% of all 1st guest visits, they generate about 5% of the revenue.

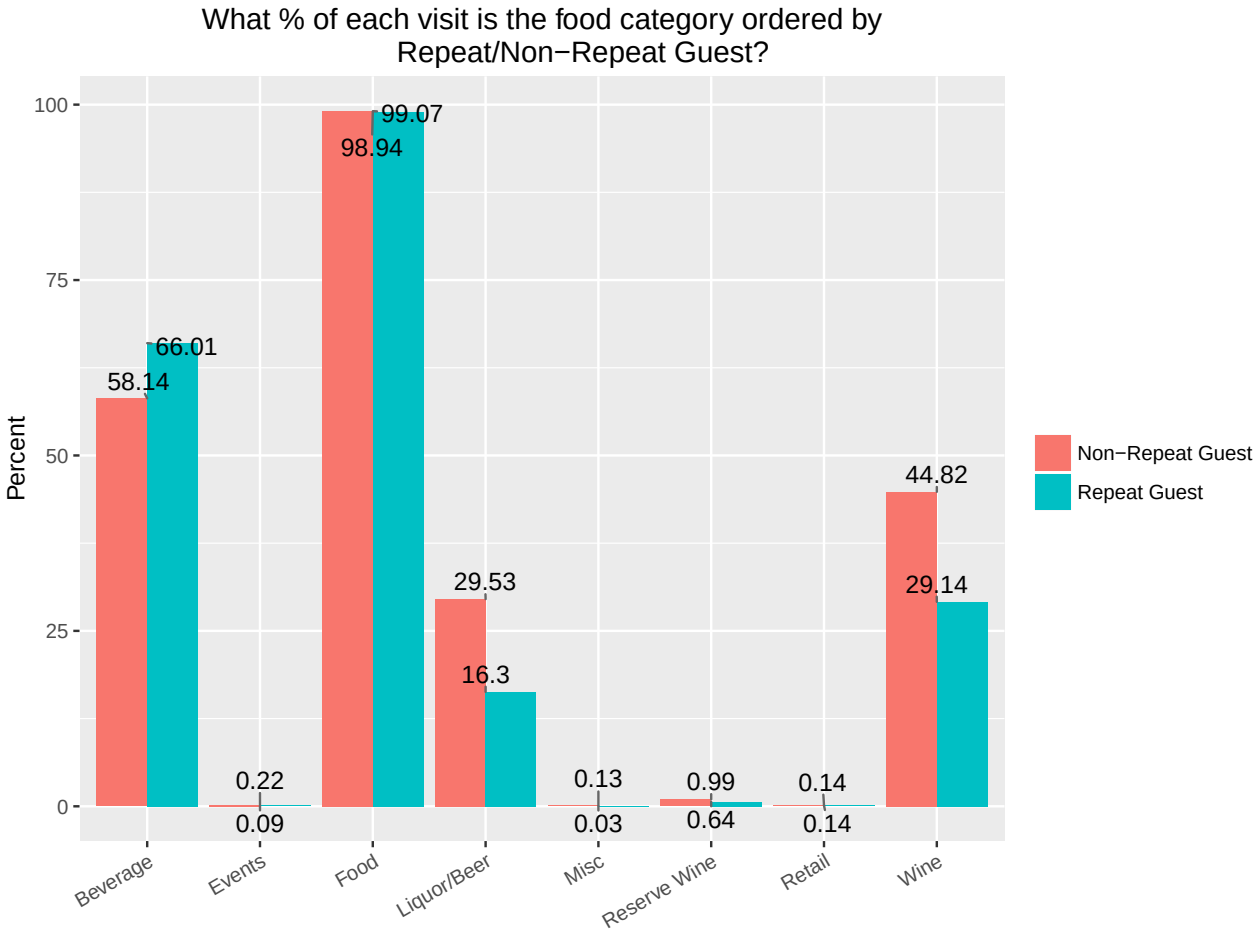
1.9 Food Category

The food category were derived from the POS item category name feature.


What's the spend by food category for repeat/non−repeat visitors (Removed Top 5% Spend & Assumes each category is ordered)?
Spend per Visit
0
10
20
30
40
50
60
70
80
90
100
Beverage
Food
Liquor/Beer
Misc
Retail
Wine
has.repeated
no
yes


What's the spend by food category for repeat/non−repeat visitors (Assumes each category is ordered)?
Spend per Visit
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
Events
has.repeated
no
yes


What's the spend by food category for repeat/non−repeat visitors (Removed Top 5% Spend & Assumes each category is ordered)?
Spend per Visit
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
Reserve Wine
has.repeated
no
yes

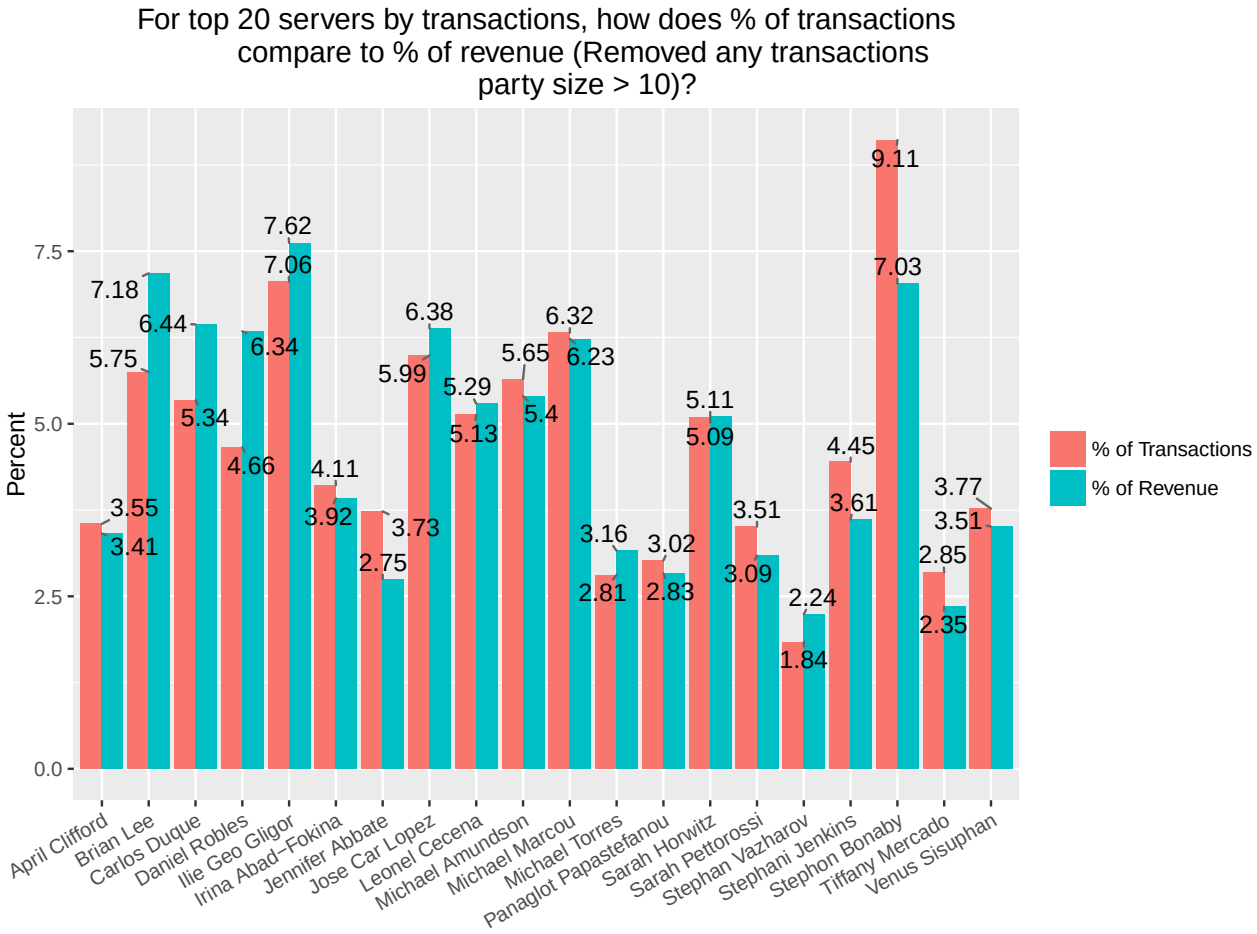


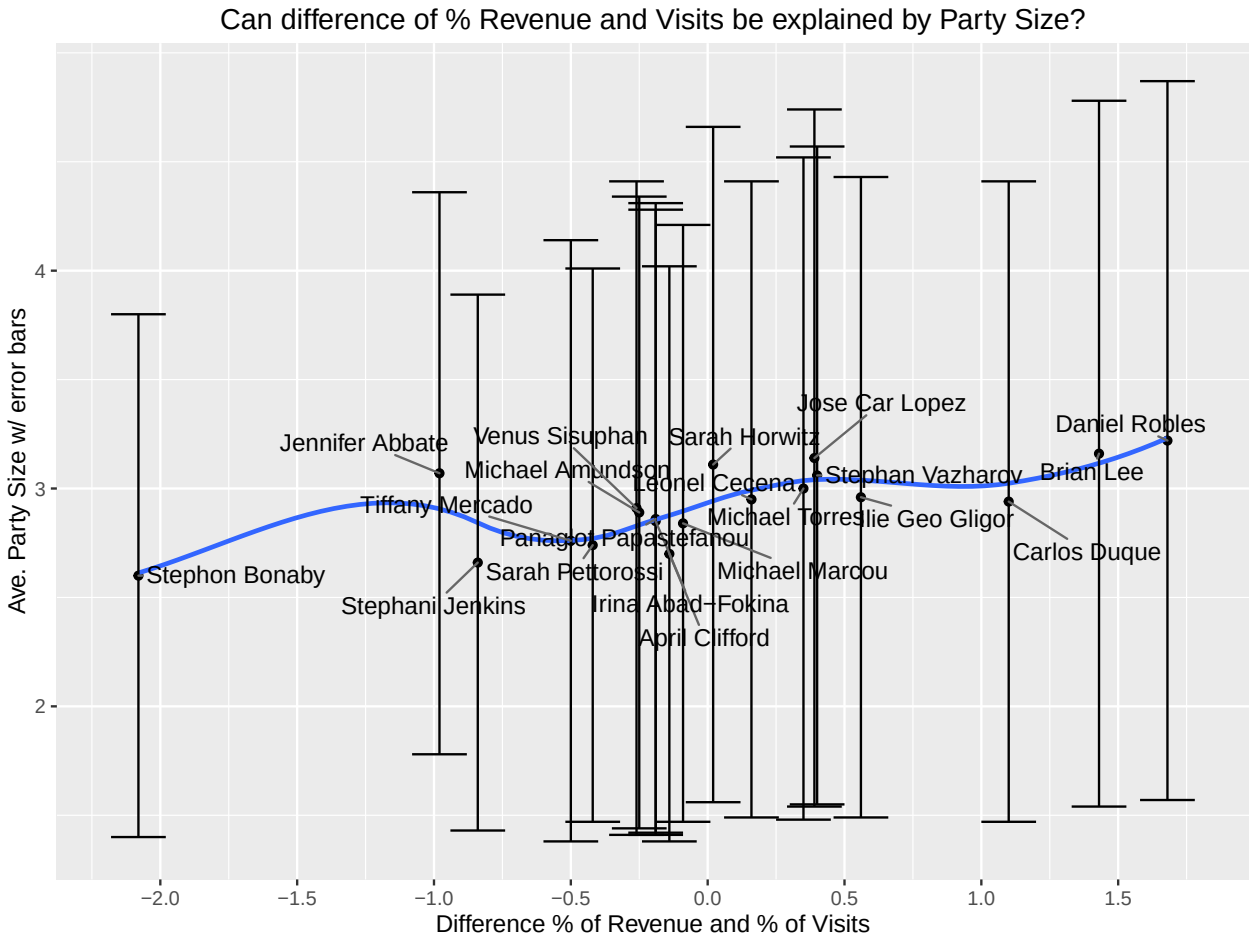
The same pattern held with the food categories: repeat guest spending less for the majority of the categories. The Misc category is generally comprised of cork fees. Repeat guests also are less likely to order beer, liquor, or wine on a visit; this likely because many are coming for lunch.

1.10 Servers

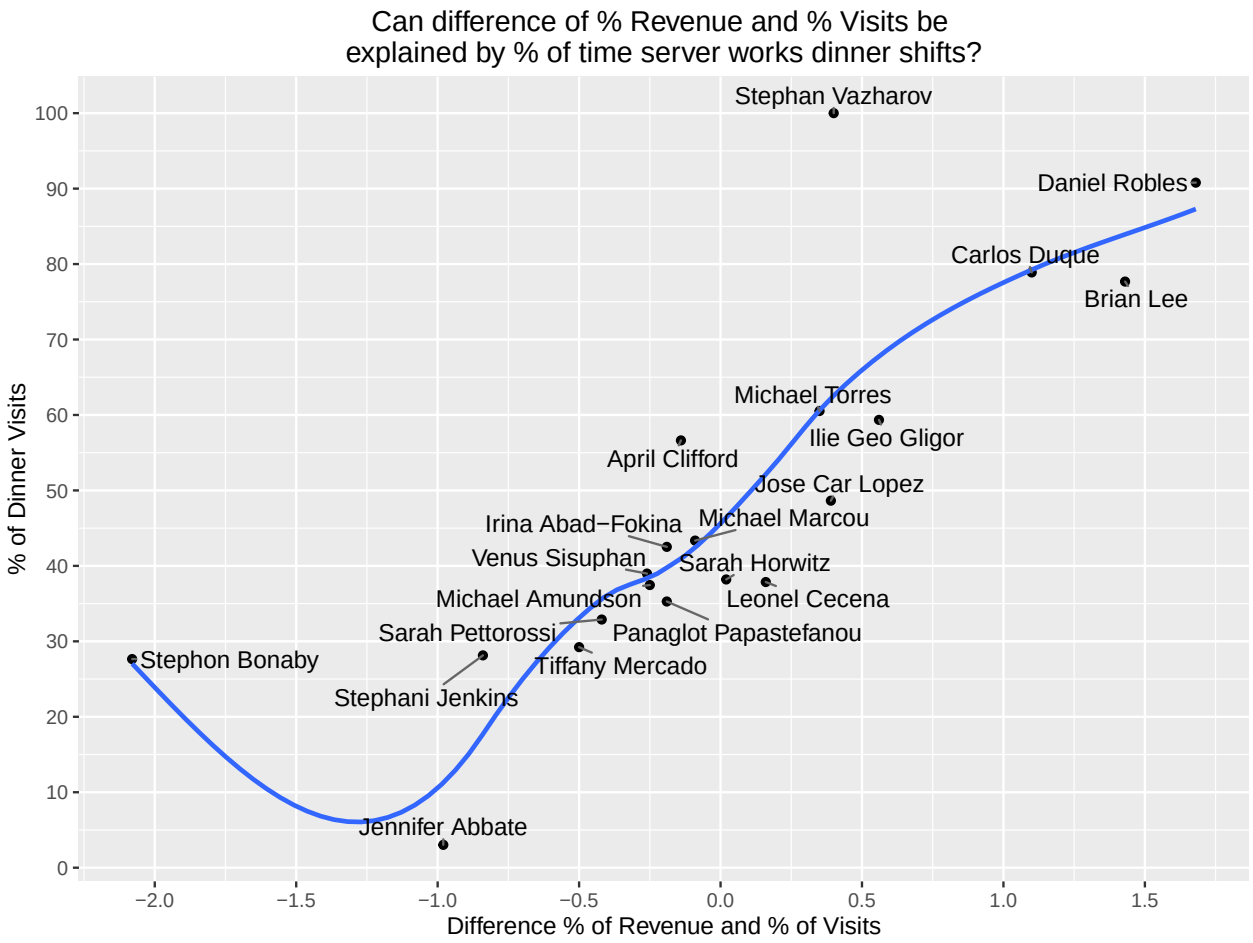
To analyze servers and tables, I removed any transactions with a party size greater than 10. Party size larger than 10 tend to be events, thus it may skewed the results. The sample size was still robust; party_size greater than 10 only accounts for 7% of the transactions.

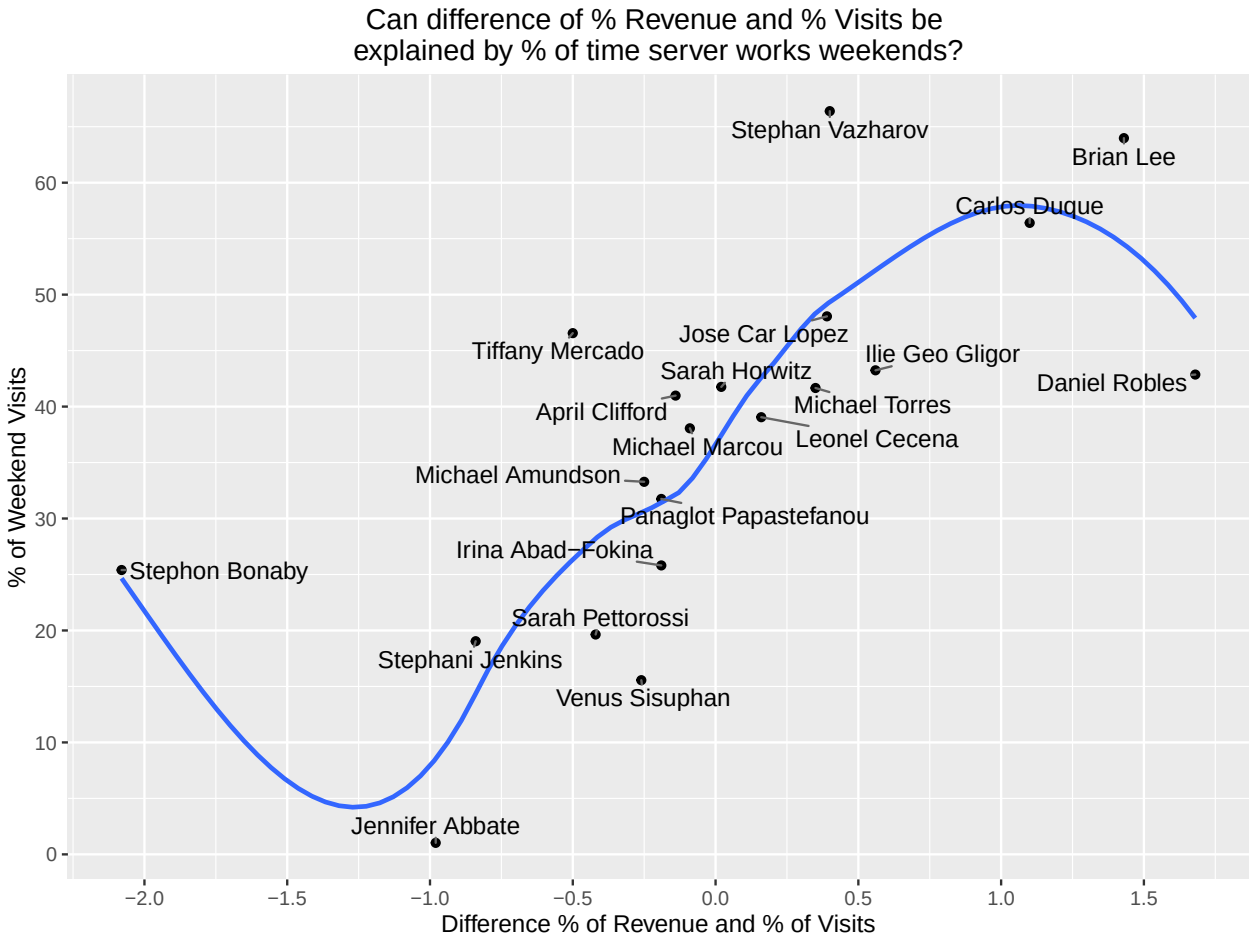
Graph below tries to answer whether difference in percent of revenue generated by server relative to percent of transactions can be explained by the average party size for the server; negative number indicates that the server generated less percentage of revenue relative to % of visits. For example, Stephon B accounted for 9.11% of the visits yet only accounted for 7.03% of the revenue, thus he's delta is -2.08%. The assumption is that over a period of time, a server should be generate revenue equal to his/her rate of transactions. Are there servers who generate more revenue relative to their number of transactions?


For top 20 servers by transactions, how does % of transactions compare to % of revenue (Removed any transactions party size > 10)?
Percent
0.0
2.5
5.0
7.5
% of Transactions
% of Revenue
April Clifford
Brian Lee
Carlos Duque
Daniel Robles
Ilie Geo Gligor
Irina Abad−Fokina
Jennifer Abbate
Jose Car Lopez
Leonel Cecena
Michael Amundson
Michael Marcou
Michael Torres
Panaglot Papastefanou
Sarah Horwitz
Sarah Pettorossi
Stephan Vazharov
Stephani Jenkins
Stephon Bonaby
Tiffany Mercado
Venus Sisuphan
3.55
3.41
5.75
7.18
6.44
5.34
4.66
6.34
7.62
7.06
4.11
3.92
2.75
3.73
5.99
6.38
5.29
5.13
5.65
5.4
6.32
6.23
2.81
3.16
2.83
3.02
5.11
5.09
3.09
3.51
1.84
2.24
4.45
3.61
9.11
7.03
2.85
2.35
3.77
3.51

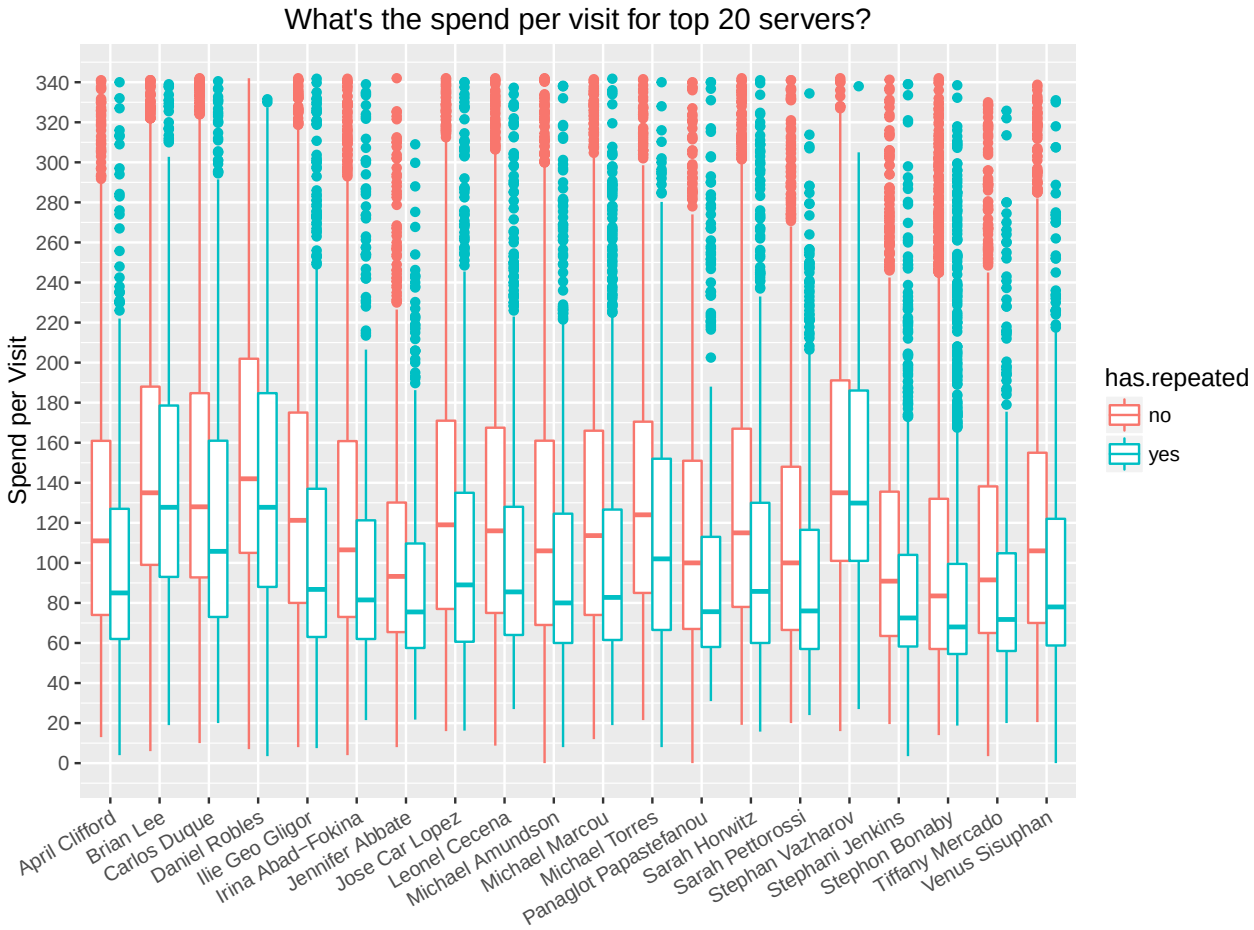


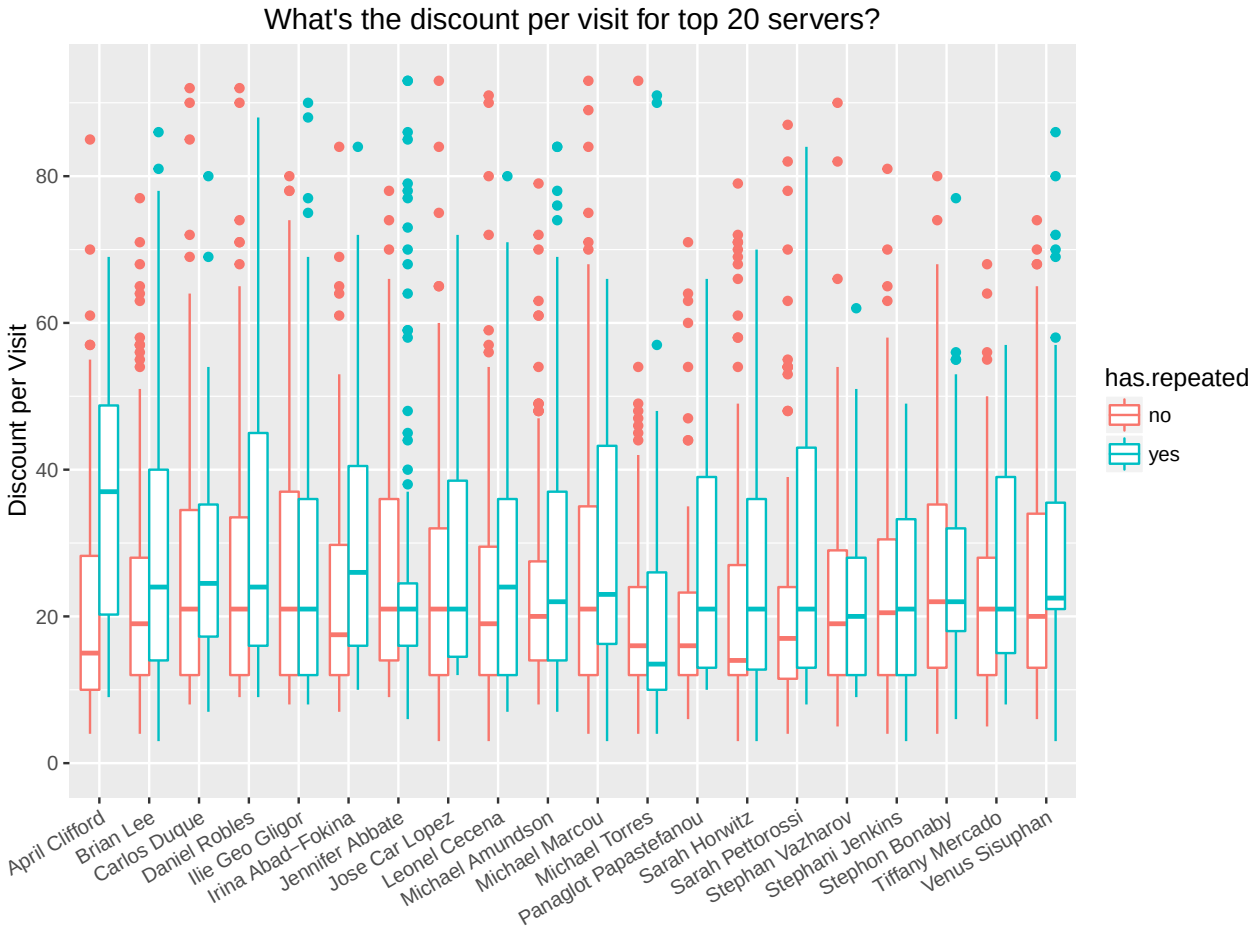
Party size cannot explain the delta between % of revenue and % of visits. Party size has some influence but not much. For example, Stephon B has -2.0 delta and Daniel R has +1.7; Stephon B servers smaller tables on average than Daniel so you would expect Daniel to generate more overall revenue. However, Sarah Horwitz servers on average the same party size as Daniel R; however, Sarah H's delta is ~ 0.


Can difference of % Revenue and % Visits be
explained by % of time server works dinner shifts?
% of Dinner Visits
Difference % of Revenue and % of Visits
Stephan Vazharov
Daniel Robles
Carlos Duque
Brian Lee
Michael Torres
Ilie Geo Gligor
April Clifford
Jose Car Lopez
Irina Abad−Fokina
Michael Marcou
Venus Sisuphan
Sarah Horwitz
Michael Amundson
Leonel Cecena
Sarah Pettorossi
Panaglot Papastefanou
Stephon Bonaby
Tiffany Mercado
Stephani Jenkins
Jennifer Abbate



Time of the day and of the week do correlate with higher revenue generations. There are still some servers who are outperforming and underperforming. For example, Jennifer A doesn't work nights or weekends yet she generates more revenue than expected. On the contrary, Stephan B works strictly nights and heavily on the weekend but he's revenue generation is much lower than it should. Perphaps Stephan B is a bartender.

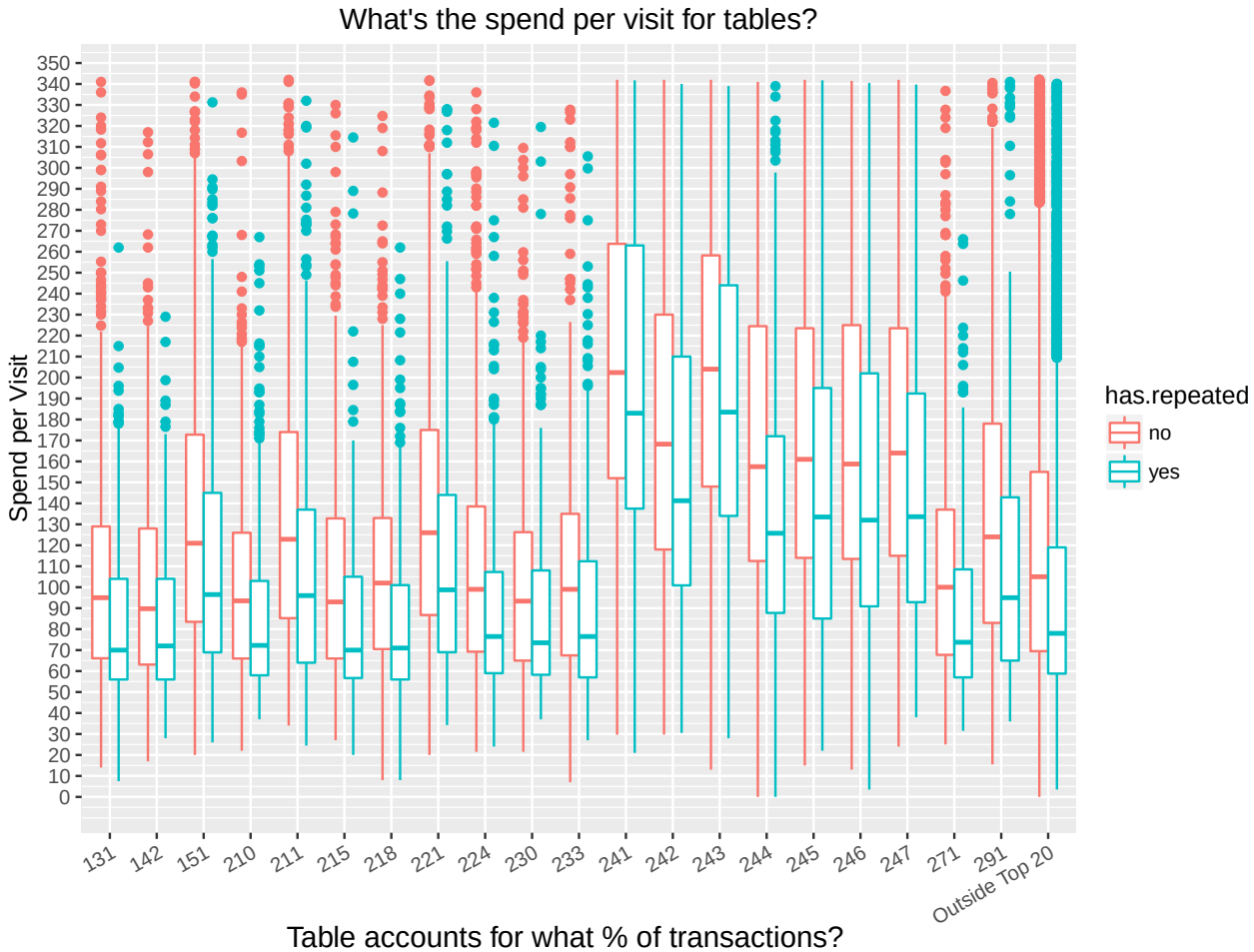

What's the spend per visit for top 20 servers?
Spend per Visit
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
April Clifford
Brian Lee
Carlos Duque
Daniel Robles
Ilie Geo Gligor
Irina Abad−Fokina
Jennifer Abbate
Jose Car Lopez
Leonel Cecena
Michael Amundson
Michael Marcou
Michael Torres
Panaglot Papastefanou
Sarah Horwitz
Sarah Pettorossi
Stephan Vazharov
Stephani Jenkins
Stephon Bonaby
Tiffany Mercado
Venus Sisuphan
has.repeated
no
yes



Same pattern as previously discussed: repeat guest spend less and get higher discounts regardless of the server.

1.11 Tables

As with the servers, I analyzed the top 20 tables by transactions. The top 20 tables account for about 43% of the transactions; ideally you want 80/20 but there are too many tables to display graphically.


What's the spend per visit for tables?
Spend per Visit
has.repeated
no
yes
131
142
151
210
211
215
218
221
224
230
233
241
242
243
244
245
246
247
271
291
Outside Top 20

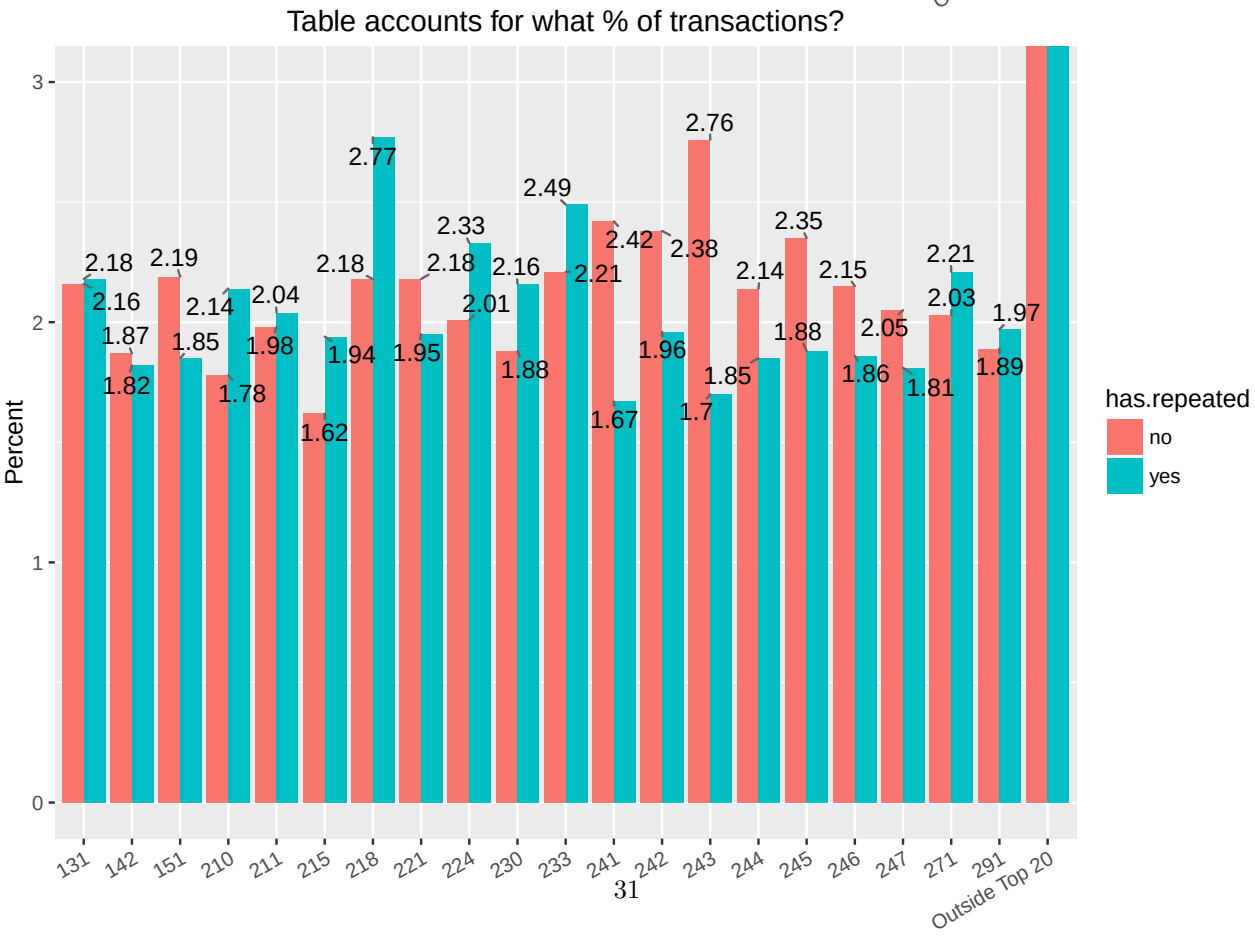


Table accounts for what % of transactions?
Percent
has.repeated
no
yes
2.18
2.16
2.19
1.87
1.82
2.14
1.85
1.78
2.04
1.98
1.62
2.18
1.94
2.77
1.95
2.18
2.33
2.01
2.16
1.88
2.49
2.21
1.67
2.42
1.96
2.38
2.76
1.7
1.85
2.14
2.35
1.88
2.15
1.86
2.05
1.81
2.21
2.03
1.89
1.97
131
142
151
210
211
215
218
221
224
230
233
241
242
243
244
245
246
247
271
291
Outside Top 20

The top 20 tables are used fairly at the same rate, but there are clear distinction in the spend for tables. Customers spend less on tables 131 through 233. Are these bar seating?

2.0 Conclusions

The profile that emerges from the analysis is that this restaurant doesn't truly have repeat guest. The median days between visits for repeats is about 70 days and the majority are repeat customer only come back once or twice for likely a businesslunch where they are likely just ordering food and soft drinks. The anomaly with guest getting a discount even though the table amount is 0 should be investigated further.