Data Analysis Assignment 4

Data Overview

	dissatisfied	neutral	satisfied	Overall
	(N=921)	(N=872)	(N=1394)	(N=3187)
Leg.room.service				
Mean (SD)	3.04 (1.32)	2.95 (1.31)	3.89 (1.13)	3.39 (1.31)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]
Arrival.Delay.in.Minutes				
Mean (SD)	17.4 (39.4)	21.1 (54.7)	11.0 (28.4)	15.6 (40.5)
Median [Min, Max]	2.00 [0, 586]	0 [0, 822]	0 [0, 350]	0 [0, 822]
Inflight.service				
Mean (SD)	3.35 (1.16)	3.32 (1.21)	4.01 (1.09)	3.63 (1.19)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]
On.board.service				
Mean (SD)	3.02 (1.26)	2.97 (1.27)	3.93 (1.09)	3.40 (1.28)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]
Cleanliness				
Mean (SD)	2.91 (1.34)	2.88 (1.33)	3.81 (1.11)	3.30 (1.32)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	3.00 [1.00, 5.00]
Inflight.wifi.service				
Mean (SD)	2.44 (0.944)	2.36 (0.936)	3.36 (1.40)	2.82 (1.26)
Median [Min, Max]	2.00 [1.00, 5.00]	2.00 [1.00, 5.00]	4.00 [1.00, 5.00]	3.00 [1.00, 5.00]
Ease.of.Online.booking				
Mean (SD)	2.69 (1.18)	2.61 (1.14)	3.18 (1.41)	2.89 (1.30)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]
Online.boarding				
Mean (SD)	2.68 (1.08)	2.69 (1.08)	4.17 (0.950)	3.34 (1.26)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]
Departure.Arrival.time.convenient				
Mean (SD)	3.25 (1.38)	3.36 (1.36)	3.11 (1.43)	3.22 (1.40)
Median [Min, Max]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]
Checkin.service				
Mean (SD)	3.04 (1.26)	3.08 (1.31)	3.64 (1.15)	3.31 (1.26)
Median [Min, Max]	3.00 [1.00, 5.00]	3.00 [1.00, 5.00]	4.00 [1.00, 5.00]	3.00 [1.00, 5.00]
Age				
Mean (SD)	38.0 (16.5)	38.5 (16.9)	42.5 (12.1)	40.1 (15.0)
Median [Min, Max]	36.0 [7.00, 80.0]	38.0 [7.00, 80.0]	43.0 [9.00, 80.0]	41.0 [7.00, 80.0]

- Table above shows the distribution of customers in the different satisfactory categories based on some important different customers features, their satisfaction with various aspects of the flight, and detail of their flight information.
- The following tables show the distribution of customers in the different satisfactory categories and also the distribution of different satisfactory categories within different customer type, gender, travel type, and travel class in the plane.
- We have **more data points at the satisfied level** compared to other levels. After running Pearson's Chi-squared test on categorical variables, including different gender, customer type, type of travel,and travel class in the plane we can tell that **customer type, type of travel, and travel class in the plane** might be very correlated to customer satisfaction (p-value < 0.01).

Var1	Freq
dissatisfied	921
neutral	872
satisfied	1394

	disloyal Customer	Loyal Customer
dissatisfied	0.45	0.26
neutral	0.35	0.26
satisfied	0.20	0.48

	Female	Male
dissatisfied	0.28	0.29
neutral	0.27	0.28
satisfied	0.45	0.43

	Business travel	Personal Travel
dissatisfied	0.22	0.46
neutral	0.19	0.47
satisfied	0.60	0.07

	Business	Eco	Eco Plus
dissatisfied	0.17	0.40	0.41
neutral	0.14	0.40	0.39
satisfied	0.69	0.19	0.19

Model

- Customers satisfaction is on an **ordered 3 scale outcome: dissatisfied, neutral, and satisfied**, which means that our response variable is ordinal. When we have ordinal response with categories dissatisfied, neutral, and satisfied, we need to use models that can reflect the ordering. Pr[yi <= dis|xi] <= Pr[yi <= neu|xi] <= Pr[yi <= sa|xi]. Here, the ordering of probabilities is the cumulative probabilities. Therefore, the model we are going to use is one type of generalized linear model, called **proportional odds model**.
- With all the predictors that we're going to use for this analysis, our model is as following:

$$\log(\Pr[\text{yi} <= \text{j}|\text{xi}]/\Pr[\text{yi} > \text{j}|\text{xi}]) = \beta_0 + \beta_1(\text{Seat. comfort}) + \beta_2(\text{Leg. room. service}) + \beta_3(\text{Arrival. Delay. in. Minutes})+$$

$$\beta_4(\text{Departure. Delay. in. Minutes})+\beta_5(\text{Inflight. service})+\beta_6(\text{On. board. service})+\beta_7(\text{Cleanliness})+\beta_8(\text{Inflight. entertainment})+$$

$$\beta_9(\text{Inflight. wifi. service})+\beta_{10}(\text{Age})+\beta_{11}(\text{Type. of. TravelPersonal})+\beta_{12}(\text{Customer. Typeloyal})+\beta_{13}(\text{Gender})+\beta_{14}(\text{ClassEco})$$

$$+\beta_{15}(\text{ClassEcoPlus})+\beta_{16}(\text{Flight. Distance})+\beta_{17}(\text{Ease. of. Online. booking})\beta_{18}(\text{Online. boarding})+\beta_{19}(\text{Food. and. drink})+$$

$$\beta_{20}(\text{Departure. Arrival. time. convenient}) + \beta_{21}(\text{Baggage. handling}) + \beta_{22}(\text{Gate. location}) + \beta_{23}(\text{Checkin. service}),$$

$$j = dissatisfied, neutral, satisfied$$

- The distributional assumption for the outcome variable is **Multinominal Distribution**, and the link function is **logit function**.

Model Results

- The influence of **Seat.comfort, Departure.Delay.in.Minutes, Inflight.entertainment, age, gender, Flight.Distance, Food.and.drink, Baggage.handling, and Gate.location** on customer satisfaction is not statistically significant (p-value > 0.05).
- **Leg.room.service**: Controlling for other predictors in the model, for every one unit increase in the score of Leg.room.service, the odds of being more satisfied(in the satisfied direction) are **1.18** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.11 and 1.26**, which is statistically significant**(p<0.01)**.
- **Arrival.Delay.in.Minutes**: Controlling for other predictors in the model, for every one minute more of Arrival.Delay.in.Minutes, the odds of being more satisfied(in the satisfied direction) are **0.99** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **0.98 and 1.00**, which is statistically significant**(p<0.01)**.
- **Inflight.service**: Controlling for other predictors in the model, for every one unit increase in the score of Inflight.service, the odds of being more satisfied(in the satisfied direction) are **1.13** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.03 and 1.24**, which is statistically significant**(p=0.01)**.
- **On.board.service**: Controlling for other predictors in the model, for every one unit increase in the score of On.board.service, the odds of being more satisfied(in the satisfied direction) are **1.20** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.11 and 1.30**, which is statistically significant**(p<0.01)**.
- **Cleanliness** : Controlling for other predictors in the model, for every one unit increase in the score of Cleanliness, the odds of being more satisfied(in the satisfied direction) are **1.12** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.01 and 1.24**, which is statistically significant**(p=0.03)**.

	Value	Std. Error	t value	p	OR	2.5 %	97.5 %
Seat.comfort	0.04	0.05	0.91	0.36	1.04	0.95	1.14
Leg.room.service	0.17	0.03	5.10	0.00	1.18	1.11	1.26
Arrival.Delay.in.Minutes	-0.01	0.00	-2.04	0.04	0.99	0.98	1.00
Departure.Delay.in.Minutes	0.01	0.00	1.45	0.15	1.01	1.00	1.01
Inflight.service	0.12	0.05	2.70	0.01	1.13	1.03	1.24
On.board.service	0.18	0.04	4.68	0.00	1.20	1.11	1.30
Cleanliness	0.11	0.05	2.21	0.03	1.12	1.01	1.24
Inflight.entertainment	0.09	0.06	1.64	0.10	1.10	0.98	1.23
Inflight.wifi.service	0.39	0.05	7.58	0.00	1.47	1.33	1.63
Age	0.00	0.00	-1.12	0.26	1.00	0.99	1.00
Type.of.TravelPersonal Travel	-1.97	0.12	-16.90	0.00	0.14	0.11	0.18
Customer.TypeLoyal Customer	1.86	0.12	15.40	0.00	6.41	5.02	8.20
GenderMale	-0.05	0.08	-0.57	0.57	0.96	0.82	1.12
ClassEco	-0.23	0.10	-2.45	0.01	0.79	0.64	0.98
ClassEco Plus	-0.54	0.15	-3.58	0.00	0.59	0.43	0.80
Flight.Distance	0.00	0.00	1.29	0.20	1.00	1.00	1.00
Ease.of.Online.booking	-0.18	0.05	-3.51	0.00	0.84	0.76	0.92
Online.boarding	0.52	0.04	11.93	0.00	1.68	1.54	1.83
Food.and.drink	-0.07	0.04	-1.67	0.10	0.93	0.85	1.02
Departure.Arrival.time.convenient	-0.08	0.04	-2.02	0.04	0.93	0.86	1.00
Baggage.handling	0.06	0.05	1.25	0.21	1.06	0.97	1.16
Gate.location	-0.04	0.04	-1.11	0.27	0.96	0.89	1.03
Checkin.service	0.20	0.03	5.95	0.00	1.22	1.14	1.31

- **Inflight.wifi.service** : Controlling for other predictors in the model, for every one unit increase in the score of Inflight.wifi.service, the odds of being more satisfied(in the satisfied direction) are **1.47** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.33 and 1.63**, which is statistically significant**(p<0.01)**.

- **Type.of.Travel**: Controlling for other predictors in the model, the odds that personal travelers of being more satisfied(in the satisfied direction) rather than being less satisfied(in the dissatisfied direction) are **0.14** times for business travelers. We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **0.11 and 0.18**, which is statistically significant**(p<0.01)**.

- **Customer.Type** : Controlling for other predictors in the model, the odds that loyal customer of being more satisfied(in the satisfied direction) rather than being less satisfied(in the dissatisfied direction) are **6.41** times the odds for disloyal customer. We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **5.02 and 8.20**, which is statistically significant**(p<0.01)**.

- **Class** : Controlling for other predictors in the model, the odds that people taking Eco/Eco_plus class of being more satisfied(in the satisfied direction) rather than being less satisfied(in the dissatisfied direction) are **0.79/0.59** times the odds for people taking business class . We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **0.64/0.43 and 0.98/0.80**, which is statistically significant**(p=0.01/p<0.01)**.

- **Ease.of.Online.booking**: Controlling for other predictors in the model, for every one unit increase in the score of Ease.of.Online.booking, the odds of being more satisfied(in the satisfied direction) are **0.84** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **0.76 and 0.92**, which is statistically significant**(p<0.01)**.

- **Online.boarding**: Controlling for other predictors in the model, for every one unit increase in the score of Online.boarding, the odds of being more satisfied(in the satisfied direction) are **1.68** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.54 and 1.83**, which is statistically significant**(p<0.01)**.

- **Departure.Arrival.time.convenient**: Controlling for other predictors in the model, for every one unit increase in the score of Departure.Arrival.time.convenient, the odds of being more satisfied(in the satisfied direction) are **0.93** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **0.86 and 1.00**, which is statistically significant**(p<0.01)**.

- **Checkin.service**: Controlling for other predictors in the model, for every one unit increase in the score of Checkin.service, the odds of being more satisfied(in the satisfied direction) are **1.22** times the odds of being less satisfied(in the dissatisfied direction). We are 95% confident that the true odds ratio of higher vs. lower categories(more satisfied vs. less satisfied) is between **1.14 and 1.31**, which is statistically significant**($p<0.01$)**.

Model Assessment

- The Accuracy of this model is **0.64**.

	dissatisfied	neutral	satisfied
dissatisfied	471	462	13
neutral	324	305	104
satisfied	126	105	1277

	x
Accuracy	0.64

Class	Sensitivity	Specificity
dissatisfied	0.49	0.76
neutral	0.23	0.86
satisfied	0.87	0.76

Multinomial vs. proportional—Compare predictions using the predictors Gender and Customer type

Checking proportional odds assumption

- When trying to compare predictions using the predictors *Gender* and *Customer type*, we create a new dataframe with different pairs of gender and customer type while keeping other variables all the same(use mean value for other numeric variables, and choose mode for categorical variables(business customer type and eco plane class)).
- The left table shows the results of prediction using the proportional odds model, and the right one shows the results of prediction using the multinomial model. We can tell that thought the overall trend is the same, the predicted probabilities of the two models are not quite the same, which means the proportional odds assumption is not perfectly met.

dissatisfied	neutral	satisfied
0.08	0.30	0.62
0.34	0.45	0.20
0.07	0.30	0.63
0.33	0.46	0.21

dissatisfied	neutral	satisfied
0.20	0.20	0.60
0.51	0.43	0.07
0.17	0.16	0.67
0.50	0.41	0.09

- **The accuracy for the multinomial model is 0.67**, which is higher than the accuracy for the proportional odds model.

	dissatisfied	neutral	satisfied
dissatisfied	489	408	78
neutral	314	363	19
satisfied	118	101	1297

	x
Accuracy	0.67

Conclusion

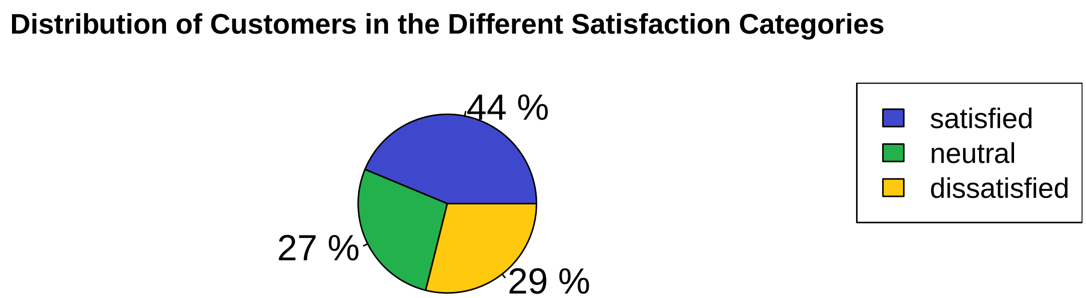
- Marketing and promotion initiatives to **improve customer loyalty and appeal to business travelers** are the most cost efficient way to improve customer satisfaction. Investments on**Leg.room.service, Inflight.service, On.board.service, Cleanliness, Inflight.wifi.service, Online.boarding, and Checkin.service** can also improve customer satisfaction.
- The ordered regression approach is more powerful and parsimonious when the outcome variable has an ordered structure and the "proportional odds" assumption holds. The proportional odds assumption means that for each term included in the model, the 'slope' estimate between each pair of outcomes across two response levels are assumed to be the same regardless of which partition we consider. Since the assumption is not perfectly met, multinomial logistic regression is more "precise". It is a tradeoff between power and assumptions. Since our response variable is ordinal, the proportional odds model is still more suitable. In addition, in order to make analysis based on the current dataset, we exclude value 0 from the original dataset since 0 means 'not applicable'. Limitations also include the fact that loyal customers are more likely to respond to the survey and are more likely to be satisfied with services, which also decreases the validity of this analysis.

Report for the Client

Introduction

In order to find services that are worth investing to improve customer satisfaction, we first need to have a look at our dataset. Overall, we have gathered 3187 data from our customer, where 1394(43.74%) of them are satisfied, 872(25.95%) of them are neutral, and 921(28.90%) of them are dissatisfied.

The following table shows the distribution of customers in the different satisfactory categories based on different customers features, including age, gender, Purpose of the passenger's flight (personal, business), travel class in the plane (business, eco, eco plus), and customer type (loyal or disloyal).



	dissatisfied	neutral	satisfied	Overall
	(N=921)	(N=872)	(N=1394)	(N=3187)
Age				
Mean (SD)	38.0 (16.5)	38.5 (16.9)	42.5 (12.1)	40.1 (15.0)
Median [Min, Max]	36.0 [7.00, 80.0]	38.0 [7.00, 80.0]	43.0 [9.00, 80.0]	41.0 [7.00, 80.0]
Gender				
Female	461 (50.1%)	433 (49.7%)	725 (52.0%)	1619 (50.8%)
Male	460 (49.9%)	439 (50.3%)	669 (48.0%)	1568 (49.2%)
Type.of.Travel				
Business travel	481 (52.2%)	415 (47.6%)	1326 (95.1%)	2222 (69.7%)
Personal Travel	440 (47.8%)	457 (52.4%)	68 (4.9%)	965 (30.3%)
Class				
Business	270 (29.3%)	227 (26.0%)	1087 (78.0%)	1584 (49.7%)
Eco	551 (59.8%)	550 (63.1%)	261 (18.7%)	1362 (42.7%)
Eco Plus	100 (10.9%)	95 (10.9%)	46 (3.3%)	241 (7.6%)
Customer.Type				
disloyal Customer	224 (24.3%)	176 (20.2%)	97 (7.0%)	497 (15.6%)
Loyal Customer	697 (75.7%)	696 (79.8%)	1297 (93.0%)	2690 (84.4%)

Methods

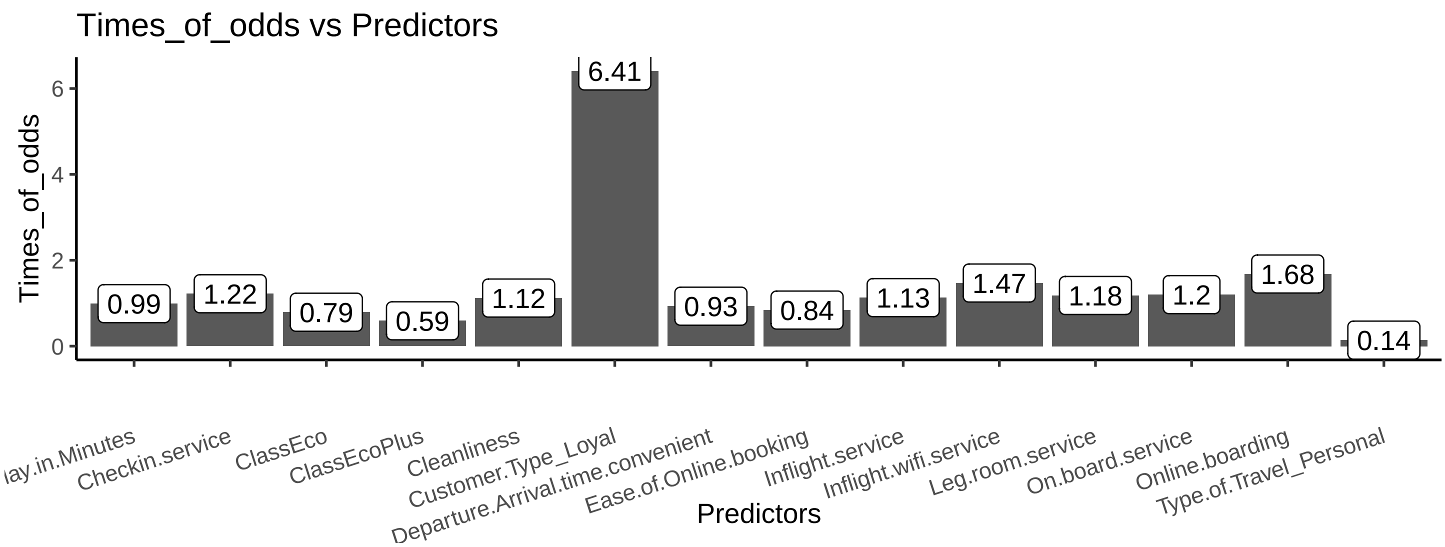
Customer satisfaction is on an **ordered 3 scale outcome: dissatisfied, neutral, and satisfied**, which means that our response variable is ordinal. When we have ordinal responses with categories dissatisfied, neutral, and satisfied, we need to use models that can reflect the ordering.

The model we are going to use is one type of generalized linear model, called **proportional odds model**.With coefficients that we get from this model, we are able to understand the times of odds of being more satisfied (in the satisfied direction) than less satisfied (in the dissatisfied direction), which can determine how significant each predictor is in predicting satisfaction level and, as a result, determine the most important elements that LaneAir can adjust going forward to enhance the customer experience.

Results

After fitting proportional odds model used important different customers features, their satisfaction with various aspects of the flight, and detail of their flight information, we found out that **Seat.comfort, Departure.Delay.in.Minutes, Inflight.entertainment, age, gender, Flight.Distance, Food.and.drink, Baggage.handling, and Gate.location** don't impact our customers satisfaction a lot, while other features play important roles in our customers satisfaction.

The following figure shows the times of odds of being more satisfied(in the satisfied direction) than being less satisfied(in the dissatisfied direction) with some features that we care about to improve our customers satisfaction.



We can tell that with times_of_odds greater than 1, increase customers satisfaction in **Leg.room.service, Inflight.service, On.board.service, Cleanliness, Inflight.wifi.service, Online.boarding, and Checkin.service** can improve customer satisfaction. Meanwhile, **loyal customer and business travelers** are more satisfied with our flight.

Conclusion

- Marketing and promotion initiatives to improve customer loyalty and appeal to business travelers are the most cost efficient way to improve customer satisfaction.
- Investments on technology and hiring more flight attendants or other staff can also be good choices to improve customer satisfaction.
- Newer plane models that improve reliability to reduce delays have no statistically significant relationship with customer satisfaction.Though larger seats can improve seat comfort and leg room and thus improve customer satisfaction, after considering the difficulties of investments, the disadvantages of reducing the number of seats per plane, and their impact on customer satisfaction, we only recommend investing in newer, larger seats when investments on the first two points above have already been made and further improvement on customer satisfaction is needed.
- Investments on online boarding and checkin service can also be good ways to improve customer satisfaction.
- Limitations: Because loyal customers are more likely to respond to the survey and are more likely to be satisfied with our services, we may need to find ways to get more balanced data to better understand the drives of customer satisfaction.

Appendix: All code for this report

```
knitr::opts_chunk$set(echo = TRUE)
library(ISLR2)
library(knitr)
library(ggplot2)
library(kableExtra)
library(lattice)
library(dplyr)
library(gt)
library(janitor)
library(flextable)
library(magrittr)
library(Hmisc)
library(caret)
library(patchwork)
library(leaps)
library(rms)
library(arm)
library(pROC)
library(e1071)
library(caret)
require(gridExtra)
library(table1)
library(nnet)
load('airline_survey')
airline$Satisfaction <- ordered(airline$Satisfaction,
                              levels=c("dissatisfied","neutral","satisfied"))
airline <- airline[airline$Seat.comfort != 0 & airline$Leg.room.service !=0 & airline$Inflight.service !=0 & ai
table1(~ Leg.room.service+Arrival.Delay.in.Minutes+Inflight.service+On.board.service +Cleanliness +Inflight.wi
t1 <- table(airline$Satisfaction)%>%round(2)
#knitr::kable(table(airline$Satisfaction),digits = 2)%>% kable_styling(position="center",full_width = FALSE,la

#loyalty/customer type
t2 <- prop.table(table(airline$Satisfaction, airline$Customer.Type), 2)%>%round(2)
#knitr::kable(prop.table(table(airline$Satisfaction, airline$Customer.Type), 2),digits = 2)%>% kable_styling(p
#chisq.test(table(airline$Satisfaction, airline$Customer.Type))

#gender
t3 <- prop.table(table(airline$Satisfaction, airline$Gender), 2)%>%round(2)
#knitr::kable(prop.table(table(airline$Satisfaction, airline$Gender), 2),digits = 2)%>% kable_styling(position=
#chisq.test(table(airline$Satisfaction, airline$Gender))

#type of travel
t4 <- prop.table(table(airline$Satisfaction, airline$Type.of.Travel), 2)%>%round(2)
#knitr::kable(prop.table(table(airline$Satisfaction, airline$Type.of.Travel), 2),digits = 2)%>% kable_styling(p
#chisq.test(table(airline$Satisfaction, airline$Type.of.Travel))

#class
t5 <- prop.table(table(airline$Satisfaction, airline$Class), 2)%>%round(2)
#knitr::kable(prop.table(table(airline$Satisfaction, airline$Class), 2),digits = 2)%>% kable_styling(position=
#chisq.test(table(airline$Satisfaction, airline$Class))

knitr::kable(list(t1, t2, t3,t4,t5))

Model1 <- polr(Satisfaction ~ Seat.comfort + Leg.room.service + Arrival.Delay.in.Minutes + Departure.Delay.in.M

#no p-value in summary
#summary(Model1)
```

```
#obtain p-value
p <- pnorm(-abs(summary(Model1)$coef[,"t value"])) * 2
ctable <- cbind(summary(Model1)$coef, p)


#knitr::kable(cbind(summary(Model1)$coef, p),digits = 2)%>% kable_styling(position="center",full_width = FALSE,
ptable <- cbind(summary(Model1)$coef, p)
ptable<-ptable[1:23,]
#confidence intervals
knitr::kable(cbind(ptable,exp(cbind(OR = coef(Model1), confint(Model1)))),digits = 2)%>% kable_styling(position
#head(Model1$fitted.values)
#head(predict(Model1))
Conf_mat <- confusionMatrix(predict(Model1), airline$Satisfaction)

#confusionMatrix(predict(Model1), airline$Satisfaction)

t1 <- Conf_mat$table
t2 <- Conf_mat$overall["Accuracy"];
#t3 <- Conf_mat$byClass[c("Sensitivity","Specificity")]

Class <- c('dissatisfied','neutral', 'satisfied' )
Sensitivity <- c( '0.49', '0.23', '0.87')
Specificity <- c('0.76','0.86','0.76')
t3 <- data.frame(Class,Sensitivity,Specificity)
#t3 %>% regulartable() %>% autofit()

knitr::kable(list(t1, t2,t3),digits = 2)%>% kable_styling(position="center",full_width = FALSE,latex_options =
newdat <- data.frame(Seat.comfort = mean(airline$Seat.comfort), Leg.room.service =mean(airline$Leg.room.service
multmod <- multinom(Satisfaction ~ Seat.comfort + Leg.room.service + Arrival.Delay.in.Minutes + Departure.Delay
p1 <- predict(Model1, newdat,type = 'probs')
p2 <- predict(multmod, newdat,type = 'probs')
knitr::kable(list(p1, p2), digits = 2)%>% kable_styling(position="center",full_width = FALSE,latex_options = c(
#Conf_mat_mult <- confusionMatrix(predict(multmod, newdat), airline$Satisfaction)
#Conf_mat_prop <- confusionMatrix(predict(Model1, newdat), airline$Satisfaction)
Conf_mat_mult <- confusionMatrix(predict(multmod), airline$Satisfaction)
t1 <- Conf_mat_mult$table
t2 <- Conf_mat_mult$overall["Accuracy"]

#t3 <- Conf_mat_prop$table
#t4 <- Conf_mat_prop$overall["Accuracy"]

knitr::kable(list(t1, t2),digits = 2)%>% kable_styling(position="center",full_width = FALSE,latex_options = c("
#knitr::kable(list(t3, t4), digits = 2)%>% kable_styling(position="center",full_width = FALSE,latex_options = c
airline$Satisfaction <- ordered(airline$Satisfaction,
                              levels=c("dissatisfied","neutral","satisfied"))
airline <- airline[airline$Seat.comfort != 0 & airline$Leg.room.service !=0 & airline$Inflight.service !=0 & ai
x <- c(1394,872,921)
piepercent = paste(round(100*x/sum(x)), "%")
cols = c("#3f48CC","#22B14C","#FFC90E")
la<-c('satisfied','neutral','dissatisfied')
pie(x, labels = piepercent, main = "Distribution of Customers in the Different Satisfaction Categories", cex =
legend("topright", la, cex=1.2,fill=cols)
table1(~  Age + Gender + Type.of.Travel + Class + Customer.Type| Satisfaction, data=airline)
Model1 <- polr(Satisfaction ~ Seat.comfort + Leg.room.service + Arrival.Delay.in.Minutes + Departure.Delay.in.M

#no p-value in summary
#summary(Model1)

#obtain p-value
p <- pnorm(-abs(summary(Model1)$coef[,"t value"])) * 2
```

```
ctable <- cbind(summary(Model1)$coef, p)


#knitr::kable(cbind(summary(Model1)$coef, p),digits = 2)%>% kable_styling(position="center",full_width = FALSE,
ptable <- cbind(summary(Model1)$coef, p)
ptable<-ptable[1:23,]
#confidence intervals
#knitr::kable(cbind(ptable,exp(cbind(OR = coef(Model1), confint(Model1)))),digits = 2)%>% kable_styling(positio
#Predictors <- c('Seat.comfort','Leg.room.service', 'Arrival.Delay.in.Minutes' , 'Inflight.service' , 'On.board
#Times_of_odds <-c('1.23','1.18','0.99','1.26','1.29','1.25','1.30','0.11','5.92')

df <- data.frame(Predictors = c('Leg.room.service', 'Arrival.Delay.in.Minutes' , 'Inflight.service' , 'On.board
Times_of_odds = c(1.18,0.99,1.13,1.20,1.12,1.47,0.14,6.41,0.79,0.59,0.84,1.68,0.93,1.22))

ggplot(df,aes(x=Predictors, y=Times_of_odds)) +
  geom_bar(stat = 'identity') +
  theme_classic() +
  labs(title="Times_of_odds vs Predictors",x="Predictors",y="Times_of_odds")+
  geom_label(data = df, aes(label = Times_of_odds))+
  theme(legend.position="none",axis.text.x = element_text(angle = 18, vjust = 0.5, hjust=1))
#plot(x = as.factor(Predictors), y = Times_of_odds)
```