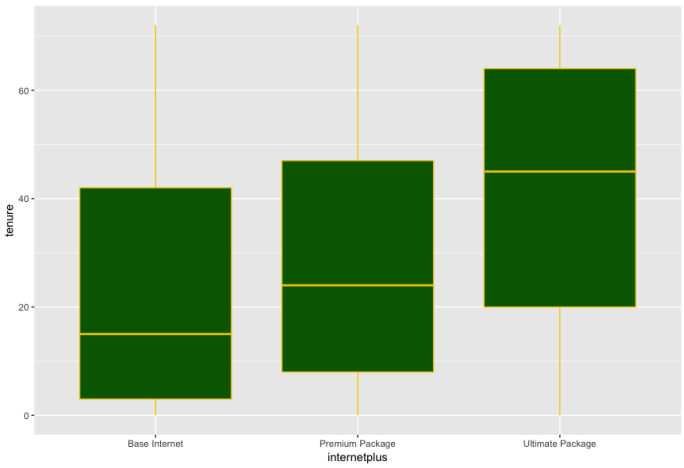
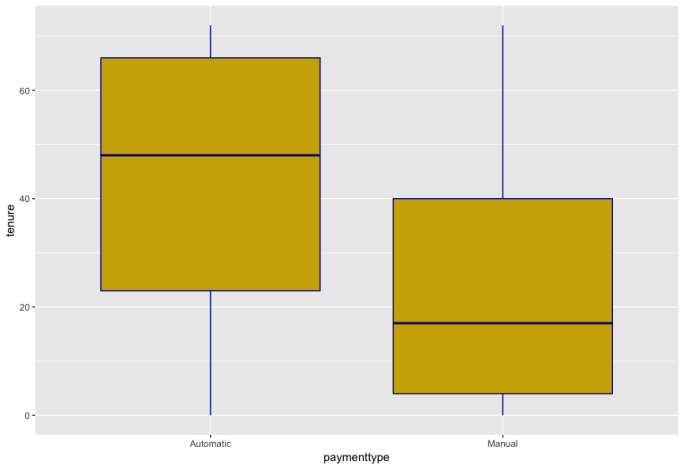
Dalton Anderson

Describe the process by which you cleaned, processed, and partitioned data as necessary. (2 points)

There were 11 NAs in total charges that were dropped. After doing some reading online it showed most of churn is caused by failed payments either people forget to update their new card or mail the payment. I did some feature engineering to create payment type from payment method to create automatic and manual. I also thought for internet service churn may be reduced by customers having extra services. I created a column called internet plus to denote the amount of additional services people had. For the dependent variable, I converted churn into a number for logit model.





What predictors do you think contributes to the churn of (i) only telephone customers, (ii) only Internet service customers, and (iii) customers who subscribe to both phone and Internet services? Explain the rationale for your answer. (2 points)

Telephone customers	Telephone customers churn is effected by amount the of lines, partnered, payment type.
Internet service customers	Internet service customers churn is effected how deep the customer is in the ecosystem. When a customer is just subscribed to only internet services they churn at this rate (.). When a customer is subcribed to many services under the internet product they churn at this rate (. .)
Phone and Internet customer	The churn effect is a combination of the effects from above.

Predictor Table for A7. Telco Churn

predictor	effect	rationale
Target (y)		
Churn	dv	The period a patient survivals
Predictors		
SeniorCitizen	+	Senior Citzens might be more cost sensitive as their income is fixed.
Partner	-	The higher number of people on a plan the more hassle it is to move to another company.
Dependents	-	The higher number of people on a plan the more hassle it is to move to another company.
tenure	+	The longer someone stays the higher likelihood they may stay.
PhoneService	+/-	Just because someone has a a phone services doesn't mean they may or may not churn.
MultipleLines	-	A family may move less then a single person.
InternetService	+/-	Just because someone has a a internet services doesn't mean they may or may not churn.
OnlineSecurity	-	More services less likely the costomer will churn.
OnlineBackup	-	More services less likely the costomer will churn.
DeviceProtection	-	More services less likely the costomer will churn.
TechSupport	-	More services less likely the costomer will churn.
StreamingTV	-	More services less likely the costomer will churn.
StreamingMovies	-	More services less likely the costomer will churn.
internet_plus	-	The customers that have many services plus their normal internet. I don't think these customers will leave as often.
ultimate_plus	-	The customers that have many services plus their normal internet and phone service. I don't think these customers will leave as often.
Contract	+	Monthly paying customers may lead to more churn as it is easier to leave.
contract_type	+	Monthly vs yearly.
MonthlyCharges	+	For monthly contract customers this may have more of a effect.
TotalCharges	+/-	I am not sure.
PaymentMethod	+	After doing some reading it seems one of the main reason for churn is failed payments.
paymenttype	+	Automatic vs non-automatic.
Excuded		
customerID	NA	Not needed.
PaperlessBilling	NA	The effect of this is variable is not clear on churn.

gender	NA	Gender shouldn't matter.

What predictors do you think contributes to the churn of (i) only telephone customers, (ii) only Internet service customers, and (iii) customers who subscribe to both phone and Internet services? List reasoning for your answer. No points without reasoning.

```
============================================================
                          Dependent variable:
                      ----------------------------
                                churnnum
                         (1)       (2)       (3)
------------------------------------------------------------
seniorcitizen             0.256**    1.518**   0.691**
                         (0.099)    (0.615)   (0.304)

partnerYes                -0.083     -0.031     0.025
                         (0.089)    (0.274)   (0.255)

tenure                   -0.040***  -0.049*** -0.027***
                         (0.003)    (0.013)   (0.008)

multiplelinesYes          0.297***    1.176
                         (0.096)    (1.158)

monthlycharges            0.034***   -0.195    -0.012
                         (0.003)    (0.221)   (0.023)

internetplusPremium Package  -0.610***            0.010
```

	(0.126)	(0.414)	
internetplusUltimate Package	-0.510***		0.857
	(0.150)	(0.600)	
contracttypeYearly	-1.065***	-1.659***	-1.632***
	(0.129)	(0.391)	(0.379)
paymenttypeManual	0.242***	-0.130	0.397
	(0.091)	(0.282)	(0.286)
Constant	-1.926***	2.829	-0.299
	(0.237)	(4.418)	(0.682)
---	---	---	---
Observations	3,620	1,160	502
Log Likelihood	-1,789.075	-243.235	-213.817
Akaike Inf. Crit.	3,598.149	502.470	445.634

Both Services	marginal effect	Effect meaning
contracttypeYearly	-19%	Customers with yearly contacts churn 19% less than customer who are on monthly contracts.
multiplelinesNo	14%	Phone customers without multiplelines churn 14% more than customer with multiplelines.
internetplusPremium	-7%	Internet customers churn 7 less then subcribed to 3 services.
Phone Only		
contracttypeYearly	-19%	Customers with yearly contacts churn 19% less than customer who are on monthly contracts.
multiplelinesNo	13%	Phone customers without multiple lines churn 13% more than customer with multiple lines.

Seniorcitizen	5%	If the customer is a senior citizen they churn 5% more than non-seniors.
Internet Only		
contracttypeYearly	-19%	Customers with yearly contacts churn 19% less than customer who are on monthly contracts.
Seniorcitizen	5%	If the customer is a senior citizen they churn 5% more than non-seniors.
internetplusPremium	-5%	Internet customers churn 5 less then subcribed to 3 services.

Fit your models using test data, and compute recall, precision, F1-score, and AUC values for each of your three models. Create a table with these values. (2 points)

Model Name	Recall	Precision	F1
test_phone_model_out	51%	68%	58%
internet_model_out	49%	67%	57%
internet_model_out	51%	69%	59%