Apoorva Srinivas
Prerana Patil
Tai-Hua Chung
Yi-Yun Su

Background

- Challenges of the Consumer Electronics Industry
  - Low profit margin due to high competition
  - Difficult to charge premium as product lifetime is short
  - Access to competitor prices online increases probability of customer switching to a different retailer
  - Manufacturers have their own direct sales channels

- Best Buy Management Strategy
  - Focus on selling bundled products
  - Focus on extended warranty sales
    - The Geek Squad Protection Plan
    - Warranties offer substantial profits

Our Goal

1. Identify warranty customers
2. Develop marketing strategy

Research Question

Which customers have a higher propensity to purchase the warranty?

Steps

1. Variable Selection
2. Conceptual Model
3. Final Model
4. Analysis Result & Interpretation
5. Implication – Marketing Strategy
6. Limitations & Suggestions

Variable Selection

- Dependent Variable
  - Warranty

- Independent Variables

	Customer Behavior		***Customer Demographics***
+	Price Category	+/-	Income
+	Product Generation	+	Married
+	My BestBuy Credit Card Holder	?	Hispanic
-	Appliances	+	Family Size
		+	Age

Final Model

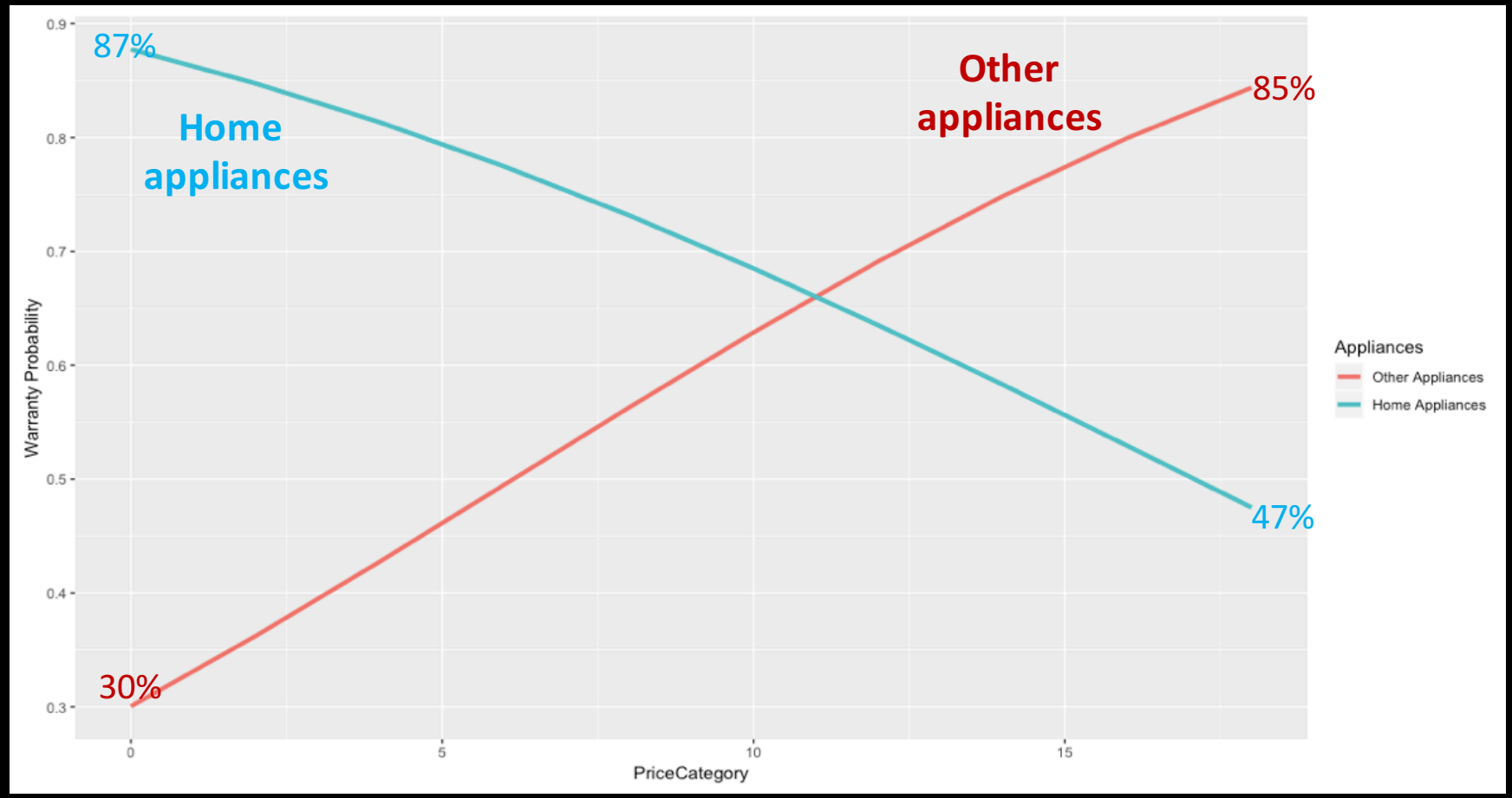
$$\text{Warranty Purchase Rate} = \beta_0 + \beta_1 \text{Price Category} + \beta_2 \text{Appliances} + \beta_3 \text{My Best Buy} + \beta_4 \text{Age} + \beta_5 \text{Income} + \beta_6 \text{Hispanic} + \beta_7 \text{Married} + \beta_8 \text{Price Category*Appliances}$$

Marginal Effect

Marginal Effects

	Dependent variable:
	Warranty
	Marg.Eff.w/RobStdEr
PriceCategory	0.0323***
	(0.0052)
appliances	0.6004***
	(0.0524)
MyBestBuy	0.0316
	(0.0195)
age	-0.0044
	(0.0026)
income	0.0445***
	(0.0129)
hisp	-0.3314***
	(0.0353)
married	0.2200***
	(0.0228)
PriceCategory:appliances	-0.0580***
	(0.0062)
Observations	3,206
Log Likelihood	-1,945.6640
Akaike Inf. Crit.	3,909.3270
Note:	*p<0.05; **p<0.01; ***p<0.001

Purchase Propensity- Appliances Type & Price



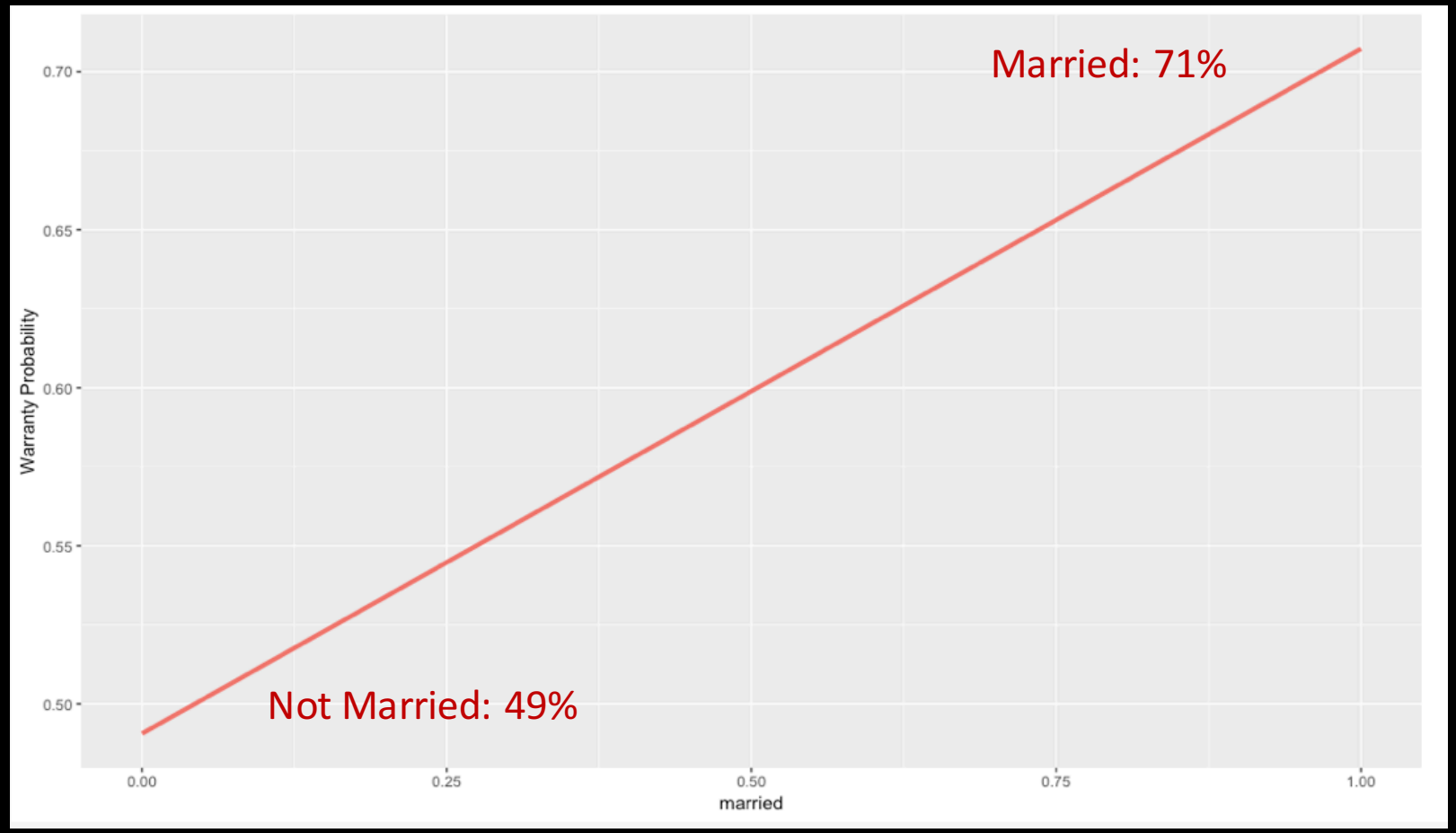
Price Category: Low ⟶ High

As Price Category increases the effect of **home appliances** on the probability to buy warranty **decreases**

As Price Category increases the effect of **other appliances** on the probability to buy warranty **increases**

Purchase Propensity- Marriage

Warranty Purchase Rate

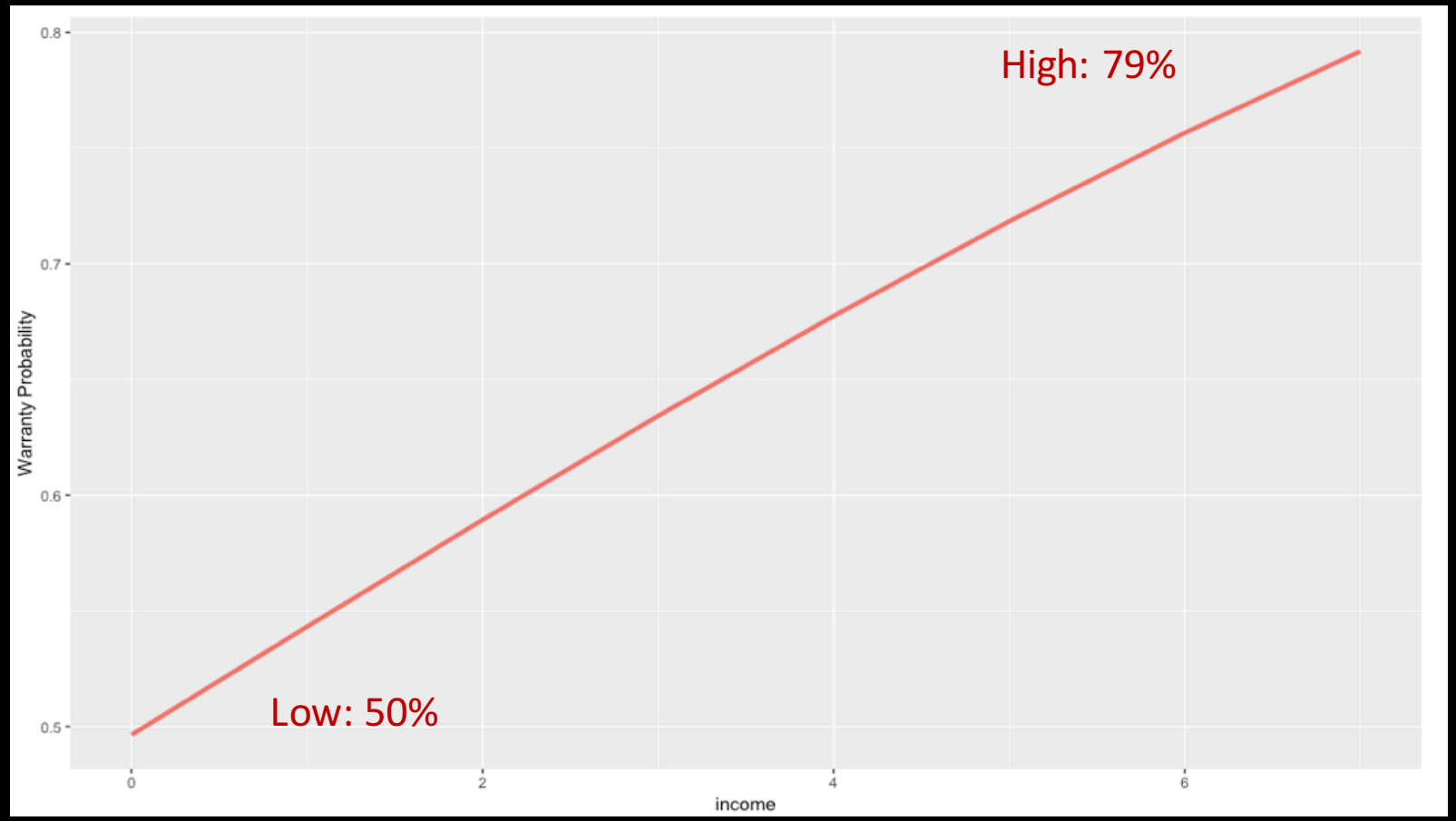


Marriage

The probability of buying the warranty **increases** for a customer who is **married**

Purchase Propensity- Income

Warranty Purchase Rate

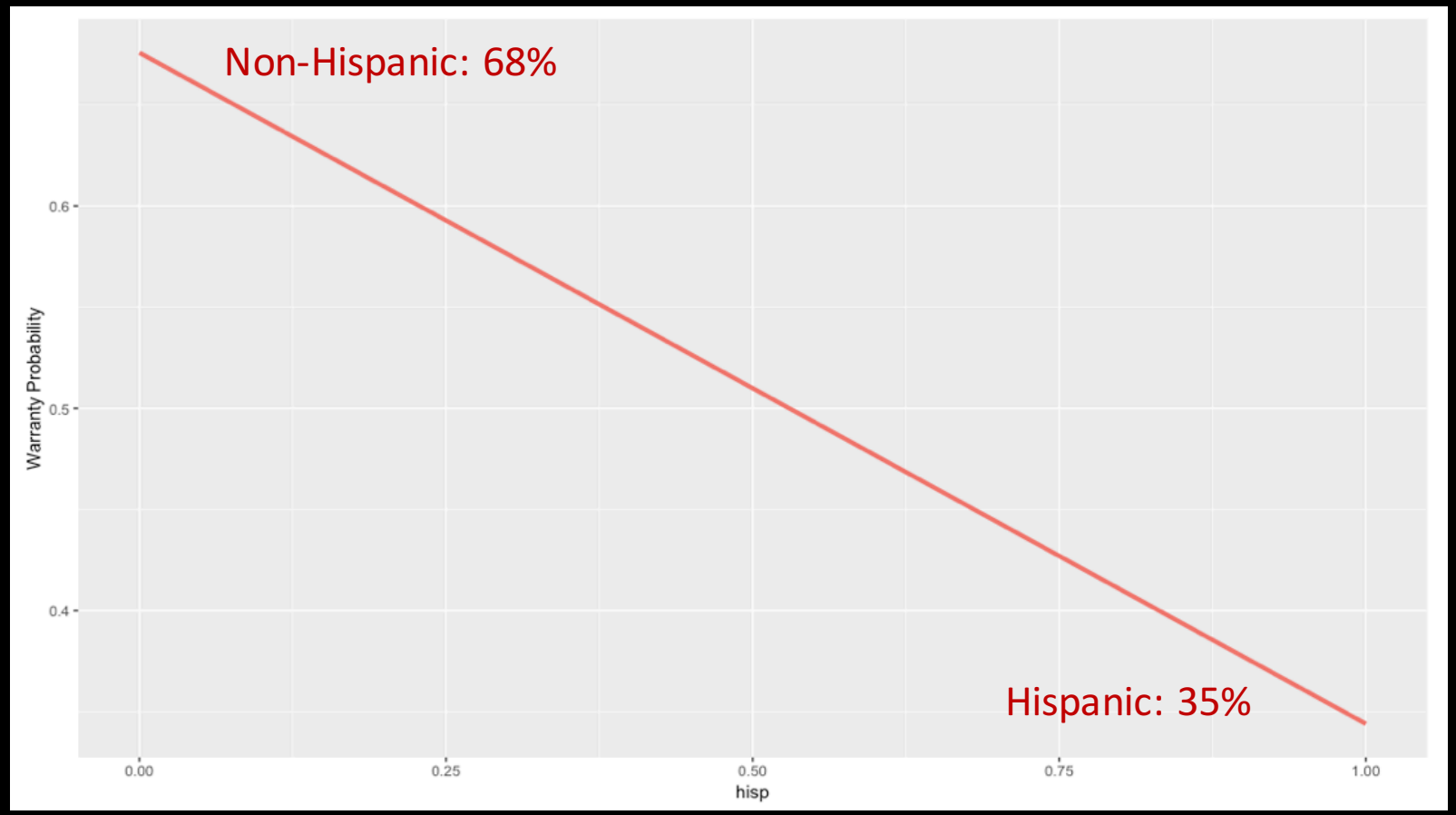


Income

The probability of buying the warranty **increases** for a customer who earns **higher income**.

Purchase Propensity- Hispanic

Warranty Purchase Rate



Hispanic

The probability of buying the warranty **decreases** for a customer who is **Hispanic**.

Implications – Marketing Strategy

Focus	Segment	Propensity	Increment
Product	Low Price Home Appliances	87%	+ 35%
	High Price Other Appliances	85%	+ 40%
Customer	High Income	79%	+ 29%
	Married	71%	+ 22%
	Non-Hispanic	68%	+ 33%

Limitations & Suggestions

Limitations	**Suggestions**
Potential Omitted Variables • Preferred Product Usage Period • Salesperson Skill • Discounts	Regarding given data: • Average age of customers – 67 years • Data pertaining to Non-Hispanic population
	Additional data that would strengthen our analysis: • Past Purchase of Warranty • Price of Warranty

Appendix

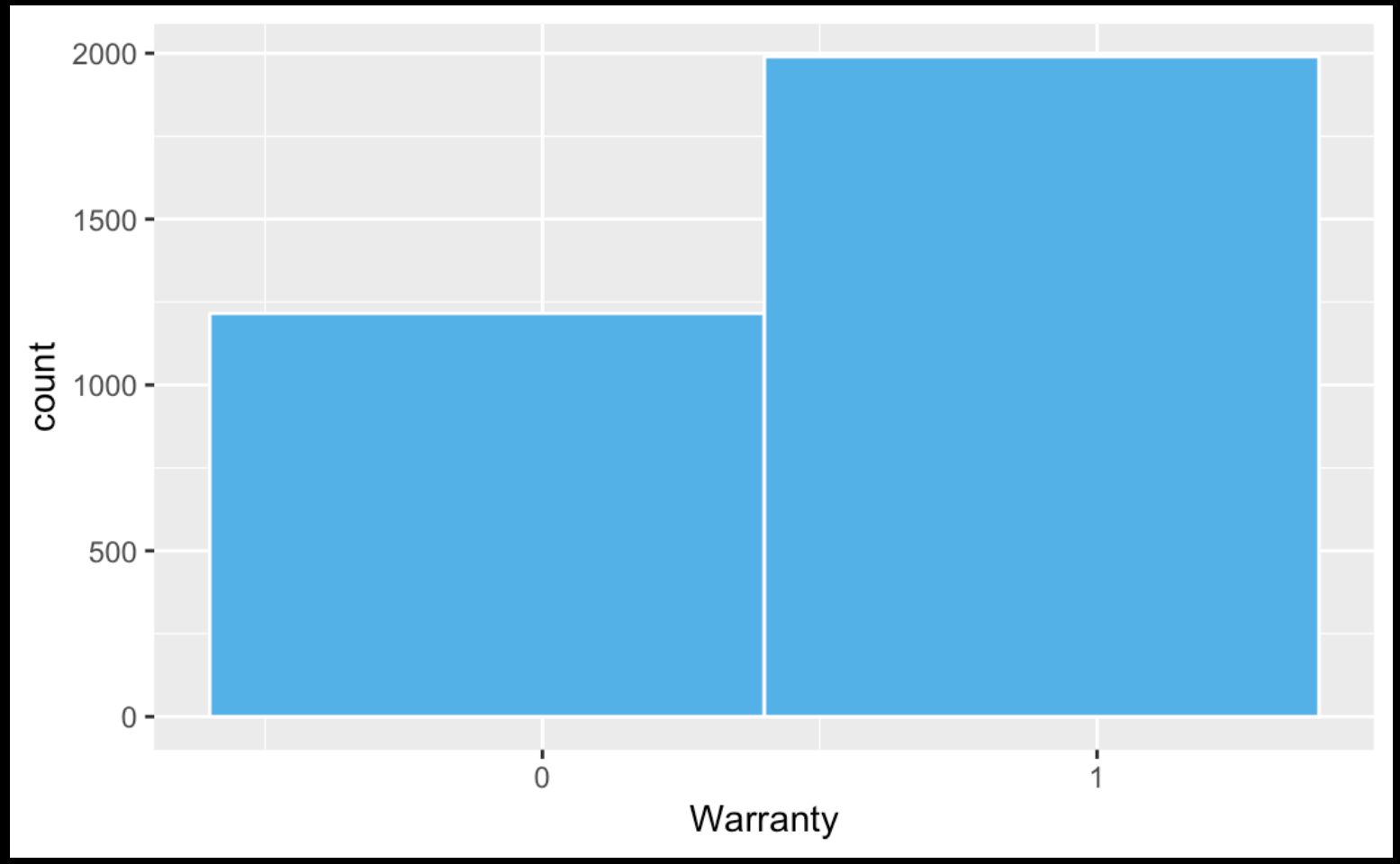
Preliminary Analysis

Descriptive Statistics

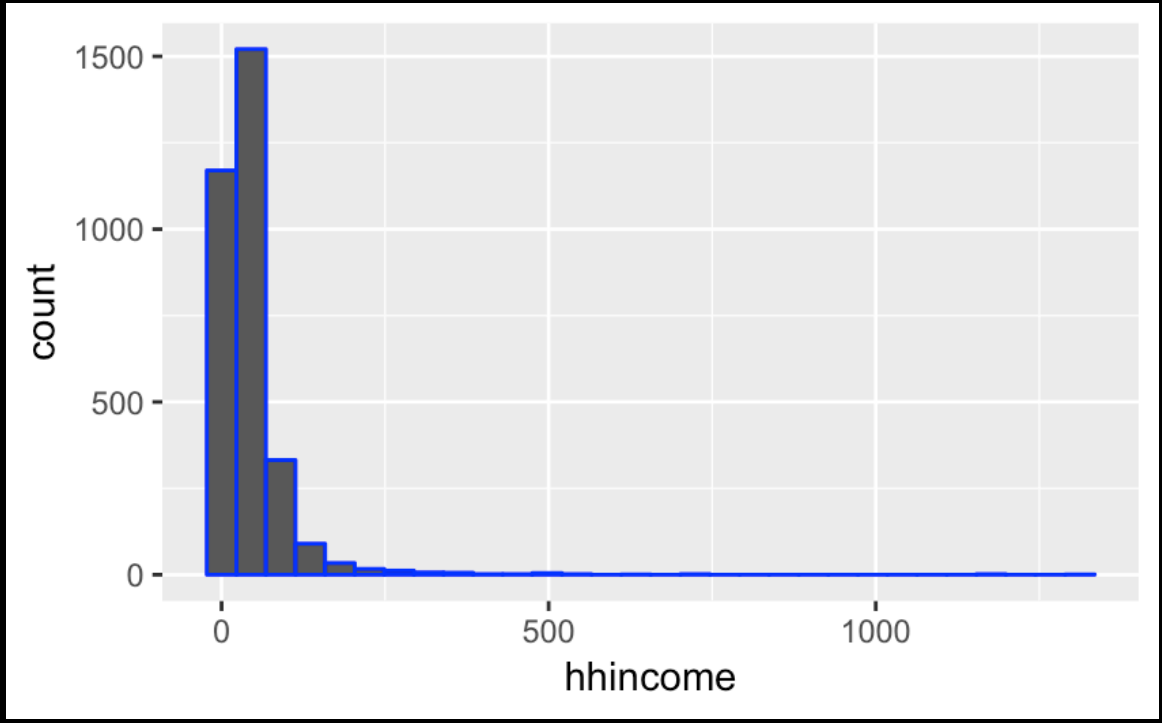
Descriptive Statistics

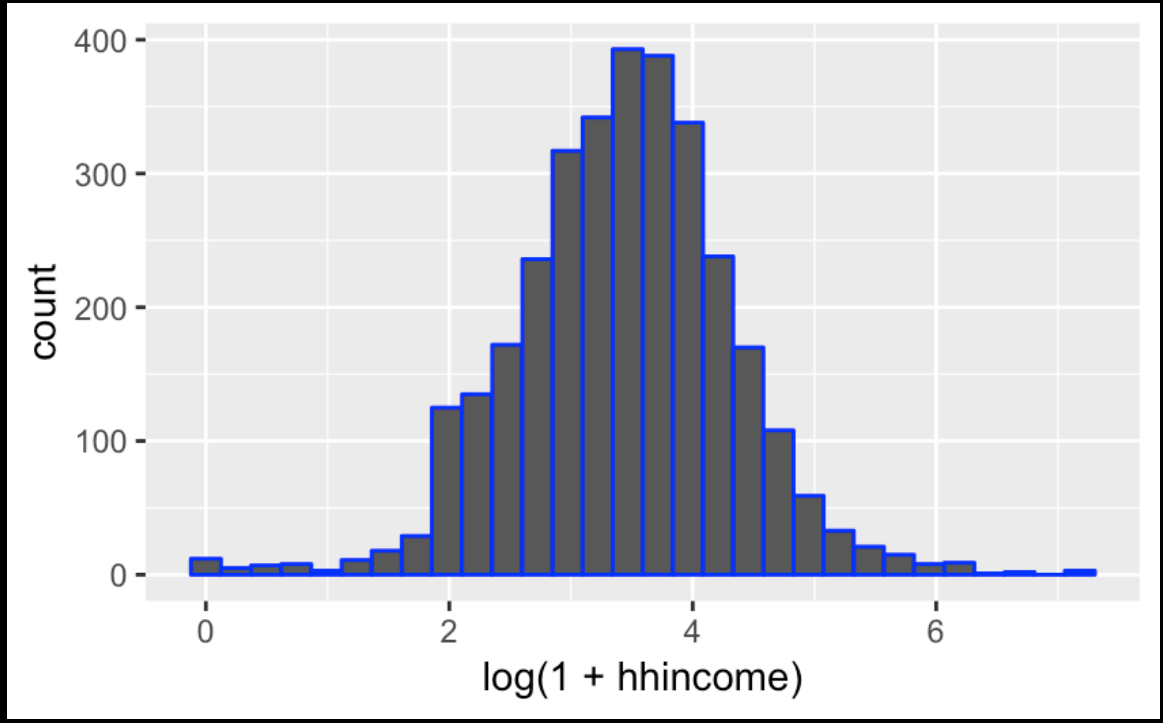
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
personid	3,206	48,620,356.0	23,174,550.0	10,004,010	31,884,768.0	47,241,515	71,483,028.0	99,564,010
age	3,206	66.9	3.7	52	65	67	69	86
hisp	3,206	0.1	0.3	0	0	0	0	1
PriceCategory	3,206	11.9	3.3	0	10	12	14	17
married	3,206	0.7	0.4	0	0	1	1	1
MyBestBuy	3,206	0.6	0.5	0	0	1	1	1
hhincome	3,206	45.3	64.3	0	17.0	31.1	52.8	1,312
appliances	3,206	0.7	0.5	0	0	1	1	1
Warranty	3,206	0.6	0.5	0	0	1	1	1
familysize	3,206	3.0	1.0	1	2	3	4	4
productgeneration	3,206	7.4	1.8	1	7	8	8	11
newcustomer	3,206	0.7	0.5	0	0	1	1	1
weekend	3,206	0.4	0.5	0	0	0	1	1

Variable Distribution: Warranty

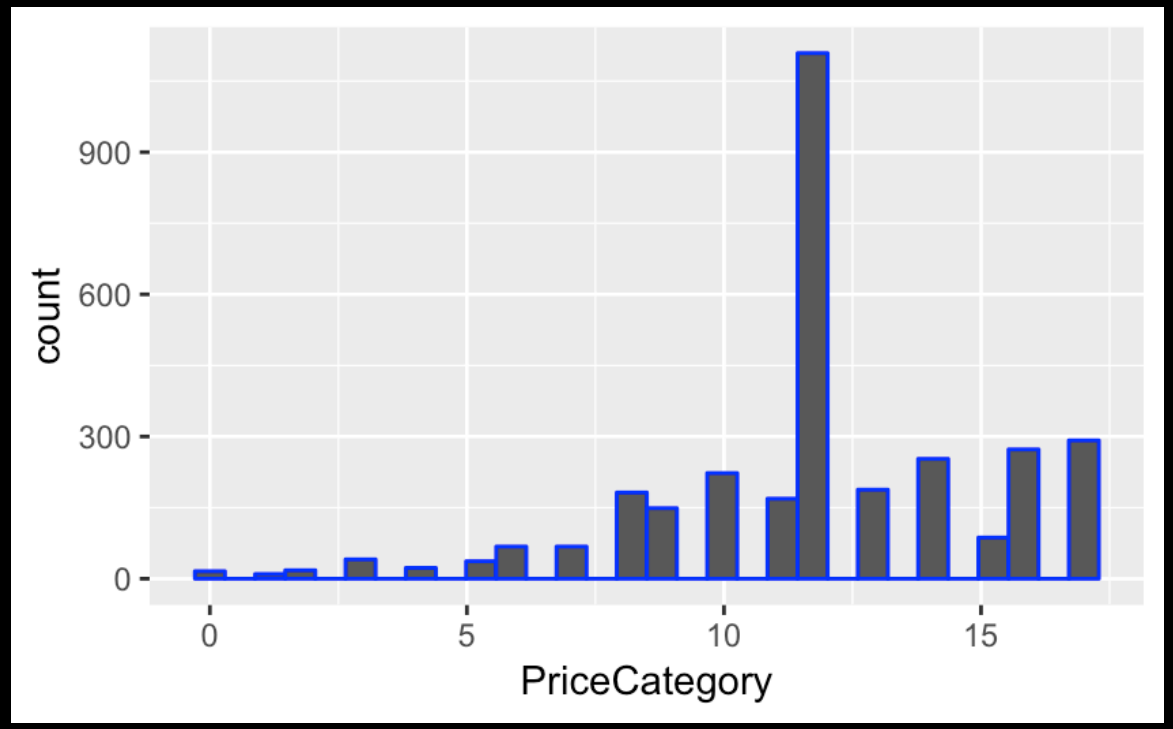


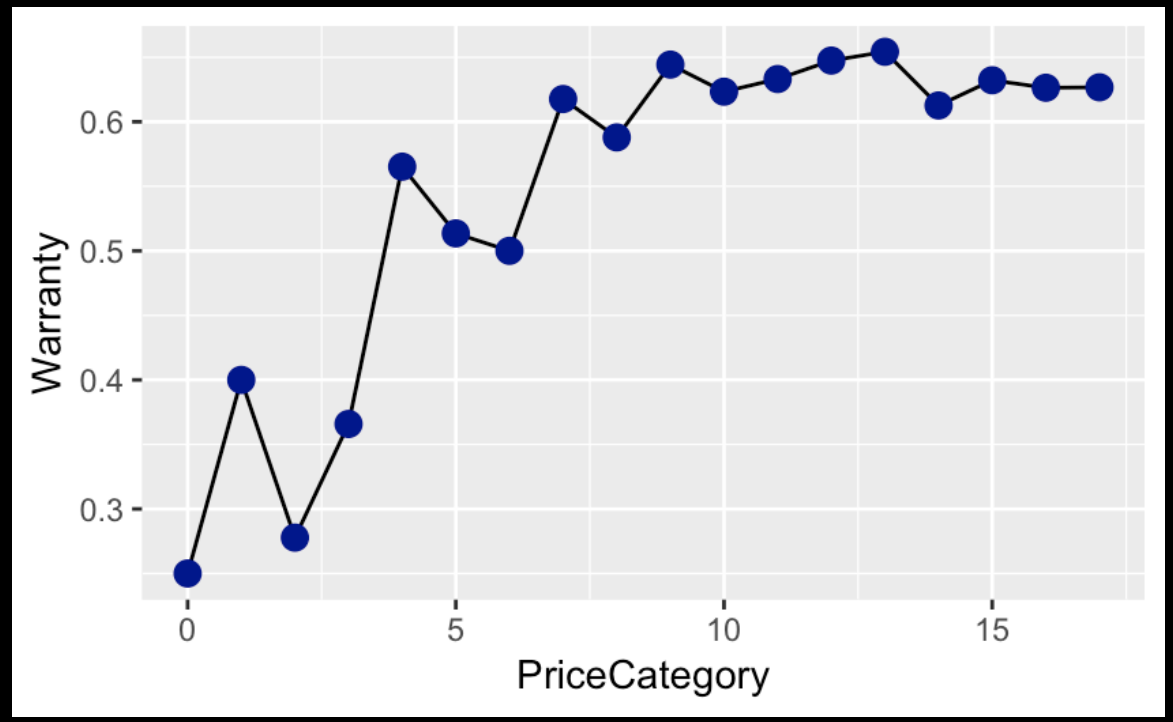
Variable Distribution: Income



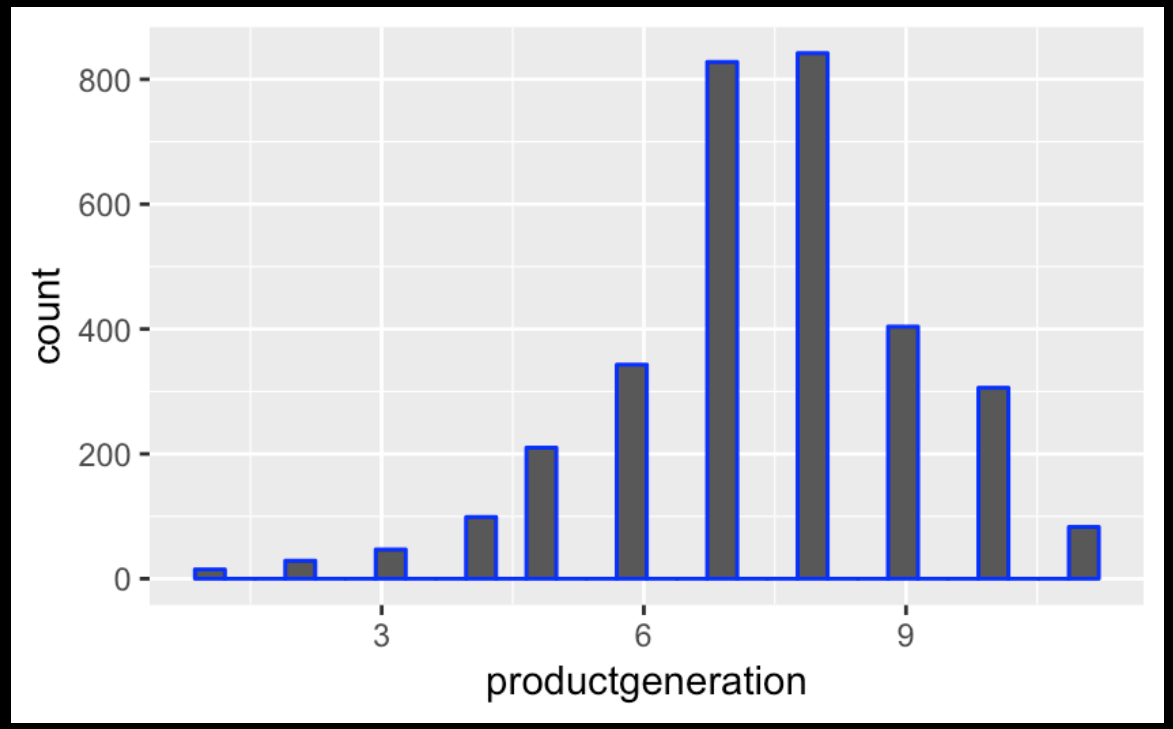


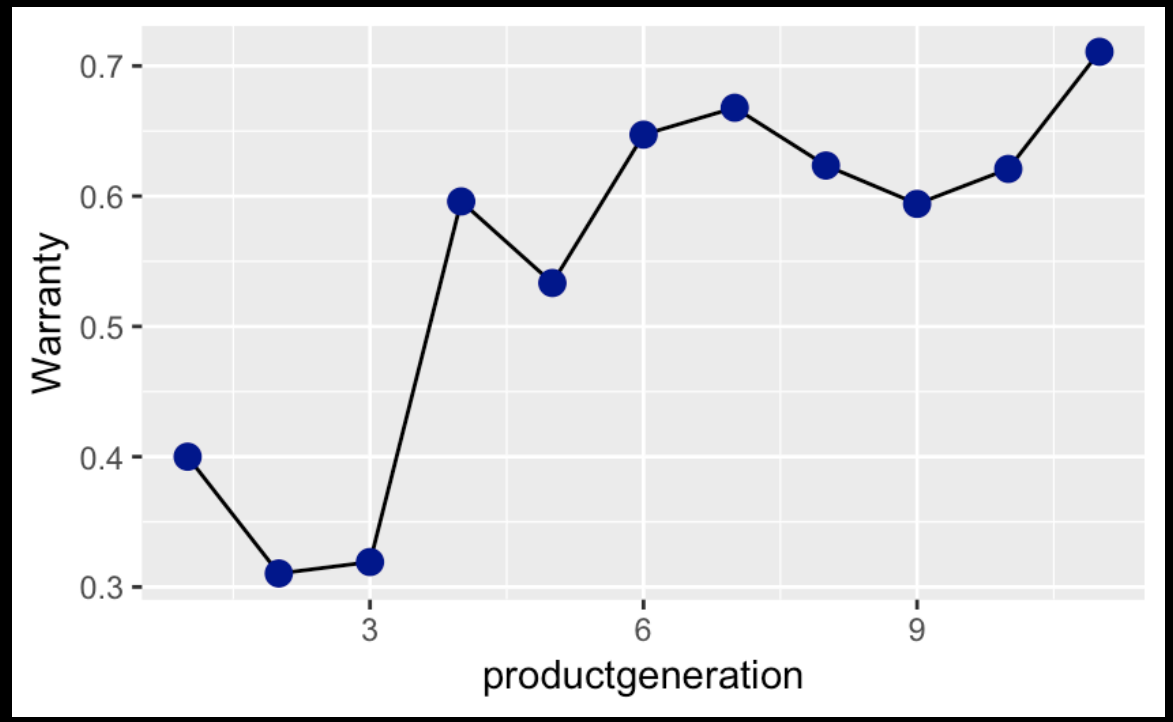
Variable Distribution: Price Category



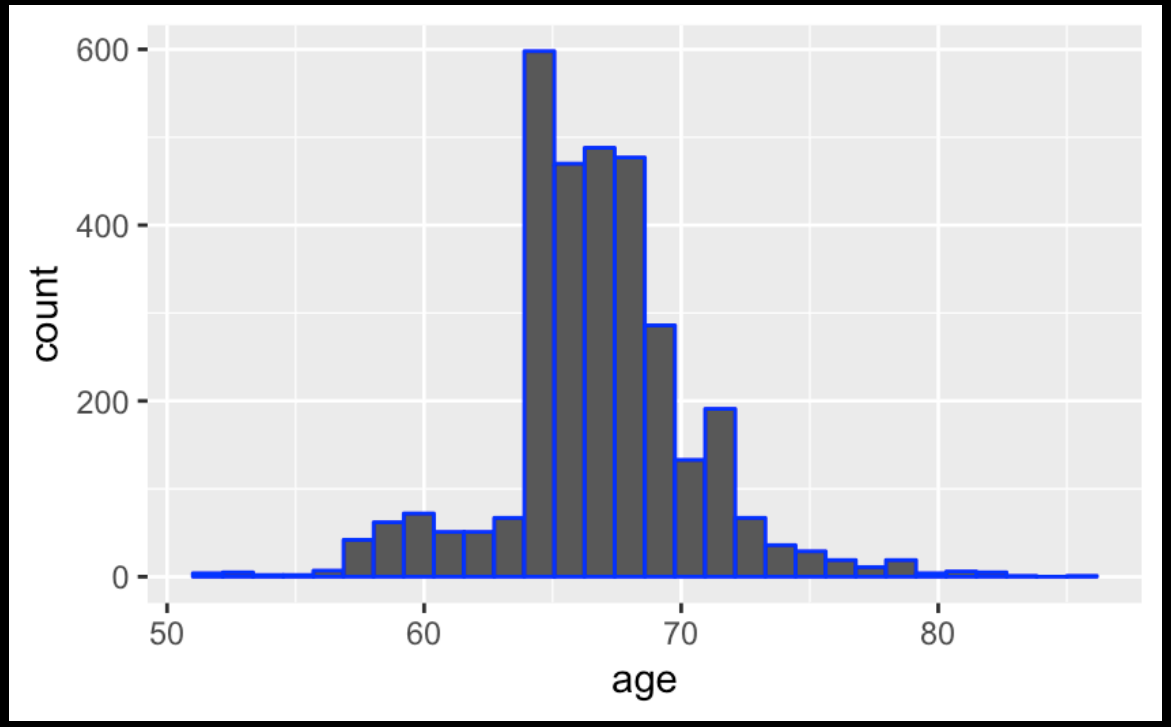


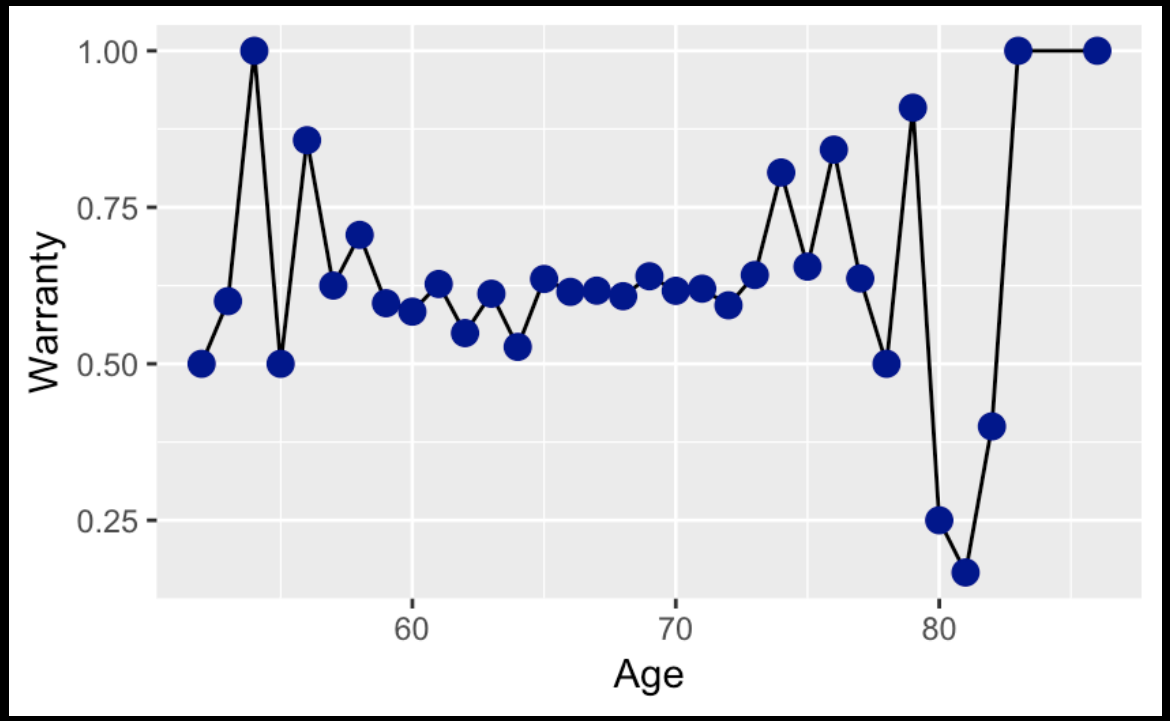
Variable Distribution: Product Generation



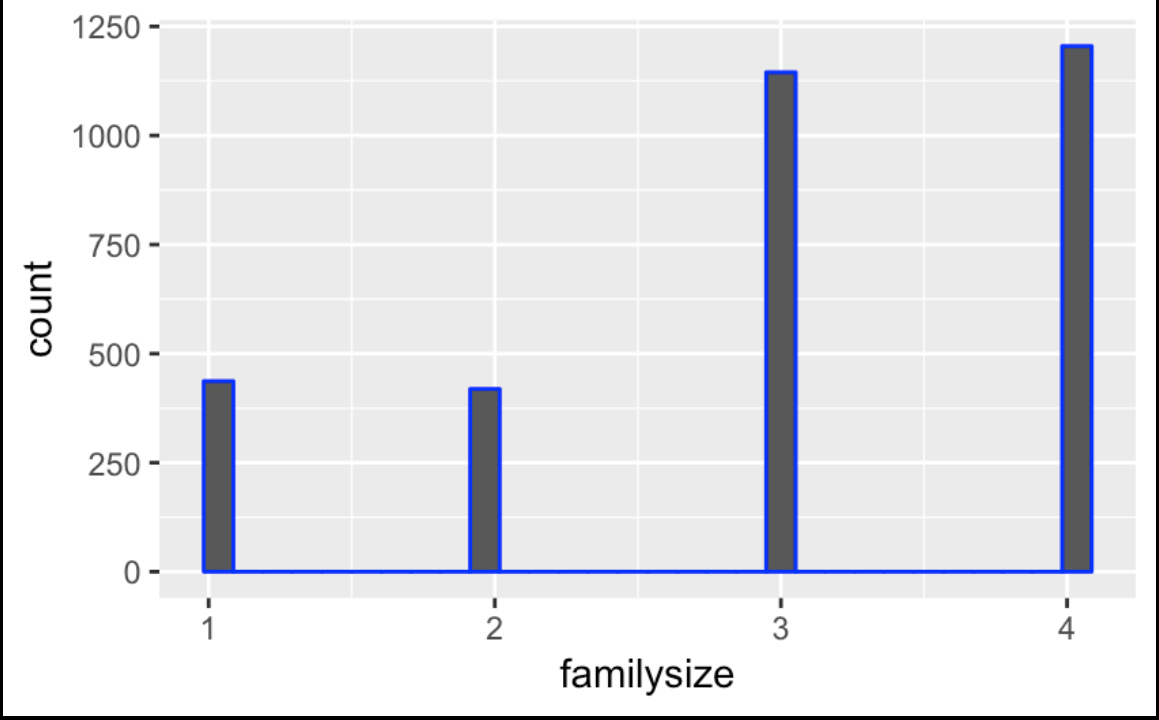


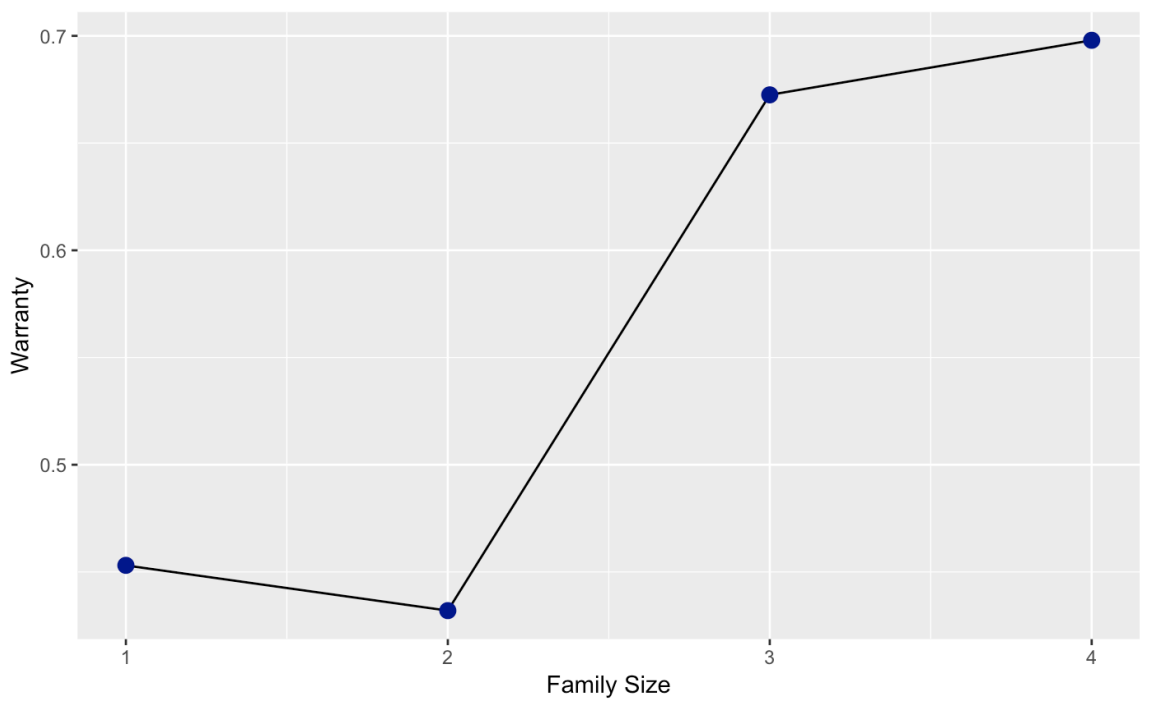
Variable Distribution: Age



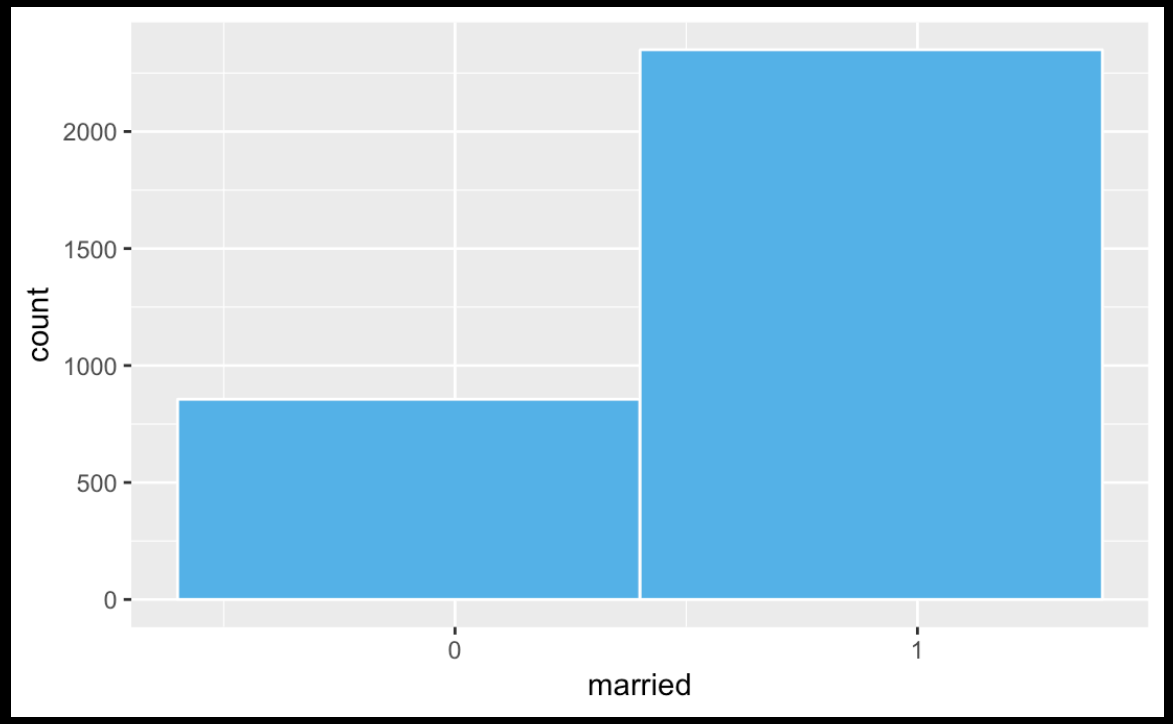


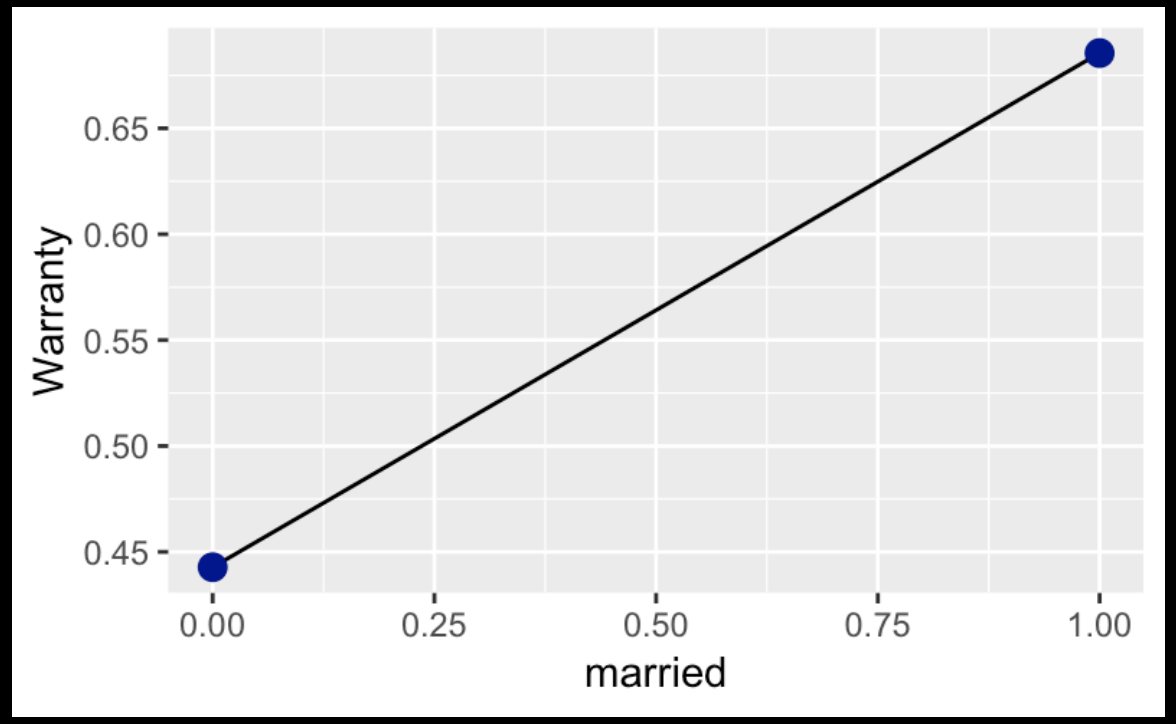
Variable Distribution: Family Size



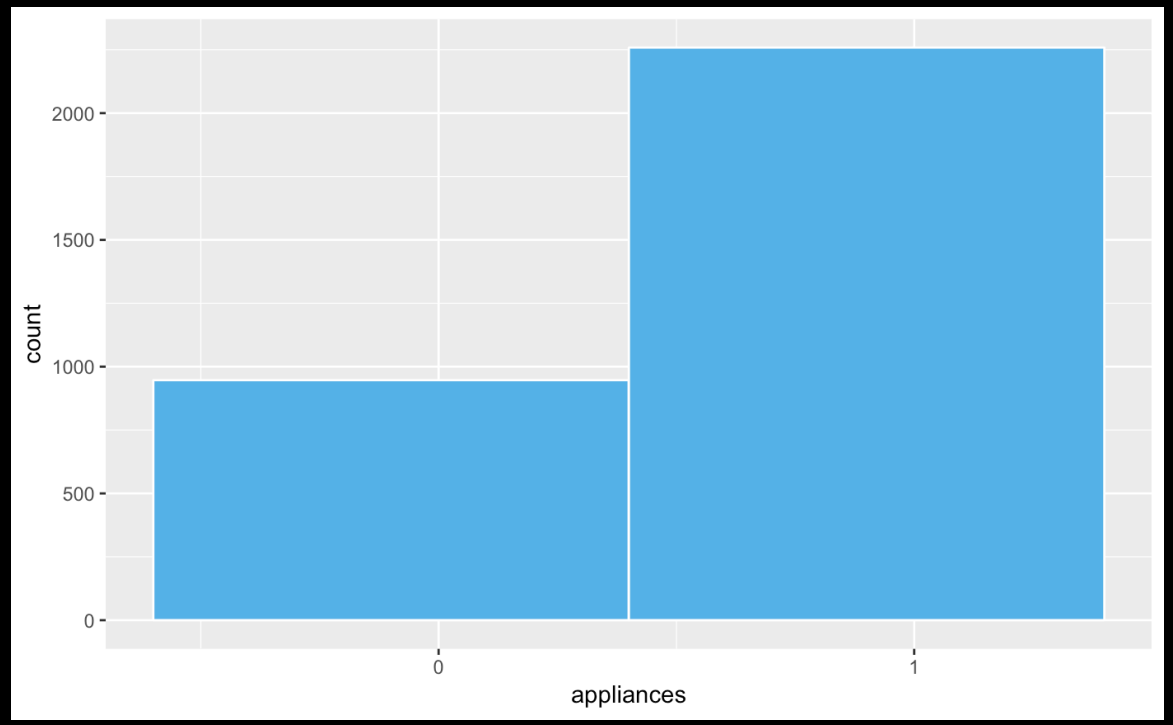


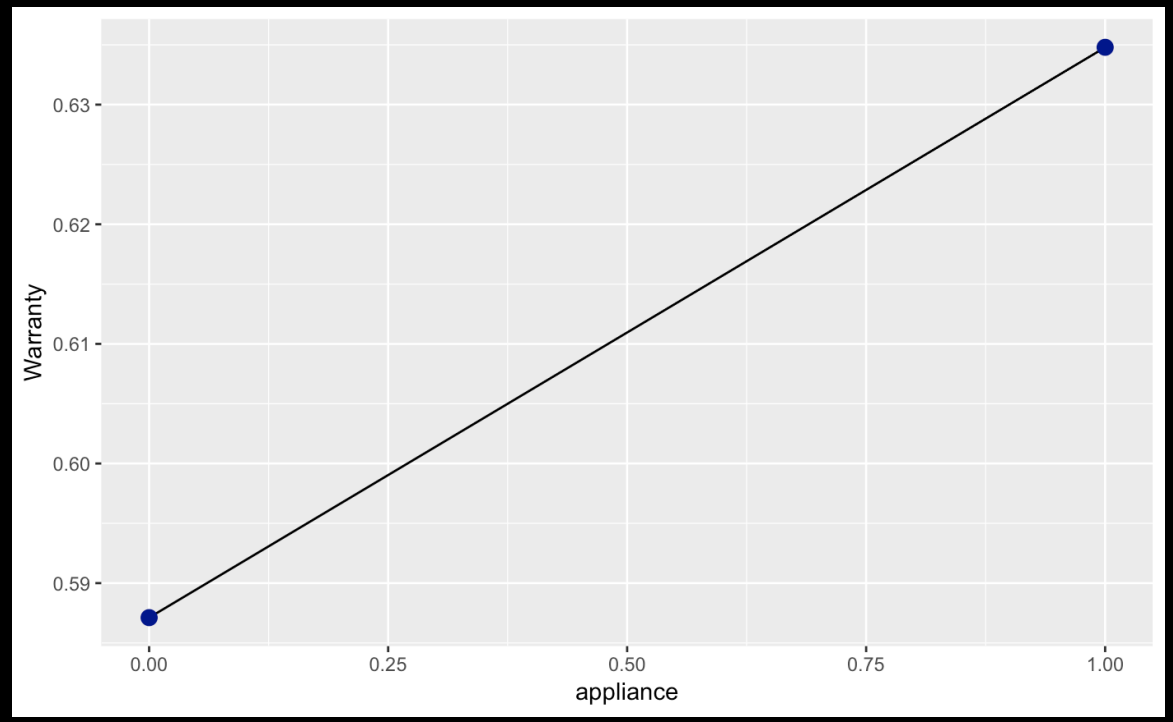
Variable Distribution: Married



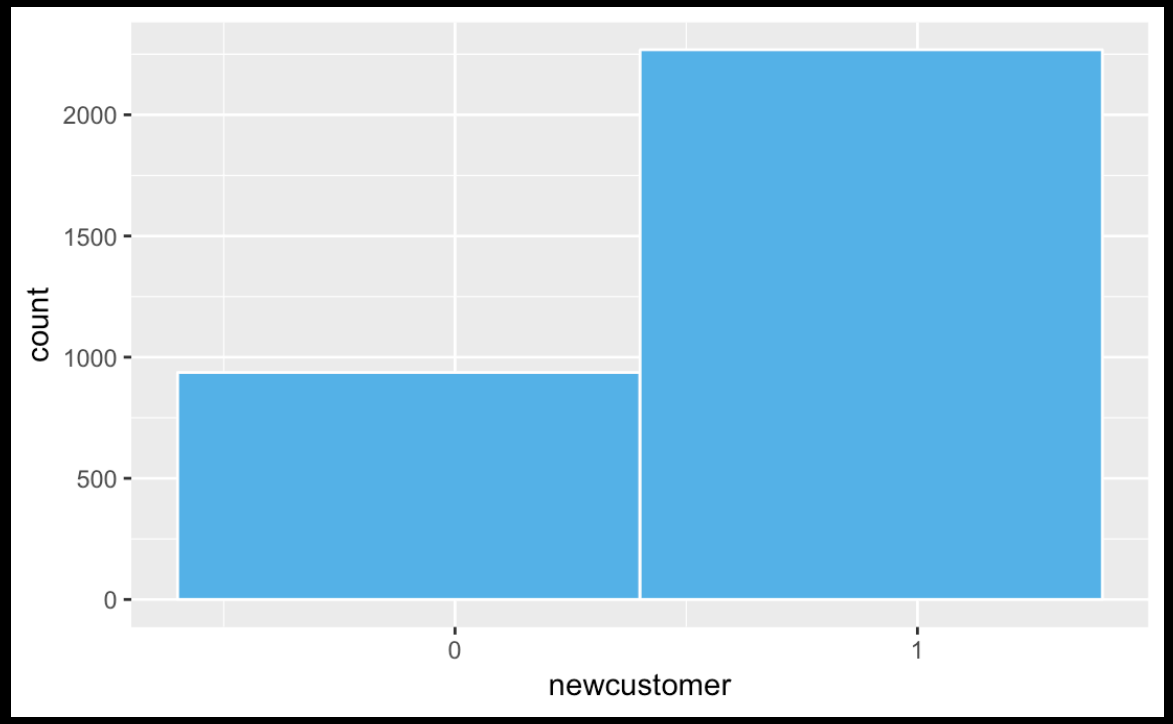


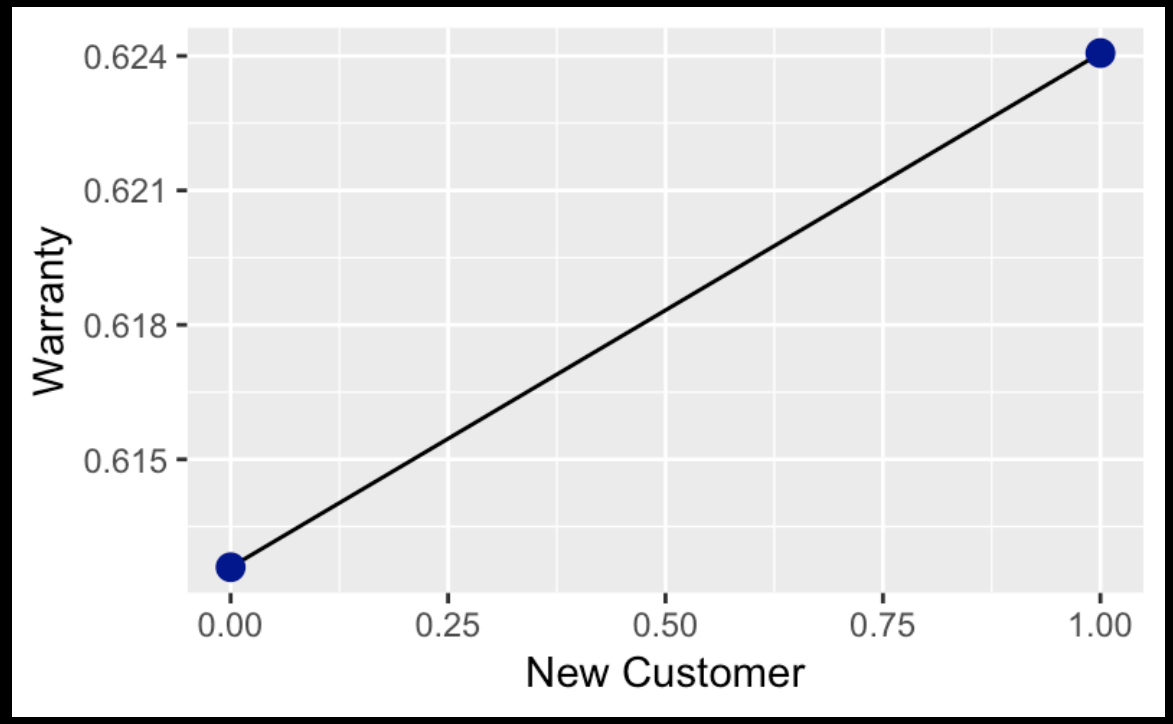
Variable Distribution: Home Appliance



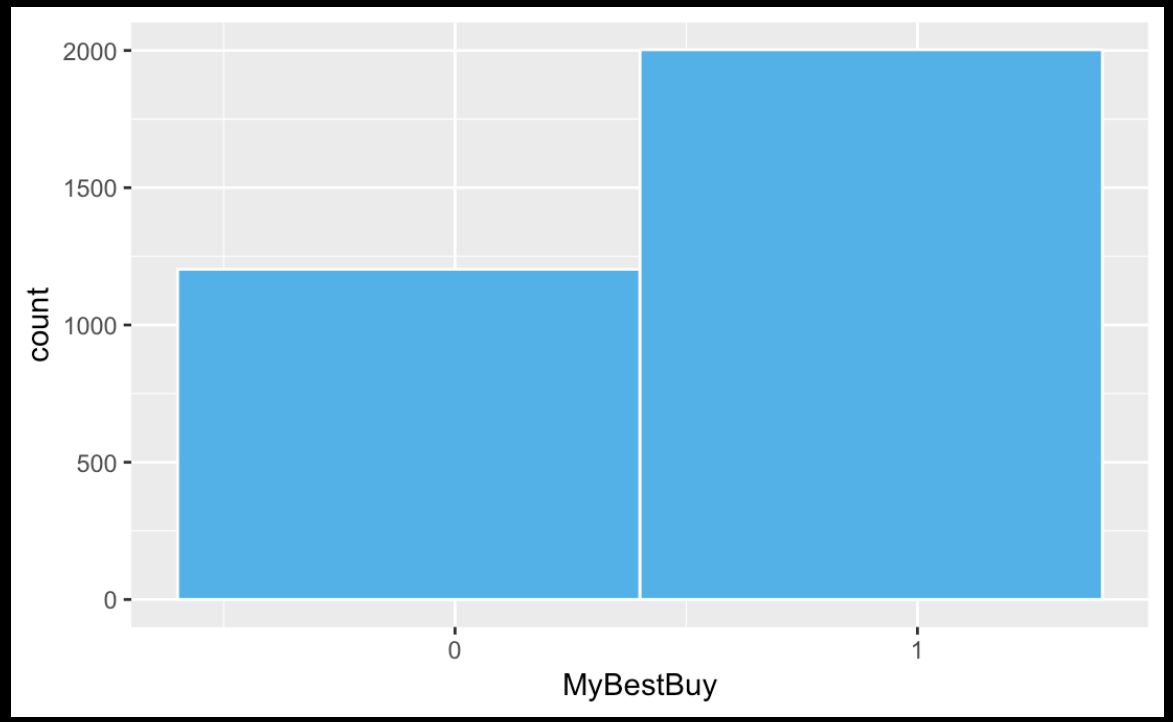


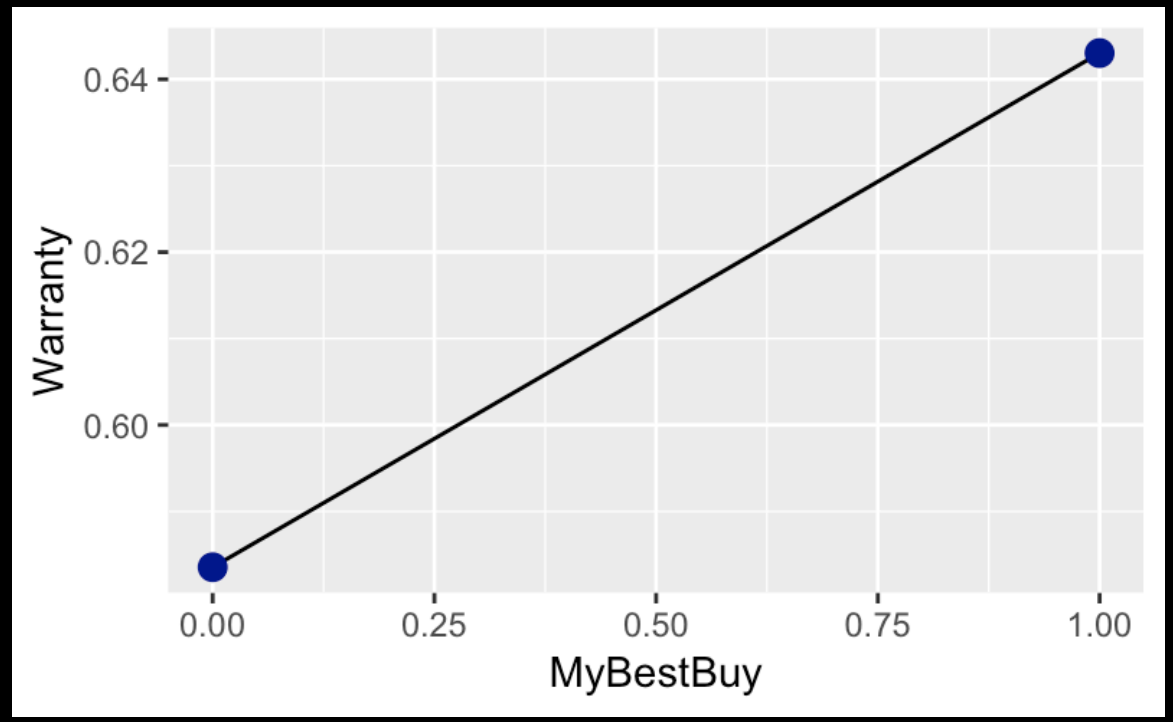
Variable Distribution: New Customer



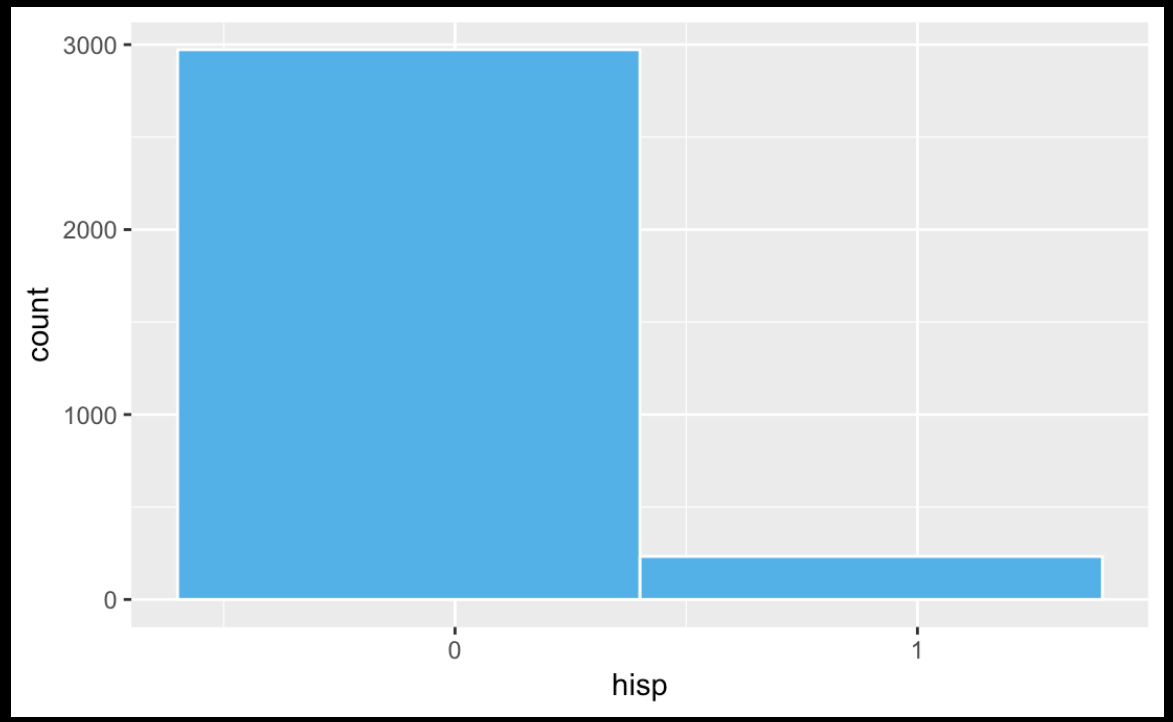


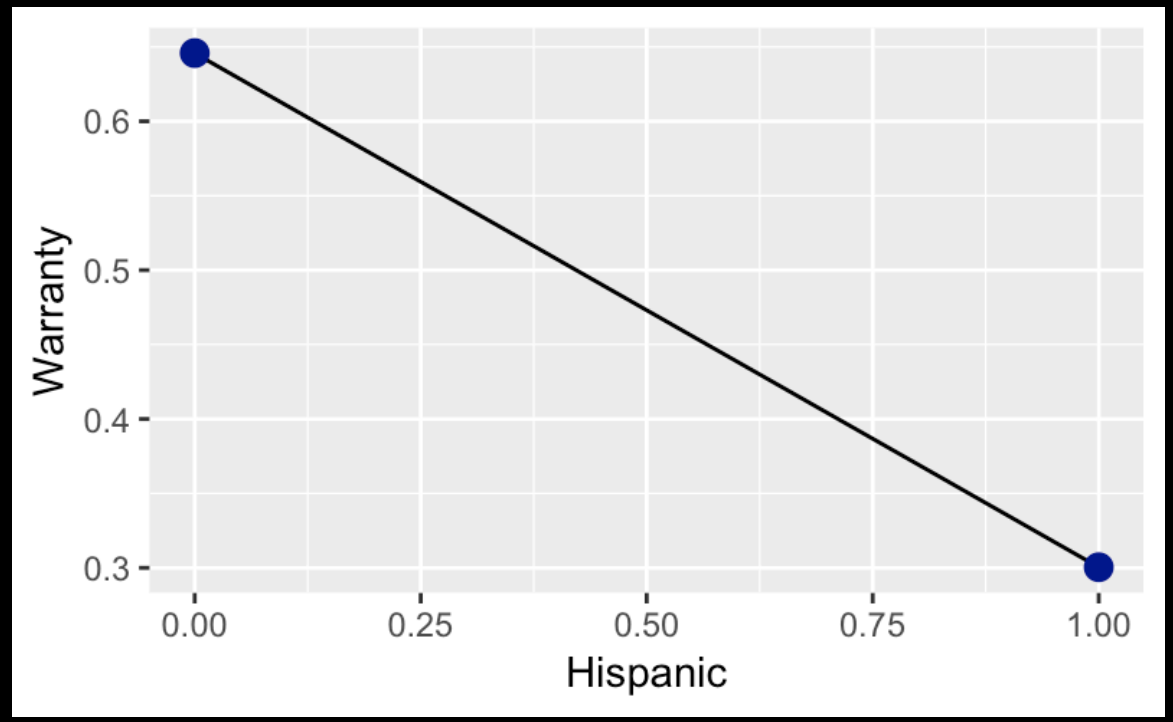
Variable Distribution: MyBestBuy Credit Card



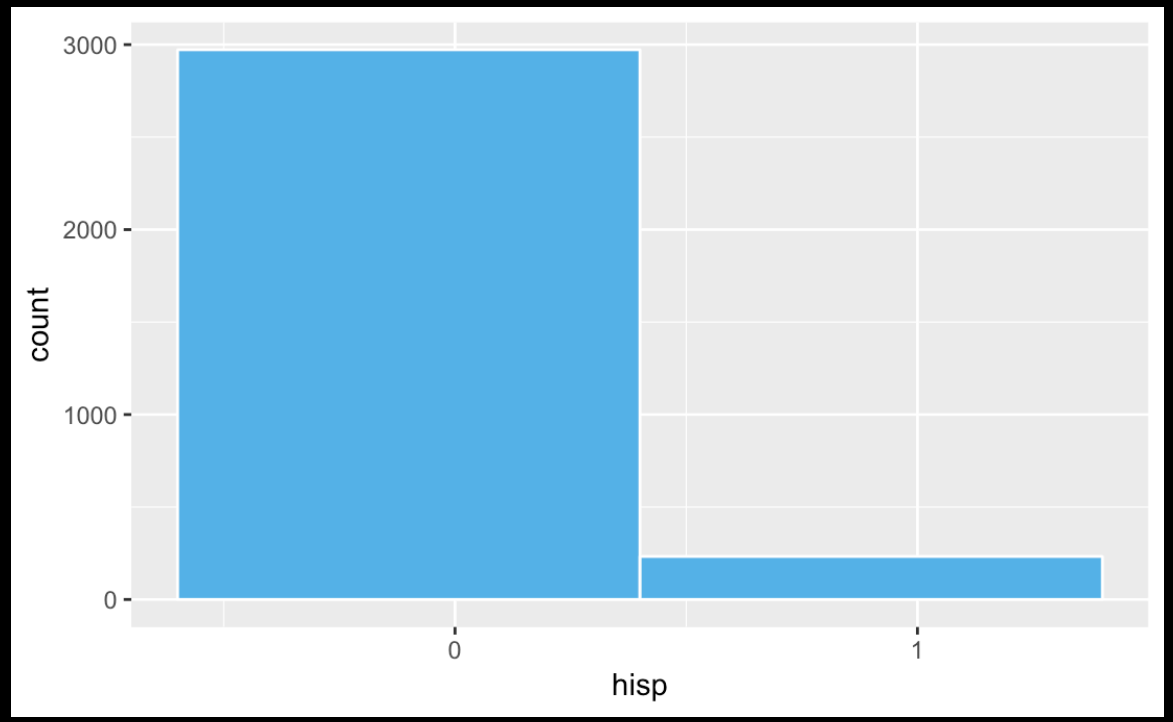


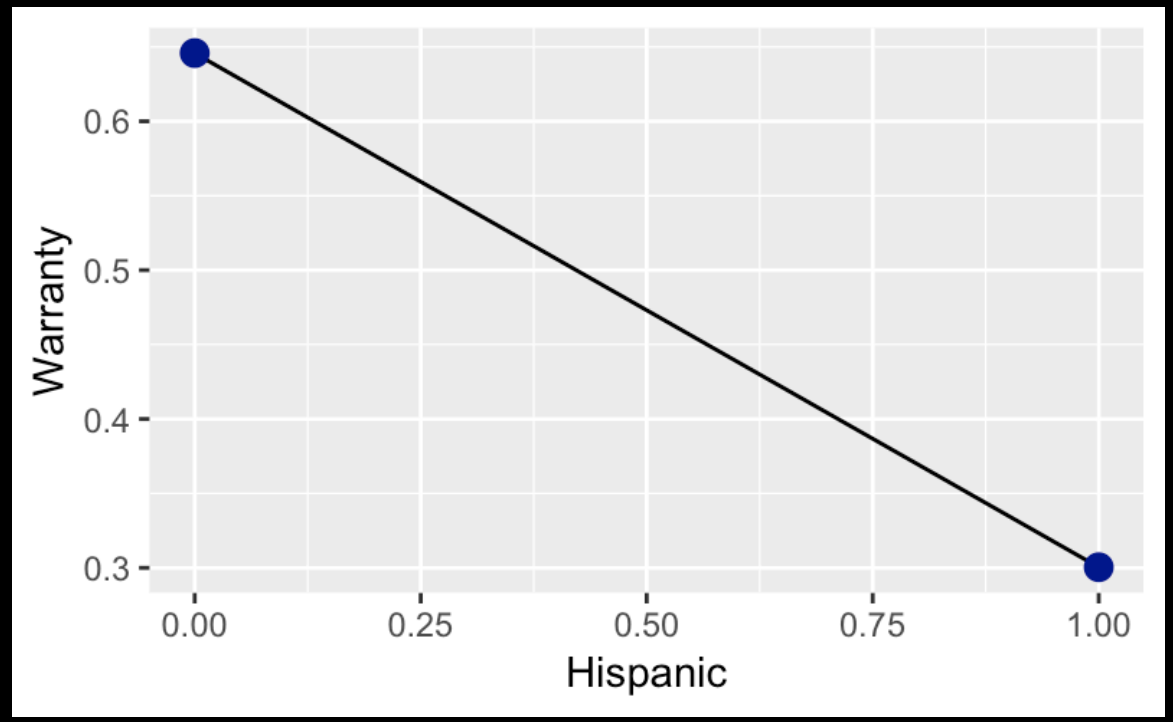
Variable Distribution: Hispanic





Variable Distribution: Hispanic





Model Selection

Probit Model

- The dependent variable is a dummy variable here, meaning it takes the values of either 0 or 1.
- Since Y is also a 'Latent Variable', I.e., the actual prices of the Warranty is not known but only the outcomes are known (if the customes have purchased the warranty or not), we can go with Logit or Probit models.
- As, the outcomes of the two models are identical, we can go with either.
- So, we chose ***Probit***.

Conceptual Model

$$\text{Warranty Purchase Rate} = \beta_0 + \beta_1 \text{Price Category} + \beta_2 \text{Appliances} + \beta_3 \text{My Best Buy} + \beta_4 \text{Age} + \beta_5 \text{Income} + \beta_6 \text{Hispanic} + \beta_7 \text{Family Size}$$

Multicollinearity

Correction Matrix

	familysize	married	age	newcustomer	PriceCategory
familysize	1.00000000	0.87308567	0.1456522428	0.023601740	0.064004686
married	0.87308567	1.00000000	0.1575621439	0.015226049	0.074948382
age	0.14565224	0.15756214	1.0000000000	0.012005404	0.039916952
newcustomer	0.02360174	0.01522605	0.0120054043	1.000000000	-0.005186286
PriceCategory	0.06400469	0.07494838	0.0399169525	-0.005186286	1.000000000
productgeneration	0.06164029	0.07437216	0.0542402527	-0.011593103	0.944045453
MyBestBuy	0.10869629	0.12202874	0.3115567007	-0.047584715	0.135125725
appliances	0.06402652	0.09605320	0.1967221171	-0.002667570	0.309767740
income	0.37362179	0.44114233	0.1534979304	0.018100583	0.451643709
hisp	-0.01477214	-0.02658656	-0.0006346689	0.013466952	-0.328141800

	productgeneration	MyBestBuy	appliances	income	hisp
familysize	0.06164029	0.10869629	0.06402652	0.37362179	-0.0147721428
married	0.07437216	0.12202874	0.09605320	0.44114233	-0.0265865550
age	0.05424025	0.31155670	0.19672212	0.15349793	-0.0006346689
newcustomer	-0.01159310	-0.04758471	-0.00266757	0.01810058	0.0134669523
PriceCategory	0.94404545	0.13512572	0.30976774	0.45164371	-0.3281418001
productgeneration	1.00000000	0.13372877	0.29397319	0.43432705	-0.3080954399
MyBestBuy	0.13372877	1.00000000	0.12800930	0.10315184	-0.0659347592
appliances	0.29397319	0.12800930	1.00000000	0.31228935	-0.1163422707
income	0.43432705	0.10315184	0.31228935	1.00000000	-0.1972426565
hisp	-0.30809544	-0.06593476	-0.11634227	-0.19724266	1.0000000000

Variance Inflation Factor

familysize	married	PriceCategory	productgeneration	age
4.336342	4.449807	9.441651	9.198624	1.168817
newcustomer	MyBestBuy	appliances	hhincome	hisp
1.004346	1.148815	1.162433	1.207024	1.125299

Modeling Result

Threshold Test

Min [1216/8, 1990/8]

= Min[152, 248.75]

= 152 > 20

Final Model

Regression Results

	Dependent variable:	
	Warranty	
	Normal SE (1)	HW-Robust SE (2)
PriceCategory	0.0852***	0.0852***
	(0.0149)	(0.0135)
appliances	1.6853***	1.6853***
	(0.1995)	(0.1856)
MyBestBuy	0.0831	0.0831
	(0.0512)	(0.0518)
age	-0.0116	-0.0116
	(0.0069)	(0.0070)
income	0.1174***	0.1174***
	(0.0341)	(0.0337)
hisp	-0.8570***	-0.8570***
	(0.1025)	(0.0973)
married	0.5687***	0.5687***
	(0.0581)	(0.0589)
PriceCategory:appliances	-0.1532***	-0.1532***
	(0.0175)	(0.0159)
Constant	-0.5552	-0.5552
	(0.4793)	(0.4783)
Observations	3,206	3,206
Log Likelihood	-1,945.6640	-1,945.6640
Akaike Inf. Crit.	3,909.3270	3,909.3270
Note:	*p<0.05; **p<0.01; ***p<0.001	

Model fit assessment: significant

```
> lrtest(probit1f, null_probit)
Likelihood ratio test

Model 1: Warranty ~ PriceCategory * appliances + MyBestBuy + age + log(1 +
    hhincome) + hisp + married
Model 2: Warranty ~ 1
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   9 -1945.7
2   1 -2127.9 -8 364.41  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Heteroscedasticity: significant

```
> # Check heteroscedasticity
> gqtest(probit1f) # p-value = 0.8474   Insignificant Goldfeld-Quandt test does not indicate heteroscedasticity

        Goldfeld-Quandt test

data:  probit1f
GQ = 0.98594, df1 = 1594, df2 = 1594, p-value = 0.6113
alternative hypothesis: variance increases from segment 1 to 2

> bptest(probit1f) # p-value = 0.00000001508   Significant Breusch-Pagan test indicates heteroscedasticity

        studentized Breusch-Pagan test

data:  probit1f
BP = 49.377, df = 8, p-value = 0.00000005384
```

Accuracy Rate = 68.22%

```
> pred = predict(probit1f, data=mydata,type="response")
> Warranty_prediction <- ifelse(pred >= 0.5,1,0)
> misClasificError <- mean(Warranty_prediction != mydata$Warranty)
> print(paste('Accuracy',1-misClasificError)) # the correct classif
[1] "Accuracy 0.682158452900811"
```

Marginal Effect

Marginal Effects

	Dependent variable:
	Warranty
	Probit
PriceCategory	0.0323***
	(0.0052)
appliances	0.6004***
	(0.0524)
MyBestBuy	0.0316
	(0.0195)
age	-0.0044
	(0.0026)
log(1 + hhincome)	0.0445***
	(0.0129)
hisp	-0.3314***
	(0.0353)
married	0.2200***
	(0.0228)
PriceCategory:appliances	-0.0580***
	(0.0062)
Observations	3,206
Log Likelihood	-1,945.6640
Akaike Inf. Crit.	3,909.3270
Note:	*p<0.05; **p<0.01; ***p<0.001

Marginal Effects

	Dependent variable:
	Warranty
	Marg.Eff.w/RobStdEr
PriceCategory	0.0323***
	(0.0052)
appliances	0.6004***
	(0.0524)
MyBestBuy	0.0316
	(0.0195)
age	-0.0044
	(0.0026)
income	0.0445***
	(0.0129)
hisp	-0.3314***
	(0.0353)
married	0.2200***
	(0.0228)
PriceCategory:appliances	-0.0580***
	(0.0062)
Observations	3,206
Log Likelihood	-1,945.6640
Akaike Inf. Crit.	3,909.3270
Note:	*p<0.05; **p<0.01; ***p<0.001

Modeling Output

Regression Results

	Dependent variable:				
			Warranty		
	probit1	probit1b	probit1d	probit1e	probit1f
	(1)	(2)	(3)	(4)	(5)
PriceCategory	-0.016	0.012	0.080***	0.085***	0.085***
	(0.008)	(0.026)	(0.014)	(0.014)	(0.014)
MyBestBuy	0.102*	0.098	0.092	0.083	0.083
	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
appliances	-0.001	-0.004	1.659***	1.688***	1.685***
	(0.055)	(0.055)	(0.188)	(0.189)	(0.189)
age	-0.013	-0.013	-0.011	-0.012	-0.012
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
log(1 + hhincome)	0.137***	0.243*	0.155***	0.117***	0.117***
	(0.033)	(0.099)	(0.033)	(0.034)	(0.034)
hisp	-0.885***	-0.873***	-0.852***	-0.855***	-0.857***
	(0.096)	(0.097)	(0.098)	(0.099)	(0.099)
familysize	0.212***	0.211***	0.213***		
	(0.024)	(0.025)	(0.025)		
PriceCategory:log(1 + hhincome)		-0.009			
		(0.008)			
factor(familysize)2				-0.045	
				(0.088)	
factor(familysize)3				0.510***	
				(0.078)	
factor(familysize)4				0.583***	
				(0.077)	
married					0.569***
					(0.059)
PriceCategory:appliances			-0.150***	-0.153***	-0.153***
			(0.016)	(0.016)	(0.016)
Constant	0.288	-0.053	-0.889	-0.521	-0.555
	(0.447)	(0.540)	(0.472)	(0.476)	(0.474)
Observations	3,206	3,206	3,206	3,206	3,206
Log Likelihood	-1,999.385	-1,998.736	-1,954.854	-1,944.659	-1,945.664
Akaike Inf. Crit.	4,014.769	4,015.473	3,927.709	3,911.317	3,909.327
Note:					*p<0.05; **p<0.01; ***p<0.001

```
> AIC(probit1, probit1a, probit1b, probit1d, probit1e, probit1f)
         df      AIC
probit1   8 4014.769
probit1a  8 4013.294
probit1b  9 4015.473
probit1d  9 3927.709
probit1e 11 3911.317
probit1f  9 3909.327
> BIC(probit1, probit1a, probit1b, probit1d, probit1e, probit1f)
         df      BIC
probit1   8 4063.351
probit1a  8 4061.876
probit1b  9 4070.128
probit1d  9 3982.364
probit1e 11 3978.118
probit1f  9 3963.982
```

Modeling Output

Regression Results

	Dependent variable:				
			Warranty		
	probit1	probit1b	probit1d	probit1e	probit1f
	(1)	(2)	(3)	(4)	(5)
PriceCategory	-0.016	0.012	0.080***	0.085***	0.085***
	(0.008)	(0.026)	(0.014)	(0.014)	(0.014)
MyBestBuy	0.102*	0.098	0.092	0.083	0.083
	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
appliances	-0.001	-0.004	1.659***	1.688***	1.685***
	(0.055)	(0.055)	(0.188)	(0.189)	(0.189)
age	-0.013	-0.013	-0.011	-0.012	-0.012
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
log(1 + hhincome)	0.137***	0.243*	0.155***	0.117***	0.117***
	(0.033)	(0.099)	(0.033)	(0.034)	(0.034)
hisp	-0.885***	-0.873***	-0.852***	-0.855***	-0.857***
	(0.096)	(0.097)	(0.098)	(0.099)	(0.099)
familysize	0.212***	0.211***	0.213***		
	(0.024)	(0.025)	(0.025)		
PriceCategory:log(1 + hhincome)		-0.009			
		(0.008)			
factor(familysize)2				-0.045	
				(0.088)	
factor(familysize)3				0.510***	
				(0.078)	
factor(familysize)4				0.583***	
				(0.077)	
married					0.569***
					(0.059)
PriceCategory:appliances			-0.150***	-0.153***	-0.153***
			(0.016)	(0.016)	(0.016)
Constant	0.288	-0.053	-0.889	-0.521	-0.555
	(0.447)	(0.540)	(0.472)	(0.476)	(0.474)
Observations	3,206	3,206	3,206	3,206	3,206
Log Likelihood	-1,999.385	-1,998.736	-1,954.854	-1,944.659	-1,945.664
Akaike Inf. Crit.	4,014.769	4,015.473	3,927.709	3,911.317	3,909.327
Note:					*p<0.05; **p<0.01; ***p<0.001

```
> AIC(probit1, probit1a, probit1b, probit1d, probit1e, probit1f)
         df      AIC
probit1   8 4014.769
probit1a  8 4013.294
probit1b  9 4015.473
probit1d  9 3927.709
probit1e 11 3911.317
probit1f  9 3909.327
> BIC(probit1, probit1a, probit1b, probit1d, probit1e, probit1f)
         df      BIC
probit1   8 4063.351
probit1a  8 4061.876
probit1b  9 4070.128
probit1d  9 3982.364
probit1e 11 3978.118
probit1f  9 3963.982
```

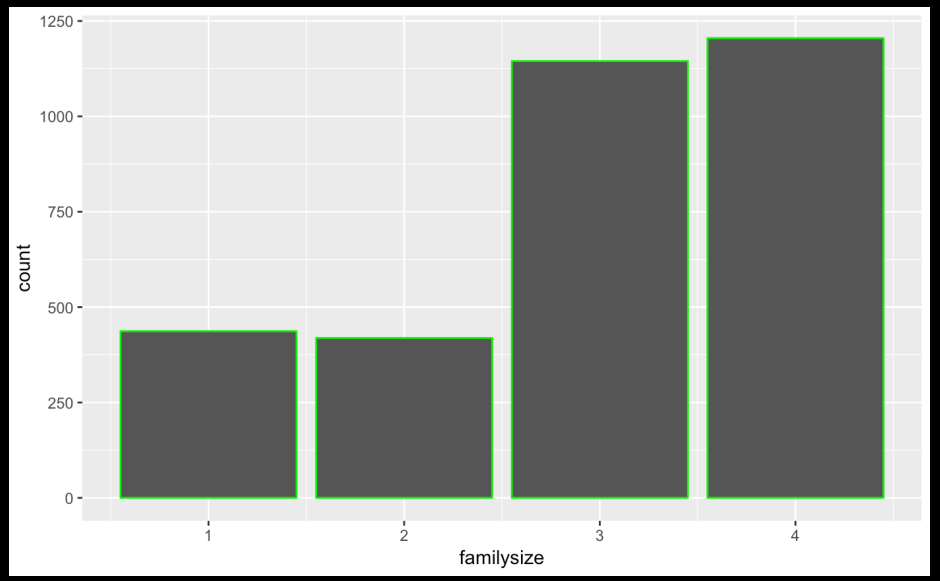
Predictive Analysis

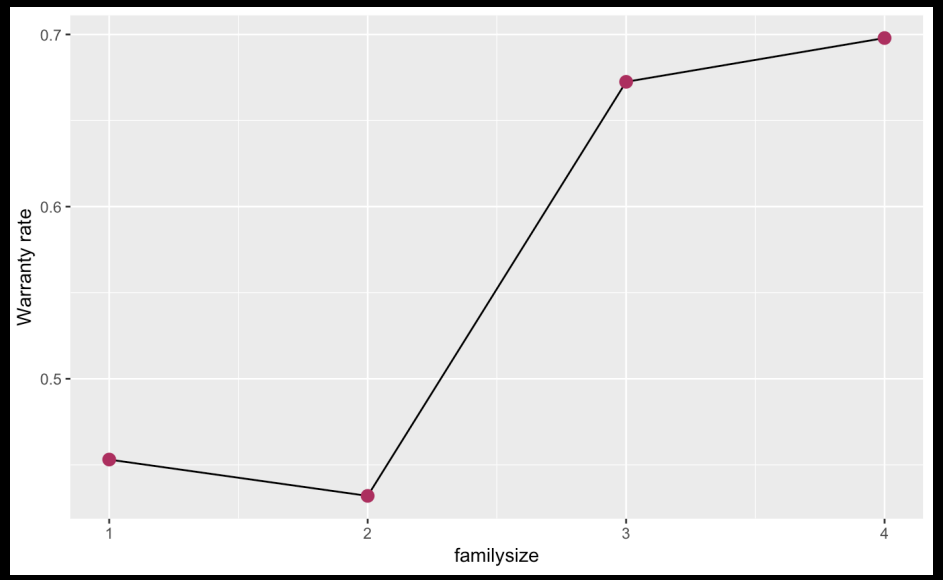
PriceCategory <dbl>	married <dbl>	MyBestBuy <dbl>	income <dbl>	age <dbl>	appliances <dbl>	hisp <int>	hispW <dbl>
11.89863	0.7330006	0.6247661	3.427251	66.91391	0.7046163	0	0.6756860
11.89863	0.7330006	0.6247661	3.427251	66.91391	0.7046163	1	0.3440951

PriceCategory <dbl>	hisp <dbl>	MyBestBuy <dbl>	income <dbl>	age <dbl>	appliances <dbl>	married <int>	marriedW <dbl>
11.89863	0.07267623	0.6247661	3.427251	66.91391	0.7046163	0	0.4906517
11.89863	0.07267623	0.6247661	3.427251	66.91391	0.7046163	1	0.7071976

Family Size

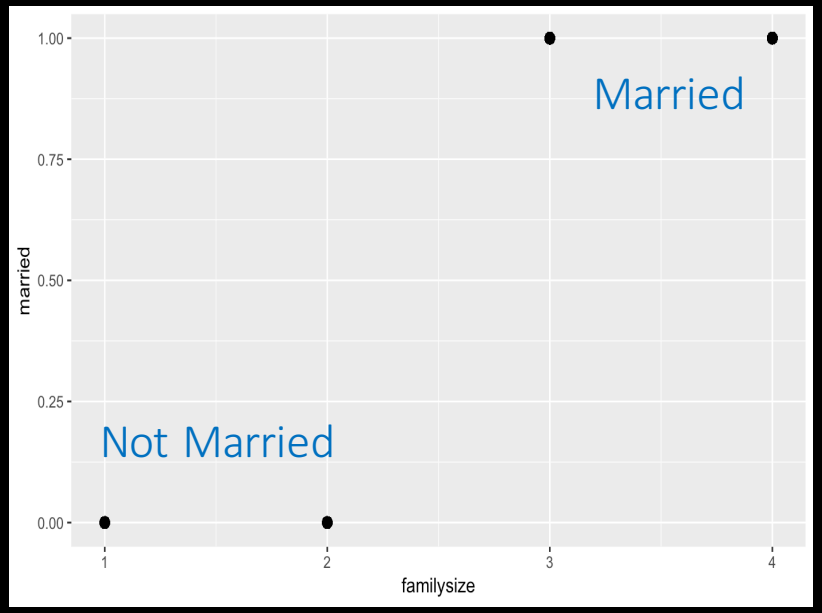
1 2 3 4





Family Size

1 2 3 4



Model Fit – Likelihood Ratio Test

Final model fits significantly better than the null model

```
Likelihood ratio test

Model 1: Warranty ~ PriceCategory + married + MyBestBuy + income + hisp +
    appliances + PriceCategory * appliances
Model 2: Warranty ~ 1
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   8 -1947.1
2   1 -2127.9 -7 361.57  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Model Fit – Correct Classification

The correct classification rate is 68%

Heteroskedasticity

Test indicates heteroscedasticity, therefore, produces Huber-White robust standard errors

```
        studentized Breusch-Pagan test

data:  probit4b4b
BP = 48.577, df = 7, p-value = 0.00000002745
```

```
        Goldfeld-Quandt test

data:  probit4b4b
GQ = 0.98811, df1 = 1595, df2 = 1595, p-value = 0.5944
alternative hypothesis: variance increases from segment 1 to 2
```

	Dependent variable:	
	Warranty	
	Normal SE	HW-Robust SE
	(1)	(2)
PriceCategory	0.087***	0.087***
	(0.014)	(0.013)
married	0.561***	0.561***
	(0.059)	(0.059)
MyBestBuy	0.059	0.059
	(0.049)	(0.049)
income	0.113***	0.113***
	(0.034)	(0.034)
hisp	-0.860***	-0.860***
	(0.099)	(0.097)
appliances	1.680***	1.680***
	(0.189)	(0.186)
PriceCategory:appliances	-0.154***	-0.154***
	(0.016)	(0.016)
Constant	-1.303***	-1.303***
	(0.163)	(0.159)
Observations	3,206	3,206
Log Likelihood	-1,947.085	-1,947.085
Akaike Inf. Crit.	3,910.169	3,910.169
Note:	*p<0.05; **p<0.01; ***p<0.001	

Logit Model

Logit Code

#Logit model

- logit1 <- glm(Warranty~PriceCategory+MyBestBuy+appliances+age+loghhincome2+PriceCategory*appliances+married+hisp, data=mydata1, family = 'binomial')

#log odds of warranty

- stargazer(logit1, title="Regression Results", type="text", column.labels=c("Logit-1"), df=FALSE, digits=2, star.cutoffs = c(0.05,0.01,0.001))

#odds ratios

- stargazer(logit1, apply.coef = exp, t.auto=F, p.auto = F, title="Regression Results", type="text", column.labels=c("OddsRatios"), df=FALSE, digits=4, star.cutoffs = c(0.05,0.01,0.001))

#create a null model

- logit2 <- glm(Warranty~1, data=mydata1)

#compare our model with null model

- lrtest(logit1,logit2)

#pvalue is < 0.001 which means our model is better than the null model.

logit1b<- logitmfx(Warranty ~ PriceCategory + MyBestBuy + appliances + PriceCategory*appliances + loghhincome2 + hisp + married, data=mydata1)

marginaleffects_lg <- logit1b$mfxest[,1]

marg.std.err_lg <- logit1b$mfxest[,2]

stargazer(logit1, omit=c("Constant"),coef = list(marginaleffects_lg), se = list(marg.std.err_lg),title="Marginal Effects", type="text", column.labels=c("Logit"),df=FALSE, digits=4, star.cutoffs = c(0.05,0.01,0.001))

Marginal Effect

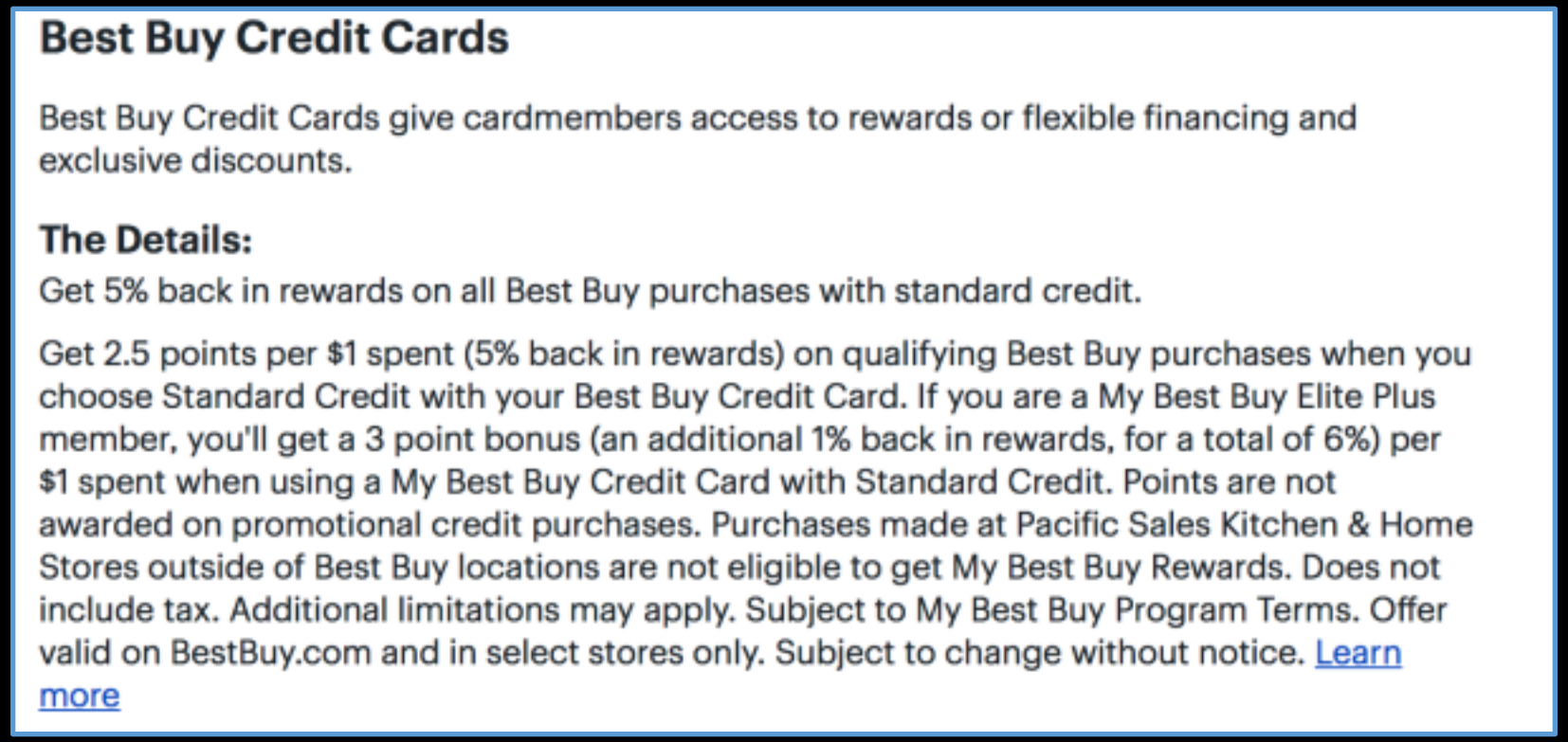
Marginal Effects

	Dependent variable:
	Warranty Logit
PriceCategory	0.0334***
	(0.0055)
MyBestBuy	0.0221
	(0.0190)
appliances	0.6007***
	(0.0503)
age	
loghhincome2	0.0435***
	(0.0131)
married	0.2190***
	(0.0230)
hisp	-0.3381***
	(0.0357)
PriceCategory:appliances	-0.0595***
	(0.0065)
Observations	3,206
Log Likelihood	-1,945.9930
Akaike Inf. Crit.	3,909.9850
Note:	*p<0.05; **p<0.01; ***p<0.001

Company Info

My Best Buy Credit Card – warranty discount

Get 5% back in rewards on all Best Buy purchases with standard credit.



Best Buy Credit Cards

Best Buy Credit Cards give cardmembers access to rewards or flexible financing and exclusive discounts.

The Details:

Get 5% back in rewards on all Best Buy purchases with standard credit.

Get 2.5 points per $1 spent (5% back in rewards) on qualifying Best Buy purchases when you choose Standard Credit with your Best Buy Credit Card. If you are a My Best Buy Elite Plus member, you'll get a 3 point bonus (an additional 1% back in rewards, for a total of 6%) per $1 spent when using a My Best Buy Credit Card with Standard Credit. Points are not awarded on promotional credit purchases. Purchases made at Pacific Sales Kitchen & Home Stores outside of Best Buy locations are not eligible to get My Best Buy Rewards. Does not include tax. Additional limitations may apply. Subject to My Best Buy Program Terms. Offer valid on BestBuy.com and in select stores only. Subject to change without notice. Learn more

https://www.bestbuy.com/site/help-topics/best-buy-credit-cards/pcmcat2044000 50012.c?id=pcmcat20440 0050012