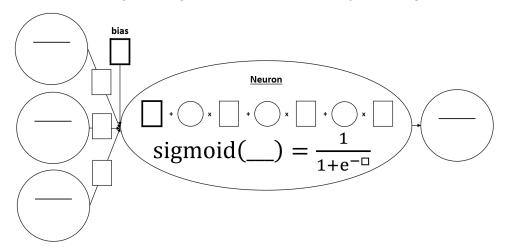
Data Analysis In-Class Worksheet #14: Logistic Regression

Student Name	Section
Part #1: Logistic Regression Fundamentals	
What does the "odds" measure?	
If the probability of winning a 30% coupon at Kohl's is 10%, then winning a 30% coupon?	vhat's the odds of
$log_{10}(10000) =$	
Why do we need to use log(odds), not probabilities, in estimating	logistic regression?
Reason #1:	
Reason #2:	
Suppose we have this logistic regression model.	
$\hat{Y} = 2 + 1 * BED$	
If BED = 3, then $\hat{Y} = \underline{\hspace{1cm}}$	
Ŷ is probability odds In(odds)	
Convert Ŷ back into probability using the function =	 =
(Reminder: You can use the formula set up in Google Sheets to co	mpute the value.)
Why do we need to convert Ŷ back into probability?	
Make a binary decision using the probability: Convert any probability > .5 into YES, and any probability <= .5 in is	nto NO. The decision

Illustrate this logistic regression model as a deep learning neuron with a sigmoid unit:



Part #2: Logistic Regression Application

Does Property Age predict the Property's Zipcode Location?

To answer this question, we can rephrase it as a pair of hypotheses:

H₀: _____

Express these hypotheses in terms of the regression β_{1} slope

H₀: _____

H₁: _____

For H_0 : $\hat{Y} =$ ______

For H_1 : $\hat{Y} =$ _____

In our class dataset, Zipcode has only two possible outcomes: 23185 and 23188. Therefore, Zipcode is a ______ variable.

- ☐ Binary (categorical)
- ☐ Numeric

•	. The odds	ing in 23185 is 80 of being in 2318 –	•			_	
	The odds	ing in 23185 is 50 of being in 23185 is					
Coefficients							
				Wald	Test		
	Estimate	Standard Error	Z	Wald Statistic	df	р	
(Intercept)		0.327	0.466	0.217	1	0.641	
Age	-0.029	0.011 coded as class 1.	-2.532	6.412	1	0.011	
Probabilities Binary cates Odds	gories	rom a logistic re	gression i	S			
Natural log	of odds (ln(odds) or logit)					
The intercept is	s, wh	nich tells us					
The slope is	, which to	ells us					
When a propert	y is a year	older, its odds of	f being in	23188 (Class 1)) wou	ld	
Abdullah's $\hat{Y} = 0$).429. Conv	ert it into odds:					
Convert odds in	to probabil	ity value:					

Use the probability value to predict if the property is in 23185 or	23188:
Abdullah's property is located in 23188. This is a correct incorrect prediction. If we get a confusion matrix like this one, calculate the overall a	ccuracy rate:

		Predicted		
		23185	23188	
Actual	23185	4	6	
	23188	16	24	

APA Writeup

A logistic regression was performed to ascertain	the effect of on the
likelihood of The logistic regression mo	del was statistically (in)significant, x^2
() =, p, McFadden's R ² =	. The model correctly classified
% of cases.	